Post-Production Performance Calibration in Analog/RF Devices

Nathan Kupp*, He Huang*, Petros Drineas†, and Yiorgos Makris*
*Department of Electrical Engineering, Yale University, New Haven, CT 06511
†Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY 12180

***Abstract*—In semiconductor device fabrication, continual demand for high performance, high yield devices has caused designers to look to post-production tunable circuits as the next logical step in analog/RF design and test development. These approaches have not yet achieved the maturity necessary for industrial adoption, primarily due to complexity and cost. In this work, we develop a general model which systematically outlines several key observations constraining the complexity of performance calibration in analog/RF devices. Moreover, we develop a detailed cost model permitting direct comparison of performance calibration methods to industry standard specification testing. Our analysis is demonstrated on a tunable RF LNA device simulated in 0.18μm RFCMOS.**

I. Introduction

In modern analog/RF device fabrication, circuits are typically designed conservatively to ensure high yield. Otherwise, yields may be low due to process variation driving devices beyond specifications. Thus, analog designers often find themselves doubly constrained by performance and yield concerns.

However, the demand for high performance, high yield analog/RF devices is relentless. As such, recent interest has been shown in producing analog/RF devices that are tunable after fabrication by introducing "knobs" (post-production tunable components) into the circuit design. By adjusting the knobs, some devices that would simply be discarded under the traditional analog/RF test regime can be tuned to meet specification limits and thereby function correctly. These tunable devices would permit analog/RF designers to create aggressively high-performance integrated circuits (ICs) with expectations of reasonable yield. Alternatively, conservative designs could be produced with nearly-perfect yields.

To date, post-production performance calibration has not achieved widespread use due to the perceived complexity and cost of implementation. This is not an unreasonable perception: knobs have apparently complex interdependent effects on performances, and iterative specification test-tune cycles to explore the large space of knob settings are prohibitively costly. In this work, we outline several key observations which appropriately constrain the free parameters of performance calibration methodologies to enable straightforward cost-effective implementation. Moreover, we develop and present a cost model which permits direct comparison of performance calibration to specification test and other state-of-the-art practices.

II. Test and Performance Calibration Methods

In this section, we present the state-of-the-art in analog/RF device testing and contextualize performance calibration within this domain.

A. Specification Testing

The current industry-standard practice for determining the functional health of analog/RF devices is specification testing. As shown in Figure 1, each fabricated device under test (DUT) undergoes a series of tests designed to compare a measured set of performances to a corresponding list of specification limits. Circuits are classified as passing if they perform within these specification limits.

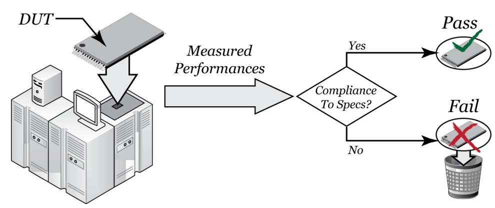


Fig. 1. Specification Test

B. Cost Reduction via Test Compaction and Alternate Test

One of the biggest limitations of specification testing is cost—analog/RF device specifications are often complex derivations which are expensive to measure, and require costly automated testing equipment (ATE) to obtain at operating frequencies. This cost has motivated development of test-cost reduction methods, which largely fall into two categories. The first is *test compaction*, which aims to reduce the number of specification tests which must be explicitly measured by leveraging correlations amongst the specification tests [1]–[5]. The second is *alternate test*, where expensive specification tests are replaced with low-cost "alternate tests" [6], as shown in Figure 2. These alternate tests are specifically designed to be well-correlated with the specification tests while consuming significantly fewer test resources to collect. This method requires a pre-production stage where we set aside a small training set of devices on which we collect *both* the low-cost alternate tests and the performances of the devices in the training set. From this training data, regression models correlating alternate tests to performances are constructed. In production, only the alternate tests are explicitly measured on every device and used in conjunction with the trained regression models to predict performances.

Although alternate test substantially reduces cost, it has not achieved widespread industry adoption due to the incurred misclassification error: small errors in the predicted values of performances occassionally result in defective devices being labeled as passing (test escapes) or passing devices being labeled as failing (yield loss). In general, alternate and machine learning-based test methods that have been proposed to date all involve some exploration of this tradeoff between cost and error [7].

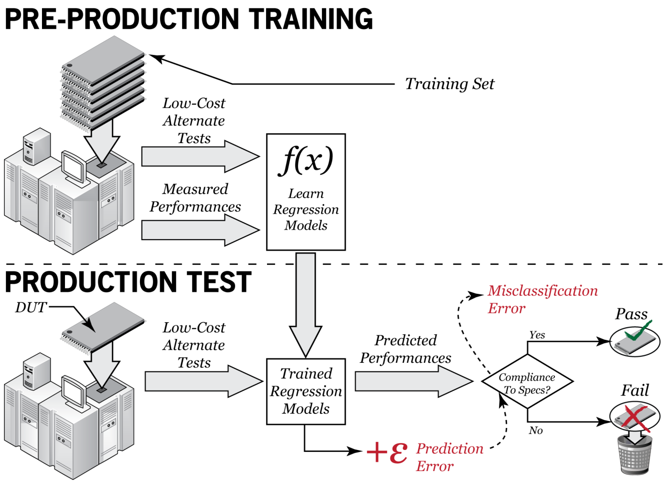


Fig. 2. Alternate Test

C. Performance Calibration and Healability

Alternate test finds an immediate home in performance calibration, however. As noted previously, the constraints of modern analog/RF design require high-yield, high-performance devices in the presence of increasing process variation, motivating a search for better control of process variation beyond traditional approaches. To this end, designers have looked to healable architectures to provide additional control over process variation effects. With performance calibration, devices without catastrophic faults can be recovered to meet specification limits by performing post-manufacture optimization of their performances.

Implementation of performance calibration requires selection of key parameters of the circuit (voltage, capacitance, etc.) as knobs. Additional circuit elements are added to enable post-production modulation and on-chip storage of these parameters. By setting knob values, the performances of the circuit can be dynamically modified and improved to meet specification limits.

