Massive MIMO and Waveform Design for 5th Generation Wireless Communication Systems

Arman Farhang*, Nicola Marchetti*, Fabrício Figueiredo† and João Paulo Miranda†

*CTVR - The Telecommunications Research Centre, Trinity College, Ireland. Email: [farhanga, marchetti]@tcd.ie
†CPqD - Research and Development Center on Telecommunication, Brazil. Email: [fabricio, jmiranda]@cpqd.com.br

Abstract—This article reviews existing related work and identifies the main challenges in the key 5G area at the intersection of waveform design and large-scale multiple antenna systems, also known as Massive MIMO. The property of self-equalization is introduced for Filter Bank Multicarrier (FBMC)-based Massive MIMO, which can reduce the number of subcarriers required by the system. It is also shown that the blind channel tracking property of FBMC can be used to address pilot contamination – one of the main limiting factors of Massive MIMO systems. Our findings shed light into and motivate for an entirely new research line towards a better understanding of waveform design with emphasis on FBMC-based Massive MIMO networks.

I. INTRODUCTION

Ongoing societal developments are changing the way we use communication systems. On-demand data is increasingly being delivered over mobile and wireless communication systems. These developments will lead to a big rise of traffic volume. Applications, such as broadband telephony and Machine-Type Communications (MTC), have diverse needs that fifth generation (5G) systems will have to support, e.g. stringent latency and reliability requirements, and a wide range of data rates

One of the main challenges is to meet these goals while at the same time addressing the growing cost pressure. A popular view is that the required increase in data rate will be achieved through combined gains in extreme network densification (to improve area spectral efficiency), increased bandwidth (by exploiting mmWaves and making better use of unlicensed spectrum), and increased spectral efficiency (through advances in Multiple Input Multiple Output (MIMO) techniques).

In the quest for bandwidth, particular challenges that need to be addressed in the context of 5G are fragmented spectrum and spectrum agility. It is unlikely that these challenges can be satisfied using Orthogonal Frequency Division Multiplexing (OFDM), and new waveforms are required. Some researchers have also started to question the working assumption of strict synchronism and orthogonality in cellular networks, as a way to relax strict time domain requirements in case of sporadic traffic generating devices (e.g. MTC devices) or applications requiring ultra-low latency, such as the Tactile Internet.

Waveform design is quite a hot topic now, as it sheds light into how candidate waveforms perform in cellular environments, and discusses how they fare regarding specific aspects of 5G systems. In [4], for instance, OFDM is compared to Filter Bank Multi-Carrier (FBMC), Time Frequency-packed Signaling (TFS), and Single-Carrier Modulation (SCM); OFDM is preferred in terms of ease of hardware implementation; SCM is the best candidate to reduce latency and peak-to-average power ratio (PAPR); FBMC is most robust against synchronisation errors; all waveforms but TFS can be used in mm-Wave bands, and all can be adopted in a Massive MIMO setup, i.e. systems using arrays with at least an order of magnitude more antennas than conventional (Multi-user MIMO) systems.

A consequence of the powerful signal processing enabled by the large number of antennas is that most of the scheduling and PHY layer control issues in general are automatically resolved in Massive MIMO systems – which of course is not the case for Multi-user MIMO systems with just a moderate number of antennas. In recent past, Massive MIMO has gained significant momentum as potential candidate to increase capacity in multi-user networks. In the limit, as the number M of antennas at the Base Station (BS) tends to infinity, the system processing gain tends to infinity. As a result, the effects of both noise and multi-user interference are removed. Massive MIMO also enables significant latency reduction on the air interface, a welcome feature for delay-constrained applications.

In general, it appears that any modulation technique, either single- or multi-carrier, can be used in combination with large antenna arrays. According to [4], it is reasonable to foresee that a similar behavior with respect to the vanishing of inter- and inter-symbol interference (ISI) can be observed for any modulation format when the number of receiving antennas M is sufficiently high. However, since not all waveforms have equal advantages, the benefits of large antenna arrays can make a certain Massive MIMO-specific waveform combination more attractive than others. FBMC offers lower out-of-band (OOB) emissions, and allows less expensive and more flexible carrier aggregation than OFDM, but traditionally had the problem of non easy applicability of MIMO to it [3]. By scaling up the number of antennas, the combination of Massive MIMO and FBMC can benefit from the former’s gains while still retaining the good properties of the latter [5].

