

Large-Scale Antenna Systems and Massive Machine Type Communications

Felipe A. P. de Figueiredo^{*‡}, Fabbryccio A. C. M. Cardoso[†], João Paulo Miranda[†], Ingrid Moerman^{*},
Claudio F. Dias[‡], and Gustavo Fraidenraich[‡]

^{*}Ghent University - imec, IDLab, Department of Information Technology, Ghent, Belgium.

[†]CPqD – Research and Development Center on Telecommunications, Brazil.

[‡]DECOM/FEEC – State University of Campinas (UNICAMP), Brazil.

Email: ^{*}[felipe.pereira, ingrid.mmoermen]@ugent.be, [†][fcardoso, jmiranda]@cpqd.br, [‡][aplrx, gf]@decom.feec.unicamp.br

Abstract—In this paper, we identify issues and possible solutions in the key area of large-scale antenna systems, also known as Massive Multiple Input Multiple Output (MIMO) systems. Additionally, we propose the use of Massive MIMO technology as a means to tackle the uplink mixed-service communication problem. Under the assumption of an available physical narrowband shared channel (PNSCH), the capacity of the MTC network and, in turn, that of the whole system, can be increased by grouping Machine-Type Communication (MTC) devices into clusters and letting each cluster share the same time-frequency physical resource blocks. We study the feasibility of applying sub-optimal linear detection to the problem of detecting a large number of MTC devices sharing the same time-frequency resources at the uplink of a base station (BS) equipped with a large number of antennas, M . In our study, we derive the achievable lower-bound rates for the studied sub-optimal linear detectors and show that the transmitted power of each MTC device can be reduced as M increases, which is a very important result for power-constrained MTC devices running on batteries. Our simulation results suggest that, as M is made progressively larger, the performance of sub-optimal linear detection methods approach the matched filter bound, also known as perfect interference-cancellation bound.
Index Terms—G sub-optimal linear detection Massive MIMO MTC IoTG sub-optimal linear detection Massive MIMO MTC IoT5

I. INTRODUCTION

We have recently witnessed dramatic changes in the way communications systems are used. These changes are, in part, due to the big rise in on-demand data consumption over mobile and wireless networks. One issue associated with the task of accommodating such changes consists of finding solutions that can meet the diverse needs of use cases regarded as market drivers for Fifth Generation (5G) networks. The development of 5G networks has been driven by a number of use cases aimed at supporting innovative applications and services [1]. International Telecommunication Union (ITU) has divided 5G network applications and services into three main categories (also referred sometimes as use cases): enhanced Mobile Broadband (eMBB); Ultra-Reliable Low-Latency Communications (URLLC); and massive Machine Type Communications (mMTC) [1, 2]. A non-exhaustive list of 5G drivers per service category includes:

- **eMBB**: broadband telephony with gigabit wireless connectivity for public safety and immersive multimedia

applications, such as high-resolution video, virtual reality and gaming [3];

- **URLLC**: the Tactile Internet for real-time applications posing ultra-low latency requirements [4];
- **mMTC**: massive deployment of Internet of Things (IoT) devices for machine-centered communications in dense networks of sporadic, bursty traffic generating devices [5].

Applications within the IoT driver's scope, which we shall refer to as MTC throughout the remainder of this paper, range from infrastructure monitoring to smart cities [6, 7], and from mobile health – including telemedicine, sports and fitness – to Advanced Driver Assistance Systems (ADAS) [8, 9]. Reliability in smart grid and critical infrastructure monitoring, for instance, is often achievable only via dedicated landlines [10, 11]. Telemedicine involves the diagnostic through medical records stored in the cloud, calling for both real-time, low-latency access and high-capacity infrastructure capable of handling data of voluminous nature, *e.g.*, magnetic resonance imaging and computerized axial tomography [12, 13]. Infotainment, pre-crash sensing/mitigation, and vehicular cooperation in ADAS also need support for high-speed, low-latency car-to-car and car-to-infrastructure communications [14]–[16].

As the discussion above attests, 5G requirements can indeed be quite diverse even within a single market driver. Another issue raised by the MTC is scalability, as the current premise is that hundreds to hundred thousands of low-cost MTC devices will be served by a sole Base Station (BS) [17]. While scalability issues have been addressed using different (sometimes complementary) approaches, such as lessons learned from duty-cycled Wireless Sensor Networks [18], waveform design for asynchronous signaling in the uplink [19], and sparse signal processing strategies [20], less is understood to date about the application of Massive MIMO techniques in the context of MTC networks. In Massive MIMO systems, a BS, equipped with a large number of antennas, serves a myriad of single-antenna users over the same time-frequency resources, where these users can be UEs and/or MTC devices.

A popular view is that the required increase in data rate will be achieved through *combined gains* [21] in extreme network densification (to improve area spectral efficiency, *i.e.*, bits/s/Hz/Km²), increased bandwidth (by exploit-

ing mmWaves), making better use of unlicensed spectrum bands [22], and increased spectral efficiency (through advances in MIMO techniques, *e.g.*, Massive MIMO). A consequence of the powerful signal processing enabled by large sizes of antenna arrays is that most of the scheduling and physical layer control issues in general are automatically resolved in Massive MIMO systems – which of course is not the case for systems with just a moderate number of antennas. In recent past, Massive MIMO has gained significant momentum as potential candidate to increase the capacity in multi-user networks [23]. In the limit, as the number M of antennas at the BS tends to infinity, the system processing gain tends to infinity. As a result, the effects of multi-user interference (*i.e.*, cross-talk), fast fading and noise disappear [23]. In [24] it is pointed out that, in a Massive MIMO system, because of the law of large numbers, the channel hardens so that each subcarrier within a coherence interval will substantially have the same asymptotically (with M) deterministic channel gain. Another advantage of Massive MIMO systems is that, thanks to their many spatial degrees of freedom and the favorable propagation phenomenon, the same time-frequency resources can be simultaneously used to serve a great number of users [23]. This plus the channel hardening render frequency-domain scheduling no longer needed, making most of the physical layer control signaling redundant [25]. Additionally, the adoption of massive MIMO systems also improve frequency reuse (due to the possibility of reduced radiated power), simplify power control (power control coefficients depend only on the large-scale fading coefficients due to channel hardening) and decreases multi-user interference (due to favourable propagation and the possibility of having very narrow beams as M increases) [23, 26].

This is an extension of our previous work [27, 28]. Differently from [27, 28], where we have only considered the Bit Error Rate (BER) analysis for perfect channel knowledge, the current paper presents additional studies and results that shed light on the application of massive MIMO technology and sub-optimal linear detection to the uplink mixed-service communication problem, where a BS has to serve not only Human Type Communications (HTC) devices but also a possible massive number of MTC devices. Although the subject being studied in this work is the same as the one in [29], *i.e.*, the application of Massive MIMO to the machine type communications problem, the results are fundamentally different. The results in [29] are related to *imperfect channel knowledge*, where the BS has to estimate the channels based on pilot sequences sent by the MTC devices. In [29], all the discussion is based on the imperfect channel knowledge assumption. On the other hand, our current work deals with the *perfect channel knowledge* assumption, where we assume the BS has perfect knowledge of the channels and therefore, it doesn't need to estimate them. Compared to [29], the current paper presents a conceptual study, where we wanted to assess the best possible scenario, *i.e.*, when the channel is fully known. The main contributions of this paper are the following:

- A survey on existing related work that identifies main issues and candidate solutions in the key area of Massive

MIMO systems. By presenting this survey, we aim at illustrating the crucial role this technology is envisioned to play in the context of 5G wireless communication systems.

- Investigation of the feasibility of applying Massive MIMO as means to address the so-called uplink mixed-service communication problem, where a *single* BS simultaneously delivers services to *both* narrowband MTC devices *and* Fourth Generation (4G) wideband services to User Equipment (UEs), *i.e.*, HTC devices. In this problem, the BS has to serve HTC devices and a possible massive number of MTC devices. The problem can be split in to two sub-problems, namely, random access and data transmission problems. The first one deals with the congestion and overloading issues brought about by the massive number of devices (*i.e.*, HTC and MTC devices) trying to get access to the network during the random access phase [30]. The second problem tackles the challenge of accommodating the data transmissions of this huge number of devices, which might tremendously impact on the operations and quality of the provided services of a mobile network [30, 31]. The focus of our work is on the data transmission phase, and in this paper we propose a massive MIMO-based scheme where the data transmissions of a possible huge number of MTC devices are served through the same time-frequency resources by a BS equipped with a large number of antennas. Treating MTC devices as regular UEs turns out to be an issue, as scheduling Physical Resource Blocks (PRBs) in extremely crowded networks is a nontrivial task made harder in the presence of retransmissions and intrinsic uplink synchronization procedures [32]–[35]. Under the assumption that a PNSCH, devised to consume the data traffic generated by MTC devices, is available, the capacity of the MTC network – and, in turn, the mixed-service system's – can be increased by clustering MTC devices and letting clusters share the same time-frequency PRBs. The underlying idea behind the PNSCH is the exploitation of the channel's geometric scattering characteristics to spread MTC signals in the spatial domain. Individual data streams conveyed by spatially spread MTC signals can be separated thanks to the powerful processing gain of the Massive MIMO setup [36], where the size of the antenna array used at the BS is at least one order of magnitude larger than the number of served MTC devices.
- Based on perfect Channel State Information (CSI), we assess and discuss the feasibility of employing simple and sub-optimal linear detection schemes (*e.g.*, MRC, ZF and MMSE) instead of highly complex and optimal non-linear schemes at a BS equipped with a large number of antennas. We derive closed-form lower-bound expressions on the uplink achievable rates for each one of the studied linear detectors for a finite number of antennas, M . Additionally, we show that even when using simple and sub-optimal linear detectors, the transmitted power of each MTC device can be reduced when the number of antennas grows. This is a very important result for

power-constrained MTC devices running on batteries.

- We present several simulation results showing that (i) the bit error rate (BER) of sub-optimal linear detection techniques approaches the perfect interference-cancellation bound [37], as the antenna array size progressively increases, (ii) the derived achievable lower-bound rates of the studied linear detectors are tight, (iii) the transmitted power of each MTC device can be reduced with M . Based on our study, we conclude that aspects like antenna array size, performance-complexity tradeoff, and balance between interference suppression and noise enhancement dictate, as expected, the performance of a given detector.

The remainder of the paper is organized as follows. Section II provides a brief but comprehensive overview of Massive MIMO, its challenges, and solutions available at the time of this writing. Section III presents the proposed system model and provides mathematical descriptions for each one of its functional blocks: signal generation & transmission, channel model and signal detection. We also present in this section a discussion on the capacity lower and upper bounds in favourable propagation. Section IV briefly discusses contemporary solutions for signal detection in Massive MIMO systems, making the case for sub-optimal linear detection methods. In section V we derive lower-bound expressions for the uplink achievable rates when linear detection (*i.e.*, MRC, ZF and MMSE) is employed with perfect CSI knowledge at the BS. The results of our simulation work are presented and discussed in Section VI, while Section VII wraps up the paper with concluding remarks and suggestions for future work.

II. MASSIVE MIMO CHALLENGES

This section discusses issues regarded as most challenging in the Massive MIMO literature. Table I lists such issues and their available solutions, each presented alongside with its side effects, *i.e.*, new issues brought about by their adoption [27, 28].

A. Impairments due to Low-cost Hardware

Large-scale multiple antenna arrays will likely be built using low-cost components to ease the introduction and leverage the penetration of the Massive MIMO technology into the market. This calls for solutions capable of circumventing hardware imperfections that manifest themselves as I/Q imbalance or phase noise. The latter issue is of particular concern because low-cost power amplifiers often have relaxed linearity requirements, which in turn translate into the need for reduced PAPR on a per antenna element basis [24].

Savings in radiated power result from using excess antennas to simultaneously send independent data to different users, but the total power consumption should also be taken into account. In this context, an interesting research path is hardware architectures for baseband signal processing [24]. Another path of interest is experimentation, as testbeds currently available only demonstrate basic capabilities, and do not take constrained BS real estates into consideration [38]. Experimentation can also be rewarding in that experimental findings can be fed back into theory, thus rendering the development of testbeds, prototypes,

and proof-of-concept experiments of utmost importance to a better understanding about the massive MIMO technology.

B. Mutual Coupling and Front-back Ambiguity

One assumption often made when modeling antenna arrays is that the separation among antenna elements is large enough to keep mutual coupling at negligible levels. This is not entirely realistic, especially in the case of a large number of antenna elements deployed as an array of constrained size and aperture. Under such practical conditions, mutual coupling is known to substantially impact the achievable system capacity [39]. Multiport impedance matching RF circuits can cancel out such coupling effects [40], but they diminish output port bandwidth [41] and increase ohmic losses [42, Chapter 10].

Two- or three-dimensional arrays have been reported to be able to avoid front-back ambiguity. A side effect of dense implementations is that the larger the number of adjacent elements, the larger the increase of coupling effects [43]. Another fundamental shortcoming specific to 3-D settings is the incapability of extracting additional information from the elements inside the array, *i.e.*, only elements on the array surface contribute to the information capacity [44]. The optimal densities above which performance deteriorates no matter how large is the number of elements are studied in [45] for indoor Massive MIMO BSs.

C. RF Propagation and Channel Modeling

Realistic performance assessments call for appropriate channel characterization and modeling. The Massive MIMO channel behavior, including its correlation properties and the influence of different antenna arrangements, cannot be captured otherwise. The interest raised by this issue has been (and still is) experiencing a fast-paced growth, and the community has already managed to contribute towards a better understanding on the matter. In [46], channel measurements are carried out to identify and statistically model the propagation characteristics of interest. These are then fed back into an existing channel model, extending its applicability to large-scale antenna arrays.

Performance assessments should ideally be conducted using a standardized or widely accepted channel model. Some models for Massive MIMO are presented and discussed in [47]. See, *e.g.*, [48], for a discussion on modeling methods, channel categories, and their underlying properties.

