Fast EVM Tuning of MIMO Wireless Systems Using Collaborative Parallel Testing and Implicit Reward Driven Learning

Suhasini Komarraju* and Abhijit Chatterjee*
*School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA
Email: skomarraju3@gatech.edu, abhijit.chatterjee@ece.gatech.edu

2020 IEEE International Test Conference (ITC) | 978-1-7281-9113-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/ITC44778.2020.9325270

***Abstract*—Modern 5G and projected 6G wireless systems deploy massive MIMO systems with antenna arrays and novel RF transceiver architectures that admit RF beamforming. Testing and tuning of the underlying transceiver arrays on a per-transceiver basis is expensive and can be expedited through the use of parallel testing and tuning techniques that stimulate the entire array transceiver system concurrently. State of the art parallel testing techniques require frequency separation between the tones applied to individual RF chains due to combining of RF signals before down-conversion in analog beamforming MIMO systems. Test schemes that allow some frequency overlap are limited to testing only third order distortion. In this paper, we first present a parallel testing scheme for testing large MIMO transceiver arrays that is amenable to higher order distortion (upto fifth order) in the RF chains considered. Second, we propose a tuning scheme for the entire MIMO array which *implicitly tunes for EVM* system specifications *without explicit knowledge* of the relationship between the system test response, the system tuning knobs and the corresponding EVM and SINR specification values. A cost metric is formulated that allows such a solution using reinforcement (multi-arm bandit) learning driven system tuning. Significant yield improvement using this approach is demonstrated by simulation experiments.**

***Index Terms*—Massive MIMO, Testing, Tuning, Multi-Arm Bandit**

I. Introduction

Fueled by emerging 5G and 6G wireless communications technologies, data rates will increase by up to 1000X by 2020 as compared to the state of the art [1]–[3]. To address this, multiplexing, diversity, coding and antenna gains along with interference suppression through the use of antenna beamforming techniques have fueled a new generation of wireless systems architectures and infrastructure. However, with the dramatically increasing levels of circuit complexity and higher operating speeds, the underlying electronics becomes highly susceptible to manufacturing process variations and the effects of device non-idealities on overall system performance. Power consumption also becomes a major issue since a large number of transmitter and receiver chains are involved.

There has been significant work in the past on built-in self-test and tuning of SISO and MIMO wireless systems [4]–[6]. The 'golden response tuning' approach of [5] initially developed for SISO systems is leveraged in this research. In [7], a self-compensating built-in test scheme for phased array RF chains is presented. The parameters of the BIST circuitry are not assumed to be known, neither are precise input signal amplitudes. The gains and relative phases of each chain are determined with high accuracy. In [8], a built-in self test technique is presented for phased arrays. The individual amplitudes and phases of all elements are determined using code modulation. A low cost scheme for measuring third order distortion and phase imbalance in MIMO RF chains is presented in [9]. This work presents a novel frequency multiplexing scheme that allows the nonideal parameters of individual MIMO chains to be computed from a single combined and down-converted time domain response to test tones input to each of the individual chains. The method uses three tones to test for the Gain, IIP3 and phase specifications of a pair of RF chains with frequency separation across pairs of concurrently tested RF chains. An advancement to testing of phased array systems that includes gain and phase mismatches of combined antenna and RF chain systems is discussed in [10]. These mismatches are de-embedded from near-field measurements of transmitter electromagnetic fields resulting in contact-less testing of phased array systems. A methodology for testing and tuning MIMO beamforming systems with frequency overlapped testing tones as test inputs to individual RF chains was presented in [11]. The method was limited to testing of third order nonlinearity in the RF chains. Further, supervised learning methods were used to learn the relationships between the test response of the system to the applied tests, the corresponding EVM and SINR values and the optimal tuning knob configurations for tested devices. Such supervised learning across large antenna arrays and their corresponding RF chains requires the acquisition of significant numbers of devices and advanced learning algorithms and outlier detection methods. This incurs significant costs in device handling and data management and in training and update of relevant learners and databases.

In contrast, a new testing approach for massive-MIMO systems is presented in this research that:

- Allows parallel tuning of up to 5th order nonlinearity as well as gain and phase mismatches of MIMO RF systems with a *minimum of test frequencies* across all the chains, thereby minimizing the complexity of test instrumentation. Gain and phase estimation is 1.7X more

978-1-7281-9113-3/20/$31.00 ©2020 IEEE

frequency-efficient than other comparable test methods.

- Uses an *implicit tuning metric* that does away with the need to accurately assess system specification values at each tuning iteration, as with other current post-manufacture tuning techniques, without compromising tuning accuracy. This eliminates significant testing costs associated with current MIMO testing methods while accelerating tuning speed.
- A novel multi-arm bandit reinforcement learning formulation of the post-manufacture tuning problem is used to improve device yield with respect to EVM and SINR specifications. Rewards are computed on the implicit tuning metric described above. Fast convergence is achieved through a response clustering based initialization of the learning process.
- The tuning solution for a die is used as the starting tuning knob configuration for tuning of adjacent die on the wafer allowing for significant savings in tuning time. This uses the knowledge that process parameters are partially correlated in neighboring regions of a wafer and reduces the number of iterations to reach convergence by about 10X.

In the following, we first describe the key contributions and approach in Section II. This is followed by a description of the core circuit infrastructure in Section III. We present the proposed concurrent testing approach in Section V. Finally, the proposed tuning algorithm and method for achieving faster convergence and experimental results are discussed in Sections VII, VIII and IX respectively. Finally, conclusions are presented in Section X.

II. Approach and contributions

We consider a two-dimensional (4 by 4) MIMO receiver array as a test vehicle for the proposed tuning approach. An antenna array as shown in Figure 1 is assumed to provide inputs to the system shown in Figure 2. Testing of the antenna itself is not considered in this research, rather the goal and focus of this work is on rapid tuning of the active RF chains across a range of beam steering angles (elevation, azimuth). We assume test access to the RF system through RF switches (See switch in Figure 2) or signal coupling mechanisms such as described in [7]. Given such a test architecture, the following steps define the proposed parallel testing and implicit tuning approach.

Step 1. RF Phase Offset Determination for Different Beam Steering Angles: Prior to the start of the tuning process, for each beam steering direction, the required phase offsets of all the 16 RF chains in the 4 by 4 array of receivers are determined. These serve as targets for parallel phase tuning of the receiver array.

Step 2. Parallel Gain and Phase Estimation: All the RF chains are stimulated in parallel with predetermined signals. The range of test frequencies needed is significantly reduced over prior techniques [11]. The relative phase and gain values of each RF chain are determined from the combined baseband signal obtained as $s_{out}(t)$ given by Equation 7. These values are compared against known amplitude and phase values for validation purposes (Section V).