After the knobs are in place, a method for circuit tuning must be devised. In this work, we focus specifically on addressing this problem. A naïve approach is shown in Figure 3, where all performances are iteratively tested across all knob settings until a setting that results in a passing device is found. Knob setting selection can also be performed by attempting to find some optimum for the circuit, i.e. searching for the lowest possible power setting. While such exhaustive specification test approaches are unrealistically expensive, we collect these results in our experiments as a ground-truth reference point: with exhaustive specification test we know *a priori* the true healable/unhealable statistics of the device population. This enables us to establish an upper bound on yield improvement due to performance calibration methods if cost were not a factor.

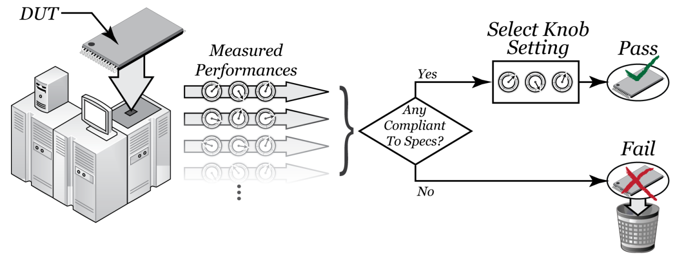


Fig. 3. Performance Calibration: Exhaustive Specification Test

D. Alternate Test-Based Performance Calibration

Clearly, these approaches are suboptimal given the high cost of specification test. As modern analog/RF device testing can comprise a substantial percentage of total device production cost, exhaustive specification testing is not economically feasible. Herein lies the distinct benefit of adopting alternate test during the performance calibration process: alternate tests can be an order of magnitude less costly to perform, yet offer reasonable accuracy.

We observe that the straightforward naïve approach to alternate test-based performance calibration would be to exhaustively apply tests as before, substituting alternate tests for specification testing, as shown in Figure 4. As in Section II-B we employ a training set, building a large number of regression models (one per knob setting) instead of the single model used in the knob-free alternate test methodology. In production, we no longer exhaustively measure performances, instead collecting alternate test measurements at various knob settings until we find a knob with performances predicted to produce a passing device. This provides a second point of reference for more sophisticated methods by sidestepping the need to model knob variation; each knob setting is given an individual alternate test regression model.

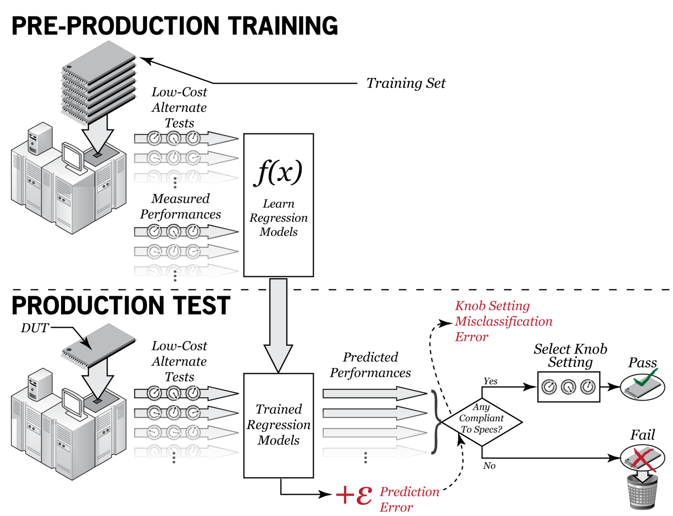


Fig. 4. Performance Calibration: Exhaustive Alternate Test

Note that as with traditional alternate test, employing this method for performance calibration again introduces a tradeoff: lower cost at the expense of added misclassification error, in this case comprised of three components:

1) Test Escapes: An unhealable device may be labeled as healable.
2) Yield Loss: A healable device may be labeled as unhealable.
3) Incorrect Heal: A predicted-to-heal knob setting may be chosen which does not actually heal the device.

E. Proposed Method: Midpoint Alternate Test-Based Performance Calibration

In this work we propose a novel performance calibration method, entitled midpoint alternate test-based performance calibration. To manage the cost of alternate test-based performance calibration, we must modify exhaustive test to reduce the large number of measurements (alternate test or otherwise) which must be collected on a given device. We do this by making an important observation: knob variation and process variation *orthogonally* act on device performances. Thus, we can separately model each axis of variation and build a composite model which accounts for both.

We have already stated that alternate tests are designed to correlate well with device performances. Implicitly, this means that we can already model process variation from the alternate tests. To model knob variation, we must understand knob effects in a process variation-free space: the ideal device, a process variation-free simulated device. A detailed account of how these axes are modeled is provided in Section IV.

By appropriately modeling the knob- and performance-variation axes, we are able to achieve an extreme reduction in the number of alternate tests which must be collected. Instead of collecting alternate test measurements at *all* knob settings, we only explicitly measure alternate tests at a *single* knob setting, where all knobs are set to their respective midpoint values. This set of midpoint alternate test measurements can be used to predict performances at *all* of the knob settings, as shown in Figure 5.

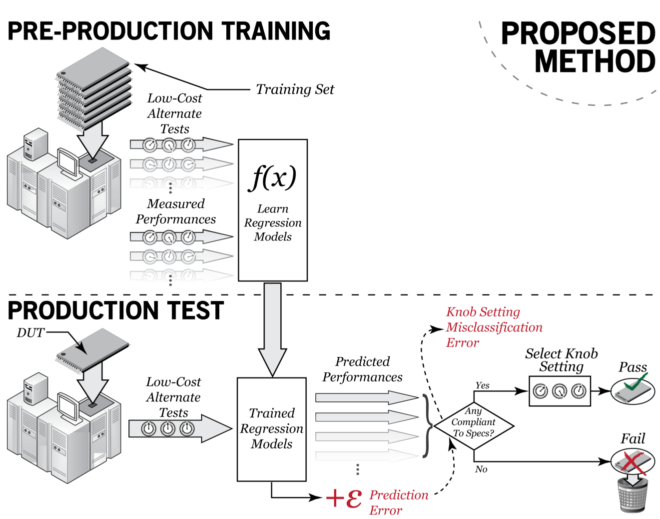


Fig. 5. Proposed Method: Midpoint Alternate Test Performance Calibration

Note that midpoint alternate test-based performance calibration is also subject to the three sources of error outlined in Section II-D due to the use of alternate test. By only collecting a single set of alternate tests instead of the naïve approach of exhaustively testing for each knob setting, we introduce a slight increase in error, which is well-justified by the extremely large reduction in test cost achieved.

