# An introduction for Massive MIMO system: A survey

Yuanyuan Tang
Electrical & Computer Engineering, University of Virginia, yt5tz@virginia.edu

***Abstract*—Massive MIMO is considered to be one of the most promising technologies for wireless communication. By deploying hundreds or thousands of antennas at the base station, Massive MIMO can provide higher energy efficiency and spectral efficiency. With reliable channel state information at the base station estimated from uplink training sequences, three linear precoding methods can achieve near-optimal performance. Furthermore, numerical results show that three precoding methods achieve good bit-error rate (BER) with respect to the signal-to-noise ratio (SNR) and the number of antennas at the base station. Finally, We conclude some limiting factors and future research to further improve the performance and future applications.**

## I. Introduction

In recent years, the significantly increasing number of mobile devices such as cellphones has posted great challenges to wireless communication [1], [2]. Because of the high throughput, efficiency and reliability, the multiple-input multiple-output (MIMO) technology has been studied for more than ten years and is widely used in the third generation (3G), IEEE 802.11ac/n (WiFi), IEEE 802.16e (WiMAX), LTE/LTE-A (4G), and other protocols and systems [1]–[3].

Originated from the point-to-point MIMO, Massive MIMO, also called large-scale antenna system or multi-user MIMO (MU-MIMO), is proposed to keep the benefits of MIMO and further improve the system performance [3], [4]. It is considered a core component for the fifth generation (5G) network [5]. In Massive MIMO system, the base station (BS) deployed with hundreds of antennas can simultaneously serve tens of single-antenna terminals at the same frequency domain and without extra power consumption [2]–[4], [6]–[8], showed in Figure 1. In general, there are two assumptions for Massive MIMO: 1) the number of users the BS serves simultaneously is less than the number of antennas at the base station; 2) The Massive MIMO system works in a time-division duplexing (TDD) operation [1]. In the frequency-division duplexing (FDD) mode, the amount of time required by BS to obtain channel state information (CSI) depends on the number of antennas at the base station [1]. In TDD operation, the time required to acquire CSI depends on the number of users in the cell. In the massive MIMO system, since the number of users is obviously smaller than that of antennas at the base station, the TDD operation is a better candidate.

By deploying hundreds of antennas at the BS, Massive MIMO provides many advantages for wireless communication:

- **High multiplexing gains.** Massive MIMO with hundreds of antennas at BS provides progressive spatial multiplexing gains, which can increase the capacity 10 times [4], [7].
- **Comparatively low deployment cost.** Since hundreds of antennas are only deployed at the expensive BS, single-antenna terminals can achieve high multiplexing gain. Then the cheap devices make Massive MIMO attractive for users [2].
- **Orthogonal user channels.** Since the BS has hundreds of antennas, based on large-number theorem, channels between the BS and different users are asymptotically orthogonal. Then the orthogonality can significantly reduce the inter-user interference and improve the performance [6].
- **High energy efficiency.** Without increasing the transmitting power, $N$ antennas at BS can generate $N$ times stronger signal at receivers [8]. If each of $K$ users is allocated with $1/k$ of the original radiated power, a receiver still gets higher signal if $N \gg K$ [8]. Hence, Massive MIMO improves the energy efficiency by the multiplexing gains [4], [7], [8].
- **High spectral efficiency.** Since the system capacity is improved within the same frequency domain, Massive MIMO achieve higher spectral efficiency [4], [7], [8].
- **Increased robustness and reliability.** By suitable precoding methods, the data is sent to intended users with beamformings. With the number of antennas increasing to infinity, the influence of small-scale fading and uncorrelated noise vanish [4], [7].

To achieve the advantages of Massive MIMO system, the CSI estimate and precoding methods play a very important role. In Massive MIO system, the CSI of the downlink is estimated based on training sequences in

