

Article

Simplified Antenna Group Determination of RS Overhead Reduced Massive MIMO for Wireless Sensor Networks

Byung Moo Lee 

School of Intelligent Mechatronics Engineering, Sejong University, Seoul 05006, Korea; blee@sejong.ac.kr;
Tel.: +82-2-6935-2520

Received: 1 December 2017; Accepted: 27 December 2017; Published: 29 December 2017

Abstract: Massive multiple-input multiple-output (MIMO) systems can be applied to support numerous internet of things (IoT) devices using its excessive amount of transmitter (TX) antennas. However, one of the big obstacles for the realization of the massive MIMO system is the overhead of reference signal (RS), because the number of RS is proportional to the number of TX antennas and/or related user equipments (UEs). It has been already reported that antenna group-based RS overhead reduction can be very effective to the efficient operation of massive MIMO, but the method of deciding the number of antennas needed in each group is at question. In this paper, we propose a simplified determination scheme of the number of antennas needed in each group for RS overhead reduced massive MIMO to support many IoT devices. Supporting many distributed IoT devices is a framework to configure wireless sensor networks. Our contribution can be divided into two parts. First, we derive simple closed-form approximations of the achievable spectral efficiency (SE) by using zero-forcing (ZF) and matched filtering (MF) precoding for the RS overhead reduced massive MIMO systems with channel estimation error. The closed-form approximations include a channel error factor that can be adjusted according to the method of the channel estimation. Second, based on the closed-form approximation, we present an efficient algorithm determining the number of antennas needed in each group for the group-based RS overhead reduction scheme. The algorithm depends on the exact inverse functions of the derived closed-form approximations of SE. It is verified with theoretical analysis and simulation that the proposed algorithm works well, and thus can be used as an important tool for massive MIMO systems to support many distributed IoT devices.

Keywords: antenna group; massive MIMO; reference signal (RS)

1. Introduction

Massive multiple-input multiple-output (MIMO) system is a powerful technology that can increase both spectral efficiency (SE) and energy efficiency (EE), and it has been actively discussed to be included in the 3rd generation partnership project (3GPP) standard as core technology for 5G systems [1–3]. The massive MIMO system uses a large amount of transmitter (TX) antennas and serves limited number of user equipments (UEs) and/or Internet of things (IoT) devices/sensors, so it is a combination scheme of multi-user (MU) MIMO and beamforming, and its drastic performance gain has already been proven in various literature [4–8]. Recently, it has also been proven that massive MIMO is quite effective in supporting many Internet of things (IoT) and/or Industrial Internet of things (IIoT) devices, and thus can be used for a core technology to configure wireless sensor networks [9–15]. The calibration of transceiver impairment is very important problem to support distributed IoT devices using massive MIMO [9,10], because without getting accurate channel information, it is impossible to support the distributed IoT devices. In [9], authors proposed very effective channel calibration method to reduce the channel mismatch due to the radio frequency (RF) impairment. In their method,

they used power headroom which is reported periodically and/or aperiodically from distributed IoT devices to base station (BS), to deliver the RF impairment information of IoT devices. Based on the RF impairment information which is embedded in power headroom, BS performs precoding with the compensation of channel mismatch. This is very important scheme to support distributed IoT devices using massive MIMO equipped data center. A more general work for transceiver impairment calibration was also proposed in [10]. In [10], authors investigated massive MIMO-based distributed detection with general transceiver impairments at both a massive-antenna data center and multiple single-antenna sensors. They first derived closed-form expressions to show the dependence of both the probability of detection and the probability of false alarm on the transceiver impairments, then they showed that hardware impairments create finite ceilings on the detection performance. In addition, they formulated an optimization problem to maximize the probability of detection, while maintaining a constant false alarm probability and a sum sensor reporting power budget. By exploiting the inherent structures in the problem formulation, they developed an iterative algorithm to solve the problem via invoking the alternating direction method of multipliers. In addition with the results, there has been a high interest to equip the sensor fusion center with massive MIMO framework [11–15]. In [11], authors considered a decentralized multi-sensor estimation problem where sensor nodes observe noisy versions of a correlated random source vector. The sensors amplify and forward their observations over a fading coherent multiple access channel to a fusion center. The fusion center is equipped with a massive MIMO and adopts a minimum mean-square error (MMSE) approach for estimating the source. They optimized the transmission power at each sensor node to increase energy efficiency in various scenarios. In [12,13], authors studied channel-aware decision fusion over MIMO channels, in the presence of a massive MIMO at the fusion center. They tried to develop linear fusion rules, and presented a wide choice of low-complexity sub-optimal rules which efficiently exploit massive MIMO benefits and are able to achieve near-optimal performance. In [14], authors considered the uplink detection and estimation of a zero mean Gaussian signal in a wireless sensor network when the fusion center is configured with a massive MIMO. For the detection problem, they studied the Neyman-Pearson (NP) detector and energy detector (ED), and found optimal values for the sensor transmission gains. According to the results, while bounds derived for NP detection shows performance gains for a fusion center with massive MIMO in low sensor transmit power scenarios, the benefit is shown to disappear when the transmit power is high. However, for the ED, having massive MIMO at the fusion center provides a significant advantage even when the sensors have high power. In [15], authors considered distributed detection in wireless sensor networks with a massive MIMO fusion center. Using the large deviation principle and random matrix theory, they analyzed the asymptotic detection performance of optimal hypothesis testing in terms of error exponents for the false alarm and miss detection probabilities which enable us to predict how difficult it will be to attain a certain level of detection reliability.

Various studies already have shown that massive MIMO is quite effective to support distributed IoT devices. However, one of big obstacles to the realization of massive MIMO systems is reference signal (RS) overhead, which increases as the number of transmitter (TX) antennas and/or user entities (UEs) increases. The RS overhead reduction is a classical issue in research field of wireless communication, and numerous related technologies have been introduced [16–20]. In general, for massive MIMO, it is very difficult to apply orthogonal RS to each antenna due to the excessive amount of TX antennas, thus antenna group-based RS overhead reduction scheme is generally applied [20,21]. In this kind of scheme, in each antenna group, the same RS is shared. In the case, there is enough correlation among antennas, antenna group can be decided based on beam groups and locations of distributed UEs. If there is not enough correlation in each antenna, several RS superposition schemes can be applied. A blind channel estimation scheme using the specific statistical property of the signal and the channel were introduced in [22,23]. A training based super-imposed signal, which locates the RS stream and data stream in the same resources, and can also be applied to super-imposed RS streams, were introduced in [24,25]. It is already proven that the antenna group-based RS overhead

reduction scheme is quite effective. One of important factors in designing the antenna group-based RS overhead reduction scheme is how many antennas will be grouped in each group. Several techniques that can effectively be applied to antenna group-based RS overhead reduction scheme have been proposed [20,21], while little work has been done for the selection criterion of the number of antennas needed in each group, and related system design [16–19,22–25]. All of the work described above focuses solely on the methodology of how to reduce RS overhead. In real situation, if we assume we have the RS overhead reduction scheme at hand based on any kind of existing methodology, before applying the scheme, it is necessary to decide the target amount of RS overhead reduction in given situations. In addition, it is also necessary to design the systematic operational step and apparatus for real implementation. As far as we know, these have not been actively discussed in the literature yet.

In this paper, we propose a simplified determination scheme of the number of antennas needed in each group for massive MIMO to support wireless sensor networks, and present related system structure for the operation. Basically, our contribution in this paper can be divided into two parts. First, we analyze the performance of RS overhead reduced massive MIMO systems with antenna grouping and related channel estimation error. Even though antenna grouping-based RS overhead reduction scheme is quite effective for the operation of massive MIMO and related IoT device support, channel estimation error is inevitable. The closed-form approximations of achievable SE, which includes related channel estimation error factor, are derived. Based on the channel estimation error factor, both RS overhead performance and the seriousness of channel estimation error can be adjusted. The seriousness of channel estimation error which is used in this paper will be shown in Section 5. We also show that the derived closed-form approximations are quite simple and effective, and in good agreement with the simulation results. Second, based on the derived closed-form approximation of SE, and using the exact inverse of the SE closed-form approximation, we propose a simplified determination scheme of the number of antennas needed in each group. Without this scheme, it is very difficult to apply antenna grouping-based RS overhead reduction scheme in a real system. The change of parameters must be reflected in real time to effectively determine the number of antennas in each group. The presented system block diagram and flow chart can be greatly helpful to real implementation of the proposed scheme. In a word, based on closed-form approximation of SE, to support distributed IoT devices, we propose a simplified determination scheme of the number of antennas needed in each group. More specifically, our main contributions are summarized as follows:

- We present the closed-form approximations of achievable SE with channel estimation error for two representative linear precodings. We used the error factor which reflects the degree of channel estimation error and the error matrix which has the same statistical characteristic but independent of the channel matrix. The range of error factor is between 0 and 1, and it is a function of the number of antennas in each antenna group, and the adjustable factor which reflects the seriousness of the channel estimation error depending on the applied reference signal (RS) overhead reduction technique. We prove the derived SEs are quite simple and effective using extensive simulation results.
- Based on the closed-form approximations of achievable SEs, we derive the determination criteria of the number of antennas needed in each antenna group. The determination criteria are derived from the inverse functions of achievable SEs.
- We provide the system block diagram and algorithmic steps to apply the proposed determination scheme. The system block diagram is consisted of 5 main blocks, The central management unit, operation parameters unit, minimum data rate unit, decision unit, and UE. There are 3 thresholds we should consider for the systematic operation, i.e., power consumption, data rate, and the number of iterations.
- We provide the extensive simulation results of the proposed scheme with various parameters, and show that the derived approximations are very effective and well-matched with the simulation results.

The paper is organised as follows: the system description is presented in Section 2. Massive MIMO model, related precoding techniques, and RS overhead model are given in Section 2. In Section 3, the performance analysis and closed-forms approximation of achievable SE are presented. In Section 4, the determination scheme of the number of antennas in each group is shown. In Section 5, the proposed model is numerically analyzed. With monte-carlo (MC) simulations, we show that the analysis is well-matched with simulation results. Finally, concluding remarks are given in Section 6.

