Ultra Reliable Low Latency Communications in Massive Multi-Antenna Systems

Bana, A.-S.; Xu, G.; Carvalho, E.D.; Popovski, P.

Published in:
2018 52nd Asilomar Conference on Signals, Systems, and Computers

DOI (link to publication from Publisher):
10.1109/ACSSC.2018.8645402

Publication date:
2019

Document Version
Accepted author manuscript, peer reviewed version

Ultra Reliable Low Latency Communications in Massive Multi-Antenna Systems

Alexandru-Sabin Bana, Guixian Xu, Elisabeth De Carvalho and Petar Popovski
Department of Electronic Systems, Aalborg University, Denmark
Email: {asb, gx, edc, petarp}@es.aau.dk

***Abstract*—Large number of antennas in Massive MIMO offer a significant spatial diversity, which makes them an attractive possibility for use in wireless settings that require very high reliability. However, in 5G ultra-reliability is coupled with low latency into ultra-reliable low-latency communications (URLLC). This is very challenging as an efficient use of Massive MIMO depends critically on training, which consumes significant resources when the latency requirement is very tight. In this paper we address this problem by exploiting the sparsity of the propagation channel and therefore rely on estimation of a small number of instantaneous channel coefficients. This leads to robust beamforming and departs from the conventional use of the instantaneous channel state information (CSI) at each transmit antenna. We compare the performance of maximum ratio transmission based on the conventional least-squares estimation of all channel coefficients and the one based on the estimation of the fading coefficients of the channel features i.e. the singular vectors of the covariance matrix. The singular vectors are assumed known and unchangeable over a long term. The results show that this approach makes massive MIMO a feasible technology in URLLC scenarios.**

I. Introduction

Massive antenna arrays at the base station (BS) or at the terminal of a wireless system are seen as a prominent enabler of more efficient communication and as a key enhancement for 5G wireless networks [1]. Massive MIMO is essential in achieving the high throughput and spectral efficiency required for next generation networks [2]. However, massive antenna arrays are also seen as means of improving the other two 5G services, namely massive machine type communications [3] and ultra-reliable low-latency communications (URLLC) [1], [4]–[6]. Massive MIMO sets a new paradigm and performance expectations from the physical layer, which is expected to complement higher layer techniques in order to achieve the stringent requirements of URLLC [7]–[10].

Massive arrays offer a large number of spatial degrees-of-freedom (DoF) [9], which results in two properties that are directly relevant for URLLC.

1) High signal-to-noise ratio (SNR) links due to the *array gain*, and links that become quasi-immune to fast fading, due to a phenomenon known as *channel hardening* [11]. It should be briefly mentioned that the latter can be achieved only in channels with rich scattering, and can heavily reduce the need for packet retransmissions or frequency diversity, as the channel becomes less likely to be subject to deep fades.
2) High spatial multiplexing capabilities provide lower and more evenly experienced latency if several terminals need to be served simultaneously, which is an important property, particularly for Tactile Internet [12].

In general, the benefits of massive antenna systems are conditioned on the acquisition of the channel state information (CSI) [13], be it in the form of instantaneous CSI for each antenna or long-term statistical CSI (channel covariance matrix). The most critical acquisition occurs at a massive array when the CSI is used for transmission mode (CSIT). In time division duplexing (TDD), the uplink and downlink channels are assumed reciprocal so that the channel coefficients at all antennas are estimated using uplink training and downlink transmission is performed based on the resulting channel estimates. However, when the channel is spatially sparse, channel estimation becomes better when the sparsity of the channel is exploited, as the number of instantaneous channel coefficients to be estimated is smaller. This assumes that the structure of the channel, i.e. the propagation directions, can be estimated, as done in sparse channel estimation methods.

In URLLC, training is critical as it might take a significant portion of the packet, leaving less time to transmit the data and the quality of the channel estimation has a major impact. This is the reason why it becomes a primary concern to exploit the channel structure in URLLC.

