Machine Learning-based Noise Classification and Decomposition in RF Transceivers

D. Neethirajan*, C. Xanthopoulos*, K. Subramani*, K. Schaub†, I. Leventhal‡ and Y. Makris*
*Department of Electrical and Computer Engineering, The University of Texas at Dallas, Richardson, TX 75080, USA
†Advantest America, Inc., 12401 Research Boulevard, Austin, TX 78759, USA
‡Advantest America, Inc., 3061 Zanker Rd, San Jose, CA 95134, USA

***Abstract*—We propose a machine learning-based solution for noise classification and decomposition in RF transceivers. Wireless transmitters are affected by various noise sources, each of which has a distinct impact on the signal constellation points. The proposed approach takes advantage of the characteristic dispersion of points in the constellation by extracting key statistical and geometric features that are used to train a machine learning model. The trained model is, then, capable of identifying the noise source fingerprint, comprised by single or multiple noise sources, for each affected device. Effectiveness of the model has been verified using constellation measurements from a combined set of simulated and actual silicon devices.**

I. Introduction

Recent advancements in System-on-Chip (SoC) technologies have facilitated the fabrication of integrated single-chip Radio Frequency (RF) transceivers for various wireless applications. While advanced technology nodes enable production of devices that are more powerful and yet have a smaller form factor, operation of such devices is more susceptible to process, voltage and temperature (PVT) variations. As a result, the device operation is more likely to deviate from its optimal point, resulting in performance degradation and yield loss.

Cartesian transmitters, which are commonly referred to as In-phase Quadrature (IQ) transmitters, are widely used in modern wireless systems such as point-to-point radios, broadband communication systems, cellular, etc. Figure 1 shows the block diagram of a typical Cartesian transmitter. When fabricating such a transmitter topology using the latest technology nodes, the individual blocks in the transmitter chain are susceptible to the impact of process variation. As a result, the RF front-end of practical wireless devices can be affected by various impairments, resulting in quality degradation of the transmitted signal. Therefore, several solutions have been proposed in the past to improve performance and yield of fabricated wireless devices [1]–[6]. These solutions can be broadly classified into (i) post-production calibration and (ii) digital pre-distortion.

Post-production calibration is performed on Automated Test Equipment (ATE) which measure the performances of each device under test (DUT). When significant deviations are detected, calibration techniques are used to tune the device performances within their respective specification limits. Specifically, performance-control knobs present in the RF front end of each DUT are adjusted and the corresponding knob settings are stored on-chip in a non-volatile memory and used during device operation.

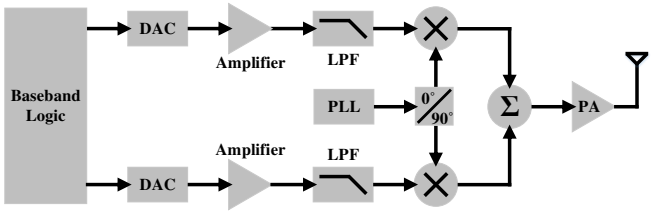

Fig. 1: Block diagram of an IQ transmitter

On the other hand, digital pre-distortion can only be performed in wireless devices that have a built-in loop-back receiver. This correction technique takes advantage of the additional receiver present in the device to down-convert and decode the transmitted messages. This allows the wireless device to observe the impairments present in its hardware and appropriately compensate the baseband IQ signals to cancel out the effects of the impairments. This solution is widely used in wireless devices that require periodic in-field calibration to self-correct performance drift.

Due to the high cost of incorporating a full loop-back receiver and the corresponding calibration logic on-chip, most wireless devices use post-production calibration as a means of correcting noise impairments. Existing calibration solutions use observation parameters such as Error Vector Magnitude (EVM) and Bit Error Rate (BER) to characterize the impact of the various noise impairments. While these parameters provide an overview of the signal quality, they do not provide the necessary information to uniquely attribute the observed performance deviation to the corresponding impairment(s).