Notation: In the rest of the paper, boldface lower-case and upper-case characters denote vectors and matrices, respectively. The operators $(\cdot)^H$ and $\mathbb{E}[\cdot]$ denote conjugate transpose and expectation, respectively. The $N \times N$ identity matrix is denoted $\mathbf{I}_N$, and the $N \times N$ zero matrix is denoted $\mathbf{0}_N$. $\mathbf{X} \sim N(\mathbf{0}_N, \mathbf{V}_N)$ is the complex Gaussian distributed vector with mean zero and covariance $\mathbf{V}_N$. $\log_2(\cdot)$ denotes the common logarithm and $\|\cdot\|_F$ represents the Frobenius norm operator. We use $\circ$ to denote the componentwise product of the matrices. we use $\text{diag}[d_1, \dots, d_N]$ for $N \times N$ diagonal matrix with d_i as the i th diagonal element.

2. System Description

2.1. Massive MIMO Model

We consider a downlink massive MIMO system with N_t TX antennas, and K single antenna UEs and/or IoT devices/sensors. The received signal vector at UEs can be represented as follows:

$$\mathbf{y} = \sqrt{\rho_t} \mathbf{G} \mathbf{s} + \mathbf{n}, \quad (1)$$

where $\mathbf{y}$ is the $K \times 1$ received signal vector for K UEs, ρ_t is the total TX power for forward link, $\mathbf{G}$ is the $K \times N_t$ channel matrix between the transmitter with N_t TX antennas and K UEs, $\mathbf{s}$ is the $N_t \times 1$ TX signal vector, and $\mathbf{n}$ is the $K \times 1$ additive white Gaussian noise (AWGN) vector at the UEs (i.e., $\mathbf{n} \sim CN(\mathbf{0}_K, \sigma_d^2 \mathbf{I}_K)$) where σ_d^2 is the variance of AWGN. $\mathbf{G}$ consists of both a small scale fading channel matrix, $\mathbf{H}$ and large scale fading channel matrix, $\mathbf{B}$, i.e., $\mathbf{G} = \mathbf{H} \circ \mathbf{B}$ where $\circ$ indicates the componentwise product. $\mathbf{H}$ is the independent and identically distributed (i.i.d) Rayleigh fading channel matrix with zero mean and unit variance which can be represented as:

$$\mathbf{H} = \begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,N_t} \\ h_{2,1} & h_{2,2} & \cdots & h_{2,N_t} \\ \vdots & \vdots & \vdots & \vdots \\ h_{K,1} & h_{K,2} & \cdots & h_{K,N_t} \end{pmatrix}, \quad (2)$$

$\mathbf{B}$ is the large scale fading channel matrix which can be represented as:

$$\mathbf{B} = \begin{pmatrix} \sqrt{\beta_{1,1}} & \sqrt{\beta_{1,2}} & \cdots & \sqrt{\beta_{1,N_t}} \\ \sqrt{\beta_{2,1}} & \sqrt{\beta_{2,2}} & \cdots & \sqrt{\beta_{2,N_t}} \\ \vdots & \vdots & \vdots & \vdots \\ \sqrt{\beta_{K,1}} & \sqrt{\beta_{K,2}} & \cdots & \sqrt{\beta_{K,N_t}} \end{pmatrix}, \quad (3)$$

where $\beta_{i,j}$ is the path loss component from j th antennas in BS to i th UE. Then, the combined channel matrix $\mathbf{G} = \mathbf{H} \circ \mathbf{B}$ can be represented as:

$$\mathbf{G} = \begin{pmatrix} h_{1,1} \sqrt{\beta_{1,1}} & h_{1,2} \sqrt{\beta_{1,2}} & \cdots & h_{1,N_t} \sqrt{\beta_{1,N_t}} \\ h_{2,1} \sqrt{\beta_{2,1}} & h_{2,2} \sqrt{\beta_{2,2}} & \cdots & h_{2,N_t} \sqrt{\beta_{2,N_t}} \\ \vdots & \vdots & \vdots & \vdots \\ h_{K,1} \sqrt{\beta_{K,1}} & h_{K,2} \sqrt{\beta_{K,2}} & \cdots & h_{K,N_t} \sqrt{\beta_{K,N_t}} \end{pmatrix}. \quad (4)$$

Since the path loss components from 1th, 2nd, $\dots$, N_t th antenna in BS to antenna in i th UE are almost the same, for the path loss component from BS to i th UE, β_i , we can represent as follows:

$$\beta_i = \beta_{i,1} = \beta_{i,2}, \dots, \beta_{i,N_t}. \quad (5)$$

where $i = 1, 2, \dots, K$.

Note that, to get the benefit of channel hardening effect, usually the massive MIMO antenna systems should satisfy the condition of $N_t > 10K$ [26,27]. There is no strict criterion to get the channel hardening effect, but we can asymptotically observe the effect when N_t is larger than $10K$.

Figure 1 presents the simulation to show the channel hardening effect. The x-axis of the figure represents the ordered eigenvalue of $\mathbf{H}\mathbf{H}^H$, and y-axis represents the cumulative distribution function (CDF). The figure shows the case in which the number of UEs is fix to $K = 4$, and the number of TX antennas increases as $N_t = 4, 40$, and 400 . As N_t increases, the randomness of the channel significantly reduces. Likewise, due to the law of large numbers, as N_t increases, the level of interference also converges to a certain constant. Figure 2 presents the normalized interference power versus number of trials. Figure 2 is the result when we use the matched filtering (MF) precoding which will be introduced in the next subsection. As observed, as N_t increases, the variation of interference reduces. It is very close to a constant when $N_t = 10K$. From Figures 1 and 2, we can observe that asymptotically the channel hardening effect is reliable enough when $N_t > 10K$, but as we mentioned, this is just asymptotic approach and there is not strict rule for the criterion of channel hardening effect.

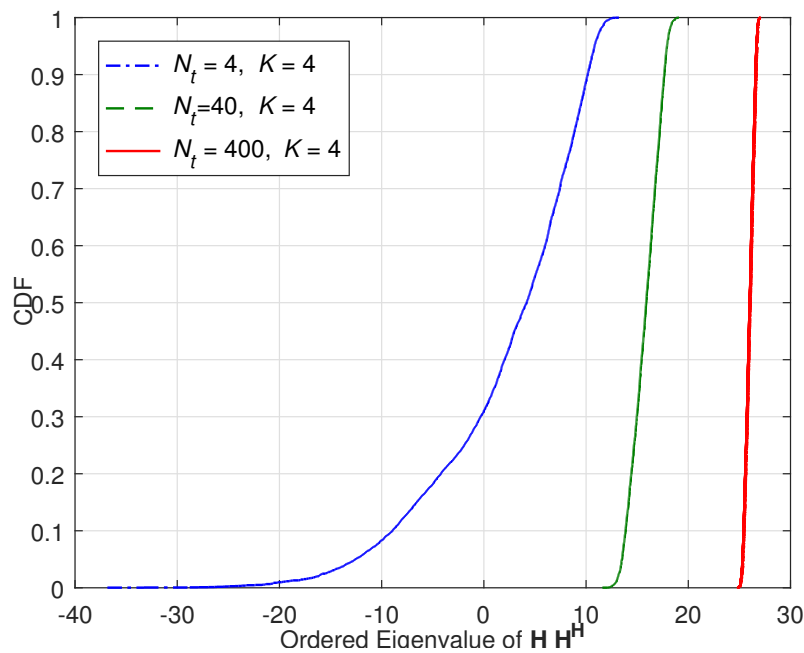


Figure 1. Example of channel hardening effect, when $N_t = 4, 40, 400$ and $K = 4$, CDF: cumulative distribution function.

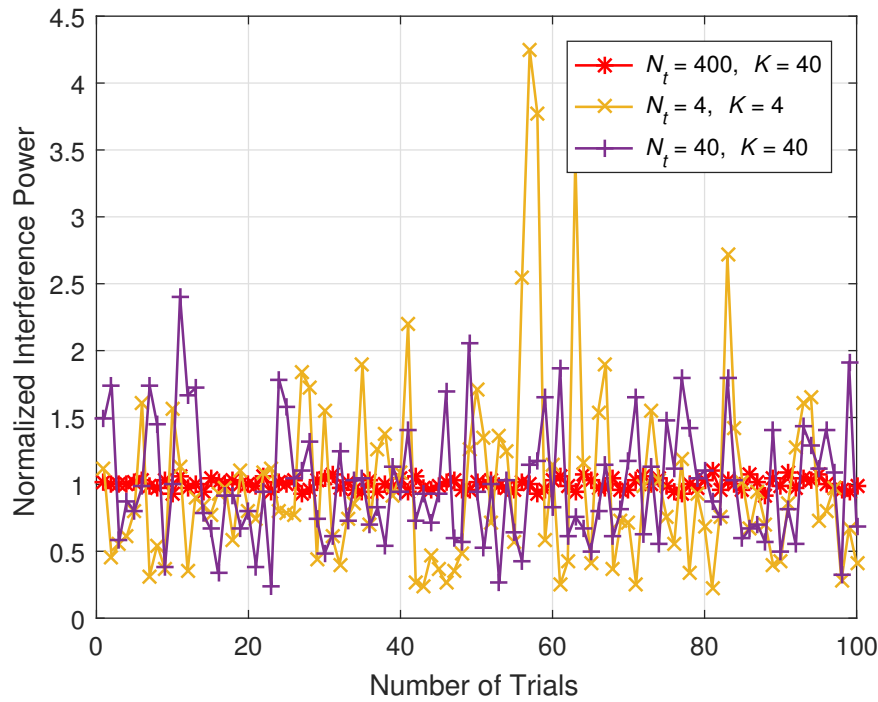


Figure 2. Normalized Interference Power versus Number of Trials: As N_t increases, the interference level converges to a certain point that can be easily estimated.

2.2. Precoding

Since N_t is much larger than K , we should map K message signal to each antenna. In general, we call it precoding, and it is applied to the transmitter part to reduce inter-user interference (IUI). Since N_t is very large, the linear precoding should be used for real systems [28–30]. For regular MIMO systems, both nonlinear and linear precoding scheme could be considered. Nonlinear precoding techniques, such as dirty paper coding (DPC) [31], vector perturbation (VP) [32], and lattice-aided methods [33] are important techniques when N_t is not much larger than K . However, with an increase in the number of antennas at the BS, linear precoders are shown to be near-optimal [28,29]. When N_t is, say, two times K , this gap is only 3 dB [29]. It is shown that with linear precoding, a sum rate as high as 98% of that of the DPC scheme can be achieved for two single antenna users served by 20 BS antennas [30]. Thus, it is more practical to use low-complexity linear precoding techniques in massive MIMO systems. Therefore, we mainly focus on linear precoding techniques.