Step 3. Parallel Tests for Determining Upto 5th Order Nonlinearities: Using two tones per RF chain is necessary for determining upto 5th order distortion effects. To estimate the same for several RF chains simultaneously, requires the tones to be separated in frequency. We propose a technique to choose the frequency tones judiciously without compromising the accuracy of with which RF chain nonlinearities can be determined. This is done by allowing some of the tones generated due to higher order nonlinearities to overlap in frequency without compromising individual contributions of tones from the combined output signal.

Step 4. Use of an Implicit Reward Driven Reinforcement Learning Based Tuning Procedure: As opposed to prior techniques [11], we do not determine the EVM, SINR specifications of the transceiver in each tuning iteration. Instead, the response of an ideal device to the applied tests described above is computed (time-domain). We first show that a distance based metric that defines the difference between the observed response of the receiver system and the response of the ideal device (called the 'golden' response) is a good proxy for the EVM and SINR of the tested device. By minimizing this distance based metric implicitly, EVM and SINR specifications are improved. Subsequently, a multi-arm bandit or Q-learning reinforcement learning approach using the distance metric as reward is used to rapidly tune the receiver array.

Step 5. Rapid Tuning Via Tuning Parameter Initialization from Neighboring Die: The tuning solution for one die is used as the starting condition for the tuning processes of neighboring die. Since process parameters of die in close proximity on a wafer are partially correlated [12], significant savings in the numbers of tuning iterations is achieved.

In the following, each of the above steps is described in greater detail along with an additional description of the test setup and simulation models.

III. RF Phase Offset Determination Across Beam Steering Angles

We use a simulation model of a 4x4 phased array system in this research. The antenna array is a square array where each element is separated by a distance 'd' from its four nearest neighbors. We assume that the main beam of the antenna array as shown in Figure 1 points in the particular direction (elevation, azimuth) = (θ, ϕ). If β_x, β_y represent progressive phase shifts of the transmitted or received waves at grid points along the x and y axes, d_x, d_y represent the distance between elements along the x and y directions, respectively, and k represents the wave number given by $k = 2\pi/\lambda$ where λ is the wavelength of transmission, then Equation 1 describes the relationships amongst all the above parameters that must be satisfied [13] from physics considerations. For such a system with specified elevation(θ) and azimuth(ϕ), the array factor is given by Equation 2. If m and n represent the x and y coordinates of individual elements in the array, respectively, then each individual antenna element angles must satisfy Equation

3 for the main lobe of individual elements to point in the direction given by (θ, ϕ) [13]. The individual phase shifts determined by parameters β_x, β_y are obtained by solving the system of equations above using Matlab and the sinusoidal input represented by $x_i(t)$ given by Equation 4 is applied to the ith RF chain. The entire receiver (antenna) array is simulated for the angles of θ 0°, 30°, 60° and 90°. The values of ϕ are varied from 0° to 180° in steps of 45°. For each combination of θ and ϕ, the phases ψ_i of Equation 4 for each receiver in the 4 by 4 receiver array are calculated. *These phase values are later used as target phase values* to which each of the RF chains is tuned.

$$\tan(\phi) = \frac{\beta_y d_x}{\beta_x d_y}$$
$$\sin(\theta) = \sqrt{(\frac{\beta_x}{kd_x})^2 + (\frac{\beta_y}{kd_y})^2} \tag{1}$$

$$\mathrm{AF} = I_0 \sum_{m=0}^{M-1} e^{kd_x \sin(\theta)\cos(\phi)+\beta_x} + \sum_{n=0}^{N-1} e^{kd_y \sin(\theta)\cos(\phi)+\beta_y} \tag{2}$$

$$kd_x \sin(\theta)\cos(\phi) + \beta_x = 0$$
$$kd_y \sin(\theta)\cos(\phi) + \beta_y = 0 \tag{3}$$

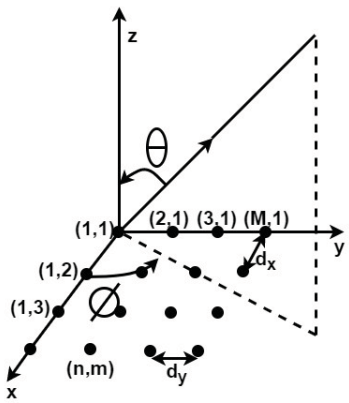

Fig. 1: Rectangular array

IV. Test Setup and Simulation Models

The input baseband signals, $x_i(t)$ in Figure 2 are multiplied with the modulating signal of frequency f_{modi} given by Equation 5. The resulting signal obtained after modulation with $m_i(t)$ is then applied to the ith RF chain of the system. The resulting chain output is then combined and demodulated using the signal of frequency f_{demod} given by Equation 6. In this case, it is assumed that the f_{modi} and f_{demod} are equal. The resulting signal is given by Equation 7 where M_i is dependent on path gain, A_i, A_{modi} and A_{demod}.The overall architecture of the system is given in Figure 2. Since the algorithms presented here do not depend on the number of RF chains involved, the same can be extended for 16 by 16 or larger antenna array systems. The analyses performed here are based on LNA, mixer and phase shifter designs in 130 nm technology which are described next.

$$x_i(t) = A_i \sin(2\pi f_i t + \psi_i) \tag{4}$$

$$m_i(t) = A_{modi} \sin(2\pi f_{modi} t) \tag{5}$$

$$d(t) = A_{demod} \sin(2\pi f_{demod} t) \tag{6}$$

$$s_{out}(t) = \sum_{i=1}^{N} M_i \sin(2\pi f_i t - \psi_i) \tag{7}$$

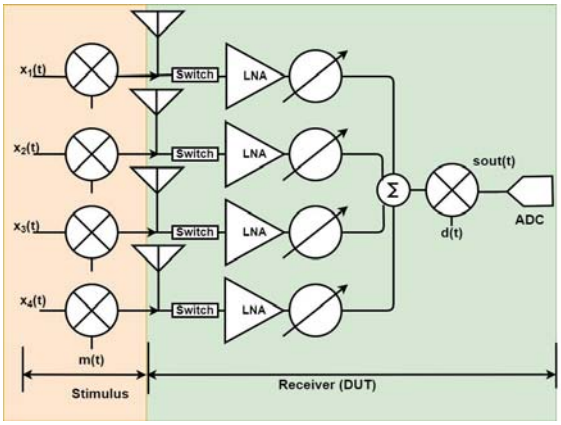

Fig. 2: Test Setup and MIMO system architecture

A single input to differential output LNA with noise cancellation presented in Figure 3 is considered here. The detailed design of the LNA is given in [14]. Since the output from LNA is differential, a fully balanced Gilbert cell type mixer is designed based on the design calculations in [15]. The circuit schematic of the mixer is as shown in Figure 3. The current sources presented in the design in [15] are removed and balanced with common source NMOS transistors to decrease power consumption. To increase the conversion gain of the system, the PMOS loads are replaced with resistances in our design. The phase shifter circuit shown in Figure 4 is implemented in 130nm technology. A quadrature all pass filter is initially used to obtain I/Q signals from the input signal. In this design, the relative strengths of the DI and DQ currents are used to determine phase shift value given by Equation 8.

