

Received July 27, 2020, accepted August 9, 2020, date of publication August 20, 2020, date of current version August 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3018393

Achieving Phase Coherency and Gain Stability in Active Antenna Arrays for Sub-6 GHz FDD and TDD FD-MIMO: Challenges and Solutions

REZA MONIR VAGHEFI¹, (Member, IEEE), RAMESH CHEMBIL PALAT¹, GIOVANNI MARZIN¹, KIRAN BASAVARAJU², YIPING FENG¹, AND MIHAI BANU¹, (Life Fellow, IEEE)

¹Blue Danube Systems Inc., Santa Clara, CA 95054, USA

²Motorola Mobility, Chicago, IL 60654, USA

Corresponding author: Reza Monir Vaghefi (vaghefi@vt.edu)

ABSTRACT Full-dimension MIMO (FD-MIMO) using planar active antenna systems (AAS) is considered a critical technology for fifth-generation (5G) cellular systems to improve network capacity. An AAS is typically subject to hardware impairments that negatively impact network capacity. Hence, this article focuses on impairments that cause phase and magnitude errors between radio frequency (RF) chains and shows why they are particularly difficult to avoid in practical AAS. Although previous investigations show these impairments to degrade performance, they are not useful in deriving measurable impairment margins for practical FD-MIMO deployments. Knowing impairment limits are critical for system designers to make hardware design tradeoffs such as AAS configuration, component selection, implementation complexity, and cost. Moreover, it also helps set conformance limits for critical lab verification. Therefore, the paper first investigates the impact of the impairments and derives their practical limits for FD-MIMO by explicitly considering the cumulative effects of the channel model, inter-cell interference, link adaptation, and channel aging due to feedback delays. It is shown that a lower number of digitized RF chains can be a better choice under lower impairments. Next, the sources of impairments are investigated by using measurements carried out in the lab and the field during live operation in a commercial LTE network. Phase drift from local oscillators (LO) and internal temperature variations are identified as two significant sources. The tradeoffs and shortcomings of some of the existing solutions in massive MIMO literature are discussed. Finally, in order to address the shortcomings, a novel and practical coherent LO distribution architecture and array calibration mechanism are proposed. This solution is shown to be applicable to both TDD and FDD FD-MIMO. Measurement results are provided to prove the high degree of coherency and stability achieved on a unique array architecture called high definition active antenna system (HDAAS).

INDEX TERMS Massive MIMO, FD-MIMO, calibration, phase and magnitude error, array coherency.

I. INTRODUCTION

Massive multiple-input and multiple-output (MIMO) using large active antenna arrays are viewed as a critical solution to solve capacity demands in Fifth Generation New Radio (5G-NR) cellular networks operating in sub-6 GHz bands. Full-dimension MIMO (FD-MIMO) defined in the 3rd Generation Partnership Project (3GPP) [1] has adopted principles of massive MIMO. It specifies the use of planar rectangular active antenna arrays at the base station (BS) to support simultaneous data transmission to multiple user equipment (UE)

The associate editor coordinating the review of this manuscript and approving it for publication was Wenjie Feng.

terminals (referred to as multi-user MIMO or MU-MIMO) to achieve significant improvement in network capacity. This topic is of active research in academia and industry [2].

A. CHALLENGES IN FIELD OPERATION DUE TO HARDWARE IMPAIRMENTS AND RELATED WORK

Antenna installations on tower tops and buildings for a macro-cell network are typically limited in space and weight due to wind loading constraints. Active antenna systems (AAS) proposed for FD-MIMO are also required to operate under similar constraints. Therefore, from a practical AAS design point of view [3], some of the key challenges

include: routing a large number of high-speed digital I/O and corporate feeds; high-speed digital processing required to support digital front end (DFE) and lower-level physical layer processing [4]; power consumed by the analog front end (AFE), data converters, digital processors and high speed I/O; and thermal design to control heat dissipation. Operating such a complex system under space and thermal constraints subjects it to non-ideal behavior from radio frequency (RF) components that degrades performance. Additional circuitry and computing resources required to compensate RF impairments only further exacerbate the above challenges. Hence, it is critical to find efficient AAS solutions that are robust against RF impairments under the above operating challenges and conditions.

The effect of RF impairments is broadly classified into residual additive noise, multiplicative noise, and thermal noise [5], [6]. Solutions to address residual additive noise such as crest factor reduction, digital pre-distortion, and IQ imbalance correction are readily available [7]. On the other hand, multiplicative noise resulting from non-ideal phase and magnitude response across RF chains is harder to mitigate, which is the focus of this article.

Theoretical aspects of massive MIMO are well studied in the literature. Proof of concept testbeds in academia [8]–[11] and field trials conducted by industry participants [12]–[14] have shown the feasibility of massive MIMO systems to achieve high capacity gains. However, these trials were conducted under tightly controlled test setups, which may not be feasible in commercial deployment. Therefore, careful assessment of these systems under more realistic conditions is needed.

FD-MIMO provides a flexible air-interface to support all types of AAS architectures, including analog, digital, and hybrid beamforming [15]. It also supports both TDD and FDD. This flexibility is enabled through flexible channel state information (CSI) acquisition mechanisms that use both channel reciprocity and feedback. The cost of using a higher number of high-speed data converters increases with larger occupied bandwidths proposed in 5G-NR [16]. This makes hybrid beamforming with limited digitized RF configurations more attractive in the near term. Furthermore, each AAS architecture poses its own unique set of challenges in implementation and controlling hardware impairments. Therefore choosing the right design requires a comparison of performance under the impact of the RF impairments.

Prior investigations [17]–[25] have analyzed the impact of channel non-reciprocity caused by RF impairments. Although the results provide useful insights into the nature of performance degradation, the system models considered are much simpler than actual deployment. For example, [17]–[24] only consider single-cell with small-scale fading. Multi-cell with large scale fading, channel estimation errors, and vector normalization was only considered in [25]. However, even in this case and others, the effects of temporal and spatial multipath clustering [26], [27], the effects of channel aging from feedback or scheduling delays, and link

adaptation are not considered. Practical FD-MIMO operation cannot ignore these effects. They are also essential to derive realistic phase and magnitude error tolerance limits for practical AAS. Knowledge of tolerance limits is highly valuable to hardware designers as it enables measurement in a lab to guarantee minimum performance in the field. It also helps make RF component choices based on their sensitivity to factors affecting frequency response. Unfortunately, these limits cannot be directly derived from previous work, and 3GPP does not specify them either. It is difficult to derive the performance analytically under the cumulative effect of the factors mentioned above. Monte-Carlo simulations are better suited to aid such analysis.

It is essential to identify and characterize the sources of phase and magnitude error to understand the limits of an AAS design configuration and find effective mitigation mechanisms. There is a lack of experimental characterization of the phase and magnitude response variations over large antenna array implementations for massive MIMO. Hence, investigations in the past [5], [18], [19] have mostly focused on frequency domain local oscillator (LO) phase noise as the primary source of multiplicative noise. However in practice, LO phase drift [28] and internal temperature variations [29] are also known to affect the stability of multi-antenna systems which have not received attention.

LO phase drift, especially with cheaper oscillators, affects phase variation between RF chains when the delay between training and downlink transmissions is not small [30]. This delay in Long-Term Evolution (LTE) or 5G New Radio (NR) based air-interface can be on the order of a few 10's of milliseconds. It impacts the choice of the LO generation method, which was overlooked in previous studies. In the past, [31]–[33] investigated the effect of temperature fluctuations over fewer RF chains (8) on MIMO performance. In comparison, an FD-MIMO AAS consists of a much larger number of digitized RF chains (8-64) with many more radiating elements (16-512) operating inside a sealed enclosure [3]. This density of analog circuitry and components makes the system even more susceptible to internal temperature variations. However, to the best of the authors' knowledge, there has been no prior characterization of temperature-related effects on FD-MIMO related AAS operation. Understanding the dynamic nature of the fluctuations is critical to appreciate the complexity of implementing calibration to meet error tolerance targets and tradeoff against using cheaper RF components whose frequency response drift with temperature.

Several antenna calibration schemes have been proposed and analyzed in the literature for massive MIMO [20], [34]. However, there is very little experimental confirmation, such as [35], regarding their feasibility in practical systems. There are no verification results available in the literature on the impact of the duration or frequency of calibration on radiation patterns or transmission. Cellular FD-MIMO systems are required to operate without any disruption of data transmission. Unfortunately, most calibration schemes proposed in the literature require the use of time-frequency resources

from active transmission. Hence, a careful analysis of existing methods is required to understand their drawbacks and devise a practical scheme. Such a calibration scheme should ideally apply to both TDD and FDD based FD-MIMO in addition to solving transmission-related challenges and meeting error tolerance limits under the worst-case scenario.

In this article, we address the issues discussed above, as summarized by the next subsection.

B. KEY CONTRIBUTIONS

The main contributions of the paper are:

- 1) In the first part of the paper, we provide an analysis of the effect of phase and magnitude errors on the average sum-throughput of the FD-MIMO system. The analysis uses link- and system-level simulations that model the combined effects of 3GPP defined 3D-channel model [27], multi-cell deployment accounting for inter-cell interference, link adaptation and channel aging due to feedback and scheduling delays. 3D-channel model defined in 3GPP specifically addresses collocated planar arrays that satisfy tower top size constraints with radio propagation in both the azimuth and elevation directions. It takes into account LOS probability, path loss, shadowing, fast fading, probability of a user being indoor or outdoor, and large scale correlated parameters to model multipath that exhibits clustering in time and spatial domain. Industry participants in 3GPP use this model to analyze the performance of LTE or 5G-NR air-interface and calibrate their results [36]. This method makes it an excellent choice to analyze the effects of RF impairments. One exception to the analysis considered in this article is the assumption about perfect channel estimation of the uplink pilot signals. This assumption isolates the effect on performance to RF impairments alone. Channel estimation and channel prediction errors are implementation-dependent [37]. These errors can be modeled separately under ideal hardware conditions, which can be used to further adjust impairment margins. Hence, analysis of the joint effect of channel estimation error and hardware impairment is out of scope for this article.
- 2) Next, we address the issue of characterizing the significant sources of phase and magnitude errors. First, we study the short term LO phase drift between RF chains affecting phase error. We provide recommendations for choosing the right architecture to meet LO coherency requirements. Second, we characterize the effect of internal temperature variations on phase and magnitude. Temperature measurements from an AAS operating in a commercial macro-cell LTE network is provided to demonstrate the dynamic nature of the variation and the need for a suitable calibration mechanism.
- 3) In the last part, we study the calibration mechanisms investigated in literature and discuss some of the

drawbacks. To solve these and the coherency problem, we propose an AAS solution that implements a novel coherent LO distribution method that provides a precise phase and frequency synchronized reference to each RF chain. A novel continuous calibration mechanism that depends on the availability of this stable reference with no interruption to data transmission or other drawbacks is also described. We provide lab and field measurement results to show the accuracy and coherency achieved with the proposed solution.