D. Acquisition of Channel State Information

In conventional Frequency-division Duplexing (FDD) systems, the BS cannot harness beamforming gains until it has established a communication link with the terminals. Firstly, the BS broadcasts pilots based on which the terminals estimate their corresponding channel responses. These terminal estimates are then quantized and fed back to the BS. Such FDD approach finds limited application in Massive MIMO systems in that the amount of time-frequency resources needed for pilot transmission in the downlink (DL) scales with the number of antennas, and so does the number of channel responses that must be estimated on the part of each terminal. In systems

TABLE I
SUMMARY OF CHALLENGES & SOLUTIONS IN LARGE-SCALE MULTIPLE ANTENNA SYSTEMS FOR 5G.

Research Area	Issue	Candidate Solutions	Shortcomings and “Side Effects”	Refs	
Hardware	Phase noise	Smart PHY transceiver algorithms	Efficacy yet to be demonstrated	[24]	
	Power consumption	Parallel, dedicated baseband processing	Open research question		
	Proof-of-Concept	Experiments, testbeds & prototypes	Only basic capabilities demonstrated	[38]	
Antenna Aspects	Antenna coupling	Multipoint impedance matching RF circuits	<ul style="list-style-type: none"> • Diminishes bandwidth • Introduces ohmic losses • Not fully understood for large M 	[39]–[42] [45]	
	Front-back ambiguity	Dense multidimensional implementations	<ul style="list-style-type: none"> • Increases coupling effects • Limited to indoor environments • 3D arrays have restricted usefulness 	[39, 44]	
Propagation	Channel modeling	<ul style="list-style-type: none"> • Realistic empirical models • Sophisticated analytical models 	Currently under development	[46]–[48, 65]	
	Cluster resolution	No solution known to date	Open research question	[43]	
Transceiver Design	CSI acquisition	BS sends pilots to terminals via FDD	Limited by the channel coherence time	[23, 49]	
		Terminals send pilots to BS via TDD	Channel reciprocity calibration	[51]–[54]	
			Pilot contamination problem	[55]–[60]	
	Precoding	Linear precoding methods	<ul style="list-style-type: none"> • ZF • MMSE 	<ul style="list-style-type: none"> • Computationally heavy for large M • Higher average transmit power 	[43, 67]
			<ul style="list-style-type: none"> • MF 	<ul style="list-style-type: none"> • Has an error floor as M increases • Higher M required for a given SINR 	[43]
			<ul style="list-style-type: none"> • BD 	Cost-effective strategies are needed	[64]
		Nonlinear precoding methods	<ul style="list-style-type: none"> • DPC 	Extremely costly for practical deployments	[61]
			<ul style="list-style-type: none"> • THP • VP 	Increased complexity is hard to justify	[62] [63]
			<ul style="list-style-type: none"> • MRC 	<ul style="list-style-type: none"> • Does not treat interference suppression 	[23, 50]
	Detection	Linear filtering	<ul style="list-style-type: none"> • ZF • MMSE 	<ul style="list-style-type: none"> • Does not treat noise enhancement • More complex than MRC & ZF 	[37] [43, 66]
			<ul style="list-style-type: none"> • MMSE-SIC • BI-GDFE 	Computationally heavy for large M	[68]
			<ul style="list-style-type: none"> • TS • LAS 	More complex than MMSE-SIC	[69] [70]
Tree-based algorithms		<ul style="list-style-type: none"> • SD 	Complexity grows exponentially in M	[73]	
		<ul style="list-style-type: none"> • FCSD 	<ul style="list-style-type: none"> • 1,000x more complex than TS • Best suitable for the $M \approx K$ case 	[74]	

with large antenna arrays, pilot transmission time may well exceed the coherence time of the channel [23, 49].

An alternative for Massive MIMO systems is to let the terminals send pilots to the BS via Time-division Duplexing (TDD). The TDD approach relies on channel reciprocity, where uplink (UL) channels serve as estimate of DL channels. This leads to training requirements independent of the number of antennas, M [51], and eliminates the need for Channel State Information (CSI) feedback. TDD’s drawbacks are reciprocity calibration and pilot contamination: the former is a need raised by different transfer characteristics of DL/UL processing chains (*e.g.*, amplifiers, filters, local oscillators, etc. present different characteristics); the latter arises in multi-user multi-cell scenarios where the use of non-orthogonal pilot sequences causes the intended user’s channel estimate to get contaminated by a linear combination of other users’ channels sharing that same pilot. Reciprocity calibration and pilot decontamination are studied in [52]–[54] and [55]–[60], but optimal solutions are unknown to date.

E. Precoding

Multi-user interference can be mitigated at the transmit side by modifying standard single-stream beamforming techniques

to support multiple streams. Precoding based on Zero-Forcing (ZF) or Minimum Mean Square Error (MMSE) is simple for a moderate number of antennas. However, reliance on channel inversions may take its complexity and power burdens to a point hard to accommodate within very large arrays [43, 67]. Matched Filtering (MF), which comprises Maximum Ratio Transmission (MRT) in the DL and Maximum Ratio Combining (MRC) in the UL, is known to be the simplest method [23].

Nonlinear precoding methods, such as Dirty Paper Coding (DPC) [61], Tomlinson-Harashima Precoding (THP) [62], and Vector Perturbation (VP) [63], also have appealing features (DPC is theoretically optimal) but are either too costly for practical deployment or offer gains hard to justify in view of their increased computational complexity. Recalling that the array size required to achieve a given Signal-to-Interference and Noise Ratio (SINR) with MF is at least two orders of magnitude larger than with ZF [43], further work on cost-effective solutions is needed, *e.g.*, as illustrated in [64] for Block Diagonalization (BD) algorithms.

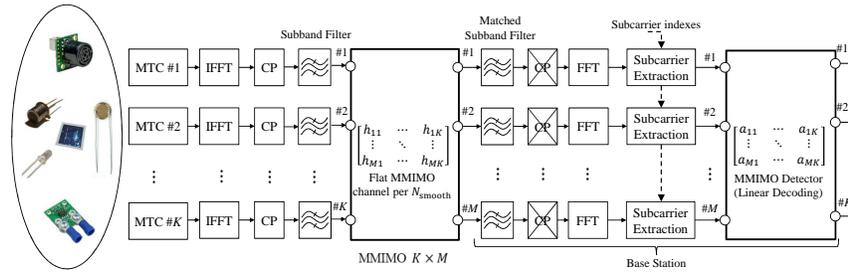


Fig. 1. Exemplary block diagram of a Massive MIMO uplink for mixed networks, where the BS simultaneously delivers narrowband services to MTC devices and wideband services to regular UEs. The cluster of MTC devices seen at the transmit side share the same PRBs in frequency and time dimensions, while the sole BS at the receive side is equipped with an antenna array at least one order of magnitude larger than the number of served MTC devices.

F. Detection

When it comes to data stream separation in conventional systems, Maximum Likelihood (ML) detection is the optimum solution but its complexity grows exponentially with the number of streams (this makes it hard to implement in MTC networks where hundreds to thousands of devices are envisioned). This is the reason why parameter estimation and detection are key problems in Massive MIMO systems. Sub-optimal linear filtering detectors with reduced computational complexity, such as MRC, ZF, and MMSE [37], offer lower costs (that do not depend on the number of streams/users and modulation order), but are not capable of achieving the full receive-diversity order of ML detection and consequently, they do not achieve the channel sum capacity for cases where the number of streams/users is approximately equal or equal to the number of antennas [67]. This performance-complexity tradeoff led to the development of several alternative detection methods, some of them are discussed in the sequel.

The first class of interest is iterative linear filtering, which encompasses MMSE with Successive Interference Cancellation (MMSE-SIC) and Block-iterative Generalized Decision Feedback Equalization (BI-GDFE) [68]. A shortcoming common to such iterative detectors is that their reliance on repeated matrix inversions may render them computationally heavy for large array sizes. Tabu Search (TS) [69] and Likelihood Ascent Search (LAS) [70] belong to a class of matrix-inversion free detectors known as random step search detection methods. Regrettably, the performance-complexity tradeoff comes into play also here, as both TS and LAS are known to be outperformed by MMSE-SIC [43]. Additionally, MMSE-SIC is known to achieve the sum capacity of the fast fading MIMO multiple-access channel [71, 72]. The last relevant class, referred to as tree-based detection algorithms, has in Fixed Complexity Sphere Decoding (FCSD) one of its most prominent methods [73, 74]. Notwithstanding the improvements of FCSD over standard sphere decoding, the method is still 1,000 times more complex than TS.

III. SYSTEM MODEL

This section describes the system depicted in Figure 1 in terms of its underlying functional blocks. In what follows, we assume that the transmitted signals of a cluster with K single-antenna MTC devices are detected by a Massive MIMO BS

equipped with M receive antennas, where $M \gg K$. Single-antenna MTC devices are simple, inexpensive, power-efficient, and each device normally has low to moderate throughput.

A. Signal Generation & Transmission

Consider a PNSCH that is available and exclusively dedicated to services related to sporadic MTC traffic. The K MTC devices map data into a set of continuous PRBs in the frequency domain, with the subcarrier indexes providing the spectral position of the PNSCH at the physical layer level.

Random access is not the focus of this work. In this work we focus on the issue posed by the necessity of the BS to serve hundreds to thousands of MTC devices with limited number of resources. However, several works in the literature deal with the problem of the random access of a large (or massive) number of devices [75]–[78]. Therefore, in this work we assume that all MTC devices being served by a BS are already synchronized and connected to it before accessing the PNSCH, *i.e.*, all the MTC devices being served have already performed random access and attach procedures before any data is sent through the PNSCH.

The PNSCH is configured at the BS via broadcasting system information blocks (SIB), just like with the Physical Random Access Channel (PRACH) used in current 4G systems (see, *e.g.* [79] and the references therein). This allows the number of PNSCH transmission opportunities in the uplink to be scheduled while taking into consideration discrepancies between the (likely different) capacities of MTC devices and regular UEs. The PNSCH time-frequency resources are semi-statically allocated by the BS, and are repeated periodically. Additionally, the SIB messages can carry, for example, information about the pilot sequence length, which in turn, dictates the useful capacity of the PNSCH as it will, consequently, determine the remaining period of time destined to data symbols. The fraction of pilot and data symbols can be selected based on the network traffic characteristics of a cell or set of them, *i.e.*, the PNSCH configuration can be modified to increase the number of served MTC devices or the throughput of the ensemble (*i.e.*, the data rate of all MTC devices served through the same time-frequency resources) by increasing the number of OFDM symbols used for pilot or data transmission accordingly. Therefore, the pilot sequence length can be varied so that more MTC devices can be simultaneously served by the BS at the expense of smaller data capacity.

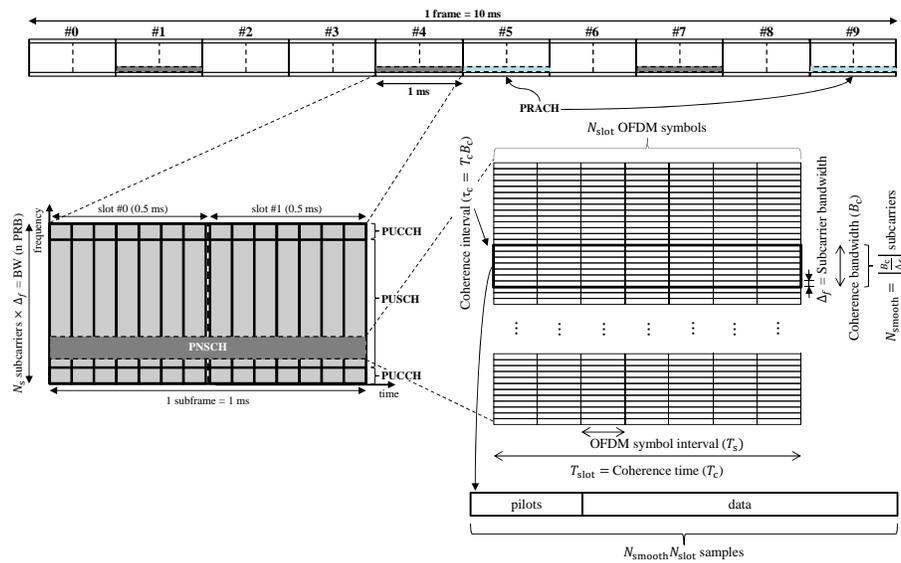


Fig. 2. Uplink frame structure with PNSCH time-frequency plane.

We assume the utilization of Orthogonal Frequency-Division Multiplex (OFDM) block-based transmissions where each MTC device transmits its signal (*i.e.*, allocated pilot sequence and data) by taking the Inverse Fast Fourier Transform (IFFT) of the mapped information (*i.e.*, pilot and data symbols), and subsequently adding a Cyclic Prefix (CP). It is important to remember that all MTC devices transmit their signals at the same time and frequency resources. We denote the OFDM symbol interval by T_s , the subcarrier spacing by Δ_f , the useful symbol duration by $T_u = 1/\Delta_f$, and the guard interval (*i.e.*, duration of the cyclic prefix) by $T_g = T_s - T_u$. As in [36], we call the reciprocal of the guard interval, when measured in subcarrier spacings, the *frequency smoothness interval*,

$$N_{\text{smooth}} = \frac{1}{T_g \Delta_f} = \frac{T_u}{T_g}, \quad (1)$$

where N_{smooth} represents the number of subcarriers over which the channel frequency response is considered smooth, *i.e.*, approximately constant [80].

A total of τ_p OFDM symbols are entirely used for transmitting pilot sequences. The remaining symbols, τ_u , within the same coherence interval (or coherence block) are used for data transmission. A coherence interval is a time-frequency space with duration equal to the coherence time, T_c , and bandwidth equal to the coherence bandwidth, B_c , see Figure 2. In general, the channel response is constant over N_{smooth} consecutive subcarriers and, therefore, the BS can estimate the channel for a total of $K_{\text{max}} = \tau_p N_{\text{smooth}}$ terminals. We assume that a coherence interval consists of N_{smooth} subcarriers and $\tau_p + \tau_u$ OFDM symbols, *i.e.*, $N_{\text{smooth}} \times (\tau_p + \tau_u)$ subcarriers, over which the channel response can be approximated as being constant and flat-fading [26].

The modulated symbols (*i.e.*, the symbols carrying data of a MTC device) are assumed to be randomly and independently drawn from a digital modulation alphabet (*e.g.*, QPSK, 16QAM, etc.) with normalized average energy. The modulated symbols are mapped into τ_u OFDM symbols. The total number

of data symbols that can be transmitted during a coherence interval is equal to $\tau_u N_{\text{smooth}}$.