It is well-known in the literature that zero-forcing (ZF) and regularized zero-forcing (RZF) are effective linear precoding techniques [34]. Also, in [4], T. Marzetta suggested even a simpler precoding technique, matched filtering (MF) precoding. The TX signal vector $\mathbf{s}$ and the message signal vector $\mathbf{x}$ are related as follows:

$$\mathbf{s} = \zeta \mathbf{F} \mathbf{x}, \quad (6)$$

where ζ is the normalization factor of TX power and $\mathbf{F}$ is a precoding matrix. Then, (1) can be represented as follows:

$$\mathbf{y} = \sqrt{\rho_t} \mathbf{G} \zeta \mathbf{F} \mathbf{x} + \mathbf{n}, \quad (7)$$

where $\mathbf{F}$ is the $N_t \times K$ precoding matrix. The three representative precoding matrices are denoted in Table 1.

Table 1. The precoding matrices of MF (matched filtering), ZF (zero-forcing), and RZF.

Precoder	MF	ZF	RZF
F	$\mathbf{H}^H$	$\mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}$	$\mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \nu\mathbf{I}_K)^{-1}$

Where superscript “ H ” denotes conjugate transpose, $(\cdot)^{-1}$ is the inverse operator, and $\mathbf{I}_K$ is the $K \times K$ identity matrix.

MF precoding matrix is simply the conjugate transpose of the channel matrix. This is conceptually simplest approach to reduce the IUI. In MF precoding, it is seen that the signal-to-interference and noise ratio (SINR) can be made as high as desired by increasing the number of antennas. However, the MF precoding exhibits an error floor in the practical number of TX antennas. The main advantage of MF precoding is its low computational complexity. Another advantage is that it is effective for distributed antenna systems, because the Massive MIMO signal processing can be performed locally at each antenna [5]. Another approach for precoding is to invert the channel by means of the pseudo-inverse. This is referred to as ZF precoding. When $N_t > K$, ZF precoding completely removes the IUI. Moreover, when SNR is high and/or $N_t \gg K$, ZF precoding can achieve nearly optimum performance [29]. However, a disadvantage of ZF is that processing cannot be done distributedly at each antenna separately. With ZF precoding, all data must instead be collected at a central node that handles the processing [29]. MF precoding can outperform ZF precoding when SNR is very low, while ZF precoding outperform MF precoding when SNR is relatively high. When SNR is not so high, ZF precoding cannot give any meaningful performance. RZF precoding can combat this problem. When SNR is high, both ZF precoding and RZF precoding shows very similar performance, while when SNR is low, RZF precoding outperforms ZF precoding. When SNR is very low, RZF precoding show similar performance with MF precoding. A brief summary of the advantages and disadvantages of the MF, ZF, and RZF is shown in Table II of [30].

The normalization factor should be determined such that the total transmit power becomes ρ_t . It can be expressed as follows:

$$\|\zeta \mathbf{F} \mathbf{x}\|_F^2 = 1, \quad (8)$$

where $\|\cdot\|_F$ stands for Frobenius norm.

ζ for k th UE, ζ_k is approximated as $\zeta_{k,mf} \approx \frac{1}{\sqrt{N_t K}}$ for MF precoding and $\zeta_{k,ZF} \approx \sqrt{\frac{N_t - K}{K}}$ for ZF precoding [35].

Both ZF and RZF precodings show very similar performance, even in the relatively practical number of massive TX antennas (i.e., $N_t \approx 10K$).

2.3. RS Overhead Model

As a reference model, the current 3GPP LTE systems use following types of RSs; common reference signal (CRS), channel state information reference signal (CSI-RS), demodulation reference signal (DM-RS), multicast-broadcast single-frequency network (MBSFN) reference signal, positioning reference signal (PRS). In this paper, we only consider CRS, CSI-RS, and DM-RS, because the three signals take most of resources for RS [36,37]. PRS are only transmitted in resource blocks over downlink subframes that are configured for PRS transmission, and The MBSFN reference signals are only transmitted in the MBSFN region of MBSFN subframes [36,37], thus PRS and MBSFN are not included in general operation mode and we disregard both cases for simplicity.

CRS is called as a cell specific reference signal, and has been in the LTE system from Release 8. The role of CRS can be defined as the cell search and initial acquisition, the downlink channel estimation for coherent demodulation/detection at the UE, and the downlink channel quality measurements. CSI-RS has been introduced from Release 10, and used by UE to estimate the channel and report channel quality information (CQI) to BS. DM-RS is usually called as an UE specific reference signal, and the role of it is for the demodulation of the signal.

In this paper, assuming frequency division multiplexing (FDD) mode, we use the following RS overhead factor for analysis [20]:

$$\chi(\%) = (\eta_{CRS} + N_t + K \cdot \psi_{DM-RS}) / \eta_{RB_{tot}} \times 100, \quad (9)$$

where $\chi(\%)$ is expressed as a percentage of RS overhead in a given total resources, η_{CRS} is the number of CRS in a given resource, ψ_{DM-RS} is the DM-RS proportional factor of K which we use 10, and $\eta_{RB_{tot}}$ is the total resource elements for a given resource block for a coherence time. In the current 3GPP LTE-A systems, the number of resource elements available in two resource blocks (1 ms) is 168 (12 (frequency tones) $\times$ 14 (time symbols)). Then, $\eta_{RB_{tot}} = 840$ resource elements assuming 5 ms coherence time. The CRS takes 14.8% of available resource elements, which is not a small portion. We do not know how technology will be evolved from an RS perspective. However it is quite obvious that RS overhead increases as the number of TX antennas increases. In this respect, (9) is reliable enough to use. The expected RS overhead reduction factor can be written as follows [20]:

$$\chi_r(\%) = (\eta_{CRS} + (N_t + K \cdot \psi_{DM-RS}) / N_g) / \eta_{RB_{tot}} \times 100, \quad (10)$$

where N_g is the RS overhead reduction factor to represent the number of antennas in each group, and the performance of RS overhead reduction scheme. Here we assume that the CRS is irreducible for cell search and initial acquisition, downlink channel estimation for coherent demodulation/detection at the UE, and downlink channel quality measurements, while all of the other RSs are ideally reducible by the RS overhead reduction scheme. Since resources of the CSI-RS and DM-RS are shared in the same group of antenna elements, it is divided by N_g . In other word, N_g number of antenna elements share the same CSI-RS and DM-RS resources, and they are not orthogonal in each group, thus in each group, CSI-RS and DM-RS resources are reduced by factor of N_g , but CRS is irreducible and not divided by N_g , because it requires high accuracy. Reducing RS resources causes channel estimation error.

3. Closed-Form Approximation of Achievable Spectral Efficiency with Channel Estimation Error

In this Section, we present the closed-form approximation of achievable SE with channel estimation error. We should keep in mind that antenna grouping-based RS overhead reduction scheme causes channel estimation error. As N_g increases, channel estimation error increases. The estimated channel, $\hat{\mathbf{H}}$, is modeled as follows:

$$\hat{\mathbf{H}} = \zeta \mathbf{H} + \sqrt{1 - \zeta^2} \mathbf{E}, \quad (11)$$

where $\zeta \in [0, 1]$ is the error factor, which reflects the degree of channel estimation error, and $\mathbf{E} \in \mathbb{C}^{K \times N_t}$ is the error matrix with the same statistical characteristic but independent of the channel, $\mathbf{H}$. ζ can also be modeled as follows:

$$\zeta = \sqrt{\hat{\zeta} \left(1 - \frac{N_g}{\varepsilon} \right)}, \quad (12)$$

where $\hat{\zeta}$ is the channel estimation error regardless of antenna group, and ε is the adjustable factor to reflect the seriousness of the channel estimation error depending on the applied technique.

For the performance analysis, the maximum achievable SE can be derived by using i.i.d. Rayleigh channel with zero mean and unit variance. From (1), and (7), the symbol received by k -th user is given by

$$y_k = \sqrt{\rho_t} \zeta_k (\mathbf{h}_k \circ \mathbf{b}_k) \mathbf{f}_k x_k + n_k + \sqrt{\rho_t} \zeta_l \sum_{l \neq k} (\mathbf{h}_k \circ \mathbf{b}_k) \mathbf{f}_l x_l, \quad (13)$$

where $\mathbf{h}_k$ and $\mathbf{b}_k$ are the $1 \times N_t$ channel vectors for k -th user, and $\mathbf{f}_k$ is the $N_t \times 1$ precoding vector for k -th UE. The last term of (13) is the IUI. ζ_k is a normalized factor for the precoding process of k -th UE, and it is approximated as $\zeta_{k,mf} \approx \frac{1}{\sqrt{N_t K}}$ for MF precoding and $\zeta_{k,ZF} \approx \sqrt{\frac{N_t - K}{K}}$ for ZF precoding [35] as we mentioned in the previous section. Assuming all the path loss component to UE k is the same,

which is true in a real system, the effective SINR at the receiver (RX) for user k , γ_k , can be expressed as follows:

$$\gamma_k = \frac{\rho_r |\zeta_k \mathbf{h}_k \mathbf{f}_k|^2}{\rho_r \sum_{l \neq k} |\zeta_l \mathbf{h}_k \mathbf{f}_l|^2 + 1}, \quad (14)$$

where $\rho_r = \frac{\rho_t \beta_k}{N_0 B}$ is the received signal-to-noise ratio at RX, β_k is the large scale fading or path loss component between TX and k th UE, and $N_0 B$ is the noise power in the given bandwidth, B . As was shown in the previous section, path loss components from 1st, 2nd, 3rd, $\dots$, N_t th antenna in BS to antenna in k th UE are the same ($\beta_i = \beta_{i,1} = \beta_{i,2}, \dots, \beta_{i,N_t}$). This is because the distances among antennas in BS are much shorter than the distances between antennas in BS and antennas in UEs.

The data rate, R of single isolated cell can be represented as follows:

$$R = \alpha \cdot B \cdot \sum_{k=1}^K \mathbb{E} [\log_2 (1 + \gamma_k)], \quad (15)$$

where α is the scaling factor for RS overhead [4], and $\mathbb{E}[\cdot]$ is the expectation operation. In the ideal case, i.e., perfect blind channel estimation with no channel estimation error, α is equal to 1. In a real situation, however, by using (10), we can say that $\alpha = (1 - \frac{\lambda_r}{100})$, when a RS overhead reduction scheme is applied. Since the scaling factor should be more than 0.5 in the system design perspective, the spectral efficiency (SE) based on capacity analysis can be represented as follows:

$$SE \approx \begin{cases} \alpha \cdot \sum_{k=1}^K \mathbb{E} [\log_2 (1 + \gamma_k)], & \alpha > 0.5. \\ 0.5 \cdot \sum_{k=1}^{K'} \mathbb{E} [\log_2 (1 + \gamma'_k)], & \alpha \leq 0.5. \end{cases} \quad (16)$$

where K' and γ'_k represent corresponding constant when overhead is 0.5.