In [7] it is pointed out that, due to the law of large numbers, the channel hardens so that each subcarrier in a Massive MIMO system will have substantially the same channel gain. Such a property has been also reported in [9] in the context of FBMC, where the authors name it self-equalization, and can lead to a reduction in the number of subcarriers required by the system. Another advantage of Massive MIMO is that, thanks to its many spatial degrees of freedom, the same frequency band can be reused for many users. This plus the channel hardening render frequency domain scheduling no
Cost-effective strategies are needed

Channel reciprocity calibration

Currently under development

Experimental assessments, testbeds & prototypes

More complex than MMSE-SIC

Limited by the channel coherence time

Dense multidimensional implementations

Candidate Solutions

Extremely costly for practical deployment

Complexity is exponential in

No solution known to date

Open research question

BS sends pilots to terminals via FDD

Limited by the channel coherence time

Terminals send pilots to BS via TDD

Pilot contamination problem

Channel modeling

Realistic empirical models

Sophisticated analytical models

No solution known to date

Open research question

Candidate Solutions

Extremely costly for practical deployment

Complexity is exponential in

No solution known to date

Open research question

Phase noise

Smart PHY transmission and receiver algorithms

Efficacy yet to be demonstrated

Experimental assessments, testbeds & prototypes

Only basic capabilities demonstrated

Power consumption

Parallel, dedicated baseband signal processing

Open research question

Proof-of-Concept (PoC)

Experimental assessments, testbeds & prototypes

Only basic capabilities demonstrated

TABLE I

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

| Research Area       | Issue                                      | Candidate Solutions                                          | Shortcomings and "Side Effects"                                                                 | Refs |
|---------------------|--------------------------------------------|-------------------------------------------------------------|--------------------------------------------------------------------------------------------------|------|
| Antenna Aspects     | Antenna coupling                           | Multiport impedance matching RF circuits                    | • Diminishes bandwidth                                                                         | [10] |
|                     |                                            |                                                             | • Introduces ohmic losses                                                                       | [11] |
|                     |                                            |                                                             | • Not fully understood for large $M$                                                            | [12] |
|                     | Front-back ambiguity                        | Dense multidimensional implementations                     | • Increases coupling effects                                                                  | [13] |
|                     |                                            |                                                             | • Limited to indoor environments                                                               | [14] |
|                     |                                            |                                                             | • 3D arrays have restricted usefulness                                                          | [15] |
| Propagation         | Channel modeling                           | Realistic empirical models                                  | Currently under development                                                                  | [16] |
|                     |                                            | Sophisticated analytical models                            |                                                                                                 | [17] |
|                     | Cluster resolution                          | No solution known to date                                   |                                                                                                 | [18] |
|                     |                                            |                                                             |                                                                                                 | [19] |
| Transceiver Design  | CSI acquisition                             | BS sends pilots to terminals via FDD                        | Limited by the channel coherence time                                                          | [20] |
|                     |                                            | Terminals send pilots to BS via TDD                        | Channel reciprocity calibration                                                               | [21] |
|                     | Precoding (in the DL)                      | Linear precoding methods                                   | • ZF                                                                                           | [22] |
|                     |                                            |                                                             | • MMSE                                                                                          | [23] |
|                     |                                            |                                                             | • Computationally heavy for large $M$                                                            | [24] |
|                     |                                            |                                                             | • Higher average transmit power                                                                 | [25] |
|                     |                                            |                                                             | • MF                                                                                           | [26] |
|                     |                                            |                                                             | • Has an error floor as $M$ increases                                                            | [27] |
|                     |                                            |                                                             | • Higher $M$ required for a given SIR                                                             | [28] |
|                     |                                            |                                                             | • BD                                                                                           | [29] |
|                     |                                            |                                                             | • Cost-effective strategies are needed                                                            | [30] |
|                     |                                            |                                                             | • DPC                                                                                           | [31] |
|                     |                                            |                                                             | • Extremely costly for practical deployment                                                       | [32] |
|                     |                                            |                                                             | • Increased complexity is hard to justify                                                        | [33] |
|                     | Nonlinear precoding methods                | Iterative linear filtering                                 | • MMSE-SIC                                                                                      | [34] |
|                     |                                            |                                                             | • BI-GDFE                                                                                       | [35] |
|                     |                                            |                                                             | • Computationally heavy for large $M$                                                            | [36] |
|                     |                                            | Random step search methods                                 | • TS                                                                                           | [37] |
|                     |                                            |                                                             | • More complex than MMSE-SIC                                                                       | [38] |
|                     |                                            | Tree-based algorithms                                      | • SD                                                                                           | [39] |
|                     |                                            |                                                             | • Complexity is exponential in $M$                                                                | [40] |
|                     |                                            |                                                             | • FCSD                                                                                          | [41] |
|                     |                                            |                                                             | • 1000x more complex than TS                                                                      | [42] |
|                     |                                            |                                                             | • Best suitable for the $M \approx K$ case                                                       | [43] |
| Hardware            | Phase noise                                | Smart PHY transmission and receiver algorithms              | Efficacy yet to be demonstrated                                                                | [44] |
|                     | Power consumption                          | Parallel, dedicated baseband signal processing              | Open research question                                                                          | [45] |
|                     | Proof-of-Concept (PoC)                     | Experimental assessments, testbeds & prototypes             | Only basic capabilities demonstrated                                                            | [46] |