F. Optional Step: Validation

As we have discussed, both exhaustive alternate test and midpoint alternate test performance calibration methods can result in misclassification error. Depending on the constraints of the specific application, the alternate-test based performance calibration stage of testing can be followed by a final specification testing stage to eliminate error due to test escapes or incorrect heals, albeit at higher cost. The validation step is illustrated in Figure 6.

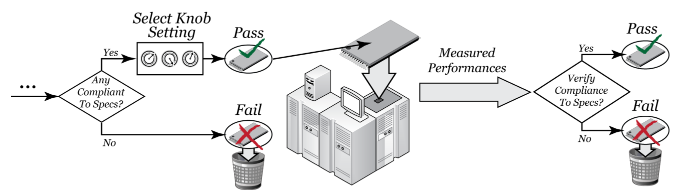


Fig. 6. Performance Calibration: Validation

If deemed necessary, this validation stage should be performed for only high-risk devices, with risk of misclassification determined by a confidence estimation method, as employed in [3], [4]. Only a small fraction of devices actually require validation with this constraint, thereby moderating the added cost.

III. Cost Model

As we have observed, one of the most significant roadblocks to performance calibration adoption is cost. Unless the cost-benefit ratio of deploying tunable architectures is on par with or lower than that of existing design and test methods, it will not be implemented. Here, we develop an inclusive cost model which compares our midpoint alternate test-based performance calibration method to the various baseline methods: specification testing, alternate testing, and exhaustive testing. The notation of Table I is employed for the remainder of this section.

Variable	Definition
C_0	Baseline cost of device development and production.
C_D	Design cost to add knobs and implement device as a tunable architecture.
N_T'	Number of devices in the training set.
N_T	Number of devices in the test set.
N_K	Number of knob settings.
P	Relative cost for measuring all performances.
A	Relative cost for measuring all alternate tests.
F	Fraction of devices undergoing validation test.

TABLE I
Cost Model Notation

A complete list of the cost models is presented in Table II. The reference case for cost is specification testing, where only baseline design cost C_0 and the cost of performing specification test once on every device $N_T P$ are included.

Configuration	Cost Model				
Specification Testing	$C = C_0$	+	$N_T P$		
Alternate Test	$C = C_0$	+	$N_T A$	+	$N'_T\,(A+P)$
Exhaustive Specification Test	$C = C_0 + C_D$	+	$N_T N_K P$		
Exhaustive Alternate Test	$C = C_0 + C_D$	+	$N_T N_K A$	+	$N'_T N_K\,(A+P)$
Midpoint Alternate Test	$C = C_0 + C_D$	+	$N_T A$	+	$N'_T N_K\,(A+P)$
Validation	$C = \underbrace{C_0 + C_D}_{\text{Baseline Term}}$	+	$\underbrace{N_T\,(A + F \cdot P)}_{\text{Test Set Term}}$	+	$\underbrace{N'_T N_K\,(A+P)}_{\text{Training Set Term}}$

TABLE II
COST MODELS FOR TEST AND PERFORMANCE CALIBRATION METHODS

As discussed in the previous section, alternate test replaces expensive specification tests with a set of low-cost alternate tests. Thus, our cost model for alternate test substitutes the $N_T P$ term with the cost of running alternate tests on every device $N_T A$. As the models to predict performances from alternate tests must be learned, we also require a small training set where both alternate and specification tests are performed, $N'_T\,(A+P)$. Note that typically, $N_T \gg N'_T$.

We include for reference the two exhaustive test performance calibration methods. For both of these approaches, we include C_D, the design cost of adding tunable elements to an analog/RF circuit design. The exhaustive specification test option does not require a training set, so the cost of testing performances at every knob setting for every device is simply $N_T N_K P$. Recall that for exhaustive alternate test, we construct individual regression models at every knob setting. This requires exhaustively collecting alternate tests in the test set, $N_T N_K A$, as well as collecting both alternate and specification test measurements exhaustively in the training set, $N'_T N_K\,(A+P)$.

Lastly, we include our midpoint alternate test performance calibration methodology. This maintains the knob-design cost term, C_D, but reduces the test set cost from $N_T N_K A$ to $N_T A$. As N_K scales, the exhaustive test performance calibration methods become unreasonably expensive, whereas the midpoint alternate test methodology retains constant cost in the test set. However, note that the midpoint test method maintains a training set cost $N'_T N_K\,(A+P)$ that is proportional to N_K. In Section VI-H we address constraining this test set cost via uniform sampling, decoupling training set cost from N_K.

If midpoint alternate test is paired with a validation stage of specification test to verify a correct tune, we replace the test set term $N_T A$ with $N_T(A + F \cdot P)$. This adds a cost $F \cdot P$ to the test set, proportional to the fraction F of devices which undergo traditional specification test validation.

IV. PROCESS AND KNOB VARIATION MODELING

A. Knob Effect Modeling

Test engineers are well-acquainted with variation in circuit performances due to uncontrollability in the fabrication process. By introducing knobs to a circuit design, a second dimension of variability is introduced, such that circuit performances are no longer simply a function of process variation, but knob positions as well. To effectively use tuning knobs to do performance calibration, variation in both dimensions must be understood and modeled appropriately.

To achieve the objective of healing devices without explicitly measuring device performances at every knob setting, we must make certain assumptions about how knob and process variations affect device performances. We theorize that we can capture the knob effects by studying the "ideal" device, or simulated performances of the circuit at each knob setting without process variation. As this simulated device does not contain process variation, the $N \times k$ matrix $\mathbf{P}$ of k performances across N knob settings provides us the necessary information to model how the ideal device responds to changes in knob position.

Analog/RF design closely approximates a zero-sum game, and is a careful balance of various trade-offs. Adding post-production tunable elements to a circuit simply postpones a portion of this trade-off optimization process until after device fabrication. Thus, any non-trivial knob circuit element will affect more than a single performance, some positively, others negatively. Ideally, knobs would also be designed to be totally independent. The two papers [8], [9] argue for the original knob designs to be designed to be near-completely independent so that a simple linear model will approximate knob effects on performances well.

However, the non-idealities of analog/RF design make complete independence impossible to achieve. More importantly, this is an unnecessary constraint. Although seeking knob independence remains a laudable objective, we can better model knob effects on performances by acknowledging and accommodating for knob interdependence through the inclusion of second-order knob interaction terms along with knob main effects in our model.