Figure 2 depicts the uplink frame structure devised for the PNSCH. As can be seen in the figure, we assume 1 ms long PNSCH transmission opportunities, however, it is important to point out that other multiple intervals of 1 ms could also be used, allowing more MTC devices and/or higher data rates. The figure also shows the time-frequency plane for one possible configuration of the PNSCH (*i.e.*, frequency position within the resource grid and time periodicity). It shows how pilots and data symbols are mapped into the time-frequency domain of a coherence interval. A 0.5 ms slot consists of N_{slot} consecutive OFDM symbols, where each of them has N_s subcarriers. If we assume that $T_c = T_{\text{slot}}$, then, a coherence interval is composed of N_{slot} OFDM symbols and N_{smooth} consecutive subcarriers. The number of coherence intervals in a slot when $T_c = T_{\text{slot}}$ is given by N_s/N_{smooth} . As shown in the figure, the time-frequency plane can be divided into several coherence intervals in which each massive MIMO channel is considered time-invariant and frequency-flat.

As an example of the possible PNSCH capacity, if we consider it has 6 PRB allocated to it over the interval of a 1 ms long subframe (where each PRB is equal to 12 subcarriers and the subframe contains 14 OFDM symbols), a periodicity of 10 ms, meaning that MTC devices transmit for 1 ms every 10 ms, $N_{\text{smooth}} = 12$, $\tau_p = 50$, and $\tau_u = 118$, totalizing 12 subcarriers \times 14 OFDM symbols = 168 subcarriers/coherence interval, as showed in Figure 4. Therefore, for this setup, 300 different MTC devices could be served over the 6 allocated PRBs (*i.e.*, 50 MTC devices per PRB), where each one of the devices can have a data rate of 70.8 Kbits/s considering 64QAM modulation. The throughput can be doubled (*i.e.*, 141.6 Kbits/s) if the PNSCH periodicity is decreased to 5 ms.

The PNSCH is time- and frequency-multiplexed with Physical Uplink Shared Channel (PUSCH), Physical Uplink Control Channel (PUCCH) and PRACH as depicted in Figure 2.

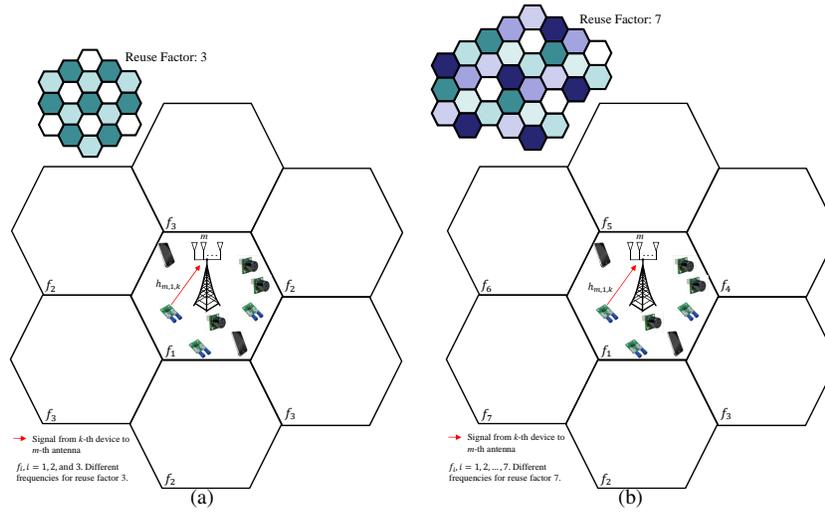


Fig. 3. Cell deployments with less aggressive reuse factors: (a) frequency reuse factor equal to 3 and (b) frequency reuse factor equal to 7.

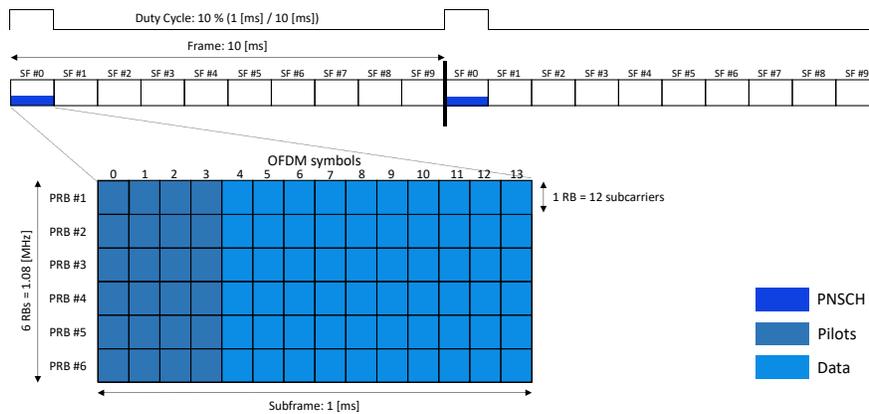


Fig. 4. PNSCH capacity example.

Therefore, as can be seen in Figure 1, filters are added to both transmission and reception chains. These filters are added to the processing chains so that out-of-band emissions (OOBE), which are intrinsic to the OFDM waveform due to the discontinuities at its edges, do not interfere with adjacent channels, *i.e.*, PRACH, PUSCH and PUCCH. Additionally, the filters help to mitigate inter-symbol interference (ISI) and inter-carrier interference (ICI) caused by asynchronous transmissions coming random access attempts happening at the PRACH [81].

Note that the received signal is passed through a matched filter, which maximizes the signal-to-noise ratio (SNR). The filter is applied to each time domain OFDM symbol (*i.e.*, after IFFT and CP insertion) to mitigate the OOBE of the PNSCH transmissions. The filters should be carefully designed to (i) maintain the complex-domain orthogonality of OFDM symbols, (ii) exhibit flat passband over the subcarriers in the PNSCH, (iii) have sharp transition band in order to reduce the guard-bands, and (iv) present sufficient stop-band attenuation [22, 82].

Additionally, we assume that inter-cell interference is negligible. Inter-cell interference can be heavily mitigated, and therefore, considered insignificant, if less-aggressive

frequency-reuse (*e.g.*, reuses of 3, 7 or higher) is adopted [83]. Figure 3 depicts reuse factors equal to 3 and 7. Inter-cell interference manifests itself in two ways, namely, coherent and non-coherent interference, being the former caused by contaminating cells (*i.e.*, cells that use the same set of pilots as the home cell, causing pilot-contamination) and the latter caused by non-contaminating cells (*i.e.*, cells that do not use the same pilots as the home cell) [80]. If a less aggressive frequency reuse is not possible or desired, then, pilot-contamination (*i.e.*, coherent interference) can be mitigated if not eliminated by making the PNSCH time-frequency resource intervals in each one of the neighbor cells different from the intervals chosen for the target cell. However, this scheme does not eliminate non-coherent interference, once the neighbor cells will be using the same frequency for other uplink channels (*e.g.*, PRACH, PUCCH, PUSCH).

B. The Massive MIMO Channel and its Key Properties

Let $h_{m,k,n}$ denote the complex channel propagation coefficient from the k -th MTC device to the m -th antenna of the BS in the n -th subcarrier

$$h_{m,k,n} = g_{m,k,n} \sqrt{d_k}, \quad (2)$$

where $g_{m,k,n}$ is the complex small-scale fading coefficient, and d_k is the amplitude coefficient that accounts for geometric attenuation and shadowing, *i.e.*, large-scale fading [36]. The large-scale fading coefficients are assumed constant with respect to both subcarrier number and BS' antenna index since the geometric and shadow fading change slowly over space [23]. Therefore, between any given BS and a MTC device, there is only one large-scale fading coefficient. Additionally, these coefficients change only when a MTC device significantly changes its position in relation to the serving BS. It is generally assumed that in the radius of 10 wavelengths, the large-scale fading coefficients are approximately constant. With respect to the small-scale fading coefficients, they significantly change as the MTC device moves by a quarter of the wavelength. Therefore, it is normally assumed that the large-scale fading coefficients change about 40 times slower than the small-scale fading coefficients [84].

In general, the bandwidth occupied by an OFDM symbol is greater than the channel coherence bandwidth, B_c , however, on the other hand, the subcarrier bandwidth (Δ_f) is smaller than B_c . Therefore, as mentioned earlier, the channel can be considered flat (*i.e.*, constant) over N_{smooth} subcarriers. Additionally, we also consider that the channel is rich in scatterers and has no Line-of-Sight (LOS) component. Therefore, it is natural to assume that the transmitted signal within a coherence interval undergoes flat i.i.d. Rayleigh fading. The justifications for the use of this model are as follows: (i) it is approximately correct under conditions of dense scattering and (ii) it enables a comprehensive performance analysis [80].

The complex random channel responses within one coherence interval are statistically identical to the ones in any other coherence interval, irrespective of whether they are separated in time and/or frequency. Hence, the channel fading can be described by a stationary ergodic random process. Therefore, hereafter, our analysis is carried out by studying a single statistically representative coherence interval [80]. We assume that the channel realizations are independent between any pair of intervals, which is known as a block/interval fading assumption in the literature [23, 80]. Consequently, for notational simplicity we suppress the dependency of $h_{m,k,n}$ on the subcarrier index, n , and rewrite it as $h_{m,k}$.

Therefore, within a coherence interval, the $M \times K$ channel matrix \mathbf{H} can be expressed as

$$\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \dots \ \mathbf{h}_K] = \mathbf{G}\mathbf{D}^{1/2} = \underbrace{\begin{pmatrix} g_{1,1} & \dots & g_{1,K} \\ \vdots & \ddots & \vdots \\ g_{M,1} & \dots & g_{M,K} \end{pmatrix}}_{\mathbf{G}} \cdot \underbrace{\begin{pmatrix} \sqrt{d_1} & & \\ & \ddots & \\ & & \sqrt{d_K} \end{pmatrix}}_{\mathbf{D}^{1/2}}, \quad (3)$$

where the elements $h_{m,k} = g_{m,k} \cdot \sqrt{d_k}$ correspond to the complex channel gains from the MTC transmit antenna to the BS receive antennas where $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}_M, b_k \mathbf{I}_M), \forall k$.

Small-scale fading is one of the major impairments in wireless communications [91]. This kind of fading is created by microscopic changes in the propagation environment and

causes the channel gain to randomly fluctuate. The random fluctuation of the channel gain will occasionally make the channel gain very small, which in consequence causes the transmitted data to be received in error, *i.e.*, the random fluctuation makes the channel unreliable [91].

As we will discuss next, the diversity obtained by transmitting a signal over several channels with independent realizations is essential to mitigate small-scale fading and noise. Therefore, the exploitation of spatial diversity becomes very interesting, once it can be achieved by simply deploying several antennas at the receiver or transmitter.

Under the assumptions of large M and that the small-scale fading coefficients experienced by each MTC device are i.i.d. random variables with zero mean and unitary variance, the column channel vector from different MTC devices becomes asymptotically orthogonal as the number of receive antennas at the BS grows without bound [36]. On the other hand, as the number of antennas increases, the channels between the MTC devices and the BS start behaving as if they were non-fading channels, *i.e.*, as if they were almost deterministic scalar channels, after combining/precoding [85]. This phenomenon is attributed to the spatial diversity obtained from having multiple receive antennas that observe independent fading realizations, which are improbable to all be zero or nearly zero simultaneously [26].

These two key properties exhibited by Massive MIMO channels are known as *favorable propagation* and *channel hardening* [83] respectively. Favorable propagation implies that noise and interference can be canceled out with simple linear detectors such as the MRC, achieving optimal performance [86]. Additionally, the favorable propagation idea provides a way to quantify the ability of a BS equipped with a large number of antennas to separate the data streams of multiple MTC devices [86]. Channel hardening makes the channel variations asymptotically reduce as the number of antennas increases, in the sense that the normalized instantaneous channel gain, $\|\mathbf{h}_k\|^2/M$, converges to the deterministic average channel gain, d_k . With channel hardening, as the small-scale fading tends to vanish, there is no need to adapt the power allocation or scheduling to its fluctuations, simplifying the resource and power allocation algorithms [87]. Moreover, channel hardening provides improved reliability due to the almost deterministic channel, which consequently results in lower latency due to a smaller number of re-transmissions. Both effects are a direct consequence of the law of large numbers, *i.e.*, they are only valid in the context of Massive MIMO when $M \rightarrow \infty$ [83]. Therefore, if the channel offers asymptotically favorable propagation and channel hardening, then we have

$$\frac{\mathbf{H}^H \mathbf{H}}{M} = \mathbf{D}^{1/2} \frac{\mathbf{G}^H \mathbf{G}}{M} \mathbf{D}^{1/2} \approx \mathbf{D}^{1/2} \mathbf{I}_K \mathbf{D}^{1/2} \rightarrow \mathbf{D}, \quad M \rightarrow \infty \quad (4)$$

where $(\cdot)^H$ denotes transpose-conjugate (Hermitian) operation. Equation (4) mathematically summarizes both phenomena showing that when $M \rightarrow \infty$, different channels become mutually orthogonal and the channel gains tend to their respective large-scale fading coefficients.

Figure 5 illustrates both the channel hardening and favor-

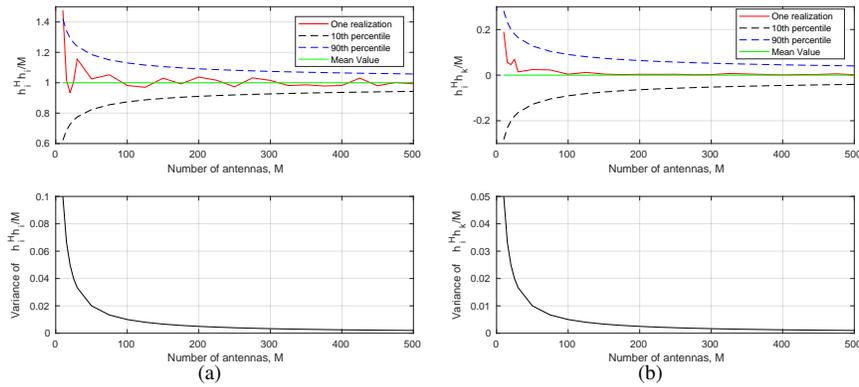


Fig. 5. Massive MIMO properties: (a) asymptotic channel hardening and (b) asymptotic favorable propagation

able propagation phenomena for an M -dimensional channel $\mathbf{h}_i \sim \mathcal{CN}(\mathbf{0}_M, \mathbf{I}_M)$, $\forall i$. The figure shows the mean values, the 10% and 90% percentiles and random realizations for both phenomena and for different numbers of antennas. As can be seen, in the channel hardening case, the normalized instantaneous channel gain $h_i^H h_i / M$ approaches its average value of 1 and the standard deviations and variance reduce as M increases. In the favorable propagation case, the inner product of the normalized channels $h_i^H h_k / \sqrt{M}$ and h_k / \sqrt{M} , tends to zero and the standard deviations and variance also reduce as M increases.

