SAND2016-1958C

Electronic Forensic Techniques for Manufacturer Attribution

Ryan L. Helinski*, Edward I. Cole Jr.*, Gideon Robertson*, Jonathan Woodbridge*, Lyndon G. Pierson
*Sandia National Laboratories
Albuquerque, New Mexico 87123
E-mail: {rhelins, coleei, garobe, jwoodbr}@sandia.gov

***Abstract*—The microelectronics industry seeks screening tools that can be used to verify the origin of and track integrated circuits (ICs) throughout their lifecycle. Embedded circuits that measure process variation of an IC are well known. This paper adds to previous work using these circuits for studying manufacturer characteristics on final product ICs, particularly for the purpose of developing and verifying a signature for a microelectronics manufacturing facility (fab). We present the design, measurements and analysis of 159 silicon ICs which were built as a proof of concept for this purpose. 80 copies of our proof of concept IC were built at one fab, and 80 more copies were built across two lots at a second fab. Using these ICs, our prototype circuits allowed us to distinguish these two fabs with up to 98.7% accuracy and also distinguish the two lots from the second fab with up to 98.8% accuracy.**

I. Introduction

Due to the globalization of the integrated circuit (IC) manufacturing industry, supply chain integrity is a major concern [1], [2], [3]. Hazards with components that result from a supply chain may include counterfeiting and subversion. Counterfeit parts may have been "knowingly misrepresented"; they may have been marked to suit the customer's order, but are may be from a different lot, a different manufacturer, may have a different qualification grade (e.g., commercial or industrial rather than military) and may have been salvaged from another system. Subverted parts may have been modified for a malicious purpose during some part of the manufacturing process (known as hardware Trojans) or have been wholly substituted with a "compatible" part. These modifications could be made with or without changes to the IC layout and could be used to transmit information, modify a part's specifications (e.g., operating temperature range) or modify its logical function [4].

The focus of this work is on characterizing and identifying the fab of origin of ICs by enabling variations in manufacturing characteristics from one or more fabs to be measured and anomalies to be detected. This may provide some quantitative level of confidence that an IC originated from the expected manufacturer and perhaps identify the production lot. We aim to address hazards in the supply chain between the fabrication and test steps of a product's life cycle. Onward, we will refer this concept as electronic circuits for forensic, non-destructive attribution (FORAB).

We designed and built three batches of 350nm technology ICs, one batch of 79 at one fab and two batches of 40 at a second fab, to test this concept. The two fabs, which we will keep anonymous and refer to only as Fab A and Fab B, are in two separate geographic locations. We describe the design of this test IC in Section IV and evaluate the usefulness of each technique in discriminating manufacturers and production lots using statistical methods, estimation and detection theories, and machine learning techniques in Section V.

II. Related Work

Physical Unclonable Functions (PUFs) leverage manufacturing process variations of built-in circuitry to generate responses that are unique for each IC, among other applications. Examples of circuits which have been used to build PUFs include ring oscillators, programmable delay lines, and capacitance and transistor saturation current mismatch [5], [6], [7], [8]. The ideal output of a PUF is a string (or strings) of uniformly distributed random bits PUF outputs so that each IC has a unique function. PUFs can mitigate hazards in the supply chain by comparing measurements between the trusted distributor or authorizing agent (test) and the customer (deployment through maintenance). However, PUF responses are not necessarily useful for understanding the inherent process variations or attributing an IC to its manufacturer.

The use of variation in path delay through functional circuits and statistical techniques to identify the fab of origin have been proposed previously [9]. In this paper, we add to this work by proposing that multiple circuits that are each affected by distinct process variations be used.