We show the reference SINR based on MF precoding. By using channel hardening effect of the massive MIMO systems [4], the reference SINR based on MF precoding, $\gamma_{k,mf}^{ref}$, can be simplified as follows [38]:

$$\gamma_{k,mf}^{ref} = \frac{\rho_r |\zeta_{k,mf} \mathbf{h}_k \mathbf{h}_k^H|^2}{\rho_r \sum_{l \neq k} |\zeta_{l,mf} \mathbf{h}_k \mathbf{h}_l^H|^2 + 1} \rightarrow \frac{N_t}{K} \left(\frac{\rho_r}{\mathcal{I}_{mf} + 1} \right) = \frac{N_t}{K} \left(\frac{\rho_r}{\rho_r + 1} \right), \quad (17)$$

where $\mathcal{I}_{mf}$ is the IUI term after MF precoding. For massive MIMO region, $\mathcal{I}_{mf}$ can be simplified to ρ_r [4]. Here we can say that $\frac{N_t}{K}$ is the channel gain of the massive MIMO system.

It is obvious that if we increase the number of TX antennas or reduce the number of UEs, we can get a more effective SINR. However, it should be noted that MF precoding does not completely remove the IUI in practical number of TX antennas. It is known that ZF precoding can completely remove IUI with the perfect channel information, but channel gain is reduced to $N_t - K$ as compared with MF precoding, N_t .

When there is a channel estimation error, the effective SINR based on MF precoding can be represented as follows:

$$\hat{\gamma}_{k,mf} = \frac{\rho_r |\zeta_{k,mf} \mathbf{h}_k \hat{\mathbf{h}}_k^H|^2}{\rho_r \sum_{l \neq k} |\zeta_{l,mf} \mathbf{h}_k \hat{\mathbf{h}}_l^H|^2 + 1}, \quad (18)$$

where $\hat{\mathbf{h}}_k$ is the $1 \times N_t$ estimated channel vector for k -th UE. Using (11) and statistical approximation, (18) can be simplified as follows:

$$\begin{aligned}\hat{\gamma}_{k,mf} &= \frac{\rho_r \left| \zeta_{k,mf} \tilde{\zeta} \mathbf{h}_k \mathbf{h}_k^H + \zeta_{k,mf} \sqrt{1 - \tilde{\zeta}^2} \mathbf{h}_k \mathbf{e}_k^H \right|^2}{\rho_r \sum_{l \neq k} \left| \zeta_{l,mf} \tilde{\zeta} \mathbf{h}_k \mathbf{h}_l^H + \zeta_{l,mf} \sqrt{1 - \tilde{\zeta}^2} \mathbf{h}_k \mathbf{e}_l^H \right|^2 + 1}, \\ &\rightarrow \frac{N_t}{K} \left(\frac{\tilde{\zeta}^2 \rho_r}{\rho_r + 1} \right),\end{aligned}\quad (19)$$

where $\mathbf{e}_l$ is the $1 \times N_t$ channel estimation error vector. For Equation (19), we use the following approximations:

$$\left| \zeta_{k,mf} \tilde{\zeta} \mathbf{h}_k \mathbf{h}_k^H + \zeta_{k,mf} \sqrt{1 - \tilde{\zeta}^2} \mathbf{h}_k \mathbf{e}_k^H \right|^2 \approx \tilde{\zeta}^2 \left| \zeta_{k,mf} \mathbf{h}_k \mathbf{h}_k^H \right|^2, \quad (20)$$

$$\sum_{l \neq k} \left| \zeta_{l,mf} \tilde{\zeta} \mathbf{h}_k \mathbf{h}_l^H + \zeta_{l,mf} \sqrt{1 - \tilde{\zeta}^2} \mathbf{h}_k \mathbf{e}_l^H \right|^2 \approx \sum_{l \neq k} \left| \zeta_{l,mf} \mathbf{h}_k \mathbf{h}_l^H \right|^2. \quad (21)$$

The approximation of (20) and (21) comes from the following equations:

$$\left| \zeta_{k,mf} \tilde{\zeta} \mathbf{h}_k \mathbf{h}_k^H \right|^2 \gg \left| \zeta_{k,mf} \sqrt{1 - \tilde{\zeta}^2} \mathbf{h}_k \mathbf{e}_k^H \right|^2, \quad (22)$$

$$\sum_{l \neq k} \left| \zeta_{l,mf} \tilde{\zeta} \mathbf{h}_k \mathbf{h}_l^H + \zeta_{l,mf} \sqrt{1 - \tilde{\zeta}^2} \mathbf{h}_k \mathbf{e}_l^H \right|^2 \approx \sum_{l \neq k} \left| \zeta_{l,mf} \tilde{\zeta} \mathbf{h}_k \mathbf{h}_l^H + \zeta_{l,mf} \sqrt{1 - \tilde{\zeta}^2} \mathbf{h}_k \mathbf{h}_l^H \right|^2. \quad (23)$$

By comparing (17) and (19), it can be seen that the SINR based on MF precoding with channel estimation error has the same characteristics, but the signal power is scaled down by the square of the error factor, $\tilde{\zeta}$.

Using (11) and statistical approximation, The effective SINR based on ZF precoding with channel estimation error can be written as follows [39]:

$$\begin{aligned}\hat{\gamma}_{k,ZF} &= \frac{\left| \zeta_{k,ZF} \mathbf{h}_k (\hat{\mathbf{H}} \hat{\mathbf{H}}^H)^{-1} \hat{\mathbf{h}}_k^H \right|^2}{\mathbf{h}_k (\hat{\mathbf{H}} \hat{\mathbf{H}}^H)^{-1} \hat{\mathbf{H}}_{[k]} \hat{\mathbf{H}}_{[k]}^H (\hat{\mathbf{H}} \hat{\mathbf{H}}^H)^{-1} \mathbf{h}_k^H \sum_{l \neq k} \zeta_{l,ZF} + 1}, \\ &\rightarrow \frac{N_t - K}{K} \left(\frac{\tilde{\zeta}^2 \rho_r}{(1 - \tilde{\zeta}^2) \rho_r + 1} \right),\end{aligned}\quad (24)$$

where $\hat{\mathbf{H}}_{[k]}^H = [\hat{\mathbf{h}}_1^H, \hat{\mathbf{h}}_2^H, \dots, \hat{\mathbf{h}}_{k-1}^H, \hat{\mathbf{h}}_{k+1}^H, \dots, \hat{\mathbf{h}}_K^H] \in \mathbb{C}^{N_t \times (K-1)}$.

For (24), the following approximations are used:

$$\begin{aligned}&\left| \zeta_{k,ZF} \mathbf{h}_k ((\tilde{\zeta} \mathbf{H}_k + \sqrt{1 - \tilde{\zeta}^2} \mathbf{E}_k) (\tilde{\zeta} \mathbf{H}_k + \sqrt{1 - \tilde{\zeta}^2} \mathbf{E}_k)^H)^{-1} \hat{\mathbf{h}}_k^H \right|^2 \\ &\approx \tilde{\zeta}^2 \left| \zeta_{k,ZF} \mathbf{h}_k \mathbf{h}_k^H \right|^2,\end{aligned}\quad (25)$$

$$\begin{aligned}
& \mathbf{h}_k((\zeta\mathbf{H}_k + \sqrt{1-\zeta^2}\mathbf{E}_k)(\zeta\mathbf{H}_k + \sqrt{1-\zeta^2}\mathbf{E}_k)^H)^{-1} \cdot \\
& ((\zeta\mathbf{H}_{[k]} + \sqrt{1-\zeta^2}\mathbf{E}_{[k]})(\zeta\mathbf{H}_{[k]} + \sqrt{1-\zeta^2}\mathbf{E}_{[k]})^H) \cdot \\
& ((\zeta\mathbf{H}_k + \sqrt{1-\zeta^2}\mathbf{E}_k)(\zeta\mathbf{H}_k + \sqrt{1-\zeta^2}\mathbf{E}_k)^H)^{-1} \cdot \mathbf{h}_k \sum_{l \neq k} \zeta_{l,ZF} \\
& \approx (1-\zeta^2) \sum_{l \neq k} \left| \zeta_{l,ZF} \mathbf{h}_k \hat{\mathbf{h}}_l^H \right|^2.
\end{aligned} \tag{26}$$

From (17) and (24), we can see that the SINR based on ZF precoding with channel estimation error is scaled down by ζ^2 and $(1-\zeta^2)$ for the signal and the interference power from reference SINR based on MF precoding, respectively. Obviously, the residual IUI term for ZF precoding is due to channel estimation error.

The SINRs for the cases of reference MF precoding, reference ZF precoding, MF precoding with channel estimation error, and ZF precoding with channel estimation error are summarized in Table 2.

Table 2. Summary of approximated SINRs.

$\gamma_{k,MF}^{ref}$	$\gamma_{k,ZF}^{ref}$	$\hat{\gamma}_{k,MF}$	$\hat{\gamma}_{k,ZF}$
$\frac{N_t}{K} \left(\frac{\rho_r}{\rho_r+1} \right)$	$\frac{N_t-K}{K} (\rho_r)$	$\frac{N_t}{K} \left(\frac{\zeta^2 \rho}{\rho_r+1} \right)$	$\frac{N_t-K}{K} \left(\frac{\zeta^2 \rho_r}{(1-\zeta^2)\rho_r+1} \right)$

If we put the SINR in the Table 2 into the Equation (16), we can get the theoretical achievable SE of RS overhead reduction scheme with channel estimation error. In a real situation, the channel estimation error due to any kind of interference can be reflected in error factor ζ . The RS overhead reduction performance can be also reflected in RS overhead reduction factor and/or the number of antenna in each group, N_g . N_g increment guarantees N_t increment. However, increasing N_g is not always beneficial. As we can see from Equation (12), N_g increment also causes ζ increment. Obviously, ζ increment can cause performance reduction.

