

Self-Aware MIMO Beamforming Systems : Dynamic Adaptation to Channel Conditions and Manufacturing Variability

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

Abstract—Emerging wireless technologies employ MIMO beamforming antenna arrays to improve channel Signal-to-Noise Ratio (SNR). The increased dynamic range of channel SNR values that can be accommodated, creates power stress on Radio Frequency (RF) electronic circuitry. To alleviate this, we propose an approach in which the circuitry along with other transmission coding parameters can be dynamically tuned in response to channel SNR and beam-steering angle to either minimize power consumption or maximize throughput in the presence of manufacturing process variations while meeting a specified Bit Error Rate (BER) limit. The adaptation control policy is learned online and is facilitated by information obtained from testing of the RF circuitry before deployment.

Index Terms—MIMO, Adaptation, actor-critic

I. INTRODUCTION

5G and beyond wireless networks employ Multiple Input Multiple Output (MIMO) beamforming antenna arrays to provide data rates in multiple Gigabits per second [1]. The wide range of wireless channel conditions seen by these systems in the field pose arguments in favor of designing the underlying electronics to adapt dynamically to such channel conditions across diverse beamsteering angles. Besides, dynamic adaptation is essential to reduce power consumption while staying within specified system bit error rate limits for corresponding wireless channels encountered. Such design is challenging due to the complex architecture of systems involved [2] and the performance versus power trade-offs of individual RF and digital modules used in the design. Power is saved by trading off linearity of individual amplifier and mixer modules of the wireless system front-end under varying channel conditions.

Systems for self-learning wireless RF MIMO receivers were presented in [3], [4] and investigated simultaneous tuning of RF circuitry and wireless channel coding parameters. These did not consider beamsteering capability and relied on the use of neural network training for learning the relationships between tuning knobs (bias voltages and currents of receiver circuitry), process parameters, wireless signal quality and system performance. The learning is performed on-line and requires dedicated on-board computing resources. Recently, there has been work on the use of multi-arm bandit reinforcement learning for tuning RF power amplifier distortion parameters [5].

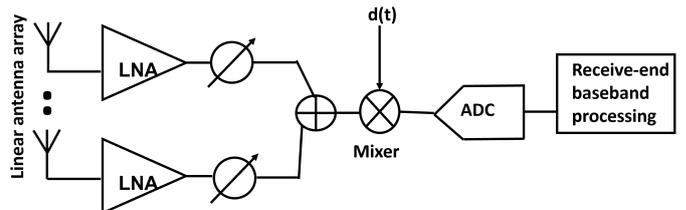


Fig. 1: Beamforming receiver system

II. PRELIMINARIES AND APPROACH

We consider a 5.9GHz linear (8x1) analog antenna array receiver system with individual low noise amplifiers (LNA) and phase shifters connected to each antenna element as shown in Figure 1. Further, we assume that there are on-board software-hardware resources to tune the receiver configuration by varying the bias voltages and currents in the receiver circuitry. The problem to be solved is defined as follows:

Problem definition: For adaptive beamsteering systems such as the one shown in Figure 1, the problem is one of determining a *control policy* for the wireless front-end circuitry that modulates its circuit-level tuning knobs (T) (bias voltages and currents), baseband word length (L), symbol modulation rate (M) and coding rate (C) in response to channel conditions (W) and desired beamsteering angle (B) in such a way as to: (a) minimize energy per bit or (b) maximize throughput per Watt of the wireless communications system under given communication bit error rate (BER) constraints.

In this work, error vector magnitude (EVM) is used as a proxy for BER as it can be estimated much faster than BER for a specified level of statistical confidence [3].

To determine the control policy, an actor-critic reinforcement learning approach is developed that allows the system to learn through exposure to real-time wireless channel conditions in field use. There are three aspects to the proposed approach in solving the above problem:

Step 1. Receiver circuitry characterization: The RF chains in the receiver are characterized by applying baseband tests to determine their gain and intermodulation specifications as a function of their configuration (T). The intermodulation effects corrupt the received signal, contributing to degraded EVM of the receiver. A measure of the receiver non-linearity (NM) as

a function of the tuning knobs T , $NM=f(T)$ is computed. A non-linear polynomial regressor is used to map to NM value. *Step 2. Wireless channel estimation:* The wireless channel quality determines the EVM of the received signal [6]. Pilot symbols associated with transmitted data frames are used to estimate channel quality using least squares estimation technique [6].