III. Distinguishing Fabs

In practice, FORAB could use a sample of "known good" ICs to establish the characteristics of a manufacturer or the profile of a manufacturer could be built up over time with assumed good ICs. This does not depend on a so-called "golden" standard or chip. It is assumed that lot-to-lot variations are at least as significant as chip-to-chip variations. Likewise, fab-to-fab variations should be at least as significant as lot-to-lot. i.e, the standard deviations are monotonically increasing $\sigma_{\text{inter-wafer}} \leq \sigma_{\text{inter-lot}} \leq \sigma_{\text{inter-fab}}$. Under this assumption, parametric measurements from ICs may enable both fabs and lots to be distinguished. Furthermore, consider some parametric value P that can be measured with on-chip circuitry. In practice, two manufacturers A and B will produce ICs that have a slightly different mean and standard deviation

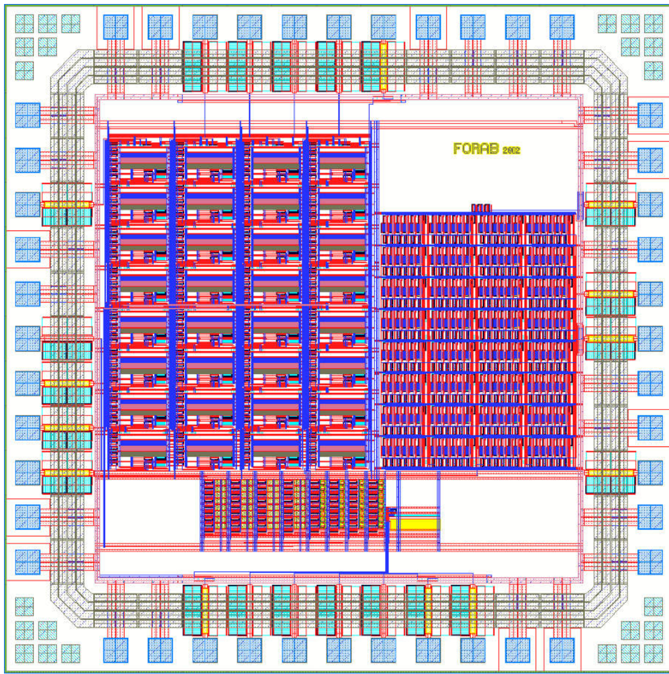

Fig. 1. Forab IC layout

for P. We choose a threshold P_{discr} on P to discriminate between manufacturers A and B. Unless the mean of P is very different relative to the standard deviation between the two manufacturers, we will get a large fraction of mislabeled ICs with a simple threshold method. To overcome this, we combine many different parametric measurements. Assuming that the means of the various parametric distributions for two specific lots from two fabs are normally distributed, the mean for fab A would be greater than that for fab B for approximately half of the parametrics. For example, we can compare P_0 and $P_{\mathrm{thres},0}$, P_1 and $P_{\mathrm{thres},1}$, P_2 and $P_{\mathrm{thres},2}$ and so on. Therefore, for each distinct manufacturing process parametric that is added, we can increase the amount of theoretical information we can gather about the fab of origin and reduce the classification error.

IV. Test IC Design

In this section, the design of the test IC shown in Figure 1, which we will call `forab`, is described. Three types of circuits were included: resistors (`meas_res`), capacitors (`oscint`) and ring oscillators (`p3`). We describe each of these in the following subsections. In each case, we included several implementations of the circuit. We were unable to obtain process variation information about the particular manufacturing process we used. This information could have included Monte Carlo simulation models that would have allowed us to predict how certain structures would vary when they were manufactured.

A. `meas_res` Resistor Circuits

In this section of the IC, we included resistors to be measured with a four-wire (Kelvin) sensing method. In this sensing method, four probes are connected to the resistor to be measured. Two probes are used to force a current and the other two are used for measuring the voltage. This scheme is a method of eliminating the effect of the intrinsic probe resistance of the probes used to measure current. NMOS and PMOS transistors were also included to measure their source-drain resistance with the gate voltage being supplied by a fifth pin. We connected each resistor to a hierarchical network of transmission gates so that the probe pins could be shared on the IC. With enough transmission gates connected in parallel, the off resistance begins to significantly affect the measurement. The network for each of the 4 probes is a hierarchy that branches our from the IC pin 4 ways, then 8 ways, then 8 ways again before it is connected to a single resistor to be measured. In this way, stacking the transmission gates increases the effective off resistance from the chip pins to "off" resistor test cells by orders of magnitude.

