Test Time Reduction in Analogue/Mixed-Signal Devices by Defect Oriented Testing: An Industrial Example

Hamidreza Hashempour† Jos Dohmen† Bratislav Tasić† Bram Kruseman†
Camelia Hora† Maikel van Beurden† Yizi Xing‡

Abstract

We present an application of Defect Oriented Testing (DOT) to an industrial mixed signal device to reduce test time and maintain quality. The device is an automotive IC product with stringent quality requirements and a mature test program that is already in volume production. A complete flow is presented including defect extraction, defect simulation, test selection, and validation. A major challenge of DOT for mixed signal devices is the simulation time. We address this challenge with a new fault simulation algorithm that provides significant speedup in the DOT process. Based on the fault simulations, we determine a minimal set of tests which detects all defects. The proposed minimal test set is compared with the actual test results of more than a million ICs. We prove that the production tests of the device can be reduced by at least 35%.

1 Introduction

Test cost for analogue and mixed-signal (AMS) products is a growing concern within the industry [1]. One of the reasons is the long test times due to the growing complexity in combination with specification (or performance) based testing. Together with expensive instruments, these time consuming tests take a larger fraction of the production cost of an IC [2][3]. Techniques such as multisite testing have been applied with success to reduce test time by sharing instruments among multiple devices under test (DUT), however, they are running out of gas and novel directions need to be explored.

While structured test methods have become common practice in digital ICs, testing AMS ICs in a structured way is still in infancy [4]. Lack of fault models for AMS ICs renders fault oriented approaches difficult and sometimes fruitless. AMS DFT and BIST techniques are also lagging. According to [3] "no proven alternative to performance-based analogue testing exists and more research in this area is needed".

1.1 Previous works

DOT has been proposed as a structured test method for AMS ICs. While details vary slightly, it generally includes defect extraction, fault modeling, fault simulation, and test application. Defects are either extracted from layout by Inductive Fault Analysis (IFA) [5] or by processing netlist or schematic. Fault models range from simple resistive models for bridge and open defect mechanisms [6] to more sophisticated models such as S-parameters for frequency domain analysis [7]. Fault simulation involves injecting faults into the netlist of the AMS IC and performing analogue simulation (DC, AC, or Transient).

DOT has not been widely used though, a major drawback with DOT is the long time associated with analogue fault simulation. A technique for trading accuracy for speedup is presented in [2] in which a mixed-mode co-simulation of HDL code and SPICE models has been proposed for fault simulation. A similar technique that uses high level behavioral modeling is presented in [8]. Fast fault simulation of analogue circuits using specialized analogue simulators have been intensely studied in the past. Analogue fault simulation in linear circuits is studied in [9][10]. The work of [11] presents an example of fault simulation in nonlinear circuits when DC point of the faulty and fault-free circuits are equal. A relaxation technique for fast calculating DC point of nonlinear faulty circuits is discussed in [12]. Examples of fast fault simulators for transient analysis in linear circuits are DRAFTS and FLYER [13][14]. Fast fault simulation for transient analysis in nonlinear circuits has also been proposed [15][16].

Computer aided test environments for AMS testing have also been the subject of research and development. They encapsulate tools (e.g., fault extractor or analogue simulator) and facilitate execution of the DOT flow [17][18]. We use NXP’s internal tool DOTSS (defect oriented test simulation system) in this work [19]. Reducing production test size has been studied in [20] where analogue fault mod-
eling is discussed followed by an analysis of specification tests and algorithms for fault-driven test selection. Reference [21] highlights redundancy within the performance tests of an RF device and proposes prediction models based on genetic algorithms to select a subset of tests.

1.2 Contribution

In this paper, we present a test time reduction technique and its application to an AMS IC that is already in volume production. The device is a CAN (controller area network) transceiver manufactured in a BiCMOS-DMOS (BCD) process. It is an automotive product and has a zero-defect quality requirement.

The proposed technique utilizes DOT; it extracts and simulates defects using production tests. The DUT netlist has 31,587 elements of which 16,859 are transistors. Although it is small compared to large System-on-Chip designs, it is quite challenging from a defect simulation point of view. For example, if simulating hundreds of tests takes one hour, it still takes several years to perform these simulations for 10,000 defects. We propose a novel analogue fault simulation algorithm that reduces time to weeks or days. Based on the simulation results, we determine a subset of tests that covers all detectable defects. This approach can be used on new DUTs to estimate and improve fault coverage as well as existing DUTs to determine a minimal test set and improve test time. We apply the latter approach. To validate that the minimal subset performs well, the production test data logs of 10 lots (over a million DUTs) are used to evaluate how many failing dies would have escaped if the selected tests were used instead of the original (full) set. We discover that the production tests can be reduced by 35% while maintaining quality. To our best knowledge, this is the first publication in which the DOT method is applied to a volume production AMS IC.