C. PAPER ORGANIZATION

The paper is organized as follows: Section II presents a brief overview of 3GPP based FD-MIMO and terminology used to specify different AAS configurations. It then describes the analysis of the effect of phase and magnitude error on FD-MIMO performance. Section III presents the effect of LO phase shift and temperature variations on phase and magnitude errors. Section IV provides a review of the calibration methods found in literature and their drawbacks. In Section V, we describe a new approach to building a coherent antenna array solution using a novel LO distribution method and associated calibration with measurements that validate the desired performance. Section VI concludes the paper with final remarks.

II. EFFECT OF PHASE AND MAGNITUDE ERRORS ON FD-MIMO PERFORMANCE

In this section, we present an analysis of the impact of phase and magnitude errors on the FD-MIMO system's performance with different hybrid beamforming-based AAS configurations. We first provide a brief overview of the channel feedback and AAS configurations specified in 3GPP.

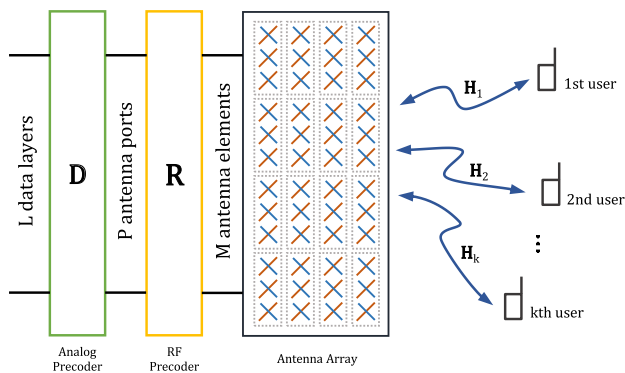
A. CHANNEL FEEDBACK MECHANISMS AND AAS CONFIGURATION IN 3GPP BASED FD-MIMO

FD-MIMO supports both channel reciprocity and channel feedback mechanisms to determine the downlink precoder. 3GPP specifies two underlying CSI feedback mechanisms for FD-MIMO [1], [38]; non-precoded codebook and precoded or beamformed channel sounding using CSI-reference signals (CSI-RS). In the non-precoded codebook feedback mechanism, the UE selects the best index from a codebook based on channel sounding from individual antenna ports at the transmitter. In the beamformed CSI-RS approach, a grid-of-beams, each with a distinct set of CSI-RS signals, is transmitted, followed by the UE feedback report indicating the best beam or a linear combination of beams to help the BS choose a precoder [14]. In this article, we consider the non-precoded codebook feedback mechanism for analysis.

AAS configuration in 3GPP is classified based on the mapping of radiating elements to data converters, as described in [36] and [38]. We will use the same terminology throughout the rest of the paper. An AAS with M rows, N columns, and P differently polarized antennas is represented by (M, N, P) configuration. An example AAS with $(12, 4, 2)$

TABLE 1. Summary of simulation parameters.

System Level	Assumption	Link Level	Assumption
Duplex method	FDD and TDD	Downlink scheduler	Round robin
System bandwidth	10 MHz (FDD), 20 MHz (TDD)	Traffic model	Full buffer
Center frequency	2 GHz	DL channel estimation	Non-ideal using DM-RS
Cellular layout	19 sites, 57 cells	Downlink receiver	MMSE
Network layout	3D hexagonal grid, wraparound	Interference modeling	Explicit intercell interference
Inter-Site distance	500 m	Max number of layers	4 (max 1 layer per user)
Users per cell	10 (uniformly in a 3D space)	Link adaptation	MCS based on reported CQI
BS antenna	(12, 4, 2)	CSI (FDD)	Class A non-precoded PMI
BS element gain	5 dBi	CSI (TDD)	Uplink SRS
BS transmit power	46 dBm	CSI periodicity	5 ms
User noise figure	9 dB	Feedback delay	6 ms

**FIGURE 1.** A massive MIMO system with an antenna array of 128 elements. The antenna configuration is (12,4,2) and port configuration is (4,4).

configuration is shown in Fig.1. This specific antenna array consists of 12 rows and 4 columns with 48 dual-polarized antennas in total. Antenna Port configuration is represented by (V, H), where V and H are the number of ports in the vertical and horizontal direction. An antenna port is a logical port that transmits the same information typically associated with a unique pilot or reference signal (RS). For example, in the case of (4,4) port configuration, 3 radiating antenna elements from each column of a single polarization are combined and mapped to a digital stream (digitized via a data converter). In this case, each data converter is mapped one-to-one with an antenna port. Hence, each column is configured with 4 antenna ports per polarization resulting in a total of 32 antenna ports for the entire array. Note that analog beamforming weights can be optionally used across the radiating elements to create specific radiation patterns per antenna port.

B. SIMULATION SETUP AND RESULTS

We resort to extensive Monte-Carlo simulations to evaluate the average sum-throughput of a multi-cell FD-MIMO network. The AAS configuration used for the simulations is (12,4,2) operating at a carrier frequency of 2 GHz with antenna element spacing of 0.5λ , where λ is the wavelength. The AAS antenna element gain is set to 5 dBi. The specific configuration is chosen since it satisfies the maximum

acceptable size limit imposed on typical commercial tower top deployment at 2 GHz. The bandwidth of the system considered is 10 MHz in FDD and 20 MHz in TDD (with 50% time allocated to downlink). UEs are assumed to have two receive antenna elements with an element gain of 0 dBi.

A summary of other assumptions is provided in Table 1. The simulation is divided into two levels: system level and link level. In the system-level simulation, a regular hexagonal cellular network with 19 sites and 57 cells is generated. 10 users per cell are randomly placed in a 3D area covered by the cells. Large- and small-scale channel parameters between each cell and user are generated following the 3D-channel model described in [27], [39]. The average received signal strength at each user is determined based on the effect of large-scale parameters. This signal strength is then used to determine the best serving cell for that user. Channel small-scale parameters are generated between each antenna element of the base station and each antenna element of the user. In the link-level simulations, waveforms similar to LTE downlink and uplink are generated and passed through the 3D channel defined by the small-scale parameters. The downlink throughput at each user is then calculated based on the received signal from the serving cell and interfering signals from neighboring cells. The average sum-throughput per cell then follows by taking the sum across the scheduled UEs. The users are scheduled in Round Robin fashion. The simulation also implements a dynamic switch between SU-MIMO and MU-MIMO transmission, where up to 4 UEs can be spatially multiplexed. When conditions do not favor improved performance compared to MU-MIMO, the system falls back to SU-MIMO. In all cases, each UE supports a maximum of 1 MIMO layer in the downlink.

Simulation process consists of two steps: channel sounding and data transmission. In FD-MIMO systems, BS uses a downlink precoder (or steering matrix) to separate users and direct the energy toward them to improve the link quality. In the TDD case, the channel sounding is carried out by the Sounding Reference Signals (SRS) transmitted in the uplink. The uplink channel estimates are assumed to be perfect in the simulations. The downlink precoder at the BS is then determined using the zero-forcing (ZF) algorithm [40], assuming channel reciprocity. In the case of FDD, the channel sounding

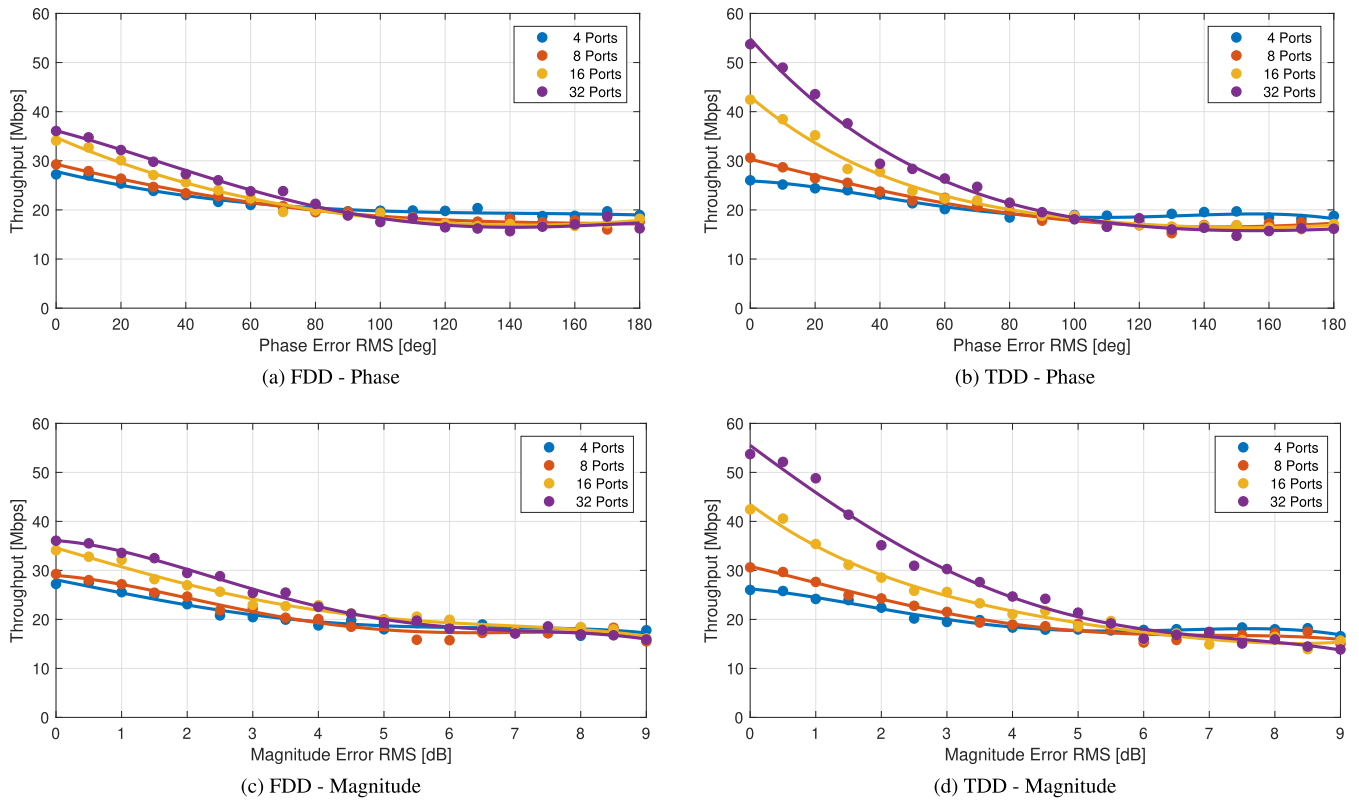


FIGURE 2. Sum-throughput of a FDD FD-MIMO system under phase and magnitude calibration error. Circles shows the simulation data points and solid lines show the best fitting line.

is done using CSI-RS as pilot signals in the downlink. The UE selects a Rank-1 Precoding Matrix Indicator (PMI) from the sounded channel and reports it to the BS. The simulations assume that CSI-RS channel sounding is perfect. The UE uses 3GPP Release-15 non-precoded CSI-RS codebook defined in [41] for PMI feedback. In both TDD and FDD cases, the UE also selects channel quality indicator (CQI) using the pilots transmitted in the downlink (CSI-RS). This feedback enables the base station to perform link adaptation by adjusting the Modulation Coding Scheme (MCS). During the data transmission stage, the BS uses a downlink precoder obtained in the first stage to precode multiple data layers to multiple users. Users use DeModulation Reference Signal (DM-RS), which is also precoded similar to data symbols, for downlink channel estimation and coherent reception of precoded data symbols. A minimum mean squared error (MMSE) receiver is used at the UE to decode data without any knowledge of co-scheduled users in the MU-MIMO case.

