This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2016.2639509, IEEE Wireless Communications Letters

Quasi-Orthogonal Combining for Reducing RF Chains in Massive MIMO Systems

Abhay Kumar Sah and A. K. Chaturvedi, *Senior Member, IEEE*

***Abstract*—We address the issue of reducing the number of radio frequency (RF) chains in large/massive MIMO systems while retaining their detection performance. For this, we use a hybrid combining approach and formulate a signal to interference plus noise ratio (SINR) maximization problem. We argue that using an orthogonal basis of the channel matrix to maximize the numerator of the SINR is a worthwhile approach to pursue and use it to propose a quasi-orthogonal combining (QOC) scheme. Simulation results show that the proposed scheme can provide a detection performance and sum rate close to that of a system utilizing a dedicated RF chain for each receive antenna with less number of RF chains and without increasing the overall computational complexity.**

***Index Terms*—Large MIMO, Massive MIMO, Hybrid Combining, RF chains.**

I. Introduction

Large/Massive multiple input multiple output (MIMO) systems are considered to be a key technology for future 5G wireless systems [1], [2]. Driven by the need for higher data rates, the number of transmit/receive antennas in future MIMO wireless systems may increase up to hundreds. This will require a large number of radio frequency (RF) chains, thereby leading to a higher implementation cost, more power requirement, and lower energy efficiency.

In the literature, two different approaches exist for reducing the number of RF chains. One uses antenna selection [3] and the other uses hybrid (analog and digital) combining/precoding [4]–[9]. Between these two approaches, hybrid combining/precoding is a recent development and has been shown to be a preferred choice due to its sum rate maximization capability [9]. However, these works do not investigate the effect of combining/precoding on the detection performance.

In this work, we focus on determining the combining required to reduce the number of RF chains without compromising the detection performance and sum rate. Thus, we consider the most recent combining architecture available for the uplink scenario [9], and formulate a signal to interference plus noise ratio (SINR) maximization problem. The optimal solution to this problem requires an exhaustive search which increases exponentially with the number of receive antennas and RF chains.

To handle this maximization problem with reasonable complexity, we argue that if an orthogonal basis is used for the channel matrix, a maximization based only on the numerator of the SINR will provide a good solution. The resulting problem is similar to a lattice search [10], which can be solved by a tree search algorithm in polynomial complexity. However, the search differs from the conventional lattice search in the sense that it searches for the farthest lattice instead of the closest one. Combining the above ideas, we propose a quasi-orthogonal combining (QOC) scheme and an improved quasi-coherent (IQC) combining scheme. It may be noted that IQC is a variant of the existing quasi-coherent (QC) combining [9]. We show that even with very few RF chains the detection performance and the sum rate of QOC are close to that of a system using a dedicated RF chain for each receive antenna i.e. a system which does not need any combining. This is achieved without any increase in the overall computational complexity.

The authors are at the Department of Electrical Engineering, Indian Institute of Technology Kanpur, UP, 208016 INDIA (e-mail: abhaysah@iitk.ac.in; akc@iitk.ac.in).

This work was supported by Tata Consultancy Services Limited.

II. Problem Formulation

A signal received at a MIMO receiver with N_r receive antennas can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{y} \in \mathcal{C}^{N_\text{r}\times 1}$ is the received signal vector and $\mathbf{x} \in \mathcal{C}^{N_\text{t}\times 1}$ is the transmitted signal vector, where each element of X is taken from a set of M complex symbols. The channel matrix $\mathbf{H} \in \mathcal{C}^{N_\text{r}\times N_\text{t}}$ with each coefficient $h_{i,j} \sim \mathcal{CN}(0,1)$ and $\mathbf{n} \in \mathcal{C}^{N_\text{r}\times 1}$ represents an i.i.d. additive white Gaussian noise (AWGN) vector with each $n_i \sim \mathcal{CN}(0,\sigma^2)$. It may be worthwhile to mention here that this model captures the uplink scenario for both single-user large MIMO systems as well as multi-user massive MIMO systems. The only difference is that in the first case, the transmit vector $\mathbf{x}$ constitutes data of a single user while in the second case, it constitutes data of multiple users.