longer needed. Issues related to Massive MIMO and waveform design for 5G are discussed in Section II and III, respectively. Section IV presents some recent results of the authors on the combination of Massive MIMO and FBMC. Section V concludes the paper.

II. LARGE-SCALE MULTIPLE ANTENNA SYSTEMS

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.

A. Mutual Antenna 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 a negligible level. This assumption is not entirely realistic, especially if a large number of elements is to be deployed as an array of constrained size and aperture. Under such practical conditions, mutual coupling is known to substantially impact the achievable system capacity [10]. Multiport impedance matching RF circuits can cancel out such coupling effects [11], but they diminish output port bandwidth [12] and increase ohmic losses [13].

Two- or three-dimensional arrays have been reported 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 [5]. 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 [15]. The optimal densities above which performance deteriorates no matter how large is the number of elements are studied in [14] for indoor Massive MIMO BSs.

B. 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 [16], 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. At the time of writing, no such a model seems to exist for Massive MIMO. See, e.g. [17], for a discussion on modeling methods, channel categories, and their underlying properties.

C. Channel State Information, Precoding & Detection

In conventional 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 Frequency-division Duplexing (FDD) finds limited application in Massive MIMO systems in that the amount of time-frequency resources needed for pilot transmission in the
downlink (DL) scales as the number of antennas, and so does the number of channel responses that must be estimated on the part of each terminal. In large arrays, pilot transmission time may well exceed the coherence time of the channel.[6][18].  

An alternative for Massive MIMO systems is to let terminals send pilots to the BS via Time-division Duplexing (TDD). TDD relies on channel reciprocity, where uplink (UL) channels serve as estimate of DL channels. This leads to training requirements independent of \( M \)[19], 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; the latter arises in multi-user 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[20][22] and[23]-[26], but optimal solutions are unknown to date.  

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 system to support multiple streams. Precoding based on Zero-Forcing by modifying standard single-stream beamforming techniques and matrix-inversion free methods such as TS and LAS are(FCSD)[35][36]. Repeated matrix inversions in MMSE-SIC(BI-GDFE)[32], Tabu Search (TS)[33], Likelihood Ascent Block-iterative Generalized Decision Feedback Equalization(MRT)[30], have appealing features (DPC is theoretically optimal) but are not capable of achieving the full receive-diversity order of ML detection[27]. Maximum Ratio Transmission (MRT) in the DL and Maximum Ratio Combining (MRC) in the UL, is known to be the simplest method[6].  

Dirty Paper Coding (DPC)[29], Tomlinson-Harashima Precoding (THP)[30], and Vector Perturbation (VP)[31], 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 Ratio (SIR) using MF is at least two orders of magnitude larger than using ZF precoding[5], further work on cost-effective solutions is needed. This is illustrated in[33] for the case of Block Diagonalization (BD) algorithms.  

When it comes to separation of data streams in conventional systems, Maximum Likelihood (ML) detection is the optimal solution but its complexity grows exponentially with the number of streams. This is the reason why parameter estimation and detection are key problems in Massive MIMO systems. Suboptimal ZF, MMSE, and MF offer lower costs (that do not depend on the modulation), but are not capable of achieving the full receive-diversity order of ML detection[27].  