To this end, in this work, we propose a machine learning-based noise classification and decomposition technique for RF transceivers. The proposed method enables the unique identification of the impairments that affect the transmitter and predicts the corresponding magnitude of each such noise source. Effectiveness of the proposed approach has been evaluated using constellation measurements from a combination of simulated and actual silicon devices.

The remainder of this paper is organized as follows. In Section II, we discuss the impact of noise impairments on the performance of a Cartesian transmitter. Existing techniques for correcting these noise impairments are presented in Section III. The proposed approach is introduced in Section IV, where we address the feature extraction and implementation details of the machine learning-based model. Then, in Section V, we in-

978-1-7281-1170-4/19/$31.00 ©2019 IEEE

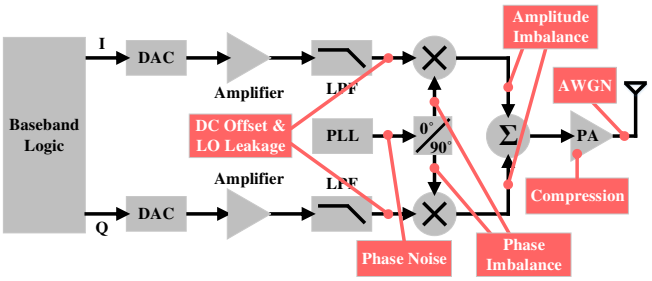

Fig. 2: IQ transmitter with impairments

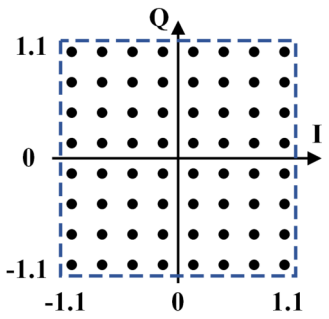

Fig. 3: 64-QAM constellation

troduce a simulation model that was constructed for generating synthetic devices, which were used along with measurements from actual silicon devices to evaluate the performance of the proposed noise classification and decomposition technique. Finally, conclusions are drawn in Section VI.

II. Impairments in a Cartesian Transmitter

Cartesian transmitters have gained popularity in modern wireless devices due to their image rejection capability, power efficiency and ease by which they modulate the baseband data bits onto the RF carrier signal. Figure 1 shows the block diagram of a typical Cartesian transmitter. Here the baseband logic, which implements the wireless standard and the corresponding signal processing blocks, generates two data streams, namely In-Phase (I) and Quadrature (Q). The generated data streams, which are in the digital domain, are first converted to analog using Digital-to-Analog (DAC) converters. Next, the translated signals are processed by dedicated data paths in the RF front-end, where they are amplified, filtered and up-converted to RF frequencies. The two RF signals are then combined and the resulting signal is amplified by a power amplifier and transmitted over-the-air using an antenna.

In recent years, however, transmitter performances have been significantly affected by increased process variation, which introduces undesired impairments in the fabricated devices. Figure 2 illustrates a number of different types of such impairments resulting from imperfections at various transmitter blocks. To understand the influence of these impairments, consider an example where the device modulates the transmitted signal using a 64-Quadrature Amplitude Modulation (QAM) scheme. Figure 3 shows the ideal constellation points for the 64-QAM on the two-dimensional Cartesian I-Q space. In the presence of noise, however, the received constellation demonstrates a characteristic dispersion of the points which deviate from their ideal positions.