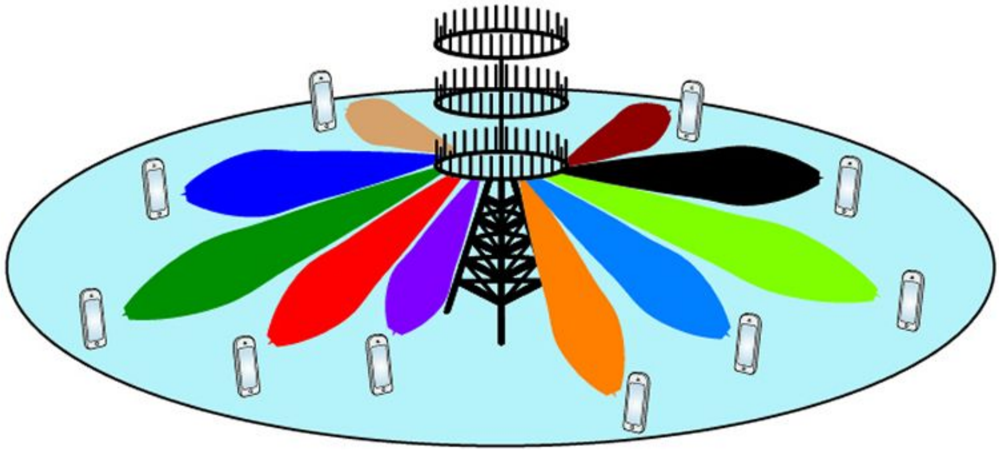

Figure 1. A single-cell Massive MIMO system [5]

the uplink [1], [4]. With reliable CSI in the downlink, precoding methods can significantly reduce the inter-user interference and send data to intended terminals by beamforming [1]. Fortunately, in the downlink, some linear precoders can achieve good performance, including matched filter (MF), also called maximum ratio transmission (MRT) beamforming, Zero-forcing, and regularized zero-forcing (RZF), also called minimum mean square error (MMSE) [1], [8].

This paper is a brief survey of Massive MIMO system. Section I introduces the concept, the advantages, and two important challenges of Massive MIMO system. The channel model of the downlink is introduced in Section II. With reliable CSI, three precoding methods are introduced in Section III and the numerical results are showed in Section IV. After that Section V shows some limitations and future research for Massive MIMO. Finally, Section VI concludes the paper.

## II. Channel models

Following the two assumptions in Section I, this section introduces the CSI in the downlink. In TDD operation, based on the reciprocity of the channels, the CSI of the downlink can be estimated by training sequences in the uplink from terminals [4]. To simplify the channel model, this model only consider the single cell scenario. Furthermore, the estimated CSI in the downlink is considered perfect.

In a single-cell Massive MIMO system, one BS equipped with $N$ antennas serves $K$ single-antenna users simultaneously [2]. With the help of the uplink training sequences, the CSI from the BS to $K$ single-antenna users can be considered as $N \times K$ matrix $\mathbf{H}$. Based on [2], the $i$-th antenna of the BS to the $j$-th user has the channel coefficient $h_{i,j}$ that can be written as

$$h_{i,j} = g_{i,j}\sqrt{d_j},$$

where $g_{i,j}$ denotes the complex small-scale fading factor, and $d_j$ represents the large-scale fading factor for user $j$. Since the large-scale fading coefficients are related to the distance between users and the BS, they stay the same for different antennas at the BS to the same user, but vary among different users [3]. Furthermore, the small-scale fading coefficients are assumed to be different for different pairs among antennas in the BS and users [3]. Hence, then the channel matrix from the BS to $K$ users can be written as

$$\mathbf{H} = \mathbf{G}\mathbf{D}^{1/2},$$

where $\mathbf{G} \in \mathcal{C}^{N\times K}$ denotes the small-scale fading matrix with $\mathbf{G}_{i,j} = g_{i,j}$, and the diagonal matrix $\mathbf{D}^{1/2} = \mathrm{diag}\{d_1, \ldots, d_K\} \in \mathcal{C}^{K\times K}$ denotes the large-scale fading matrix. Since hundreds of antennas are deployed at the BS, the channel vectors from the BS to different users are asymptotically orthogonal [2]. Therefore, the columns of the channel matrix $\mathbf{H}$ to different users are asymptotically orthogonal as the number of antennas at the BS, $N$, grows to infinity. Then we have

$$\mathbf{H}^H\mathbf{H} = \mathbf{D}^{1/2}\mathbf{G}^H\mathbf{G}\mathbf{D}^{1/2} = N\mathbf{D},$$

where $\mathbf{H}^H$ denote the conjugate transpose of the channel matrix $\mathbf{H}$.