$$\theta = tan^{-1}(\frac{Q}{I}) \tag{8}$$

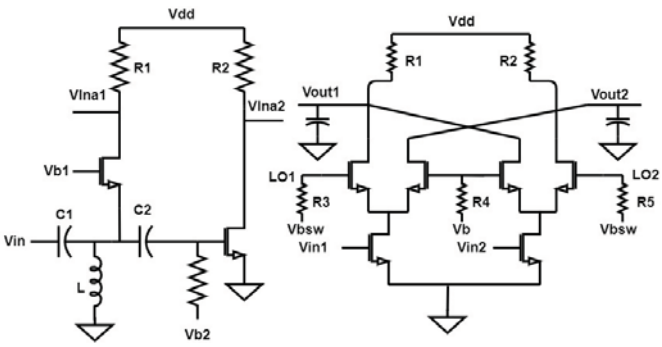

Fig. 3: Low Noise Amplifier and Mixer

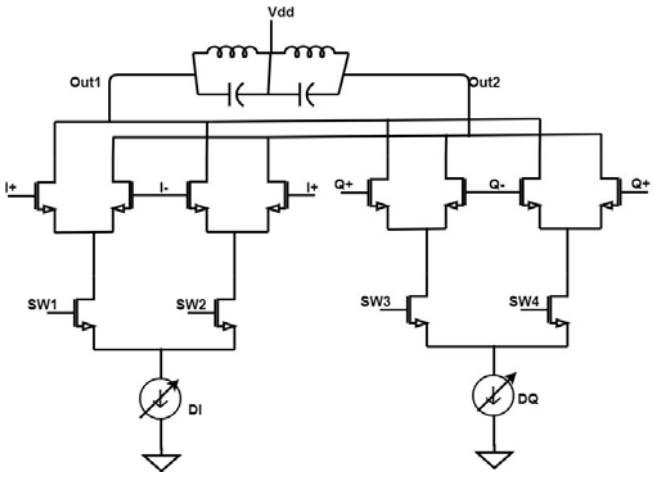

Fig. 4: Phase shifter circuit

TABLE I: Performance specifications of sample RF chain

Parameter	Achieved specification
VDD	2.5V
DC power consumption	45mW
S_{11}	$\leq$-11dB
Frequency bandwidth	2.4-8GHz
Peak conversion gain	31.2dB
Noise figure	4.2dB
Input P1dB	-15dBm
P_{IIP3}	-10dB
Inductance used	2.9nH

Each RF chain can be tuned for different values of gain, phase shift and non-linearities by varying appropriate parameters in the design. The Vb1 and Vb2 voltages shown in Figure 3 are varied to tune the performance of the LNA. Conversion gain of the mixer is tuned using the Vb and Vbsw voltages shown in Figure 3. The performance parameters obtained for a RF chain with designed LNA, phase shifter and mixer is given in Table I. The gain of each RF chain varies with power as shown in Figure 5.

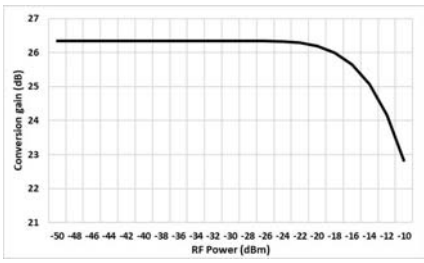

Fig. 5: Plot of gain versus power

V. Fast parallel testing methodology

Gain and phase estimation: Consider an array of N receiver chains where each chain is assumed to exhibit non-linearities up-to fifth order (α_1, α_2, α_3, α_4, α_5). When an input signal with frequency f is applied to a nonlinear system, the output frequencies of higher order are produced. Further, the non-linearity co-efficient of kth order (k=1,2,3,4,5) generates a frequency which is the kth multiple of the input frequency. When a sinusoidal input of frequency f is upconverted to RF by Equation 5 and applied to an RF system with 5th order distortions, frequency tones produced by different distortion co-efficients are shown in Table II.

TABLE II: Frequency tones affected due to non-linearity coefficients

Non-linearity co-efficient	Frequency tones affected
α_1	f
α_2	2f
α_3	f, 3f
α_4	2f, 4f
α_5	f, 3f, 5f

TABLE III: Frequency tones on first 4 RF chains for parallel gain/phase testing

RF chain number	Frequency tone
1	f
2	6f
3	7f
4	8f

From Table II, it can be seen that given an input with frequency f to a non-linear system with nth order distortion, the output contains tones which are upto *n* integer multiples of *f*. To test N fifth order non-linear systems simultaneously, the input frequencies to the N chains along with the output frequencies generated by each chain must not overlap in the spectrum of the combined output. It follows that the input frequency tone applied to the second chain must not overlap with f_1, $2f_1$, $3f_1$, $4f_1$ and $5f_1$ where f_1 is the frequency of the baseband input applied to the first chain. Following a similar methodology, for a system of 4 RF chains, each exhibiting upto fifth order nonlinearities, input tones to the RF chains can be allocated as shown in Table III. We have implemented an algorithm for computing the tones above. This is not discussed here for brevity.