Figure 4 shows the received constellation for the various impairments introduced in Figure 2. Each plot exhibits some unique characteristics, that can be decomposed to several basic features. These features include the rotation, shift, and compression of the entire constellation, as well as the linear or radial dispersion pattern for each of the sub-constellations. The type and degree by which these features are present in each plot is indicative of the noise type and its impact on the device operation. As mentioned before, these variations significantly

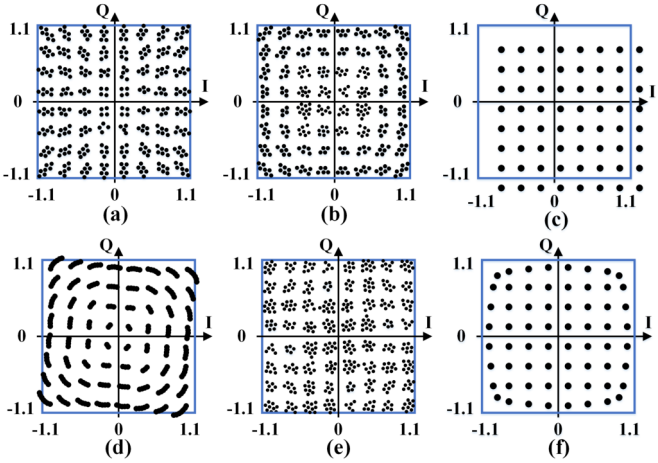

Fig. 4: Received constellation from a transmitter affected by: (a) amplitude imbalance, (b) phase imbalance, (c) DC Offset and LO leakage, (d) phase noise, (e) AWGN and (f) gain compression

affect the performance and yield of the fabricated wireless device. In order to improve the yield, in the example case of a 64-QAM transmitter without a built-in loop-back receiver, post-production calibration is necessary.

III. Related Work

Unlike digital circuits, analog and RF ICs are more susceptible to the impact of process variations. As a result, several post-production calibration techniques have been proposed over the years [7]–[11], with the aim of simultaneously optimizing performance and yield of the targeted device. However, when dealing with complex systems such as a wireless transceiver, calibration becomes an extremely challenging task. This is evident in a Cartesian transmitter, where the various impairments within the device manifest in the transmitted signal, thereby affecting its quality.

In the past, several studies have investigated the impact of these impairments on the target device [2], [12]–[17]. In [2], [14], [15], the authors propose a loop-back testing approach, where the DUT is excited using a carefully selected test signal, such that the device response contains sufficient information to determine its performance characteristics. Based on the observed performance traits, digital pre-distortion is used to

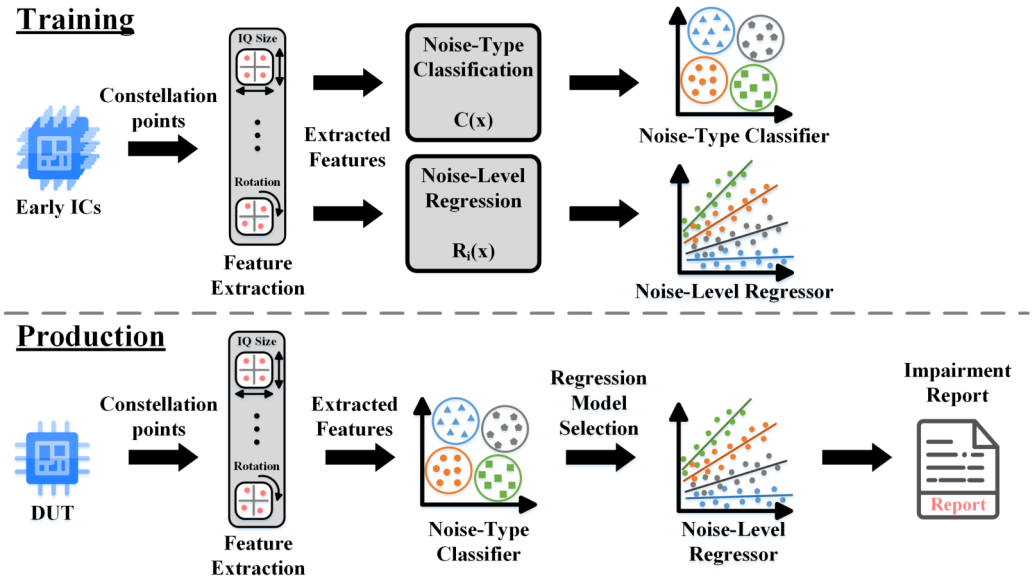

Fig. 5: Machine learning-based noise classification and decomposition

correct the impact of the impairments. Effectiveness of this method relies on identifying and generating the test signal with certain characteristics, which may not always be easy to establish or may depend on the performance parameter under study. In [13], [18], [19], the authors use on-chip sensors for calibration, where the sensor outputs act as low-cost measurements for predicting the RF performances of the device. Efficacy of such techniques relies on sensor performance, which may also be affected by process variation.