In the TDD operation, if the perfect CSI in the downlink can be estimated, the received vector $\boldsymbol{y} \in \mathcal{C}^{N_r\times 1}$ can be written as

$$\boldsymbol{y} = \sqrt{\rho_d}\mathbf{H}^T\boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

where $\rho_d$ denotes the transmitted power, $\boldsymbol{x} \in \mathcal{C}^{N\times 1}$ is the transmitted signal after precoding, and $\boldsymbol{n} \in \mathcal{C}^{K\times 1}$ is a zero-mean noise vector with complex Gaussian distribution and identity covariance matrix [2]. As mentioned before, based on the uplink pilot sequences, the BS usually has imperfect CSI to all users in the downlink.

Therefore, performing power allocation at the BS is possible to maximize the sum transmission rate. With power allocation, the sum capacity for Massive MIMO system is [2]

$$\begin{aligned} C &= \max_{\mathbf{P}} \log_2(\mathbf{I} + \rho_d \mathbf{H}^H \mathbf{P} \mathbf{H}) \\ &= \max_{\mathbf{P}} \log_2(\mathbf{I} + \rho_d N \mathbf{P} \mathbf{D}), \end{aligned} \tag{2}$$

where $\mathbf{P} \in R^{K \times K}$ denotes the diagonal power matrix with $\sum_{i=1}^{K} p_{i,i} = 1$. The power allocation algorithms such as the *water-pouring* algorithm can maximize the channel capacity.

## III. Precoding methods

In Section II, equation (2) shows the optimal channel capacity of Massive MIMO system. However, in practical scenarios, the high throughput and spatial multiplexing gains in the downlink highly rely on the precoding methods at the BS [1]. Without precoding, a signal is uniformly radiated into space, showed in Figure 1 (a). By applying suitable precoding methods, BS can focus the signal to their perspective users by beamforming, significantly reducing the inter-user interference, showed in Figure 2 (b) [8].

For point-to-point MIMO system, non-linear precoding methods such as *dirty-paper-coding* (DPC) outperforms linear precoding methods [2]. However, with the number of antennas deployed at BS approaching infinity, linear precoding methods achieves near-optimal performance [2].

In TDD operation, we assume that the BS has reliable CSI of the downlink. Since Massive MIMO system has hundreds of antennas, this section introduces three fundamental linear precoding methods, matched filter (MF), zero-forcing (ZF) precoding, and regularized zero-forcing (RZF) precoding. For the three precoding methods, if the BS acquires the reliable CSI, the transmitted signal in the downlink to users can be written as

$$\boldsymbol{x} = \mathbf{W} \boldsymbol{s},$$

where $\mathbf{W} \in \mathcal{C}^{N \times K}$ is the linear precoding code, $\boldsymbol{s} \in \mathcal{C}^{K \times 1}$ the transmitted information before precoding [1]. Furthermore, the signal power is normalized, i.e.$|s| = 1$. To satisfy the power constraint at the BS, the precoding matrix satisfies $\mathrm{trace}(\mathbf{W}^H \mathbf{W}) = 1$. Note that, the precoding matrix $\mathbf{W}$ is a function of channel matrix $\mathbf{H}$.

### *A. Matched Filter*

The MF precoder, also called maximum ratio transmission (MRT), tries to maximize signal gain or signal-to-interference-noise ratio (SINR) at the perspective users. Since it is similar to the maximal-ratio combining (MRC) decoding for uplink, the MF precoding can be written as

$$\mathbf{W}_{MF} = \sqrt{\alpha} \mathbf{H}^*,$$

where $\alpha$ is a scaling factor to normalize the power of the precoding matrix [1]. Based on the precoder, the received vector at $K$ users can be expressed as

$$\boldsymbol{y} = \sqrt{\rho_d \alpha} \mathbf{H}^T \mathbf{H}^* \boldsymbol{s} + \boldsymbol{n}.$$

If the number of antennas is comparatively small, the MF precoding cannot migtigate the inference of different users [1]. With the number of antennas at the BS going to infinity, i.e., $N \to \infty$, the term $\mathbf{H}^H \mathbf{H} \to N \mathbf{I}$. Then the MF precoding achieves near-optimal performance.