The types of resistors are as follows. `rm1` through `rm4` are the first through fourth metal layers. `rpo1` is a polysilicon layer used for routing. `via1` is a vertical serpentine structure which changes many times between the first and second metal layers utilizing a large number of vias. `nch` and `pch` are chains of 167 stacked transistors. 32 of each of type of resistors was included. The resistors were designed to give at least 500Ω of resistance so that they could be easily measured in the lab with a Keithley 2400 source meter. However, they could have been made smaller and measured accurately with more sensitive equipment.

B. `oscint` Capacitor Circuits

The next type of device we wanted to measure was the capacitor. There are many ways to measure on-chip capacitance[10]. Our implementation is based on a technique used to measure current from a radiation sensor[11] which uses a current source to charge or discharge a selected capacitor, in what is called an integrating oscillator (`oscint`). The capacitor voltage is monitored with a comparator. When the voltage is less than a particular threshold (greater than V_{SS}), the circuit begins to charge the capacitor. Likewise, when the voltage rises above a second threshold (less than V_{DD}), the circuit begins discharging the capacitor again. The state of this circuit is a square wave (charging or discharging). Assuming that the connected capacitance is significantly larger than the intrinsic capacitance of this circuit, the frequency of that square wave is a strong function of the connected capacitance. Measuring an absolute capacitance requires calibrating the operating frequency of this circuit to a known capacitor value. However, calculating the absolute magnitude of the capacitor values is not necessary because our objective is to characterize only the relative manufacturing variations in these capacitors.

There are two types of capacitors available in the process we used: PIP and MOS. The PIP capacitors are specialized polysilicon-insulator-polysilicon capacitors. The MOS capacitors are an NMOS transistor with the gate serving as one terminal, and the source and drain are connected together to serve as the other. 32 of each of these types of capacitors

were included. As reported later, we made measurements that we call the "intrinsic", "intrinsic MOS" and "intrinsic PIP" in addition to the basic measurements. These "intrinsic" measurements are made with no transmission gates switched on to connect the integrating oscillator circuit to a specific capacitor. The "intrinsic MOS" and "intrinsic PIP" measurements are made with the transmission gate network partially switched on, connecting the oscillator to only the bank of transmission gates near the MOS and PIP capacitors, respectively, but not to a specific capacitor. In these intrinsic configurations, the circuit is essentially a free-running oscillator.

C. *p3 Ring Oscillator Circuits*

`p3` is the third and final circuit on the IC. This circuit uses nothing more than the features provided by the IC manufacturing process and only digital input and output. The novel concept is to start with a base layout of a cell used for measuring path delay. This cell has been scaled up in width and length. Then, we derive various "flavors" of this cell by reducing one specific feature at a time to minimum size. In the base cell, the effects of manufacturing process variations are reduced significantly. For example, metal line edge roughness has less effect on a metal interconnect as the interconnect is widened. The dimensions of the various layers affect manufacturing process variations differently, but there is a general trend that scaling up reduces variation in the circuit's performance. By reducing a feature to minimum size, both the effect of that feature and its manufacturing process variations on the circuit's performance become more significant. We chose a five-stage ring oscillator (RO) as the basis of the design because the speed of the ring oscillator is dependent upon the transistor performance, metal and via resistance, and various intrinsic capacitances.

Our "nominal" ring oscillator called `nom` which was scaled up by $2\times$. From this, we derived 4 other flavors: `cc`, `m1`, `poly` and `m2`. In the `cc` flavor, the density of contacts between the silicon substrate and the first metal layer is reduced by a factor of 3. In the `m1` flavor, the first metal layer interconnects are reduced to minimum width including a serpentine. In the `poly` flavor, the polysilicon interconnects are reduced to minimum width. Finally, in the `m2` flavor, the width of a serpentine interconnect in the second metal layer which connects the output of the RO back to the input is reduced to minimum width. To support and balance the `m1`, `poly` and `m2` flavors, serpentine patterns in the corresponding layers were added beforehand in the `nom` design.

V. RESULTS

Our test IC, described in Section IV, consists of 512 measurable circuits: 256 resistors of 8 types (`meas_res`), 64 capacitors of 2 types (`oscint`) and 192 ring oscillators of 6 types (`p3`). In this section, we present example distributions of the circuits we measured, metrics on the differences between these distributions and the results of using a machine learning algorithm to predict the accuracy with which the fab of origin of an IC could be classified.