When input sinusoids with baseband frequencies, as per the above method, are applied to a system with *N* RF chains exhibiting upto fifth order distortion, the received signal contains harmonic distortions due to all the chains. To obtain accurate estimates of the amplitudes and phases of the input baseband signals from the combined demodulated output, a function giving the sum of *N* sinusoidal functions with unknown amplitudes and phases as shown in Equation 9 is fit to the received signal using the Matlab curve fitting toolbox. In Equation 9, A_i, ψ_i represent the gain and phase of the ith RF chain while f_i represents the frequency tone applied to that particular chain.

$$y = \sum_{i=1}^{N} A_i \sin(2\pi f_i t + \psi_i) \tag{9}$$

For this optimization, the initial values are obtained by taking the FFT of the received signal as shown in Figure 6 for the particular case of 4 RF chains. This shows that the distortions

TABLE IV: Reference, FFT and Optimized Amplitude and Phase for system of 4 RF chains

Chain	Ref A	Ref ϕ	FFT A	FFT ϕ	Opt A	Opt ϕ
1	1.50	5.60	1.50	5.60	1.5033	5.60
2	1.51	25.68	1.53	25.68	1.51	25.68
3	1.52	48.69	1.52	48.70	1.52	48.69
4	1.49	77.72	1.47	77.71	1.48	77.72

produced due to nonlinearity in one chain do not affect the measurement of amplitudes and phases of other chains. The gain and phase difference of each chain is individually simulated and used as reference for measuring the accuracy of this estimation. The corresponding result showing the comparison of reference values of amplitude and phase with those obtained from the FFT and those obtained from optimization after the FFT is presented in Table IV. It can be seen that the amplitude and phase of each chain is estimated with high accuracy. Besides, this method is an improvement in terms of frequencies used for parallel testing over the method of using co-prime frequencies for the same demonstrated in [11]. Note that for a 4x4 system, this method requires test tones over a bandwidth of *31f* for estimating the individual gains and phases of 16 RF chains while the method proposed in [11] requires a bandwidth of *53f* in comparison. Hence, the proposed approach is 1.7X more frequency-efficient than the scheme proposed in [11] for a 4x4 RF system.

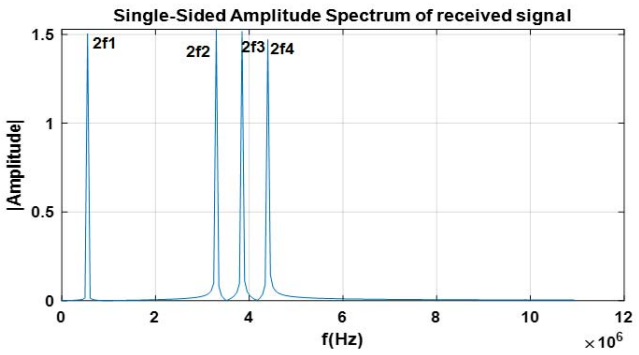

Fig. 6: FFT of the received signal

Estimation of upto 5th order distortion: Using the method described above, the phase and gain (α_1) is estimated for all RF chains in parallel. For estimation of higher order non-linearities of the transformation given by Equation 10 where x is the input signal, two tones are required per chain. In the presence of two tones, the output response of the system contains intermodulation tones. When the sum of two sinusoidal signals with frequencies f_1 and f_2 as shown in Equation 11, is applied to a non-linear system, the output contains various linear combinations of the input frequencies as shown in Table V. Table V shows the tones produced due to each of the non-zero co-efficients α_1, α_2, α_3, α_4 and α_5. When the tones generated by all the chains do not overlap, the non-linearities can be estimated by the measurement of amplitudes of intermodulation tones in the demodulated signal. However, when there are overlapping tones, the output amplitude and phase of the resulting tone is formed by superposition of the individual overlapping tones. From the FFT of the combined signal s_{out}(t) of Figure 2, the amplitude A and phase ϕ of two combined (overlapped in the frequency domain) tones from two different RF chains can be determined. Also, their individual phases ϕ_1 and ϕ_2 are known from prior gain and phase estimation steps discussed earlier. Consequently, using the two relations of Equation 12, the individual amplitudes A_1 and A_2 of the two tones can be calculated.

Since the amplitudes and the phases of the input fundamental tones are now determined, the non-linearities can be estimated from the combined output when there is an overlap of upto two tones.

$$y = \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \alpha_4 x^4 + \alpha_5 x^5 \tag{10}$$

$$x = \cos(2\pi f_1 t) + \cos(2\pi f_2 t) \tag{11}$$

$$\begin{aligned} A &= \sqrt{(A_1^2 + A_2^2 + 2A_1A_2\cos(\phi_1 - \phi_2))} \\ tan(\phi) &= \frac{A_1 \sin\phi_1 + A_2\cos\phi_2}{A_1\cos\phi_1 + A_2\sin\phi_2} \end{aligned} \tag{12}$$

Hence, by allowing an overlap of contributions from 2 RF chains at any given frequency, Table VI presents the frequency tone values used to calculate the amplitudes of tones generated by 5th order distortion in each RF chain. From Table VI, it can be seen that the frequencies in A and B columns never overlap. These are the input fundamental tones with known amplitude and phase values. The frequencies in columns B and D overlap. The frequencies in columns C and F overlap and those in H and J overlap.The last row in Table VI gives the values of frequency for k^{th} RF chain in a total of *N* RF chains. Hence, this method can be extended to any number of RF chains and is limited only by the total available frequencies for testing. Since the maximum distortion tone overlap that can happen at any given frequency is two, the individual amplitudes can be computed using Equation 12. Once the amplitudes of all tones corresponding to *N* RF chains are computed, the IIP3, IIP5 and P1dB are calculated as shown in Equations 13, 14 and 15 respectively. From Equations 13 and 14, the third and fifth order intermodulation co-efficients of every RF chain are estimated.

$$IIP3_{channel} = A_{in}\sqrt{\frac{V_{f1}}{V_{2f1-f2}}} = \sqrt{\frac{4\alpha_1}{3\alpha_3}} \tag{13}$$

$$IIP5_{channel} = A_{in}\sqrt{\frac{V_{f1}}{V_{3f2-2f1}}} = \sqrt[4]{\frac{8\alpha_1}{5\alpha_5}} \tag{14}$$

TABLE V: Frequency tones due to nonlinearities in system

Nonlinearity	Frequency tones
α_1	f_1, f_2
α_2	$2f_1, 2f_2, f_1 + f_2$
α_3	$2f_1 - f_2, 2f_2 - f_1$
α_4	$3f_1 - f_2$
α_5	$4f_1 - f_2, 3f_1 - 2f_2, 3f_2 - 2f_1$