Thus, we model the performance responses of the ideal device as functionally dependent on the knob settings via a model that is linear in the parameters but includes the pairwise quadratic interaction terms of explanatory variables in the design matrix:

$$\hat{P} = \hat{\beta}_0 + K^T \hat{\beta} + \varepsilon \qquad (1)$$

where $\hat{\beta}_{\mathbf{0}}$ is an intercept term representing the variation-free performances of the device, K is the $1 \times \left(p + \binom{p}{2}\right)$ vector of knob settings and interaction terms:

$$K = (\underbrace{K_1, K_2, \ldots, K_p}_{\text{main effects}}, \underbrace{K_1K_2, K_1K_3, \ldots, K_{p-1}K_p}_{\text{interaction terms}}) \qquad (2)$$

and $\hat{\beta}$ is the the knob effect parameter vector, also $\left(p + \binom{p}{2}\right) \times$

1, estimated by our model:

$$\hat{\beta} = (\underbrace{\beta_1, \beta_2, \ldots, \beta_m}_{\text{main effects}}, \underbrace{\beta_{1:2}, \beta_{1:3}, \ldots, \beta_{(p-1):p}}_{\text{interaction terms}}) \tag{3}$$

Including $\hat{\beta}_0$ in our vector of parameters $\hat{\beta}$ and prepending a constant 1 to our vector K permits us to formulate the equation as traditional least-squares:

$$\hat{P} = f(K) = K^T \hat{\beta} + \varepsilon \tag{4}$$

where we solve for the knob effects using the standard least-squares solution:

$$\hat{\beta} = (\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T P \tag{5}$$

given the $N \times \left(p + \binom{p}{2}\right) + 1$ design matrix $\mathbf{K}$. This process is repeated for all k of the ideal device performances, allowing us to individually model each of the device performances.

Justification was provided earlier for including second order interaction terms in the knob effect model. The astute reader may inquire why third-order or higher-order interaction terms were not considered. For our work, we found that second order interaction terms are sufficient for successful performance calibration; moreover, we observe that the number of free parameters $|\beta|$ becomes very large as we add higher-order terms to the model:

$$\lim_{n} |\beta| = \lim_{n} \left[1 + p + \binom{p}{2} + \binom{p}{3} + \cdots + \binom{p}{n} \right] = \infty \tag{6}$$

exponentially increasing model complexity and substantially reducing model interpretability.

B. Process Variation Modeling

Were we to apply the knob-effect model described in the previous section to data from a real device, the prediction error would be large as the model does not account for process variation at all. We argue a surprising result: given the orthogonality of process variation and knob variation, process variation is a *constant offset* from the presented knob effect model. That is, we can jointly model *both* knob and process variation effects by adding a single term to our model:

$$\hat{P} = f(D, K) = \hat{\beta}_0 + \hat{\beta}_1 D + K^T \hat{\beta} + \varepsilon \tag{7}$$

where we have simply added an additional device-specific scalar value D and a parameter $\hat{\beta}_1$ to represent process variation-induced perturbation of performances.

The immediate challenge is obtaining an appropriate estimate for D. For this, we look to alternate tests. Each alternate test A gives us a measure of the magnitude of process variation effects. Of course, each performance shows high correlation with different subsets of the alternate test set. Thus, we can include all of the alternate tests collected to improve our estimate, resulting in the following complete model:

$$\hat{P} = f(A, K) = \hat{\beta}_0 + A^T \hat{\beta}' + K^T \hat{\beta}^\dagger + \varepsilon \tag{8}$$

Finally, we concatenate the vectors A and K as X (following convention and prepending a constant term as before), and concatenate $\hat{\beta}'$ and $\hat{\beta}^\dagger$ as $\hat{\beta}$ to arrive at the equation:

$$\hat{P} = f(X) = X^T \hat{\beta} + \varepsilon \tag{9}$$

and solve for $\hat{\beta}$ via the least-squares solution in Equation 5. Thus, Equation 9 provides a complete joint model for knob and process variation effects on a single performance. We follow this approach to generate individual models for each of the k performances $P \in \mathbf{P}$.

C. Incremental Model Improvements

To further improve performance on actual device data, we make several enhancements to our model. First, we slightly modify the design matrix: Instead of predicting device performances from knob settings, knob interaction terms, and a set of alternate test measurements, we replace the knob terms with a data frame of ideal device performances. Thus, when building the models on a training set, we construct the model frame shown in Table III for each device. This model frame is then repeated and concatenated row-wise to form the model training set $\mathbf{X}$ on which we regress $\mathbf{P}$.

	Ideal Device	**Test Set Device**
Knob Setting	Power and Performances	Alternate Tests
1		↑
	Simulated	Measured
⋮	Values	Constant Columns
N		↓

TABLE III
MODEL FRAME DIAGRAM

For our reported experimental results we employ a second change, replacing ordinary least-squares with the Multivariate Adaptive Regression Splines (MARS) algorithm [10]. This choice was driven by an experimentally-observed small improvement in residual error versus the least-squares regression models. However, using MARS is not ideal, as the coefficients and split points of spline basis functions are not easily interpretable. The least-squares model in Equation 9 should be constructed alongside any more sophisticated models to provide the test engineer with the important parameter estimates $\hat{\beta}$, so that the knob effect models can be understood.

D. Knob Setting Selection

Once we have modeled both knob effects and process variation effects, the models must be employed to inform knob setting selection decisions on each device in the test set, where limited information about the device is available to us. In our work, we have chosen to predict performances for each knob setting for every device. This allows us to accomplish two things. First, we partition the test set into unhealable and healable regions by determining for every device in the test set whether at least one knob setting is available which will heal it. Second, for every healable device in the test set we predict a family of knob settings which will heal it.

Given this family of predicted-to-heal knob settings, we need to employ some method to make an appropriate selection. As a baseline reference case, we report a probability of correct heal by random selection, picking a knob setting uniformly at random:

$$Pr(\text{Correct Heal}) = \frac{\#PH \bigcap \#AH}{\#PH} \tag{10}$$

That is, the number of knob settings predicted to heal ($\#PH$) which actually heal ($\#AH$) as a percentage of the total number of knob settings predicted to heal. We intuitively expect that we can devise some knob setting selection method which improves on simple random selection. Here, we present two such methods.