In the case of RZF precoding, the SINR is shown in Equation (19) of [39] and Equation (17) of [40]. Substituting the SINR of RZF, $\gamma_{k,RZF}$ into Equation (16), one can obtain the achievable SE of RZF precoding. Since the expressions in [39,40] are too complex to use in real system analysis, we show the simplified expression using the SINRs of MF and ZF precoding. Basically in low power regime, the SINR of RZF precoding is similar to MF precoding. In high power regime, the SINR of RZF is similar to ZF precoding. Based on these facts, we can approximate the SINR of RZF with channel estimation error as follows:

$$\hat{\gamma}_{k,RZF} \approx \begin{cases} \frac{(N_t-K)\zeta^2 \left(1-\frac{N_g}{\epsilon}\right) \rho_r}{K \left(\left(1-\zeta^2 \left(1-\frac{N_g}{\epsilon}\right)\right) \rho_r + 1 \right)}, & N_t \rho_r \gg K, \\ \frac{N_t \zeta^2 \left(1-\frac{N_g}{\epsilon}\right) \rho_r}{K(\rho_r+1)}, & N_t \rho_r \ll K. \end{cases} \tag{27}$$

Since we assume the system with $N_t > 10K$, due to the distinct channel gain, the SINR of RZF precoding is similar to the SINR of ZF precoding. We will show this in Section 5. Readers can refer to [39,40] for more details of RZF precoding.

4. N_g Determination and Related System Structure

In this Section, we propose how to determine N_g in an antenna group-based RS overhead reduced massive MIMO. Based on analysis in Section 3, ZF and MF precoded SE, SE_{ZF} and SE_{MF} can be approximated as follows:

$$SE_{ZF} \approx \begin{cases} \left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \cdot K \cdot \left[\log_2 \left(1 + \frac{(N_t - K) \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K \left(\left(1 - \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right)\right) \rho_r + 1 \right)}\right) \right], & \left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) > 0.5 \\ 0.5 \cdot K' \cdot \left[\log_2 \left(1 + \frac{(N'_t - K') \hat{\xi} \left(1 - \frac{N'_g}{\epsilon}\right) \rho_r}{K \left(\left(1 - \hat{\xi} \left(1 - \frac{N'_g}{\epsilon}\right)\right) \rho_r + 1 \right)}\right) \right], & \left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \leq 0.5 \end{cases} \quad (28)$$

$$SE_{MF} \approx \begin{cases} \left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \cdot K \cdot \left[\log_2 \left(1 + \frac{N_t \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K(\rho_r + 1)}\right) \right], & \left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) > 0.5 \\ 0.5 \cdot K' \cdot \left[\log_2 \left(1 + \frac{N'_t \hat{\xi} \left(1 - \frac{N'_g}{\epsilon}\right) \rho_r}{K'(\rho_r + 1)}\right) \right], & \left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \leq 0.5 \end{cases} \quad (29)$$

where K' , N'_t , and N'_g are corresponding K , N_t , and N_g when $\left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) = 0.5$.

A few observations are in order.

Observation 1: As N_g increases, bandwidth increases due to RS overhead reduction.

Observation 2: As N_g increases, SINR decreases due to channel estimation error.

Increasing N_g is beneficial to the system when system is working in bandwidth limited regime, but if system is working in power limited regime and/or interference limited regime, increasing N_g worsens system performance.

Two strategies can be considered to choose N_g . First, N_g can be determined to maximize SE which can be represented as follows for ZF precoding and MF precoding respectively.

$$\begin{aligned} \max_{N_g} & \left(\Theta \cdot \sum_{k=1}^K E \left[\log_2 \left(1 + \frac{(N_t - K) \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K \left(\left(1 - \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right)\right) \rho_r + 1 \right)} \right) \right] \right) \\ \text{s.t.} & \quad \Theta \geq 0.5 \\ & \quad \rho_r \leq \rho_{r,\max}, \end{aligned} \quad (30)$$

$$\begin{aligned} \max_{N_g} & \left(\Theta \cdot \sum_{k=1}^K E \left[\log_2 \left(1 + \frac{N_t \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K(\rho_r + 1)} \right) \right] \right) \\ \text{s.t.} & \quad \Theta \geq 0.5 \\ & \quad \rho_r \leq \rho_{r,\max}, \end{aligned} \quad (31)$$

where $\Theta = \left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right)$ and $\rho_{r,\max}$ is the maximum SNR which reflect the maximum allowable TX power. This is a typical setting for wireless communication systems up to now. However, to support distributed sensors and/or sensor networks, energy efficiency can be more important than SE depending on applications. In a certain application, sensor devices do not require high data rate, but maintaining moderate data rate can be enough. In this situation, N_g can be determined based on the minimum required SE in a given circumstances. In particular, for the energy efficient sensor networks, if the minimum required SE is satisfied, the system can operates efficiently with high energy efficiency and a moderate data rate. Also, if determination of N_g causes lower SE than the requirement, it could be problematic to the proper operation of sensor networks. It is particularly important to maintain the required data rate for industrial IoT and/or sensor networks. If N_g can be derived as a closed-form equation, it can be very useful to reduce the computational complexity and delay of the system. The inverses of Equations (28) and (29) can provide N_g with given required SE, however the exact closed-form inverse equations of Equations (28) and (29) are impossible, because N_g terms exist both inside and outside of log term. Instead, by keeping the $\frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}$ as 0.5, the minimum

value of data portion, we can get N_g with satisfying the minimum required value of SE. The proposed scheme is can be established based on following Proposition.

Proposition 1. $\left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \cdot K \cdot \left[\log_2 \left(1 + \frac{(N_t - K) \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K \left(\left(1 - \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right)\right) \rho_r + 1 \right)} \right) \right] \geq \Delta \cdot K \cdot \left[\log_2 \left(1 + \frac{(N_t - K) \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K \left(\left(1 - \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right)\right) \rho_r + 1 \right)} \right) \right]$ iff $\Delta \leq 0.5$ and $\left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \geq 0.5$.

Proof. Since $\left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \geq 0.5$, if $\Delta \leq 0.5$, then

$$\frac{\left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \cdot K \cdot \left[\log_2 \left(1 + \frac{(N_t - K) \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K \left(\left(1 - \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right)\right) \rho_r + 1 \right)} \right) \right]}{\Delta \cdot K \cdot \left[\log_2 \left(1 + \frac{(N_t - K) \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K \left(\left(1 - \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right)\right) \rho_r + 1 \right)} \right) \right]} = \frac{\left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right)}{\Delta} \geq 1,$$

thus Proposition 1 is satisfied. $\square$

From Proposition 1, the maximum value of Δ is 0.5. This means, if we fix $\Theta = 0.5$ and derive the N_g , the minimum required SE can be satisfied. Proposition 1 also can be applied to the MF precoding case.

Corollary 1. $\left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \cdot K \cdot \left[\log_2 \left(1 + \frac{N_t \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K(\rho_r + 1)} \right) \right] \geq \Delta \cdot K \cdot \left[\log_2 \left(1 + \frac{N_t \hat{\xi} \left(1 - \frac{N_g}{\epsilon}\right) \rho_r}{K(\rho_r + 1)} \right) \right]$ iff $\Delta \leq 0.5$ and $\left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) \geq 0.5$.

Proof. Proof of Corollary 1 can follow the same way of that of Proposition 1. $\square$

Let $\left(1 - \frac{\eta_{CRS} + (N_t + K \cdot \psi)}{\eta_{RB_{tot}} N_g}\right) = \delta$, then the closed-form equation for N_g for ZF precoding and MF precoding with the minimum required SE_{req} can be represented as follows:

$$N_{g,ZF} \rightarrow \frac{\epsilon \left(K + K\rho_r - 2K\rho_r \hat{\xi} + N_t \rho_r \hat{\xi} - K2^{\left(\frac{\widetilde{SE}_{ZF}}{\delta K}\right)} - K\rho_r 2^{\left(\frac{\widetilde{SE}_{ZF}}{\delta K}\right)} + K\rho_r \hat{\xi} 2^{\left(\frac{\widetilde{SE}_{ZF}}{\delta K}\right)} \right)}{\rho_r \hat{\xi} \left(-2K + N_t + K2^{\left(\frac{\widetilde{SE}_{ZF}}{\delta K}\right)} \right)}. \quad (32)$$

$$N_{g,MF} \rightarrow \frac{\epsilon K + \epsilon K\rho_r + \epsilon N_t \rho_r \hat{\xi} - \epsilon K 2^{\left(\frac{\widetilde{SE}_{MF}}{\delta K}\right)} - \epsilon K \rho_r 2^{\left(\frac{\widetilde{SE}_{MF}}{\delta K}\right)}}{N_t \rho_r \hat{\xi}}. \quad (33)$$

where $\widetilde{SE}_{ZF}$ and $\widetilde{SE}_{MF}$ are required minimum SE for ZF and MF precoded systems. By fixing $\delta = 0.5$, the lowest bound, we can get N_g that satisfy to $\widetilde{SE}_{ZF}$ and $\widetilde{SE}_{MF}$ with some margin.

The proposed scheme can be used in various ways. In particular, it would be very effective to reduce the power consumption of massive MIMO-based wireless sensor networks. Figure 3 shows example of a system block diagram for the proposed scheme. It consists of several system blocks.

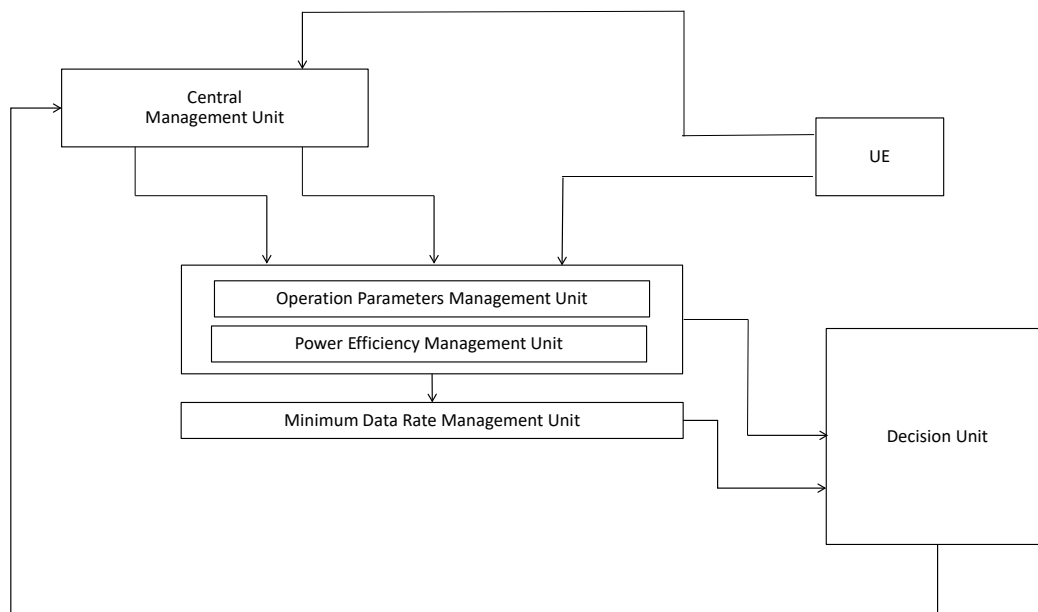


Figure 3. Example of system block diagram for proposed scheme.