In contrast to all the studies noted above, the proposed method: (i) does not rely on any traits of the input test signal, and (ii) does not require additional logic, such as sensors or monitors, to be added on the device. Instead, it uses existing capabilities of an ATE to measure the constellation data points, from which it predicts the magnitude of the various impairments using the proposed statistical model.

IV. Proposed Approach

The proposed approach aims to complement the existing calibration techniques that are based on EVM and BER, by providing the type and magnitude for each impairment affecting the performances of a wireless device. To achieve this, various machine learning-based models are used to infer the impairment types and their corresponding levels as a function of the constellation points. Figure 5 provides an overview of the proposed two-phased approach. During the first phase, the underlying supervised learning models are trained using data from a set of early manufactured devices. The trained models are then used during production for the identification and decomposition of the noise source into its constituents and, thereby, for guiding device calibration.

The first step in both phases of the proposed method is feature extraction, which aims to create a feature vector representative of the noise sources affecting the performance of each device. As shown in Figure 4, the constellation plots carry enough information to be used as fingerprints revealing the type of impairments exhibited by the device.

Once the feature vector is selected, a classifier is trained to identify the noise source based on a labeled dataset from early devices whose noise source has been manually identified or synthetic devices which have been purposely generated to represent their respective noise source. Once trained, the confidence estimate produced by the classifier can be used for determining the noise source contributing to the constellation variation based on the measured constellation of a device.

Similar to the classifier, the same feature vector can be used to train regression models for each observed impairment in order to estimate the degree of noise that is present. We note that a separate regression model is needed per impairment, as the input features must be uniquely weighted based on the impairment type whose magnitude we seek to predict.

Sections IV-A and IV-B describe the details of each of the two major steps in the proposed approach.

A. Feature Extraction

Effectiveness of the proposed approach relies on the selection of features that can uniquely identify the impact of the various impairments on the constellation. For this purpose, we selected several statistical and geometric features representing i) alignment of each sub-constellation, ii) dispersion for each sub-constellation, iii) constellation shift, iv) constellation rotation, and v) total size of the constellation. To measure sub-constellation alignment, we fit a line within each sub-constellation using *least-squares regression*. The angle of the fitted line is recorded as a feature called θ, which reflects the alignment of the sub-constellation with respect to the I axis. The size of each sub-constellation is measured using the *smallest enclosing circle* algorithm, where the diameter of the circle corresponds to the sub-constellation size, λ. For extracting the remaining features, a common step is

the identification of the sub-constellation centers. To this end, the constellation points are first clustered using the *K-Nearest Neighbors (K-NN)* algorithm and labeled based on their respective sub-constellation. For each sub-constellation point-cloud, the I and Q coordinates of the center are then calculated. Based on these coordinates, any potential shift from the ideal constellation points can be calculated and used as a feature. Next, the sub-constellation center that has the maximum magnitude within each quadrant is identified and used for the calculation of the next feature, namely the angle of rotation of the overall constellation with respect to the origin. Finally, the two features representing the total size of the constellation are calculated by finding the distance between the maximum and minimum coordinate for each axis.

Altogether, for 64-QAM, we extract 133 features from the constellation of each device, which are used to train the two types of machine learning-based models from devices in the training set and to obtain the impairment decomposition report for devices in production.

B. Modelling

Details of the classification and regression models used in this work are provided below.