In order to promote reliability and low latency, the general purpose of the beamforming design is to rely as much as possible on the structure of the channel (i.e., the propagation paths, varying on a large scale basis) and as less as possible on the small scale fading properties, while still benefiting from the properties brought by the massive number of antennas. Previous works have considered transmission using the long-term CSIT, [14]–[18], where the authors consider beamforming along the singular vectors of the channel covariance matrix, with several power allocation criteria between the singular vectors. In [19], the authors consider the combination of long-term channel statistics and instantaneous channel norm, fed back in quantized form, to be sufficient for precoding in wide area scenarios. Multiplexing users with either statistical CSI or imperfect instantaneous CSI has been treated in [20], where the interference between the two types of users is investigated. The authors of [17] have shown that a beamforming design relying on combining partial mean and covariance CSIT is beneficial in order to improve the expected weighted sum rate. Further, it is shown in [18] that even imperfect channel covariance information can improve the spectral efficiency in massive MIMO.

This paper further investigates the means of exploiting the multiple antennas at the transmitter to ensure URLLC, for which it appears preferable to combine instantaneous CSI with long-term statistical CSI. In particular, we compare the performance of maximum ratio transmission based on the conventional least-squares estimation of all channel coefficients and the one based on the estimation of the fading coefficients of each channel features, here the singular vectors of the covariance matrix. We assume here that the singular vectors are perfectly known.

Throughout the paper we will use boldface small ($\mathbf{x}$) and boldface capital letters $\mathbf{X}$ to denote vectors and matrices, respectively. The superscripts $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote the conjugate, transpose and the conjugate transpose operations, respectively.

II. System model and Multi-Antenna Transceiver Designs

A. Scenario and Channel Model

The basic scenario is displayed as a factory-type environment in Fig. 1. An access point is equipped with an array comprising a very large number of antennas (M) which serves single-antenna terminals (actuators). The terminals are served in a time division multiplexing manner.

We adopt a cluster-based channel model where each cluster is characterized by a group of localized propagation paths defined by their angle of departure (AoD) and their angle of arrival (AoA). Each propagation path is affected independently by an attenuation factor that follows a certain distribution. Furthermore, if we consider a power decay throughout the paths, this is also modeled with the attenuation factor. The attenuation factor is then used as a standard deviation in order to generate the random complex fading coefficients α_i in (1).

The channel from the access point to a terminal is described as the sum of the propagation paths over all the clusters (in the sum, we make no distinction between clusters):

$$\mathbf{h} = \sum_{i=1}^{N_P} \alpha_i \mathbf{s}_i. \tag{1}$$

Here, $\mathbf{h}$ is a column of size M, N_P is the total number of paths and the column vector $\mathbf{s}_i$ is the normalized steering vector (SV) for each path, defined as

$$\mathbf{s}_i = \frac{1}{\sqrt{M}} \exp\left(-j2\pi(k-1)\frac{\delta}{\lambda}\Omega\right), k = 1 : M \tag{2}$$

where $\Omega = \cos\left(\frac{\pi}{M}(k-1) + \frac{\pi}{M}\mathrm{rand}(0,1)\right)$ is the random angle of each beam. The direction of the propagation paths correspond to long-term statistics, meaning that for a localized movement of the terminal the directions remain unchanged, whereas the coefficients $\{\alpha_i\}$ correspond to small scale fading and vary for small movements. Furthermore, we assume that each steering vector has a relative power component, such that we can define a decay parameter between the strongest and the weakest steering vector ($0\,\mathrm{dB}$ decay means equal power throughout the SVs, whereas higher decay means exponentially decreasing power between the strongest and the weakest SVs that is equal to the decay value).

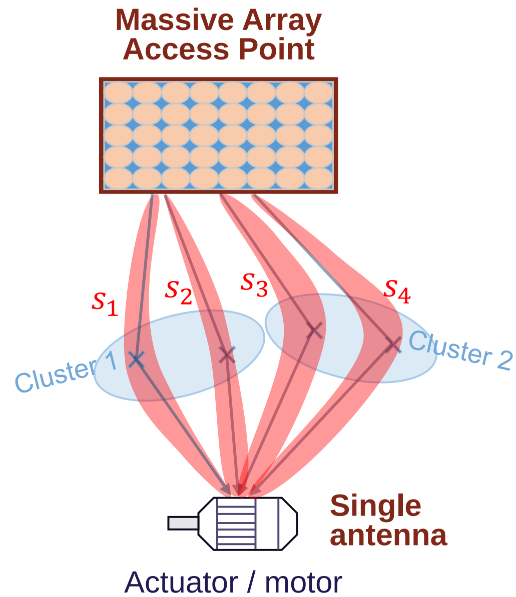


Fig. 1: Factory Scenario: one massive MIMO access point beamforming towards a terminal using a coherent combination accross the structure of the channel.

