IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 19, 2017, accepted May 11, 2017, date of publication May 18, 2017, date of current version June 28, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2705561

The World's First Real-Time Testbed for Massive MIMO: Design, Implementation, and Validation

STEFFEN MALKOWSKY[1], (Student Member, IEEE), JOÃO VIEIRA[1], (Student Member, IEEE), LIANG LIU[1], (Member, IEEE), PAUL HARRIS[2], (Student Member, IEEE), KARL NIEMAN[3], (Member, IEEE), NIKHIL KUNDARGI[3], (Member, IEEE), IAN C. WONG[3], (Senior Member, IEEE), FREDRIK TUFVESSON[1], (Fellow, IEEE), VIKTOR ÖWALL[1], (Member, IEEE), AND OVE EDFORS[1], (Senior Member, IEEE)

[1]Department of Electrical and Information Technology, Lund University, 223 63 Lund, Sweden
[2]Communication Systems and Networks Group, University of Bristol, Bristol, BS81UB, U.K.
[3]National Instruments Corporation, Austin, TX 78759, USA

Corresponding author: Steffen Malkowsky (steffen.malkowsky@eit.lth.se)

This work was supported in part by the Swedish Foundation for Strategic Research SSF, VR, in part by the Strategic Research Area ELLIIT, and in part by the EU Seventh Framework Programme (FP7/2007-2013) under Grant 619086 (MAMMOET).

ABSTRACT This paper sets up a framework for designing a massive multiple-input multiple-output (MIMO) testbed by investigating hardware (HW) and system-level requirements, such as processing complexity, duplexing mode, and frame structure. Taking these into account, a generic system and processing partitioning is proposed, which allows flexible scaling and processing distribution onto a multitude of physically separated devices. Based on the given HW constraints such as maximum number of links and maximum throughput for peer-to-peer interconnections combined with processing capabilities, the framework allows to evaluate modular HW components. To verify our design approach, we present the Lund University Massive MIMO testbed, which constitutes the first reconfigurable real-time HW platform for prototyping massive MIMO. Utilizing up to 100 base station antennas and more than 50 field programmable gate arrays, up to 12 user equipments are served on the same time/frequency resource using an LTE-like orthogonal frequency division multiplexing time-division duplex-based transmission scheme. Proof-of-concept tests with this system show that massive MIMO can simultaneously serve a multitude of users in a static indoor and static outdoor environment utilizing the same time/frequency resource.

INDEX TERMS 5G, system design, testbed, outdoor measurement, indoor measurement, software-defined radio, TDD.

I. INTRODUCTION

In massive MIMO (MaMi) an unconventionally high number of base station (BS) antennas (hundreds or even higher) is employed to serve *e.g.,* a factor of ten less user equipments (UEs). Due to the excess number of BS antennas, linear signal processing may be used to spatially focus energy with high precision, allowing to separate a multitude of UEs in the spatial domain while using the same time/frequency resource [1]. MaMi theory promises a variety of gains, *e.g.,* increase in spectral and energy efficiencies as compared with single antenna and traditional MU-MIMO systems [2], [3], thereby tackling the key challenges defined for 5G.

Although MaMi is a promising theoretical concept, further development requires prototype systems for proof-of-concept and performance evaluation under real-world conditions to identify any further challenges in practice. Because of its importance, both industry and academia are making efforts in building MaMi testbeds, including the Argos testbed with 96-antennas [4], Eurecom's 64-antenna long-term evolution (LTE) compatible testbed, Samsung's Full-Dimension (FD) MIMO testbed and Facebook's Project Aries. Nevertheless, publications systematically describing the design considerations and methodology of a MaMi testbed are missing and real-time real-scenario performance evaluation of MaMi systems using testbeds have not been reported yet. At Lund University, the first real-time MaMi testbed, the Lund University MaMi (LuMaMi) testbed, showing successful MaMi transmission on the up-link (UL), was built [5]. Ever since, many testbeds have been constructed based on identical HW utilizing the same generic design principle, *e.g.,* the MaMi testbeds at the University of Bristol [6], Norwegian University of Science and Technology in Trondheim and University of Leuven in Belgium. The LuMaMi testbed provides a fully reconfigurable platform for testing MaMi under

real-life conditions. To build a real-time MaMi testbed many challenges have to be coped with. For example, shuffling data from 100 or more antennas, processing large-scale matrices and synchronizing a huge number of physically separated devices. All this has to be managed while still ensuring an overall reconfigurability of the system allowing experimental hardware and software solutions to be tested rapidly.

This paper discusses how implementation challenges are addressed by first evaluating high-level HW and system requirements, and then setting up a generic framework to distribute the data shuffling and processing complexity in a MaMi system based on the given HW constraints for interconnection network and processing capabilities. Taking into account the framework and requirements, a suitable modular HW platform is selected and evaluated. Thereafter, a thorough description of the LuMaMi testbed is provided including system parameters, base-band processing features, synchronization scheme and other details. The LuMaMi testbed constitutes a flexible platform that supports prototyping of up to 100-antenna 20 MHz bandwidth MaMi, simultaneously serving 12 UEs in real-time using orthogonal frequency division multiplexing (OFDM) modulation in time-division duplex (TDD) transmission mode. Bit Error Rates (BERs) and constellations for real-time UL and downlink (DL) uncoded transmission in a static indoor and static outdoor scenario are presented. Our first real-life proof-of-concept measurement campaigns show, that MaMi is capable of serving up to 12 UEs in the same time/frequency resource even for high user density per unit area. The gathered results suggest a significant increase in spectral efficiency compared to traditional point-to-point MIMO systems. By building the LuMaMi testbed we now have a tool which supports accelerated design of algorithms [7] and their validation based on real measurement data, with the additional benefit of real-world verification of digital base-band solutions.

Our main contributions can be summarized as follows:

- We provide overall and thorough analysis for MaMi systems, especially from a signal processing perspective, and identify design requirements as well as considerations on building up a MaMi testbed.
- We propose signal processing breakdown and distribution strategy to master the tremendous computational complexity in a MaMi system and introduce general hardware architecture for a MaMi testbed.
- We present the world's first real-time 100-antenna MaMi testbed, built upon Software-Defined Radio (SDR) technology.
- We validate the MaMi concept and its spatial multiplexing capability in real-life scenarios (both indoor and outdoor) with over-the-air transmission and real-time processing.

II. MASSIVE MIMO BASICS

In this section, the basic key detection and precoding algorithms utilized in MaMi are presented. Implementation specific details required to apply these algorithms, such as channel state information (CSI) estimation, are discussed in Sec. V. A simplified model of a MaMi BS using M antennas while simultaneously serving K single antenna UEs in TDD operation in a propagation channel $\boldsymbol{B}$ is shown in Fig. 1. To simplify notation, this discussion assumes a base-band equivalent channel and expressions are given per subcarrier, with subcarrier indexing suppressed throughout.

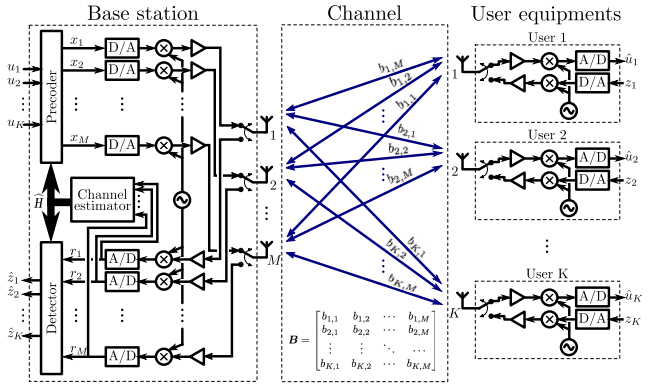


Fig. 1. A MaMi system model. Each antenna at the BS (left side) transmits a linear combination of K user-intended data symbols $u_k{}_{k=1}^{K}$. After propagation through the DL wireless channel B, each user antenna receives a linear combination of the signals transmitted by the M BS antennas. Finally, each of the K users, say user k, produces an estimate of its own intended data symbol, *i.e.*, u_k. Similar operation is employed for UL data transmission. Here, reciprocity for the propagation channel is assumed, *i.e.*, $B = B^{\mathsf{T}}$.

A. UP-LINK

The UL power levels used by the K UEs during transmission build the $K \times K$ diagonal matrix $\boldsymbol{P}_{\text{ul}}$. By collecting the transmitted UE symbols in a vector $\boldsymbol{z} \triangleq (z_1, \ldots, z_K)^{\mathsf{T}}$, the received signals $\boldsymbol{r} \triangleq (r_1, \ldots, r_M)^{\mathsf{T}}$ at the BS are described as

$$\boldsymbol{r} = \boldsymbol{G}\sqrt{\boldsymbol{P}_{\text{ul}}}\boldsymbol{z} + \boldsymbol{w}, \tag{1}$$

where $\boldsymbol{G}$ is the $M \times K$ UL channel matrix,[1] $\sqrt{\boldsymbol{P}_{\text{ul}}}$ an elementwise square-root, and $\boldsymbol{w} \sim \mathcal{CN}(0, \mathbf{I}_M)$ is independent and identically distributed (iid) circularly-symmetric zero-mean complex Gaussian noise. The estimated user symbols $\widehat{\boldsymbol{z}} \triangleq (\hat{z_1}, \ldots, \hat{z_K})^{\mathsf{T}}$ from the K UEs are obtained by linear filtering of the received vector $\boldsymbol{r}$ as

$$\widehat{\boldsymbol{z}} = f_{\text{eq}}(\boldsymbol{G})\boldsymbol{r}, \tag{2}$$

where $f_{\text{eq}}(\cdot)$ constructs an appropriate equalization matrix.

B. DOWN-LINK

On the DL, each UE receives its corresponding symbol $\hat{u}_k$ which are collected in a vector $\widehat{\boldsymbol{u}} \triangleq (\hat{u}_1, \ldots, \hat{u}_K)^{\mathsf{T}}$, representing the symbols received by all UEs. With this notation, the received signal becomes

$$\widehat{\boldsymbol{u}} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{w}' \tag{3}$$

[1] $\boldsymbol{G}$ is the up-link radio channel capturing both, the propagation channel $\boldsymbol{B}^{\mathsf{T}}$ and the up-link hardware transfer functions.

TABLE 1. Linear precoding/detection matrices.