The CSI-RS and SRS periodicity are set to 5 ms with a feedback delay of 6 ms. This makes the CSI information at the transmitter to be stale by the time downlink data transmission is carried out. Hence, the simulation also accounts for channel aging that would be typical in LTE or 5G-NR air-interface.

The simulation setup, as described above, accounts for the cumulative effects of the 3D-channel model, inter-cell interference, inter-user interference, channel aging due to feedback delay, and feedback quantization in the case of

FDD. The performance of the FD-MIMO system is evaluated for a different number of digitized antenna ports: 4, 8, 16, and 32 with (2,1), (2,2), (4,2), and (4,4) AAS (V, H) configurations respectively. It is assumed that even under phase errors the UEs are able to lock on to the carrier frequency and perform carrier frequency offset correction and synchronization using broadcast channels as we only evaluate the impact of impairments on the data channel.

The FDD FD-MIMO system's average sum-throughput performance as a function of root mean square (RMS) of the phase error and the number of digital ports is illustrated in Fig. 2a. We define phase error as the deviation in the phase of the signal from that of the desired phase corresponding to the chosen precoder for downlink transmission over each transmit chain. The phase error is modeled as a zero-mean Gaussian distribution with a standard deviation equal to the target RMS phase error. The average sum-throughput with MU-MIMO under ideal conditions with 32 and 16 ports is equivalent to a cell average spectral efficiency of 3.6 bps/Hz and 3.4 bps/Hz respectively. These results are very close to those provided in [38] with CSI-RS based Class-A feedback mechanism using the LTE codebook. The results also fall in the range of that provided in [36]. Hence, the simulations are well-calibrated to follow those provided in the standards body.

It can be observed that the FDD FD-MIMO system with MU-MIMO improves the average sum-throughput by about

two times over SU-MIMO as the number of ports increases to 32 under ideal conditions. As the number of ports increases, the throughput of the system improves over the 4-port (2,1) configuration. The improvement is mainly due on the additional resolution provided by a more significant number of ports in the horizontal or vertical dimension. As the RMS phase error increases, the throughput decreases in all cases due to the increase in inter-user interference. Above 80° RMS phase error, the performance in all cases hit a floor as the BS always selects SU-MIMO for transmission. At these levels of phase error across antenna ports, the precoder chosen for downlink transmission gets highly distorted resulting in worst-case performance. The 32-port (4,4) system starts performing worse than a 16-port (4,2) system with no phase error at 20° RMS phase error and worse than an 8-port (2,2) system with no phase error at 50° phase error. This performance suggests that as the number of ports increases, the improvement in sum-throughput is only assured if the RF impairments are tightly controlled. For example, from a design tradeoff perspective, a 16-port system that achieves lower than 10° of RMS phase error is a better choice than a 32-port system that has higher phase error with higher implementation complexity.

Fig. 2b shows the average sum-throughput performance of TDD FD-MIMO system with ZF precoder as a function of the RMS phase error and number of antenna ports. The average spectral efficiency per cell under ideal conditions corresponds to 5.3 bps/Hz with a 32-port system and 4.2 bps/Hz with a 16-port system. These are slightly better results compared to 5.85 bps/Hz with 64 ports and 3.55 bps/Hz with 16 ports reported in [36] for SRS based TDD FD-MIMO systems. These performance numbers again show that the simulation is well-calibrated against those provided in the 3GPP standards body. Similar to the FDD case, better performance can be achieved with a greater number of digitized ports. Compared to the FDD case, the TDD system shows better performance due to channel reciprocity, which provides better CSI than quantized codebook-based feedback. However, as the phase error increases, the performance of TDD system degrades rapidly and approaches that of an ideal FDD system beyond 20° RMS. The beamforming and spatial multiplexing gains degrade with increasing phase errors. Beyond 80° RMS phase error, the system only supports SU-MIMO with a distorted precoder resulting in low performance. This shows that the reciprocity-based systems provide higher performance gains only at very high coherency.

The higher sensitivity of TDD compared to FDD to phase errors can be explained by looking at the sensitivity of the nulls of a beam pattern to phase errors. Analysis in [42] showed that an ideal null depth of -25.62 dB from the peak power at the receiver locations of co-scheduled users with the ZF precoder requires the transmit nodes to be within 3° RMS phase error. Using the same analysis, the null depth at 20° RMS phase error decreases to -9.14 dB. This decrease results in a 15 dB increase in interference from each spatially multiplexed user from the ideal case. In comparison, the FDD

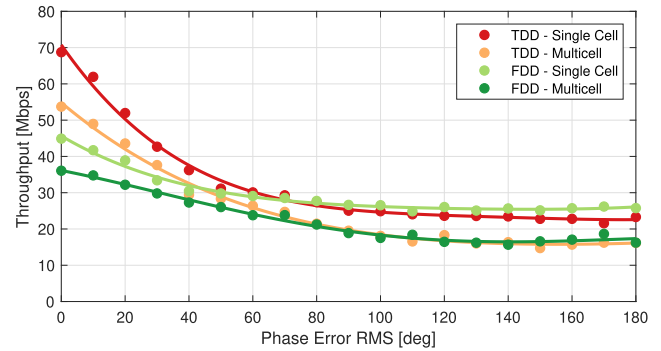


FIGURE 3. Comparing single-cell and multi-cell scenarios. Circles shows the simulation data points and solid lines show the best fitting line.

system can be viewed as a system that provides separation between the main lobes of the users rather than the null placement. Since the gain from the main lobe of a beam is less sensitive to phase errors [42], the FDD system degrades less as a percentage from the peak. The results from simulations confirm this effect. The percentage loss in sum-throughput at 20° RMS error is about 20% under TDD and 3% under FDD when compared to the ideal case.

Fig. 2c and Fig. 2d depict the performance of a FD-MIMO system in the presence of magnitude errors for similar simulation settings. The system is assumed to have perfect phase alignment as magnitude error varies from 0 dB to 9 dB where 0 dB corresponds to zero magnitude error in linear scale. The magnitude error is modeled as an additive zero-mean Gaussian random variable with a standard deviation of $10^{0.2/10} - 1$. Fig. 2 shows that when the magnitude error increases, the sum-throughput decreases. The advantage of using more digital ports disappear when the error increases beyond 4 dB RMS. Sum-throughput of a 32-port system with RMS magnitude error greater than 1.5 dB gets worse than an ideal 16-port system. Like the phase error case, adding more antenna ports only improves the performance, if the magnitude error can be kept the same. The TDD system is again more sensitive to magnitude errors than FDD, as creating interference nulls for MU-MIMO becomes difficult. Its performance becomes similar to an ideal FDD system at errors greater than 1 dB RMS. Beyond 5 dB RMS magnitude error, the BS always selects SU-MIMO with a precoder that is distorted. This results in the worst-case Rank-1 SU-MIMO performance.

Fig. 3 compares the sum-throughput of the single-cell and multi-cell scenarios for both FDD and TDD 32-port systems against phase errors. In the single-cell case, the performance is only affected by degradation from intra-cell inter-user interference that increases with phase error. In the multi-cell case, the ideal hardware performance without phase errors already includes performance degradation due to inter-cell interference. Adding phase error degrades this further at a slower rate compared to the single-cell case as the inter-cell interference does not increase at the same rate as inter-user interference between 0° to 20° . As phase errors increase and inter-user interference become dominant, the performance of

single and multi-cell approach each other between 20° to 80° . At very high phase errors beyond 80° , the system switches completely to Rank-1 SU-MIMO with distorted precoder. Hence, at this point, the inter-cell interference again becomes the dominant factor. As expected, the sum-throughput performance gap between single and multi-cell again becomes evident.

Non-ideal hardware typically causes a combination of phase and magnitude errors. Based on the above results, limiting phase error to 10° RMS and magnitude error to 0.5 dB RMS for FD-MIMO systems can ensure performance within 10 – 15% of ideal hardware performance. These limits can be used as a minimum specification for phase and magnitude error margins for RF conformance tests. The phase and magnitude errors between RF chains can be measured in a lab to conform to these limits to assure minimum performance in the field. In the next section, we describe two significant contributors to RF impairments that make it challenging to design an AAS to meet these limits for commercial deployments.

III. CHARACTERIZING RF IMPAIRMENTS AFFECTING PHASE AND MAGNITUDE STABILITY

The frequency response across RF chains in an AAS can differ from one another and also vary in time. Static variations result from factors such as manufacturing tolerances and mutual coupling between antennas. Non-linear changes in frequency response of RF components during operation results in variation over time. The static variation between RF chains can be measured and compensated using test equipment along with near/far-field measurements in a factory setting before deployment [43], [44]. On the other hand, the dynamic variations require correction during operation to avoid performance degradation. In this section, we investigate the major contributors of phase and amplitude variations that include LO phase drift and variation in frequency response across RF chains from internal temperature variation. We provide measurements from both lab bench tests and a fielded system to characterize the behavior.

A. PHASE VARIATION DUE TO LO PHASE DRIFT

Change in phase response due to unequal phase variations across LO outputs used for mixing at each RF chain is one of the significant sources of phase error in active antenna arrays. Three different approaches for LO generation and distribution have been proposed in the literature for massive MIMO:

- i Separate LO (SLO) where a LO signal is generated by an independent oscillator locally for each transceiver [5], [18].
- ii Phase-locked loop (PLL) based where a low-frequency reference is first distributed to each transceiver and a PLL is then used to generate desired LO frequency [19] as shown in Fig. 4.
- iii Common LO (CLO) where the desired high-frequency LO signal is directly distributed from a common source

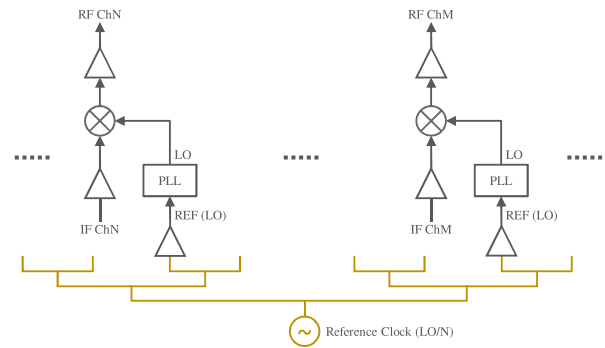


FIGURE 4. PLL-based LO distribution.

to all transceivers [18] instead of using a low frequency signal shown in Fig. 4.