Step 3. Learning the control adaptation policy: The wireless channel estimated in Step 2 along with the weights of the regression function determined in Step 1 are passed as inputs to the learning algorithm. The learning algorithm is based on actor-critic reinforcement learning. In this, choice of the actor determines the probability that a particular action (defined as vector v with $[T, M, C, L]$ defined above) is selected for a given receiver state. The state of the receiver is defined by the current estimated wireless channel and beamsteering angle. The critic restricts the actions taken based on given EVM constraints. The critic evaluates the taken action based on EVM, power consumption and Energy/bit (or Throughput/Watt) values of the receiver after taking the action. It provides feedback to the actor so that the probability of an action given a receiver state, is updated. As the system experiences state-action combinations and learns over time, given any state, the optimal action for minimizing (or maximizing) Energy/bit (or Throughput/Watt) is determined.

III. RECEIVER CIRCUITRY CHARACTERIZATION

At 5.9GHz frequency of operation, a noise cancellation equipped low noise amplifier (LNA) along with the bias voltages ($Vb1$, $Vb2$) used for tuning are shown in Figure 2 is designed in 130 nm technology [7]. A fully balanced Gilbert cell type mixer implemented based on the work in [8] is also given in Figure 2 and is tuned by varying the highlighted Vb and Vb_{sw} voltages. The phase shifter circuit used is given in Figure 3. The DI and DQ currents determine the phaseshift value as $\tan^{-1}(DQ/DI)$ and are used for tuning it.

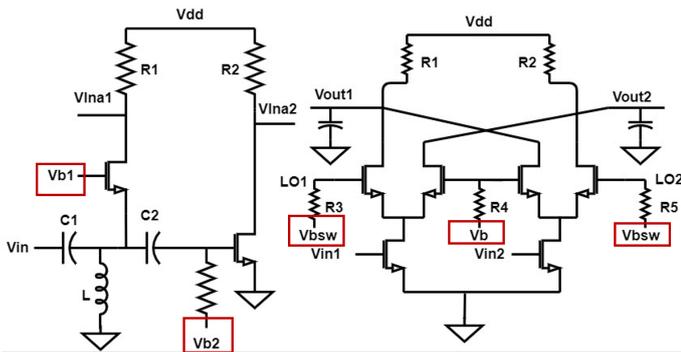


Fig. 2: Low Noise Amplifier and Mixer

For an ideal RF chain, the input $x(t)$ and output $y(t)$ of the system are related linearly as $y_{ideal}(t) = \alpha_1 x(t)$. A real RF chain exhibits higher order distortions as shown in Equation 1. The α_1 , α_3 and α_5 represent the gain, third order and fifth order distortions of a RF chain. The research presented in [9] shows that the tuning parameters (circuit biases mentioned

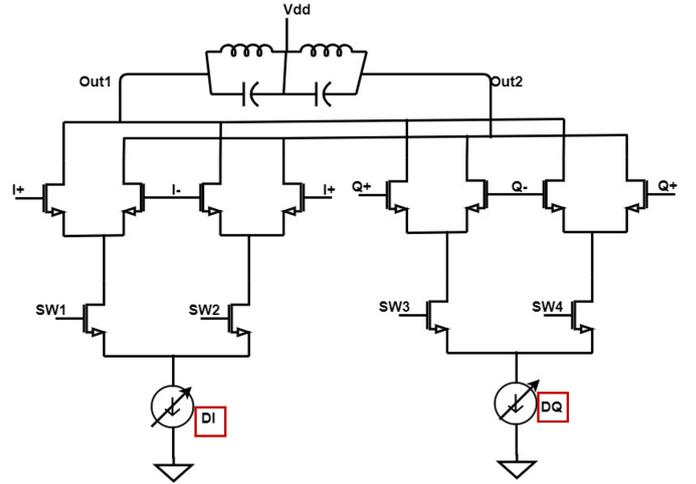


Fig. 3: Phase shifter

above) affect the gain and non-linearities. Table I presents the input sinusoid frequencies applied to each RF chain in an eight element receiver array to estimate α_1 , α_3 and α_5 . In table I, f is the minimum baseband frequency tone possible. FFT is applied to the combined receiver output and the $\alpha_1, \alpha_3, \alpha_5$ values are calculated using the amplitudes of the frequency tones in the resultant output. The input frequencies in table I are chosen such that, when frequencies given in a column are applied concurrently to the receiver system, the combined output has minimal distortion [10].