1) Noise Type Classification: A single classification model is used to identify the impairment type based on the various features described above. Specifically, a Support Vector Machine (SVM) with a linear kernel is used for this purpose. The classifier is trained with devices that have been affected only by a single noise type. As a result, we rely on the probability estimate of each label to identify the contribution of multiple noise impairments. This approach simplifies the collection and labeling of training samples, as it does not require all combinations to be present in the training set. On the other hand, using the probability estimates of a single classifier can result in having the prominent noise source dominate the others. For the purpose of this work, this limitation does not hinder the effectiveness of the proposed approach as it would result in the identification of the dominant noise source, the calibration of which must be prioritized. After such calibration, the dominant noise impairment will be eliminated and the newly measured constellation will reveal the impact of other weaker noise impairments in the combination. This sequence of prediction and calibration will continue until there are no more impairments in the DUT.

2) Noise Magnitude Regression: For the estimation of the noise magnitude of each impairment, we use regression models trained using the same feature vector as the classifier mentioned in Section IV-A. Here, a Multivariate Adaptive Regression Splines (MARS) model is used for this purpose [20]. During production, the probability estimates of the noise-type classifier will be used to select the appropriate regression models. More specifically, every noise-type label that has a probability estimate value above 0% invokes its corresponding trained regressor to predict the magnitude of that noise. Existing calibration techniques can then take advantage of the predicted noise impairment magnitude to robustly calibrate the performance of the DUT.

Noise Impairment	Range
Amplitude Imbalance	0.1dB : 1.0dB
Phase Imbalance	1deg : 10deg
AWGN	15dB : 35dB

TABLE I: Simulated noise impairments and their range

V. Experimental Results

In this section, we first introduce the simulation model used for creating our dataset and we evaluate the accuracy of the proposed method on the synthetically generated devices. Then, we also assess its effectiveness on actual silicon devices.

A. Simulation Model

In our experiments, we model an Orthogonal Frequency-Division Multiplexing (OFDM) transmitter as the DUT and an OFDM receiver as the ATE. Performances of the individual blocks in the system, such as input-output return loss, gain, noise figure, non-linearity, phase noise, etc., are based on off-the-shelf components, whose specification margins have been imported from the corresponding datasheets. Therefore, when running Monte-Carlo simulations for 100 devices, performance characteristics for the individual blocks are selected randomly from their specification margins and the device performance traits can be reproduced across simulations by setting the seed for the Monte-Carlo run.

In our setup, we consider three different noise impairments, namely amplitude imbalance, phase imbalance and Additive White Gaussian Noise (AWGN). The effect of these impairments on the received constellation are studied for both individual noise conditions, as well as combinations thereof. Table I specifies the range of each noise impairment. The first impairment, namely amplitude imbalance, is caused by gain mismatch between the I and Q data paths. Effects of this mismatch are simulated by adjusting the gain of the blocks that constitute the two data paths and the corresponding mismatch level is represented in dB. The second impairment, namely phase imbalance or quadrature skew, is caused by imperfections in the LO splitter, where the phase difference between its outputs deviates from 90°. Therefore, to simulate this impairment, the phase difference between the I and Q mixer's LO input has been varied and its value is denoted in degrees. Random fluctuations and noise inherent in the measurement setup determine the repeatability of the experiments. The third and final impairment, has been modeled using an AWGN block at the output of the transmitter chain and the noise level is represented in terms of Signal-to-Noise Ratio (SNR). Using this system model, constellation data corresponding to different noise impairment combinations have been collected and are used to train and evaluate the effectiveness of the proposed approach.