	MRT/MRC	ZF	RZF
DL	$\boldsymbol{G}^*$	$\boldsymbol{G}^*(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{G})^{-\mathsf{T}}$	$\boldsymbol{G}^*(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{G}+\beta_{\mathrm{reg}_{\mathrm{pre}}}\boldsymbol{I}_K)^{-\mathsf{T}}$
UL	$\boldsymbol{G}^{\mathsf{H}}$	$(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{G})^{-1}\boldsymbol{G}^{\mathsf{H}}$	$(\boldsymbol{G}^{\mathsf{H}}\boldsymbol{G}+\beta_{\mathrm{reg}_{\mathrm{dec}}}\mathbf{I}_K)^{-1}\boldsymbol{G}^{\mathsf{H}}$

where the $K \times M$ matrix $\boldsymbol{H}$ is the DL radio channel,[2] $\boldsymbol{w}' \sim \mathcal{CN}(0, \mathbf{I}_K)$ is an iid circularly-symmetric zero-mean complex Gaussian receive noise vector with covariance matrix $\mathbf{I}_K$, and $\boldsymbol{x} \triangleq (x_1, \ldots, x_M)^{\mathsf{T}}$ is the transmit vector.

As explicit DL channel estimation is very resource consuming, it is not considered practical in a MaMi setup [1]. Taking into account that the propagation channel $\boldsymbol{B}$ is generally agreed on to be reciprocal [7], the estimated UL channel matrix $\boldsymbol{G}$ can be utilized to transmit on the DL. However, differences due to analog circuitry in the UL and DL channels, $\boldsymbol{G}$ and $\boldsymbol{H}$, need to be compensated. Thus, a possible construction for $\boldsymbol{x}$ is of the form

$$\boldsymbol{x} = f_{\mathrm{cal}}(f_{\mathrm{pre}}(\boldsymbol{G}))\boldsymbol{u}, \tag{4}$$

where $\boldsymbol{u} \triangleq (u_1, \ldots, u_K)^{\mathsf{T}}$ is a vector containing the symbols intended for the K UEs, $f_{\mathrm{pre}}(\cdot)$ is some precoding function, and $f_{\mathrm{cal}}(\cdot)$ is a reciprocity calibration function to be discussed next.

C. RECIPROCITY CALIBRATION

In most practical systems, the UL and DL channels are not reciprocal, i.e. $\boldsymbol{G} \neq \boldsymbol{H}^T$. This is easily seen by factorizing $\boldsymbol{G}$ and $\boldsymbol{H}$ as

$$\boldsymbol{G} = \boldsymbol{R}_{\mathrm{B}}\boldsymbol{B}^{\mathsf{T}}\boldsymbol{T}_{\mathrm{U}}, \quad \text{and } \boldsymbol{H} = \boldsymbol{R}_{\mathrm{U}}\boldsymbol{B}\boldsymbol{T}_{\mathrm{B}}, \tag{5}$$

where the two $M \times M$ and $K \times K$ diagonal matrices $\boldsymbol{R}_{\mathrm{B}}$ and $\boldsymbol{R}_{\mathrm{U}}$ model the non-reciprocal hardware responses of BS and UE receivers (RXs), respectively, and the two $M \times M$ and $K \times K$ diagonal matrices $\boldsymbol{T}_{\mathrm{B}}$ and $\boldsymbol{T}_{\mathrm{U}}$ similarly model hardware responses of their transmitters (TXs). Thus, in order to construct a precoder based on the UL channel estimates, the non-reciprocal components of the channel have to be calibrated. Previous calibration work showed that this is possible by using

$$\boldsymbol{C}f_{\mathrm{pre}}(\boldsymbol{G}) = f_{\mathrm{cal}}(f_{\mathrm{pre}}(\boldsymbol{G})), \tag{6}$$

where $\boldsymbol{C} = \boldsymbol{R}_{\mathrm{B}}\boldsymbol{T}_{\mathrm{B}}^{-1}$ is the, so-called, calibration matrix which can be estimated internally at the BS [7]. Such calibration is sufficient to cancel inter-user interference stemming from non-reciprocity [8].

D. LINEAR DETECTION & PRECODING SCHEMES

Table 1 shows a selection of weighting matrices used in linear precoding and detection schemes, with non-reciprocity compensation included in the form of the $M \times M$ diagonal matrix $\boldsymbol{C}$ as defined above. The maximum ratio transmission (MRT) precoder and the maximum ratio combining (MRC) decoder maximize array gain without active suppression of interference among the UEs [1]. The zero-forcing (ZF) precoder and ZF combiner employ the pseudo-inverse, which provides inter-user interference suppression with the penalty of lowering the achievable array gain. A scheme that allows trade-off between array gain and interference suppression is the regularized ZF (RZF) precoder and RZF combiner. This is achieved by properly selecting the regularization constants $\beta_{\mathrm{reg}_{\mathrm{pre}}}$ and $\beta_{\mathrm{reg}_{\mathrm{dec}}}$. If $\beta_{\mathrm{reg}_{\mathrm{pre}}}$ and $\beta_{\mathrm{reg}_{\mathrm{dec}}}$ are selected to minimize mean-square error (MSE) $E\|\boldsymbol{u} - \frac{1}{\sqrt{\rho}}\hat{\boldsymbol{u}}\|^2$, where ρ is a scaling constant, we obtain the minimum MSE (MMSE) precoder/detector [9].

[2] $\boldsymbol{H}$ is the down-link radio channel capturing both, the propagation channel $\boldsymbol{B}$ and the down-link hardware transfer functions.

III. SYSTEM DESIGN ASPECTS

Having discussed the MaMi basics, we move on to system design aspects. These include modulation scheme, frame structure and hardware requirements.

A. MODULATION SCHEME

While many different modulation schemes can be used with MaMi, this paper focuses on OFDM, employed in many modern wireless communication systems. Properly designed OFDM renders frequency-flat narrowband subcarriers, facilitating the single channel equalization strategy used here.

For ease of comparison and simplicity, LTE-like OFDM parameters, as shown in Table 2, are used throughout this discussion. The more common parameters with LTE, the easier it is to evaluate how MaMi as an add-on would influence current cellular systems.

TABLE 2. High-level system parameters.

Parameter	Variable	Value
Bandwidth	W	20 MHz
Sampling Rate	F_{s}	30.72 MS/s
FFT Size	N_{FFT}	2048
# Used subcarriers	N_{used}	1200
Cyclic prefix	N_{cp}	144 samples
OFDM symbol length	t_{OFDM}	71.4 µs

B. TDD VERSUS FDD

Current cellular systems either operate in frequency-division duplex (FDD) or TDD mode. FDD is, however, considered impractical for MaMi due to excessive resources needed for DL pilots and CSI feedback. TDD operation relying on reciprocity only requires orthogonal pilots in the UL from the K UEs, making it the feasible choice [10]. For this reason, we focus entirely on TDD below.

C. RECIPROCITY

To allow operation in TDD mode, differences in the TX and RX transfer functions on both, the BS and UEs have to be calibrated as discussed in Sec. II-C. Drifts over time are mainly caused by HW temperature and voltage changes,

and thus, the calibration interval depends on the operating environment of the BS.

D. FRAME STRUCTURE

The frame structure defines among other things, the pilot rate which determines how well channel variations can be tracked and, indirectly, the largest supported UE speed.

1) MOBILITY

The maximum supportable mobility, *e.g.,* the maximum speed of the UEs is defined by the UL pilot transmission interval. In order to determine this constraint, a 2D wide-sense stationary channel with uncorrelated isotropic scattering is assumed. For the contributions from the different BS antennas to add up coherently high channel correlation is required and, as an approximation to formulate the final requirement, a correlation of 0.9 was used to ensure sufficient channel coherency. Further discussions on such modeling assumption are found in [11]. Although these assumptions may not be completely valid for MaMi channels, they allow an initial evaluation based on a maximum supported Doppler frequency, $\nu_{\max}$, by solving

$$J_0(2\pi\,\nu_{\max} T_{\mathrm{p}}) = 0.9, \tag{7}$$

for $\nu_{\max}$, where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, stemming from a standard Jakes' fading assumption, and T_{p} the distance between pilots in time. Hence, the maximum supportable speed of any UE may be evaluated using

$$\mathrm{v}_{\max} = \frac{c\nu_{\max}}{f_{\mathrm{c}}}, \tag{8}$$

once a specific frame structure is provided. In (8) $\mathrm{v}_{\max}$ is the maximum supported speed of a UE, c the speed of light and f_c the chosen carrier frequency.

2) PROCESSING LATENCY

The frame structure has to be designed for the highest speed of UEs to be supported which requires a high pilot rate for high mobility scenarios. Within two consecutive UL pilot symbols, all UL data, DL data and guard symbols have to be accommodated which in turn decreases the available time between UL pilot reception and DL transmission. In a high mobility scenario this poses tight latency requirements for TDD transmission as CSI has to be estimated in order to produce the precoding matrix to beamform the DL data.

To formulate the TDD precoder turnaround time, Δ, all HW units introducing a delay must be taken into account. This includes the analog front-end delays for the TX $\Delta^{\mathrm{rf,TX}}$ and RX $\Delta^{\mathrm{rf,RX}}$, the processing latency for OFDM modulation/demodulation (including cyclic prefix (CP) and guard band operation) Δ^{OFDM}, the time for processing UL pilots to estimate CSI Δ^{CSI}, and the processing latency for precoding $\Delta^{\mathrm{precode}}$ including reciprocity compensation. Additional sources of latency include overhead in data routing, packing, and unpacking, *i.e.,* Δ^{rout} such that the overall TDD precoder turnaround time may be formulated as

$$\Delta = \Delta^{\mathrm{rf,TX}} + \Delta^{\mathrm{rf,RX}} + \Delta^{\mathrm{OFDM}} + \Delta^{\mathrm{CSI}} + \Delta^{\mathrm{precode}} + \Delta^{\mathrm{rout}}. \tag{9}$$

Depending on the specific arrangement of the OFDM symbols and the pilot repetition pattern in the frame structure, base-band processing solutions, especially Δ^{CSI} and $\Delta^{\mathrm{precode}}$, have to be optimized to not violate the given constraint, *i.e.,* Δ.

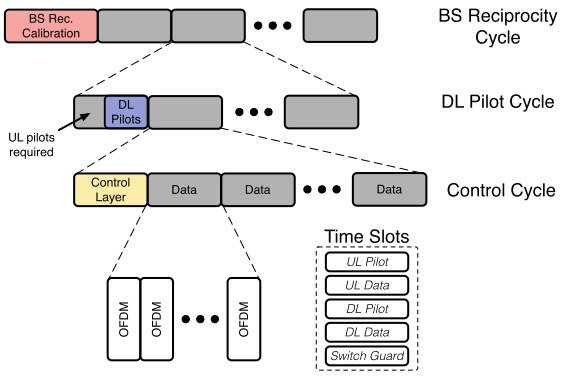


Fig. 2. Generic frame structure of a LTE like TDD-based MaMi system. Within one BS reciprocity cycle the BS operates using the same reciprocity calibration coefficients. A certain number of DL pilot cycles are integrated as UEs suffer from faster changing environments. Each control cycle contains a control layer to perform, for example over-the-air synchronization and within these the data transmission slots are encapsulated.