Phase noise in these architectures were evaluated from a frequency domain perspective and were shown to satisfy massive MIMO performance gains. However, the phase errors arising from the effect of short term time stability issues of LOs were ignored.

The effect of short term time stability of LOs on distributed-beamforming was theoretically evaluated in [45]. It was experimentally verified in [28]. The results in these studies show that independent LOs (TCXOs) similar to the SLO case, drift out of phase within a few milliseconds, drastically degrading beamforming gain. To maintain beamforming gain, frequent training and calibration is required with very high overhead. This discrepancy between the two cases is further explained below with a numerical example.

Phase variation caused by an LO is generally modeled by a Wiener process. The variance of the phase error at time duration equal to the time elapsed between uplink channel estimation and data transmission is given by [46]

$$\sigma_{\phi}^2(t, \tau) = 4\pi^2 f_c^2 c \tau = 4\pi^2 f_c^2 c T_s \frac{\tau}{T_s} = \sigma_{\phi T_s}^2 l, \quad (1)$$

where f_c is the center carrier frequency, c is the phase noise constant of the free-running oscillator, τ is the time elapsed between training and downlink transmission, T_s is the OFDM sample time duration, $l = \tau/T_s$ is the number of samples occupying the elapsed time, and $\sigma_{\phi T_s}^2$ is the variance computed from the frequency domain phase error spectra of the LO. This frequency-domain model does not capture the short-term time-domain frequency stability of an LO [30], especially in low-cost crystal oscillators such as VCXOs and TCXOs, which are more commonly used in cellular radios [47].

Allan variance [30], on the other hand, characterizes the phase error variance from an LO such as TCXO due to time-domain short term stability. It is the two-sample variance of an oscillator output for a given gating period τ . A generic model for LO phase error that includes both the frequency domain and short term time domain factors is described in [28], [45], [48]. This model uses a state-space model where multiple Wiener processes, each with different variances, are

considered. Each Wiener process is mapped to a dominant noise source depending on the gating period considered for Allan variance measurement and the type of oscillator.

In the case of a TCXO and a gating period ranging from 100 μ s to 1 s, the underlying noise is dominated by flicker frequency and white frequency modulation noise [47]. In this time interval the mean squared phase error for the Wiener process is given by [48]

$$\sigma_{\phi}^2(t, \tau) = 4\pi^2 f_c^2 \sigma_{Adev}^2(\tau) \tau^2, \quad (2)$$

where $\sigma_{Adev}(\tau)$ is the Allan deviation. Allan deviation for TCXOs is typically specified as 10^{-9} [49].

Now to understand the difference between the two methods, consider an LTE air-interface with a bandwidth of 20 MHz at a carrier frequency of 2 GHz. For the gating period of 100 μ s to 1 s, the product $\sigma_{Adev}(\tau)\tau$ is a constant [30]. Hence for an elapsed time of 1 ms, the variance in phase error computed using the time domain method is given by

$$\sigma_{\phi}^2(t, \tau) = 4\pi^2 f_c^2 (10^{-9})^2 = 158 \text{ rad}^2. \quad (3)$$

This corresponds to an RMS phase error of 720° . In contrast, the frequency domain phase noise from an LO is computed as twice the one-sided phase error spectra and is expressed as

$$\sigma_{\phi Ts}^2 = 2 \int_{f_L}^{f_H} \mathcal{L}(f) df, \quad (4)$$

where $\mathcal{L}(f)$ is the one-sided power spectrum of phase error, f_L is 15 KHz and f_H is 10 MHz for the LTE example. For LTE systems, the integrated phase error is typically limited to -40 dBc corresponding to an error vector magnitude (EVM) of 1% [50]. This is equivalent to a frequency domain integrated phase noise of $\sigma_{\phi Ts}^2 = 10^{-4} \text{ rad}^2$. The variance in phase error over an elapsed time of $\tau = 1$ ms similar to that considered in the Allan variance case from (1) is computed as $\sigma_{\phi}^2(t, \tau) = 10^{-4} \times 30720 = 3 \text{ rad}^2$. This corresponds to a RMS phase error of 99° which is much lower than that computed using the time domain approach. Hence, it is important to consider time domain short term stability of LOs to evaluate phase coherency of massive MIMO when the time intervals between training and data transmission are large.

Based on results in section II-A, 10° RMS phase error is a good target for relative phase coherency in the array. Note that this limit is for the entire RF chain, and hence the LO related error margin has to be much more stringent. To achieve 10° RMS phase error, the integrated phase noise only needs to be less than -60 dBc based on (4). This requirement can be satisfied by the most commonly used LOs (TCXOs) for cellular application [47]. To meet the same specification using the time-domain method, the Allan deviation based on (3) needs to be 10^{-11} , which can only be satisfied by higher quality LOs such as OCXOs. OCXOs as LO on every RF chain for SLO type architecture would make the solution cost-prohibitive for commercial use.

The PLL based methods involve additional intermediate PLL ICs or distribution clocks [9], [51] with low-skew buffering circuits and transmission lines with matching lengths to

distribute reference clock signal. This requires more careful design to address phase noise from the intermediate ICs to achieve phase coherency with constant phase offsets between RF chains. Achieving phase alignment between RF chains further requires periodic calibration and digital adjustments resulting in more overhead. Hence, a PLL based solution is challenging in terms of higher implementation complexity.

The CLO architecture is known to achieve phase coherency better than 1° RMS [52]–[54]. However, it is also very challenging to implement across a large array. Distributing a high-frequency signal (> 1 GHz) is more challenging than a low-frequency reference signal used in the PLL method due to higher sensitivity to impairments at smaller wavelength. The challenges related to this method are well-covered in [55].

Based on the above discussion, it is clear that it is harder to meet phase error performance with a cost-effective solution. Therefore it is highly desirable to explore new LO generation and synchronization methods for practical FD-MIMO AAS systems.

B. PHASE AND MAGNITUDE RESPONSE VARIATIONS IN RF CHAINS DUE TO INTERNAL TEMPERATURE FLUCTUATION

Fig. 5 shows simplified typical transmitter chains of an active antenna array. Receiver chains are similar and are not shown here. Notice that despite the usual label of “digital radio” blocks, these transmitter chains are made of pure analog/RF circuits such as amplifiers, mixers, filters, electronic attenuators, voltage regulators, and data converters. The fabrication process, bias, and temperature variations affect the frequency response of most analog/RF components; therefore, phase and amplitude errors will normally occur in any of these analog/RF blocks. These errors are in addition to the LO phase drifts we covered earlier.

There are many techniques available to minimize or even eliminate the phase and magnitude errors occurring in analog/RF components, but these techniques usually come with a heavy penalty on cost. Process variations between

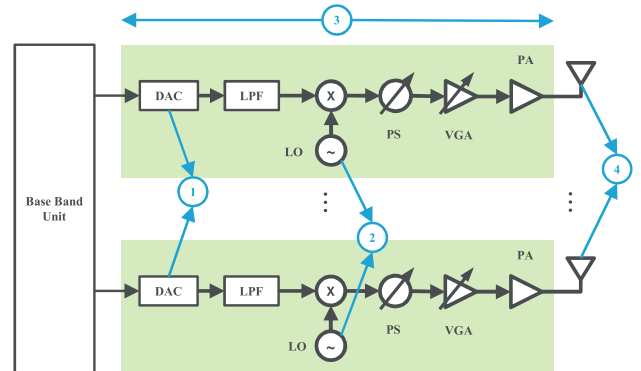


FIGURE 5. Sources of calibration errors in an active antenna array on the transmit side.

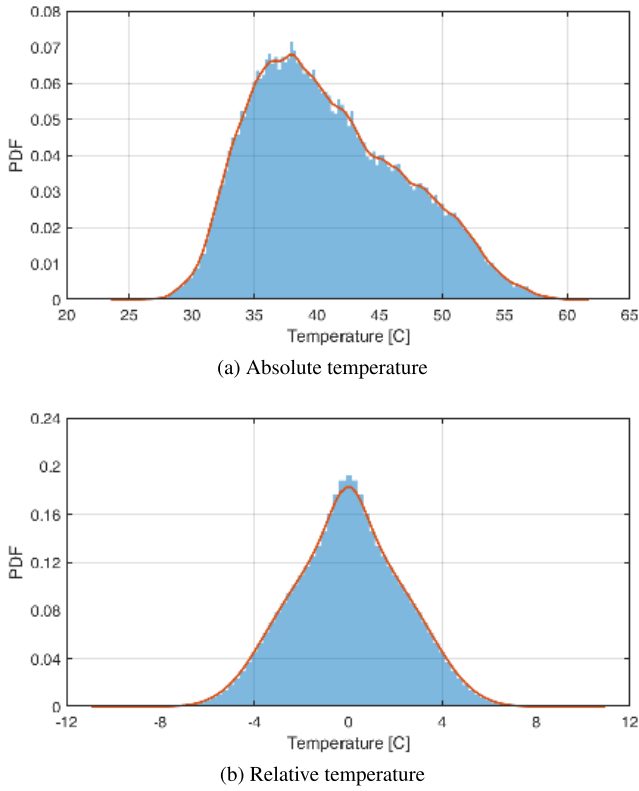


FIGURE 9. Distribution of absolute and relative temperature.

temperatures across the array move in the same direction. The relative temperature variation w.r.t. a reference antenna is shown in Fig. 8b. The relative temperature on any tile w.r.t. a specific reference tile is as high as $\pm 9^\circ\text{C}$. We further plot the probability density function (PDF) of both the absolute temperature over a single tile and relative temperatures between any two tiles measured over a longer duration of 4 weeks in Fig. 9. The absolute temperature varies as much as 35°C . This corresponds to a maximum of 50° in phase error and 5 dB in magnitude error over a single RF chain. However, the relative temperature at any instance has lower variation under $\pm 11^\circ\text{C}$. This corresponds to a maximum relative phase error of 22° and a maximum relative magnitude error of 2.2 dB based on Fig. 7. If these values are considered to be the peak at roughly 4σ , then the corresponding RMS phase and magnitude errors are 5.5° and 0.5 dB respectively. Note that this result only considers the effect of change in the frequency response of the RF chain to temperature and does not include other sources of error such as LO or sampling offset.