### *B. Zero-forcing*

By nulling signal in other directions, the ZF precoding transmits the signal toward the intended users and eliminate the inter-user interference. Based on [1], the ZF precoder is obtained by

$$\mathbf{W}_{ZF} = \sqrt{\alpha} \mathbf{H}^* (\mathbf{H}^T \mathbf{H}^*)^{-1}.$$

Based on the precoding matrix, then the received vector $\boldsymbol{y} \in \mathcal{C}^{K \times 1}$ is

$$\begin{aligned} \boldsymbol{y} &= \sqrt{\rho_d \alpha} \mathbf{H}^T \mathbf{H}^* (\mathbf{H}^T \mathbf{H}^*)^{-1} \boldsymbol{s} + \boldsymbol{n} \\ &= \sqrt{\rho_d \alpha} \boldsymbol{s} + \boldsymbol{n}, \end{aligned} \tag{3}$$

Where the matrix $\mathbf{H}^T \mathbf{H}^*$ denotes the mutual-correlation of different channels. Based on (3), with a cost of throughput of the channel capacity, the ZF precoding decorrelates the channels [1]. In particular, if the channel has no noise, the ZF is the optimal precoding method. However, if the channel contains noise, the ZF precoding could amplify the noise at the intended users.

### *C. Regularized Zero-Forcing*

Regularized zero forcing (RZF) precoder, also called MMSE or eigenvalue-based beamforming, is considered as the state-of-the-art linear precoder for Massive MIMO [1]. In [2], the RZF precoding matrix can be written as

$$\mathbf{W}_{RZF} = \sqrt{\alpha} \mathbf{H}^* (\mathbf{H}^T \mathbf{H}^* + \delta \mathbf{I})^{-1}.$$

where $\delta > 0$ is the regularized factor. Based on $\mathbf{W}_{RZF}$, the received vector $\boldsymbol{y}$ is

$$\boldsymbol{y} = \sqrt{\rho_d \alpha} \mathbf{H}^T \mathbf{H}^* (\mathbf{H}^T \mathbf{H}^* + \delta \mathbf{I})^{-1} \boldsymbol{s} + \boldsymbol{n}. \tag{4}$$

Based on (4), the RZF precoding approaches MF precoding if $\delta \to \infty$, while the RZF precoding approaches ZF precoding if $\delta \to 0$ [1]. To achieve the trade-off of the advantages of MF and ZF precoding, the computational complexity of RZF is higher.

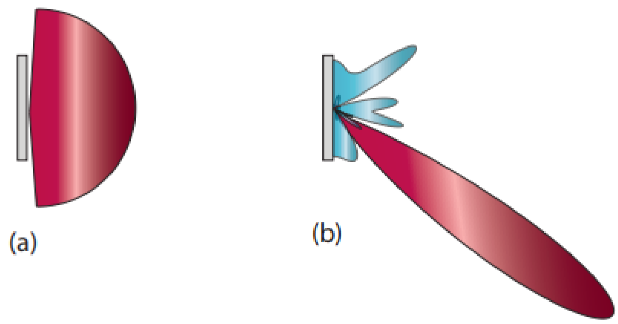


Figure 2. Beamforming achieved by digital signal processing [8]

## IV. Numerical results

Based on the Matlab simulation, we analyze the bit-error rate (BER) of three fundamental precoding methods in Section III. In the Massive MIMO (MU-MIMO) system, a single cell with a BS serves $K = 8$ users simultaneously. The BS sends $2 \times 10^6$ bits of data to each user, and the BPSK modulation and demodulation are applied at the transceivers. For $8$ users, the large-scale factors $d_k, k \in \{1, 2, \ldots, K\}$, are assumed to be larger than $0.3$. For RZF precoding, the regularized factor is $40$.

Figure 3 shows the BER of three precoding methods with respect to SNR. The BS is deployed with $M = 100$ antennas. For this figure, the $X$ axis denotes the SNR while $Y$ axis denotes the BER of Massive MIMO system. With the increasing of SNR (transmitting power) at the BS in the range $[-10, 20]$ dB, the BER of the three precoding methods decreases rapidly. Furthermore, the MF precoding achieves the worst performance while the ZF precoding achieves the best performance. By setting $\delta = 40$, the RZF precoding achieves trade-off between two methods in BER.

Figure 4 shows the BER of three precoding methods with respect to the number of antennas at the BS. The SNR at the transmitter is $10$ dB. For this figure, the $X$ axis denotes the number of antennas at the BS while $Y$ axis denotes the BER of Massive MIMO system. With the increasing of antennas at the BS from $20$ to $160$, the BER of the three precoding methods decreases rapidly. Furthermore, the MF precoding achieves worst performance and the ZF precoding achieves the best performance. By setting $\delta = 40$, the RZF precoding also achieves a trade-off between the other two methods.

## V. Limitations and future research

Based on Section II and Section III, we have introduced the downlink channels and some precoding methods to achieve the good performance in application. However, there are still a lot limiting factors to be solved, including channel estimate accuracy and pilot contamination. Furthermore, the advantages of Massive MIMO can lead to more applications in the future.