We refer the interested reader to [26, 80, 88] for a detailed discussion on these phenomena, and to [89] for experimental evidence supporting the assumption of i.i.d. Rayleigh small-scale fading coefficients with zero mean in Massive MIMO settings. In [80] the authors demonstrate that two fundamentally different channel models, namely, independent Rayleigh fading (*i.e.*, isotropic scattering) and Uniformly Random LOS, offer approximately favorable propagation and channel hardening properties. These two models represent quite extreme cases, and actually, the real channel is more likely to be something (*i.e.*, exhibit characteristics) between these two extremes. Therefore, it is reasonable to expect that the assumptions of favorable propagation and channel hardening hold for most of the practical cases. This has been experimentally confirmed by several independent measurement campaigns [24, 43, 93]–[96].

C. Capacity Lower and Upper Bounds in Favorable Propagation

The favorable propagation phenomenon does not only provide optimal performance with sub-optimal linear processing (see equation (4)) but also constitutes the most valuable scenario from the perspective of maximizing the achievable rate. In order to find lower and upper bounds for a channel offering favorable propagation we employ a capacity argument where we consider the uplink direction and a fixed and deterministic channel, \mathbf{H} . The sum capacity for this channel is defined by [80]

$$R = \log_2 |\mathbf{I}_M + \rho \mathbf{H} \mathbf{H}^H|, \quad (5)$$

where we assume the BS knows the channel \mathbf{H} and that the MTC devices know their respective individual rates [80].

In order to determine the lower and upper bounds that (5) can assume for distinct characteristics of \mathbf{H} , we rewrite it as

$$\begin{aligned} R &= \log_2 |\mathbf{I}_M + \rho \mathbf{H} \mathbf{H}^H| \\ &\stackrel{(a)}{=} \log_2 |\mathbf{I}_K + \rho \mathbf{H}^H \mathbf{H}| \\ &\stackrel{(b)}{\leq} \log_2 \left(\prod_{k=1}^K [\mathbf{I}_K + \rho \mathbf{H}^H \mathbf{H}]_{kk} \right) \\ &= \sum_{k=1}^K \log_2 (1 + \rho \|\mathbf{h}_k\|^2), \end{aligned} \quad (6)$$

where in (a) we used Sylvester's determinant theorem and in (b) we used the Hadamard inequality, and ρ is the power transmitted per MTC device. The equality in (b) of (6) holds if and only if $\mathbf{H}^H \mathbf{H}$ is a diagonal matrix (*i.e.*, the channel matrix \mathbf{H} has mutually orthogonal columns), which is the case when the channel exhibits favorable propagation [26]. Next, we find lower and upper bounds for (6) when favorable propagation is assumed and the constraint $\|\mathbf{H}\|_F^2 = \sum_{k=1}^K \|\mathbf{h}_k\|^2 = \sum_{k=1}^K \lambda_k^2 = MK$ (*i.e.*, the columns of \mathbf{H} have the same norm) is applied to the channel matrix, where $\{\lambda_k\}$ are the singular values of the channel matrix \mathbf{H} . Notice that when we have favorable propagation, the singular values, $\{\lambda_k\}$, are equal to the channel norms, $\{\|\mathbf{h}_k\|\}$, which are in consequence equal to $\{Md_k\}$. Here we consider that the channel gains, $\{d_k\}$, are all equal to 1.

The lower bound of (6), given that we have favorable propagation, is achieved when \mathbf{H} has its rank equal to one, which corresponds to an LOS channel. The rank of \mathbf{H} is equal to the number of non-zero singular values of \mathbf{H} and determines how many data streams can be simultaneously multiplexed over the channel. This is expressed by, for example, when $\|\mathbf{h}_1\|^2 = MK$, $\|\mathbf{h}_2\|^2 = \dots = \|\mathbf{h}_K\|^2 = 0$, which meets the constraint $\|\mathbf{H}\|_F^2 = MK$. Consequently, the lower bound of (6) is given by

$$R_{\text{lower bound}} = \log_2 (1 + \rho MK). \quad (7)$$

The upper bound of (6), given that we have favorable propagation, is achieved when \mathbf{H} has full-rank, *i.e.*, its rank is equal to K , and $\|\mathbf{h}_k\|^2 = M$, $\forall k$, meaning that the channel matrix is well-conditioned [26], once $\lambda_{\min} = \lambda_{\max} = \sqrt{M}$, which are the minimum and maximum singular values of

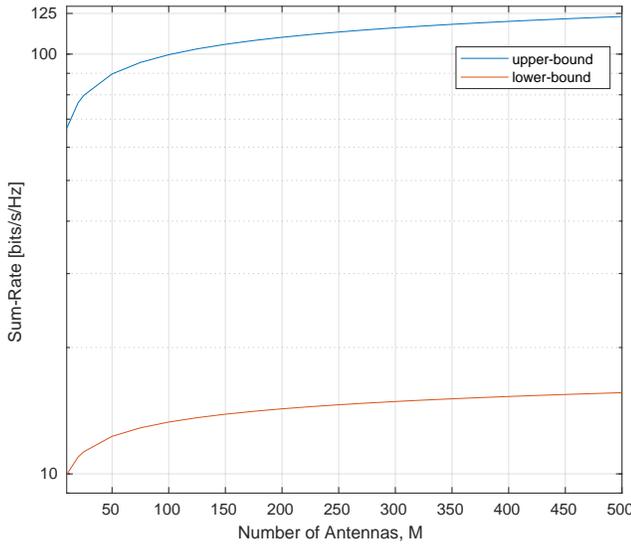


Fig. 6. Capacity lower and upper bounds for several number of antennas, M , when the channel offers favorable propagation.

$\mathbf{H}^H \mathbf{H}$ respectively [86]. Therefore, as can be seen, a full-rank and well-conditioned channel matrix attains the highest possible capacity. Next, in order to find the upper bound for the capacity, we apply Jensen's inequality and the constraint $\|\mathbf{H}\|_F^2 = \sum_{k=1}^K \|\mathbf{h}_k\|^2 = MK$ to equation (6),

$$\begin{aligned} R &\leq \sum_{k=1}^K \log_2(1 + \rho \|\mathbf{h}_k\|^2) = K \left[\frac{1}{K} \sum_{k=1}^K \log_2(1 + \rho \|\mathbf{h}_k\|^2) \right] \\ &\leq K \log_2 \left(1 + \frac{\rho}{K} \sum_{k=1}^K \|\mathbf{h}_k\|^2 \right) = K \log_2 \left(1 + \frac{\rho}{K} \|\mathbf{H}\|_F^2 \right) \\ &= K \log_2(1 + \rho M). \end{aligned} \quad (8)$$

Therefore, the upper bound,

$$R_{\text{upper bound}} = K \log_2(1 + \rho M), \quad (9)$$

is achieved only when the columns of \mathbf{H} are mutually orthogonal (*i.e.*, the channel offers favorable propagation), making the equality in the first line of (8) hold true, and when $\{\mathbf{h}_k\}$ have the same norm, making the equality in the second line of (8) hold true as well.

Finally, summarizing, under the constraint $\|\mathbf{H}\|_F^2 = MK$, the achievable rate for a channel exhibiting favorable propagation is bounded as

$$\log_2(1 + \rho MK) \leq R \leq K \log_2(1 + \rho M). \quad (10)$$

Figure 6 shows the capacity lower and upper bounds for several number of antennas, M , $K = 10$ MTC devices, a channel offering favorable propagation and the application of the constraint $\|\mathbf{H}\|_F^2 = MK$ to the columns of the channel matrix.

D. Signal Detection

The vector received at the BS can be written as [24, 83]

$$\begin{aligned} \mathbf{y} &= \sqrt{\rho} \mathbf{H} \mathbf{x} + \mathbf{w} \\ &= \sqrt{\rho} \sum_{k=1}^K \mathbf{h}_k x_k + \mathbf{w}, \end{aligned} \quad (11)$$

where ρ is the average uplink transmit power of each MTC device (UL Tx power), $\mathbf{y} \in \mathbb{C}^{M \times 1}$, and $\mathbf{w} \in \mathbb{C}^{M \times 1}$ is a zero-mean noise vector with complex Gaussian distribution and identity covariance matrix, *i.e.*, $\mathcal{CN}(\mathbf{0}_M, \mathbf{I}_M)$. As the noise variance of all elements of \mathbf{w} are assumed to be equal to 1, thus, ρ can be interpreted as a normalized transmit SNR and consequently is dimensionless. There exist M PNSCH signal versions in (11) for each one of the K MTC devices. Hence, the task of the BS consists of detecting K simultaneous MTC transmissions on the basis of estimates of the channel coefficients in (3). Here we consider that the BS has full knowledge on the channel coefficients. Therefore, detection techniques need to be employed in order to separate each of the data streams transmitted by the various devices in a Massive MIMO system.

Maximum Likelihood (ML) multi-user detection is the optimum detection technique for the uplink but it is highly complex. Its complexity grows exponentially with the number of MTC devices, K , making it hard to implement in our case where hundreds to thousands of MTC devices are envisioned to be served by the BS over the same time-frequency resources. To circumvent this limitation, we discuss in the next section a couple of sub-optimal alternatives with reduced computational complexity [37].

IV. SUB-OPTIMAL MASSIVE MIMO DETECTION

When it comes to separation of data streams in conventional MIMO systems (*i.e.*, where M is small), ML multi-user detection is the optimal solution but its complexity grows exponentially with the number of MTC devices. This detector works by trying all possible transmitted data vectors, \mathbf{x} , and selects the one that minimizes the following equation:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in X^K} \|\mathbf{y} - \sqrt{\rho} \mathbf{H} \mathbf{x}\|^2, \quad (12)$$

where X is the set with all possible data symbol vectors, \mathbf{x} . The BS has to search over all possible data vectors the one that minimizes (12), which as can be noticed, exponentially increases with the number of MTC devices, K (*e.g.*, for 64 QAM modulation with an alphabet of 6 symbols and 10 MTC devices the BS would have to check $6^{10} = 60466176$ possibilities). This is the reason why signal detection is a key problem in Massive MIMO systems. In the sequel we overview the literature on the subject, and justify our choices for the detectors used in the simulation work presented later on in Section VI.

A. Linear Detection Methods

As mentioned before, we consider here the case where the BS has perfect Channel State Information (CSI), *i.e.* \mathbf{H} is perfectly known at the BS, and that the channels are i.i.d.

Rayleigh fading. Let \mathbf{A} be an $M \times K$ linear detector matrix that depends on the channel \mathbf{H} . By using a linear detector, the received signal can be separated into different data streams using \mathbf{A}^H as follows

$$\mathbf{r} = \mathbf{A}^H \mathbf{y}, \quad (13)$$

where the vector \mathbf{r} collects the data streams received at the BS, *i.e.*, the symbols of all K single-antenna MTC devices, and \mathbf{A} is a receive matrix that depends on the specific linear detector used at the BS. After matched filtering, CP removal, FFT processing and subcarrier extraction within each one of the OFDM symbols, as seen in Figure 1, a $M \times 1$ vector consisting of transmissions from all the K MTC terminals in the cell undergoes linear detection in order to retrieve the data symbols from all devices.

Inspection of (3) reveals that \mathbf{D} is a diagonal matrix, which means we can use Maximum Ratio Combining (MRC) in the uplink to separate the signals from different MTC devices into different streams with asymptotic no inter-user interference as $M \rightarrow \infty$ [36]. Thereby each MTC device's transmission can be seen as the signal of a single device passing through a Single Input Single Output (SISO) channel. In the limit, this implies that MRC is optimal when the number of receive antennas is much larger than the number of transmit antennas, *i.e.* $M \gg K$, $M \rightarrow \infty$ – as can be seen from (4). In MRC the linear detection matrix \mathbf{A} is chosen using

$$\mathbf{A}_{\text{MRC}} = \mathbf{H}, \quad (14)$$

where the dominant computation is due to matrix transposition. For the MRC detector, the BS aims at maximizing the received SNR of each one of the K streams, but ignoring the effect of multi-user interference. The associated complexity is of only $\mathcal{O}(MK)$ multiplications. As mentioned earlier, as M increases, MRC asymptotically becomes optimum in the sense that multi-user interference is totally removed. More specifically, the received signal is multiplied by the conjugate-transpose of the channel vector, \mathbf{h}_k as follows

$$r_k = \mathbf{h}_k^H \mathbf{y} = \sqrt{\rho} \|\mathbf{h}_k\|^2 x_k + \sqrt{\rho} \sum_{i=0, i \neq k}^K \mathbf{h}_k^H \mathbf{h}_i x_i + \mathbf{h}_k^H \mathbf{w}. \quad (15)$$

One advantage of the MRC detector is that its the signal processing is very simple since the BS just multiplies the received vector with the conjugate-transpose of the channel matrix \mathbf{H} . On the other hand, one disadvantage is that once it disregards the effect of multi-user interference it performs poorly in interference-limited scenarios.

