

## Investigation of Ideal Number User Terminals with Spectrum Efficiency in Next Generation Wireless Communication Systems



Osman Dikmen\*, Selman Kulaç

Department of Electrical Electronics, Faculty of Engineering, Duzce University, Duzce 81620, Turkey

Corresponding Author Email: [osmandikmen@duzce.edu.tr](mailto:osmandikmen@duzce.edu.tr)

<https://doi.org/10.18280/ts.380112>

### ABSTRACT

**Received:** 18 December 2020

**Accepted:** 2 February 2021

#### Keywords:

*massive MIMO, spectrum efficiency, multicellular system, user equipment, pilot reuse factor, 6G*

Suppose that a multi-user multiple-input multiple-output (MIMO) system is developed from scratch to equally envelop a defined region with optimal spectrum efficiency (SE) in next generation wireless communication systems such as sixth-generation (6G) and beyond networks. What are the ideal number of user terminals  $U$ , number of base stations antennas, and used pilot reuse factor? The purpose of this paper is to address this specific issue. Three interference levels are specified for this. Based on these interference levels, signal-to-interference-and-noise ratios (SINRs) are extracted. Closed-form spectrum efficiency equations are thus obtained. As a function of the base station (BS) antenna number, simulations are carried out considering multiple pilot reuse factors and diverse processing schemes such as Maximum Ratio Combining (MRC) and Zero-Forcing (ZF). From the results, it is understood that  $U$  varies according to the processing schemes. Therefore, evaluating the results considering the fixed number of users  $K$  will not give an accurate result in determining the design parameters for the next generation communication systems. In general, these results are useful statements that spectrum efficiency is maximized when the ideal number of users  $U$  is used in multi-cell systems.

## 1. INTRODUCTION

Massive MIMO system draws great attention for the next generation wireless communication technologies [1]. In the Massive MIMO system, an antenna array generally serves multiple terminals simultaneously. However, the number of antennas in the base station is quite high compared to terminals. Such a system is studied by Marzetta [2]. As can be seen in the study, Time Division Duplex (TDD) scenario has been taken into consideration. In this scenario, the time frame is divided between "reverse link pilots" and "forward link data transmission". These pilots are used in channel estimation. As a result, pre-coding is used to be used in data transmission [1]. Although the number of terminals has an effect on the time required for pilots, the number of antennas on the base station has no effect [3]. The coherence time is the interval based on the assumption that the channel is constant. This parameter depends on the variable mobility of the user and especially on the carrier frequency. Coherence time has an important place in Massive MIMO, especially in channel estimation. In Massive MIMO, channel estimation is carried out very intelligently. Time division duplex, uplink pilots and channel reciprocity are used to ensure that the estimated overhead is independent of the number of antennas. Therefore, Massive MIMO's coherence time will not limit the number of antennas. However, channel coherence is absolutely important in determining how often we need to predict the channel over time and frequency. The number of channels used per consistency block is the product of the coherence time and consistency bandwidth. The number of terminals to be served is limited by the coherence time. The coherence time depends on the mobility of the terminals [4]. In LTE systems, the

channels are estimated at a regular time interval. This interval is independent of the coherence time of the users. However, it is expressed as the shortest time of coherent users can have. The coherence time is inversely proportional to the user speed. For this reason, the time interval is specified as the maximum mobility speed [5]. Thus, a better channel estimation can be achieved. However, there is a situation where the coherence time is high for most users. This will lead to more predictions of channel estimation than usual. As a result, excessive use of radio channels will occur. Although better results can be obtained in conditions with a small number of users, more efficient studies are required for Massive MIMO systems where the number of users per cell is high.