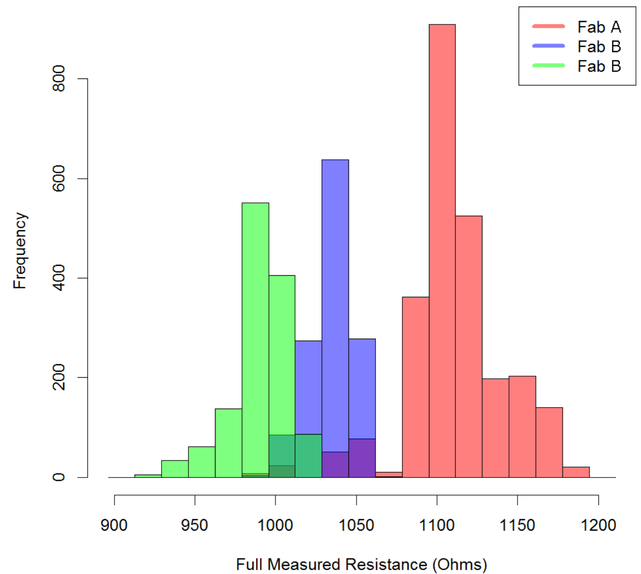

Fig. 2. Histogram of poly-1 resistance measurements, colored by lot.

As mentioned in Section I, we had three batches of chips made at Fabs A and B. We will refer to the lot from the first fab as "Lot 1" and the two lots from the second fab as "Lot 2" and "Lot 3". There are 80 chips in Lot 1, 40 chips each in Lots 2 and 3; 160 chips total. One of the chips from Lot 1 was defective, and the results from it are not included. Lots 1 and 2 were used to quantify inter-fab variations. Lots 2 and 3 were used to quantify inter-lot variations, serving as our experimental control.

A. *Distribution Separation Metrics*

If we plot the histograms for a specific type of circuit for the three lots, we hope to see the the two Fab B lots look similar and the Fab A lot look different. If all the lots have similar distributions, or if the two lots from Fab B look more different than the inter-fab difference, then that circuit type may not be useful for identifying the fab of origin. Figure 2 shows an ideal set of histograms. The histogram bars are drawn semi-transparent so that the overlap can be seen. The distributions of the two lots from Fab B look different, but not as different as either of these lots compared to the lot from Fab A. Also, the populated bins of the lots from Fab B are almost distinct from the those of the lot from Fab A. For some of the other circuit types, the histograms were less than ideal. For example, sometimes there was a large overlap between the distributions from the two lots from Fab B, and sometimes all of the distributions looked essentially the same.

To quantify the difference between two distributions, we define a fitness function as $f(X_a, X_b)$ which takes two random variables X_a, X_b and returns a number between 0 and 1. 0 represents the worst case where the two distributions overlap exactly, and 1 represents the best case where there is perfect separation. We considered two such fitness functions.

TABLE I
DISTRIBUTION SEPARATION METRICS FOR LOT 1 VS. 2

Measurement Path	20-bin Histogram Fitness	Bhattacharyya Distance
meas_res/rm1	0.88	0.96
meas_res/rm2	0.76	0.88
meas_res/rm3	0.70	0.85
meas_res/rm4	0.95	0.43
meas_res/rpo1	0.92	0.38
meas_res/via1	0.94	0.24
meas_res_mos/nch Vg=0.0	0.99	0.62
meas_res_mos/nch Vg=1.65	0.74	0.97
meas_res_mos/nch Vg=3.3	0.86	0.96
meas_res_mos/pch Vg=0.0	0.80	1.00
meas_res_mos/pch Vg=1.65	0.81	1.00
meas_res_mos/pch Vg=3.3	0.99	0.46
oscint/pip	0.91	0.54
oscint/mos	0.90	0.81
oscint/intrins	0.88	0.93
oscint/intrins pip	0.91	0.57
oscint/intrins mos	0.93	0.52
p3/nom	0.88	0.97
p3/cc	0.88	0.93
p3/m1	0.88	0.97
p3/m2	0.88	0.95
p3/poly1	0.95	0.81
Average	0.884	0.735

