Exploring Test and Reliability Challenges in the Realm of Approximate Circuitry

Sandeep Kumar Mishra,
Assistant Professor,
Dept.of Electronics & Telecommunication Engineering
College of Engineering Bhubaneswar

***Abstract*—Physical realization of approximate computing (AC) systems is becoming more and more relevant as it becomes increasingly clear that AC is a critical enabler for next-generation computer architectures. This letter explores an important but little-studied topic: the connection between the AC paradigm and hardware malfunctions. We outline the potential for effective testing techniques to raise the yields of the underlying hardware blocks for AC components as well as the related theoretical and practical difficulties. We distinguish between circuits with insufficient provisions and incomplete designs, and we outline the drawbacks of current testing techniques for these two categories of AC block types. We also go over the complex interactions between approximation circuitry and malfunctions that arise over the course of the device's lifetime.**

I. Introduction

VARIOUStypesofapproximatecomputing(AC)wereshowntoofferbenefitsinapplicationsincludingtelecommunications[6],real-timeimageprocessing[2],neuralnetworksimulation[3],arithmeticcalculations[20],andnumericalanalysis [17].SeveralAC-enabledcomputerarchitectures,includingStanford'sERSA[10]andIntel'sRunnemede[4],havebeenproposed.ACconceptsarebeingdiscussedasacoretechnologyfornext-generationdeep-learningaccelerators[21].WithACgettingclosetopracticalapplication,thephysicalrealizationofapproximatecircuitryshiftsintothefocus.Onekeyaspectoftoday'snanoscaletechnologiesistheirvulnerabilitytodefectsthatareintroducedincourseofcircuitfabrication(chieflyaddressedbytesting)andtofailuresduringthecircuit'slifetime(chieflyaddressedbyreliability-enhancementtechniques,suchasonlineerrordetectionandrecovery).Aswewillshowinthisletter,therelationshipofACwithstate-of-the-arttestandreliability approachesisinsufficientlyunderstoodandneedsmore attention.

AC is defined across abstraction layers and combines concepts on system, software, and hardware levels [1]. However,test and reliability challenges are naturally more pronouncedforapproximatehardwareblocks,alsoknownbythename "under signed and opportunistic circuitry" [6]. In general,approximate circuit blocks are allowed to deviate from their reference behavior in a well-defined way. Depending on the specific AC variant, this deviation might be due to incompletedesign(e.g.,anadderthatdoesnotfunctioncorrectlyforafewinput combinations) [11], under provisioning (voltage or fre-quency overscaling) [7], [9], defects (error-tolerance) [8] or soft errors(transienterrortolerance)[15].

Theacceptablebehaviorisusuallyformalizedbymetricsbasedonerrormagnitude,errorrate,ortheircombination.*Errormagnitude*isthemaximalallowedextentofdeviation,e.g.,maximalnumericaldifference[8],peaksignaltonoiseratio[5],orstructuraldissimilarity[9]inimage-processingapplications,orframeerrorrateinwirelesstelecommunication[19].*Errorrate*isthefrequencywithwhicherroneous outputs are produced (which, under simplifyingassumptions,correspondstotheprobabilityoferror)[18].In this letter, we discuss the challenges and the opportunitiesfortestandreliabilityapproachesincontextofapproximatecircuitblocks.Wedistinguishbetweentwobasictypesof AC hardware:incompletedesignsversusunderprovisionedcircuitry, becausetheirrespectivehandlingfromtestandreliability perspectiveisquitedifferent.Theremainderofthisletter is organizedasfollows. Testingof incompletedesignsandunderprovisionedcircuitsarediscussedinthenexttwosections,andreliabilityissuesofapproximatecircuitsarethe focusofSectionIV.SectionVconcludesthisletter.