The DOT flow is presented in Figure 1. It has 4 main steps, defect extraction, defect simulation, test selection, and validation.

![Figure 1: Basic steps in the DOT flow.](image)

The rest of this paper is organized as follows: Section 2 presents details of the defect extraction in the DOT flow. Section 3 discusses the details of our fast simulation algorithm. Section 4 presents details of the test selection including combinatorial formulation and the utilized heuristics. Section 5 presents how the selected tests are validated by the production test data followed by Conclusions in section 6.

2 Defect Extraction

Different defect mechanisms can occur in the DUT. According to the production test data, bridge defects are the most dominant one followed by dislocation defects. Open defects are less likely in this technology so they are not used in the current flow.

To identify bridge defects, an IFA like approach is utilized; the DUT design database (layout) is passed through a parasitic capacitor extractor and the capacitor list is used in DOTSS to enlist bridge defects together with their relative likelihood. For dislocation defects, the cross section of library models that contain PN junctions are studied and the potential dislocation candidates are determined for each model. A script parses the DUT netlist and when it finds an instance whose model has potential dislocation defects, writes them into a dislocation defect list. We extracted about 38,000 bridge defects and about 1,500 dislocation defects. This paper deals with the 38,000 bridge defects, however, work is in progress to include dislocation and open defects in the near future.

3 Fault Simulation

The extracted defects are modeled as faults and injected one by one in the netlist of the DUT. The fault model is a simple resistor whose value is set according to the defect class (bridge) and the process information. The faulty netlists are then simulated. The full-chip analogue simulation of about 38,000 faulty netlists can be extremely time consuming. We use a novel fast simulation algorithm called fault sensitivity analysis (FSA) that is embedded within PSTAR. We are also discussing this methodology with EDA vendors for its implementation within commercial tools such that a wider community can utilize it. We also use other internal tools of NXP (AnalogueShell and DOTSS) to facilitate handling the large volume of simulation data corresponding to 38,000 defects and 197 production tests as well as fault injection and simulator launches.

3.1 Fault Sensitivity Analysis

Fault sensitivity analysis (FSA) is a novel methodology for establishing the detection status of defects in the DOT flow. FSA allows very efficient analogue fault simulation with reasonable accuracy for transient analysis in nonlinear circuits. FSA is based on two features of fault simulation in DOT: 1) establishing the detection status of a fault is more important than the actual faulty circuit output value, and

---

2The in-house analogue simulator of NXP.
2) the output value is often measured at few time points compared to many time points of a transient simulation.

Figure 2 (a) shows time points for fault simulation of a given bridge with known resistance (therefore known conductance $G = \frac{R}{2}$). The transient simulation includes many time points but the test measurement points are only few as shown in Figure 2 (b). The idea is to calculate the faulty circuit output only at the measurement time points.

The question now is how to calculate the faulty circuit output at only the measurement points and avoid the rest of time points? This is achieved by using a discrete bridge model instead of the standard bridge model. The discrete bridge model has zero conductance for all time points, except for the test points; it is assumed that the discrete bridge influences the circuit only at the test points.

Hence, for all time points, a standard transient analysis of the golden circuit is performed and for the test points an additional analysis is performed for each fault. This analysis repeats the golden time step for the test point and reuses the golden data of the previous time point. The faulty circuit solution $V_{\text{fault}}$ is computed from

$$F(V_{\text{fault}}) + G_{\text{fault}}V_{\text{fault}} = 0 \quad (1)$$

where the golden solution satisfies the nonlinear equation $F(V_{\text{golden}}) = 0$ and the linear term $G_{\text{fault}}V_{\text{fault}}$ is the contribution due to the conductance of a fault ($G_{\text{fault}}$). Equation (1) is solved using the Newton-Raphson method, which can be expressed as

$$V_{i+1} = V_i - (J_i + G_{\text{fault}})^{-1}(F_i + G_{\text{fault}}V_i) \quad (2)$$

where $i$ is the iteration index and $J_i$ is the Jacobian of $F$ at $V_i$. The golden solution is used as a start of the Newton-Raphson process and therefore the first iteration approximates the solution by linearizing the faulty circuit at the golden value $V_{\text{golden}}$.