### A. Limiting factors

The first limiting factor is the channel reciprocity. The CSI in downlink is estimated based on CSI uplink in TDD operation [7]. The accuracy of the estimation relies on two properties of the channel: time-invariant and channel reciprocity [4]. However, since the channels are not perfectly time-invariant, and the hardware chains are not reciprocal, the estimated CSI is imperfect. Fortunately, if the antenna array is properly calibrated, Massive MIMO with imperfect CSI can still achieve good channel gains [4].

The second limiting factor is the *pilot contamination*. In TDD operation, the channel reciprocity and training symbols (pilots) in Uplink contribute the CSI in downlink. In practical application, the orthogonal or non-orthogonal pilot sequences are reused in multiple cells, then the estimated CSI of a specific terminal is contaminated by users with the same training sequence in the adjacent cells [4], [7]. To solve this problem, some methods have been proposed to mitigate the pilot contamination, including optimal pilot waveform allocation, channel estimation algorithms, and network-based precoding [4], [7], [8].

The third limiting factor is the mobility of users. As mentioned earlier, the good performance of Massive MIMO is tightly related to the accuracy of the CSI estimate. If users have a comparatively high mobility, the estimated CSI by the uplink pilots cannot match the CSI in the downlink. Hence, Massive MIMO is more suitable for communication cells with comparatively low mobility [6].