First, the 20-bin histogram fitness function f_{h} measures how different two histograms appear, and is defined by

$$f_{\mathrm{h}}(X_a, X_b) = \sum_{i=0}^{n} \left(1 - \frac{\max_i(c_a(i), c_b(i)) - |c_a(i) - c_b(i)|}{\max_i(c_a(i), c_b(i))} \right)$$

where $c_a(i), c_b(i)$ are the ith histogram bin counts of distributions a and b, respectively. This represents the sum of the overlaps of each histogram bin for the two distributions. We could have also used the Wasserstein metric (a.k.a. the Earth Mover's Distance).

Second, the Bhattacharyya Distance (BD) fitness function f_{BD} is defined as

$$f_{\mathrm{BD}}(X_a, X_b) = 1 - \frac{2 \arctan\left(\mathrm{BD}(\mu_a, \mu_b, \sigma_a, \sigma_b)\right)}{\pi},$$

where μ_a, μ_b are the means, and σ_a, σ_b are the standard deviations, of distributions a and b, respectively. The BD measures the similarity of two Gaussian probability distributions. However, many of our measurements have multi-modal distributions even on the same chip.

These two metrics are shown in Table I for each type of circuit. These metrics help us evaluate the circuit types among each other, but are not conclusive because we have nothing to which to compare them. For this reason, we evaluate the accuracy with which we can predict the fab given the measurements from a particular IC in Section V-C.

B. Distribution Overlap Errors

The "intersect" is the point where two estimated Gaussian PDFs intersect. This point can be used to create the threshold for a classifier. In the case that the mean for class a is less than the mean for class b, $\mu_{\mathrm{a}} < \mu_{\mathrm{b}}$, this classifier would label values falling below the intersect as class a, and labels those falling above the intersect as class b. In the case that $\mu_{\mathrm{a}} > \mu_{\mathrm{b}}$, the opposite applies.

The Gaussian PDFs are estimated for each circuit class using a statistical technique called bootstrapping which can improve the stability and accuracy of classifiers [12]. Given a training set of samples D of size n, bootstrapping generates m new training sets each of size $n' \leq n$ by sampling from D using a uniform random distribution and with replacement. In this way, a model of the PDF can be created for each of the m new training sets and these model PDFs can be averaged together to produce a more accurate overall PDF model. In our analysis, we used bootstrapping to estimate our PDFs based on 10,000 sets of samples of the same size as our original set.

Using these model PDFs for the two parametric distributions, the intersection is computed by equating the PDF functions and solving for the x-axis value τ which makes them equal. Given the intersect τ, the misclassification error areas can be estimated using either Equation 1 or 2, depending which of the sample means, μ_a and μ_b, is the lesser. $P(a \rightarrow b)$ is the probability that a sample from fab a is mis-labeled as a sample from fab b. Likewise, $P(b \rightarrow a)$ is the probability that a sample from fab b is mis-labeled as a sample from fab a. Note that $P(a \rightarrow a) + P(a \rightarrow b) = 1$ and $P(b \rightarrow a) + P(b \rightarrow b) = 1$. With these two overlap errors computed, they can be combined using the average to yield an overall classification error because the total area under each PDF is equal to 1. We will report each of these errors as the "percent incorrect".

$$P(a \rightarrow b) = \begin{cases} 1 - F_{\mu_a,\mu_a}(\tau) & \text{if} \mu_a < \mu_b \\ F_{\mu_b,\sigma_2}(\tau) & \text{if} \mu_a > \mu_b \end{cases} \quad (1)$$

$$P(b \rightarrow a) = \begin{cases} F_{\mu_b,\sigma_2}(\tau) & \text{if} \mu_a < \mu_b \\ 1 - F_{\mu_a,\sigma_1}(\tau) & \text{if} \mu_a > \mu_b \end{cases} \quad (2)$$