The above data shows that temperature variation can be a significant contributor to RF impairments. However, these results are specific to HDAAS implementation. The internal temperature gradient can depend on several factors such as the size of the array, thermal design, the heat-dissipating components such as power amplifiers (PAs), digital processors and data converters, amount of data traffic,

ambient temperature and environmental factors outside the array. Hence, temperature-related change in frequency response depends heavily on the construction and design of the AAS. In the case of HDAAS, there is a linear relationship between temperature and phase or magnitude variation, as observed in Fig. 7. The PDF of relative temperature variation shows a probability distribution that is closer to Gaussian distribution. Therefore as described in section II-A, it is a good approximation to model the relative phase errors using a Gaussian distribution and the magnitude errors using a log-normal distribution.

Calibration is necessary to compensate for phase and magnitude deviations observed above. The frequency at which calibration routines need to be executed depends on the rate of the dynamic variations. If internal temperature fluctuation is considered as the dominant factor, then the rate of change of relative temperature can be used as a rough estimate of calibration frequency. In the case of HDAAS, based on Fig. 8b, the worst case appears to be 1 hour where relative temperature jumps by $6\text{--}8^\circ$. However, this is highly design-specific. Frequent calibrations can be detrimental to capacity improvement, especially if transmission resources need to be reserved for calibration. Hence, we evaluate the effectiveness of different calibration methods described in literature before presenting a novel AAS architecture and related calibration.

IV. REVIEW OF CALIBRATION METHODS

In typical macro-cell deployments, the Radio Unit (RU) implements the RF front end and the digital front end on a tower top. The digital baseband unit (BBU) is located at the bottom of the tower or in a data center connected to fiber. The RU and BBU are both required to operate coherently. This coherency is achieved by making the RU time synchronized with the BBU using a tight synchronization protocol over a front-haul interface [57]. An FD-MIMO AAS is expected to operate similarly. Coherency hence implies that the entire system from the antenna to the BBU is coherent. To achieve this tight coordination between the calibration process at the RU and transmission resource scheduling at the BBU [58].

A good overview of different calibration methods investigated for massive MIMO systems is provided in [20] and references thereof. In general, relative calibration [8], [35] requires calibration coefficients to be calculated as the ratio of the frequency responses measured between a reference RF chain and other RF chains. This is calculated as $c_i = \frac{t_i r_{\text{ref}}}{r_i t_{\text{ref}}}$ where t_i and r_i are the transmit and receive frequency responses of the i th RF chain and t/r_{ref} is the reference RF chain. Let $h_{k \rightarrow i}$ be the channel response between the k th user and i th receive antenna. The CSI estimate using the calibration coefficient is computed as

$$\hat{h}_{i \rightarrow k} = c_i r_i h_{k \rightarrow i} = \frac{t_i r_{\text{ref}}}{r_i t_{\text{ref}}} r_i h_{k \rightarrow i} = \frac{r_{\text{ref}}}{t_{\text{ref}}} t_i h_{k \rightarrow i}. \quad (5)$$

It can be observed that $t_i h_{k \rightarrow i}$ is the desired CSI and $\frac{r_{\text{ref}}}{t_{\text{ref}}}$ is a constant offset that is applied to all RF chains. Hence using the calibration coefficient the modified CSI estimate only

differs from the desired CSI by a constant offset. This offset gets compensated by the downlink channel estimation at the UE using DMRS during precoded transmission. This makes the linear precoder adjustment determined after calibration still remain valid. The performance of the system heavily relies on the accuracy of the calibration coefficient. Each calibration run requires either measurement across all or partial set of RF chains. This can be time consuming and make the RF chains involved in calibration go out of commission affecting performance.

In order to understand the impact on transmission we next examine specific methods proposed in the literature to carry out relative calibration:

- i Circuit based calibration: This involves transmitting pilot signals or a tone into a calibration circuit network covering the entire active antenna array to determine calibration coefficients [35]. The sounding signal power needs to be split equally across different branches of the network to ensure similar signal distribution and maintain response identical to that observed during transmission. The circuit based calibration requires routing over transmission lines, directional couplers, switches, and dividers across multiple boards [3] with careful considerations to avoid interference or superposition. This type of calibration can be quite challenging and complex. In the case of partial array calibration schemes that try to operate in parallel with active transmission, high isolation can be hard to achieve between the calibration network and other receive paths. Therefore in most cases, the active transmission would either need to be blanked out for the duration of calibration or calibration would need to be carried out on certain reserved frequency resources. Both methods have the disadvantage of degrading the capacity and adding additional overhead associated with the coordination of user-resource allocation in the baseband and calibration process.
- ii Calibration using over-the-air measurements: Channel sounding based on mutual coupling with over-the-air measurement between adjacent antennas was suggested in [59]. Relative calibration using mutual coupling requires sequential channel sounding over different pairs of transmit and receive chains in order to avoid mutual interference. Poor isolation during calibration can affect uplink receiver performance. Channel sounding during calibration can also pose challenges satisfying downlink RF conformance requirements related to adjacent channel leakage ratio (ACLR) and EVM.

The relative calibration methods also have the disadvantage of only applying to reciprocity based TDD system. It is also not easy to implement for hybrid beamforming architectures [34]. Hence, novel calibration mechanisms that cause minimal degradation to data transmission are highly desirable. In the next section we describe an absolute calibration method and its implementation over HDAAS architecture to solve the inadequacies describe above.

V. ABSOLUTE ARRAY CALIBRATION BASED ON A NEW APPROACH FOR COHERENT LO GENERATION

First, we explain the basic principle behind absolute calibration before providing details of its implementation on HDAAS. Absolute calibration depends on the availability of a common reference for phase and magnitude invariant across RF chains over temperature and time. If such a reference is available, then the relative phase and magnitude offset calculation w.r.t. a reference RF chain is not required to calibrate the array. Instead, channel sounding can be carried out per transmit or receive RF chain and the result compared to the common reference to make corrections. This operation can be carried out on-demand or in a periodic manner, depending on the observed stability of the array coherence. If the compensation is handled at RF continuously at each radiating element, then the downlink precoder can be applied directly without any corrections as determined by channel estimation or feedback in the digital domain. This is because the array continuously maintains coherency and operates close to the ideal response at the radiating element level. The absolute calibration method is not dependent on channel reciprocity and, hence, applies to FDD as well as TDD. The coherence of the reference across the array is critical. This reference also needs to be phase coherent with the clock used for baseband processing to achieve end-to-end coherency with BBU.

Next, we describe a new approach to manufacturing a low-cost coherent Massive MIMO system based on our large active array concept called High Definition Active Antenna System or HDAAS along with its coherent LO generation and absolute calibration mechanisms. The key system properties of HDAAS are:

- a) Large RF coherent, active aperture with no data-flow interruptions for calibration
- b) Digitization that extends support to flexible hybrid beamforming architectures including sub-array and full connection models [36]
- c) Ability to control phase and magnitude at every radiating element at RF and digital precoding over digitized RF chains
- d) Agile beam steering and beam switching at RF using stored beam pattern coefficients at each radiating element
- e) Flexibility of digital control of the AAS from either baseband or any network element for coordination across cells.

A. HDAAS ARCHITECTURE

The high-level architecture of HDAAS is shown in Fig. 10, constructed of power supplies and three main circuit boards: an active IF/RF/digital front-end board called a Tile, a large passive board called Master-Board (MB) and a Digital Front-End (DFE) board. The Tile contains standard radio circuits such as antenna elements, connectors, RF couplers, RF filters, up/down conversion mixers, phase rotators, variable-gain amplifiers, PAs, LNAs, power regulators,

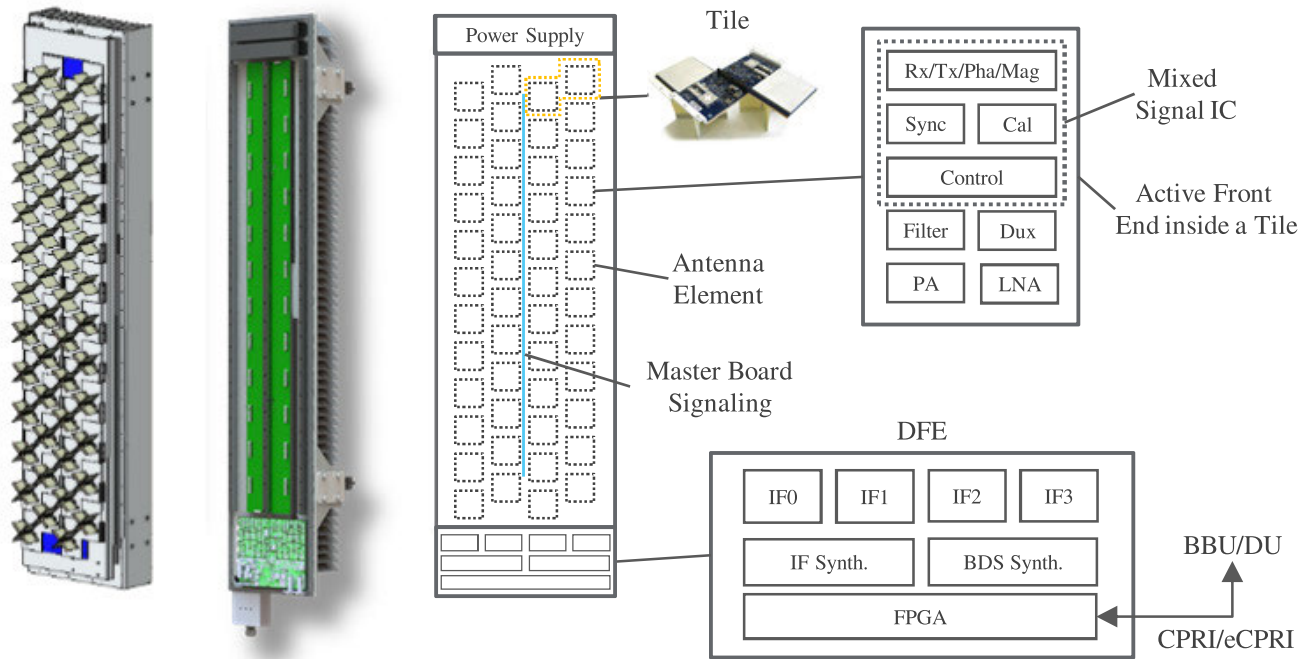


FIGURE 10. HDAAS: a novel low-cost coherent massive MIMO design.

memory, digital microcontroller and novel mixed-signal circuits for RF synchronization and system calibration. Most of these standard and novel circuits are integrated into a small, custom mixed-signal IC fabricated in an old, very low-cost Si technology (currently 180 nm BiCMOS). This integration is essential because HDAAS has a very large number of phase rotators and variable gain amplifiers supporting each transmit and receive beam that can be digitally controlled on each Tile. Typically, a Tile contains circuits for two cross-polarized active antenna pairs, but other Tile designs supporting more active antenna pairs are possible.