TABLE VI: Frequency assignment for non-linearity estimation

	Primary		Secondary			Third order		Fourth order	Fifth order		
	A	B	C	D	E	F	G	H	I	J	K
Channel	f_1	f_2	$2f_1$	$2f_2$	f_1+f_2	$2f_1-f_2$	$2f_2-f_1$	$3f_1-f_2$	$4f_1-f_2$	$3f_1-2f_2$	$3f_2-2f_1$
1	f	6f	2f	12f	7f	-4f	11f	-3f	-2f	-9f	16f
2	7f	12f	14f	24f	19f	2f	17f	9f	16f	-3f	22f
3	13f	18f	26f	36f	31f	8f	23f	21f	34f	3f	28f
k	6k-5	6k	12k-10	12k	12k-5	6k-10	6k+5	12k-15	18k-20	6k-15	6k+10

$$P1dB_{channel} = \sqrt{0.145\frac{\alpha_1}{\alpha_3}} \tag{15}$$

Using the FFT of s_{out}(t) for estimation of amplitude and phase poses a limit on the frequency separation between between adjacent tones for a given ADC sampling rate (Figure 2). Hence, the number of channels that can be tested simultaneously depends on the total test frequency bandwidth required for testing and the number of orthogonal tones that can be supported by the baseband processor for the appropriate communication protocol employed (OFDM in our case). In a 20MHz OFDM band, the frequency separation required to distinguish two adjacent channels is 0.3125MHz. Using the same separation, bandwidth requirements for a testing scheme with no overlap of frequency tones, scheme proposed with frequency overlap for a 3rd order system [11] and the current testing method proposed with frequency overlap for a 5th order system is presented in Figure 7.

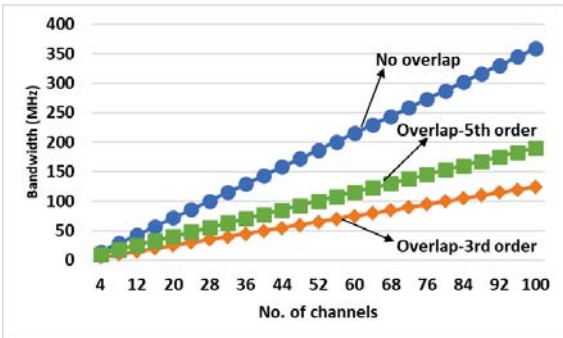

Fig. 7: Bandwidth requirements

VI. Implicit reward metric: Correlation to EVM and SINR

To speed up the RF chain array tuning process, we propose the use of an *implicit reward metric* for tuning as opposed to explicit computation of EVM and SINR in each tuning iteration. In the following, we show that the proposed implicit reward metric accurately tracks EVM and SINR and can be used as a proxy for the same. We first discuss how EVM and SINR are calculated and then discuss the proposed reward metric and show correlation between the two to justify use of the reward metric for tuning purposes.

Determination of EVM and SINR: The distortions of all RF chains are estimated using the prior discussed test procedures. These measurements are mapped to EVM and SINR parameters using analytical models. When a modulated symbol given by Equation 16 is phase shifted, it is transformed to S_p given by Equation 17. Due to the nonlinear distortions, the symbol is transformed as S_n^i given by Equation 18. The received symbols from all RF chains are combined as shown in Equation 19. These symbols are demodulated and filtered to obtain the received signal as shown in Equation 20.

$$\begin{aligned} &S_{mod} = I\sin(2\pi ft) + Q\cos(2\pi ft) = A\sin(2\pi ft+\psi) \\ &\text{where} \quad A = \sqrt{(I^2+Q^2)}, \quad \tan(\psi) = Q/I \end{aligned} \tag{16}$$

$$S_p^i = A\sin(2\pi ft + \psi + \phi_i) \tag{17}$$

$$S_n^i = \alpha_1^i S_p + \alpha_3^i S_p^3 + \alpha_5^i S_p^5 \tag{18}$$

$$S_{com} = \sum S_n^i \tag{19}$$

$$\begin{aligned} &I_{received} = S_{com}(\beta_1^I LO_I), Q_{received} = S_{com}(\beta_1^Q LO_Q) \\ &\quad LO_I = \sin(2\pi ft) \quad and \quad LO_Q = \cos(2\pi ft), \\ &S_{received} = I_{received} + jQ_{received} \end{aligned} \tag{20}$$

When N symbols are used to determine EVM, the EVM is obtained using the Equation 21.

$$EVM = 100 * \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(S^i - S^i_{received})^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(S^i)^2}} \tag{21}$$

The energy of received symbols given by Equation 20 is calculated as the square of the amplitude. The square of the amplitude of the received signal is taken as the received signal energy. The energy of noise A_{noise} at the receiver end is obtained by subtraction of the received symbol energy calculated from Equation 20 from the received signal energy. The received SINR is calculated using Equation 22.

$$SINR = \frac{A^2_{sym}}{A^2_{noise}} \tag{22}$$

$$y(t) = \sum_{i=1}^{N} \alpha_{1i} x(t), \text{ N is Number of RF chains} \tag{23}$$

Formulation of the implicit reward metric and correlation with EVM and SINR: The challenge in tuning a massive MIMO systems lies in estimating the EVM of the complete receiver array for every possible tuning combination. Hence, the first step is to correlate EVM with a test evaluation metric that is easier to calculate than EVM and is strongly correlated with the same. EVM results from distortions due to non-linearities and noise. In the absence of distortions, the system can be represented

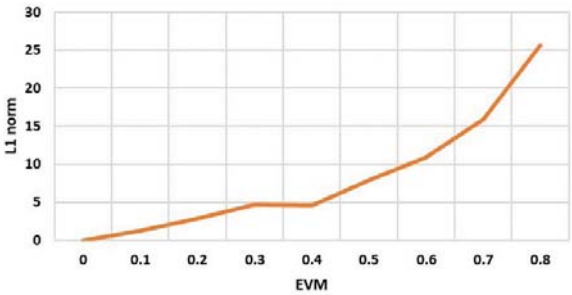

Fig. 8: Plot of L1 norm between received signal and golden response versus EVM

as shown in Equation 23 which is the response to an ideal distortion-less system and is called the golden response of the system. Here, α_{1i} represents the gain of ith RF chain. Hence, if the L1 norm of the difference (defined in Equation 25) between the golden response of the system and the received baseband signal obtained after demodulation is calculated, the EVM of the system should correlate with this parameter. The same is verified using simulations and presented in Figure 8. A positive correlation of around 95% is observed between EVM and the L1 norm indicating that as the system performance degrades from the ideal system corresponding to the golden response, the EVM also degrades (i.e. increases). A monotonic relationship between the L1 norm and EVM is observed in Figure 8 further cementing the reliability of the L1 norm as a proxy for EVM during optimization. The SINR of the system is also similarly correlated with the L1 norm and is found to exhibit a negative correlation of around 93.4%. Hence, the L1 norm of the difference between golden response of the system and the system output can be considered as a reliable metric for tuning the system. Lower values of the L1 norm correspond to higher rewards during tuning optimization. Similarly, higher values correspond to higher regret (lower reward).