3) PILOT PATTERN

In general, to acquire CSI at the BS, the K UEs transmit orthogonal pilots on the UL. Different approaches are, *e.g.,* distributed pilots over orthogonal subcarriers [12] or sending orthogonal pilot sequences over multiple subcarriers [13]–[15] but also semi-blind and blind techniques have been proposed [16].

Fig. 2 shows a generic frame structure capturing the aforementioned aspects in a hierarchical manner assuming all UEs transmit their pilots within one dedicated pilot symbol. At the beginning of each BS reciprocity cycle, reciprocity calibration at the BS is performed and within these a certain number of DL pilot cycles are encapsulated where precoded DL pilot symbols are transmitted. The length of the BS reciprocity cycle is determined by the stability of the transceiver chains in the BS. As the reciprocity calibration at the BS side only compensates for BS transceivers, DL pilots are necessary to compensate for transceiver differences at the UE side. Their frequency depends on the stability at the UE side and can be considered significantly smaller than for the BS as UEs are subject to faster changes in their operational environment, *e.g.,* thermal differences when having the UE in a pocket or using it indoors or outdoors. To be able to send precoded pilots on the DL, transmission of UL pilots is required beforehand. Several control cycles are embedded inside each DL pilot cycle carrying a certain number of data time slots. Time

slots contain five different OFDM symbol types for physical layer implementation. These are *(i)* UL Pilot where the UEs transmit orthogonal pilots to the BS, *(ii)* UL Data where all UEs simultaneously send data to the BS, *(iii)* DL Pilot where the BS sends precoded pilots to all UEs, *(iv)* DL Data where the BS transmits data to all UEs and *(v)* Switch Guard, which idles the RF chains to allow switching from RX to TX or vice versa.

E. HARDWARE REQUIREMENTS

To illustrate the required HW capabilities for the testbed, the values from Table 2 are used to estimate the Gops/s[3] and the data shuffling on a per OFDM symbol basis for the general case and a specific case assuming $M = 100$ and $K = 12$.

1) PROCESSING CAPABILITES

Table 3 summarizes the overall number of real-valued arithmetic operations. For the processing estimates, it is assumed that each complex multiplication requires four real multiplications. Close to the antennas, M fast-Fourier transforms (FFTs) or inverse FFTs (IFFTs) are needed equating to 126 Gops/s. Data precoding and detection as well as reciprocity compensation require large matrix and vector multiplications, for instance, an $M \times K$ matrix with a $K \times 1$ vector leading to up to 80 Gops/s.

TABLE 3. Processing requirements in a MaMi system.

Function	General	Specific
	Gops/s	Gops/s
FFT/IFFT	$4M \log_2(N_{\text{FFT}}) N_{\text{FFT}} / t_{\text{OFDM}}$	126
Detection	$4MKN_{\text{used}} / t_{\text{OFDM}}$	80
Precoding	$4MKN_{\text{used}} / t_{\text{OFDM}}$	80
Recip. Cal.	$4MKN_{\text{used}} / t_{\text{OFDM}}$	80
Pseudo-inv.	$4N_{\text{used}} \left(2MK^2 + K^3\right) / \left(2t_{\text{OFDM}}\right)$	1080

Finally, when using ZF, the pseudo-inverse matrix is required which includes the calculation of the Gram matrix requiring MK^2 multiplications with the $K \times K$ matrix inversion adding another K^3 in complexity assuming a Neumann-Series approximation [17] or a QR decomposition. The last multiplication of the inverse with the Hermitian of the channel matrix $\boldsymbol{H}$ needs another MK^2 multiplications which combined with a requirement of finishing within two OFDM symbols leads to approximately 1 Tops/s for the overall pseudo-inverse calculation.

2) DATA SHUFFLING CAPABILITIES

Table 4 summarizes required interconnect bandwidth and number of links. Communication paths to each antenna transfer at the sampling rate of $F_{\text{s}} = 30.72$ MS/s which is decreased to the subcarrier rate $F_{\text{sub}} = 16.8$ MB/s by performing OFDM processing $\left(F_{\text{s}} \cdot N_{\text{used}} / (N_{\text{FFT}} + N_{\text{cp}})\right)$.

[3]Gops/s is used here, but these can be seen as GMACs/s, *i.e.,* the number of multiply-accumulate operations, as almost all operations involve matrix-matrix and matrix-vector calculations.

TABLE 4. Data shuffling requirements in a MaMi system.

Purpose	General	Specific
	#	#
Links to cent. proc	$2M$	200
	MB/s	MB/s
Antenna Rate	$w_{\text{ant}} M F_{\text{s}}$	w_{ant} 3,072
Subcarrier Rate	wMF_{sub}	w 1,680
Information rate	$K \cdot F_{\text{sub}}$	201.6

Considering M antennas, the overall subcarrier data rate is $M \cdot w \cdot 16.8$ MB/s, with w being the combined wordlength for the in-phase and quadrature components in bytes. The information rate in an OFDM symbol carrying data is $K \cdot 16.8$ MB/s assuming 8 bit per sample, *i.e.,* 256−QAM as highest modulation. Assuming separate links between centralized processing and the antenna units on UL and DL, $2M$ peer-to-peer (P2P) links[4] are needed between the antennas and the centralized MIMO processing.

3) RECONFIGURABILITY

The testbed has to be reconfigurable and scalable, to support different system parameters, different processing algorithms and adaptive processing. It is also crucial to have the possibility to integrate in-house developed HW designs for validation and performance comparison of algorithms. Variable center frequencies, run-time adjustable RX and TX gains as well as configurable sampling rates are highly desirable to be able to adapt to other parameters than the ones presented in Table 2.

IV. GENERIC HARDWARE AND PROCESSING PARTITIONING

In this section a generic HW and processing partitioning is presented to explore the parallelism in MaMi, which needs consideration of processing together with data transfer requirements (throughput, latency, # of P2P links), and at the same time provides scalability.

A. HIERARCHICAL OVERVIEW

To be able to build a MaMi testbed with modular HW components, a hierarchical distribution as shown in Fig. 3 is proposed. The main blocks are detailed as follows:

1) SDR

SDRs provide the interface between the digital and radio-frequency (RF) domain as well as local processing capabilities.

2) SWITCHES

Switches aggregate/disaggregate data between different parts of the system, *e.g.,* between SDRs and the co-processors.

[4]In this discussion, each interconnection transferring data between physically separated devices is denoted a peer-to-peer (P2P) link.

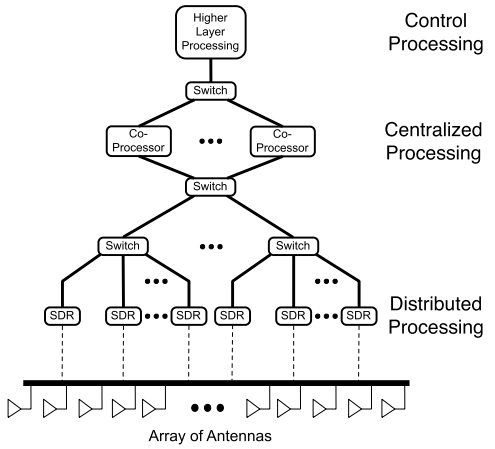


Fig. 3. Hierarchical overview of a MaMi BS built from modular HW components.

3) CO-PROCESSING MODULES

Co-processing modules provide a centralized node to perform MIMO processing.

4) HIGHER LAYER PROCESSING

Higher layer processing controls the system, configures the radios, and provides run-time status metrics of the system.

B. PROCESSING AND DATA DISTRIBUTION

For proper base-band processing partitioning, throughput constraints of HW components have to be taken into account. Assuming each SDR supports n_{ant} antennas, the required number of SDRs becomes $\lceil M/n_{\text{ant}} \rceil$ for an M-antenna system.

1) SUBSYSTEMS

As shown in Fig. 4, RF-Front End, OFDM processing and reciprocity compensation are performed on a per-antenna basis using the SDRs. This distributes a large fraction of the overall processing and reduces the data rate before transferring the acquired samples over the bus. Still, the number of direct devices on a bus is limited, and thus, setting up $2M$ P2P links directly to the co-processors would most likely exceed the number of maximum P2P links for any reasonable number of MaMi antennas. To reduce this number, data can be aggregated using the concept of grouping. The different data streams from several SDRs are interleaved on one common SDR and then sent via one P2P link. Therefore, subsystems are defined, each containing n_{sub} SDRs. Data from all antennas within a subsystem is aggregated/disaggregated on the outer two SDRs and distributed to the n_{co} co-processors using high-speed routers.

At closer look, Fig. 4 reveals that the SDRs on the outer edges which realize the $(n_{\text{ant}}n_{\text{sub}})$ to (n_{co}) and (n_{co}) to $(n_{\text{ant}}n_{\text{sub}})$ router functionalities, require the highest number of P2P links, and thus have to deliver the highest throughput. Hence, the following inequalities have to be fulfilled for the subsystems not to exceed the constraints for maximum number of P2P links ($\text{P2P}_{\text{SDR,max}}$) and maximum

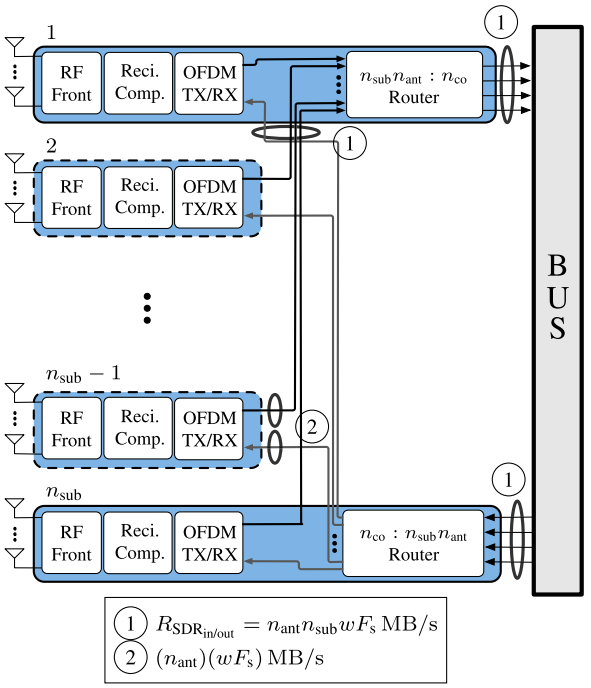


Fig. 4. A subsystem consisting of n_{sub} SDRs where the two outer SDRs implement an antenna combiner / BW splitter and an antenna splitter / BW combiner, both implemented using high-speed FPGAs routers. Inter-SDR and SDR to central processor connections utilize a bus for transferring the samples.

bidirectional throughput ($R_{\text{SDR}_{\text{max}}}$):

$$R_{\text{SDR}_{\text{max}}} > R_{\text{SDR}_{\text{out}}} = R_{\text{SDR}_{\text{in}}} = n_{\text{ant}} \cdot n_{\text{sub}} \cdot w \cdot F_{\text{sub}} \quad (10)$$

$$\text{P2P}_{\text{SDR,max}} > \text{P2P}_{\text{SDR}} = n_{\text{co}} + n_{\text{sub}} \quad (11)$$

where it is assumed that if an SDR employs more than one antenna, the data is interleaved before it is sent to the router on the outer SDRs. The constraints given in equation (10)-(11) can be used to determine the maximum number of SDRs per subsystem (n_{sub}) such that hardware constraints are not exceeded.