II. TestingIncompleteDesigns

Inthisletter,"incompletedesign"referstoanapproximatecircuitwhichdeviatesfromitsspecificationinadeterministicandrepeatablemanner.(Incontrast,underprovisionedcircuitsmayfailoccasionally,governedbyarandomprocess.)Oneexampleisanadderwhich performscorrectadditionforallpossiblepairsofoperandsappliedtoitsinputsexceptasmallsetSofsuchoperandpairs.Thespecificationoftheincompleteaddermaydemandthatadditionworks"tosomeextent"(e.g.,with4-bitinsteadof32-bitprecision)evenforoperandpairsfrom*S*,oritmayallowarbitrarymisbehaviorforsuchpairs.Foraconventionaldigitalcircuit,onesingleerroneousbitthat was observed during testing is a sufficient reason to rejectthiscircuit,asitcontainsamanufacturingdefect. s, testingchecksthatamanufacturingdefect*d*turnedthecircuit's regular function f into f_d. Applying a test t can exposethedefectthroughthemeasurement$f_d(t) / f(t)$.Incaseofanapproximateblock,wemayoptagainstrejectingadefective

circuit as long as its function $f_d$still meets the loosened specification. In fact, one variant of AC, error-tolerance [5], [8],was chiefly motivated by improving manufacturing yield byretaining "good-enough" defective chips about which someguaranteescouldbemade.Forinstance,suchchipscouldbe sold for use in less-critical application, for a lower pricethanfullyfunctional,defect-freecircuits.Thisdifferentiationis somewhat related to "speed binning," where circuits man-ufactured in the same lot have slightly different performance(attainable clock frequency) due to random process variations.The speed of all circuits from the lot is measured, and fastercircuitsareusedforhigher-performanceproducts.

While the possible yield improvement is clearly beneficial,there are also strong arguments against approximate circuitshaving manufacturing defects. First, the test set applied dur-ing manufacturing test is rather compact and cannot cover thecomplete circuit functionality. Even if only acceptable behav-ior,withintheboundsoftheloosenedcircuitspecification,wasobserveduringtest,thereisnoguaranteethatthiswillstillbethecaseforarbitrarysequenceswhichwillbeappliedtothe circuit in application. This problem does not occur in con-ventional testing, because a circuit with any defect, however,small,isrejectedwithoutfurtherexamination.

Second, the defect can interfere with critical circuit infrastructure,suchasclockdistributionnetwork,powergrid,orvoltageandfrequencyscalingcontrol.Suchinterfer-encemaycauseminimalimpactduringtherathershorttest,butmaycompletelydisruptthecircuitoperationdur-ing its actual use. Third, many defects can deteriorate overtime and impact neighboring circuit structures due to factorslike electromigration [13] or causing excessive power-supplynoise [14]. Detrimental influence of infrastructure and agingalso occur for conventional, nonapproximate circuits, but thepresenceofatoleratteddefectmagnifiestheseissuesbyfar.

A. StructuralTestingforMagnitude-OrientedMetrics

One promising idea is the extension of structural, model-basedtesting into the AC domain.This idea,whichcan beseen as a generalization of threshold testing [8] to more gen-eral approximate blocks, is illustrated in Fig. 1. Let the circuitbespecifiedbyareferencebehaviormodel(e.g.,acircuitwith-out any approximations) and a maximal extent of deviation τ(threshold) according to an *error magnitude metric* Δ (e.g.,structural dissimilarity of images). The circuit is acceptable aslong as every value produced at its output deviates from thereferencebehaviorbyatmostτ.Considera(modelof)defect$d$inthe circuit.Createthreeblocks:oneforreferencebehavior(producing output o), one for circuit with defect d (producingoutput o_d), and one which calculates the difference $\Delta(o, o_d)$andcomparesitwithτ.