B. Transceiver Designs

We assume the transceiver to operate in TDD mode. Therefore, the channel is estimated at the access point using uplink pilots sent by the terminal. The uplink training is followed by a switching time, necessary due to propagation delay and the processing time at the access point, as shown in Fig. 2.

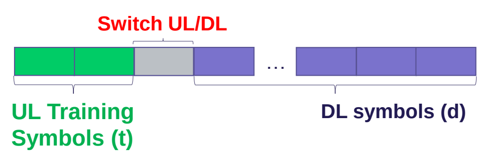


Fig. 2: Transceiver operation and frame structure.

The channel coefficients are estimated from UL training symbols using the least-squares method. The UL received training signal is then the M by t matrix $\mathbf{Y}_t = \mathbf{hp} + \mathbf{N}$, with $\mathbf{p}$ being the row vector of pilot symbols used for training, and $\mathbf{N}$ being complex additive white Gaussian noise with zero mean and σ_n^2 variance. Therefore, the complex channel estimate coefficients $\hat{\mathbf{h}}$ can be expressed as

$$\widehat{\mathbf{h}} = \frac{1}{N_t}\mathbf{Y}_t\mathbf{p}^H. \tag{3}$$

The least squares channel estimate can also be expressed in terms of the true value of the channel and the estimation error as

$$\widehat{\mathbf{h}} = \mathbf{h} + \mathbf{e}, \tag{4}$$

where $\mathbf{e} \sim \mathcal{CN}(0, \sigma_e^2 \mathbf{I})$ and $\sigma_e^2 = \sigma_n^2/(\sigma_h^2 N_t)$.

C. Beamforming with maximum ratio transmission (MRT)

Maximum ratio transmission (MRT) relies on matched filtering using the instantaneous estimated channel coefficients at each antenna. Therefore, the transmitter uses the filter coefficients $\mathbf{w}_{\text{MRT}} = \widehat{\mathbf{h}}^H/\|\widehat{\mathbf{h}}\|$, such that the received signal at the terminal is

$$y = \frac{\widehat{\mathbf{h}}^H \mathbf{h} x}{\|\widehat{\mathbf{h}}\|} + n. \tag{5}$$

In this case, the downlink signal-to-noise ratio (SNR) can be expressed as

$$\gamma_{\text{MRT}}^{\text{DL}} = \frac{\sigma_h^2}{\sigma_n^2} \frac{\left|\widehat{\mathbf{h}}^H \mathbf{h}\right|^2}{\left\|\widehat{\mathbf{h}}\right\|^2}. \tag{6}$$

D. Beamforming exploiting channel sparsity

Here we review how beamforming is performed based on the covariance matrix of the received signal of the massive array. We assume that the covariance matrix is perfectly known. In practice, its estimation is done on a longer-term than the instantaneous CSI.

The covariance matrices can be written in two different ways: as a function of the steering vectors and using the singular value decomposition. The covariance matrix of the received signal at the BS is:

$$\mathbf{R} = \mathbf{V} \Lambda \mathbf{V}^H. \tag{7}$$

Equation (7) describes the singular value decomposition of $\mathbf{R}$: the columns of $\mathbf{S}$ comprise the singular vectors denoted as $\mathbf{v}_i$ and Λ is a diagonal matrix grouping the non zero singular values. Without loss of generality, we assume that the dimension of the signal subspace is equal to the number of paths N_p, so we can express the channel coefficients as

$$\mathbf{h} = \sum_{i=1}^{N_p} \beta_i \mathbf{v}_i. \tag{8}$$

The beamforming is performed both using the second order statistics from the covariance matrix and the instantaneous channel estimate $\widehat{\mathbf{h}}$ from (4) in the following manner. The instantaneous CSI estimate $\widehat{\mathbf{h}}$ is projected on the subspace spanning *all* of the singular vectors (SVs) of the channel, in order to get an estimate of the fading coefficients $\widehat{\alpha_i} = \widehat{\mathbf{h}}^H \mathbf{v}_i$. The fading coefficients estimates for all paths can be expressed as a row vector $\widehat{\boldsymbol{\alpha}} = \widehat{\mathbf{h}}^H \mathbf{V}$, which is projected onto the channel propagation subspace, resulting in the projected estimated channel being $\mathbf{h}_p = \widehat{\mathbf{h}}^H \mathbf{V} \mathbf{V}^H$. The beamforming is the corresponding normalized matched filter:

$$\mathbf{w}_{\text{allSV}} = \frac{\mathbf{V} \mathbf{V}^H \widehat{\mathbf{h}}^*}{\left\|\mathbf{V} \mathbf{V}^H \widehat{\mathbf{h}}^*\right\|}. \tag{9}$$