Zero Forcing (ZF) detection is an alternative linear detection method that chooses \mathbf{A} with the aim of completely eliminating inter-user interference, regardless of noise enhancement, *i.e.*, it takes the inter-user interference into account, but neglect the effect of noise. Specifically, the ZF detector chooses \mathbf{A} constrained to $\mathbf{A}\mathbf{H} = \mathbf{I}$

$$\mathbf{A}_{\text{ZF}} = \mathbf{H}(\mathbf{H}^H \mathbf{H})^{-1}, \quad (16)$$

which is of complexity $\mathcal{O}(MK + MK^2 + K^3)$ [89]. One drawback of ZF is that it insists in forcing interference to zero independent of the interference strength, *i.e.*, any energy

of the signal of interest that lies in the interference subspace is discarded. More specifically, the received vector, \mathbf{y} , is multiplied by the pseudo-inverse of the channel matrix \mathbf{H} , as follows

$$\mathbf{r} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{y} = \sqrt{\rho} \mathbf{x} + (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{w}. \quad (17)$$

For the ZF detector to work, it requires that $M \geq K$, otherwise $(\mathbf{H}^H \mathbf{H})$ is not invertible. It is noticeable that each stream of \mathbf{r} in (17) is free of multi-user interference. One advantage of ZF detection is that it also has simple signal processing. As disadvantages we can mention that once it neglects the effect of noise, it performs poorly in noise limited scenarios and its implementation complexity is higher when compared to MRC, once it has to compute the pseudo-inverse of the channel gain matrix.

A better strategy would be to choose \mathbf{A} so as to balance the signal energy lost with the increased interference. From this point of view, it is much better to accept some residual interference provided that this allows the detector to capture more of the desired signal's energy [37]. Additionally, if the channel, \mathbf{H} , is ill-conditioned (*i.e.*, it has a high *condition number* [26] where (13) has no or an infinite set of non-unique solutions), then the pseudo-inverse in (16) augments the noise power's, decreasing significantly ZF's performance.

One last linear detector that, together with MRC and ZF, poses complexity cost that does not depend on the number of MTC devices, K , is the Minimum Mean Squared Error (MMSE). As the name suggests, the MMSE detector chooses the \mathbf{A} that minimizes $e = E[\|\mathbf{A}^H \mathbf{y} - \mathbf{x}\|^2]$ without any additional constraints

$$\mathbf{A}_{\text{MMSE}} = \mathbf{H} \left(\mathbf{H}^H \mathbf{H} + \frac{1}{\rho} \mathbf{I} \right)^{-1}. \quad (18)$$

In contrast to ZF, which minimizes multi-user interference but fails to treat noise, and to MRC, which minimizes noise but fails to treat interference, MMSE achieves an optimal balance between interference suppression and noise enhancement at a cost similar to ZF [66, 89]. At high SNR (*i.e.*, high ρ), the ZF detector approaches the performance of the MMSE detector, while at low SNR, the MRC detector performs as well as MMSE detector [89]. The MMSE detector exhibits computational complexity similar to that of the ZF detector when the SNR, which is represented by its inverse in the second term, $\frac{1}{\rho} \mathbf{I}$, inside the parentheses of (18), is known. Otherwise, the complexity of the algorithm used to estimate the SNR has also to be taken into account.

B. Discussion

Table I summarizes the solutions discussed in this section as potential candidates for data stream separation (*i.e.*, detection) in the uplink of Massive MIMO systems. The shortcomings listed under iterative detection, random step search, and tree-based methods suggest that these detection classes perform really well for $M \approx K$ but are too complex to be practical in the cases where hundreds to thousands of devices are served by the BS and/or $M \gg K$. This indicates that more work is

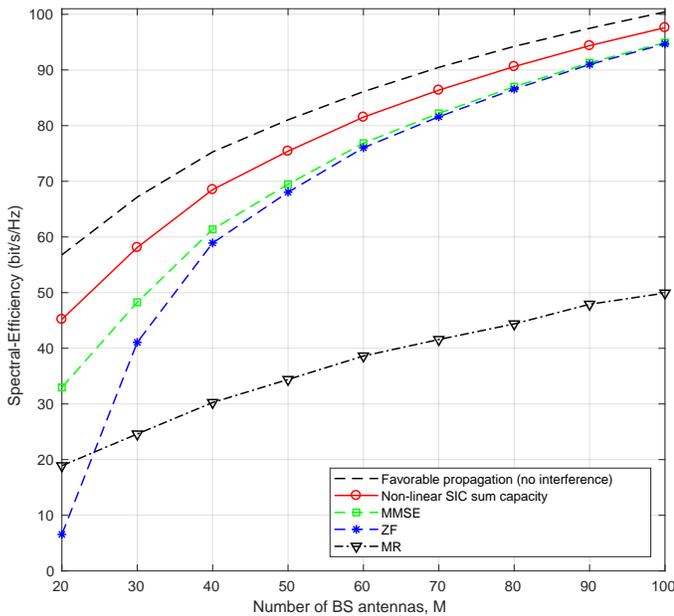


Fig. 7. Comparison of the sum spectral efficiency for i.i.d. Rayleigh fading channels with non-linear and linear signal processing and variable number of antennas, M .

needed on this matter, perhaps towards turbo codes or Low-density Parity-check (LDPC) codes in iterative detection and decoding settings [67].

Under perfect CSI knowledge, non-linear signal detection schemes such as SIC are known to achieve the uplink sum capacity [25, 71, 72]. SIC achieves the optimal performance by iteratively removing interference during the decoding process [72]. However, as mentioned earlier, this scheme requires precise CSI and a considerable number of computations. If CSI is not accurate enough, the interference removal process brings more problems than benefits [25]. On the other hand, linear detectors, such as MRC, ZF and MMSE, can only reject interference through the application of linear transformations (*i.e.*, projections) to the received signal.

Therefore, one pertinent question that comes to our mind is how better is the performance of optimal non-linear SIC detection when compared to that of sub-optimal linear detection schemes such as MRC, ZF and MMSE? In order to assess this question, we provide a quantitative numerical comparison in Figure 7. In the figure, we consider the sum capacity of a single-cell (*i.e.*, no inter-cell interference) under the availability of perfect CSI at the BS. The BS is equipped with M antennas and simultaneously serves $K = 20$ MTC devices. The channels are modeled as i.i.d. Rayleigh fading ($\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}_M, \mathbf{I}_M), \forall k$) and the SNR is set to -5 dB. The figure depicts the average sum spectral efficiency for different number of antennas, M , achieved by MRC, ZF and MMSE sub-optimal linear detectors and the sum capacity-achieving SIC non-linear detector.

The results in Figure 7 show that linear processing is outperformed by non-linear SIC processing when $M \approx K$. As can be seen, MMSE detection exhibits performance that is far better than that of MRC and ZF detectors when $M \approx K$, however, it is still more than 12 bits/s/Hz lower than that presented by

the SIC detector. Nevertheless, the performance gap between linear and non-linear detection schemes decreases quickly with M , once the channels become mutually orthogonal due to the favourable propagation (FP) phenomenon (*i.e.*, the channels decorrelate and then noise and interference among MTC devices vanish). As can be seen, as M increases, all the curves asymptotically approach the FP curve. Additionally, it is also important to notice that non-linear detection only outperforms linear detection when $M \approx K$, while the gap becomes small when $M/K > 2$, approaching the FP curve as $M \rightarrow \infty$ [36]. These results demonstrate that a BS equipped with a large number of antennas (*i.e.*, $M \gg K$) could serve each one of the K MTC devices as if they were operating alone in the cell. It is interesting to see that the same capacity as with SIC can be achieved by using ZF and MMSE detection with the addition of a few more antennas (*e.g.*, around 10 additional antennas for this simulation setup). The addition of a few more antennas is a low price to pay for the reduced complexity of ZF or MMSE detectors when compared to the complexity of, for example, the MMSE-SIC detector [25]. Finally, it is also important to mention that the results showed in Figure 7 are also valid for the downlink with DPC, due to the duality explained in [90]. Therefore, as linear detectors asymptotically achieve the sum capacity as $M \rightarrow \infty$, we consider them in our simulation work.

Another important point of discussion is the comparison of the achievable capacity against the transmission power. Figure 8 depicts the achievable sum-rates for MRC, ZF, and MMSE detectors versus different SNR values, *i.e.*, ρ , with $M = 20$ and $K = 20$. These curves are computed by considering that the elements of \mathbf{H} are i.i.d. Gaussian distributed with zero mean and unit variance. As expected, the MMSE detector performs better than ZF and MRC detectors for all SNR values. Comparing these results with those in Figure 7 we notice that the spectral efficiency can be hugely improved by just adding antennas to the BS without the necessity to increase the transmitted power, *e.g.*, the SE for the ZF detector with $M = 20$ is 6.56 bits/s/Hz, however, just by adding 10 more antennas is grows to 41 bits/s/Hz. As we will shown in the next section, the transmitted power can, in fact, be decreased as the number of antennas grows.

V. ACHIEVABLE RATES

In this subsection we present lower bounds on the achievable rate for the linear detectors discussed in the previous section when perfect channel-state information (CSI) is assumed. Considering that a less-aggressive frequency-reuse factor (*e.g.*, reuse factor of 3, 4, 7 or higher) is adopted, we can analyze and derive the achievable rates as if the target cell was a single-cell system, emphasizing the fact that inter-cell interference is inexistent or negligible and therefore, do not need to be accounted for. The rationale behind the single-cell Multi User (MU)-MIMO analysis is that its results are readily comprehended, they bound the performance of multicell systems and that the single-cell performance can be actually achieved if less aggressive frequency-reuse is adopted. For the following derivations we use standard capacity bounding techniques from the massive MIMO literature [80, 83, 97].

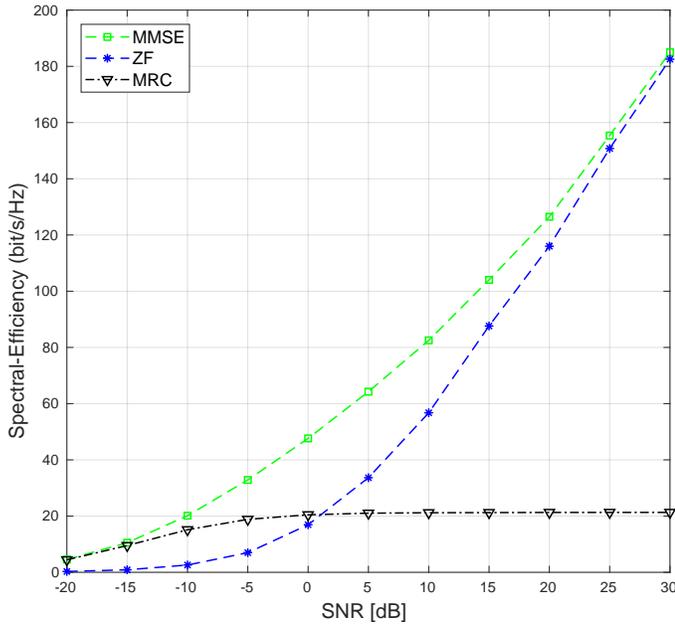


Fig. 8. Performance of linear receivers for different SNR values.

The received signal vector at the BS can be rewritten as

$$\mathbf{r} = \mathbf{A}^H (\sqrt{\rho} \mathbf{H} \mathbf{x} + \mathbf{w}) = \sqrt{\rho} \mathbf{A}^H \mathbf{H} \mathbf{x} + \mathbf{A}^H \mathbf{w}. \quad (19)$$

The received signal associated with the k -th MTC device is given by

$$r_k = \sqrt{\rho} \mathbf{a}_k^H \mathbf{H} \mathbf{x} + \mathbf{a}_k^H \mathbf{w} \\ = \underbrace{\sqrt{\rho} \mathbf{a}_k^H \mathbf{h}_k x_k}_{\text{desired signal}} + \underbrace{\sqrt{\rho} \sum_{l=1, l \neq k}^K \mathbf{a}_k^H \mathbf{h}_l x_l}_{\text{intra-cell interference}} + \underbrace{\mathbf{a}_k^H \mathbf{w}}_{\text{effective noise}}, \quad (20)$$

where \mathbf{a}_k , and \mathbf{h}_l are the l -th columns of \mathbf{A} , and \mathbf{H} , respectively. As shown in (20), the last two terms in the equation are considered as intra-cell interference and effective noise respectively. Therefore, an achievable rate for uplink transmission for the k -th MTC device is defined by

$$R_k = \mathbb{E} \{ \log_2 (1 + \text{SINR}_k) \}, \quad (21)$$

$$\text{SINR}_k = \frac{\rho |\mathbf{a}_k^H \mathbf{h}_k|^2}{\rho \sum_{l=1, l \neq k}^K |\mathbf{a}_k^H \mathbf{h}_l|^2 + \|\mathbf{a}_k\|^2}. \quad (22)$$

By following a similar line of reasoning as in [80, 83] we obtain lower bounds on the achievable rate for the MRC, ZF and MMSE linear detectors for the perfect CSI case.

1) *MRC detector*: The received SINR of the k th stream for the MRC detector is defined as

$$\text{SINR}_k^{\text{MRC}} = \frac{\rho \|\mathbf{h}_k\|^4}{\rho \sum_{l=1, l \neq k}^K |\mathbf{h}_k^H \mathbf{h}_l|^2 + \|\mathbf{h}_k\|^2} \\ \rightarrow \frac{\|\mathbf{h}_k\|^4}{\sum_{l=1, l \neq k}^K |\mathbf{h}_k^H \mathbf{h}_l|^2}, \text{ as } \rho \rightarrow \infty. \quad (23)$$

Considering i.i.d. Rayleigh fading, and MRC processing, the achievable rate for the k -th MTC device is lower bounded by

$$\tilde{R}_k^{\text{MRC}} = \log_2 \left(1 + \frac{\rho (M-1) d_k}{1 + \rho \sum_{l=1, l \neq k}^K d_l} \right). \quad (24)$$

If M grows without limit and we consider that the transmit power of each MTC device can be scaled with M according to $\rho = P/M$, where P is fixed, then (24) becomes

$$\tilde{R}_k^{\text{MRC}} = \log_2 \left(1 + \frac{\frac{P}{M} (M-1) d_k}{1 + \frac{P}{M} \sum_{l=1, l \neq k}^K d_l} \right) \\ \rightarrow \log_2 (1 + P d_k), M \rightarrow \infty. \quad (25)$$

This result shows that with a large number of antennas and perfect CSI available at the BS, the intra-cell interference vanishes and the transmit power of each MTC device can be reduced proportionally to $1/M$. The result in (25) is equal to the performance of a Single-Input Single-Output (SISO) system with transmit power P , with no small-scale fading (*i.e.*, fast-fading) and intra-cell interference. Additionally, the use of a large number of antennas at the BS increases the spectral efficiency of the system K times by simultaneously serving K MTC devices over the same time-frequency resources.