Figure 3 (a) illustrates eq. (1) in a graphical way. The faulty solution $V_{\text{fault}}$ is symbolically plotted versus the fault conductance. The bold curve points to the circuit behavior as a function of fault conductance given by eq. (1) and the thin curve is the behavior of the linearized faulty circuit. The value of the thin curve at the conductance of a given fault $f_1$ is referred to as the linear FSA result of $f_1$. The error of the linear FSA is also shown. The subsequent Newton-Raphson iterations are corrections of the nonlinear FSA which will eventually converge to the actual solution $V_{\text{fault}}$ by applying eq. (2). An impression of the nonlinear FSA iterations is presented in Figure 3 (b).

Table 1 compares the standard (non-FSA) simulation with linear as well as nonlinear sensitivity analysis by applying different number of iterations. The FSA provides at least 173x speedup (corresponding to the nonlinear FSA with 10 iterations) over the standard simulation. These results are obtained for 38,000 defects and 75 tests with only one CPU used.

![Figure 2: Time points reduction: (a) simulation calculates circuit at each and every time point (b) test only needs circuit results at the two (generally few) measurement points.](image1)

![Figure 3: Fault Sensitivity Analysis: (a) Linear (b) Non-linear](image2)

### Table 1: Fault Sensitivity Analysis (FSA) performance

<table>
<thead>
<tr>
<th>Simulation Methodology</th>
<th>Speedup</th>
<th>Wall clock time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>1</td>
<td>24,228h (1010 days)</td>
</tr>
<tr>
<td>Linear FSA</td>
<td>1050</td>
<td>23h (1 day)</td>
</tr>
<tr>
<td>Nonlinear FSA (1 iteration)</td>
<td>515</td>
<td>47h (2 days)</td>
</tr>
<tr>
<td>Nonlinear FSA (5 iterations)</td>
<td>238</td>
<td>102h (4.2 days)</td>
</tr>
<tr>
<td>Nonlinear FSA (10 iterations)</td>
<td>173</td>
<td>140h (5.8 days)</td>
</tr>
</tbody>
</table>

FSA performs well for both catastrophic and non-catastrophic faults. In the former case, the impact is often large thus detection status is identified correctly with FSA.

### 3.2 FSA Accuracy

FSA aims to predict the outcome of fault detection correctly rather than the actual faulty value. Deviations from the actual faulty value are acceptable as long as FSA predicts the same detect status that the accurate non-FSA simulation (e.g. SPICE) would do. Nevertheless, there are cases for which FSA fails to predict the detect sta-
tus correctly. An experiment is devised to check this phenomenon: about 8,000 analogue-only defects (that is, both bridged nets are analogue nets/lines) are simulated for about 150 (of 197) tests using the standard non-FSA (accurate) simulation. The choice of defects and tests is to keep the simulation time manageable and the experiment practical. The detect statuses are established and checked versus those obtained by FSA. The percentage of defects whose status is correctly predicted by FSA (with reference to the standard non-FSA simulation) is called matching and used as a figure of merit for FSA. Figure 4 (a) presents the matching plots as a function of tests. Figure 4 (b) shows the distribution of tests versus matching. On average matching is 95.1% over all tests while about 87.4% of tests have matching more than 90%. The worst case matching is 82.92%; our investigation shows this worst case is because the nonlinear solver does not converge. Using a more robust algorithm for the nonlinear solver will improve the matching significantly.

Figure 4: The matching results of FSA: (a) matching versus tests (b) histogram of the tests versus matching.

4 Test Selection

Test selection is the process by which a small (minimal) subset of the production tests is made that covers a required percentage of (detectable\(^3\)) defects. Since the subset is smaller than the original set, test reduction and subsequently test time reduction can be achieved. The percentage of the covered defects is user adjustable, however by default, the flow covers all detectable defects (100% defect coverage).

Input to the test selection is a detection matrix whose entries specify if a test (column) detects a defect (row) \([22]\). In \([22]\) the entries are either 0 or 1 (undetect or detect) due to the digital nature of the DUT and its tests. In this work however, the entries are continuous numbers \(\geq 0\). For each test there are two limits: if a test result falls within the limits, it is assumed passed, else failed. Similarly, if a test simulation result falls between limits, it is assumed undetected (0); otherwise detected (\(>0\)). The detection matrix entries show how far the simulation or measurement results are from the test limits, see Figure 5.

Figure 5: Calculation of detection matrix entries for a test and three faults.