The construction of Fig. 1 can be reduced to an automatictest pattern generation (ATPG) instance (in this case, the threeblocks are represented by digital circuit models) or, equiv-alently, to a Boolean satisfiability (SAT) formula (thentheblocksaremappedtoconjunctivenormalforms).Running anATPGalgorithm,oranSATsolver, thefaultycircuitexceedsthemaximumallowed

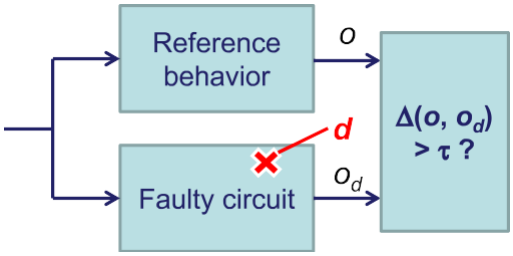

Fig.1. Model-basedtestgenerationinanapproximatecircuit.

deviation. If the analysis shows that no such t exists, this constitutes a formal proof that the circuit with defect d is stillacceptable. Note that threshold testing [8] is a special case,where the outputs are interpreted as numerical values witherror magnitude $\Delta(o, o_d)$ set to their difference$|o-o_d|$. The"miter circuit" construction used in conventional circuit test-ing is an even more restricted variant, with $\Delta(o, o_d)$ beingsimply1if$o_d$$\neq$$o$and0otherwise.

Two reasons prevent a wide-spread application of structuraltesting to approximate circuits. First, the need to model elabo-ratedeviationmetricsgreatlyincreasesthecomputationalcomplexity of analysis. Already the simple threshold testing [8]was associated with much larger run times than conventionaltest generation. Incorporating complex error magnitude met-rics, such as perceptual image or audio models, will make thetestgenerationintractable.Moreover,thecomplexityincreases further if the approximate block is embedded into a larger systemandthetestgeneratedfortheblockneedstobetransferredtothesystem'stoplevel.

Perhaps even more severe, it is impossible to imply theexact defect (or set of defects) present in a circuit from themeasurements performed during manufacturing test. Even ifwe know for some well-defined defects $d_1, \ldots, d_n$(e.g., somestuck-at or simple bridging faults) that a circuit with thesedefects is acceptable, we cannot be sure whether a specificcircuit instance has one of these defects or some other, per-haps unmodeled, defects. There are diagnostic methods whichderive likely defect locations from tester data, but they do notprovide guarantees and often have to be complemented byphysicalfailureanalysisinpractice.

B. TestingforError-RateOrientedMetrics

Anapproximatecircuitisacceptablewithrespecttoan*error-rate bound* ε[0, 1] if the fraction of its producedoutputs that deviate from the reference behavior is ε or less.Existing error-rate test approaches [18] make the assumptionthat the input distribution is uniform; then, the circuit withinput space I is acceptable if the input subspace S for which itdeviates from the reference behavior fulfills$S \leq I\varepsilon$. Error-ratetestgeneration[18]aimsatdeciding,foragivendefect$d$, whether a circuit affected by $d$will still satisfy the error-bound. This is done by generating a test set Tthat isindicative for the error-rate of the entire input space I(whichisnormallyexponentialinthenumberofcircuitinputs).In otherwords,$|S \cap T|/|T|$shouldapproximate$|S|/|I|$astight

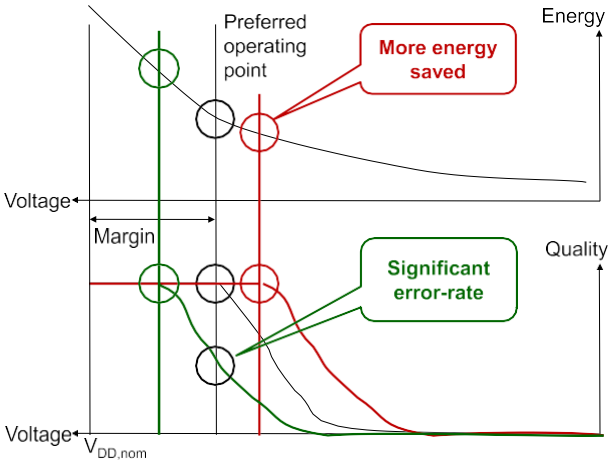

Fig.2.Effectsofvoltagescalingonenergyconsumptionandoutputqualityforafast(red),nominal(black),andslow(green)circuitinstance.

aspossible.Suchtestsetstendtobelargerthanconventionaldetectiontests,andtheyarevalidonlywithaconfidence.