Recall that many copies of each measurement circuit are included on each IC so that the effects of manufacturing process variation on each circuit can be characterized. In our analysis, we also use subsets of these copies in order to draw various conclusions. We use the mean of each type of circuit to minimize intra-chip manufacturing process variations and assess the best case for each circuit to detect fab-to-fab differences. We also used the first 3 and a random set of 3 of each circuit type. We used these two analyses to assess the performance of the circuits if a minimal number of copies of each circuit type was included on the IC. In our measurement sequence, the first 3 of each circuit type are near one another. We used this to assess the performance of the circuits to detect fab-to-fab differences if the area of the chip to which FORAB is constrained is small and contiguous. We used the a random subset of 3 circuits of each type in order to assess the detection performance of the circuits when they are distributed widely across the IC. To prevent bias from a specific choice of the random subset, we used the average performance of 10 random subsets for this analysis. We will call these analyses the "all", "mean", "first 3" and "random subset of 3" below.

For brevity, we present the complete intersect and percent incorrect results for only the "mean" analysis in Table II,

TABLE II
PDF INTERSECTION AND PERCENT CLASSIFIED INCORRECT WHEN THE MEAN OF EACH TYPE OF CIRCUIT IS USED.

Circuit Class	Intersect	Percent Incorrect Lots 2&3	Lot 1	Average
oscint/cap8_pip	2.19e+07	2.972	4.639	3.805
oscint/cap32_mos_intrin	5.17e+07	7.729	6.246	6.988
oscint/cap32_pip_intrin	5.19e+07	12.970	9.286	11.128
oscint/cap8_mos	1.62e+07	5.052	25.580	15.316
meas_res_blk/pch	1.20e+06	47.095	1.757	24.426
meas_res_blk/nch	1.23e+06	51.369	4.472	27.920
p3/poly1	4.36e+06	2.456	58.485	30.470
oscint/intrinsic	6.95e+07	49.843	15.762	32.803
p3/cc	8.23e+06	60.203	10.860	35.531
p3/m2	8.27e+06	20.929	53.349	37.139
p3/nom	8.29e+06	50.506	28.589	39.547
meas_res_blk/rm3	1.22e+04	67.290	15.247	41.268
meas_res_blk/rpo1	1.23e+04	67.314	16.882	42.098
meas_res_blk/rm4	1.71e+04	68.291	17.702	42.996
meas_res_blk/via1	1.90e+04	69.910	18.785	44.347
p3/m1	8.33e+06	77.425	11.427	44.426
meas_res_blk/rm1	1.19e+04	71.794	18.295	45.045
meas_res_blk/rm2	5.09e+02	29.508	61.682	45.595

which is sorted from best to worst by the average percent incorrect. The PIP capacitors described earlier performs the best in this analysis, the MOS intrinsic measurements rank second and third. The rankings for the "first 3" and "random 3" analyses are similar.

C. Machine Learning Predictions

In this section, we take the analysis further and apply machine learning techniques to estimate the fab prediction error based on the measurements we made of our test ICs.

k-fold cross-validation is a technique for estimating how accurately a predictive model will perform in practice. In this technique, a sample is randomly partitioned into k equal-size subsamples. One of the k subsamples is kept as a validation sample and the rest are used as training data. To reduce variability and increase the confidence of the result, the cross-validation is performed k times with each of the k subsamples being used as the training data exactly once and the results are averaged over k rounds of cross-validation. In the following results, we performed 10-fold cross-validations and used Freund & Schapire's Adaboost M1 method [13]. Note that cross-validation can overestimate the prediction accuracy, also that threshold-based classifiers do not take into account the apparent variation (the width of the distributions) of multiple measurements on a single IC.

We present our classification accuracy results in a confusion matrix which shows how each IC from the two samples were classified by listing the predicted class in the columns and the true class in the rows. Ideally, the number of ICs in each sample would appear along the diagonal of this 2-by-2 confusion matrix, and zeros would appear in the other positions. This would indicate that all of the ICs from sample a and b were classified correctly. Instead, some ICs from sample a will be classified as being from sample b ($a \rightarrow b$), and some ICs from sample b will be classified as being from sample a ($b \rightarrow a$).