First, the central management unit can decide the power consumption and/or power efficiency of the system. Then, BS measures current power efficiency. BS also has various operation parameters, such as N_t , K , B , and ρ_t and so on. Based on the power efficiency and operation parameters, the minimum data rate that must be satisfied for the system can be derived. If the system is designed for Industrial internet, there would be a threshold for data rate, and the system should set the minimum data rate higher than the threshold. Then, from the minimum data rate and other parameters that are necessary, N_g can be determined. The BS can receive feedback from UEs to check if they need a higher data rate depending on the situation. Wireless channel is not stable, so there could be several situations that can arise. Then, the procedures we mentioned can be repeated for a given number of iterations. There could be an interval for the determination of N_g . If the interval is small, then the performance would be better, but complexity could be a problem or vice versa.

Example of operation flow chart is presented in Figure 4. First, BS can get system parameters, then it gets required power efficiency. Based on system parameters and power efficiency, the required minimum data rate can be derived. The N_g can be derived using the proposed closed-form equation. The RS can be designed and used based on N_g . Then, communications between massive MIMO and distributed IoT devices would be conducted. If power consumption is higher than the threshold, we can reduce power consumption by any kind of parameter adjustment. Typically, TX power consumes a lot of power, thus it can be reduced to increase power efficiency. However, TX power reduction accompanies data rate reduction. If the data rate is higher than threshold even though TX power is reduced, then system is in good shape and the procedure can be finished, but if data rate is lower than the threshold due to the TX power reduction, then in a given number of iterations, we should reduce required minimum data rate to satisfy power efficiency. The determination of thresholds for the number of iterations, power consumptions, data rate are the system designer's choice. For the better understand and implementation, we provide the detailed algorithm as pseudo-code type in Algorithm 1.

The required number of iterations for the system can be changed based on system parameters. There are several parameters including the threshold for the number of iterations, thus system designer can decide it between performance and delay/complexity trade-off.

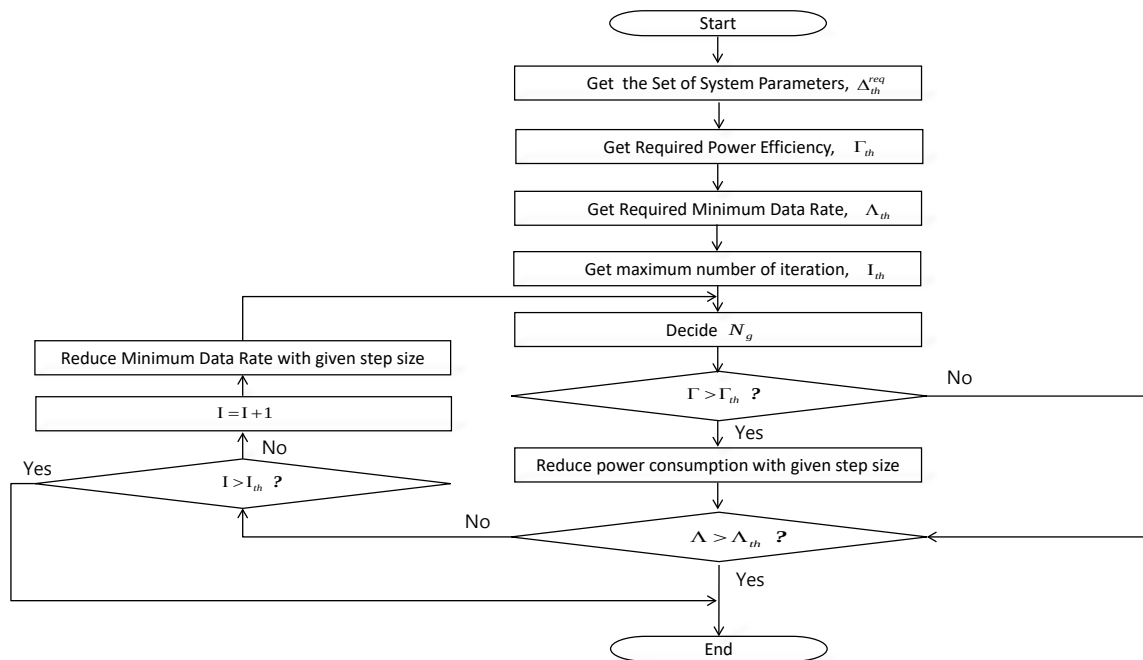


Figure 4. Example of operation flow chart.

Algorithm 1: Operation of proposed scheme

```

1 Start;
2 Get the Set of System Parameters,  $\Delta_{sp}^{req}$ ;
3 Get Required Power Efficiency,  $\Gamma_{th}$ ;
4 Get Required Minimum Data Rate,  $\Lambda_{th} = B \cdot SE$ ;
5 Get Maximum Number of Iterations,  $I_{th}$ ;
6 Decide  $N_g$  based on Equations (32) and (33);
7 if  $\Gamma > \Gamma_{th}$ ;
8 then
9   Reduce  $\Gamma$  to  $\Gamma_{th}$ ;
10  if  $\Lambda > \Lambda_{th}$ ;
11  then
12    Go to the End;
13  else
14    Check the number of iterations,  $I_{th}$ ;
15    if  $I > I_{th}$ ;
16    then
17      Go to the End;
18    else
19       $I = I + 1$ ;
20      Reduce  $\Lambda_{th}$ ;
21      Go to 6;
22 else
23   Go to 10;
24 End;
  
```

5. Numerical Results and Discussion

In this Section, we verify the proposed scheme using monte-carlo (MC) simulations with 10^4 iterations, and present the related discussions. The simulation parameter for this section is shown in Table 3.

Table 3. Simulation Parameters.

Parameter	Value
Coherence Time, τ_c	5 ms
Coherence Bandwidth, BW_c	180 kHz
max RS time, τ_p	2.5 ms
Total symbols in coherence interval, S_{tot}	840
TX power, ρ_t	80 W
Signal Bandwidth, B	20 MHz
Carrier Frequency, f_c	2 GHz
Path Loss Model	ETSI
Number of TX antennas, N_t	100~1000
Max. number of serviced UEs for downlink,	100
Error factor without RS overhead reduction, $\hat{\zeta}$	1
Precoding	ZF, MF

We used $K = 0.1N_t$ to meet the minimum requirement of channel hardening effect. The noise power N_0B is normalized to unity with $\rho_t = 80$ W and $B = 20$ MHz bandwidth. Assuming we use a macro-cell type setup with 2GHz carrier frequency, the path loss in dB is modeled as $128.1 + 37.6 \log(d)$ with distance d in kilometers [41]. Then, the setup is equivalent to average user distance of 1.194 km. We set $\hat{\zeta} = 1$ and $\epsilon = 300$ in Equation (12). Obviously, the parameter can be changed depending on the applied technique and circumstances.

Before showing the numerical results, we present the seriousness of channel estimation error in Figure 5.

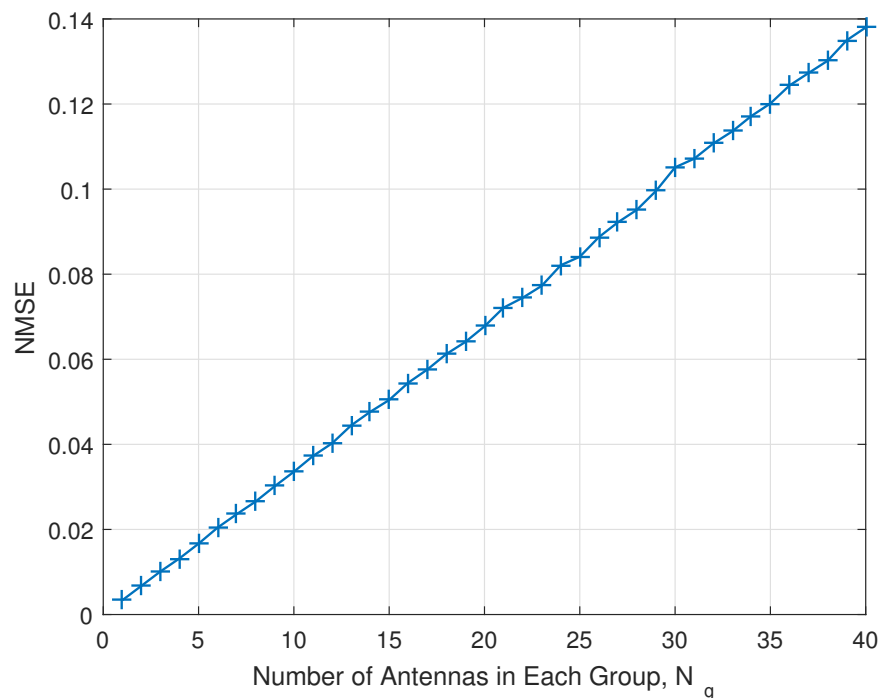


Figure 5. Seriousness of channel estimation error: normalized mean square error (NMSE) versus N_g .

We show the seriousness of channel estimation error using normalized mean square error (NMSE) which is defined as:

$$NMSE = \frac{\|\mathbf{H} - \hat{\mathbf{H}}\|_2^2}{\|\mathbf{H}\|_2^2}. \quad (34)$$

$\hat{\mathbf{H}}$ is the estimated channel and it is a function of N_g as we shown in the previous section. NMSE is linearly increased as N_g increases. When N_g is 40, NMSE has reached around 14%, while we can get bandwidth benefit thanks to the RS overhead reduction.

Figure 6 shows the numerical results of SE versus N_t . Red ‘*’s and black ‘x’ indicate the simulation results, and lines are plotted based on theoretical analysis. Our theoretical analysis is well-matched with the simulation result. When N_g is small which is less than 4, due to the excessive RS overhead which reaches the maximum allowable RS overhead, SE is not increased, even though N_t is increased. However, when N_g is larger than 8, the SE increases, as N_t increases up to 1000. RS overhead reduction technique is definitely necessary to increase N_t and SE, and fully enjoy the benefit of massive MIMO. However, increasing N_g causes channel estimation error as we have seen from previous Section. Increasing N_g gives a bandwidth advantage by giving more room for data, but due to channel estimation error, SINR becomes worse.