Analternativeapproachtotheerror-ratetestingproblemistogenerateasetofacceptabledefects$D_{acc} = \{d_1, \ldots, d_n\}$andtoperformdiagnosisforaspecificfailingcircuit.IfthecircuithasoneofthedefectsfromD_{acc},itcanbeconsideredacceptabledespitebeingdefective.Problemslistedformagnitude-orientedmetrics(insufficientdiagnosticresolution,unmodeleddefects,andreliabilityimpact)stillapplyinthiscase.

III. TestingUnderprovisionedCircuits

Underprovisioned circuits are operated at voltage-frequencypoints for which their correct functioning is not guaranteed.Thisis illustrated inFig.2forthecaseofsupplyvoltageV_{DD}. Suppose that the circuit's V_{DD} is gradually reduced, i.e.,movedtotherightofitsnominalvalue$V_{DD,nom}$,whileitsclock frequency (cycle duration) is kept constant. The powerand the energy consumed by the circuit is reduced roughlyquadraticallywithV_{DD}decrease,asindicatedintheupperpart of Fig. 2. At the same time, the switching speed of logicgates reduces, and at some point one or multiple paths throughthe circuit will become slower than the cycle time and errorswilloccuroncircuitoutputs.

Circuits are usually designed with a margin, i.e., the difference between the cycle duration and the delay of the longestpath in the circuits exceeds zero. Because of natural processvariations, the margins are different for specific manufacturedcircuit instances. It is advantageous to operate the circuit withjust the minimal V_{DD}before failure, because this minimizestheenergyconsumption.Fig.2showsthepreferredoperat-ingpointsforthreecircuitsofdifferentspeed.Itisimportantto select individual operating points rather than one precalculated for the nominal circuit, because otherwise faster circuitscannot realize maximal energy benefits, and slower circuitsenter the regime beyond their margin and produce excessiveerror-rate.

Note that temperature changes and V_{DD} fluctuations mayshiftthecurvesfromFig.2dynamically.Thiscanbeaddressedbyperformingintrospectionatrun-timebydetectingerrorsandadapting(increasevoltageorreducefrequency)upontheir occurrence [9]. As an alternative, it is possible to avoidadaptation and continue operation in the beyond-the-marginregime [7]. However, in many practical circuits no guaranteesabouterrormagnitudesorratesinthisregimecanbemade.

Self-adapting underprovisioned approximate circuits poseconceptual challenges to testing. One key question is whetherthe self-adaptation features should be enabled or disabled dur-ingtest.Ifself-adaptationisenabledonacircuitwhichhasa small-delay defect just outside the normal extent of variabil-ity, it may compensate this defect by, e.g., increasing V_{DD} tothe maximum possible level, such that the circuit passes thetest. This circuit is then operated but has no resilience reservesand will fail during the next temperature increase or occur-rence of power-supply noise. If self-adaptation is disabled, thetestermayconfuse"regular"transientfailures,fromwhichitis designed to recover, with defects and reject circuits whichcouldbenormallyoperated,thisincreasingtheyieldloss.

Further issues in testing underprovisioned approximate circuitsareparameter-dependentdefectsanddefectsintheadaptation infrastructure itself. Parameter-dependent defectsaredefectswhichmanifestthemselvesonlyunderspecificvolt-age/frequencyconditions(they may be seen as irregularitieson what test engineers known as "Shmoo plots"). The mostobvioustechniquetoaddressthemistotestthecircuitundera number of different voltage/frequency combinations, which,however, increases test time, complexity, and cost. Defects inself-adaptation circuitry are difficult to observe during test,because they are disconnected from the functional part of thecircuit. Reproducing all possible self-adaptation scenarios dur-ing test is infeasible, and special design-for-testability featureswill be typically required. This problem generally arises forconventional circuits as well, but the comparatively large mar-gins of such circuits make them less prone to self-adaptationrelated failures than underprovisioned circuits and soften therequirementsontestprocedures.

It is logical to apply error-rate metrics discussed above tounderprovisionedcircuits.However,theirnondeterminismandcomplex, intermittent nature of failure patterns poses knownerror-rate testing concepts inapplicable to them. Any error-ratestatements for such circuits should perhaps be obtained bymore elaborate probabilistic characterization procedures, suchas tomographic testing originally developed for quantum andotherprobabilisticcircuits[12].