1) Mahalanobis Distance: In terms of error, the most conservative approach is to order potential knob settings on the basis of Mahalanobis distance from specification planes in a normalized performance space[1]. We then select the knob setting which maximizes the Mahalanobis distance from specification planes. Optimality is contingent on the type of specification limits provided: For single-sided limits, optimality is distance maximization with respect to the plane iteself, whereas optimality for double-sided limits is minimization of distance with respect to the midpoint of the specification limits. Our metric m is therefore a composite of these two cases, given a mixed set of single- and double-sided specification limits for the K performances $s = \{s_1, s_2, \ldots, s_K\}$:

$$m = \max_k \sqrt{\sum_{i=1}^{K} d_i^2} \tag{11}$$

where each d_i is given by:

$$d_i = \begin{cases} (p_{ik} - s_i), & \text{1-Sided} \\ 1/\left[p_{ik} - \left(s_{i_{lower}} + s_{i_{range}}/2\right)\right], & \text{2-Sided} \end{cases} \tag{12}$$

in normalized space, and each k is a predicted-to-heal setting for the given device, generated via our prediction model. This approach minimizes the probability of a mistake due to marginal prediction error at the specification limit boundaries, at the expense of tending towards large increases in power consumption.

2) Power: An ever-important constraint of analog/RF design is power. Given a set of predicted-to-heal knob settings for a device, power is a natural optimizer for selection.

However, following the logic presented earlier, exhaustively testing power at every knob setting is impractical. Therefore, to use power as a knob setting selection metric, we must add power to the list of predicted device performances during the model-construction stage of our midpoint alternate test-based performance calibration. This enables us to predict device power consumption for every knob setting of every device in the test set. Significantly, we found that the prediction error for power was very low, such that predicted power and actual power produced the same knob setting rankings. This enables us to use predicted power for ranking knob settings.

Once we have used our trained regression models to predict power values, we employ two selection metrics: minimum power and median power. Minimizing power while meeting specification limits would appear to be the global optimum; indeed, this would be the case were we to have performed exhaustive specification test giving ground truth pass/fail for every knob setting. However, using statistical models introduces slight errors in the pass/fail boundary. By minimizing power, performances are pushed closer to their specification limits, thereby increasing the apparent misclassification error. By optimizing for median power, we mitigate some of this error, while avoiding the high power consumption of the Mahalanobis distance metric presented previously.

V. Comparison to Prior Work

Several previous publications [8], [9], [11]–[14] attack the problem of performance calibration in analog/RF devices. These papers fall into two categories, namely *optimization* and *prediction*. The former is employed in [11]–[14], where gradient descent-based methods are used to iteratively perform test-tune-test cycles and heal the device. This method operates on the assumption that knob effects cannot be characterized in closed-form requiring use of iterative optimization methods. As we demonstrate, we can make much stronger assertions about how knobs interact with device performances. Moreover, using an iterative approach is overly expensive, requiring multiple test-tune-test cycles, whereas our proposed midpoint alternate test-based performance calibration methodology requires only a single test-tune step. The latter category, which is employed in [8], [9], recognizes that knob effects can be approximately characterized by first-order linear models, and a series of models are built to perturb baseline alternate test MARS model predictions. The effective cost of such methods is equivalent to our proposed method. However, these models are built on the assumption that designers can effectively build knobs that are near-perfectly independent. As complete independence is not achievable, we avoid the error introduced by this oversimplification and include knob interaction effects in our model. Moreover, instead of implementing a two-model approach (MARS and linear regression), we handle knob and process variation jointly in a single model.

VI. Experimental Validation

To validate our methods, we designed a cascode low-noise amplifier (LNA) in 0.18μm CMOS. In this section, we document our design choices and show experimental results for the proposed midpoint alternate test-based performance calibration method.

A. Tunable Low-Noise Amplifier

We selected the RF LNA as our platform for experimental validation, as it is one of the most frequently used components in commercial transceiver RFICs. Among the numerous possible LNA architectures, we chose one of the most widely-adopted designs, the cascode topology.

[1]Mahalanobis distance is simply a covariance-scaled version of Euclidean distance. We use Mahalanobis distance instead of Euclidean distance as our distance metric to ensure each specification is uniformly weighted.

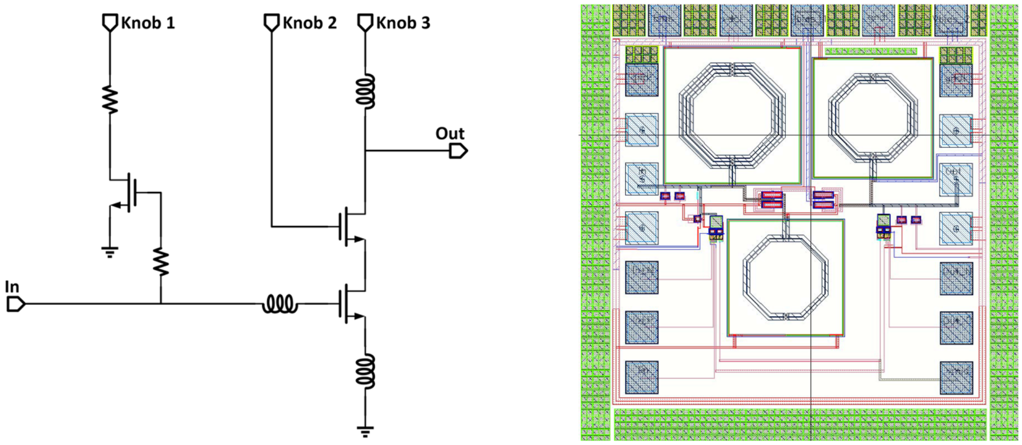
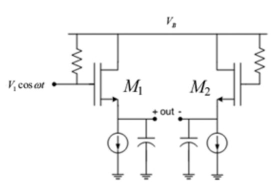


Fig. 7. LNA Design

To perform post-production performance calibration, we must modify the DUT to include tunable circuit elements. In our device design, we selected three key bias voltages to include as tuning knobs, as these provided maximal control over performances. The device schematic, including our selected knobs, is shown in Figure 7 along with the layout of the DUT.