After analyzing (24) and (25), we see that if we decrease the transmit power proportionally to $1/M^\alpha$, where $\alpha > 1$, then the SINR of the uplink transmission from the k -th device will go to zero as $M \rightarrow \infty$. When $\alpha < 1$ the SINR of the uplink transmission from the k -th device grows without bound as $M \rightarrow \infty$. This means that $1/M$ (*i.e.*, $\alpha = 1$) is the fastest rate at which we can decrease the transmit power of each device and still maintain a fixed rate.

2) *ZF detector*: The received SINR of the k th stream for the ZF detector is defined as

$$\text{SINR}_k^{\text{ZF}} = \frac{\rho}{\left[(\mathbf{H}^H \mathbf{H})^{-1} \right]_{kk}}. \quad (26)$$

Considering i.i.d. Rayleigh fading, ZF processing, and $M \geq K + 1$, the achievable rate for the k -th MTC device is lower bounded by

$$\tilde{R}_k^{\text{ZF}} = \log_2 (1 + \rho (M - K) d_k). \quad (27)$$

If M grows without limit and we make $\rho = P/M$, then (27) becomes

$$\tilde{R}_k^{\text{ZF}} = \log_2 \left(1 + \frac{P}{M} (M - K) d_k \right) \\ \rightarrow \log_2 (1 + P d_k), M \rightarrow \infty. \quad (28)$$

3) *MMSE detector*: The received SINR of the k -th stream for the MMSE receiver is defined as

$$\text{SINR}_k^{\text{MMSE}} = \frac{1}{\left[(\mathbf{I}_K + \rho \mathbf{G}^H \mathbf{G})^{-1} \right]_{kk}} - 1 \\ \rightarrow \frac{\rho}{\left[(\mathbf{G}^H \mathbf{G})^{-1} \right]_{kk}}, \text{ as } \rho \rightarrow \infty. \quad (29)$$

The second line of (29) shows that as the SNR increases, ZF and MMSE detectors will present the same spectral efficiency (see Figure 8).

Considering i.i.d. Rayleigh fading, and MMSE processing, the achievable rate for the k -th MTC device is approximately lower bounded by

$$\tilde{R}_k^{\text{MMSE}} = \log_2(1 + (\alpha_k - 1)\theta_k), \quad (30)$$

where

$$\alpha_k = \frac{(M - K + 1 + (K - 1)\mu)^2}{M - K + 1 + (K - 1)\kappa}, \quad (31)$$

$$\theta_k = \frac{M - K + 1 + (K - 1)\kappa}{M - K + 1 + (K - 1)\mu} \rho d_k, \quad (32)$$

where μ and κ are obtained by solving the following two equations:

$$\mu = \frac{1}{K - 1} \sum_{l=1, l \neq k}^K \frac{1}{M \rho d_l \left(1 - \frac{K-1}{M} + \frac{K-1}{M} \mu\right) + 1} \quad (33)$$

$$\begin{aligned} \kappa & \left(1 + \sum_{l=1, l \neq k}^K \frac{\rho d_l}{\left(M \rho d_l \left(1 - \frac{K-1}{M} + \frac{K-1}{M} \mu\right) + 1\right)^2}\right) \\ & = \sum_{l=1, l \neq k}^K \frac{\rho d_l \mu + 1}{\left(M \rho d_l \left(1 - \frac{K-1}{M} + \frac{K-1}{M} \mu\right) + 1\right)^2} \end{aligned} \quad (34)$$

VI. SIMULATION RESULTS

In this section, we assess the performance of MRC, ZF, and MMSE linear detectors in terms of their bit error rate (BER), sum-spectral efficiency and transmitted power efficiency with perfect CSI knowledge. We assume a typical hexagonal single-cell structure with a radius of 1000 meters as depicted in Figure 3. The MTC devices are uniformly placed at random within the cell. Additionally, we consider that no device positioned closer than 100 meters to the BS. The large-scale fading coefficients $\{d_k\}$ are independently generated by $d_k = \psi / \left(\frac{r_k}{r_0}\right)^v$, where $v = 3.8$, $10 \log_{10}(\psi) \sim \mathcal{N}(0, \sigma_{\text{shadow,dB}}^2)$ with $\sigma_{\text{shadow,dB}} = 8$, $r_0 = 100$ meters and r_k is the distance of the k -th MTC device to the BS. Both, the path loss exponent, v , and the standard deviation of the log-normal shadowing, $\sigma_{\text{shadow,dB}}$, are common values for outdoor shadowed urban cellular radio environments [91, 92]. We assume $K = 10$ MTC devices for all simulations in this section.

Next, we present simulations to assess how the BER behaves when a large number of antennas and sub-optimal linear detectors are employed at the BS over a range of uplink transmit power (UL Tx power), ρ , values. For the following simulation results, we assume OFDM parameters similar to LTE: (i) a symbol interval of $T_s = 500/7 \approx 71.4 \mu\text{s}$, (ii) a subcarrier spacing of $\Delta_f = 15$ KHz, (iii) a useful symbol duration $T_u = 1/\Delta_f \approx 66.7 \mu\text{s}$, and (iv) a cyclic prefix interval (guard interval) $T_g = T_s - T_u \approx 4.76 \mu\text{s}$. The

frequency smoothness interval is approximately $N_{\text{smooth}} = 14$ subcarriers. We assume a coherence time of 1 ms (equivalent to 14 OFDM symbols), of which 1 symbol is used to send uplink pilots and the remaining symbols, *i.e.*, 13 symbols, are used to send uplink data. Therefore, in this case, the maximum number of MTC devices, $K = \tau_p N_{\text{smooth}} = 1 \times 14 = 14$. Additionally, we consider uncoded QPSK uplink data transmissions. We also consider perfect time and frequency synchronism among the multiple MTC devices and the BS.

We use the Matched Filter Bound (MFB) as benchmark for the BER comparisons. The MFB is also known in the literature as the perfect interference-cancellation bound [37]. As the name suggests, MFB performs as the k -th user of a matched-filter receiver in the absence of no other sources of interference (*i.e.*, no cross-talk interference) [37]. Our motivation for this choice is that for $M \gg K$ both multi-user interference and noise tend to vanish (thanks to the favourable propagation phenomenon), so the performance of the $K \times M$ MIMO channel, which is assumed to be i.i.d. flat Rayleigh fading inside a coherence interval, approaches the MFB as $M \rightarrow \infty$. The simulation results discussed in the sequel were averaged over 10^{10} realizations. The simulation type is Monte-Carlo with a bit error counting procedure that compares the transmitted bit vector to the received bit vector.

Figure 9 shows the BER of linear detectors for a fixed number of $K = 10$ MTC devices and BS array sizes in the range of $20 \leq M \leq 160$ antennas. As expected, and due to its better balance between interference suppression and noise enhancement, MMSE outperforms MRC and ZF in all cases studied. The performance gap inherent to MRC detection becomes evident from the results shown in this figure. The performance gap between ZF and MMSE, which is small enough to be considered negligible for $M \leq 40$, entirely vanishes as the BS array size is grown to $M = 80$ or above. Additionally, as can also be seen, MMSE detection is the best sub-optimal linear detector for cases where $M \leq K$.

Figure 10 shows that the performance of the MRC detector can be asymptotically increased at the expense of larger array sizes at the BS. This suggests that even low-complex MRC has the potential to approximate MFB in case M can be made large enough. In fact, the main conclusion drawn from the plots is that MRC, ZF, and MMSE all approach the performance of MFB as M grows without bound, but the gap between the perfect interference-cancellation bound and ZF/MMSE decreases at a faster pace than in case of MRC.

Figure 11 compares the lower bounds and simulated values of the spectral efficiency for MRC, ZF, and MMSE linear detectors for different number of BS antennas and perfect CSI under random large-scale fading $\{d_k\}$. The spectral efficiency is averaged over 1×10^6 realizations of $\{d_k\}$. In this result, there are $K = 10$ MTC devices, the transmit power per terminal is $\rho = 10$ dB, and the propagation channel parameters were $\sigma_{\text{shadow}} = 8$ dB, and path loss exponent, $\gamma = 3.8$ (Suburban Area). As can be seen, with this UL Tx power, the spectral efficiency for $M = 500$ is in the order of 32 - 42 bits/s/Hz, corresponding to a spectral efficiency of 3.2 - 4.2 bits/s/Hz per MTC device. These values correspond to practical values, for instance, 64-QAM modulation, where a

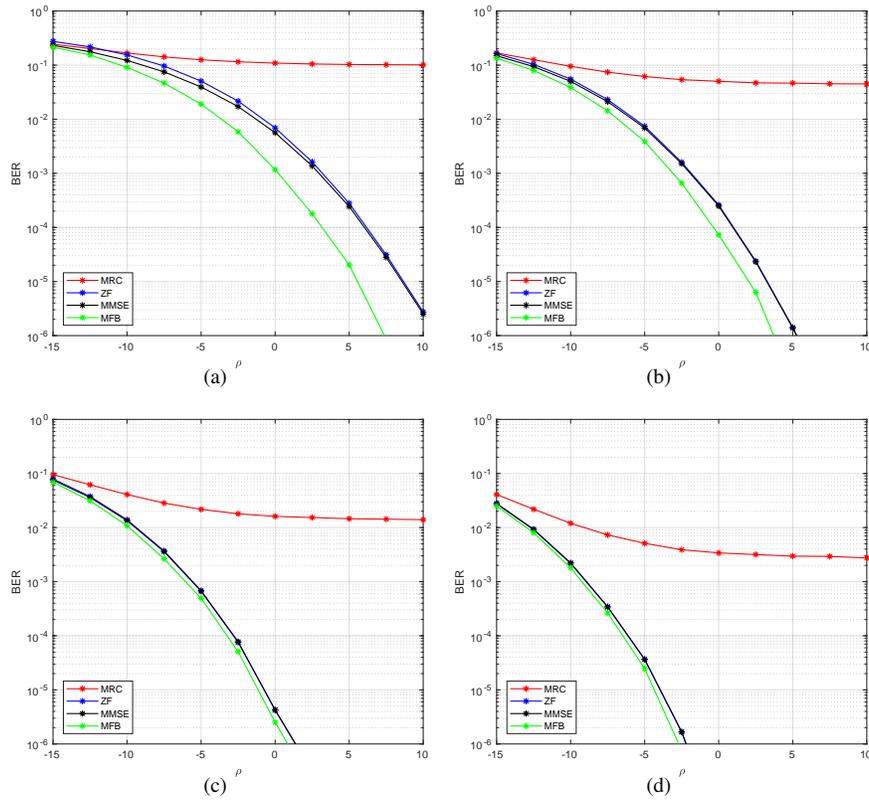


Fig. 9. BER performance of different linear detection methods for $K = 10$ single-antenna MTC devices over a range of different uplink transmit power, ρ values and different array sizes deployed at the BS. MFB is provided as a benchmark for comparisons. (a) $M = 20$ antennas. (b) $M = 40$ antennas. (c) $M = 80$ antennas. (d) $M = 160$ antennas.

symbols carries 6 bits, with a channel coding rate of $3/4$ corresponds to spectral efficiency of 4.5 bits/s/Hz. As can be noticed from the figure, the MMSE detector always presents better performance than the MRC and ZF detectors, however, the ZF's detector performance asymptotically approaches the MMSE when M increases. Therefore, we conclude that even with simple and suboptimal linear processing algorithms such as MRC and ZF, it is possible to achieve high spectral efficiencies and consequently serve more MTC devices through the same time/frequency resources.

In Figure 12 we show the transmit power per MTC device that is needed to reach a fixed spectral efficiency. The figure shows the normalized power, ρ , required to achieve 1 and 2 bits/s/Hz per MTC device as a function of M . For the 1 bit/s/Hz case, it can be seen that by doubling the number of antennas, M , the transmit power can be cut back by approximately 3 dB. When M is large, the difference in performance between MRC and ZF (or MMSE) is less than 0.65 dB. This difference increases when the target spectral efficiency is increased. For the 2 bits/s/Hz case, the cross-talk interference, *i.e.*, interference from other devices, is more significant (relative to thermal noise) and therefore, ZF and MMSE receivers perform relatively better. The difference in performance between MRC and ZF (or MMSE) for $M = 500$ antennas is approximately 2.86 dB. This characteristic of massive MIMO is very important to MTC devices, where stringent power constraints have to be respected.

In Figure 13 we illustrate the power-scaling laws. The figure

shows the spectral-efficiency versus the number of antennas M for $\rho = P/M$ and $\rho = P/\sqrt{M}$ and with MRC, ZF and MMSE linear detectors respectively. In the upper part of the figure we set P to 20 dB. At this SNR, the spectral efficiency is in the order of 10-30 bits/s/Hz, which corresponds to a spectral efficiency per MTC device of 1-3 bits/s/Hz. As expected, with $\rho = P/M$, when the number of antennas, M , grows, the spectral efficiency approaches a constant value. However, when $\rho = P/\sqrt{M}$, the spectral-efficiency grows without bound, *i.e.*, logarithmically fast with M , when $M \rightarrow \infty$. Next, in the lower part of the figure we set P to 5 dB. This figure provides the same insights as the upper one. The gap between the performance of MRC and that showed by the ZF and MMSE detectors is reduced compared the the upper part of the figure. This is due to the fact that the interference caused by cross-talk, *i.e.*, the interference coming from other MTC devices, is smaller in this figure when compared to the thermal noise. These results attest that the transmit power of each MTC device can be scaled down as P/M . If we compare the MRC detector with ZF and MMSE detectors in the upper part of the figure, it is possible to see that when the radiated power by MTC device is made inversely proportional to \sqrt{M} , the power level is not sufficiently low to make the MRC detector perform as well as the ZF and MMSE detectors. However, when the radiated power is made inversely proportional to M , then, in this case, the MRC detector performs closely to the ZF and MMSE detectors for a large number of antennas,

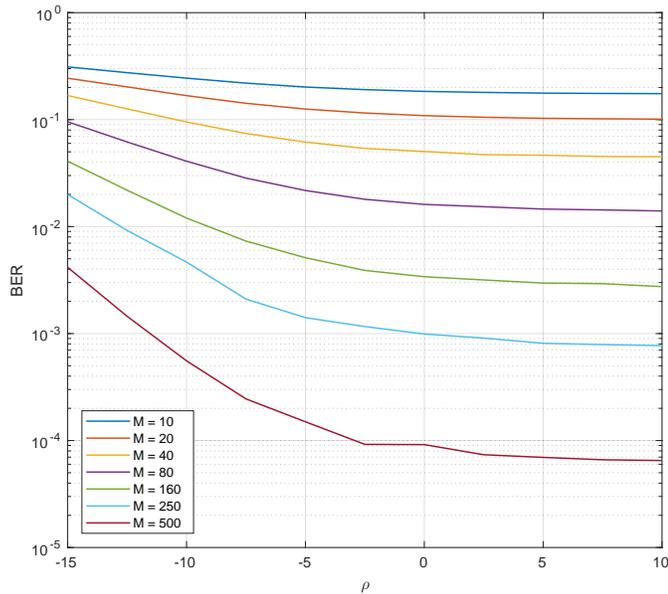


Fig. 10. BER performance for the MRC detector and several different number of antennas, M .