Thus, the downlink SNR for this method (which we denote SV - since it is based on coherently beamforming across the singular vectors) can be written as

$$\gamma_{\text{SV}}^{\text{DL}} = \frac{\sigma_h^2}{\sigma_n^2} \frac{\left|\mathbf{h}^T \mathbf{V} \mathbf{V}^H \widehat{\mathbf{h}}^*\right|^2}{\left\|\mathbf{V} \mathbf{V}^H \widehat{\mathbf{h}}^*\right\|^2}. \tag{10}$$

E. Post processing SNR - large scale analysis

By projecting the estimated instantaneous CSI $\hat{\mathbf{h}}$ onto the subspace spanned by the singular vectors $\mathbf{S}$ we essentially perform a denoising operation, by removing the estimation error components that are outside of the projection subspace.

In this subsection, we provide approximate DL SNR expressions for the scheme relying only on a *partial set of SVs* (denoted by SV) in (11), then generalize it to the case of using *all* of the SVs (allSV) available in (12).

Substituting $\widehat{\mathbf{h}}$ from (4) in (10) and neglecting the cross-terms when expanding the numerator and denominator in (10) we get

$$\gamma_{\text{SV}}^{\text{DL}} \approx \frac{\sigma_h^2}{\sigma_n^2} \left(\mathbf{h}^H \mathbf{V}_s \mathbf{V}_s^H \mathbf{h}\right) \frac{\mathbf{h}^H \mathbf{V}_s \mathbf{V}_s^H \mathbf{h} + \sigma_e^2}{\mathbf{h}^H \mathbf{V}_s \mathbf{V}_s^H \mathbf{h} + N_s \sigma_e^2} \tag{11}$$

where N_s is the number of SVs used, $\mathbf{V}_s$ is the matrix containing only the columns with the used SVs, and which for the case of using all SVs becomes

$$\gamma_{\text{allSV}}^{\text{DL}} \approx \frac{\sigma_h^2}{\sigma_n^2} \left(\mathbf{h}^H \mathbf{h}\right) \frac{\mathbf{h}^H \mathbf{h} + \sigma_e^2}{\mathbf{h}^H \mathbf{h} + N_p \sigma_e^2}. \tag{12}$$

The large scale analysis can be applied to the case of MRT as well, in which case the SNR can be approximated similarly to (12), except that N_p becomes M, which is the reason for worse performance. The approximation reads

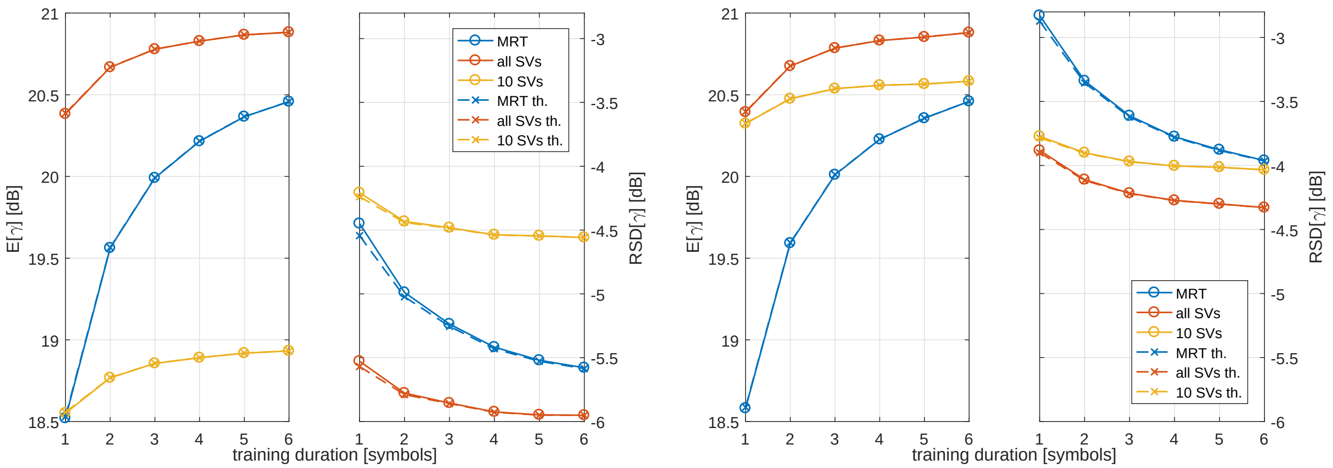
$$\gamma_{\text{MRT}}^{\text{DL}} \approx \frac{\sigma_h^2}{\sigma_n^2} \left(\mathbf{h}^H \mathbf{h}\right) \frac{\mathbf{h}^H \mathbf{h} + \sigma_e^2}{\mathbf{h}^H \mathbf{h} + M \sigma_e^2}. \tag{13}$$