VII. RL ASSISTED TUNING METHODOLOGY

The gain, phase and distortions of RF chains can be tuned by modulating various circuit tuning knobs. The available tuning knobs per RF chain are the voltages Vb1, Vb2, Vb, Vbsw and the currents DI and DQ presented in Section IV. For the 4 by 4 RF array discussed in this research, the total number of tuning knobs is 96. This represents a large tuning space as distortions in different chains interact with each other for minimum EVM. Explicitly modeling the relationship between EVM and the tuning knobs is difficult owing to the dimensions of the sample space in which the optimized tuning values are to be found. For this purpose, we use a reinforcement learning algorithm which learns the underlying correlations between the receiver output and the distortions introduced due to noise and non-linearities in the system.

The first step in demonstrating the tuning process is to create a set of process varied devices by varying process parameters such as threshold voltage(V_{th}), width(W), oxide thickness(t_{ox}) of transistors in the circuits described in Section IV. The corresponding RF chains are simulated and the L1 norm values determined for default (reference) values of all the tuning knobs of the complete system. Based on the L1 norm values obtained, the given devices are classified into M clusters using nearest-neighbor classification. This is a supervised algorithm for classification as presented in [16].

The benefit of clustering is that the tuning process for devices mapped to a cluster can start with known solutions for other devices in the cluster, thereby minimizing the number of tuning iterations. Consequently, for an untuned new device, the first step is to map the device to one of the clusters already identified by the clustering procedure. We next describe the tuning process.

We propose to use a reinforcement learning multi arm bandit algorithm for solving the tuning problem. The problem to be solved can be defined as shown in Equation 24 where A_n is the set of tuning knob values chosen for the nth RF chain, Q(a) is the estimated regret R_n at that step. Here, the regret R_n is defined as the L1 norm of the integral of the difference between the received system response($r(t)$) and the golden response of the system($g(t)$) over a period of time T, where T is taken as the time duration of the response signal r(t). This is presented in Equation 25. Choosing the set A_n which gives the minimum Q(a) at each step repeatedly works well only after there is enough understanding of the search space of tuning possibilities. To do this, a multi arm bandit algorithm is used.

$$Q(a) = \mathbb{E}[R_n | A_n = a] \tag{24}$$

$$R_n = \left\| \int_0^T [r(t) - g(t)] dt \right\| \tag{25}$$

The multi arm bandit algorithm involves selection of tuning knob values (called *arm*) that gives the maximum reward or minimum regret for the whole system as explained in [17]. The algorithm has an agent which chooses an action corresponding to a selection of the set of tuning knob values for each RF chain and calculates the L1 norm as described earlier. In describing the algorithm below, we associate the L1 norm with regret rather than reward because larger values of the L1 norm result in higher regret values. Actions with lower regrets are selected with higher probability in future tuning iterations. The ϵ-greedy algorithm is used with $\epsilon = 0.1$, meaning that random tuning knob values are selected in each tuning iteration with a probability of 0.1. The agent is trained to choose the values for all RF chains and update the expected regret accordingly. After a sufficient number of iterations, the algorithm converges to the best possible tuning settings. The steps involved can be explained as shown below:

Lines 1 to 12 in the algorithm define all the arms as the tuning settings. Lines 13 to 21 in the algorithm allow the agent to choose one of the defined arms and the function returns the chosen arm settings. Lines 22 to 24 in the algorithm define the golden response of every cluster. Lines 25 to 28 in the algorithm calculate the regret value after choosing a particular arm. Finally, the lines 30 to 39 update the received signal and

Input: RF chain to be tuned
Output: Tuned settings for RF chain

```
Function tune_chain(rxchains):
    n_clusters=M;
    n_chains=N;
    n_arms=K;
    niterations=Num;
    ϵ=0.1;
    for i ← 0 to M − 1 1 do
        for j ← 0 to K − 1 1) do
            arm[i][j]=i^th tunesettingsof j^th cluster
        end
    end
return
Function Choosearm(clusterj):
    r=randomvalue;
    if r≤ ϵ then
        p=random(K-1);
        chosenarm=arm[p][clusterj]
    else
        chosenarm=self;
    end
return chosenarm
for i ← 0 to M − 1 1 do
    refarmsig[i]=Golden_response
end
Function Calreg(rxchainsig,refarmsig):
    temp=rxchainsig-refarmsig;
    temp=Norm(temp);
return temp
recsignal=0;
Function run():
    for j ← 0 to niterations 1 do
        for i ← 0 to N − 1 1 do
            rx[i]=Choosearm(clusterj);
            chosensettings[i]=rx[i];
            recsignal=recsignal+rx[i];
        end
        regretvalue=Calreg(recsignal,refarmsig)
    end
return chosensettings, regretvalue
```

Algorithm 1: Multi arm bandit algorithm for tuning

call the previously defined functions to calculate and update the regret value.

The value of ϵ is to be chosen so that the expected regret is minimized in the least possible amount of time. For this purpose, we varied it's value as 0.1, 0.2 and 0.3 and recorded the results after every 10 iterations for around 500 iterations. It is seen that the minimum regret is obtained by using ϵ value of 0.1 as shown in Figure 9.

Using this algorithm, the optimized set of tuning values for each RF chain are obtained. In our case, we implemented the algorithm for 16 RF chains. With R_n (Equation 25) as the parameter for regret of the algorithm, the time for tuning is around 2s (on the average, 100 iterations of multi-arm bandit

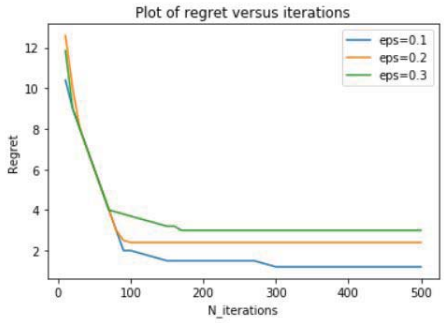

Fig. 9: Performance of algorithm for different ϵ values

algorithm) and with EVM as the parameter for regret, the time for tuning is around 63s. This is mainly because the tuning approach requires transmission of many symbols and estimation of EVM in every iteration of the tuning process as explained in Section VI. The same algorithm can be extended to massive MIMO systems with more than 16 RF chains and its complexity increases as a multiple of number of RF chains.