The MB is a PCB that contains transmission lines for analog IF signals, LO signals, a few synchronization/calibration signals, DC supply lines, and digital control signals. It also contains connectors appropriately placed such that many Tiles and one DFE can be plugged in easily to construct the HDAAS.

The hardware and firmware on the DFE board support a digital front-haul interface (CPRI to BBU in 4G or eCPRI to the decentralized unit (DU) in 5G), digital processing to implement digital front end functions and some physical layer functions, digitization at IF for all RF chains and system synchronization/calibration support. The DFE software functionality includes HDAAS system management support for hardware control, beamforming, and calibration. The DFE board includes an FPGA to implement firmware along with ARM-based processors to support software, although other designs, for example, with SoCs are also possible.

When the HDAAS assembled as discussed above is powered up, the DFE has full control over the entire system,

including over the phase and magnitude settings of each radio signal path on each Tile. These settings can be changed every millisecond or even faster while maintaining synchronization with the timing derived from the baseband unit over the front-haul interface. The system RF calibration may be programmed to run continuously or on-demand basis in the background to maintain RF coherency over the entire aperture. The digital precoding and beamforming functions can be controlled by the DFE locally as an independent process. This can be driven by the FD-MIMO baseband unit or controlled by a network element management system (EMS) residing in the cloud.

1) SYSTEM RF SYNCHRONIZATION

One of the most critical design features of HDAAS is the accurate RF synchronization (within a few degrees of error) of the Tiles w.r.t. each other. An RF system reference clock in the DFE is transmitted over the MB serially such that two instances of the clock arrive at each Tile over two separate paths. These paths are constructed such that the average arriving time between the two paths is a constant [60], even on a low-cost PCB. This is accomplished by placing these paths on the MB very close to each other. Such placement makes any local variations of the board materials, traces, and temperature cancel each other to a high order. A simple version of this technique was used in the VLSI clock distribution method described in [61].

A unique PLL circuit on the Tile extracts the average time of the two paths with high accuracy, and each Tile generates a local clock signal, which by construction is precisely in phase

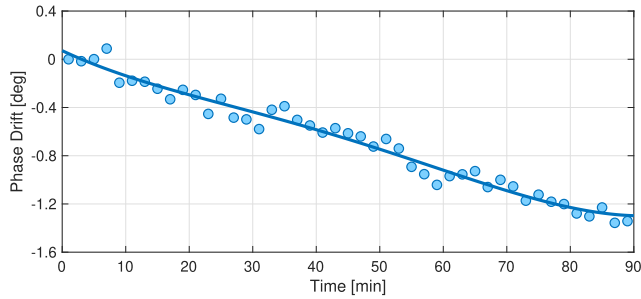


FIGURE 11. PLL phase drift over time with BDS method.

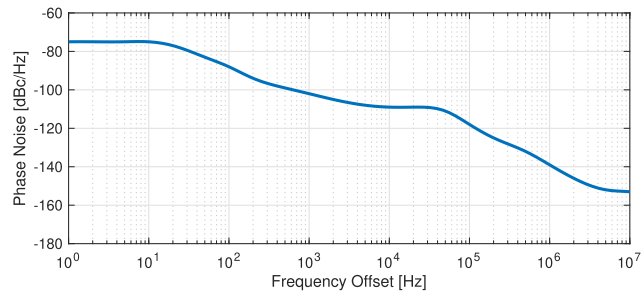


FIGURE 12. Phase noise power spectral density with BDS method.

(with insignificant error) with the local clock of all the other Tiles. Since these clock signals are already at RF, there is no need to use local frequency multiplication (often used in PLL clock generation), which is notoriously prone to generating significant phase errors. Once the Tiles are synchronized at RF, maintaining RF coherency over the entire aperture (all Tiles) is simplified significantly. For example, all LO signals on the Tiles are derived from this local RF clock signal and are automatically in almost perfect phase alignment with very low phase noise.

Fig. 11 shows the measured LO phase variation of this method for 90 minutes at an ambient temperature of 21 °C. The total phase drift due to secondary effects in the Tile is only 1.3°, which is insignificant and is removed during Tile calibrations. This method performs as well as the CLO method mentioned earlier. A significant additional advantage is the insensitivity to temperature variations, which means the LO distribution does not need re-calibration as the PLL or SLO approaches. The power spectral density of the LO signal extracted is shown in Fig. 12. The integrated phase noise is about -83 dBc which is well below -60 dBc requirement discussed earlier.

The LO generation and distribution method can be implemented at a very low cost, and hence it is an attractive solution for FD-MIMO array implementation. This method can also be extended to provide coherent clocking across multiple high-speed data converters when a significant number of RF chains need to be digitized.

2) TILE CALIBRATION

The Tiles contain extensive standard-quality, low-cost IF and RF circuits, subject to large fabrication and operational

drifts due to temperature, humidity, and aging. The brute force solution, often adopted by classical military phased arrays, would replace these standard-quality low-cost circuits with very high-performance but very expensive components with stable electrical characteristics. Additionally, we would have to employ expensive high accuracy factory calibrations. Instead, HDAAS uses commercial off the shelf quality low-cost components and coarse low-cost factory calibrations. The key to obtaining and maintaining RF coherency is a sophisticated background Tile calibration method, which monitors with high precision the RF path drifts in phase and magnitude. It corrects them automatically against the known coherent reference RF signal available at each Tile. During AAS operation, each active antenna is electrically removed from the system one at a time, calibrated for both transmit and receive paths individually, and reconnected to the system. The phase and magnitude deviation measured on a benchtop after applying the correction by the calibration scheme is shown in Fig. 7. The calibration method accurately estimates and corrects the deviation for different temperature-related variations. This method includes analog circuit mechanisms in the Tile for phase and magnitude drift sensing, digital firmware in the Tile for correcting the drifts, and software managing the calibration frequency. Notice that unlike the relative calibration methods, HDAAS can calibrate one or a few active signal paths independently at a time without the need to interrupt the operation of the rest of the system. When one or even a few Tiles are removed electrically, the system performance degradation is practically imperceptible. This result is also the consequence of the fact that the differential electrical characteristics of the MB traces are quite insensitive to operational drifts as described next.

3) MB CALIBRATION

As the MB is fabricated in a low-cost PCB technology, the various transmission lines it contains are not guaranteed to be mutually matched after fabrication. For proper HDAAS operation, it is necessary to ensure all IF paths in the MB are equal in electrical length. This consistency is achieved with a similar calibration scheme to the Tile calibration. The only difference is that while the Tiles must be calibrated very often (e.g., every half hour), the MB calibration is typically done only once at system boot time. As mentioned earlier, the MB differential drifts with temperature and other operational conditions are insignificant. Data collected over multiple HDAAS operational units in the field over long periods including very hot summers, and very cold winters have shown no measurable performance degradation without re-calibrating the MB.

B. LAB AND FIELD TESTS

In order to validate the performance of the coherent LO distribution and the proposed calibration scheme, we measured the radiated beam patterns in both a near field anechoic chamber and far-field radiated testing range. The phase and magnitude coefficients applied to the mixed-signal ICs correspond

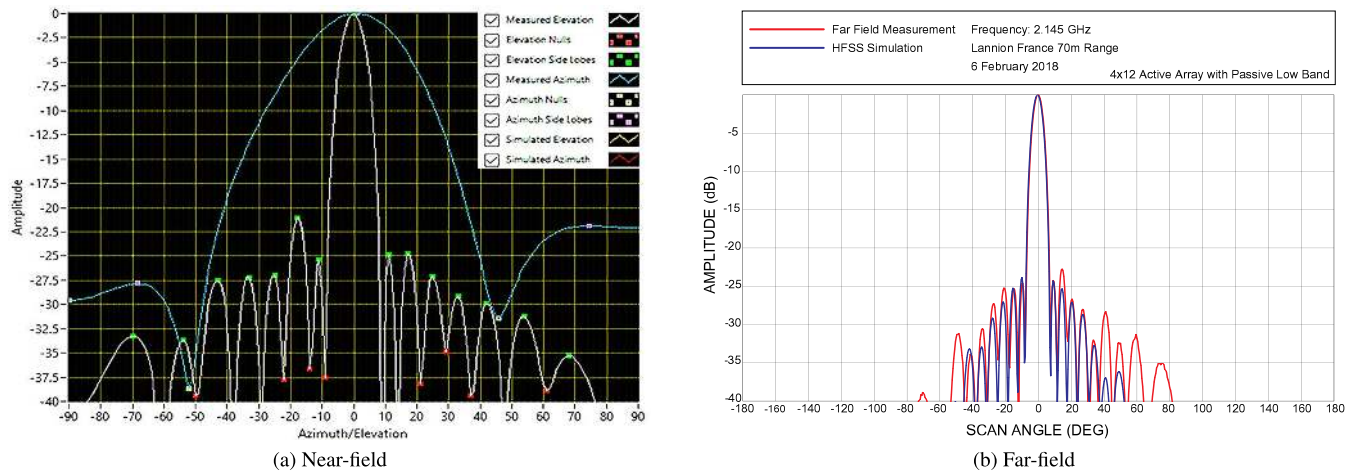


FIGURE 13. Comparing near and far field boresight measured radiation patterns against simulation.

to amplitude taper with a bore-sight beam. RF calibration is activated in a continuous sequential mode to assess its impact on the radiated pattern. Near-field test result for the generated narrow beam is shown in Fig. 13. The array under test corresponds to the system specified earlier for HDAAS. Hence, the radiation pattern corresponds to a 12×4 array with 48 elements in one polarization. Both Azimuth and Elevation cuts of the 3D radiation pattern are shown. The figure also includes plots generated from high-frequency structure simulator (HFSS) using the exact models of antenna elements and spacing. These correspond to the ideal expected beam patterns from the antenna array and are labeled as Simulated Elevation and Azimuth. Near-field measured results are overlaid on the same plot. It can be observed that there is an excellent match over the main lobes, the side-lobes, and even the null-depths. The match in the null-depths as explained in section II-B imply much better performance than 5° RMS phase error and 0.5 dB RMS magnitude error. The radiated beam from an outdoor far-field test is shown in Fig. 13b. This result shows the azimuth cut of the measured radiation pattern. The blue plot corresponds to the expected beam pattern result from HFSS. The red plot corresponds to the measured result. There is an excellent match over the main-lobe. There is less than 5 dB mismatch over the side-lobe levels closer to the main lobe. More importantly, from an MU-MIMO user separation perspective, there is an excellent match over the null depths, which demonstrates the stability achieved by the system while calibration is continuously performed in the background.