The combining architecture is shown in Fig. 1 [9], where the signals received by the low noise amplifiers (LNA) are combined using switches, adders and analog phase shifters (APS), which are connected to N_c ($\geq N_\text{t}$) RF chains. Mathematically, the receive vector after combining can be expressed as

$$\mathbf{r} = \mathbf{C}^H\mathbf{y} = \mathbf{C}^H\mathbf{H}\mathbf{x} + \mathbf{C}^H\mathbf{n}, \tag{2}$$

where $\mathbf{C} \in \mathcal{C}^{N_\text{r}\times N_\text{c}}$ is a combining matrix and $\mathbf{r} \in \mathcal{C}^{N_\text{c}\times 1}$ is the received vector which will be processed for detection. The individual columns of the matrix $\mathbf{C}$ are expressed as $\mathbf{c}_k = \mathbf{S}_k\mathbf{p}\,\forall k = 1,2,\cdots,N_\text{c}$. Here $\mathbf{p} = \left[1\ \exp\left(\frac{-j2\pi}{N_\text{p}}\right)\ \cdots\ \exp\left(\frac{-j2\pi(N_\text{p}-1)}{N_\text{p}}\right)\right]^T$ denotes an array of N_p phase shifters and $\mathbf{S}_k$'s are the $N_\text{r}\times N_\text{p}$ switching matrices. These switching matrices contain the information about the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2016.2639509, IEEE Wireless Communications Letters

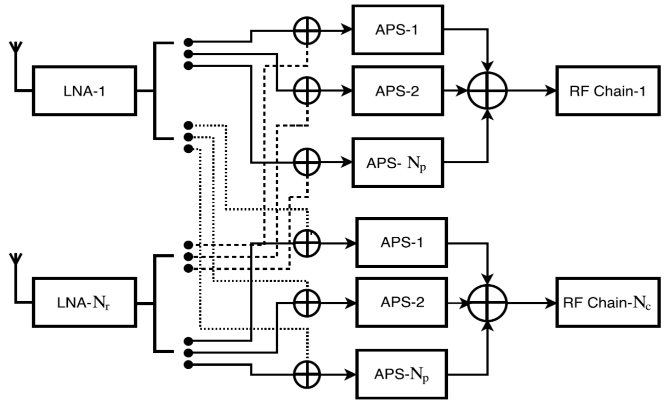

Fig. 1. Massive MIMO combining architecture for reducing RF chains [9].

states of different switches, 1 represents an *on* state while 0 represents an *off* state and each switch can be connected to at most one of the N_p APS. This means the elements of $\mathbf{c}_k$ can take only the phase values which are shown as elements of $\mathbf{p}$.

Assuming that the number of RF chains is lesser than the number of receive antennas i.e. $N_\text{c} < N_\text{r}$, the goal is to find the switch positions and the phase shifts i.e. $\mathbf{C}$ such that the SINR is maximized at the receiver i.e.

$$\mathbf{C} = \underset{\mathbf{C}\in\Psi^{N_\text{r}N_\text{c}}}{\operatorname{argmax}} \sum_{k=1}^{N_\text{c}} \frac{\left|\mathbf{c}_k^H \mathbf{h}_k\right|^2}{\sigma^2\|\mathbf{c}_k\|^2 + \sum_{\substack{i=1\\ i\neq k}}^{N_\text{t}} \left|\mathbf{c}_k^H \mathbf{h}_i\right|^2} \quad (3)$$

where $\Psi = \left\{\exp\left(\frac{-j2\pi(i-1)}{N_\text{p}}\right) \middle| \forall\, i = 1, 2, \cdots, N_\text{p}\right\}$.

III. Quasi-Orthogonal Combining

Finding the optimal solution of the problem in (3) requires an exhaustive search over $N_\text{p}^{N_\text{r}N_\text{c}}$ possible solutions, which is computationally prohibitive for large systems. The expression in (3) involves a dot product of vector $\mathbf{c}_k$ with $\mathbf{h}_k$ in the numerator and with the remaining columns of $\mathbf{H}$ in the second term of the denominator. If we are able to choose $\mathbf{c}_k$ such that the dot product between $\mathbf{c}_k$ and $\mathbf{h}_k$ is high and at the same time the dot product with the remaining columns of $\mathbf{H}$ is low, it will simultaneously increase the numerator and decrease the denominator, thereby helping the maximization of (3). Since the magnitude of the dot product is maximum for collinear vectors and minimum for orthogonal vectors, we propose to use $\widetilde{\mathbf{H}}$ - a matrix the columns of which constitute an orthogonal basis of $\mathbf{H}$, to seek the $\mathbf{C}$ which will maximize $|\mathbf{c}_k^H\widetilde{\mathbf{h}}_k|\,\forall\, k = 1,\, 2,\, \cdots,\, N_\text{c}$. This is because this $\mathbf{C}$ will also reduce the second term in the denominator. However, the complexity of finding $\mathbf{C}$ using an exhaustive search is still of the order of $N_\text{p}^{N_\text{r}}$. As an example, let us consider a receiver with 32 antennas, 8 RF chains, and 4 APS. For such a system, the complexity of exhaustive search is reduced from $4^{32\times 8} = 2^{512}$ to $4^{32} = 2^{64}$ possible solution vectors, which is still a very high number. We address this issue next.