VIII. Fast tuning based on tuning solutions for neighboring die

Once the tuning knob values for a given RF array system (die) are obtained, the same can be used to obtain the values of tuning knobs for RF chains on an adjacent die. This is due to the assumption that adjacent die on a wafer exhibit process parameter values that are partially correlated with those exhibited by the given die . Hence, the tuning knob values obtained for a previously tuned RF system on a given die can be used as an initial condition for tuning knob settings on an adjacent die to facilitate fast convergence. The same is verified as shown in Figure 10. In this, the systematic component of process variations is assumed to vary linearly from one die to the adjacent die across a 1% to 5% range and the random component of process variability is assumed to vary from 10% to 50% of the nominal process parameter values in a similar manner. The same tuning algorithm is used with an initial condition taken from the tuning knob solution for the neighboring die. As a result, it is observed that the *tuning algorithm converges to optimal solution within 6 iterations* (0.2sec total test time as compared to 2sec without initialization). This is a significant improvement as without such explicit initialization, almost 100 multi-arm bandit iterations are needed to achieve convergence on the average. When the die-to-die tracking hits disconnects, the same algorithm will take higher number of iterations for relevant die even with initialization due to the random fluctuations of process parameters.

IX. Experimental results

In this section, the effect of the tuning parameters considered for each device in RF chain on the performance is presented. This is followed by the results of tuning algorithm based on EVM and SINR specifications. The parameters Vb1 and Vb2 in Figure 3 can be tuned to vary the performance of the RF chain in terms of gain, IIP3, IIP5 as shown in Figures

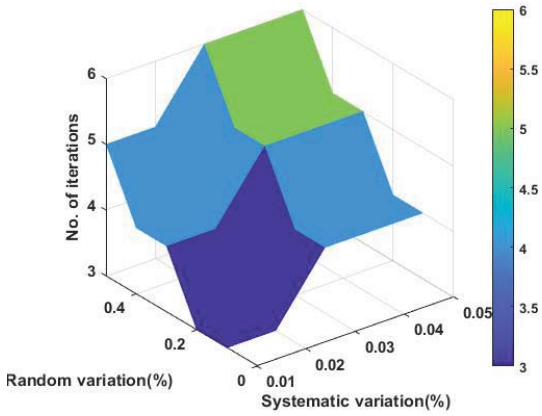

Fig. 10: Performance of algorithm with initial solution

11, 12 and 13 respectively. It can be seen from the graphs that choosing values of Vb1 and Vb2 that reduce IIP3 and IIP5 close to 0dB, reduce the gain of the RF chain.

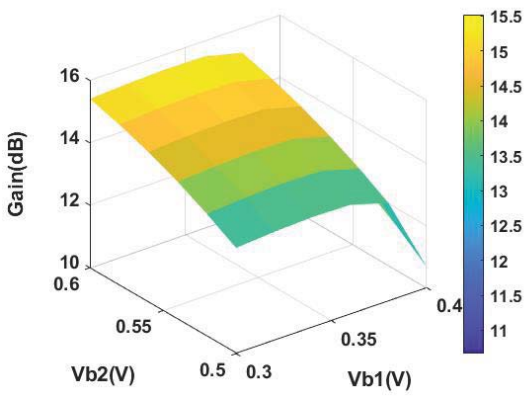

Fig. 11: Effect of Vb1 and Vb2 on gain

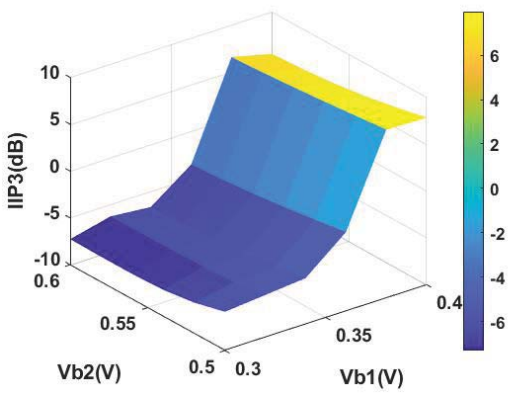

Fig. 12: Effect of Vb1 and Vb2 on IIP3

The parameters Vb and Vbsw in Figure 3 can be tuned to vary the performance of the RF chain in terms of gain as shown in Figure 14 respectively. The effect of the same range

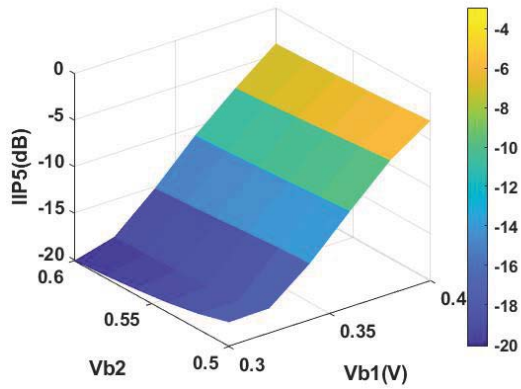

Fig. 13: Effect of Vb1 and Vb2 on IIP5

of Vb and Vbsw on IIP3 is shown in Figure 15 while the effect on IIP5 is negligible for the whole range ($\leq 2dB$). Further, it should be noted that when the devices are connected as a chain, the influence of following stages on whole system IIP3 and IIP5 increases with increased gain of preceeding stages thus creating a tradeoff between the gain and the distortions of the system.

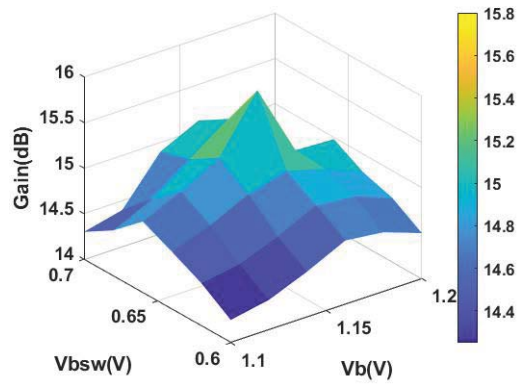

Fig. 14: Effect of Vb and Vbsw on gain

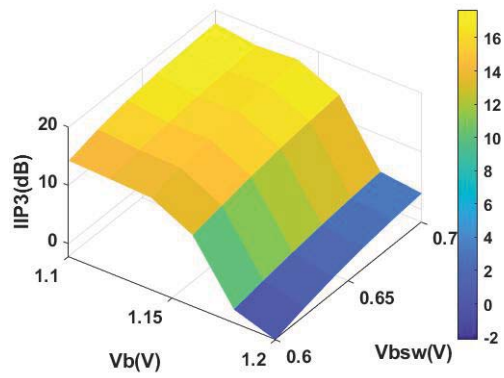

Fig. 15: Effect of Vb and Vbsw on IIP3

The performance of the tuning method explained in previous

section is presented by considering 500 MIMO receiver systems with 16 RF chains each as described in Section IV. The testing and tuning process is repeated for all beam-steering angles. The distribution of devices based on their EVM and SINR values before and after tuning is shown in Figure 16 and Figure 17 respectively. Acceptance boundaries for EVM and SINR are set as 5% and 8dB respectively. The yield is improved from 65% to 81% after tuning. The total number of tuning bits per RF chain is 32 (LNA:6, Mixer:6, Phase shifter:16, VGA:4 bits). For a system with a 4x4 phased array, the total number of tuning bits is 512 with 16 RF chains. The tuning algorithm presented here works on 512 bits simultaneously. The tuning process takes around 5 seconds using Jupyter (Python).