2) CO-PROCESSORS

As shown in Fig. 5, detection, precoding, CSI acquisition, symbol mapping and symbol demapping are integrated in the centrally localized co-processor modules which collect data from all SDRs. Using CSI estimated from UL pilots, MIMO processing as discussed in Sec. II and symbol mapping/de-mapping is performed.

Based on the selected OFDM modulation scheme the subcarrier independence can be exploited allowing each of the n_{co} co-processors to work on a sub-band of the overall 20 MHz bandwidth. This efficiently circumvents issues with throughput and latency constraints in the MIMO signal processing chain. The co-processors aggregate/disaggregate data from all the antennas in the system using reconfigurable high-speed routers, as shown in Fig. 5 for a system having $\lceil M/(n_{\text{sub}}n_{\text{ant}}) \rceil$ subsystems and n_{co} co-processors.

Similarly to the SDRs, the two main constraints for the co-processors are the maximum number of P2P links denoted $\text{P2P}_{\text{CO,max}}$ and the maximum throughput denoted $R_{\text{CO}_{\text{max}}}$.

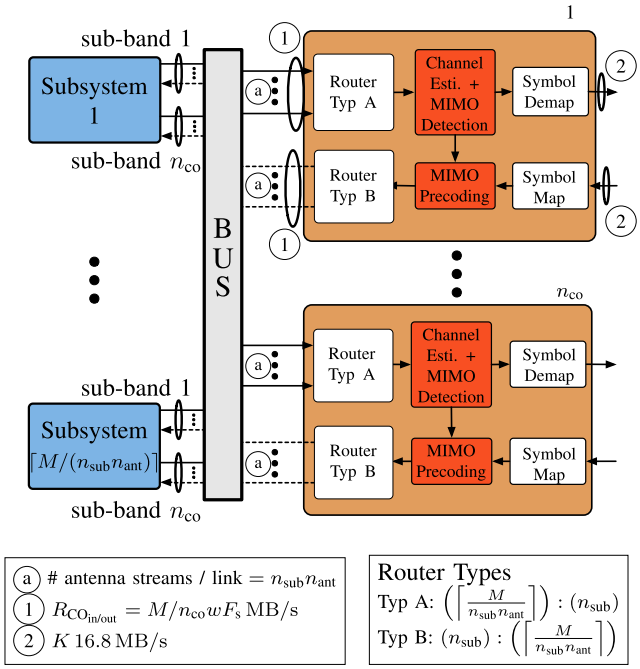


Fig. 5. Shuffling data from the $\lceil M/(n_{sub}n_{ant})\rceil$ subsystems to the n_{co} co-processors. The routers use a simple round robin scheme to combine/distribute the data from/to corresponding subsystems.

The following inequalities have to hold for the co-processor not to exceed these constraints:

$$R_{CO_{max}} > R_{CO_{out}} = R_{CO_{in}} = \left(\frac{M \cdot w + K}{n_{co}}\right) \cdot F_{sub} \quad (12)$$

$$P2P_{CO,max} > P2P_{CO} = 2 \cdot \lceil M/n_{sub} \rceil + 2. \quad (13)$$

Using this modular and generic system partitioning, HW platforms built using modular components can be evaluated. Note, that expressions (10) - (13) may also be used with other system parameters, *e.g.,* by redefining F_s and F_{sub}.

V. LuMaMi TESTBED IMPLEMENTATION

In this section the LuMaMi specific implementation details are discussed based on the aforementioned general architecture. The LuMaMi system was designed with 100 BS antennas and can serve up to 12 UEs simultaneously. Based on these parameters, the selected modular HW platform is presented and given constraints are evaluated. Consequently, the specific frame structure and other features of the system including base-band processing, antenna array, mechanical structure and synchronization are briefly described. Before providing details, the authors would like to emphasize, that this is the initial version of the LuMaMi testbed and that add-ons and further improvements are planned for the future.

A. SELECTED HARDWARE PLATFORM

The hardware platform was selected based on requirements discussed in Sec. III. Table 5 shows the selected off-the-shelf modular hardware from National Instruments used to implement the LuMaMi testbed. The SDRs [18] allow up to 15 P2P links ($P2P_{SDR,max} = 15$) with a bidirectional throughput of $R_{SDR_{max}} = 830$ MB/s, support a variable center frequency from 1.2 GHz to 6 GHz and have a TX power of 15 dBm. Each SDR contains two RF chains, *i.e.,* $n_{ant} = 2$, and a Kintex-7 FPGA. Selected co-processors [19] allow a bidirectional P2P rate of $R_{CO_{max}} = 2.4$ GB/s with up to $P2P_{CO,max} = 32$ P2P links and employ a powerful Kintex-7 FPGA with a reported performance of up to 2.845 GMACs/s [20]. This is sufficient for a 100 BS antenna MaMi testbed due to the fact that n_{co} co-processors can be utilized in parallel. Interconnection among devices is achieved using 18-slot chassis [21] combined with per-slot expansion modules [22]. Each chassis integrates two switches based on Peripheral Component Interconnect Express (PCIe) using direct memory access (DMA) channels which allow inter-chassis traffic up to 7 GB/s and intra-chassis traffic up to 3.2 GB/s.

The host [23] is an integrated controller, running LabVIEW on a standard Windows operating system and is used to configure and control the system. The integrated hardware/software stack provided by LabVIEW provides the needed reconfigurability as it abstracts the P2P link setup, communication among all devices and allows FPGA programming as well as host processing using a single programming language. An additional feature of LabVIEW is the possibility to seamlessly integrate intellectual property (IP) blocks generated via Xilinx Vivado platform paving a way to test in-house developed IP.

To be able to synchronize the full BS, a Reference Clock Source [24] and Reference clock distribution network [25] are required. Their functionalities will be later discussed when presenting the overall synchronization method.

B. SUBSYSTEMS AND NUMBER OF CO-PROCESSORS

To build the LuMaMi testbed with $M = 100$ antennas, 50 SDRs are necessary. The maximum possible subsystem size is chosen to minimize the utilization of available P2P links at the co-processors. By using (10) and an internal fixed-point wordlength of $w = 3$ corresponding to a 12-bit resolution on the I- and Q-components, n_{sub} is found to be 8. As this is not an integer divider of 50, the last subsystem only contains two SDRs.

Based on Table 4, the combined subcarrier rate for all antennas is $wMF_{sub} = 5$ GB/s and another $K \cdot F_{sub} = 200$ MB/s are needed for information symbols. To not exceed $R_{CO_{max}}$ at least three co-processors must be utilized. To further lower the burden on the design of the low-latency MIMO signal processing chain, $n_{co} = 4$ is chosen such that each co-processor processes 300 of the overall 1200 subcarriers.

Table 6 summarizes the LuMaMi testbed parameters and shows that constraints are met according to (10)-(13). It can also be seen that the design is still within the constraints if scaling up the number of BS antennas to $M = 128$, which has been done in subsequent designs based on the same hardware, *e.g.,* [6].

TABLE 5. Selected hardware from national instruments.

Type	Model	Features
Host	PXIe-8135	2.3 GHz Quad-Core PXI Express Controller Up to 8 GB/s system and 4 GB/s slot bandwidth
SDR	USRP RIO 294xR / 295xR	2 RF Front Ends and 1 Xilinx Kintex-7 FPGA Center frequency variable from 1.2 GHz to 6 GHz 830 MB/s bidirectional throughput on up to 15 DMA channels
Co-Processor	FlexRIO 7976R	1 Xilinx Kintex-7 410T FPGA 2.4 GB/s bidirectional throughput on up to 32 DMA channels
Switch	PXIe-1085	Industrial form factor 18-slot chassis 7 GB/s bidirectional throughput per slot 2 switches per chassis with inter-switch traffic up to 3.2 GB/s Links between chassis bound to 7 GB/s bidirectional
Expansion Module	PXIe-8374	PXI Express (x4) Chassis Expansion Module Software-transparent link without programming Star, tree, or daisy-chain configuration
Reference Clock Source	PXIe-6674T	10 MHz reference clock source with < 5 ppb clock accuracy 6 configurable I/O connections
Ref. Clock Distribution	OctoClock	10 MHz 8-channel clock and timing distribution network

TABLE 6. System parameters and validation of constraints in the LuMaMi testbed.

Parameters		Rates MB/s
M	100	$R_{\text{SDR}_{\text{max}}} = 830 > R_{\text{SDR}_{\text{out}}} = R_{\text{SDR}_{\text{in}}} = 806.4$
K	12	$R_{\text{CO}_{\text{max}}} = 2,400 > R_{\text{CO}_{\text{out}}} = R_{\text{CO}_{\text{in}}} = 1,460$
n_{ant}	2	**P2P Links**
n_{sub}	8[a]	$\text{P2P}_{\text{SDR,max}} = 15 > \text{P2P}_{\text{SDR}} = 12$
n_{co}	4	$\text{P2P}_{\text{CO,max}} = 32 > \text{P2P}_{\text{CO}} = 18$

[a] Note, that the last subsystem only consists of two SDRs.

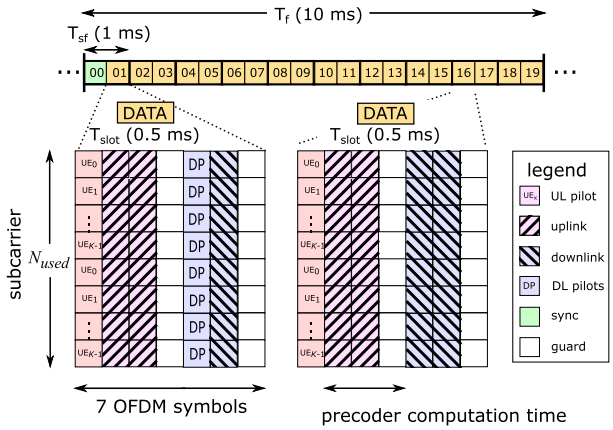


Fig. 6. The default frame structure used in the LuMaMi testbed.