The larger the entry is, the better is the detection; for example, 0.715 indicates a weaker detection compared to 1.429. The user can also set a threshold to control how strong of detections he wants to include in the selection process. This is useful in accounting for process variations because out of limit measurements do not necessarily point to a faulty DUT, the DUT could be suffering from process variations. An example of the detection matrix is presented in Figure 6.

Figure 6: Example of detection matrix: detection of 'Defect 25' by 'Test 122' is a lot stronger than 'Defect 26' by the same test.

The detection matrix is mapped to an instance of set-covering combinatorial problem. In set-covering several sets are given as input that may have common elements. One must select a minimal number of these sets so that the selected sets contain all the elements that are in the union of the input sets. In the mapping, tests represent input sets and the defects they detect are the set elements (some defects may be detected by multiple tests). We must select a minimal number of tests (sets) such that they cover all defects detectable by the original tests. An example of test selection is shown in Figure 7.

Set-covering is an NP-Complete problem but heuristics
The heuristic is presented in Algorithm 1.

```
T ← Tests; /* original tests */
D ← Defects; /* detectable defects */
S ← ∅; /* Minimal set, initially empty */
T ← T − T;
while t ∈ T do
    if no condition of t ∈ S then
        S ← S ∪ {t}; /* select condition c of t */
    end
end
```

Algorithm 2: Selects the minimal test set S and preserves spec tests.

```
c ← the supply condition that appears most in S;
```

```
for t ∈ T do
    if no condition of t ∈ S then
        S ← S ∪ {t}; /* select condition c of t */
    end
end
```

Algorithm 3: Selects the minimal test set S, preserves spec tests, and ensures at least one supply condition of any test is selected.

The production tests of our vehicle DUT include digital, analogue, and IDDQ tests. Non-digital tests (analogue and IDDQ) dominate the test time so we focus on the 197 such tests. Included in the 197 tests are up to 3 voltage supply conditions.

## 5 Validation

During validation, the production test data logs of 10 lots, containing 1.3 million dies, are analyzed. In this analysis continue-on-fail is used, hence results for all tests are stored. All tests detect some defects although their efficiency varies widely. Some tests detect more than half of the rejected devices while others catch far less than 1% of the rejected ICs. Most of the tests are performed at all the 3 supply voltages. At each of them, we find unique defects. Hence, test time reduction is not as trivial as determining the best voltage setting to improve detectability.

4The test equipment does not stop after the first failing test and continues test application until the last test.
The analysis procedure to validate the algorithms (selections) has two steps: 1) apply the test program limits for 197 original tests and set the pass/fail boundary for all 1.3 million dies; 2) remove the dies that fail at least one test in the selected test set; any remaining die that fails any of the original tests escapes the selected tests (would-escape-die). The analysis results for the three selections associated with Algorithm 1, 2, and 3 are presented in Table 2. The number of dies that escape if the selected tests are used as well as test reduction and test time reduction figures are presented. The last column presents the actual test time reductions. This can be different from the test reduction since some tests can take longer to apply than others.

<table>
<thead>
<tr>
<th>Selection</th>
<th>Escapes</th>
<th>Test Reduction</th>
<th>Test Time Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm 1</td>
<td>14</td>
<td>46% (107/197)</td>
<td>37.3%</td>
</tr>
<tr>
<td>Algorithm 2</td>
<td>6</td>
<td>41% (116/197)</td>
<td>33.2%</td>
</tr>
<tr>
<td>Algorithm 3</td>
<td>0</td>
<td>35% (128/197)</td>
<td>27.2%</td>
</tr>
</tbody>
</table>

As seen, 14 and 6 dies would escape selection 1 and selection 2 respectively but selection 3 guarantees no escapes. Normalized with respect to the 1.3 million dies, the defect level (ppm) figures are at 11, 4.6, and 0 ppm for the three selections respectively. It is noteworthy that a would-escape-die does not necessarily mean a customer return. Such dies may suffer from process variation or only marginally fail the test.

6 Conclusions

A DOT approach is utilized to optimize production tests of an AMS automotive product with stringent quality requirements. The results are promising: using a fast analogue simulation algorithm enables us to simulate an extensive set of defects. Three selection algorithms are proposed and evaluated. These algorithms enable us to reduce the number of tests by 35% to 46%; corresponding to the actual test time reduction of 27% to 37%. Algorithm 3 is capable of detecting all outliers and fits the zero-defect automotive requirements. The other two algorithms result in more test time improvement but will miss a number of outliers, which are seen as a potential quality risk for automotive products.

References