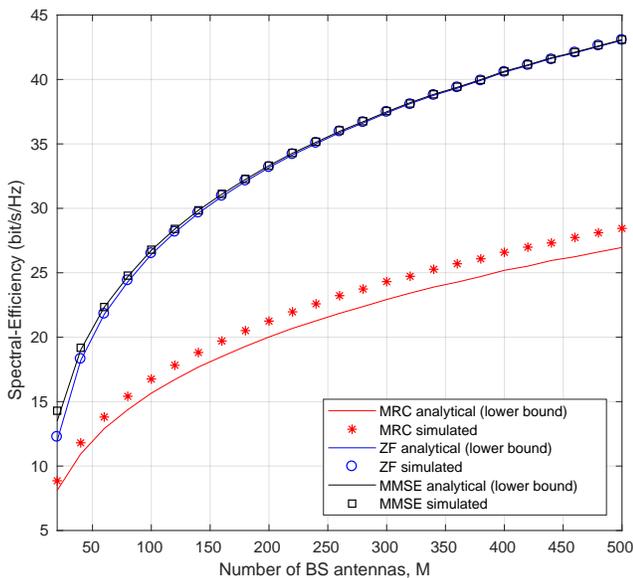


Fig. 11. Lower bounds and numerically evaluated values of the spectral efficiency for different BS antennas for MRC, ZF, and MMSE linear detectors and perfect CSI.

M . Therefore, these results show that the radiated power by MTC device can be made inversely proportional to the number of BS antennas, M , with no reduction in performance. Additionally, as the radiated power levels are reduced, the cross-talk interference caused by the use of the inferior MRC detector eventually drops below the noise level, making this simple linear detector a practical option, once it performs as well as ZF and MMSE detectors for the large M regime.

VII. CONCLUSIONS

This paper has proposed the use of a Massive MIMO technology as means to address the uplink mixed-service communication problem. In our study, we considered K single-

antenna MTC devices served by a BS equipped with a large number of antennas, M . Under the assumption of an available PNSCH, the capacity of the MTC network and, in turn, that of the whole system, can be increased by grouping MTC devices into clusters and letting each cluster share the same time-frequency resources. We show that the individual data streams conveyed by spatially spread MTC signals can be separated thanks to the powerful processing gain provided by the adoption of Massive MIMO technology.

As the size of the antenna array at the BS is made progressively larger, our simulation results suggest that the BER performance of sub-optimal linear detection methods approach the matched filter bound (also known as perfect interference-cancellation bound). ZF and MMSE detectors approach the bound at a faster pace than the simpler and less complex MRC detector. Due to its better balance between interference suppression and noise enhancement, MMSE outperforms MRC and ZF in all studied cases. The gap in the performance of ZF, however, is negligible for array sizes around 40 antennas, and entirely vanishes when the BS is equipped with $M \geq 80$ antennas.

We provide results showing that the adoption of massive MIMO technology increases the spectral efficiency of the whole system by allowing a great number of MTC devices to be served through the same time-frequency physical resources. Additionally, we also present results demonstrating that the transmit power of the MTC devices can be made inversely proportional to the number of available antennas.

It is also important to notice that (4) is an approximation that only becomes an equality when $M \gg K$, $M \rightarrow \infty$ and the channel offers favorable propagation and channel hardening, which explains why the MRC detection is not optimum for the results presented here.

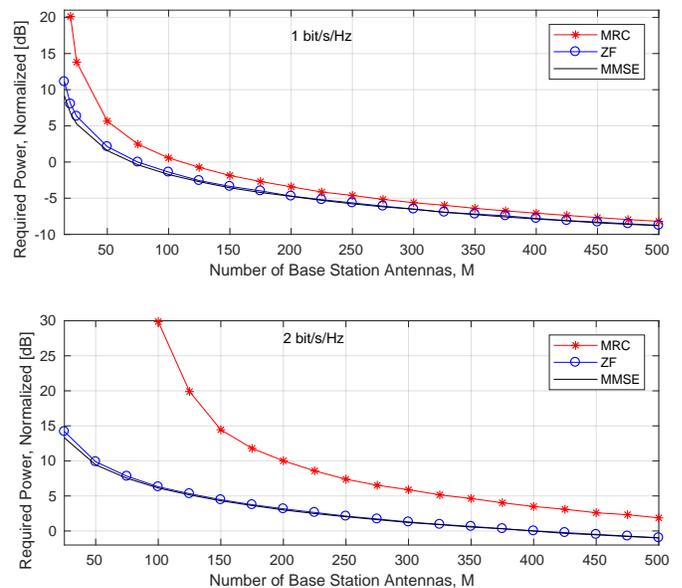


Fig. 12. Transmit power required to achieve 1 and 2 bits/s/Hz per MTC device for MRC, ZF and MMSE linear receivers as a function of the number of antennas M . The number of MTC devices is set to $K = 10$, and the propagation parameters are $\sigma_{\text{shadow}} = 8$ dB and $\nu = 3.8$.

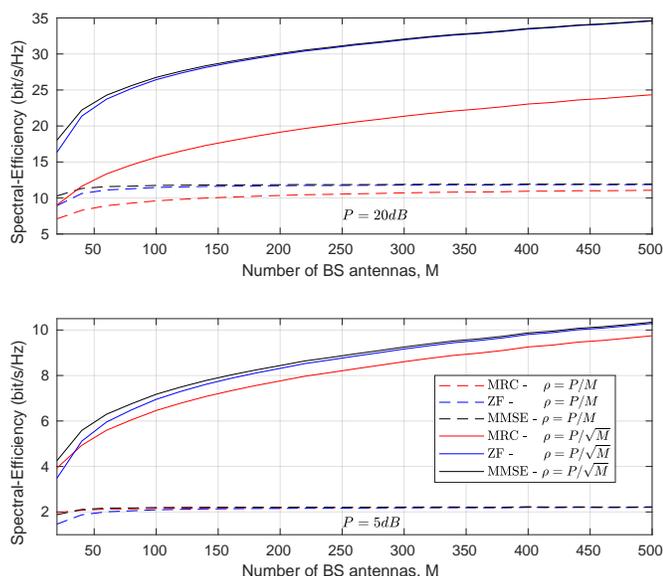


Fig. 13. Spectral-efficiency versus number of antennas, M , for linear detectors: MRC, ZF and MMSE. In these examples $K = 10$ MTC devices are served simultaneously and the transmit power is $P = 20$ dB for the upper figure and $P = 5$ dB for the lower figure and the propagation parameters are $\sigma_{\text{shadow}} = 8$ dB and $v = 3.8$.

In our future work, we will relax the assumption of perfect channel knowledge at the BS and assess its impact on bit error rate and spectral efficiency. Additionally, we also plan to analyze the effects of inter-cell interference on the system's performance.

FUNDING

This research article has received no funding.

DATA AVAILABILITY

There is no public data related to this research article.

REFERENCES

- [1] ITU-R, Recommendation ITU-R M.2083, *IMT Vision - Framework and overall objectives of the future development of IMT for 2020 and beyond*, September 2015.
- [2] Huawei Technologies CO., LTD, *5G Network Architecture - A High-Level Perspective*, White Paper, Online. [Available]: <https://www.huawei.com/en/industry-insights/outlook/mobile-broadband/insights-reports/5g-network-architecture>, 2016.
- [3] 4G Americas, *4G Americas' Recommendations on 5G Requirements and Solutions*, Online. [Available]: <http://www.4gamericas.org/en/resources/white-papers/>, October 2014.
- [4] G. Fettweis and S. Alamouti, *5G: Personal Mobile Internet beyond What Cellular Did to Telephony*, IEEE Communications Magazine, vol. 52, no. 2, pp. 140-145, February 2014.
- [5] G. Wunder et al., *5GNOW: Non-Orthogonal Asynchronous Waveforms for Future Mobile Applications*, IEEE Communications Magazine, vol. 52, no. 2, pp. 97-105, February 2014.
- [6] D. Boswarthick, O. Elloumi, and O. Hersent, *M2M Communications: A Systems Approach*, Wiley, 2012.
- [7] A. Elmangoush, H. Coskun, S. Wahle, N. Blum and T. Magedanz, *Promoting M2M Application Development for Smart City*, 29th meeting of the Wireless World Research Forum (WWRF), October 2012.
- [8] J. M. Corchado, J. Bajo, D. I. Tapia, and A. Abraham, *Using heterogeneous wireless sensor networks in a telemonitoring system for healthcare*, IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 2, pp. 234-240, March 2010.
- [9] P. Papadimitratos, A. de La Fortelle, K. Evensen, R. Brignolo, and S. Cosenza, *Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation*, IEEE Communications Magazine, vol. 11, no. 1, pp. 84-95, November 2009.
- [10] K. Moslehi and R. Kumar, *A reliability perspective of the smart grid*, IEEE Transactions on Smart Grid, vol. 1, no. 1, pp. 57-64, June 2010.
- [11] Elias Kyriakides and Marios Polycarpou, *Intelligent Monitoring, Control and Security of Critical Infrastructure Systems*, Springer, 1st Ed. 2015.
- [12] C. S. Pattichis, E. Kyriacou, S. Voskarides, M. S. Pattichis and R. Istepanian, *Wireless telemedicine Systems: An Overview*, IEEE Antennas and Propagation Magazine, vol. 44, no. 2, pp.143-153, April 2002.
- [13] Anita Majerowicz and Susan Tracy, *Telemedicine: Bridging Gaps in Healthcare Delivery*, Journal of AHIMA, vol. 81, no. 5, pp. 52-53, May 2010.
- [14] G. Araniti, C. Campolo, M. Condoluci, A. Iera and A. Molinaro, *LTE for Vehicular Networking: A Survey*, IEEE Communications Magazine, vol. 51, no. 5, pp. 148-157, May 2013.
- [15] Tamer ElBatt, Siddhartha K. Goel, Gavin Holland, Hariharan Krishnan, and Jayendra Parikh, "Cooperative collision warning using dedicated short range wireless communications, International Workshop on Vehicular Ad Hoc Networks (VANET), pp. 1-9, September 2006.
- [16] Christoph Ide, Fabian Kurtz and Christian Wietfeld, *Cluster-Based Vehicular Data Collection for Efficient LTE Machine-Type Communication*, IEEE Vehicular Technology Conference (VTC), September 2013.
- [17] Francis da Costa, *Rethinking the Internet of Things: A Scalable Approach to Connecting Everything*, Apress, 1st Ed. 2014.
- [18] B. C. Villaverde et al., *Service Discovery Protocols for Constrained Machine-to-Machine Communications*, IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 41-60, 2014.
- [19] G. Wunder, M. Kasparick, and P. Jung., *Spline Waveforms and Interference Analysis for 5G Random Access with Short Message Support*, arXiv:1501.02917, Online. [Available]: <http://arxiv.org/pdf/1501.02917.pdf>, January 2015.
- [20] G. Wunder, P. Jung, and C. Wang, *Compressive Random Access for Post-LTE Systems*, IEEE ICC - Workshop on Massive Uncoordinated Access Protocols, pp. 539-544, June 2014.
- [21] J.G. Andrews et al., *What Will 5G Be?*, IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1065-1082, June 2014.
- [22] Felipe A. P. de Figueiredo et. al., *A Spectrum Sharing Framework for Intelligent Next Generation Wireless Networks*, IEEE Access, vol. 6, pp. 60704-60735, December 2018.
- [23] Marzetta, T. L., *Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas*, IEEE Transactions on Wireless Communications, vol. 9, no. 11, pp. 3590-3600, October 2010.
- [24] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta., *Massive MIMO for next generation wireless systems*, IEEE Communications Magazine, vol. 52, no. 2, pp. 186-195, February 2014.
- [25] Emil Bjornson, Erik G. Larsson, and Thomas L. Marzetta, *Massive MIMO: Ten Myths and One Critical Question*, IEEE Communications Magazine, vol. 54, no. 2, pp. 114-123, February 2016.
- [26] Emil Bjornson, Jakob Hoydis, and Luca Sanguinetti, *Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency*, Foundations and Trends in Signal Processing: vol. 11, no. 3-4, pp. 154-655, DOI: 10.1561/20000000093, 2017.
- [27] Felipe A. P. de Figueiredo, Fabbryccio A. C. M. Cardoso, Renato R. Lopes and Joao Paulo Miranda, *On the Application of Massive MU-MIMO in the Uplink of Machine Type Communication Systems*, International Workshop on Telecommunications (IWT), June 2015.
- [28] Joao Paulo Miranda, Arman Farhang, Nicola Marchetti, Felipe A. P. de Figueiredo, Fabbryccio A. C. M. Cardoso, and Fabrício Figueiredo, *On massive MIMO and its applications to machine type communications and FBMC-based networks*, EAI Endorsed Transactions on Ubiquitous Environments, vol. 2, no. 5, July 2015.
- [29] Felipe A. P. de Figueiredo, Fabbryccio A. C. M. Cardoso, Ingrid Moerman, Gustavo Fraidenraich, *On the Application of Massive MIMO Systems to Machine Type Communications*, IEEE Access, vol. 7, no. 12, pp. 2589-2611, December 2018.
- [30] Andres Laya, Luis Alonso, and Jesus Alonso-Zarate, *Is the Random Access Channel of LTE and LTE-A Suitable for M2M Communications? A Survey of Alternatives*, IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 4-16, First Quarter 2014.
- [31] Stefania Sesia, Issam Toufik and Matthew Baker, *LTE - The UMTS Long Term Evolution: From Theory to Practice*, John Wiley & Sons, ISBN:9780470697160, February 2009.
- [32] S.-Y. Lien, K.-C. Chen, and Y. Lin, *Toward Ubiquitous Massive Accesses in 3GPP Machine-to-machine Communications*, IEEE Communications Magazine, vol. 49, no. 4, pp. 66-74, April 2011.