C. FRAME STRUCTURE

The default frame structure for the LuMaMi testbed is shown in Fig. 6. One frame is $T_{\text{f}} = 10$ ms and is divided in ten subframes of length $T_{\text{sf}} = 1$ ms. Each subframe consists of two slots having length $T_{\text{slot}} = 0.5$ ms, where the first subframe is used for control signals, *e.g.,* to implement over-the-air synchronization, UL power control and other control signaling. The 18 slots in the other nine subframes encapsulate seven OFDM symbols each. Comparing to Fig. 2, a reciprocity calibration cycle is defined over the whole runtime of the BS for simplicity and due to the fact that there is no large drift after warming up the system in a controlled environment [5]. The DL pilot cycles and control cycles are both set to be the length of one frame. Each frame starts with one control subframe followed by one subframe with one DL pilot and one DL data symbol whereas all others use two DL data symbols.

D. MOBILITY

The pilot distance in time in the default frame structure given in Fig. 2 is $T_{\text{p}} \approx 430\,\mu$s or six OFDM symbols. Thus, $\nu_{\text{max}} \approx 240$ Hz for a correlation of 0.9. Due to availability from a network operator, a carrier frequency of $f_{\text{c}} = 3.7$ GHz is selected. Using (8), $\text{v}_{\text{max}} = 70$ km/h is found as maximum supported speed.

E. TDD TURNAROUND TIME

The pre-coding turnaround time requirement for the implementation can be analyzed based on (9). The analog front-end delay of the SDRs was measured to be about 2.25 µs. Taking the frame structure in Fig. 6 (assuming $\Delta^{\text{rf,TX}} = \Delta^{\text{rf,RX}}$ which is not necessarily true), the latency budget for base-band processing is as follows: Overall time for pre-coding after receiving the UL pilots is 214 µs (3 OFDM symbols). The 2048 point FFT/IFFT (assuming a clock frequency of 200 MHz) requires around 35 µs $\times$ 2 $=$ 70 µs in total for TX and RX (including sample reordering). As a result, the remaining time for channel estimation, MIMO processing, and data routing is around 140 µs, which is the design constraint for this specific frame structure.

An analysis of the implemented design showed that the latency is far below the requirement for the default

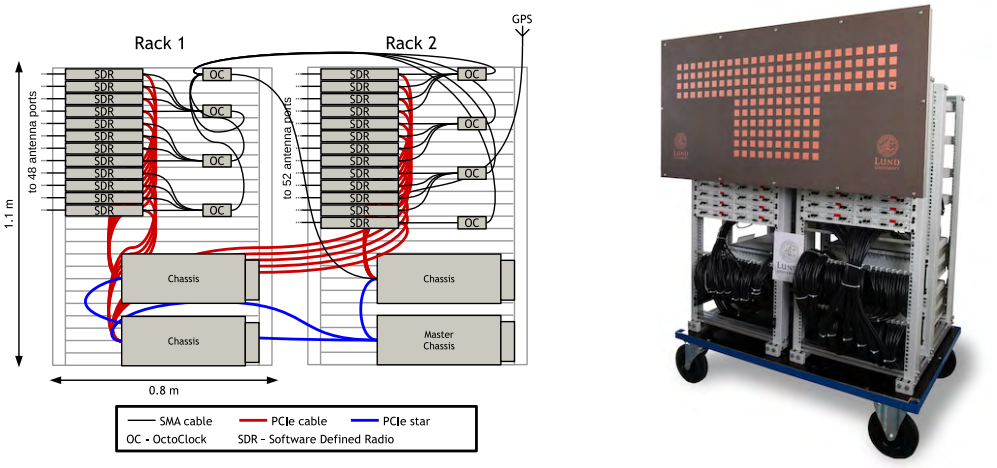


Fig. 7. Left: Side view of the mechanical assembly of the BS. The two racks sit side by side (not as shown) with the SDRs facing the same direction (towards the antenna array). Two columns of USRP SDRs are mounted in each rack, totaling 50 of them. Right: The assembled LuMaMi testbed at Lund University, Sweden.

frame structure which makes it possible to use the testbed for higher mobility scenarios from this point of view [26].

F. IMPLEMENTATION FEATURES

1) BASE-BAND PROCESSING

On the LuMaMi testbed, each UE sends pilots on orthogonal subcarriers, *i.e.,* each UE uses every K-th subcarrier with the first UE starting at subcarrier 0, the second at subcarrier 1 etc., overall utilizing a full OFDM symbol. It was shown that performance does not suffer significantly compared to a full detector calculated for each subcarrier using this method [12]. Moreover, it efficiently remedies processing requirements and reduces the required memory for storing estimated CSI matrices by a factor of K. A least-square CSI estimation algorithm with zeroth-order hold over $K = 12$ subcarriers was implemented, however, better estimates could be obtained by on-the-fly interpolation between the estimated subcarriers. Overall, utilizing this approach reduces the required detection matrix throughput to one matrix every 12 subcarriers, *i.e.,* 16.8×10^6 subcarriers/s/12 $= 1.4 \times 10^6$ Detection Matrices/s.

Two versions for detection were implemented. The first one based on a QR decomposition of the channel matrix augmented with the regularizations factors to a matrix of size $2M \times K$. This is then formulated into a partial parallel implementation employing a systolic array [27]. The latter one based on a Neumann-series [17]. In the QR decomposition, each column is processed using the discrete steps of the modified Gram-Schmidt algorithm. The logic on the co-processors can be reconfigured so that the same hardware resources that provide the RZF decoder can also provide the ZF and MRC decoders, *i.e.,* the detection / precoding schemes discussed in Sec. II are supported with run-time switching. The Neumann-series based ZF detector utilizes the unique property that in MaMi, the Gramian matrix shows dominant diagonal elements if UEs use UL power control, or if scheduling is performed to serve UEs with similar power levels in the same time/frequency block to mitigate the influence of path loss differences. This, allows the matrix inversion to be approximated with low overall error [17]. The utilizations for the two FPGA designs are shown in Table 7. Clearly, overall processing complexity and resource utilization can be significantly reduced by exploiting the special properties of MaMi.

TABLE 7. FPGA utilization for two different MIMO processing implementations.

Implementation	Registers	LUT	RAMs	DSP48
QRD	46470 (9.1%)	49315 (20.3%)	171 (21.5%)	596 (38.7%)
Neumann-Series	16000 (3.1%)	28700 (11.8%)	6 (0.75%)	176 (11.4%)

At this point, the regularizations factors $\beta_{\text{reg}_{\text{pre}}}$ and $\beta_{\text{reg}_{\text{dec}}}$ are not run-time optimized but set manually, however, implementation of this feature is planned in future. For a more detailed discussion of the low-latency signal processing implementation on the testbed we refer to [26].

2) HOST-BASED VISUALIZATION AND DATA CAPTURING

The available margin of 1 GB/s and 14 P2P links to the corresponding maximum values on the co-processors are used for visualization and system performance metrics. The host receives decimated equalized constellations and raw subcarriers for one UL pilot and one UL data symbol per frame. These features add another

$$\frac{300 \cdot 2\text{bytes} + 2 \cdot 300 \cdot 4\text{bytes}}{10\text{ms}} = 300\,\text{MB/s}$$

of data flowing in and out of the co-processor. The raw subcarriers are used to perform channel estimation and UL data detection on the host computer with floating point precision and allow fast implementation of different metrics,

like constellation, channel impulse response, power level per antenna and user. Another 12 P2P links available are utilized to transmit and store real-time BERs for all 12 UEs.

Moreover, to be able to capture dynamics in the channel for mobile UEs, CSI can be stored on a ms basis. An integrated 2 GB Dynamic Random Access Memory (RAM) (DRAM) buffer on each of the co-processors was utilized for this since direct streaming to disk would exceed the P2P bandwidth limits. Snapshots can either be taken for 60 s in a 5 ms interval or over 12 s in a 1 ms interval, both corresponding to 2 GB of data for 300 subcarriers per co-processor.

3) SCALABILITY/RECONFIGURABILITY

Before startup, the number of deployed BS antennas can be arbitrarily set between 4 and 100. This is achieved by introducing zeros for non-existing antennas within the lookup-table (LUT)-based reconfigurable high-speed routers on the co-processors, thereby allowing to evaluate effects of scaling the BS antennas in real environments [26]. Additionally, all 140 OFDM symbols in a frame can be rearranged arbitrarily before start-up while each frame always repeats itself. For instance, we can choose to set the first symbol as UL pilots and all others as UL data in a static UL only scenario.

4) RECIPROCITY CALIBRATION

Estimation of the reciprocity calibration coefficients was implemented on the host, mainly for two reasons: *(i)* the host can perform all operations in floating-point which increases precision and *(ii)* the drift of the hardware is not significant once the system reached operating temperature [5]. Estimated reciprocity coefficients are applied in a distributed manner on the SDRs [26].

G. MECHANICAL STRUCTURE AND ELECTRICAL CHARACTERISTICS

Two computer racks containing all components measuring $0.8 \times 1.2 \times 1$ m were used, as shown Fig. 7. An essential requirement for the LuMaMi testbed is to allow tests in different scenarios, *e.g.,* indoor and outdoor. Therefore, the rack mount is attached on top of a 4-wheel trolley.

H. ANTENNA ARRAY

The planar T-shaped antenna array with 160 dual polarized $\lambda/2$ patch elements was developed in-house. A 3.2 mm Diclad 880 was chosen for the printed circuit board substrate. The T upper horizontal rectangle has 4×25 elements and the central square has 10×10 elements (see Fig. 7 right). This yields 320 possible antenna ports that can be used to explore different antenna array arrangements, for example 10×10 or 4×25 with the latter one being the default configuration. All antenna elements are center shorted, which improves isolation and bandwidth. The manufactured array yielded an average 10 dB-bandwidth of 183 MHz centered at 3.7 GHz with isolation between antenna ports varying between 18 dB and 28 dB depending on location in the array.

I. USER EQUIPMENT

Each UE represents a phone or other wireless device with single antenna capabilities. One SDR serves as two independent UEs such that overall six SDRs are required for the 12 UEs. The base-band processing, *i.e.,* OFDM modulation/demodulation and symbol mapping/demapping are essentially identical to the BS implementation. A least-square CSI acquisition is performed on precoded DL pilot followed by a ZF-equalizer. The DL pilots occupy a full OFDM symbol. The UEs may be equipped with any type of antenna using SMA connectors.