B. Results for Synthetic Data

Table II shows the various noise impairment combinations and their alias number that will be used hereinafter. For the

Noise Impairment Combination	Alias
No Impairment	(0)
Amplitude Imbalance	(1)
AWGN	(2)
Phase Imbalance	(3)
Amplitude Imbalance + Phase Imbalance	(4)
Amplitude Imbalance + AWGN	(5)
Phase Imbalance + AWGN	(6)
Amplitude Imbalance + Phase Imbalance + AWGN	(7)

TABLE II: Noise combinations and their alias

Noise Impairment	(0)	(1)	(2)	(3)
(0)	100	0	0	0
(1)	0	200	0	0
(2)	0	0	200	0
(3)	0	0	0	200

TABLE III: Confusion matrix for classification of synthetic devices - single noise impairments

constellation data collected from the theoretical model, the 133 features mentioned in Section IV-A were first extracted for all noise combinations. Next, the features corresponding to single noise impairments were used to train an SVM-based multi-class classifier. Table III shows the classification accuracy of the multi-class classifier when tested using single-noise impairments. Here, the test dataset for every single noise impairments consisted of 200 devices, for which the model achieved 100% accuracy for all three single-noise impairments (1), (2) and (3). For the second step of the proposed approach, MARS regression was performed to identify the magnitude of the noise impairments. Figure 6 shows the prediction accuracy of the regression models for the three single-noise impairments. As can be observed through the fitted 45° line, the predicted values accurately track the actual measurements.

In order to evaluate the effectiveness of the classifier, an equivalent of the confusion matrix needs to be defined for the various combinations of noise impairments. In other words, we need to define what constitutes a correct and incorrect

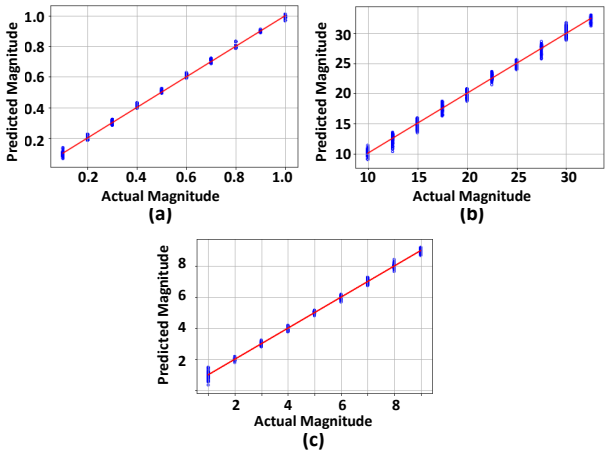

Fig. 6: Predicted magnitude for single noise impairment: (a) amplitude imbalance, (b) AWGN and (c) phase imbalance

Noise Impairment	Misclassification Percentage
(4)	0
(5)	0
(6)	7.33
(7)	0

TABLE IV: Classification report for synthetic devices with combinations of noise impairments

Noise Impairment	Mean Relative Error
(4)	10.99
(5)	7.00
(6)	1.20
(7)	9.40

TABLE V: Noise magnitude prediction error for combination of noise impairments

classification in all cases. As explained in Section IV, in the presence of a dominant noise impairment, identification of the weaker noise impairments in the combination may not always be possible. In such cases, in order to replicate what would happen in production, we use simulated calibration. This means that the model iteratively classifies the IQ data and calibrates the noise based on the impairment report, until the DUT is classified under the no-impairment class. If the model can successfully predict the complete succession of noise impairments, then there will be no misclassification in the proposed approach. In other words, a misclassification will occur if, at any iteration, the model predicts a noise impairment that is not present in the combination or prematurely terminates the simulated calibration.

Given the above, the two trained models were also evaluated using different combinations of noise impairments. Table IV shows the classification accuracy of the multi-class classifier for the various noise combinations. For noise combinations (4) and (5), the classifier operates with 0% error. However, in the case of (6), a misclassification of 7.33% is observed. By visually inspecting the relevant constellation graphs, it appears that the combination of Phase Imbalance and AWGN at very high noise levels creates an impact on the constellation points which resembles the effect of an Amplitude Imbalance, thereby explaining the observed classification error.