Given the similarity of maximizing $|\mathbf{c}_k^H\widetilde{\mathbf{h}}_k|\,\forall\, k = 1,\, 2, \cdots,\, N_\text{c}$ to the lattice search problems [10], tree search algorithms are natural candidates for consideration. Tree search algorithms are categorized into depth first tree search (DFTS) and breadth first tree search (BFTS) algorithms. Here we will consider a BFTS approach due to its ability to achieve a near optimal solution with low complexity. Motivated by this, we first break the maximization of $|\mathbf{c}_k^H\widetilde{\mathbf{h}}_k|$ into the following N_r intermediate maximization problems and solve them sequentially. The ith intermediate problem can be expressed as

$$\{c_{1,k}, \cdots, c_{i,k} | c_{1,k}, \cdots, c_{i-1,k}\} = \underset{c_{j,k}\in\Psi}{\operatorname{argmax}} \left| \sum_{j=1}^{i} c_{j,k}^* \cdot \widetilde{h}_{j,k} \right|, \quad (4)$$

where $c_{j,k}$ and $\widetilde{h}_{j,k}$ are the jth element of $\mathbf{c}_k$ and $\widetilde{\mathbf{h}}_k$ respectively. The algorithm begins with $i = 1$ by determining the value of ψ which is closest to $\angle\widetilde{h}_{1,k}$. Next we obtain the product $\psi^*\widetilde{h}_{1,k}$ for this value of ψ. Then we compute the products of $\widetilde{h}_{2,k}$ with all the N_p possible values of ψ's. Based on these products, we compute the sum of products (i.e. the function in (4)) for all possible N_p combinations and select the combination which maximizes the sum. The procedure is repeated for $i = 2$ to N_r. At the end i.e. for $i = N_\text{r}$, the candidate which maximizes $|\mathbf{c}_k^H\widetilde{\mathbf{h}}_k|$ is declared as $\mathbf{c}_k$. The same procedure is repeated for obtaining the other columns of $\mathbf{C}$. The complete algorithm is provided in Algorithm-1. It may be noted that although $\mathbf{C}$ is obtained using an orthogonal basis, it need not be orthogonal. Therefore, we refer to this proposal as a quasi-orthogonal combining (QOC) scheme.

Algorithm 1: Quasi-Orthogonal Combining

Input : $\widetilde{\mathbf{H}}$, Ψ
Output: $\mathbf{C}$
for $k \leftarrow 1$ **to** N_c **do**
  $c_{1,k} \leftarrow$ select ψ for which $|\angle\widetilde{h}_{1,k} - \angle\psi|$ is minimum;
  $d = c_{1,k}^*\widetilde{h}_{1,k}$;
  for $i \leftarrow 2$ **to** N_r **do**
    for $r \leftarrow 1$ **to** N_p **do**
      $\widetilde{d}_r = d + \psi_r^* \cdot \widetilde{h}_{i,k}$;
    end
    $\lambda \leftarrow$ the index r for which $|\widetilde{d}_r|$ is maximum;
    $c_{i,k} \leftarrow \psi_\lambda$;
    Replace $d \leftarrow \widetilde{d}_\lambda$;
  end
end
return C

Interestingly, the proposed approach i.e. to use $\widetilde{\mathbf{H}}$ instead of $\mathbf{H}$, can also be used to improve the performance of the quasi-coherent (QC) scheme [9]. We refer to this variant of QC as an improved QC (IQC) scheme.