This performance-complexity tradeoff led to methods, such as MMSE with Successive Interference Cancellation (SIC), Block-iterative Generalized Decision Feedback Equalization (BI-GDFE)[32], Tabu Search (TS)[33], Likelihood Ascent Search (LAS)[34], and Fixed Complexity Sphere Decoding (FCSD)[35][36]. Repeated matrix inversions in MMSE-SIC and BI-GDFE may be computationally heavy for large arrays, and matrix-inversion free methods such as TS and LAS are both outperformed by MMSE-SIC[5]. This is an indication that more work is needed on this matter, perhaps towards turbo codes or Low-density Parity-check (LDPC) codes in iterative detection and decoding settings[27].  

D. Impairments due to Low-cost Hardware  

Large-scale multiple antenna arrays will most likely be built using low-cost components so as 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[7].  

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[7]. Another path of interest is experimentation, as testbeds currently available only demonstrate basic capabilities, and do not take constrained BS real estates into consideration[37].

III. WAVEFORM DESIGN FOR 5G

This section presents the state of the art about candidate waveforms for 5G. Comparisons in Table II are based on the best possible performance of each waveform, with OFDM with cyclic prefix (CP-OFDM) used as benchmark.  

A. The Baseline OFDM and its Enhancements

Despite the advantages that led to the near-universal adoption of CP-OFDM, it is not without its limitations[3]. High PAPR in Massive MIMO is a concern, as it sets up a tradeoff between the amplifier’s linearity and cost. MmWave deployment may also prove hard due to the difficulty to develop efficient amplifiers. Spectral efficiency[3] can be improved by means of shorter CP lengths, and Frequency and Quadrature Amplitude Modulation (FQAM) to boost DL throughput for cell-edge users. TFS and faster-than-Nyquist signaling have been claimed able to offer efficiency gains on the order of 25% over conventional OFDM (see[2][3] and references therein). Other drawbacks of CP-OFDM are sensitivity to phase noise and asynchronous signaling, poor spectrum localization, large OOB emissions, and long round-trip time (RTT).  

The amount of implementation experience and knowledge about tricky aspects of OFDM available today make it possible to modify it to create new schemes capable of circumventing most of its inherent limitations. Improved sidelobe suppression for dynamic spectrum access and fragmented spectrum use is claimed achievable using noncontiguous waveforms, such as Cancellation Carriers (CC) or Edge Windowing (EW). In[38], CC-OFDM with a single cancellation carrier is shown better than both its variant with multiple (weighted) cancellation carriers and EW-OFDM, but suppression performance degrades as the subcarrier index runs away from the gap edge.  

Another option consists of manipulating OFDM to mimic SCM[2][3] to reduce PAPR and provide robustness against frequency offsets. Discrete Fourier Transform spread OFDM (DFT-s-OFDM) enhances noise in faded channels, offers poor
spectral containment, and allows neither frequency-selective scheduling nor link adaptation. These limitations are overcome by employing zero-tail DFT-s-OFDM, exploiting receiver diversity, and applying the DFT spread at the physical resource blocks level [39][40]. Improved flexibility (dynamic overhead adaptation instead of CP hardcoding) and OOB emissions (smoother transitions between adjacent symbols) are additional advantages of zero-tail DFT-s-OFDM over CP-OFDM.

B. Filter Bank Multicarrier

FBMC introduced multicarrier techniques over two decades before the introduction of OFDM in wireless communications systems [8]. While OFDM relies on the CP to prevent ISI and to convert the channel into a set of flat-gain subcarriers, FBMC exploits the fact that narrow and numerous subcarriers can be characterized by a flat gain. The length and superior frequency localization of FBMC prototype filters allow the terminal to deal with high delay spreads and compensate frequency offsets without feedback to the BS [2][3] (at the expense of increased complexity, latency, and equalization requirements).

Fast-convolution based highly tunable multirate filter banks are investigated in [41]. Capable of implementing waveform processing for multiple single-carrier and/or multicarrier transmission channels with nonuniform bandwidths and subchannel spacings simultaneously, this method is a competitive option in terms of spectral containment and complexity.

C. “Born-to-be-5G” Waveforms

In contrast to FBMC or OFDM, which apart from enhancements like CC-OFDM and DFT-s-OFDM were not originally designed bearing 5G requirements in mind, we have recently witnessed the outbreak of waveforms crafted for MTC and the Tactile Internet. Biorthogonal Frequency Division Multiplexing (BFDM) waveforms, for instance, have been regarded as suitable to support sporadic data traffic and asynchronous signaling. One appealing feature of BFDM is time and spectral localization balancing through iterative interference cancellation, which in turn allows to control degradations due to time and frequency offsets [42].