TABLE III
CONFUSION MATRICES FOR A THE "MEAN", B THE "FIRST 3", C THE "RANDOM SUBSET OF 3" (PERFORMED 10 TIMES) AND D THE "ALL" ANALYSES FOR THE RESPONSES OF EACH CIRCUIT TYPE.

a	b	← classified as
79	1	a = Lot 2+Lot 3
3	76	b = Lot 1
Correct		97.5%

(a)

a	b	← classified as
80	0	a = Lot 2+Lot 3
2	77	b = Lot 1
Correct		98.7%

(b)

a	b	← classified as
765	35	a = Lot 2+Lot 3
35	755	b = Lot 1
Correct		95.6%

(c)

a	b	← classified as
79	1	a = Lot 2+Lot 3
1	78	b = Lot 1
Correct		98.7%

(d)

TABLE IV
INTER-LOT CONFUSION MATRIX FOR A THE "MEAN", B THE "FIRST 3", C "RANDOM SUBSET OF 3" (PERFORMED 10 TIMES) AND D THE "ALL" ANALYSES FOR THE RESPONSES OF EACH CIRCUIT TYPE.

a	b	← classified as
40	0	a = Lot 3
1	39	b = Lot 2
Correct		98.8%

(a)

a	b	← classified as
38	2	a = Lot 3
1	39	b = Lot 2
Correct		96.2%

(b)

a	b	← classified as
386	14	a = Lot 3
38	362	b = Lot 2
Correct		93.5%

(c)

a	b	← classified as
39	1	a = Lot 3
1	39	b = Lot 2
Correct		97.5%

(d)

First, Table III(a) shows the confusion matrix for the "mean" analysis, where the mean of each circuit type is used. The fraction of ICs which were correctly classified is 97.5%. Next, the confusion matrix for the "first 3" analysis is shown in Table III(b), representing a minimal number of circuits of each type packed closely together in the design. As we expect, using only three circuits of each type decreases the accuracy, but only slightly to 98.7%. Third, the average result for the "random 3" analysis is presented in Table III(c). Note that there are $10\times$ as many samples in this case because we repeated the classification with 10 different random subsets. Similar to the first-3 subset, the accuracy decreases to 95.6%. Finally, Table III(d) is the confusion matrix when we present all of the circuits of each type to the classification tool. We might expect this classifier to be the best, but this is not necessarily the case because including more of the repeated copies of each circuit type and treating them independently results in added noise. The result in this case is the same as the "first 3" analysis.

We can also quantify the accuracy with which lots 2 and 3 from Fab B can be identified. The corresponding confusion matrices for the mean, first 3, random 3 and all of each circuit type are shown in Tables IV(a), IV(b), IV(c) and IV(d), respectively. As we expected, these tables show that the measurements are not as good at distinguishing the fabs from each other, but are still all above 93.5% correct.

VI. DISCUSSION

Our proof of concept was built in one of the oldest, most mature technologies available, 350nm, which was first released

in 1995. As a result, there was a chance that all the fabs offering this process would have exactly the same characteristics and practically no lot-to-lot variation. Instead, we found that these variations are easily measured. It is not clear if the result would be the same on newer process technologies. On one hand, the level of process variations would be increased, but, on the other, the number of equipment providers is be smaller.

Our goal was to distinguish two fabs. The problem of blindly identifying the fab of an IC is different and might be more difficult without characterizing a large number of fabs.

The measurements presented in this paper were all made at controlled room temperature (25°C). Many of the circuits we have shown are affected by temperature and voltage variations, and therefore these factors may have to be controlled or accounted for in practice.

These circuits could be integrated into a functional design with low overhead. Assuming that a counter and fixed-frequency oscillator are already available, the `p3` circuits, for example, can be included with as little at 5% of a 3mm x 3mm die or less than 1% of a 7mm x 7mm die. The overhead for the `p3` circuits will be less for smaller technology nodes. Our exhaustive test suite took approximately 13 minutes per chip to run. This included repeating each measurement 10 times for noise reduction and fault characterization. With a generous integration time, each `oscint` and `p3` measurement took about 0.125 seconds to collect. This also includes time for an IDDQ measurement. The `meas_res` circuits were slower to measure at 0.190 seconds per measurement because we wanted to give the DC measurement time to settle.