Based on the classifier output, we use the appropriate noise regression model to estimate the magnitude of the noise impairment. Table V shows the prediction accuracy in terms of mean relative error (MRE) for each combination. MRE is defined as the average absolute error over the specification range. Overall, the model predicts with a MRE of less than 11%, with the highest error corresponding to noise combination (4).

C. Results for Silicon Data

The silicon dataset was provided by Advantest and consisted of measurements from four devices that were affected by two different noise impairments, namely Amplitude Imbalance and Phase Imbalance. In the first step, following the flow of the proposed approach, we use the multi-class classifier which

Noise Impairment	(0)	(1)	(3)
(0)	0	0	0
(1)	1	19	0
(3)	5	0	15

TABLE VI: Confusion matrix for classification of silicon devices

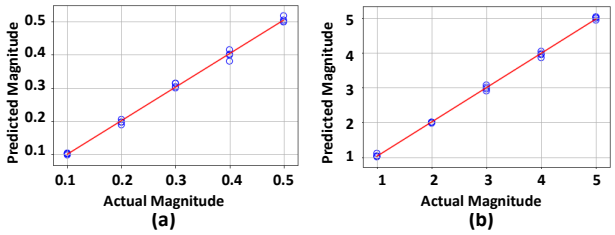

Fig. 7: Predicted magnitude for Silicon Devices: (a) amplitude imbalance and (b) phase imbalance

was trained using measurements from the synthetic devices to classify the noise impairments in the silicon dataset irrespective of the silicon devices' architecture. The classification results are shown in Table VI. As may be observed, there are a few classification errors, pushing some of the devices with the two noise impairment types (1) and (3) available in the silicon dataset towards the no impairment type (0). Upon further examination, we observed that these devices correspond to an extremely low noise level of 0.5° at which the constellation experiences negligible impact caused by the impairments, thereby explaining the misclassification.

Once again, based on the classifier output, the appropriate regression model is chosen to predict the noise magnitude level. The results for this step are shown in Figure 7 in terms of actual vs. predicted magnitude values for the two noise impairment types, with the proximity of the points to the 45° line confirming the accuracy of the predictions.

VI. Conclusion

In this paper, a machine learning-based solution for noise classification and decomposition in RF transceivers has been presented. The proposed method leverages the distinct impact caused by different noise impairments on the transmitted signal constellation by extracting certain key features from the constellation data points. Thereby, a multi-class classifier is used to determine the type of noise impairment that is affecting the transmitted signal and a regression model is used to determine the corresponding magnitude of the impairment. The proposed method has been evaluated using constellation measurements from a combined set of simulated and actual silicon devices. Experimental results show that the statistical model can classify and decompose the noise type with very high accuracy.

VII. Acknowledgment

This work was partially supported by Advantest America, Inc. and by the Semiconductor Research Corporation (SRC) task 2712.031 through The University of Texas at Dallas' Texas Analog Center of Excellence (TxACE).

References

[1] V. Natarajan, S. Sen, S. K. Devarakond, and A. Chatterjee, "A Holistic Approach to Accurate Tuning of RF Systems for Large and Small Multiparameter Perturbations," in *VLSI Test Symposium (VTS)*, 2010.

[2] A. Nassery, J. W. Jeong, and S. Ozev, "Zero-overhead Self Test and Calibration of RF Transceivers," in *International Test Conference (ITC)*, 2013.

[3] E. Acar and S. Ozev, "Digital Calibration of RF Transceivers for I-Q Imbalances and Nonlinearity," in *International Conference on Computer Design (ICCD)*, 2007.

[4] J. W. Jeong, S. Ozev, S. Sen, V. Natarajan, and M. Slamani, "Built-In Self-Test and Characterization of Polar Transmitter Parameters in the Loop-Back Mode," in *Design, Automation & Test in Europe (DATE)*, 2014.

[5] D. Lee, A. Chaterjee, J. A. Copeland, and S. Shin, "Error Tolerance in Wireless OFDM Data Transmission using Signal Quality Driven Symbol Re-mapping," in *Consumer Communications and Networking Conference (CCNC)*, 2011.