Rendered attractive for nonsynchronous burst transmissions by its block-based structure, Generalized Frequency Division Multiplexing (GFDM) was originally proposed as a nonorthogonal alternative to FBMC. GFDM can be set to mimic OFDM, although its benefits are better experienced with SCM setups, e.g. to transmit multiple symbols per subcarrier. To the best of our knowledge, GFDM is the only 5G candidate waveform for which support for High-order Modulation (HOM) and Tactile Internet has been explicitly investigated. We refer the reader to [43][44] for a recent analysis of characteristics, relevant features, performance, and implementation aspects of GFDM.

D. Universal Filtered Multicarrier

Another distinguishing aspect of OFDM and FBMC is that the former applies filtering on the whole band, while the latter works on a per subcarrier basis. Universal Filtered Multicarrier (UFMC) has been advanced as a more general solution because its filtering is applied on the level of multiple subcarriers, e.g. on a per resource block basis. As compared with OFDM, UFMC offers better spectral efficiency and robustness against time and frequency offsets [2][4]. Some advantages of UFMC over FBMC are lower latency (due to its shorter filter lengths), reduced overhead, and improved support for MTC [3][5] – although both may require more involved multi-tap equalizers.

IV. FBMC-BASED MASSIVE MIMO NETWORKS

Networks resulting from the combination of Massive MIMO and FBMC are of the utmost importance as in these systems spectrum not only can be reused by all the users, i.e. advantage of Massive MIMO, but can also be used in an efficient manner, i.e. due to the FBMC’s low OOB emissions. The application of FBMC to Massive MIMO was first considered in [9], with an interesting finding of this work being the self-equalization property of FBMC in Massive MIMO channels (in contrast with the limited applicability of FBMC to conventional MIMO.
As mentioned in [9], Cosine Modulated Multi-tone (CMT), viz. a particular FBMC form, has a blind equalization capability [50], which can be used to decontaminate erroneous channel estimates in multicellular Massive MIMO networks caused by pilot contamination. This approach, which is somewhat similar to the Godard blind equalization algorithm, is particularly suitable for the Doppler-limited uplink of Massive MIMO systems. To alleviate the performance loss caused by corrupted channel estimates, the blind equalization technique adaptively corrects the linear combiner tap weights. Starting with MMSE using noisy channel estimates, the SINR performance reaches that of MF with noise-free channel matrix within a small number of iterations, and keeps improving to get to that of MMSE with noise-free channel matrix.

This is shown in Figure 2 for a Massive MIMO network in TDD mode with seven cells and one user per cell. All users are assumed to use the same pilot sequences, $L = 256$, and $M = 128$. In some situations where the length of the UL data packets is close to the channel coherence time, the estimated CSI in the beginning of the packet may get outdated, resulting in a performance loss (as the same CSI is used for precoding in the DL). This problem can be alleviated by the utilization of blind channel tracking techniques as the one mentioned above, since these provide up-to-date CSI. In other words, the blind channel tracking techniques sweep through all the symbols in the data packet and update the CSI. The latest CSI taken from the last transmitted symbols can thus be obtained.

V. CONCLUSION

This paper has reviewed existing work and identified main challenges at the intersection of Massive MIMO and waveform design. The property of self-equalization was introduced for FBMC-based Massive MIMO systems, which can help reduce the number of subcarriers required by the system. It is also shown that the blind channel tracking property of FBMC can be used to address pilot contamination – a major limiting factor of Massive MIMO systems. Our findings shed light into and motivate for an entirely new research avenue towards a better understanding of waveform design for 5G with a particular emphasis on FBMC-based Massive MIMO networks.

REFERENCES

[1] “FP7 European Project 317669 METIS (Mobile and wireless communications Enablers for the Twenty-twenty Information Society).”, [Online]. Available: https://www.metis2020.com/
[2] J.G. Andrews et al., “What Will 5G Be?”, IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1065-1082, June 2014.
[3] G. Wunder et al., “5GNOW: non-orthogonal, asynchronous waveforms for future mobile applications”, IEEE Commun. Mag., vol. 52, no. 2, pp. 97-105, Feb. 2014.

![Fig. 1. SINR comparison: MMSE vs. MF ($K = 6$, $L = 64$, $M = 128$).](image1)

![Fig. 2. SINR comparison: Blind tracking technique vs. MF vs. MMSE ($L = 256$, $M = 128$).](image2)