IV. Simulation Results

To investigate the potential utility of QOC and IQC, we consider a 16-QAM system with 16 and 32 receive antennas for $N_\text{c} = N_\text{t} = N_\text{p} = 8$. We determine $\mathbf{C}$ using the proposed QOC and IQC schemes, and compare their respective detection performances with the QC scheme [9] in Fig. 2. Here, for the purpose of obtaining the detection performance, we consider the MMSE detector for all the three algorithms. We notice

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2016.2639509, IEEE Wireless Communications Letters

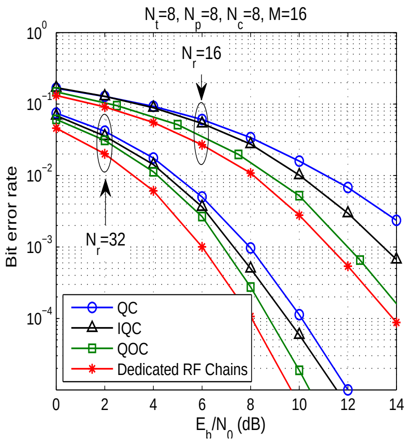

Fig. 2. Combining schemes for $N_\mathrm{r} = 16$ and 32.

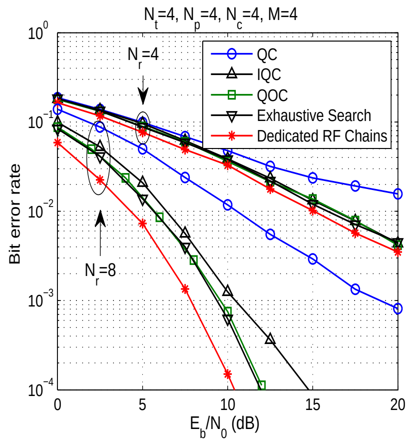

Fig. 3. Combining schemes for $N_\mathrm{r} = 4$ and 8.

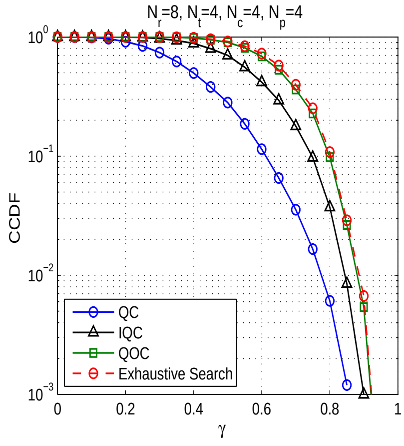

Fig. 4. CCDF of $\mathrm{Orth}(\mathbf{C})$.

that using only 8 RF chains, QOC with $N_\mathrm{r} = 32$ provides a performance which is marginally inferior to the system with 32 dedicated RF chains i.e. one RF chain for each receive antenna. It can be seen that both QOC and IQC schemes perform better than QC combining. For e.g., for $N_\mathrm{r} = 16$ and a bit error rate (BER) of 10^{-3}, QOC and IQC have a more than 3 dB and 2 dB gain over QC, respectively. Further, for the same N_r and BER, we notice that QOC with 8 RF chains is inferior to the 16 dedicated RF chains system by only 0.7 dB.

As there is a marginal loss in the error performance of QOC compared to a dedicated RF chains system, it would be interesting to examine the performance of QOC with an exhaustive search based combining. Since exhaustive search is not possible on such large systems, we consider a relatively smaller system i.e. $N_\mathrm{c} = N_\mathrm{t} = N_\mathrm{p} = 4$ and $N_\mathrm{r} = 8$ for 4-QAM. The results have been shown in Fig. 3. From the figure, we can observe that QOC overlaps with the exhaustive search based combining. For the sake of completeness, in Fig. 3 we have also compared QOC, IQC and QC, and as expected QOC and IQC perform better than QC. We further validate this for the same system with 4 receive antennas. Thus, exhaustive search per se does not provide any gain in performance compared to QOC. In any case, it is infeasible for larger systems.

This will be further corroborated using a measure of orthogonality, defined by subtracting the orthogonal deficiency [11] from unity, as follows

$$\mathrm{Orth}(\mathbf{C}) = \frac{\det(\mathbf{C}^H\mathbf{C})}{\prod_{i=1}^{N_\mathrm{c}} \|\mathbf{c}_i\|^2}. \tag{5}$$

$\mathrm{Orth}(\mathbf{C})$ is a positive number between 0 and 1, higher the value of $\mathrm{Orth}(\mathbf{C})$ better is the orthogonality of $\mathbf{C}$. We computed $\mathrm{Orth}(\mathbf{C})$ for several realizations of $\mathbf{H}$ and determined the probability of $\mathrm{Orth}(\mathbf{C})$ being greater than a threshold γ i.e.