Encouraged by these results, several field trials have been conducted in various networks carrying commercial LTE traffic supporting Rel-8/10 FDD 2×2 MIMO LTE air-interface. In these deployments, cell re-shaping by modifying the phase and magnitude coefficients of the mixed-signal ICs was carried out to manage inter-cell interference and improve downlink and uplink SINR. The results from one of these trials have been presented in [56]. Here, we reproduce one

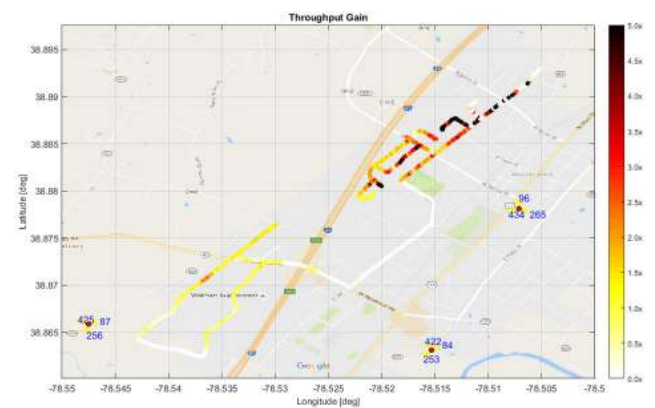


FIGURE 14. Downlink user throughput gain from cell-resaping using HDAAS in a commercial LTE 2×2 MIMO network.

of the drive-test results discussed in the paper, as shown in Fig. 14. The drive-test was conducted with a commercial UE attached to the cell that uses HDAAS for transmission. The plot shows a map of the drive test path and the gain in downlink UE throughput compared to the legacy system. The legacy system used a passive antenna that had 17 dBi antenna gain with 65° half-power beamwidth (HPBW). HDAAS was configured with a narrow beam with an array gain of 24 dBi and an HPBW of 26° . The total transmit power in both cases was 80 W. First, it should be noted that the UE was able to attach to the network and transfer data. This result would not be possible without having a coherent system. Second, the throughput gains of 2-5 times that of the legacy system were observed. This improvement is due to improvement in SINR from beamforming that resulted in the corresponding improvement in MCS through link adaptation and rank improving from 1 to 2.

Successful split-sector and cluster site field evaluations with cell shape optimization were also carried out using the same array to evaluate performance gains as described in [62].

These results validate the synchronization and coherency achieved by the array architecture and calibration scheme to enable low-cost AAS for cellular deployment. The results also show the feasibility of the novel LO distribution and calibration for all types of AAS architectures and configurations supported by FD-MIMO.

VI. CONCLUSION

Transforming any novel communications systems idea from concept to real-world requires overcoming practical challenges. In the case of the FD-MIMO system proposed in 3GPP standards, this requires overcoming non-ideal behavior due to RF impairments. In this article, we specifically addressed the challenges posed by phase and magnitude variations in RF chains. We analyzed the impact of these variations on average sum-throughput of FD-MIMO systems and derived their limits to attain satisfactory performance. We studied and characterized the main factors that impede an AAS from attaining these performance specifications. We discussed some of the drawbacks related to calibration mechanisms and presented a novel scheme to overcome them with a novel active antenna array architecture.

Key observations are summarized as follows:

- i If digitizing a large number of RF chains leads to higher phase and magnitude errors, then it may be better to deploy a system with fewer digitized antenna ports with better coherency and at a lower cost.
- ii Practical FD-MIMO systems need to operate within 10° RMS phase error and 0.5 dB magnitude error to perform within 15% of performance with ideal hardware.
- iii Generating coherent LOs across RF chains is critical, especially when the duration between training and data transmission is large. SLO and PLL based architectures pose implementation challenges requiring frequent re-calibration due to short term phase drift.
- iv Internal temperature variation is a major contributor to phase and magnitude errors in an AAS. It is important to characterize the effect of temperature-related fluctuations to determine the impact on performance from residual errors. Frequent calibration can take away time-frequency resources from data transmission. Hence, calibration methods must be designed carefully.

The proposed AAS was lab and field verified to showcase the coherency achieved and its practical viability for FD-MIMO. For future work we propose to extend the analysis and field evaluation to 5G-NR air-interface, MU-MIMO performance with FD-MIMO and synchronization of multiple adjacent AAS radio units for shared cell or distributed-MIMO application.

REFERENCES

- [1] H. Ji, Y. Kim, J. Lee, E. Onggosanusi, Y. Nam, J. Zhang, B. Lee, and B. Shim, "Overview of full-dimension MIMO in LTE-advanced pro," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 176–184, Feb. 2017.
- [2] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.
- [3] W. Honcharenko, "Sub-6 GHz mMIMO base stations meet 5G's size and weight challenges," *Microw. J.*, vol. 62, no. 2, pp. 1–5, Feb. 2019.
- [4] L. M. P. Larsen, A. Checko, and H. L. Christiansen, "A survey of the functional splits proposed for 5G mobile crosshaul networks," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 146–172, 1st Quart., 2019.
- [5] E. Bjornson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015.
- [6] A. Papazafeiropoulos and T. Ratnarajah, "Toward a realistic assessment of multiple antenna HCNs: Residual additive transceiver hardware impairments and channel aging," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9061–9073, Oct. 2017.
- [7] D. Leibholz, "Quad-channel, wideband RF transceiver platform, 200 MHz bandwidth integrated radio transceiver solution," Analog Devices, Norwood, MA, USA, Tech. Rep. ADRV9026, Jun. 2019.
- [8] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical many-antenna base stations," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw. (Mobicom)*, 2012, pp. 53–64.
- [9] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liu, I. Wong, V. Owall, O. Edfors, and F. Tufvesson, "A flexible 100-antenna testbed for massive MIMO," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 287–293.
- [10] T. Wirth, L. Thiele, M. Kurras, M. Mehlhose, and T. Haustein, "Massive MIMO proof-of-concept: Emulations and hardware-field trials at 3.5 GHz," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 1793–1798.
- [11] W. Boukley Hasan, P. Harris, A. Doufexi, and M. Beach, "Real-time maximum spectral efficiency for massive MIMO and its limits," *IEEE Access*, vol. 6, pp. 46122–46133, Aug. 2018.
- [12] L. Dong, H. Zhao, Y. Chen, D. Chen, T. Wang, L. Lu, B. Zhang, L. Hu, L. Gu, B. Li, H. Yang, H. Shen, T. Tian, Z. Luo, and K. Wei, "Introduction on IMT-2020 5G trials in China," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1849–1866, Aug. 2017.
- [13] J. Wang, A. Jin, D. Shi, L. Wang, H. Shen, D. Wu, L. Hu, L. Gu, L. Lu, Y. Chen, J. Wang, Y. Saito, A. Benjebbour, and Y. Kishiyama, "Spectral efficiency improvement with 5G technologies: Results from field tests," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1867–1875, Aug. 2017.
- [14] B. Halvarsson, A. Simonsson, A. Elgcróna, R. Chana, P. Machado, and H. Asplund, "5G NR testbed 3.5 GHz coverage results," in *Proc. IEEE 87th Veh. Technol. Conf. (VTC Spring)*, Jun. 2018, pp. 1–5.
- [15] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, Sep. 2017.
- [16] E. Dahlman and S. Parkvall, "NR—The new 5G radio-access technology," in *Proc. IEEE 87th Veh. Technol. Conf. (VTC Spring)*, Jun. 2018, pp. 1–6.
- [17] A. Pitarokoulis, S. K. Mohammed, and E. G. Larsson, "Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 711–723, Feb. 2015.
- [18] R. Krishnan, M. R. Khanzadi, N. Krishnan, Y. Wu, A. G. i Amat, T. Eriksson, and R. Schober, "Linear massive MIMO precoders in the presence of phase noise—A large-scale analysis," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3057–3071, May 2016.
- [19] A. Puglielli, G. LaCaille, A. M. Niknejad, G. Wright, B. Nikolic, and E. Alon, "Phase noise scaling and tracking in OFDM multi-user beamforming arrays," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.
- [20] J. Vieira, F. Rusek, O. Edfors, S. Malkowsky, L. Liu, and F. Tufvesson, "Reciprocity calibration for massive MIMO: Proposal, modeling, and validation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3042–3056, May 2017.
- [21] W. Zhang, H. Ren, C. Pan, M. Chen, R. C. de Lamare, B. Du, and J. Dai, "Large-scale antenna systems with UL/DL hardware mismatch: Achievable rates analysis and calibration," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1216–1229, Apr. 2015.
- [22] X. Luo, "Multiuser massive MIMO performance with calibration errors," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4521–4534, Jul. 2016.
- [23] D. Mi, M. Dianati, L. Zhang, S. Muhaidat, and R. Tafazolli, "Massive MIMO performance with imperfect channel reciprocity and channel estimation error," *IEEE Trans. Commun.*, vol. 65, no. 9, pp. 3734–3749, Sep. 2017.