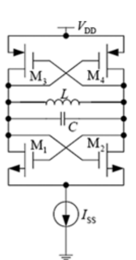
1) On-Chip Amplitude Sensor: Along with the LNA, we designed and implemented an on-chip amplitude sensor and on-chip signal generator, for collecting alternate test data [15]. With an appropriate choice of input signals, the alternate test measurements produced by the amplitude sensor/signal generator pair have been demonstrated to be well-correlated with performances. The schematic and performances of the proposed RF on-chip amplitude sensor are shown in Figure 8.



Parameter	Performance
Operation Frequency	1.575 GHz
Dynamic Range	50 dBm
Power Consumption	3.7 μW
Area Overhead	42 μm $\times$ 80 μm

Fig. 8. Amplitude Sensor Schematic

2) On-Chip Signal Generator: The schematic and performances of the proposed CMOS LC-tuned VCO are shown in Figure 9.



Parameter	Performance
Frequency Tuning Range	1.4 – 1.9 GHz
Phase Noise	-111.3 600 kHz
	-116.2 1 MHz
S11	-19.3 dB
Power Consumption	2.5 mW

Fig. 9. Signal Generator Schematic

The layout-level LNA was used to collect performance data across all knob settings of each device. For alternate test data, two amplitude detectors were added at the input and output of the LNA, and both were measured with stimuli provided by the RF signal generator. Two different frequencies of the RF signal generator were employed, for a total of 4 alternate test measurements collected per knob setting per device.

B. Dataset

For our experiments, we created 1,000 instances of the LNA with process variation effects included to simulate a production environment. The 3 knobs in the LNA designed for our experiment were assigned 3 discrete settings (i.e., 1.6V, 1.8V, 2.0V) for a total of $3^3 = 27$ possible knob positions.

On every device in our dataset, we collected four performances: S11, Noise Figure (NF), Gain, and S22. We also collected a power measurement and the four low-cost amplitude sensor (peak detector) alternate test measurements. Thus, for every device there are 9 figures of merit.

As noted in the introduction, process variation effects perturb circuit performances such that some devices exceed specification limits and fail specification test. However, the knobs in a tunable circuit also affect this 9-tuple performance vector. Clearly, this is a requirement of our knob circuit elements; otherwise the knobs are useless for performance calibration. Thus, the entire dataset is a $1,000 \times 27 \times 9$ tensor, as shown in Figure 10. In production it is infeasible to measure all circuit performances on every device at every knob setting, so only some circuit performances are explicitly measured. Essentially, all of the performance calibration methods proposed to date reduce to slicing away pieces of this 3-dimensional matrix and then determining the capability to do performance calibration with severely constrained information about the performance of a given device.

As stated previously, if we are to model the circuit response to knob and process variation, an initial training set must be generated which includes the relationships we wish to model. For example, if we wish to predict circuit performances at every knob setting, these performances must be explicitly measured for a small training set in order to construct our models. Once these models are constructed, they can be used to predict circuit performances for the remaining circuits. For

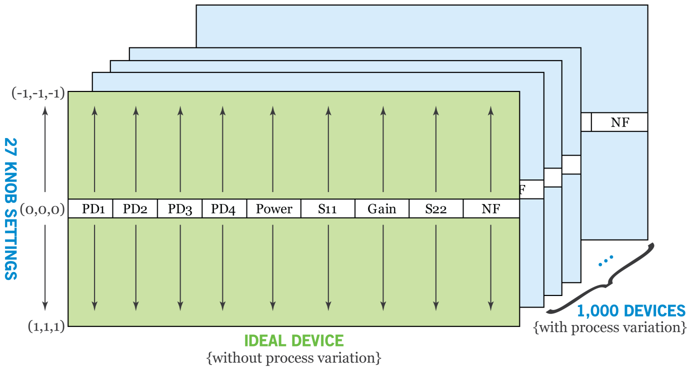


Fig. 10. Graphical depiction of dataset

the experiments which required training statistical models, we split the dataset 50/50, training on data from 500 devices and predicting on the remaining 500. We also performed 10 cross-validations to ensure statistical stability of the reported results.

C. Specification Test

As shown in Figure 11, we use the center knob position to emulate a knob-free device, and compare the performances measured at the center knob position to specification limits to obtain a pass/fail value for every device in the dataset. Of the 1,000 devices, 851 pass specification testing and 149 devices fail, translating to 85.1% yield.

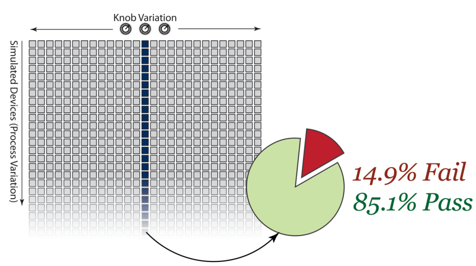


Fig. 11. Specification Test

D. Alternate Test

We also performed simple alternate test (no guard-banding or other derivative performance improvements) by considering only data from the midpoint knob setting, emulating a knob-free device. We constructed prediction models correlating each of the 4 device performances with peak detector measurements. The results of this experiment are shown in the confusion matrix of Table IV.

		Actual	
		Fail	Pass
Predicted	Fail	10.86%	1.52%
	Pass	3.54%	84.08%

TABLE IV
ALTERNATE TEST RESULTS

That is, employing standard alternate test results in a 3.54% test escape rate and a 1.52% yield loss rate. This is consistent with state-of-the-art alternate test literature, excluding sophisticated error compensation techniques such as guard banding.

E. Performance Calibration: Exhaustive Specification Test

As noted previously, exhaustive specification test provides a useful reference point for the absolute ceiling on yield improvement possible by using performance calibration techniques. As shown in Figure 12, we exhaustively measure all circuit performances to determine a ground truth pass/fail label for every knob setting for every device.