$$F(\gamma) = Prob\left(\mathrm{Orth}(\mathbf{C}) > \gamma\right). \tag{6}$$

We use this to plot the complementary cumulative distribution function (CCDF) of $\mathrm{Orth}(\mathbf{C})$ for $\gamma = 0$ to 1 in Fig. 4. From the figure one can observe that even exhaustive search is unable to achieve a value of unity for $\mathrm{Orth}(\mathbf{C})$. This is because

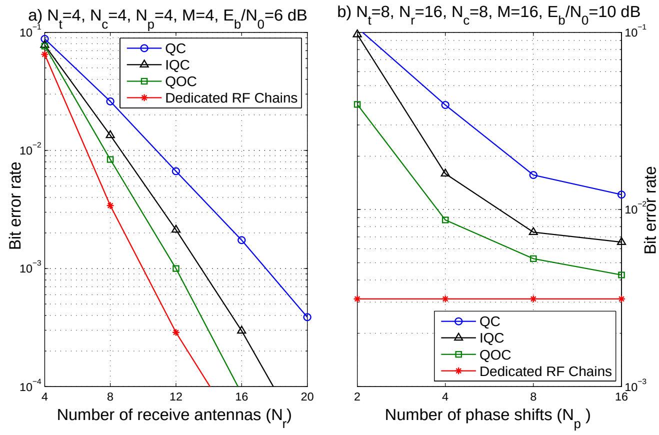

Fig. 5. BER performance of the combining schemes using MMSE detection with increasing number of receive antennas/APSs.

only finite number of phase shifts are available. This also explains the reason for the earlier referred gap between the curves corresponding to the dedicated RF chains system and exhaustive search based combining in Figs. 2 and 3. From the CCDF plot we can observe that $\mathrm{Orth}(\mathbf{C})$ is maximum for QOC, minimum for QC and IQC is in between. Interestingly, the CCDF plot for QOC almost overlaps with exhaustive search. Both these observations are in line with the simulation results reported earlier in this section.

Further, in Fig. 5 we examine the performance of QOC and IQC with increasing number of receive antennas (N_r) and APSs (N_p). From Fig. 5 (a) one can observe that it is possible to attain the BER of a system utilizing dedicated RF chains by a system using reduced number of RF chains by slightly increasing the number of receive antennas. For e.g., performance of the system consisting of 12 dedicated RF chains and 12 receive antennas is achieved by QOC using only 4 RF chains and 14 receive antennas, by IOC using 4 RF chains but 16 receive antennas, and by QC using 4 RF chains but 20 receive antennas. From Fig. 5 (b) one can view that by increasing N_p, the gap between the dedicated RF chains system and the reduced RF chains system can be brought down for all the three schemes.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LWC.2016.2639509, IEEE Wireless Communications Letters

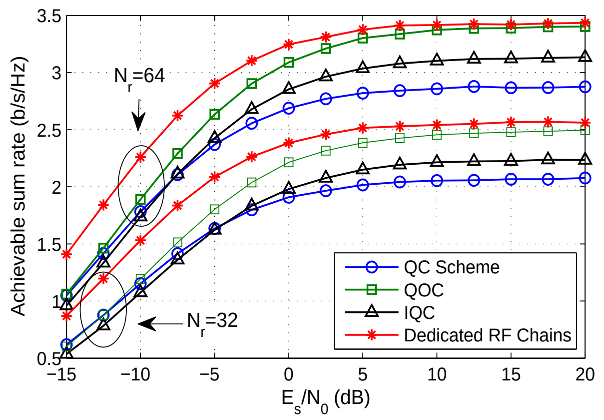

Fig. 6. Sum rate per transmit antenna for $N_c = N_t = 8$ and $N_p = 4$.

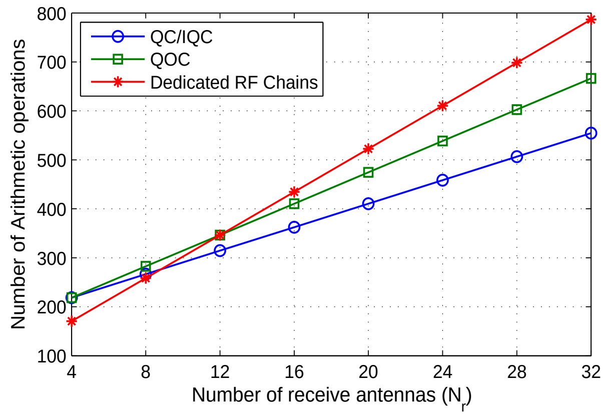

Fig. 7. Number of arithmetic operations per bit required with MMSE detection for $N_c = N_t = 4$, $N_p = 4$ and $M = 4$ at $E_b/N_0 = 6$ dB.