VII. Conclusion

We have shown that it is feasible to reliably identify the fab of origin of an IC by adding special circuits to its design. Using all of the circuits we included, we could distinguish the fab with 98.7% accuracy and the lot with 98.8% accuracy. In both fab-to-fab and lot-to-lot comparisons, the random subset of 3 performed the worst. Recall that this analysis was to determine the performance if the circuits of a single type were sparsely populated on the IC. Therefore, we conclude that there is no need to distribute the copies of each circuit type in our case. Some of these circuits required external analog measurement apparatus. If we use only the `oscint` or `p3`, then the only instrumentation that would be required would be a crystal oscillator. Note that more "flavors" of `p3` ring oscillator circuits could easily be added to the 5 we used in order to improve accuracy. Finally, note that these results may be the result of chance because we only built three lots of chips. It is possible that the results may not extend to other process technologies, or may not even be repeatable for the technology we chose. Further investigation is necessary to answer these questions.

Acknowledgment

The authors would like to thank the Lab-Directed Research & Development (LDRD) program at Sandia National Laboratories. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

References

[1] United States Defense Science Board Task Force, "High performance microchip supply," 2005.
[2] United States Defense Advanced Research Projects Agency (DARPA), "Baa06-40," 2006.
[3] J. I. Lieberman, "White paper: National security aspects of the global migration of the u.s. semiconductor industry," 2003.
[4] M. Tehranipoor and F. Koushanfar, "A survey of hardware trojan taxonomy and detection," *Design Test of Computers, IEEE*, vol. 27, no. 1, pp. 10 –25, jan.-feb. 2010.
[5] G. E. Suh and S. Devadas, "Physical unclonable functions for device authentication and secret key generation," in *Proceedings of the 44th Annual Design Automation Conference*, ser. DAC '07. New York, NY, USA: ACM, 2007, pp. 9–14. [Online]. Available: http://doi.acm.org/10.1145/1278480.1278484
[6] M. Majzoobi, F. Koushanfar, and S. Devadas, "Fpga puf using programmable delay lines," in *Information Forensics and Security (WIFS), 2010 IEEE International Workshop on*, Dec 2010, pp. 1–6.
[7] D. Roy, J. Klootwijk, N. Verhaegh, H. Roosen, and R. Wolters, "Comb capacitor structures for on-chip physical uncloneable function," *Semiconductor Manufacturing, IEEE Transactions on*, vol. 22, no. 1, pp. 96–102, Feb 2009.
[8] M. Kalyanaraman and M. Orshansky, "Novel strong puf based on nonlinearity of mosfet subthreshold operation," in *Hardware-Oriented Security and Trust (HOST), 2013 IEEE International Symposium on*, June 2013, pp. 13–18.
[9] J. B. Wendt, F. Koushanfar, and M. Potkonjak, "Techniques for foundry identification," in *Proceedings of the 51st Annual Design Automation Conference*, ser. DAC '14. New York, NY, USA: ACM, 2014, pp. 208:1–208:6. [Online]. Available: http://doi.acm.org/10.1145/2593069.2593228
[10] Skoric, Schrijen, Ophey, Wolters, Verhaegh, and van Geloven, "Experimental hardware for coating pufs and optical pufs," 2007.
[11] M. R. Shaneyfelt, T. A. Hill, T. M. Gurrieri, J. R. Schwank, R. S. Flores, P. E. Dodd, S. M. Dalton, and A. Robinson, "An embeddable soi radiation sensor," *Nuclear Science, IEEE Transactions on*, vol. 56, no. 6, pp. 3372–3380, December 2009.
[12] B. Efron and R. Tibshirani, "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," *Statistical Science*, vol. 1, no. 1, pp. pp. 54–75, 1986. [Online]. Available: http://www.jstor.org/stable/2245500
[13] Y. Freund, R. Schapire, and N. Abe, "A short introduction to boosting," *Journal-Japanese Society For Artificial Intelligence*, vol. 14, no. 771-780, p. 1612, 1999.