J. SYNCHRONIZATION

A MaMi BS requires time synchronization and phase coherence between each RF chain. This is achieved using the 10 MHz reference clock source and the reference clock and trigger distribution network (see Table 5). The reference clock is used as the source of each radio local oscillator, providing phase coherence among devices. The trigger signal is used to provide a time reference to all the radios in the system. A master provides an output digital trigger that is amplified and divided among all the radios. Upon receipt of the rising edge of the event trigger, all SDRs are started. The basic structure can be identified in Fig. 7 on the left.

To synchronize the UEs with the BS over-the-air (OTA), the LTE Zadoff-Chu Primary Synchronisation Signal (PSS) is used, which occupies the center 1.2 MHz of the overall bandwidth. OTA synchronization and frequency offset compensation are achieved by employing a frequency-shifted bank of replica filters. The process follows a two step procedure: finding a coarse candidate position by scanning over the whole radio frame followed by tracking the PSS in a narrowed window located around the coarse candidate position. Additionally, by disciplining the UE SDRs with Global Positioning System (GPS), frequency offset compensation may be avoided by lowering the frequency offset to <300 Hz.

VI. PROOF-OF-CONCEPT RESULTS

This section describes two experiments performed to validate our testbed design, the MaMi concept and its performance. The first test is performed indoors with high density of users per area unit to stress the spatial multiplexing capabilities of the system. The second test is conducted outdoors with less dense deployment of UEs and is primarily designed to test the range and multiplexing capabilities outdoors. For all tests, the default antenna configuration, *i.e.,* 4×25 was used on the BS side whereas the UEs were equipped with linear polarized ultra-wideband antennas. It has to be noted that all results shown in this section are obtained from real-time operation without UL power control.

A. INDOOR TEST

In this test real-time uncoded BER curves are measured, employing MRC/MRT and ZF as decoders/precoders. The UL BER curves are obtained by sweeping all UE TX power

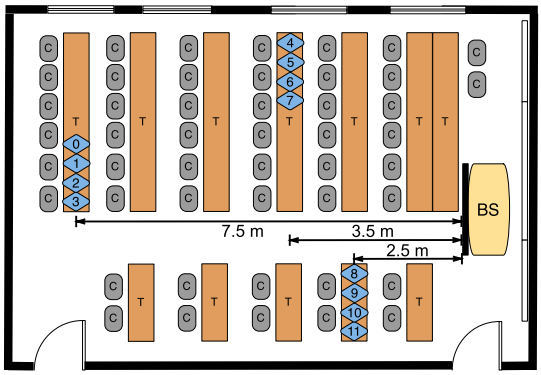


Fig. 8. The indoor measurement setup in a lecture room including the positions of the 12 UEs. The BS is shown at the right-hand side and is situated at the front of the lecture hall. The terminals are placed in groups of four on three different tables and distances to the BS.

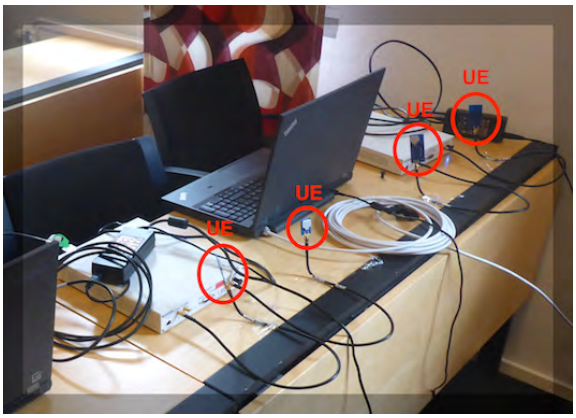


Fig. 9. One group of four UEs with a high user density per unit area to validate the spatial multiplexing capabilities of MaMi.

amplifier (PA) gains synchronously, and for the DL BER curves the PA gains of the BS TX chains while keeping other system parameters constant. Note that the initial parameterization of the system is chosen empirically, so it allows smooth BER curves starting at about 0.5. Each gain step is held constant for about 4 s corresponding to about 36×10^6 and 108×10^6 transmitted bits per step for QPSK and 64-QAM modulation, respectively.

1) SCENARIO

Twelve UEs are set up in a lecture hall at Lund University with the BS at the front as shown in Fig. 8 including the respective UE placements. All UEs are packed in groups of four resulting in a high density of UEs per area unit. One of these groups can be seen in Fig. 9.

2) UL BERs

Fig. 10, (a) and (b), show the BERs for all 12 UEs using ZF detector for QPSK and 64-QAM modulation, respectively. For both constellation sizes, the UEs furthest away, UE0 to UE3 show highest BER. UE0 and UE1 even show a sudden increase for the BER to 0.5 which was diagnosed to be due to saturation of their respective PAs. Moreover, their performance shows severe limitation compared to the other UEs, giving a clear indication that their performance is interference rather than power limited. The group closest to the BS, UE9-UE12, shows best performance although the variation within the group is still quite significant. Overall, the expected trend, increasing performance with increased transmit gain is clearly noticeable with the BER curve shapes resembling those of additive white Gaussian noise (AWGN) channels. Comparing the amplifier gain settings for QPSK and 64-QAM to achieve the same BER the differences are found to be in the range of 10 dB to 16 dB whereas a difference of 9 dB is expected for AWGN. Overall, it can be seen that all UEs except UE0 and UE1 achieve BER below 10% at an amplifier gain of 15 dB for QPSK and 25 dB for 64-QAM, respectively.

3) DL BERs

Fig. 10, (c) and (d), show the DL BERs using ZF precoder for QPSK- and 64-QAM modulation, respectively. Using QPSK modulation, the group closest to the BS, UE9-UE12, achieves a considerably better performance than the other two groups. Using 64-QAM, all UEs show an error-floor towards higher TX gain values which is likely a result of imperfect reciprocity calibration combined with leakage among UEs due to non-perfect channel knowledge resulting in interference among UEs. However, for the QPSK modulation case all UEs experience better BER rates which can be explained by the significantly higher available transmit power on the BS side, utilizing 100 active RF-chains. Comparing again the difference in amplifier gain setting for QPSK and 64-QAM, their differences are about 12 dB to 16 dB. The tests performed were mainly to prove functionality, and thus, no special care was taken to achieve best possible accuracy for the reciprocity calibration. However, individual parts are continuously tested to be improved.

4) PERFORMANCE EVALUATION

While the BERs plots in Fig. 10 nicely show the trend with increasing transmit power, they do not provide a real performance indication against signal-to-noise ratio (SNR). The current implementation of the testbed does not provide SNR estimates in real-time such that the data presented in Fig. 10 can be seen as the raw data provided during measurements. To provide an indication of the system performance the SNR of UE4 was estimated based on the received UL channel estimates. Estimated subcarriers at different time instances (about 200 ms apart) were subtracted / added to extract the noise / signal plus noise level which was then used to calculate the SNR value. However, this practice has limits as for close users interference may be stronger than the noise whereas for far away users the signal level may be too low. Therefore, UE4 was chosen which due to its placement during the measurement allowed a relatively good SNR estimation. Fig. 11 shows the BER of UE4 in comparison with the theoretical performance in AWGN and Rayleigh fading channels. It is

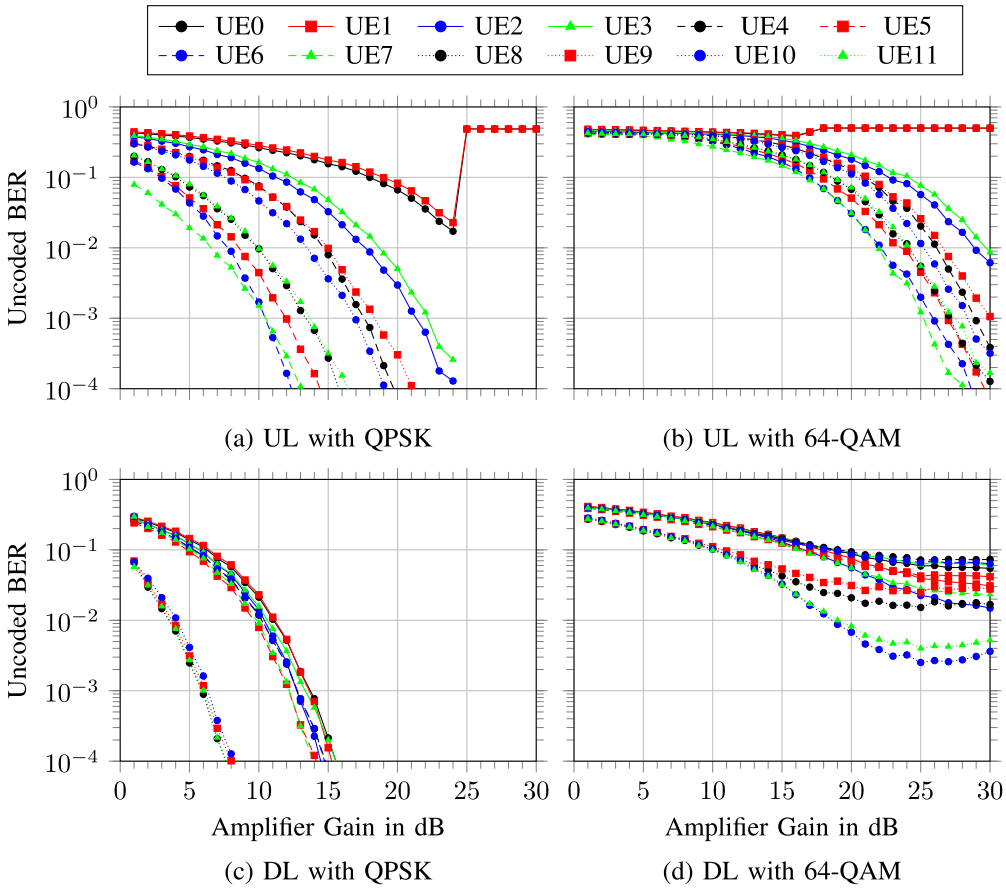


Fig. 10. UL and DL BERs for 12 UEs with ZF decoder/precoder.

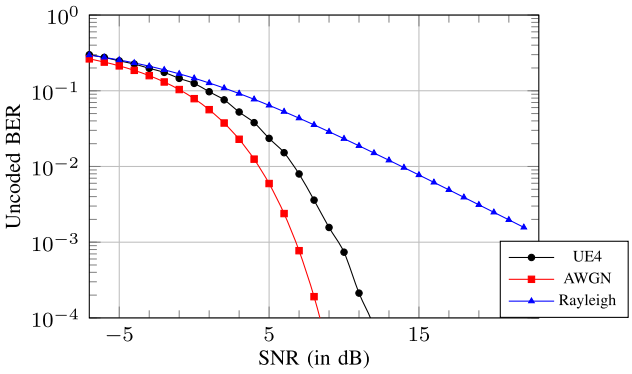


Fig. 11. Comparing the BER of UE4 to AWGN and Rayleigh fading channels.

visible that due to the excess amount of BS antennas the performance is close to the AWGN channel. To be more specific, due to the channel hardening the performance is only about 3 dB worse than for a AWGN channel which would be achieved for perfect channel hardening. On the DL the SNRs are affected by several factors including the higher overall transmit power from the 100 active RF-chains and possible inaccuracies in the reciprocity calibration coefficients. As DL precoding is performed based on UL channel estimates, SNR estimation is practically not feasible.