[6] A. Banerjee and A. Chatterjee, "Signature Driven Hierarchical Post-Manufacture Tuning of RF Systems for Performance and Power," *Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, 2015.

[7] Y. Lu, K. S. Subramani, H. Huang, N. Kupp, K. Huang, and Y. Makris, "A Comparative Study of One-Shot Statistical Calibration Methods for Analog/RF ICs," in *International Test Conference (ITC)*, 2015.

[8] G. Volanis, D. Maliuk, Y. Lu, K. S. Subramani, A. Antonopoulos, and Y. Makris, "On-Die Learning-Based Self-Calibration of Analog/RF ICs," in *VLSI Test Symposium (VTS)*, 2016.

[9] V. Natarajan, A. Banerjee, S. Sen, S. Devarakond, and A. Chatterjee, "Yield Recovery of RF Transceiver Systems Using Iterative Tuning-Driven Power-Conscious Performance Optimization," *Design & Test*, vol. 32, 2015.

[10] M. Andraud, H.-G. Stratigopoulos, and E. Simeu, "One-Shot Non-Intrusive Calibration Against Process Variations for Analog/RF Circuits," *Transactions on Circuits and Systems I: Regular Papers*, vol. 63, 2016.

[11] S. Sun, F. Wang, S. Yaldiz, X. Li, L. Pileggi, A. Natarajan, M. Ferriss, J.-O. Plouchart, B. Sadhu, B. Parker, A. Valdes-Garcia, M. A. T. Sanduleanu, J. Tierno, and D. Friedman, "Indirect Performance Sensing for On-Chip Self-Healing of Analog and RF Circuits," *Transactions on Circuits and Systems I: Regular Papers*, vol. 61, 2014.

[12] J. Tubbax, A. Fort, L. Van der Perre, S. Donnay, M. Engels, M. Moonen, and H. De Man, "Joint Compensation of IQ imbalance and Frequency Offset in OFDM systems," in *Global Telecommunications Conference (GLOBECOM)*, 2003.

[13] S. Bhattacharya and A. Chatterjee, "A Built-In Loopback Test Methodology for RF Transceiver Circuits using Embedded Sensor Circuits," in *Asian Test Symposium*, 2004.

[14] A. Halder, S. Bhattacharya, G. Srinivasan, and A. Chatterjee, "A System-Level Alternate Test Approach for Specification Test of RF Transceivers in Loopback Mode," in *International Conference on VLSI Design*, 2005.

[15] I. Bouras, S. Bouras, T. Georgantas, N. Haralabidis, G. Kamoulakos, C. Kapnistis, S. Kavadias, Y. Kokolakis, P. Merakos, J. Rudell, S. Plevridis, I. Vassiliou, K. Vavelidis, and A. Yamanaka, "A Digitally Calibrated 5.15-5.825 GHz Transceiver for 802.11a Wireless LANs in 0.18μm CMOS," in *International Solid-State Circuits Conference (ISSCC)*, 2003.

[16] A. Tarighat and A. H. Sayed, "On the Baseband Compensation of IQ Imbalances in OFDM Systems," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2004.

[17] D. Tandur and M. Moonen, "Digital Compensation of RF Imperfections for Broadband Wireless Systems," in *Symposium on Communications and Vehicular Technology in the Benelux*, 2007.

[18] M. Onabajo, F. Fernandez, J. Silva-Martinez, and E. Sánchez-Sinencio, "Strategic Test Cost Reduction with On-Chip Measurement Circuitry for RF Transceiver Front-Ends–An Overview," in *International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2006.

[19] C. Zhang, R. Gharpurey, and J. A. Abraham, "Low Cost RF Receiver Parameter Measurement with On-Chip Amplitude Detectors," in *VLSI Test Symposium (VTS)*, 2008.

[20] J. H. Friedman, "Multivariate Adaptive Regression Splines," *The Annals of Statistics*, vol. 19, 1991.