$$y_{real}(t) = \alpha_1 x(t) + \alpha_3 (x(t))^3 + \alpha_5 (x(t))^5 \quad (1)$$

TABLE I: Input frequencies applied to estimate α_1 , α_3 & α_5

RF chain	To estimate α_1	To estimate α_3 & α_5
1	f	$f, 6f$
2	$6f$	$7f, 12f$
3	$7f$	$13f, 18f$
4	$8f$	$19f, 24f$
5	$9f$	$25f, 30f$
6	$10f$	$31f, 36f$
7	$11f$	$37f, 42f$
8	$13f$	$43f, 48f$

The difference of $y_{real}(t)$ and $y_{ideal}(t)$ gives the distortion in the received output of RF chain. Note that this is dependent on the bias values (T) set for each RF chain and affects the combined receiver EVM. For a receiver system with N RF chains set at T configuration (as defined by bias values), the total non-linearity measure (NM) can be given by Equation 2 in which α_{ki} represents the k^{th} order distortion in RF chain i . The higher the value of NM , the higher is the distortion and higher is the EVM. By increasing the power consumption of the circuitry (by varying T), NM can be decreased. As a result, given any state, the learning policy must take actions that result in maximum decrease in NM while ensuring minimum increase in power consumption. Hence, a metric NP which is a function of tuning configuration (T) and length of digitized output (L) is defined in Equation 3.

$$NM(T) = \sum_{n=1}^{n=N} [\alpha_{3i} + \alpha_{5i}] \quad (2)$$

$$NP = \frac{NM(T)}{\text{Power consumption}(T, L)} \quad (3)$$

For a given device, the circuit tuning parameters (T) and word-size (L) are varied to estimate the corresponding NP values. [T,C,M,L] defined in Section II is taken as vector 'v'. For a given device, the vector 'v' is randomly sampled and the corresponding value NP is noted to form a dataset. In our receiver design there are 24 bits (3 for the LNA and mixer each, 16 bits for the phase shifter and 2 bits for length of digitized output). In a MIMO system with 8 RF chains, this evaluates to 192 bits of tuning. Hence, generating and storing the NP value for all possible combinations of 'v' is expensive. A multi-dimensional quadratic polynomial regression as shown in Equation 4 is applied on the generated dataset to estimate the corresponding value of NP as M_{reg} . In this equation, β_i ($\forall i=0$ to 2) are regression vector co-efficients, v^2 is a vector obtained by squaring each value in the vector v and ' \cdot ' represents the dot product between the vectors. The final β_i values after regression are passed as inputs to the learner. $M_{reg}(v) = \beta_0 + (\beta_1 \cdot v) + (\beta_2 \cdot (v^2))$ (4)

IV. WIRELESS CHANNEL ESTIMATION

In addition to the distortion within the receiver circuitry, the wireless channel exhibits additive noise and fading characteristics that degrade the received signal. As such, there is a requirement to estimate the wireless channel. To estimate the wireless channel quality, 'pilot symbols' are transmitted along with data symbols at periodic intervals. If the transmitted pilot symbols are represented using a diagonal matrix X and the received symbols by vector Y , the relation between X and Y is given as $Y = XH + Z$ in which Z represents the noise vector and H gives the channel matrix. Least squares(LS) channel estimation states that the estimated channel matrix \hat{H}_{LS} is one that minimizes the cost function $J(\hat{H})$ given in Equation 5 [6] and the corresponding \hat{H}_{LS} is also given by the same Equation. This channel matrix along with estimated SNR [11] are passed to the learning algorithm as inputs.

$$J(\hat{H}) = \|Y - X\hat{H}\|^2 \text{ and } \hat{H}_{LS} = X^{-1}Y \quad (5)$$

V. LEARNING THE CONTROL ADAPTATION POLICY

The adaptation control policy is learned using the actor critic approach [12]. Given any state of the wireless channel, actions are chosen according to a *policy* by an *actor*. The *critic* evaluates the system after the action is applied and updates the actor's policy such that the actions with maximum *reward*, represented by 'r', are reinforced. In this research, a state of the system is defined by the estimated wireless channel quality (W) and beamsteering angle (B). The actions are vectors of [T,M,C,L] which are defined in Section II. The actor policy $\pi_t(s, a)$ is given in Equation 6 and denotes the policy of actor at time step t when the environment is in state s. The probability that the actor chooses action a in state s and depends on M_{reg} obtained from Equation 4 as well as real-time measurements of power and throughput. In this, ζ_1, ζ_2 are weighing factors chosen so that the value of $\pi_t(s, a)$ is positive and less than 1. It is updated by the critic using the function