- [33] M.-Y. Cheng, G.-Y. Lin, H.-Y. Wei, and A.-C. Hsu, *Overload control for machine-type-communications in LTE-advanced system*, IEEE Communications Magazine, vol. 50, no. 6, pp. 38-45, June 2012.
- [34] M. Gerasimenko et al., *Energy and delay analysis of LTE-advanced RACH performance under MTC overload*, IEEE Globecom Workshops, pp. 1632-1637, December 2012.
- [35] U. Phuyal, A. Koc, M.-H. Fong, and R. Vannithamby, *Controlling access overload and signaling congestion in M2M networks*, Asilomar Conference on Signals, Systems, and Computers, pp. 591-595, November 2012.
- [36] T. L. Marzetta, *Noncooperative cellular wireless with unlimited numbers of base station antennas*, IEEE Transactions on Wireless Communications, vol. 9, no. 11, pp. 3590-3600, November 2010.
- [37] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, Springer, 3rd Ed. 2004.
- [38] Shepard, C. et al., *Argos: Practical many-antenna base stations*, International Conference on Mobile Computing and Networking (ACM MobiCom), pp. 53-64, August 2012.
- [39] Janaswamy, R., *Effect of element mutual coupling on the capacity of fixed length linear arrays*, IEEE Antennas Wireless Propagation Letters, vol. 1, no. 1, pp. 157-160, 2002.
- [40] Wallace, J. W. and Jensen, M. A., *Mutual coupling in MIMO wireless systems: A rigorous network theory analysis*, IEEE Transactions on Wireless Communications, vol. 3, no. 4, pp. 1317-1325, July 2004.
- [41] Lau, B. K. et al., *Impact of matching network on bandwidth of compact antenna arrays*, IEEE Transactions on Antennas Propagation, vol. 54, no. 11, pp. 3552-3238, November 2006.
- [42] Sibille, A., Oestges, C. and Zanella, A., *MIMO: From Theory to Implementation*, Academic Press, 2010.
- [43] Rusek, F. et al., *Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays*, IEEE Signal Processing Magazine, vol. 30, no. 1, pp. 40-60, January 2013.
- [44] Moustakas, A. L. et al., *Communication through a diffusive medium: Coherence and capacity*, Science, vol. 287, pp. 287-290, January 2000.
- [45] Artiga, X., Devillers, B. and Perruisseau-Carrier, J., *On the selection of radiating elements for compact indoor massive-multiple input multiple output base stations*, IET Microwaves, Antennas & Propagation, vol. 8, no. 1, pp. 1-9, January 2014.
- [46] Gao, X., Tufvesson, F. and Edfors, O., *Massive MIMO channels - measurements and models*, Asilomar Conference on Signals, Systems and Computers, November 2013.
- [47] Shangbin Wu, Cheng-Xiang Wang, Yang Yang, Wenjin Wang, and Xiqi Gao, *Performance comparison of massive MIMO channel models*, IEEE/CIC International Conference on Communications in China (ICCC), July 2016.
- [48] Zheng, K. Ou, S. and Yin, X., *Massive MIMO Channel Models: A Survey*, International Journal of Antennas and Propagation, vol. 2014, 10 pages, June 2014.
- [49] Jiang, Z. et al., *Achievable Rates of FDD Massive MIMO Systems with Spatial Channel Correlation*, IEEE Transactions on Wireless Communications, vol. 14, no. 5, pp. 2868-2882, May 2015.
- [50] Hien Quoc Ngo, *Massive MIMO: Fundamentals and System Designs*, Linköping University, Sweden, 2015.
- [51] Jose, J. et al., *Pilot Contamination and Precoding in Multi-Cell TDD Systems*, IEEE Transactions on Wireless Communications, vol. 10, no. 8, pp. 2640-2651, August 2011.
- [52] Kaltenberger, F. et al., *Relative channel reciprocity calibration in MIMO/TDD systems*, Future Network and Mobile Summit, June 2010.
- [53] Rogalin, R. et al., *Hardware-Impairment Compensation for Enabling Distributed Large-Scale MIMO*, Information Theory and Applications Workshop (ITA), February 2013.
- [54] Rogalin, R. et al., *Scalable Synchronization and Reciprocity Calibration for Distributed Multiuser MIMO*, IEEE Transactions on Wireless Communications, vol. 13, no. 4, pp. 1815-1831, April 2014.
- [55] Ngo, H. Q. and Larsson, E. G., *EVD-based channel estimations for multicell multiuser MIMO with very large antenna arrays*, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), March 2012.
- [56] Ashikhmin, A. and Marzetta, T. L., *Pilot contamination precoding in multi-cell large scale antenna systems*, IEEE International Symposium on Information Theory (ISIT), July 2012.
- [57] Yin, H. et al., *A coordinated approach to channel estimation in large-scale multiple-antenna systems*, IEEE Journal on Selected Areas in Communications, vol. 31, no. 2, pp. 264-273, February 2013.
- [58] Mueller, R., Vehkaperae, M. and Cottatellucci, L., *Blind Pilot Decontamination*, IEEE Journal of Selected Topics in Signal Processing, vol. 8, no. 5, October 2014.
- [59] Olakunle Elijah, Chee Yen Leow, Tharek Abdul Rahman, Solomon Nunoo, and Solomon Zakwoi Iliya, *A Comprehensive Survey of Pilot Contamination in Massive MIMO-5G System*, IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 905-923, 2016.
- [60] Felipe A. P. de Figueiredo et al., *Channel estimation for massive MIMO TDD systems assuming pilot contamination and flat fading*, EURASIP Journal on Wireless Communications and Networking, vol. 2018, no. 14, January 2018.
- [61] Caire, G. and Shamai, S., *On the achievable throughput of a multi-antenna Gaussian broadcast channel*, IEEE Transactions on Information Theory, vol. 49, no. 7, pp. 1691-1706, July 2003.
- [62] Windpassinger, C. et al., *A performance study of MIMO detectors*, IEEE Transactions on Wireless Communications, vol. 5, no. 8, pp. 2004-2008, August 2006.
- [63] Peel, C. B. et al., *A vector-perturbation technique for near-capacity multi-antenna communication, part I: Channel inversion and regularization*, IEEE Transactions on Communications, vol. 53, no. 1, pp. 195-202, January 2005.
- [64] Zu, K., de Lamare, R. C. and Haardt, M., *Generalized design of low-complexity block diagonalization type precoding algorithms for multiuser MIMO systems*, IEEE Transactions on Communications, vol. 61, no. 10, pp. 4232-4242, October 2013.
- [65] Stephan Jaeckel, Leszek Raschkowski, Kai Borner, Lars Thiele, *QuaDRiGa: A 3-D Multi-Cell Channel Model With Time Evolution for Enabling Virtual Field Trials*, IEEE Transactions on Antennas and Propagation, vol. 62, issue 6, pp. 3242-3256, June 2014.
- [66] S. Verdú, *Multiuser Detection*, Cambridge University Press, 1st Ed. 1998.
- [67] R. C. de Lamare, *Massive MIMO Systems: Signal Processing Challenges and Research Trends*, URSI Radio Science Bulletin, 2013.
- [68] Y.-C. Liang, E. Y. Cheu, L. Bai, and G. Pan, *On the relationship between MMSE-SIC and BI-GDFE receivers for large multiple-input multiple-output channels*, IEEE Transactions on Signal Processing, vol. 56, no. 8, pp. 3627-3637, August 2008.
- [69] H. Zhao, H. Long, and W. Wang, *Tabu search detection for MIMO systems*, IEEE PIMRC, pp. 1-5, September 2007.
- [70] Y. Sun, *A family of likelihood ascent search multiuser detectors: An upper bound of bit error rate and a lower bound of asymptotic multiuser efficiency*, IEEE Transactions on Communications, vol. 57, no. 6, pp. 1743-1752, June 2009.
- [71] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge, UK, Cambridge University Press, 2003.
- [72] D. Tse, and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge, UK, Cambridge University Press, 2005.
- [73] J. Jaldén and B. Ottersten, *On the complexity of sphere decoding in digital communications*, IEEE Transactions on Signal Processing, vol. 53, no. 4, pp. 1474-1484, April 2005.
- [74] L. Barbero and J. Thompson, *Fixing the complexity of the sphere decoder for MIMO detection*, IEEE Transactions on Wireless Communications, vol. 7, no. 6, pp. 2131-2142, June 2008.
- [75] Monowar Hasan, Ekram Hossain, and Dusit Niyato, *Random access for machine-to-machine communication in LTE-advanced networks: issues and approaches*, IEEE Communications Magazine, vol. 51, no. 6, pp. 86-93, June 2013.
- [76] Md Shipon Ali, Ekram Hossain, and Dong In Kim, *LTE/LTE-A Random Access for Massive Machine-Type Communications in Smart Cities*, IEEE Communications Magazine, vol. 55, no. 1, pp. 76-83, January 2017.
- [77] Han Seung Jang, Su Min Kim, Kab Seok Ko, Jiyoung Cha, and Dan Keun Sung, *Spatial Group Based Random Access for M2M Communications*, IEEE Communications Letters, vol. 18, no. 6, pp. 961-964, June 2014.
- [78] Shree Krishna Sharma, and Xianbin Wang, *Towards Massive Machine Type Communications in Ultra-Dense Cellular IoT Networks: Current Issues and Machine Learning-Assisted Solutions*, arXiv:1808.02924, August 2018.
- [79] Felipe A. P. de Figueiredo et al., *Multi-stage Based Cross-Correlation Peak Detection for LTE Random Access Preambles*, Revista Telecomunicações, vol. 15, pp. 1-7, October 2013.
- [80] Thomas L. Marzetta, Erik G. Larsson, Hong Yang, Hien Quoc Ngo, *Fundamentals of Massive MIMO*, Cambridge U.K., Cambridge University Press, 1st edition, November 2016.
- [81] Shendi Wang, John S. Thompson, and Peter M. Grant, *Closed-Form Expressions for ICI/ISI in Filtered OFDM Systems for Asynchronous 5G Uplink*, IEEE Transactions on Communications, vol. 65, no. 11, pp. 4886-4898, November 2017.

- [82] Abdoli J., Jia M., and Ma J., *Filtered OFDM: A New Waveform for Future Wireless Systems*, IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), July 2015.
- [83] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, *Energy and spectral efficiency of very large multiuser MIMO systems*, IEEE Transactions on Communications, vol. 61, no. 4, pp. 1436-1449, April 2013.
- [84] A. Ashikhmin, T. L. Marzetta, L. Li, *Interference reduction in multi-cell massive MIMO systems I: Large scale fading precoding and decoding*, arXiv preprint arXiv:1411.4182, November 2014.
- [85] Zheng Chen, and Emil Bjornson, *Channel Hardening and Favorable Propagation in Cell-Free Massive MIMO with Stochastic Geometry*, arXiv:1710.00395, June 2018.
- [86] H. Q. Ngo, Erik G. Larsson, and Thomas L. Marzetta, *Aspects of Favorable Propagation in Massive MIMO*, European Signal Processing Conference (EUSIPCO), pp. 76-80, September 2014.
- [87] H. Q. Ngo, and E. Larsson, *No downlink pilots are needed in TDD Massive MIMO*, IEEE Transactions on Wireless Communications, vol. 16, no. 5, pp. 2921-2935, May 2017.
- [88] M. Matthaiou, M. R. McKay, P. J. Smith, and J. A. Nossek, *On the condition number distribution of complex wishart matrices*, IEEE Transactions on Communications, vol. 58, no. 6, pp. 1705-1717, June 2010.
- [89] F. Rusek et al., *Scaling up MIMO: Opportunities and Challenges with Very Large Arrays*, IEEE Signal Processing Magazine, vol. 30, no. 1, pp. 40-60, January 2013.
- [90] Pramod Viswanath and David Tse, *Sum Capacity of the Vector Gaussian Broadcast Channel and Uplink-Downlink Duality*, IEEE Transactions on Information Theory, vol. 49, no. 8, August 2003.
- [91] Theodore S Rappaport, *Wireless communications: principles and practice*, Prentice Hall, 2002.
- [92] S. R. Saunders, and A. Aragon-Zavala, *Antennas and Propagation for Wireless Communication Systems*, John Wiley & Sons, Chichester, UK, 2nd edition, 2007.
- [93] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, *Massive MIMO performance evaluation based on measured propagation data*, IEEE Transactions on Wireless Communications, vol. 14, no. 7, pp. 3899-3911, July 2015.
- [94] J. Hoydis, C. Hoek, T. Wild, and S. ten Brink, *Channel measurements for large antenna arrays*, International Symposium on Wireless Communication Systems (ISWCS), pp. 811-815, August 2012.
- [95] A. O. Martinez, E. D. Carvalho, and J. O. Nielsen, *Towards very large aperture massive MIMO: A measurement based study*, Proceedings of IEEE Global Telecommunications Conference (GLOBECOM), pp. 281-286, December 2014.
- [96] Alex Oliveras Martinez, Elisabeth De Carvalho, Jesper Odum Nielsen, *Massive MIMO Properties based on Measured Channels: Channel Hardening, User Decorrelation and Channel Sparsity*, Asilomar Conference on Signals, Systems and Computers (ACSSC), November 2016.
- [97] E. G. Larsson, and H. V. Poor, *Joint beamforming and broadcasting in massive MIMO*, IEEE Transactions on Wireless Communications, vol. 15, no. 4, pp. 3058-3070, April 2016.