IV. ReliabilityAspectsof

APPROXIMATECIRCUITS

Like their conventional counterparts, approximate circuitsare potentially affected by failures that occur after manufac-turing and during the device's lifetime. Early life failures andwearout mechanisms could, e.g., slowly increase the circuit'serror rate from an acceptablevalue to one overthe thresholdε. For example, negative-bias temperature instability, or hot-carriereffectsmayincreasethelogicgatedelayswithinthe circuit, and electromigration could do the same with interconnect delays [13]. As a consequence, the circuit will exceed itscycledurationmoreoften.

Some of the specific aspects of approximate circuits makethem more and some less vulnerable to early life failures andwearout in comparison with regular circuits. Approximate cir-cuits often have some degree of error-resilience or gracefuldegradation, i.e., capability to operate (at least to some extent)in presence of errors. This feature increases their robustnessand, potentially, their lifetime, but may hinder error-detectionwhich would miss the buildup of many small, uncritical errorstoalargeandcriticalfailure.

The underprovisioned species of approximate circuits tendstobeoperatedundermorerelaxedoperatingpointsthanfunctionally identical conventional designs, reducing the probabilityofoverstressthatleadstoearlylifefailuresandexcessive aging. Atthesametime,theyareoperatedina marginal regime, with little or no noise margins. The smallestdisturbancewillleadtoafailure,andanunexpectedlylargedisturbance may exceed the capability of the self-adaptationinfrastructure to recover the circuit from the failure. It appearsessential that self-adaptation strategies and in particular intro-spectionroutinesaccountforthepossibilityofaginganddeteriorationduring the circuit's lifetime. For this purpose,theymayhavetobecombinedwithbuilt-inself-testandcharacterizationprocedureswhich notonly decidewhetherthecircuitisfaultybutalsodistinguishintermittenterrors(which the circuit is designed to tolerate) from permanentwearout-inducedfailures[16].

V. CONCLUSION

Approximate circuits must be manufactured, tested, and their post-production quality must be regulated if they are to be used. In the event of incomplete designs, conventional test methods might not fully grasp the potential of yield optimization, and they might not be accurate for circuitry that is underprovisioned. There are methods for testing circuits with looser specifications (concerning error magnitude or error rate), but they are often imprecise, predicated on faulty premises, and not very scalable. Instead of using simulations, there seems to be an urgent need for data obtained on real circuits produced in large volumes. What proportion of the flaws that were found were classified as "benign"? Did the tests designed for a basic fault model, such as stuck-at faults, accurately depict the real failures? In the context of reliability, the lack of data is even more significant.Approximate circuits need to be efficiently tested, which requires the solution of several unresolved issues. Automated proofs for approximation boundaries under the assumption of a complicated acceptability metric and test generation under such a metric are of essential interest.Better failure isolation and diagnosis during post-manufacturing testing must be made possible by design-for-testability techniques, and on-chip introspection circuitry must offer fine-grained circuit health monitoring spontaneously These characteristics already advantageous for conventional circuits, but they are crucial for approximation circuits as well. The additional expense of these solutions must be carefully weighed against the efficiency advantages brought about by the AC paradigm. Vendors of electrical design automation tools must undertake an economic analysis and a legal evaluation of the hazards associated with delivering circuits that are known to be defective in tandem with these investigations.

REFERENCES

[1] A. Agrawal *et al.*, "Approximate computing: Challenges and opportunities," in *Proc. IEEE Int. Conf. Rebooting Comput.*, San Diego, CA,USA,2016,pp.1–8.

[2] A. Alaghi, C. Li, and J. P. Hayes, "Stochastic circuits for real-timeimage-processingapplications,"in*Proc.DAC*,2013,Art.no.136.

[3] B.D.BrownandH.C.Card,"Stochasticneuralcomputation.I.Computational elements,"*IEEETrans.Comput.*,vol.50,no.9,pp.891–905,Sep.2001.