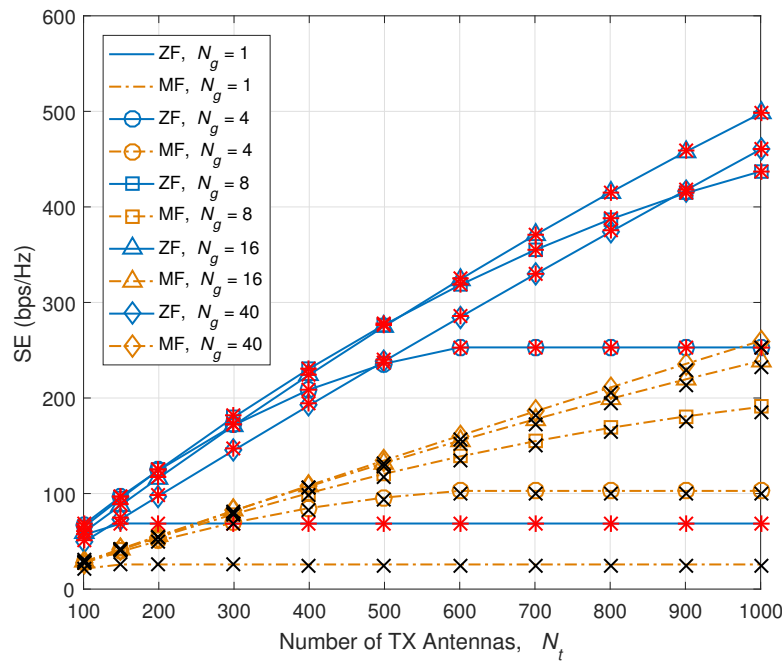


Figure 6. SE versus number of TX antennas, N_t . Red ‘*’s and black ‘x’ indicate the simulation results.

To show the relationship between SE and N_g , we present Figure 7, SE versus N_g . As observed, if N_g is larger than 16, the SE is reduced due to worsened SINR. It is noticeable in ZF precoding because as N_g grows, IUI terms grows. The IUI term in ZF precoding is almost negligible when there is little channel estimation error. In the case of MF precoding, there is no SE reduction, but SE is fixed to a certain point even though N_g increases. This is because IUI is already a dominant term in MF precoding even when there is little channel estimation error. When we use MF precoding, if N_g is larger than 16, the system is already working in a power limited regime, thus increasing N_g gives little help in increasing SE. This phenomenon is maintained even if we increase N_t .

Now, we show the effectiveness of the closed-form equation for N_g determination. Figures 8 and 9 show the determined N_g versus SE. ‘*’ marks indicate results of closed-form SE equation, and ‘x’ marks indicate results of MC simulations. Lines are plotted based on theoretical derivations. As observed, all 3 kinds of results are well-matched, and the derived closed-form equations for N_g can be used with

high accuracy. In the case of ZF precoding, the lower bound of SE is sensitive to the variation of N_g , while in the case of MF precoding, the lower bound of SE is not so sensitive to the variation of N_g .

Maintaining SE lower bounds is quite important to the energy efficient sensor network. The important metric of wireless sensor network is moving from SE to energy efficiency. However, keeping a stable SE lower bound is also very important for some kinds of sensor network, such as Industrial internet network. The proposed scheme provides the logic of keeping high energy efficiency maintaining SE lower bound, thus it can be a core technology for the massive MIMO-based energy efficient sensor network.

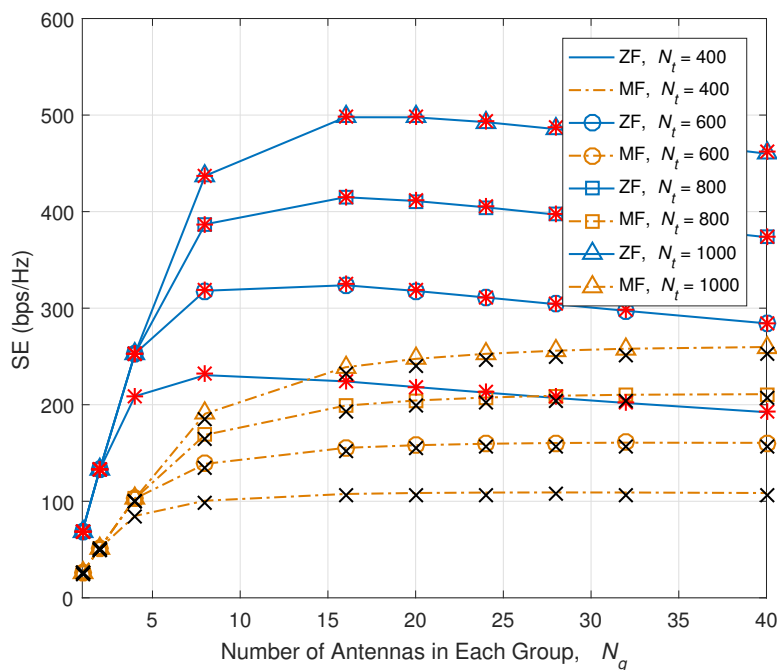


Figure 7. SE versus number of antennas in each group, N_g . Red '*'s and black 'x' indicate the simulation results.

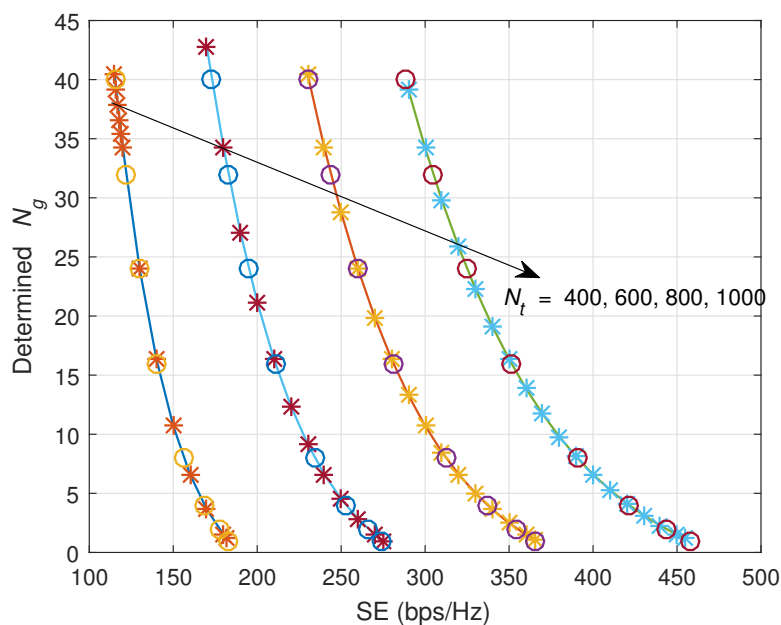


Figure 8. Determined N_g versus guaranteed SE (bps/Hz) when ZF precoding is applied. '*' marks indicate the results of closed-form SE equation, and 'O' marks indicate results of MC simulations.

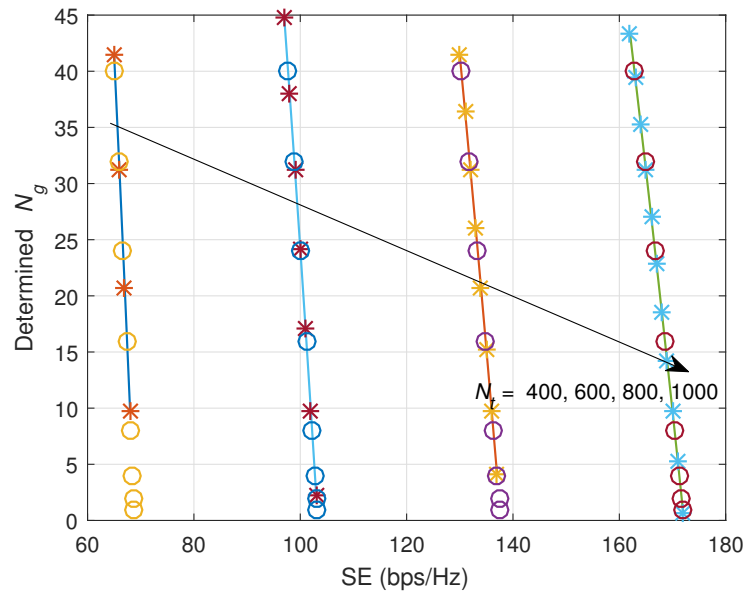


Figure 9. Determined N_g versus guaranteed SE (bps/Hz) when MF precoding is applied. “*” marks indicate results of closed-form SE equation, and ‘O’ marks indicate results of MC simulations.

As we mentioned in the previous section, for the case of RZF precoding, we can use the case of ZF precoding. Figure 10 presents the SE comparison of ZF precoding and RZF precoding. We use $\nu \triangleq \text{diag} \left[\frac{1}{\rho_r}, \dots, \frac{1}{\rho_r} \right]$ which is the downlink RX inverse SNR matrix [28,30,40]. As observed, the SE of RZF precoding also exactly matches with the analytical SE of ZF precoding. Thus, all the analysis for ZF precoding in this paper is also applicable to RZF.

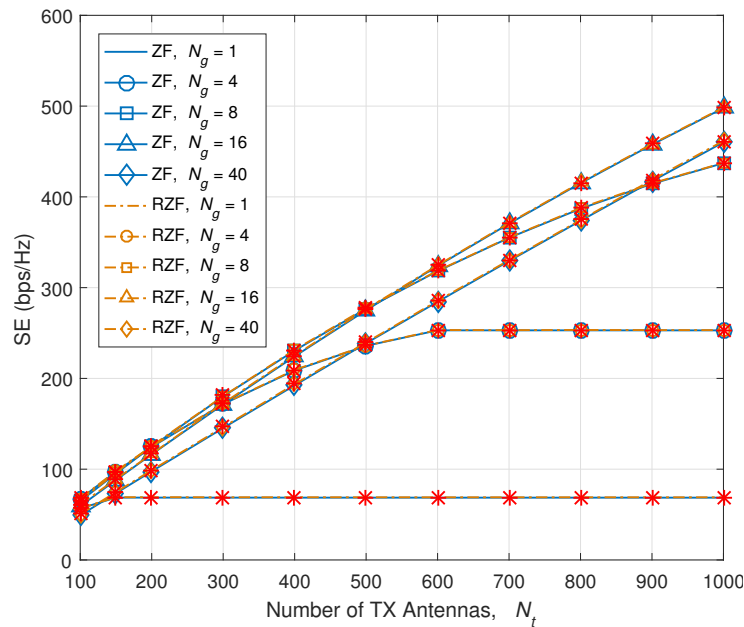


Figure 10. SE comparison of ZF and RZF precoding. “*” marks indicate results of closed-form SE equation for ZF precoding.