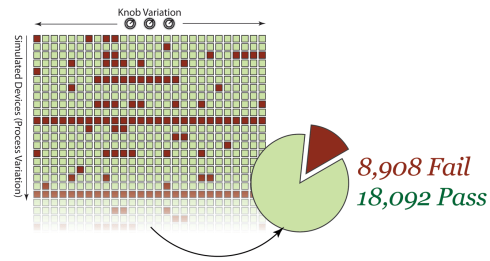


Fig. 12. Exhaustive Specification Test

Instead of simply looking at pass/fail labels for devices, using performance calibration permits us to extend the simple paradigm of pass/fail and label devices as *healable* or *unhealable*. A healable device is defined as a device with at least one knob setting which produces passing performances. We can use the matrix of Figure 12 to assign healable/unhealable labels to our test devices. For our data, 973 of the devices are healable, and 27 are unhealable. That is, by our definition of healability, 973 of the devices have at least one knob setting which enables passing all specification limits. Recall that when the tuning is not used, 851 of the devices meet specification limits and pass. Therefore, the maximum possible benefit from performance calibration methods for this device is 122 devices, or a 12.2% gain in yield. Note that in all, approximately two-thirds (18,092) of the $1,000 \times 27 = 27,000$ total number of knob settings produce passing performances, which clearly indicates that performing random knob setting selection would introduce unacceptably-high error.

		Actual	
		Unhealable	Healable
Predicted	Unhealable	1.56%	0.34%
	Healable	1.04%	97.06%

TABLE V
EXHAUSTIVE ALTERNATE TEST

F. Performance Calibration: Exhaustive Alternate Test

Here we provide results for the reference case of exhaustive alternate test approach detailed in Section II-D. Again, as this is a performance calibration method, we label devices as healable or unhealable instead of pass/fail. Recall that the use of alternate test introduces some error, and for alternate test-based performance calibration, this error presents as unhealable/healable misclassification error. The confusion matrix in Table V shows the error for unhealable and healable classification using exhaustive alternate test regression models. That is, due to the use of alternate test we introduce an approximately 1.04% test escape rate and a 0.34% yield loss rate, for a slightly over 1% total error rate.

For the devices that were correctly labeled as healable, we examined the predictions to determine the probability of correct heal vis-à-vis Equation 10. This number gives us an estimate of the quality of our knob setting pass/fail classifications across the healable devices. For exhaustive alternate test-based performance calibration, the probability of correct heal is 95.9% using a uniform random selection amongst the predicted-to-heal knob settings for each device.

G. Performance Calibration: Midpoint Alternate Test

As noted earlier, exhaustive test methods, even when alternate tests are used, are prohibitively expensive to adopt for practical performance calibration. In contrast, the proposed midpoint alternate test technique outlined in II-E is a viable option, as we demonstrate herein.

1) Knob Effect Modeling: To justify our model choice in Section IV-A, we evaluate the model using simulation data from a process variation-free ideal device, using least squares to estimate the parameters $\hat{\beta}$ to provide us with the model of Equation 4. Table VI displays the R-squared values for the ideal device models. This conclusively verifies our assertion that ideal device performances can be expressed *solely* as a function of knobs and knob interaction terms.

S11	Gain	NF	S22
0.941	0.98	0.99	0.98

TABLE VI
KNOB EFFECT MODELS: R^2

2) Process Variation Modeling: To visualize how process variation perturbation affects the knob effect model, we directly applied the ideal-device regression models of Section IV-A to data from a sampled subset of simulated devices. In Figure 13 we show predicted vs. actual gain, where each line trace represents the 27 knob settings for a single device. Note that data from only a handful of devices is presented for clarity.

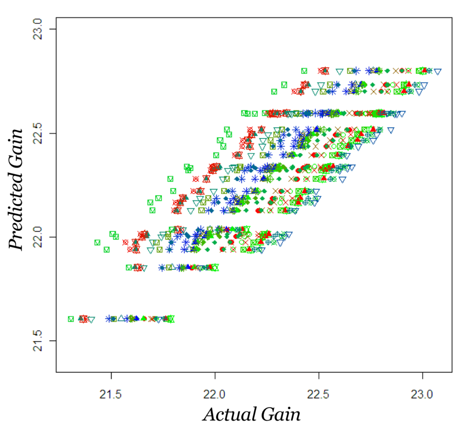


Fig. 13. Process Variation Effects

We immediately observe that process variation modulates the knob-effect model via a simple constant offset, again verifying our prior assertions that a joint model of knob and process variation simply requires adding a device constant to the knob effect model.

3) Prediction Results: Given that we can model knob effects with a linear model (Section IV-A) and process variation by modifying the linear model of knob effects and interaction terms to include a process variation term (Section IV-B), we are justified in our adoption of midpoint alternate test-based performance calibration to predict device performances across knob settings. Using the midpoint alternate test-based performance calibration methodology outlined in Section II-E, we again classified devices as healable or unhealable, with a success rate demonstrated in the confusion matrix of Table VII.

Thus, due to the use of alternate test an approximately 0.62% test escape rate and a 0.48% yield loss rate are introduced, for a slightly over 1% total error rate. We also evaluated the probability of correct heal using Equation 10 as in Section VI-F, which was reduced from 95.9% in the previous section to 94.3% for our midpoint alternate test method. Therefore, with respect to exhaustive alternate test we have achieved a lower test escape rate, a similar yield loss and probability of correct heal, and most significantly, a substantial cost reduction.

		Actual	
		Unhealable	Healable
Predicted	Unhealable	1.98%	0.48%
	Healable	0.62%	96.92%

TABLE VII
MIDPOINT ALTERNATE TEST

4) Knob Setting Selection: As we outlined in Section IV-D, once the performances have been predicted using midpoint alternate test, knob setting selection is performed by employing the Mahalanobis distance or the predicted power knob setting selection metric. Presented in Figure 14 is the power vs. correct-heal tradeoff for the knob setting selection optimality metrics: minimum power, median power, and maximum Mahalanobis distance. As can be seen from the figure, the Mahalanobis distance metric achieves a near-perfect 99.2% correct-heal rate, at the expense of high power consumption, whereas minimizing power (as expected) substantially improves power consumption, while increasing error.