F. Downlink Transmission and Outage Probability

We evaluate the reliability of the downlink transmission of a packet of a fixed size of b bits. We assume that the training takes t symbols, and due to the lantecy constraint and the symbol lost during the switching time (Fig. 2), we are left with d symbols for the data transmission. Therefore, the data rate is $R = \frac{\text{b}}{2\text{d}}$ (one symbol is regarded as 2 channel uses due to having a complex channel [21, footnote 1]), from which we can determine the SNR threshold required to successfully receive a packet as $\gamma_{\text{th}} = 2^R - 1$. Since the transmit power is limited, the SNRs expressed in (10) and (6) may fall below the threshold. We evaluate this outage probability by using

$$P_{\text{outage}} = \Pr\left[\gamma^{\text{DL}} < \gamma_{\text{th}}\right]. \tag{14}$$

The main reason for considering the outage probability as a suitable evaluation metric is because the packet errors are generally due to not meeting the SNR requirement in the downlink, rather than the random noise component at the terminal. Finite blocklength theory states that the outage probability is a correct performance metric even for short lengths [22].



(a) decay $= 0\,\mathrm{dB}$ (b) decay $= 20\,\mathrm{dB}$

Fig. 3: Dependency of the mean SNR (left subplot) and relative standard deviation (RSD) of SNR (right subplot) on the training length, for different transmission schemes, and for two channels without (3a) and with (3b) power decay between SVs. Pre-processing SNR $= 4\,\mathrm{dB}$, $N_p = 20$. Circle markers denote simulations and "x" markers denote theoretical large scale analysis.

III. Performance Evaluation

We perform evaluations of the post-processing SNR and packet outage probability via Monte-Carlo simulations, considering the parameter assumptions listed in Table I.

Packet length	32 bytes (256 bits)
Subcarrier spacing	$60\,\mathrm{KHz}$
Latency	2 slots $\times$ 14 symbols $= 0.5\,\mathrm{ms}$
BS antennas	$M = 100$
Terminal antennas	1
Moderate channel scattering	$N_p = 20$ paths
Power decay of paths	$[0, 20]\,\mathrm{dB}$

TABLE I: System parameters.

We show the mean SNR and its relative standard deviation (RSD) in Fig. 3 for channels with $N_p = 20$ paths and either equal power SVs (Fig. 3a) or with $20\,\mathrm{dB}$ decay between the strongest and weakest SVs (Fig. 3b). The first observation we can draw from Fig. 3 is that the scheme relying on coherent transmission along all the SVs outperforms MRT both in terms of mean SNR and RSD of the SNR (the lower RSD, the better). The reason for this is that the imperfect instantaneous CSI is projected on the channel structure, thereby being denoised. Furthermore, it can also be noticed that the training duration has a greater impact on the MRT scheme than on the proposed schemes, as the mean and RSD of the SNR tend to vary more depending on the number of training symbols.

Comparing the mean SNRs from Figures 3a and 3b, we notice that in the case of power decay, the mean SNR of the strongest 10 SVs scheme has improved. This is due to the fact that more power is concentrated in the 10 strongest SVs. On the other hand, on the RSD subplot we notice an increase for all schemes due to the power decay. This corresponds to a diversity loss, since some of the paths in the channel experience lower power.