To completely validate the proposed system, it is important to perform real measurement. Since we show the feasibility of the proposed system in this paper, the real measurement for the system can be performed in future work.

6. Conclusions

In this paper, we have proposed a design logic of RS overhead reduced massive MIMO for energy efficient wireless sensor networks. The contribution of this paper can be divided into two parts. First, we have presented the closed-form achievable SE of massive MIMO with channel estimation error. Second, based on the the results, we have provided an effective scheme to determine the number of antennas required in each group. Numerical results showed that the derived theoretical achievable SE and N_g are accurate enough, and thus be provided as an useful tool for the design of massive MIMO-based wireless sensor networks and many other related engineering applications.

Acknowledgments: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017R1D1A1B03028350).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ji, H.; Kim, Y.; Lee, J.; Onggosanusi, E.; Nam, Y.; Zhang, J.; Lee, B.; Shim, B. Overview of Full-Dimension MIMO in LTE-Advanced Pro. *IEEE Commun. Mag.* **2017**, *55*, 176–184.
2. Chinnadurai, S.; Selvaprabhu, P.; Jeong, Y.; Jiang, X.; Lee, M.H. Worst-Case Energy Efficiency Maximization in a 5G Massive MIMO-NOMA System. *Sensors* **2017**, *17*, 2139.
3. Selvaprabhu, P.; Chinnadurai, S.; Li, J.; Lee, M.H. Topological Interference Management for K-User Downlink Massive MIMO Relay Network Channel. *Sensors* **2017**, *17*, 1896.
4. Marzetta, T.L. Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas. *IEEE Trans. Wirel. Commun.* **2010**, *9*, 3590–3600.
5. Marzetta, T.L. Massive MIMO: An Introduction. *Bell Labs Tech. J.* **2015**, *20*, 11–22.
6. Björnson, E.; Larsson, E.G.; Marzetta, T.L. Massive MIMO: Ten myths and one critical question. *IEEE Commun. Mag.* **2016**, *54*, 114–123.
7. Gimenez, S.; Roger, S.; Baracca, P.; Martín-Sacristán, D.; Monserrat, J.F.; Braun, V.; Halbauer, H. Performance Evaluation of Analog Beamforming with Hardware Impairments for mmW Massive MIMO Communication in an Urban Scenario. *Sensors* **2016**, *16*, 1555.
8. Maruta, K.; Iwakuni, T.; Ohta, A.; Arai, T.; Shirato, Y.; Kurosaki, S.; Iizuka, M. First Eigenmode Transmission by High Efficient CSI Estimation for Multiuser Massive MIMO Using Millimeter Wave Bands. *Sensors* **2016**, *16*, 1051.
9. Lee, B.M. Calibration for Channel Reciprocity in Industrial Massive MIMO Antenna Systems. *IEEE Trans. Ind. Inform.* **2017**, doi:10.1109/TII.2017.2749431.
10. Ding, G.; Gao, X.; Xue, Z.; Wu, Y.; Shi, Q. Massive MIMO for Distributed Detection with Transceiver Impairments. *IEEE Trans. Veh. Technol.* **2017**, doi:10.1109/TVT.2017.2747772.
11. Shirazinia, A.; Dey, S.; Ciuonzo, D.; Rossi, P.S. Massive MIMO for Decentralized Estimation of a Correlated Source. *IEEE Trans. Signal Process.* **2016**, *64*, 2499–2512.
12. Ciuonzo, D.; Rossi, P.S.; Dey, S. Massive MIMO channel-aware decision fusion. *IEEE Trans. Signal Process.* **2015**, *63*, 604–619.
13. Ciuonzo, D.; Rossi, P.S.; Dey, S. Massive MIMO meets decision fusion: Decode-and-fuse vs. decode-then-fuse. In Proceedings of the 2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM), A Coruna, Spain, 22–25 June 2014.
14. Jiang, F.; Chen, J.; Swindlehurst, A.L.; Lopez-Salcedo, J.A. Massive MIMO for Wireless Sensing with a Coherent Multiple Access Channel. *IEEE Trans. Signal Process.* **2015**, *63*, 3005–3017.
15. Vien, V.M.; Jeong, Y.; Shin, H. Error exponents for distributed detection. *IEEE Commun. Lett.* **2016**, *20*, 121–124.
16. Noh, S.; Zoltowski, M.D.; Sung, Y.; Love, D.J. Pilot Beam Pattern Design for Channel Estimation in Massive MIMO Systems. *IEEE J. Sel. Top. Signal Process.* **2014**, *8*, 787–801.
17. Mohammadian, R.; Amini, A.; Khalaj, B.H. Compressive Sensing-Based Pilot Design for Sparse Channel Estimation in OFDM Systems. *IEEE Commun. Lett.* **2017**, *21*, 4–7.
18. Chen, Z.; Wang, L.; Zhao, Z.; Huang, H. ICA Filtering Based Channel Estimation for Massive MIMO TDD Systems. *IEEE Commun. Lett.* **2016**, *20*, 2253–2256.

19. He, X.; Song, R.; Zhu, W.-P. Pilot Allocation for Sparse Channel Estimation in MIMO-OFDM Systems. *IEEE Trans. Circuits Syst. II Express Briefs* **2013**, *60*, 612–616.
20. Lee, B.M.; Kim, Y. Beam Grouping Based RS Resource Reuse and De-Contamination in Large Scale MIMO Systems. *Appl. Sci.* **2017**, *7*, 96.
21. Adhikary, A.; Nam, J.; Ahn, J.-Y.; Caire, G. Joint Spatial Division and Multiplexing—The Large-Scale Array Regime. *IEEE Trans. Inf. Theory* **2013**, *59*, 6441–6463.
22. Shin, C.; Heath, R.W.; Powers, E.J. Blind channel estimation for MIMO-OFDM systems. *IEEE Trans. Veh. Technol.* **2007**, *56*, 670–685.
23. Petropulu, A.P.; Zhang, R.; Lin, R. Blind OFDM Channel Estimation through Simple Linear Precoding. *IEEE Trans. Wirel. Commun.* **2004**, *3*, 647–655.
24. Coldrey, M.; Bohlin, P. Training-Based MIMO Systems—Part I: Performance Comparison. *IEEE Trans. Signal Process.* **2007**, *55*, 5464–5476.
25. Coldrey, M.; Bohlin, P. Training-Based MIMO Systems: Part II —Improvements Using Detected Symbol Information. *IEEE Trans. Signal Process.* **2008**, *56*, 296–303.
26. Lee, B.M.; Choi, J.; Bang, J.; Kang, B.-C. An energy efficient antenna selection for large scale green MIMO systems. In Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Beijing, China, 19–23 May 2013.
27. Lee, B.M.; Kim, Y. Transmission Power Determination based on Power Amplifier Operations in Large Scale MIMO-OFDM Systems. *Appl. Sci.* **2017**, *7*, 709.
28. Lu, L.; Li, G.Y.; Swindlehurst, A.L.; Ashikhmin, A.; Zhang, R. An Overview of Massive MIMO: Benefits and Challenges. *IEEE J. Sel. Top. Signal Process.* **2014**, *8*, 742–758.
29. Rusek, F.; Persson, D.; Lau, B.K.; Larsson, E.G.; Marzetta, T.L.; Edfors, O.; Tufvesson, F. Scaling up MIMO: Opportunities and Challenges with Very Large Arrays. *IEEE Signal Process. Mag.* **2013**, *30*, 40–60.
30. Fatema, N.; Hua, G.; Xiang, Y.; Peng, D.; Natgunanathan, I. Massive MIMO Linear Precoding: A Survey. *IEEE Syst. J.* **2017**, doi:10.1109/JSYST.2017.2776401.
31. Costa, M. Writing on dirty paper. *IEEE Trans. Inf. Theory* **1983**, *29*, 439–441.
32. Hochwald, B.M.; Peel, C.B.; Swindlehurst, A.L. A vector-perturbation technique for near-capacity multiantenna communication-part II: Perturbation. *IEEE Trans. Commun.* **2005**, *53*, 537–544.
33. Windpassinger, C.; Fischer, R.F.H.; Huber, J.B. Lattice-reduction-aided broadcast precoding. *IEEE Trans. Commun.* **2004**, *52*, 2057–2060.
34. Spencer, Q.H.; Swindlehurst, A.L.; Haardt, M. Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels. *IEEE Trans. Signal Process.* **2004**, *52*, 461–471.
35. Marzetta, T.L.; Larsson, E.G.; Yang, H.; Ngo, H.Q. *Fundamentals of Massive MIMO*; Cambridge University Press: London, UK, 2016.
36. Ahmadi, S. *LTE—Advanced: A Practical Systems Approach to Understanding 3GPP LTE Releases 10 and 11 Radio Access Technologies*, 1st ed.; Academic Press: Cambridge, MA, USA, 2013.
37. Sesia, S.; Toufik, I.; Baker, M. *LTE—The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed.; Wiley: Hoboken, NJ, USA, 2011.
38. Lee, B.M.; Choi, J.H.; Bang, J.H.; Kang, B.C. An energy efficient PAPR reduction technique for mobile terminals. In Proceedings of the IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 13–16 January 2012; pp. 366–367.
39. Wagner, S.; Couillet, R.; Debbah, M.; Slock, D.T.M. Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels Under Limited Feedback. *IEEE Trans. Inf. Theory* **2012**, *58*, 4509–4537.
40. Hoydis, J.; Brink, S.; Debbah, M. Comparison of linear precoding schemes for downlink massive MIMO. In Proceedings of the 2012 IEEE International Conference on Communications (ICC), Ottawa, ON, Canada, 10–15 June 2012; pp. 2136–2139.
41. *Radio Frequency (RF) Requirements for LTE Pico Node B*; ETSI Technical Report 136.931; European Telecommunications Standards Institute: Sophia Antipolis, France, 2012.