[4] N. P. Carter *et al.*, "Runnemede: An architecture for ubiquitous high-performance computing," in *Proc. IEEE HPCA*, Shenzhen, China, 2013,pp.198–209.

[5] I.S.ChongandA.Ortega,"Hardwaretestingforerrortolerantmultimedia compression based on linear transforms," in *Proc. IEEEDFTS*,Monterey,CA,USA,2005,pp.523–531.

[6] P. Gupta *et al.*, "Underdesigned and opportunistic computing in presenceofhardwarevariability,"*IEEETrans.Comput.-AidedDesignIntegr.CircuitsSyst.*,vol.32,no.1,pp.8–23,Jan.2013.

[7] R. Hegde and N. R. Shanbhag, "Soft digital signal processing," *IEEETrans. Very Large Scale Integr. (VLSI) Syst.*, vol. 9, no. 6, pp. 813–823,Dec.2001.

[8] Z.JiangandS.K.Gupta,"Thresholdtesting:Improvingyieldfornanoscale VLSI," *IEEE Trans. Comput.-Aided Design Integr. CircuitsSyst.*,vol.28,no.12,pp.1883–1895,Dec.2009.

[9] P. K. Krause and I. Polian, "Adaptive voltage over-scaling for resilientapplications,"in*Proc.DATE*,Grenoble,France,2011,pp.1–6.

[10] L. Leem, H. Cho, J. Bau, Q. A. Jacobson, and S. Mitra, "ERSA: Errorresilientsystemarchitectureforprobabilisticapplications,"in*Proc.DATE*,Dresden,Germany,2010,pp.1560–1565.

[11] J. Miao, A. Gerstlauer, and M. Orshansky, "Multi-level approximatelogicsynthesisundergeneralerrorconstraints,"in*Proc.ICCAD*, SanJose,CA,USA,2014,pp.504–510.

[12] A.Paler,A.Alaghi,I.Polian,andJ.P.Hayes,"Tomographictestingand validation of probabilistic circuit," in *Proc. IEEE ETS*, Trondheim,Norway,2011,pp.63–68.

[13] M. Pecht, R. Radojcic, and G. K. Rao, *Guidebook for Managing SiliconChipReliability*.BocaRaton,FL,USA:CRCPress,1999.

[14] I.Polian,"Powersupplynoise:Causes,effects,andtesting,"*ASPJ.LowPower Electron.*,vol.6,no.2,pp.326–338,2010.

[15] I. Polian, J. P. Hayes, S. Kundu, and B. Becker, "Transient fault characterization in dynamic noisy environments," in *Proc. ITC*, Austin, TX,USA,2005,pp.1048–1057.

[16] L.R.Gómez*etal.*,"AdaptiveBayesiandiagnosisofintermittentfaults," *J.Electron.Test.TheoryAppl.*,vol.30,no.5,pp.527–540,2014.

[17] A. Schöll, C. Braun, and H.-J. Wunderlich, "Applying efficient fault tolerancetoenablethepreconditionedconjugategradientsolveronapproximatecomputinghardware,"in*Proc.IEEEDFTS*,2016,pp.21–26.

[18] S. Shahidi and S. K. Gupta, "ERTG: A test generator for error-ratetesting,"in*Proc.ITC*,SantaClara,CA,USA,2007,pp.1–10.

[19] V.Tomashevich,C.Gimmler-Dumont,N.When,andI.Polian,"Reliability analysis of MIMO channel preprocessing by fault injec-tion," in *Proc. IEEE WISEE Conf.*, Noordwijk, The Netherlands, 2014,pp.1–6.

[20] R. T. Uppu, R. K. Uppu, A. D. Singh, and A. Chatterjee, "A highthroughput multiplier design exploiting input based statistical distribu-tion in completion delays," in *Proc. Int. Conf. VLSI Design*, Pune, India,2013,pp.109–114.

[21] J. Yoshida. *Race for AI Chips Begins*. EE Times,SanFrancisco,CA,USA,2016.[Online].Available:http://www.eetimes.com/document.asp?doc_id=1331052