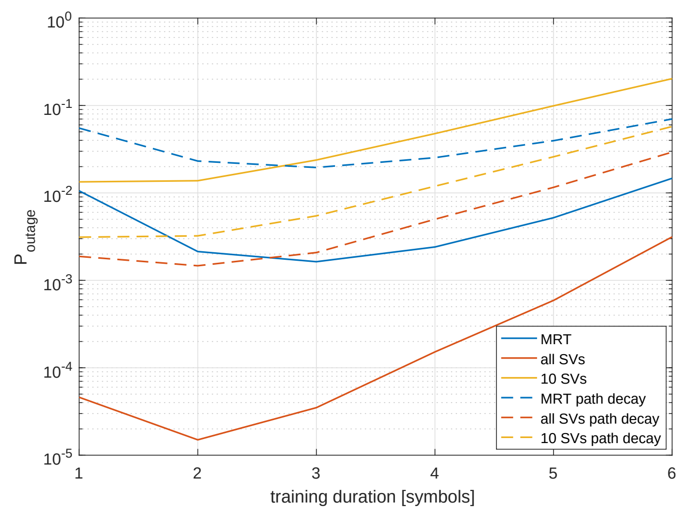


Fig. 4: Dependency of the outage on the training length, for different transmission schemes and path power decay. Solid lines denote no power decay between paths, and dashed lines denote $20\,\mathrm{dB}$ decay between the strongest and weakest path.

When looking at the outage probability in Fig. 4, we notice a similar ranking of the methods as described by the RSD in Fig. 3. However, it should be also noticed that since we deal with a latency constrained scenario, there is an optimum number of training symbols that minimizes the outage probability. The optimum number of training symbols is smaller for the schemes based on SVs than for MRT, as expected after seeing in Fig. 3 that a higher number of training symbols is not as beneficial for SV schemes as it is for MRT.

In Fig. 5 we show what is the minimum pre-processing SNR required to achieve the target reliability of $1 - 10^{-5}$. The vertical lines are the SNRs required for MRT in the case

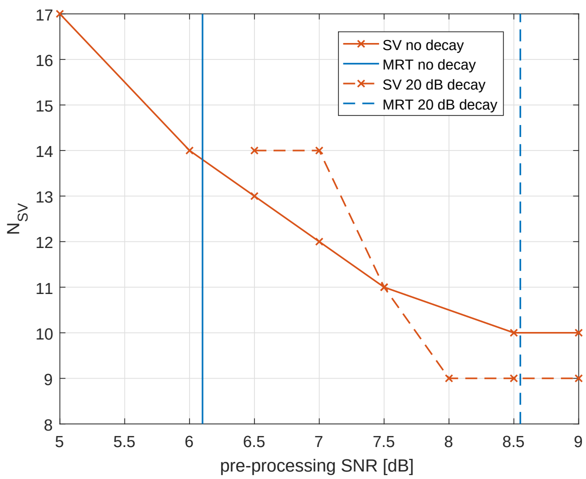


Fig. 5: Minimum number of SVs for given pre-processing SNR in order to achieve a target reliability of $1 - 10^{-5}$, shown in solid lines for a channel with no decay among paths and in dashed lines for $20\,\mathrm{dB}$ decay between the strongest and weakest path.

of no decay (solid) and $20\,\mathrm{dB}$ decay (dashed). For the scheme relying on the channel structure, we show also how many SVs are needed to use in order to achieve the reliability for a specific pre-processing SNR.

IV. Conclusion and Future work

In this paper we have investigated a beamforming design based on both imperfect instantaneous CSI and the channel structure. Specifically, we consider a precoder which projects the imperfect instantaneous CSI onto the subspace spanned by the SVs of the channel covariance matrix. This design proves to be suitable for URLLC, as it not only improves the mean SNR, but also reduces the relative standard deviation of the SNR, thereby improving reliability compared to the case of MRT.

Possible extensions of this work may consider evaluating the performance of the scheme with imperfect channel covariance matrix estimation, as well as multi-user multiplexing.

References

[1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, 2014.

[2] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for Maximal Spectral Efficiency: How Many Users and Pilots Should Be Allocated," *IEEE Trans. on Wireless Comm.*, vol. 15, no. 2, Feb. 2016.

[3] L. Liu and W. Yu, "Massive device connectivity with massive MIMO," *Int. Symp. Inf. Theory (ISIT)*, Jun. 2017.

[4] S. R. Panigrahi, N. Björsell, and M. Bengtsson, "Feasibility of Large Antenna Arrays towards Low Latency Ultra Reliable Communication," *IEEE Int. Conf. on Industrial Technology (ICIT)*, Mar. 2017.