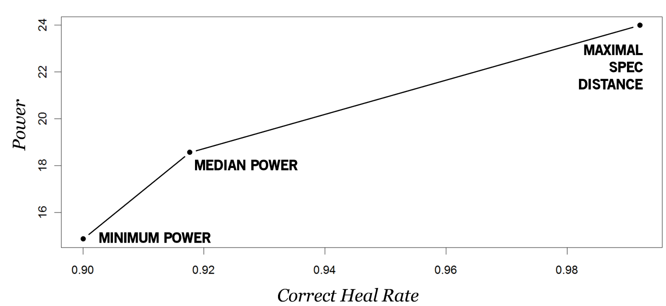


Fig. 14. Power-Prediction Quality Tradeoff

H. Training Set Cost Reduction

As discussed in Section III, the proposed midpoint alternate test-based performance calibration method incurs an initial

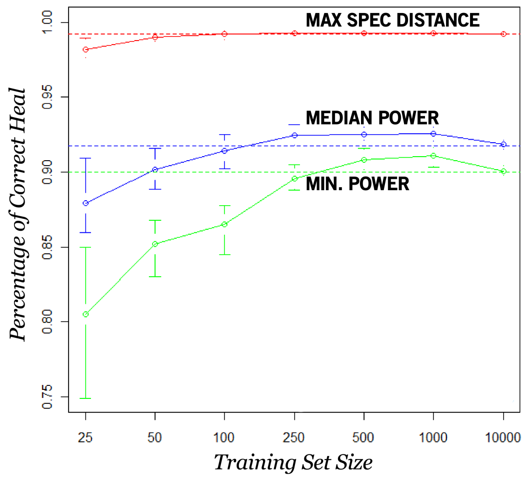


Fig. 15. Percentage of Correct Heal vs. Training Set Size

training set cost $N'_T N_K (A + P)$ proportional to the number of knob settings N_K. We found that this cost is far too pessimistic, and for real devices the number of training instances required to adequately learn the statistics of knob and process variation is actually much smaller.

To demonstrate this, we used uniform sampling to reduce the size of the training set from the initial 13,500 observations (500 devices $\times$ 27 knob settings) to 25, 50, 100, 250, 500, 1,000, and 10,000 observations. In Figure 15, we show the percentage of correct heal vs. the number of training set observations for the knob setting selection methods. For both methods, error bars are displayed for the 10 cross-validations. The horizontal dashed lines present the baseline values obtained by building models from the complete training set.

Note that in both cases, training on just 500 observations (3.7% of the original 13,500 observations) provides prediction quality on par with models constructed from the full training set. Thus, for our midpoint alternate test-based performance calibration method we can decouple the training cost from N_K, resulting in a total cost (pre-production training and production testing) that achieves low costs and total error on par with traditional alternate test, while gaining the benefits of post-production performance calibration.

VII. Conclusion

We have demonstrated that appropriate modeling of knob and process variation enables highly successful performance calibration. The proposed midpoint alternate test is a cost-effective means of introducing performance calibration methodologies into an analog/RF device test flow. Indeed, it overcomes the limitations of both iterative approaches and two-model approaches by implementing a *single model* requiring a *single alternate test measurement* step to perform tuning. This method achieves highly accurate healable/unhealable classification, with a 0.62% test escape rate and a 0.48% yield loss rate, and a 99.2% correct-heal rate using the Mahalanobis distance metric to select a knob setting on the healable devices. Finally, we have demonstrated experimentally that we can decouple the training set size from the number of knob settings N_K, requiring only a small random sample of alternate tests and performances from a handful of devices to sufficiently learn the statistics of knob and process variation.

Acknowledgments

This research has been carried out with the support of the National Science Foundation (NSF CCF-0916803 and CCF-0916415) and the Semiconductor Research Corporation (SRC-1836.029). The first author is supported by an IBM/GRC (Global Research Collaboration) graduate fellowship.

References

[1] S. Biswas and R. D. Blanton, “Statistical test compaction using binary decision trees,” *IEEE Design and Test of Computers*, vol. 23, pp. 452–462, 2006.

[2] S. Biswas, P. Li, R. D. Blanton, and L. Pileggi, “Specification test compaction for analog circuits and MEMS,” in *Design, Automation and Test in Europe*, 2005, pp. 164–169.

[3] N. Kupp, P. Drineas, M. Slamani, and Y. Makris, “Confidence estimation in non-RF to RF correlation-based specification test compaction,” in *Proceedings of the European Test Symposium*, 2008, pp. 35–40.

[4] N. Kupp, P. Drineas, M. Slamani, and Y. Makris, “On boosting the accuracy of non-RF to RF correlation-based specification test compaction,” *Journal of Electronic Testing: Theory and Applications*, vol. 25, no. 6, pp. 309–321, 2009.

[5] H.-G. D. Stratigopoulos, P. Drineas, M. Slamani, and Y. Makris, “Non-RF to RF test correlation using learning machines: A case study,” in *VLSI Test Symposium*, 2007, pp. 9–14.

[6] P. N. Variyam, S. Cherubal, and A. Chatterjee, “Prediction of analog performance parameters using fast transient testing,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 3, pp. 349–361, 2002.

[7] H.-G. D. Stratigopoulos and Y. Makris, “Error moderation in low-cost machine learning-based analog/RF testing,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 2, pp. 339–351, 2008.

[8] A. Goyal, M. Swaminathan, and A. Chatterjee, “A novel self-healing methodology for RF amplifier circuits based on oscillation principles,” in *Design, Automation and Test in Europe*, 2009, pp. 1656–1661.

[9] A. Goyal, M. Swaminathan, and A. Chatterjee, “Self-calibrating embedded RF down-conversion mixers,” in *Asian Test Symposium*, 2009, pp. 249–254.

[10] J. H. Friedman, “Multivariate adaptive regression splines,” *The Annals of Statistics*, vol. 19, no. 1, pp. 1–67, 1991.

[11] V. Natarajan, R. Senguttuvan, S. Sen, and A. Chatterjee, “ACT: Adaptive calibration test for performance enhancement and increased testability of wireless RF front-ends,” in *VLSI Test Symposium*, 2008, pp. 215–220.

[12] S.K. Devarakond, V. Natarajan, S. Sen, and A. Chatterjee, “BIST-assisted power aware self healing RF circuits,” in *Mixed-Signals, Sensors, and Systems Test Workshop*, 2009, pp. 1–4.

[13] V. Natarajan, S.K. Devarakond, S. Sen, and A. Chatterjee, “BIST driven power conscious post-manufacture tuning of wireless transceiver systems using hardware-iterated gradient search,” in *Asian Test Symposium*, 2009, pp. 243–248.

[14] D. Han, B. S. Kim, and A. Chatterjee, “DSP-driven self-tuning of RF circuits for process-induced performance variability,” *IEEE Transactions on VLSI Systems*, vol. 18, no. 2, pp. 305–314, 2010.

[15] Q. Wang and M. Soma, “RF front-end system gain and linearity built-in test,” in *IEEE VLSI Test Symposium*, 2006, pp. 228–233.