σ given as $\sigma_t(s, a) = \sigma_t(s, a) + \rho\delta_t$, where ρ is a positive parameter and δ_t is given by $\delta_t = r_{t+1} + \gamma V(s_{t+1}) - V(s_t)$. The critic first estimates the EVM and if it is not within the chosen limits, it informs the actor to set the probability of the chosen action to 0. Otherwise, it calculates the temporal difference error based on the environment before and after applying the action using the function σ defined above. Here, γ is a positive value and $V(s_t)$ is the value function in state s at time step t. The value function is defined based on objective functions (i.e., maximize Throughput/Watt or minimize Energy/Bit) as given in Equation 7.

Note that initially the learning starts with low non-linearity and higher power consumption. At each learning step, the algorithm explores lower power but higher non-linearity system configurations that meet stipulated EVM constraints.

$$\pi_t(s, a) = \frac{e^{-\zeta_1 M_{reg}(a) + \zeta_2 \sigma_t(s, a)}}{\sum_b e^{-M_{reg}(b)}} \quad (6)$$

$$V(s) = \left| \frac{\text{Throughput}}{\text{Power Consumption}} \right| \text{ or } \left| \frac{\text{Power consumption}}{\text{Throughput}} \right| \quad (7)$$

VI. EXPERIMENTAL RESULTS

The receiver circuitry was implemented using Advanced Design System (ADS) and the learning algorithm was implemented in Python using Jupyter Notebook. The tuning parameters for the receiver system under consideration are the voltage and current biases as specified in Section III, modulation rate and code rate of the wireless system and the FFT wordsize. The modulation rate of the transmitter was selected to be 16QAM, 64QAM and 128QAM while the code rate was varied between 0.5 and 0.75. The research in [13] demonstrated that by decreasing the wordsize, the power consumption in filtering and FFT operations is reduced. We correlate two wordsizes 10 and 12 with corresponding power consumptions in FFT operations, 12.1mW and 17.4mW, respectively. Similarly, the two wordsizes correspond to 17.4mW and 11.8mW power consumption in the FIR filter. Excluding the baseband power, the power consumption of nominal receiver circuitry described in Section IV is 60mW without any optimization. Thus a receiver system without any adaptation consumes around 89mW power.

The power consumption of a receiver system for the two proposed objective function optimizations (Throughput/Watt and Energy/Bit) under varying channel conditions is presented in Figure 4. It can be seen that initially, the power savings are from the receiver tuning while the wordlength is maintained at 12. Around 3.4dB (for energy/bit optimization) and 4.2dB (for throughput/Watt optimization), the word length changes from 12 to 10 resulting in power savings of around 30mW. Similarly, the resultant optimal throughput for both optimizations is plotted in Figure 5. It can be seen that the throughput improves as channel SNR improves due to change in the modulation rate. It can be observed from Figure 5 that higher throughput can be achieved with optimization even for channels with lower SNR than in the case without any optimization (2Mbps for all channel SNR values). Note that

the change in modulation rate can be observed for varying levels of throughput.