- [24] O. Raeesi, A. Gokceoglu, Y. Zou, E. Bjornson, and M. Valkama, "Performance analysis of multi-user massive MIMO downlink under channel non-reciprocity and imperfect CSI," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2456–2471, Jun. 2018.
- [25] A. Minasian, S. Shahbazpanahi, and R. S. Adve, "Distributed massive MIMO systems with non-reciprocal channels: Impacts and robust beamforming," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5261–5277, Nov. 2018.
- [26] S. Sangodoyin, V. Kristem, C. U. Bas, M. Kaske, J. Lee, C. Schneider, G. Sommerkorn, C. J. Zhang, R. Thoma, and A. F. Molisch, "Cluster characterization of 3-D MIMO propagation channel in an urban macro-cellular environment," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5076–5091, Aug. 2018.
- [27] *Study on 3D Channel Model for LTE*, document TR 36.873, 3GPP, Jan. 2018.
- [28] J. McNeill, S. Razavi, K. Vedula, and D. Richard Brown, "Experimental characterization and modeling of low-cost oscillators for improved carrier phase synchronization," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (IMTC)*, May 2017, pp. 1–6.
- [29] N. Tyler, B. Allen, and H. Aghvami, "Adaptive antennas: The calibration problem," *IEEE Commun. Mag.*, vol. 42, no. 12, pp. 114–122, Dec. 2004.
- [30] I. Galton and C. Weltin-Wu, "Understanding phase error and jitter: Definitions, implications, simulations, and measurement," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 66, no. 1, pp. 1–19, Jan. 2019.
- [31] G. Tsoulos, J. McGehean, and M. Beach, "Space division multiple access (SDMA) field trials. 2. Calibration and linearity issues," *IEEE Proc.-Radar, Sonar Navigat.*, vol. 145, no. 1, pp. 79–84, Feb. 1998.
- [32] T. Brauner, R. Kung, R. Vogt, and W. Bachtold, "5–6 GHz low-noise active antenna array for multi-dimensional channel-sounding," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Sep. 2003, pp. 297–301.
- [33] N. Tyler, B. Allen, and A. H. Aghvami, "Calibration of smart antenna systems: Measurements and results," *IET Microw., Antennas Propag.*, vol. 1, no. 3, pp. 629–638, Jun. 2007.
- [34] X. Jiang and F. Kaltenberger, "Channel reciprocity calibration in TDD hybrid beamforming massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 422–431, Jun. 2018.
- [35] G. Xu, Y. Li, J. Yuan, R. Monroe, S. Rajagopal, S. Ramakrishna, Y. H. Nam, J.-Y. Seol, J. Kim, M. M. U. Gul, A. Aziz, and J. Zhang, "Full dimension MIMO (FD-MIMO): Demonstrating commercial feasibility," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1876–1886, Aug. 2017.
- [36] *Study on Elevation Beamforming/Full-Dimension (FD) MIMO for LTE (Release 13)*, document TR 36.897, 3GPP, Jun. 2015.
- [37] H. Sahlin, "Channel prediction for link adaptation in LTE uplink," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, Sep. 2012, pp. 1–5.
- [38] F. W. Vook, W. J. Hillery, E. Visotsky, J. Tan, X. Shao, and M. Enescu, "System level performance characteristics of sub-6GHz massive MIMO deployments with the 3GPP new radio," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Aug. 2018, pp. 1–5.
- [39] B. Mondal, T. Thomas, E. Visotsky, F. Vook, A. Ghosh, Y.-H. Nam, Y. Li, J. Zhang, M. Zhang, Q. Luo, Y. Kakishima, and K. Kitao, "3D channel model in 3GPP," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 16–23, Mar. 2015.
- [40] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.
- [41] *LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures (Release 15)*, document TR 36.213, 3GPP, 2018.
- [42] D. R. Brown, P. Bidigare, S. Dasgupta, and U. Madhow, "Receiver-coordinated zero-forcing distributed transmit nullforming," in *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Aug. 2012, pp. 269–272.
- [43] H. M. Aumann and F. G. Willwerth, "Phased array calibrations using measured element patterns," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. Dig.*, Jun. 1995, pp. 918–921.
- [44] L. Kuehnke, "Phased array calibration procedures based on measured element patterns," in *Proc. IEEE APC*, vol. 2, Apr. 2001, pp. 660–663.
- [45] A. Kumar, P. A. Koch, H. E. Baidoo-Williams, R. Mudumbai, and S. Dasgupta, "An empirical study of the statistics of phase drift of off-the-shelf oscillators for distributed MIMO applications," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DYSpan)*, Apr. 2014, pp. 350–353.
- [46] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1607–1616, Aug. 2007.
- [47] G. Trudgen, "Variance as applied to crystal oscillator," Rakon UK Ltd, Harlow, U.K., Tech. Rep. TR-07-2009, Jul. 2009.
- [48] C. Zucca and P. Tavella, "The clock model and its relationship with the allan and related variances," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 52, no. 2, pp. 289–296, Feb. 2005.
- [49] J. Vig, "Introduction to quartz frequency standards," Army Research Laboratory, Adelphi, MD, USA, Tech. Rep. SLCET-TR-92-1, Oct. 1992.
- [50] O. Werther and R. Minihold, "LTE system specifications and their impact on RF & base band circuits," Rohde & Schwarz, Munich, Germany, Tech. Rep. IMA221, Apr. 2013.
- [51] P. Delos, "System-level LO phase noise model for phased arrays with distributed phase-locked loops," Analog Devices, Norwood, MD, USA, Tech. Rep. TA20865-0-11/18, Nov. 2018.
- [52] D. Hall, A. Hinde, and Y. Jia, "Multi-channel phase-coherent RF measurement system architectures and performance considerations," in *Proc. IEEE Mil. Commun. Conf.*, Oct. 2014, pp. 1318–1323.
- [53] C. Troster-Schmid and T. Bednorz, "Generating multiple phase coherent signals—Aligned in phase and time," Rohde & Schwarz, Munich, Germany, Tech. Rep. IGP108_1E, Sep. 2016.
- [54] Keysight, "Signal source solutions for coherent and phase stable multi-channel systems," Keysight, Santa Rosa, CA, USA, Tech. Rep. 5990-5442EN, Jul. 2014.
- [55] C. Pearson, "Clock skew in large multi-GHz clock trees," Analog Devices, Norwood, MA, USA, Tech. Rep. Analog Dialogue 53-01, Jan. 2019.
- [56] R. M. Vaghefi, G. Miranda, R. Srirambhatla, G. Marzin, C. Ng, F. Fayazbakhsh, S. Tarigopula, R. C. Palat, and M. Banu, "First commercial hybrid massive MIMO system for sub-6GHz bands," in *Proc. IEEE 5G World Forum (5GWF)*, Jul. 2018, pp. 357–362.
- [57] *5G Wireless Fronthaul Requirements in a Passive Optical Network Context*, document ITU-T G.Supp66, Jan. 2020.
- [58] *On the Use of Channel Reciprocity for NR*, document R1-163239, 3GPP, Apr. 2016.
- [59] X. Jiang, A. Decurninge, K. Gopala, F. Kaltenberger, M. Guillaud, D. Slock, and L. Deneire, "A framework for over-the-air reciprocity calibration for TDD massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 9, pp. 5975–5990, Sep. 2018.
- [60] M. Banu and V. Prodanov, "Method and system for multi-point signal generation with phase synchronized local carriers," U.S. Patent 8 259 884, Sep. 4, 2012.
- [61] V. Prodanov and M. Banu, "GHz serial passive clock distribution in VLSI using bidirectional signaling," in *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2006, pp. 285–288.
- [62] S. Shahsavari, S. A. Hosseini, C. Ng, and E. Erkip, "Adaptive hybrid beamforming with massive phased arrays in macro-cellular networks," in *Proc. IEEE 5G World Forum (5GWF)*, Jul. 2018, pp. 221–226.



REZA MONIR VAGHEFI (Member, IEEE) received the M.S. degree in communication engineering from the Chalmers University of Technology, Gothenburg, Sweden, in 2011, and the Ph.D. degree in electrical engineering from Virginia Tech, Blacksburg, VA, in 2014. His research interests include wireless communications and signal processing, estimation theory, convex optimization, cooperative and noncooperative localization, and sensor synchronization. He is currently an Associate Editor of the *EURASIP Journal on Advances in Signal Processing* and *Wireless Communications and Mobile Computing*.



RAMESH CHEMBIL PALAT received the B.Tech. degree in electronics and communications engineering from the University of Calicut, India, in 1998, and the master's and Ph.D. degrees from Virginia Tech, in 2002 and 2006, respectively. He has worked as a Senior Researcher with the Nokia Research Center, Berkeley, and a Staff Engineer with Qualcomm Flarion Technologies. He is currently the Director of communications systems with Blue Danube Systems. His industry experiences include system analysis, prototype development and algorithm, and DSP firmware development for commercial wireless modems. His research interests include signal processing for massive MIMO, architecture exploration for software defined radios, and signal processing for wireless communications.



GIOVANNI MARZIN was born in Latisana, Italy, in 1983. He received the M.Sc. and Ph.D. degrees in electrical engineering from the Politecnico di Milano, Milan, Italy, in 2009 and 2013, respectively. During his Ph.D. degree, he has worked on all-digital RF transmitters for 4G applications. In 2013, he was a Postdoctoral Researcher with the Politecnico di Milano, continuing his research on high-efficiency frequency synthesizers. In January 2014, he joined Blue Danube Systems Inc.,

Warren, NJ, USA, where he is currently a Senior RFIC Designer. His research interest includes cellular transceivers for active antenna systems. He was a co-recipient of the Dimitris N. Chorafas Foundation 2013 Award.



KIRAN BASAVARAJU was born in Bengaluru, Karnataka, India, in 1990. He received the B.E. degree in electronics and communications engineering from Visvesvaraya Technological University, Karnataka, in 2012, and the M.S. degree in electrical engineering–RF and microwave engineering from The University of Texas at Dallas, Texas, USA, in 2015.

After graduation, he started his career at Blue Danube Systems, Inc., as a RF Design Engineer, in 2015. He was a part of the RF Research and Development Team. His primary focus was achieving best system calibration for phase coherency and system stability. Since February 2020, he has been working with the Motorola RF Research and Development Team, concentrating on development of next-generation mobile design, mainly on receiver and transmitter designs for current and future mobile communication technologies like LTE and 5G. He is currently working as a Senior RF Engineer with Motorola Mobility, Chicago, IL, USA. His research interests include development of RF/IF circuit boards, modules for cellular basestations with HDAAS technology, and develop computer controlled automated tests for performance characterization.



YIPING FENG received the B.S. degree in electrical engineering from Shanghai Jiaotong University, Shanghai, China, in 1998, the M.S. degree in electrical engineering from Rutgers University, Piscataway, NJ, in 2005, and the Ph.D. degree in electrical engineering from Columbia University, New York, in 2009. In 1998, he joined Lucent Technologies, China, as a System Engineer with the Optical Network Group. From 2000 to 2003, he was at Nortel Networks, China, as a Senior

Technical Specialist with the Optical and Broadband Group. From 2009 to 2013, he was with MHI consulting LLC, NJ, USA, as an RFIC Design Engineer. In 2013, he co-founded Blue Danube Systems, NJ, USA. He is currently the Director of IC development with Blue Danube Systems, where he is involved in design of active antenna array systems and custom massive MIMO transceiver ICs.



MIHAI BANU (Life Fellow, IEEE) received the Ph.D. degree in electrical engineering from Columbia University, in 1984. He was with Bell Laboratories, Murray Hill, NJ, from 1980 to 2001, as a Summer Intern, a Ph.D. Grant Recipient, a MTS, the Department Head, and the Research and Development Director. From 2001 to 2004, he served as the Research and Development Director of Agere Systems, a Lucent spinoff. His Bell Labs/Agere contributions include seminal

research work in analog and RF integrated circuits and leading advanced development in IC technology and RF systems for wireless. In 2005, he started MHI Consulting LLC and, from 2005 to 2013, provided consulting services and did research work with NSF funding developing new concepts for the design of super low-cost coherent massive MIMO radio arrays. He founded Blue Danube Systems Inc., to apply these results to radio access products for 4G, 5G, and beyond. In 2013, Blue Danube Systems received Series A funding from Sequoia Capital. Other investors in Series B and C include AT and T, Northgate, and Silver Lake. He currently serves with Blue Danube as the Director, the CTO, and the VP Research and Development. He received several IEEE and Lucent awards and is the author of many publications and patents.

...