In Fig. 6 we compare the achievable sum rate [9] per transmit antenna of QOC and IQC with QC. We consider $N_r = 32$ and $N_r = 64$ for $N_c = N_t = 8$ and $N_p = 4$. From the figure, it can be observed that both QOC and IQC outperform QC. In fact, for high E_b/N_0, QOC is very close to the dedicated RF chains systems. Sum rate for the dedicated RF chains systems has been evaluated assuming conjugate beamforming as in [9] i.e. by taking $\mathbf{C} = \mathbf{H}$ in (2).

Lastly, we investigate the complexity of QOC and IQC based detection. These schemes have three steps: i) Computation of $\widetilde{\mathbf{H}}$, ii) Computation of $\mathbf{C}$ and iii) Detection. The first step can be obtained in $O(N_r N_t^2)$ computations using a QR decomposition while the second step can be obtained in $O(N_r N_p N_c)$ computations for both QOC and IQC. Considering the scenario used for simulations i.e. $N_p \leq N_t$, $N_c = N_t$ and a low complexity detector like MMSE, the complexity is of the order of $O(N_r N_t^2)$, which is same as the complexity of MMSE based detection in a dedicated RF chains system.

It is worthwhile to mention here that QR decomposition is required in many detection algorithms such as sphere decoder (SD) [12] and K-best [13]. Even for an MMSE detector, QR decomposition is helpful in reducing the complexity. Thus, step (i) is common to QC, IQC and QOC. Therefore, we compare the algorithms only on the basis of steps (ii) and (iii). The number of arithmetic operations per bit is shown in Fig. 7 as a function of N_r for the same simulation parameters as used in Fig. 5 (a). It may be noted that the complexity of IQC is same as that of QC. From the figure it can be observed that the complexities of QC/IQC and QOC are lower than the dedicated RF chains system after 8 and 12 receive antennas, respectively. This is because a reduction in the number of RF chains is accompanied by savings in detection complexity. It may be noted that the complexity gains in large/massive MIMO systems will actually be higher because they are likely to use detectors with complexities higher than that of MMSE.

V. Conclusion

We consider a combining architecture for an uplink massive MIMO scenario and formulate an SINR maximization problem. We propose two combining schemes to solve this problem, namely QOC and IQC. Using simulations, we show that with less number of RF chains and without increasing the overall complexity, the proposed schemes can provide a detection performance close to that of a system utilizing a dedicated RF chain for each receive antenna. Further, we show that the schemes are also able to achieve a higher sum rate.

References

[1] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[2] A. Osseiran *et al.*, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.

[3] S. Sanayei and A. Nosratinia, "Antenna selection in mimo systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[4] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Multi-switch for antenna selection in massive MIMO," in *Proc. 2015 IEEE Global Commun. Conf. (GLOBECOM)*, pp. 1–6.

[5] ——, "Massive MIMO in real propagation environments: Do all antennas contribute equally?" *IEEE Tran. Commun.*, vol. 63, no. 11, pp. 3917–3928, Nov. 2015.

[6] Z. Li, S. Han, and A. F. Molisch, "Hybrid beamforming design for millimeter-wave multi-user massive mimo downlink," in *Proc. 2016 IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[7] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Tran. Commun.*, vol. 64, no. 1, pp. 201–211, Jan. 2016.

[8] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Tran. Wireless Commun.*, vol. 15, no. 5, pp. 3311–3326, May 2016.

[9] A. Alkhateeb, Y. H. Nam, J. Zhang, and R. W. Heath, "Massive MIMO combining with switches," *IEEE Wireless Commun. Lett.*, vol. 5, no. 3, pp. 232–235, Jun. 2016.

[10] M. O. Damen, H. E. Gamal, and G. Caire, "On maximum-likelihood detection and the search for the closest lattice point," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2389–2402, Oct. 2003.

[11] Q. Zhou and X. Ma, "Element-based lattice reduction algorithms for large MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 274–286, Feb. 2013.

[12] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. Expected complexity," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[13] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-best sphere decoding for MIMO detection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 491–503, Mar. 2006.