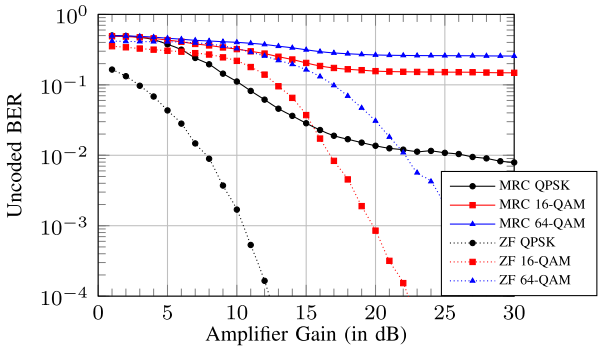
As all shown BERs curves closely resemble the shape of an AWGN channel it can be claimed that the MaMi concept works and is capable of serving 12 UEs on the same time/frequency resource even with a high UE density which in turn significantly improves the spectral efficiency compared to current cellular standards.

5) MRC/MRT VERSUS ZF

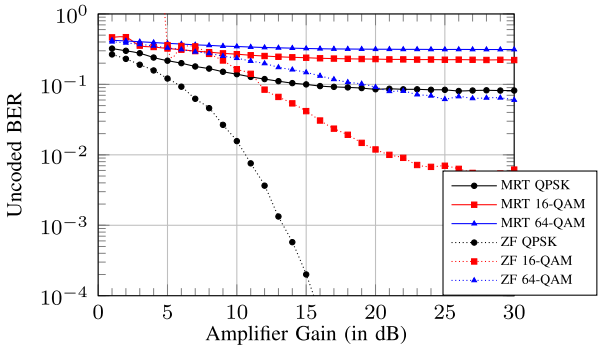
To compare the performance of MRC/MRT and ZF it is beneficial to isolate the analysis to one UE. Fig. 12a and Fig. 12b show the BER for UE7 for QPSK, 16-QAM and 64-QAM modulations while the BS employs either MRC/MRT or ZF on the UL and DL, respectively.

Overall, ZF shows an superior performance trend with increasing PA gains, while the performance of MRC appears to level off.[5] Looking in more detail, ZF is capable of achieving more than an order of magnitude lower BERs, compared to MRC. Using higher constellation sizes, 16-QAM or 64-QAM, the results for MRC show an even more significant deterioration. On the DL, ZF also outperforms MRT by far, the latter shows a significant error floor towards higher gains as in the UL case.

[5]This is expected from theory, as inter-user interference is the main source of error during data detection. The high density users setup adopted in this experiment highly contributes to this phenomena.



(a)



(b)

Fig. 12. BERs for UEs7 using QPSK, 16-QAM and 64-QAM modulation. (a) on the UL for ZF and MRC detector and (b) on the DL for ZF and MRT precoder.

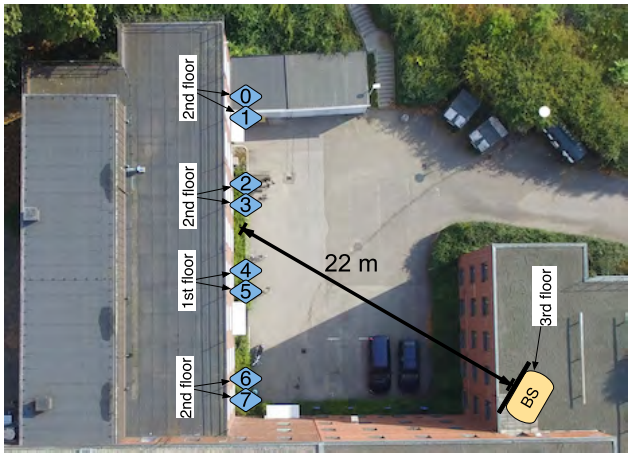


Fig. 13. Scenario for the outdoor tests. BS placed on the rooftop of the building (third floor) serving eight UEs on the opposite wing, with six UEs on second floor and two UEs on first floor.

Unfortunately, direct comparison between UL and DL results shown here is not easy to perform. This is due to the fact that on the UL, the performance is isolated to the UL transmit power only whereas on the DL a combination of UL channel estimate quality, DL transmit power and reciprocity accuracy determines overall performance.

B. OUTDOOR TEST

For the outdoor test, the testbed was placed on the rooftop of one of the wings of the department building while the UEs where placed on the opposite wing utilizing scaffolding mounted to the building. Up to eight UEs were served simultaneously in a distance of about 18 to 22 meters, six on the second floor and two on the first floor while the testbed was situated on the third floor (rooftop). The scenario is shown in Fig. 13.

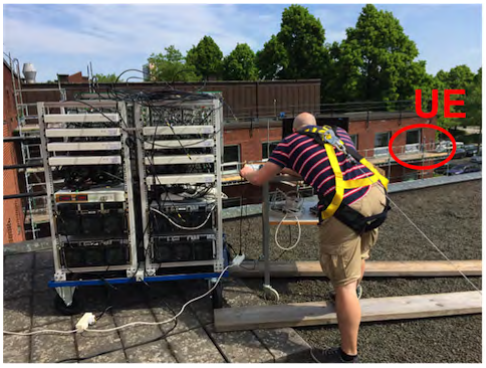
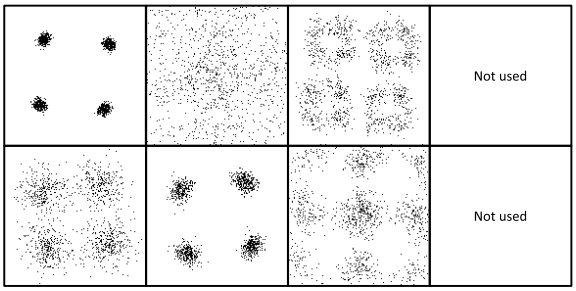
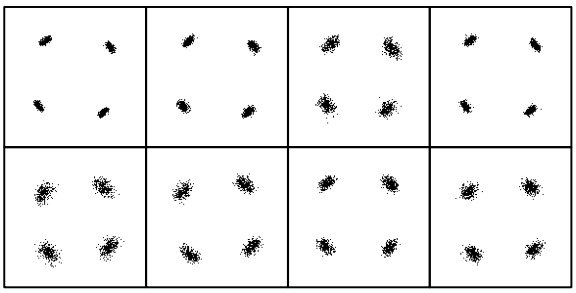


Fig. 14. The outdoor test scenario setup with the BS deployed on the rooftop of the department building marked with two UEs on the opposite building wing.



(a)



(b)

Fig. 15. UL constellations for the outdoor experiment: (a) when using MRC with 6 UEs and (b) when using ZF to serve 8 UEs.

Fig. 14 shows the BS placed on the rooftop of the department building facing towards the opposite wing. The placement for UEs 0 and 1 is also marked.

Fig. 15 shows a screenshot of the received UL QPSK constellations for this test setup when using MRC and ZF, respectively. Using MRC without error-correcting code (ECC) for this test, the six UEs show significant interference. Therefore, focus is put on the results obtained with ZF which is capable of separating up to eight UEs and shows very clear constellations, due to the interference suppression.

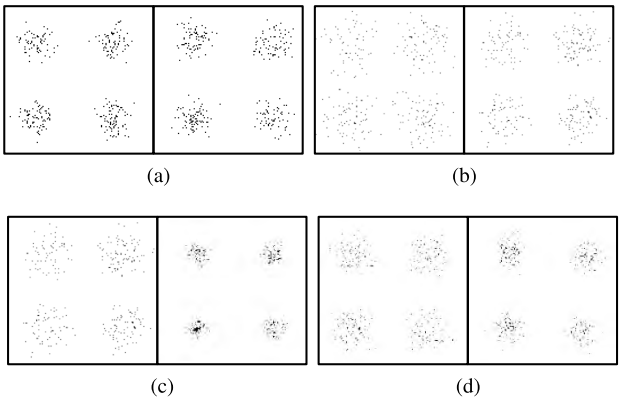


Fig. 16. Received DL constellations using ZF: (a) UE0 & UE1 (b) UE2 & UE3 (c) UE5 & UE8 and (d) UE9 & UE10.

Considering ZF on the DL, the constellations for all 8 UEs can be seen in Fig. 16. Although in-detail analysis is not provided for this test, it is clearly visible that ZF outperforms MRC which is often claimed to be sufficient in literature when analyzing performance based on iid channel models [1]. The results observed in this experiment are representative for most tests performed so far, *i.e.,* DL always showed to be the more challenging duplex case.

The LuMaMi testbed was also utilized to perform the first MaMi outdoor mobility measurements involving moving pedestrians and cars as UEs, however, a discussion of this is out of scope of this paper. Results and analysis from the mobility tests can be found in [28].

VII. CONCLUSION

This paper presented the LuMaMi testbed, which is the first fully operational real-time testbed for prototyping massive MIMO. Based on massive MIMO system requirements, system parameters were discussed and defined. Further, a detailed generic hardware partitioning to overcome challenges for data shuffling and peer-to-peer link limitations while still allowing scalability, was proposed. By grouping Software-Defined Radios and splitting overall bandwidth, implementation of massive MIMO signal processing was simplified to cope with challenges like time-division duplex precoding turnaround time and limited peer-to-peer bandwidth enforcing strict design requirements when scaling the number of base station antennas up to 100 or higher. Based on the generic system partitioning and system requirements, a hardware platform was selected and evaluated. It was shown that internal system configuration is within throughput and processing capabilities before the complete LuMaMi testbed parameters were described. Finally, field trial results including Bit Error Rate performance measurements and constellations were presented from both indoor and outdoor measurement campaigns. The results showed that it is possible to separate up to 12 user equipments on the same time/frequency resource when using massive MIMO. Having established a flexible platform for testing new algorithms and digital base-band solutions we are able to take massive MIMO from theory to real-world tests and standardization for next generation wireless systems.