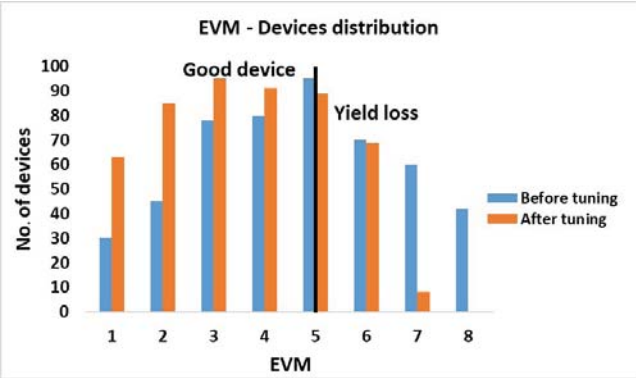

Fig. 16: Distribution of devices before and after tuning based on EVM

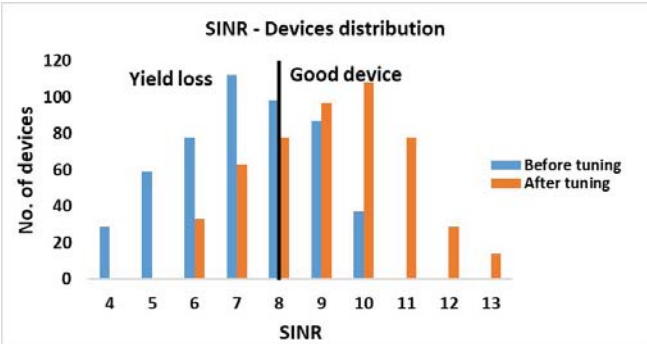

Fig. 17: Distribution of devices before and after tuning based on SINR

X. Conclusions

In this work, we have presented a fast parallel testing and tuning procedure for beamforming RF MIMO arrays which employ phased array RF chains for wireless communication. Frequency-efficient and parallel test methods for gain, phase and upto 5th order nonlinearity of RF chains are developed in this research. To speed up testing and reduce tuning costs, an implicit reward metric based reinforcement learning approach is used in this research. In relation to prior techniques, this allows a 4X4 receiver array to be tested in 0.2sec - 2sec depending on utilization of tuning information from neighboring die to speed up the tuning process. The method does not require explicit EVM or SINR tests and thereby avoids tester costs while reducing test time.

XI. Acknowledgements

This research was supported by NSF Award ECCS 1815653. We would like to acknowledge Sabyasachi Deyati of Intel Corp for helpful discussions and Kwondo Ma for his contributions.

References

[1] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5g wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[2] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan, "Mimo for millimeter-wave wireless communications: Beamforming, spatial multiplexing, or both?" *IEEE Communications Magazine*, vol. 52, no. 12, pp. 110–121, 2014.

[3] W. Saad, M. Bennis, and M. Chen, "A vision of 6g wireless systems: Applications, trends, technologies, and open research problems," *IEEE network*, 2019.

[4] D. Han, B. S. Kim, and A. Chatterjee, "Dsp-driven self-tuning of rf circuits for process-induced performance variability," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 18, no. 2, pp. 305–314, 2009.

[5] V. Natarajan, S. Sen, A. Banerjee, A. Chatterjee, G. Srinivasan, and F. Taenzler, "Analog signature-driven postmanufacture multidimensional tuning of rf systems," *IEEE Design & Test of Computers*, vol. 27, no. 6, pp. 6–17, 2010.

[6] D. Maliuk and Y. Makris, "An experimentation platform for on-chip integration of analog neural networks: A pathway to trusted and robust analog/rf ics," *IEEE transactions on neural networks and learning systems*, vol. 26, no. 8, pp. 1721–1734, 2014.

[7] J. W. Jeong, J. Kitchen, and S. Ozev, "A self-compensating built-in self-test solution for rf phased array mismatch," in *2015 IEEE International Test Conference (ITC)*. IEEE, 2015, pp. 1–9.

[8] K. Greene, V. Chauhan, and B. Floyd, "Built-in test of phased arrays using code-modulated interferometry," *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 5, pp. 2463–2479, 2018.

[9] E. Acar and S. Ozev, "Low cost mimo testing for rf integrated circuits," *IEEE transactions on very large scale integration (VLSI) systems*, vol. 18, no. 9, pp. 1348–1356, 2009.

[10] M. Shafiee and S. Ozev, "Contact-less near-field measurement of rf phased array antenna mismatches," in *2017 22nd IEEE European Test Symposium (ETS)*. IEEE, 2017, pp. 1–6.

[11] S. Deyati, B. J. Muldrey, B. Jung, and A. Chatterjee, "Concurrent built in test and tuning of beamforming mimo systems using learning assisted performance optimization," in *2017 IEEE International Test Conference (ITC)*. IEEE, 2017, pp. 1–10.

[12] J. K. Kibarian and A. Strojwas, "Using spatial information to analyze correlations between test structure data (semiconductor ic manufacture)," *IEEE Transactions on Semiconductor Manufacturing*, vol. 4, no. 3, pp. 219–225, 1991.

[13] V. Rabinovich and N. Alexandrov, *Antenna Arrays and Automotive Applications*. Springer Publishing Company, Incorporated, 2012.

[14] S. C. Blaakmeer, E. A. M. Klumperink, D. M. W. Leenaerts, and B. Nauta, "Wideband balun-lna with simultaneous output balancing, noise-canceling and distortion-canceling," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 6, pp. 1341–1350, 2008.

[15] E. Ortigueira, I. Bastos, L. B. Oliveira, J. P. Oliveira, and J. Goes, "A simplified design of a mosfet-only wideband gilbert cell," in *Proceedings of the 18th International Conference Mixed Design of Integrated Circuits and Systems - MIXDES 2011*, 2011, pp. 225–230.

[16] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theor.*, vol. 13, no. 1, p. 21–27, Sep. 2006. [Online]. Available: https://doi.org/10.1109/TIT.1967.1053964

[17] A. Slivkins, "Introduction to multi-armed bandits," 2019.