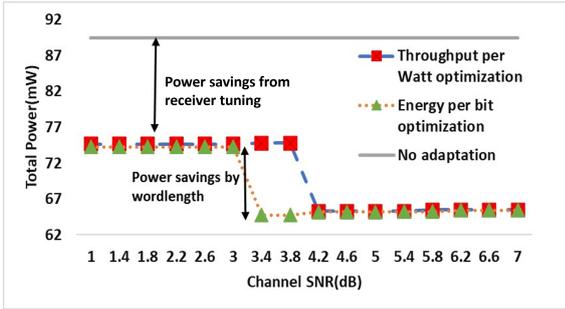


Fig. 4: Optimized total power consumption

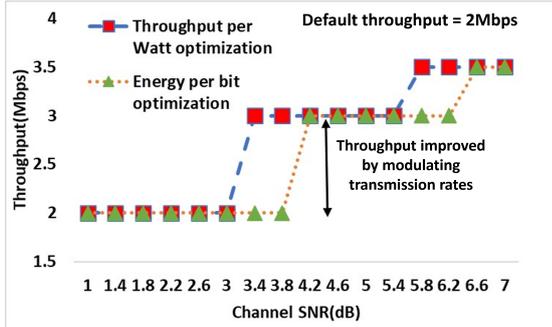


Fig. 5: Optimized throughput

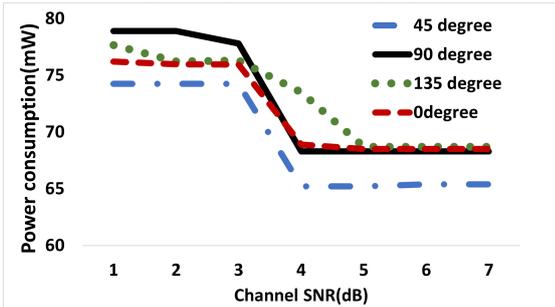


Fig. 6: Optimal power consumption across beam angles

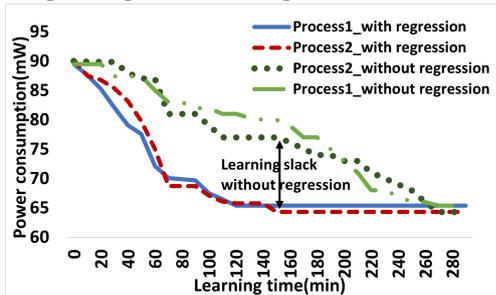


Fig. 7: Learning performance with and without prior regression for process varied devices

The optimized total power consumption of a receiver system under Energy per bit optimization is shown for four beam-steering angles (45° , 90° , 135° , 0°) in Figure 6. It can be seen that the optimized power profiles vary with varying beam steering angle. This is due to the difference in gain and power consumption patterns of phase shifters across different beam angles.

In Figure 7, the power consumption as a function of the learning time of the algorithm to optimize Energy/Bit is shown for two process varied devices. With the help of regression function guided learning, the power consumption reaches its optimal value (72% of the original) in both cases. Without the regression function, both the systems take more time to reach their optimal points as seen in Figure 7. This demonstrates the efficiency of the proposed adaptation across process varied devices as well as the importance of using regression based actor-critic.

VII. CONCLUSIONS

In this work, a reinforcement learning based approach for power-optimal cross-layer adaptation of beamforming RF MIMO arrays for dynamically changing channel conditions is developed. Results indicate that 1.3X-2X improvements are observed in selected power and energy metrics under diverse channel conditions.

VIII. ACKNOWLEDGEMENTS

This research was supported by NSF Award ECCS 1815653.

REFERENCES

- [1] W. Saad, M. Bennis, and M. Chen, "A vision of 6g wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.
- [2] T. C. W. Schenk and E. R. Fledderus, "Rf impairments in high-rate wireless systems - understanding the impact of tx/rx-asymmetry," in *2008 3rd International Symposium on Communications, Control and Signal Processing*, 2008, pp. 117–122.
- [3] D. Banerjee, S. K. Devarakond, X. Wang, S. Sen, and A. Chatterjee, "Real-time use-aware adaptive rf transceiver systems for energy efficiency under ber constraints," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, 2015.
- [4] D. Banerjee, B. Muldrey, X. Wang, S. Sen, and A. Chatterjee, "Self-learning rf receiver systems: Process aware real-time adaptation to channel conditions for low power operation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, 2016.
- [5] F. Wang, S. Xu, J. Romberg, and H. Wang, "An artificial-intelligence (ai) assisted mm-wave doherty power amplifier with rapid mixed-mode in-field performance optimization," 08 2019, pp. 1–3.
- [6] J.-J. van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, "On channel estimation in ofdm systems," in *1995 IEEE 45th Vehicular Technology Conference. Countdown to the Wireless Twenty-First Century*, vol. 2, 1995, pp. 815–819 vol.2.
- [7] 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.
- [8] 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.
- [9] 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 *IEEE International Test Conference*, 2017.
- [10] S. Komaraju and A. Chatterjee, "Fast evm tuning of mimo wireless systems using collaborative parallel testing and implicit reward driven learning," in *IEEE International Test Conference*, 2020.
- [11] A. Das and B. D. Rao, "Snr and noise variance estimation for mimo systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 3929–3941, 2012.
- [12] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: A Bradford Book, 2018.
- [13] M. M. Nisar and A. Chatterjee, "Test enabled process tuning for adaptive baseband ofdm processor," in *26th IEEE VLSI Test Symposium*, 2008.