[5] T. K. Vu, C.-F. Liu, M. Bennis, M. Debbah, M. Latva-aho, and C. S. Hong, "Ultra-Reliable and Low Latency Communication in mmWave-Enabled Massive MIMO Networks," *IEEE Commun. Lett.*, no. 99, May 2017.

[6] A.-S. Bana, M. Angjelichinoski, E. De Carvalho, and P. Popovski, "Massive MIMO for ultra-reliable communications with constellations for dual coherent-noncoherent detection," *22nd International ITG Workshop on Smart Antennas (WSA)*, Mar. 2018.

[7] H. Chen, R. Abbas, P. Cheng, M. Shirvanimoghaddam, W. Hardjawana, W. Bao, Y. Li, and B. Vucetic, "Ultra-Reliable Low Latency Cellular Networks: Use Cases, Challenges and Approaches," *IEEE Commun. Mag.*, Sep. 2018.

[8] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, "Ultra-Reliable and Low-Latency Communications in 5G Downlink: Physical Layer Aspects," *IEEE Wireless Commun. Mag.*, vol. 25, no. 3, Jun. 2018.

[9] P. Popovski, J. J. Nielsen, Č. Stefanović, E. De Carvalho, E. Strom, K. F. Trillingsgaard, A. Bana, D. M. Kim, R. Kotaba, J. Park, and R. B. Sorensen, "Wireless Access for Ultra-Reliable Low-Latency Communication: Principles and Building Blocks," *IEEE Network*, vol. 32, no. 2, pp. 16–23, Mar. 2018.

[10] P. Popovski, Č. Stefanović, J. J. Nielsen, E. De Carvalho, M. Angjelichinoski, K. F. Trillingsgaard, and A.-S. Bana, "Wireless access in ultra-reliable low-latency communication (urllc)," 2018.

[11] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in Massive MIMO," *Proc. 22nd European Sign. Proc. Conf. (EUSIPCO)*, 2014.

[12] W. Tarneberg, M. Karaca, A. Robertsson, F. Tufvesson, and M. Kihl, "Utilizing Massive MIMO for the Tactile Internet: Advantages and Trade-Offs," *IEEE Int. Conf. Sensing, Comm. and Networking (SECON Workshops)*, 2017.

[13] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, 2014.

[14] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, no. 5, May 2002.

[15] O. Simeone, U. Spagnolini, and Y. Bar-Ness, "Linear and non-linear precoding/decoding for MIMO systems using the fading correlation at the transmitter," *4th IEEE Workshop on Signal Processing Advances in Wireless Commun. (SPAWC)*, 2003.

[16] O. Simeone and U. Spagnolini, "Combined linear pre-equalization and BLAST equalization with channel correlation feedback," *IEEE Commun. Lett.*, vol. 7, no. 10, Oct. 2003.

[17] W. Tabikh, Y. Yuan-Wu, and D. Slock, "Beamforming design with combined channel estimate and covariance CSIT via random matrix theory," *IEEE Int. Conf. Commun. (ICC)*, May 2017.

[18] E. Bjrnson, L. Sanguinetti, and M. Debbah, "Massive MIMO with imperfect channel covariance information," *50th Asilomar Conf. Signals, Syst. and Computers*, Nov. 2016.

[19] E. Björnson, D. Hammarwall, R. Zakhour, M. Bengtsson, D. Gesbert, and B. Ottersten, "Feedback design in multiuser MIMO systems using quantization splitting and hybrid instantaneous/statistical channel information," *Proc. ICT Mobile and Wireless Commun. Summit*, 2008.

[20] S. Qiu, D. Chen, D. Qu, K. Luo, and T. Jiang, "Downlink Precoding With Mixed Statistical and Imperfect Instantaneous CSI for Massive MIMO Systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, Apr. 2018.

[21] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proc. IEEE*, vol. 104, Sep. 2016.

[22] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-Static Multiple-Antenna Fading Channels at Finite Blocklength," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232 – 4265, Jul. 2014.

Acknowledgment

The work has partly been supported by the Danish Council for Independent Research (DFF) (CELEST project), by the European Research Council (ERC Consolidator Grant nr. 648382 WILLOW), and partly by the Horizon 2020 project ONE5G (ICT-760809).