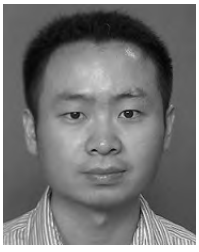
REFERENCES

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
[2] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
[3] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
[4] C. Shepard *et al.*, "Argos: Practical many-antenna base stations," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw. (Mobicom)*, New York, NY, USA, 2012, pp. 53–64. [Online]. Available: http://doi.acm.org/10.1145/2348543.2348553
[5] J. Vieira *et al.*, "A flexible 100-antenna testbed for massive MIMO," in *Proc. Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 287–293.
[6] P. Harris, S. Zang, A. Nix, M. Beach, S. Armour, and A. Doufexi, "A distributed massive MIMO testbed to assess real-world performance and feasibility," in *Proc. IEEE 81st Veh. Technol. Conf. (VTC Spring)*, May 2015, pp. 1–2.
[7] 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.
[8] R. Rogalin *et al.*, "Scalable synchronization and reciprocity calibration for distributed multiuser MIMO," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1815–1831, Apr. 2014.
[9] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul. 2014.
[10] E. Björnson, E. G. Larsson, and Thomas L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.
[11] A. Molisch, *Wireless Communications* (Wiley IEEE). Hoboken, NJ, USA: Wiley, 2010.
[12] MAMMOET (Massive MIMO for Efficient Transmission). (Jan 2016). *ICT- 619086-D3.2: Distributed and Centralized Baseband Processing Algorithms, Architectures, and Platforms*, EU-project Deliverable. [Online]. Available: https://mammoet-project.eu/publications-deliverables
[13] N. Shariati, E. Björnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 815–830, Oct. 2014.
[14] Y. Han and J. Lee, "Uplink pilot design for multi-cell massive MIMO networks," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1619–1622, Aug. 2016.
[15] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
[16] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A comprehensive survey of pilot contamination in massive MIMO—5G system," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 905–923, 2nd Quart., 2016.
[17] H. Prabhu, J. Rodrigues, O. Edfors, and F. Rusek, "Approximative matrix inverse computations for very-large MIMO and applications to linear precoding systems," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 2710–2715.
[18] National Instruments. (2014). *USRP-2943R Data Sheet*, accessed on Oct. 4, 2016. [Online]. Available: http://www.ni.com/datasheet/pdf/en/ds-538
[19] National Instruments. (Jul. 2014). *FlexRIO 7976R Data Sheet*, accessed on Oct. 4, 2016. [Online]. Available: http://www.ni.com/pdf/manuals/374546a.pdf
[20] Xilinx. (2016). *7 Series FPGAs Overview: DS180 (v2.0) Product Specification*, accessed on Oct. 4, 2016. [Online]. Available: http://www.xilinx.com/support/documentation/datasheets/ds1807SeriesOverview.pdf
[21] National Instruments. (2015). *PXIe 1085 Manual*, accessed on Oct. 4, 2016. [Online]. Available: http://www.ni.com/pdf/manuals/373712f.pdf

[22] National Instruments. (2011). *MXI-Express x4 Series User Manual*, accessed on Oct. 4, 2016. [Online]. Available: http:/0/www.ni.com/pdf/manuals/371977c.pdf
[23] National Instruments. (2013). *PXIe 8135 Manual*, accessed on Oct. 4, 2016. [Online]. Available: http://www.ni.com/pdf/manuals/373716b.pdf
[24] *PXIe-6674T User Manual: Timing and Synchronization Module for PXI Express*, National Instruments, Austin, TX, USA, 2015.
[25] Ettus Research. *USRP Hardware Driver and USRP Manual: Octo-Clock*, accessed on Oct. 4, 2016. [Online]. Available: http://files.ettus.com/manual/page octoclock.html
[26] S. Malkowsky *et al.*, "Implementation of low-latency signal processing and data shuffling for TDD massive MIMO systems," in *Proc. IEEE Int. Workshop Signal Process. Syst. (SiPS)*, Oct. 2016, pp. 260–265.
[27] D. Wubben, R. Bohnke, V. Kuhn, and K.-D. Kammeyer, "MMSE extension of V-BLAST based on sorted qr decomposition," in *Proc. IEEE 58th Veh. Technol. Conf.*, vol. 1. Oct. 2003, pp. 508–512.
[28] P. Harris *et al.* (Jan. 2017). "Performance characterization of a real-time massive MIMO system with LOS mobile channels." [Online]. Available: https://arxiv.org/abs/1701.08818

STEFFEN MALKOWSKY (S'13) received the B.Eng. degree in electrical engineering and information technology from Pforzheim University, Germany, in 2011, and the M.Sc. degree in electronic design from Lund University in 2013. He is currently pursuing the Ph.D. degree with the Digital ASIC Group, Department of Electrical and Information Technology, Lund University. His current research interests include development of reconfigurable hardware and implementation of algorithms for wireless communication with emphasis on massive MIMO. For the development of a massive MIMO testbed in collaboration with University of Bristol and National Instruments Corporation and a set spectral efficiency world record, he received five international awards from National Instruments, Xilinx and Hewlett Packard Enterprise.

JOÃO VIEIRA (S'14) received the B.Sc. degree in electronics and telecommunications engineering from the University of Madeira in 2011, and the M.Sc. degree in wireless communications from Lund University, Sweden, in 2013. He is currently pursuing the Ph.D. degree with the Department of Electrical and Information Technology, Lund University. His main research interests regard parameter estimation and implementation issues in massive MIMO systems.



LIANG LIU (M'10) received the B.S. and Ph.D. degree from the Department of Electronics Engineering (2005) and Micro-electronics (2010) from Fudan University, China. In 2010, he was a Visiting Researcher with the Rensselaer Polytechnic Institute, USA. He joined Lund University as a Post-Doctoral in 2010. Since 2016, he is currently an Associate Professor with Lund University. His research interest includes wireless communication system and digital integrated circuits design. He is a Board Member of the IEEE Swedish SSC/CAS chapter. He is a member of the technical committees of VLSI systems and applications and CAS for communications of the IEEE circuit and systems society.

PAUL HARRIS (S'14) received the degree (Hons.) in electronic engineering from the University of Portsmouth in 2013. He is currently pursuing the Ph.D. degree with the Communication Systems and Networks Group, University of Bristol. His current research interests include massive MIMO system design, performance evaluation in real-world scenarios, and the optimization of techniques, such as user grouping or power control using empirical data. Working in collaboration with Lund University and National Instruments, he implemented a 128-antenna massive MIMO test system and led two research teams to set spectral efficiency world records in 2016. For this achievement, he received five international awards from National Instruments, Xilinx, and Hewlett Packard Enterprise, and an honorary mention in the 2016 IEEE ComSoc Student Competition for Communications Technology Changing the World.

KARL NIEMAN (M'06) received the B.S. degree in electrical engineering from the New Mexico Institute of Mining and Technology in 2009, and the M.S. and Ph.D. degrees in electrical engineering from The University of Texas at Austin, in 2014 and 2011, respectively. He is currently a Senior Wireless Platform Architect with the Advanced Wireless Research Team, National Instruments. His interests are with research and standardization of 5G technology, particularly Massive MIMO architectures and signal processing. He has designed and implemented several FPGA-based real-time wireless communication systems, has made multiple contributions to 3GPP RAN1, and holds multiple issued and pending patents on 5G technologies.

NIKHIL KUNDARGI (M'08) received the Ph.D. degree in electrical engineering from The University of Texas at Austin in 2012. He is currently a Senior Wireless Platform Architect with the Advanced Wireless Research Team, National Instruments Corporation since 2012. He leads the Massive MIMO research initiative at NI. He is the 3GPP RAN1 Delegate for NI and participates with LTE-Advanced and 5G cellular standardization. His current research interests include massive MIMO, full dimension MIMO, 5G new radio, mmWave, PHY/MAC layer design and prototyping, dense LTE networks, real-time DSP, software defined radio architectures, Cognitive radio networks, spectrum sensing, dynamic spectrum access, anomaly detection and statistical signal processing. He was a member of Wireless Networking and Communications Group, UT Austin and formerly a Graduate School Fellow with the University of Minnesota from 2006 to 2010. He is an Active Member of the IEEE Austin ComSoc Chapter and served as the Vice-Chair for the Industry Forum at the IEEE Globecom 2015.

IAN C. WONG (SM'08) received the B.S. (magna cum laude) degree in electronics and communications engineering from the University of the Philippines in 2000, and the M.S. and Ph.D. degrees in electrical engineering from The University of Texas at Austin in 2004 and 2007, respectively. He is currently a Senior Manager of the Advanced Wireless Research Group, National Instruments, where he leads the companyŠs 3GPP and 802.11 wireless standards strategy and platforms for wireless system design, simulation, prototyping, and implementation. From 2007 to 2009, he was a Systems Research and Standards Engineer with Freescale Semiconductor, where he represented Freescale in the 3GPP LTE standardization efforts. He was the co-author of the Springer book *Resource Allocation for Multiuser Multicarrier Wireless Systems* has over ten patents, over 25 peer-reviewed journal and conference papers, and over 40 standards contributions. His current research interests include 5G wireless systems design and prototyping, and design automation tools for rapid algorithm development. He received the Texas Telecommunications Engineering Consortium Fellowship from 2003 to 2004, and the Wireless Networking and Communications Group Student Leadership Award in 2007. He was the Industry Program Chair for the IEEE Globecom 2015 in Austin, the Director for Industry Communities for the IEEE Communications Society from 2016 to 2017.

FREDRIK TUFVESSON (F'16) received the Ph.D. degree from Lund University, Sweden, in 2000. He joined the Department of Electrical and Information Technology, Lund University, where he is currently a Professor of Radio Systems. He has authored around 60 journal papers and 120 conference papers. His main research interests are channel modeling, measurements and characterization for wireless communication, with applications in various areas, such as massive MIMO, UWB, mm wave communication, distributed antenna systems, radio-based positioning, and vehicular communication.He received the Neal Shepherd Memorial Award for the best propagation paper in the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

VIKTOR ÖWALL (M'90) received the M.Sc. and Ph.D. degrees in electrical engineering from Lund University, Lund, Sweden, in 1988 and 1994, respectively. From 1995 to 1996, he joined the Electrical Engineering Department, University of California at Los Angeles, as a Post-Doctoral, where he was involved in multimedia simulations. Since 1996, he has been with the Department of Electrical and Information Technology, Lund University, Lund, Sweden. He was the Founder and the Director of the VINNOVA Industrial Excellence Center in System Design on Silicon, which he headed until 2014. He was a Co-Founder of the start-up company Phase Holographic Imaging, who develops microscopes utilizing digital holography. He is currently a Full Professor and since 2015, he has been the Dean of the Faculty of Engineering. His main research interest is in digital hardware implementation, especially algorithms and architectures for wireless communication and biomedical applications.

OVE EDFORS (SM'17) is currently a Professor of Radio Systems with the Department of Electrical and Information Technology, Lund University, Sweden. His current research interests include statistical signal processing and low-complexity algorithms with applications in wireless communications. In the context of massive MIMO, his main research focus is on how realistic propagation characteristics influence system performance and base-band processing complexity.