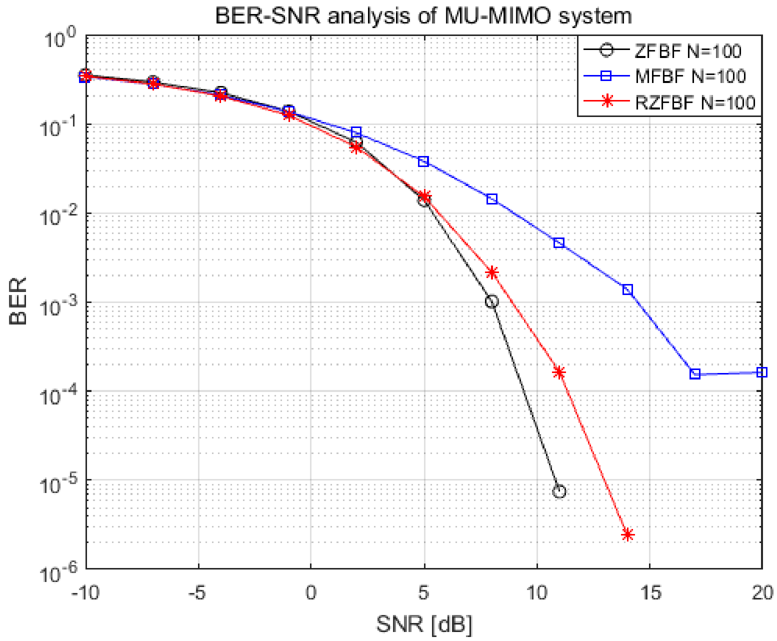


Figure 3. The BER-SNR analysis of MU-MIMO system

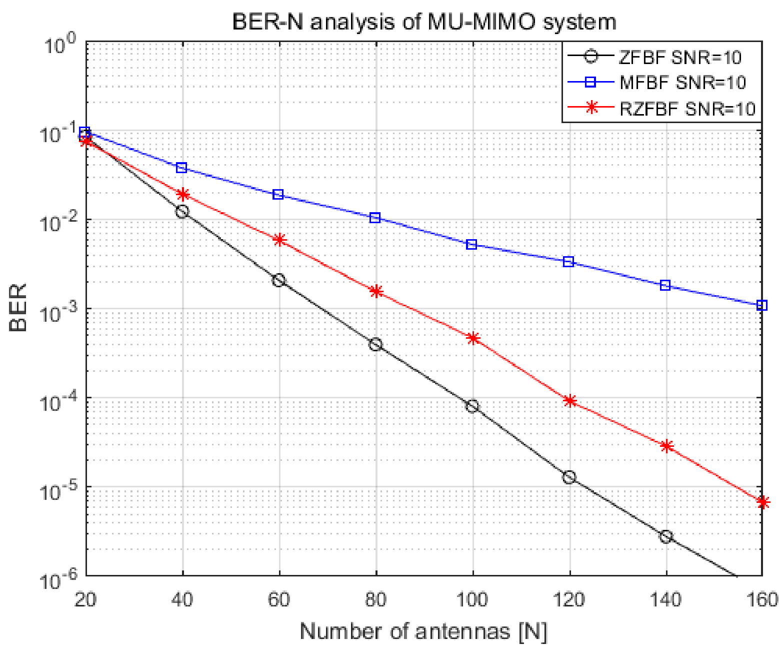


Figure 4. The BER v.s. number of antennas analysis of MU-MIMO system

### *B. Future research*

First, deploy extremely large virtual Massive MIMO antenna array. The spectral efficiency of Massive MIMO grows monotonically with the number of antennas. expect a future where hundreds or thousands of antennas are used to serve a set of users. There are, however, practical limits to how many antennas can be deployed at conventional towers and rooftop locations. For example, some compact 64-antenna array has been deployed into 2-dimensional products, but it is still far from the academic literature [8]. To further improve the performance, hundreds or thousands of antennas can be distributed over a large area and are integrated as a virtual MIMO, also called *Cell-free Massive MIMO* [9].

Second, high accurate positioning of users. For the cell wireless communication, obtaining users' spatial locations is a significant applications. In general, a user's location can be obtained by two-stage processes [8], [10]: collect measurements and compute a position estimate. For the two processes, the position accuracy is mainly determined by the underlying measurements. By deploying large antenna arrays at the base station, Massive MIMO can obtain more physical measurements, including angle-of-arrivals (AOA) in uplink and angle-of-departure (AoD) in downlink [8]. Thus, Massive MIMO provides prmosimg future for positionings.

Third, applying *machine learning* to achieve intelligent Massive MIMO [8]. Because of the good performance with many data, machine learning has been applied in many areas. By applying hundreds or thousands of antennas at the base station, machine learning can make full use of the data to improve the performance of current algorithms, channel estimation, and power control.

## VI. Conclusion

This paper gives a survey about Massive MIMO system. By deploying hundreds of antennas at the BS, Massive MIMO system achieves high energy efficiency and spectral efficiency. If the BS acquires reliable CSI of downlink by uplinking training sequences, the channel capacity can be obtained by suitable power allocation algorithms such as water-pouring. In reality, if the number of antennas at the base station is large enough, three fundamental linear precoding methods, i.e., MF, ZF, and RZF, could achieve near-optimal performance. Furthermore, for single-cell Massive MIMO system, we do simulations about the BER of three precoding methods with respect to SNR and the number of antennas at BS. The numerical results show that three precoding methods achieve great BER with suitable SNR and the number of antennas. Even though a lot of progresses have been made about Massive MIMO, there are still some limiting factors, including channel reciprocity, pilot contamination, and the mobility. Finally, the advantages of Massive MIMO also lead to more research in applications in the future, such as extremely large virtual Massive MIMO antenna array, accurate positioning of users, and intelligent Massive MIMO.

## References


[1] N. Fatema, G. Hua, Y. Xiang, D. Peng, and I. Natgunanathan, "Massive mimo linear precoding: A survey," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3920–3931, 2017.

[2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive mimo: Benefits and challenges," *IEEE journal of selected topics in signal processing*, vol. 8, no. 5, pp. 742–758, 2014.

[3] K. Zheng, S. Ou, and X. Yin, "Massive mimo channel models: A survey," *International Journal of Antennas and Propagation*, vol. 2014, 2014.

[4] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE communications magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[5] J. Mundy and K. Thomas. What is massive mimo technology? [Online]. Available: https://5g.co.uk/guides/what-is-massive-mimo-technology/

[6] E. G. Larsson and L. Van der Perre, "Massive mimo for 5g," 2017.

[7] O. Elijah, C. Y. Leow, T. A. Rahman, S. Nunoo, and S. Z. Iliya, "A comprehensive survey of pilot contamination in massive mimo—5g system," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 905–923, 2015.

[8] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive mimo is a reality—what is next?: Five promising research directions for antenna arrays," *Digital Signal Processing*, vol. 94, pp. 3–20, 2019.

[9] E. Björnson and L. Sanguinetti, "Making cell-free massive mimo competitive with mmse processing and centralized implementation," *IEEE Transactions on Wireless Communications*, 2019.

[10] R. Zekavat and R. M. Buehrer, *Handbook of position location: Theory, practice and advances*. John Wiley & Sons, 2011, vol. 27.