Massive MIMO, the use of which has recently increased in wireless communication technologies, is known as the system in which communication is carried out through multiple antennas in the receiver and transmitter parts. It allows communication with multiple terminals at the same time. Among the main benefits of using multiple antennas are higher performance achieved with diversity and greater data rate through spatial multiplexing. There is no limit to the number of antennas. In fact, this situation made it possible to carry out studies on the comparisons of the differences in the number of antennas in the receiver and transmitter. It is seen that various researches are continuing on subjects such as design, analysis and function of Massive MIMO. In the study [6], which is one of these researches, theoretical studies on network design and signal processing and signal encoding for Massive MIMO system are presented. Massive MIMO system is a concept designed especially for 5G and beyond technologies. It is planned to equip thousands of antennas in the base station. In this way, it plays a leading role, especially in improving

spectrum efficiency [7]. With this system, not only spectrum efficiency but also energy efficiency, higher capacity, higher speed and less latency can be achieved. Tests have also been carried out on 5G networks, especially in terms of increasing antenna array gain [8].

When studies on Massive MIMO are examined, it is observed that researchers generally focus on basic physical layer features such as obtaining channel state information and preventing intercellular interference [1, 9, 10]. In further studies, the contribution of the Massive MIMO to energy efficiency improvements was mentioned [11-13]. In addition, studies have been conducted on the effect of hardware failures of transceivers on Massive MIMO systems to be less than older systems [14-16]. However, the studies on the programming of the users and the ideal allocation of resources have not been carried out sufficiently. In the study [17], where the first study related to this is available, users are divided into certain groups. The basic reason here is that the limited orthogonal pilot sequences are appropriately shared between users. There are some studies in which this sharing is carried out by various methods [18, 19]. Addressing resource allocation is an important issue to maximize spectrum efficiency. One of the resource allocation studies is from the research [20]. Determining ideal number of users that must be programmed per cell is not an answer to multi-cell systems. Looking at the literature in general, there are some studies on this subject for single cell systems. One of these studies is mentioned in the study [21]. Here, although operations were carried out with a fixed number of users, the ideal value of the number of users in a given antenna number was calculated according to the number of symbols. The main contribution of this study is the answer to the question that what are the ideal antenna number and number of active user's terminals for maximum spectrum efficiency (SE) for MIMO system?

It is stated that studies that use point to point MIMO instead of Massive MIMO show a contradiction between physically realistic models and theoretical models [1]. However, in this study, terminals will be distributed randomly in multi-user MIMO systems. Here, a cellular system of hexagonal cells is considered. TDD scenario and Orthogonal Frequency Division Multiplexing (OFDM) will be taken into account and the number of antennas in the base station is  $M$ , where  $M$  is assumed to be an infinite number. Each base station serves as many terminals as  $K$ . Neither the base station nor the terminals have information about the channel. Therefore, channel state information (CSI) is obtained in a coherence time over a period of time. Orthogonal pilot sequences are assigned to each terminal in a cell. These can be reused in other cells based on pilot reuse factors. The base station calculates its own channel information thanks to the received pilot sequences. The pilot sequences here may have been disrupted by other cells. However, there is no question of sharing information between cells. Linear pre-coder is used in the uplink. In the downlink, the received antenna signals are combined.

## 2. SYSTEM MODEL

The system model considered is a cellular network topology. It is expressed as the number of  $L$  cells. The base station in the cells have  $M$  number of antenna arrays. These base stations are in communication with  $K$  mobile terminals which have single antenna. The active subunits of the terminals are expressed as  $k \in \{1, \dots, K\}$ . The region where user  $k$  is located in any cell  $l$

is considered to be  $\mathbf{z}_{lk} \in R^2$ . The main reason it is expressed in this way is the consideration of the different placements of users. Coherence time  $T_C$  in s and coherence bandwidth  $W_C$  in Hz are important variables in terms of time-frequency sources [21]. Transmission symbols are calculated with  $S = T_C \times W_C$ . It is stated here that  $T_C$  is less than the consistency period of all users and  $W_C$  is less than the consistency bandwidth of all users [22]. The Time-frequency block containing the frame structure is shown as in Figure 1 [21].

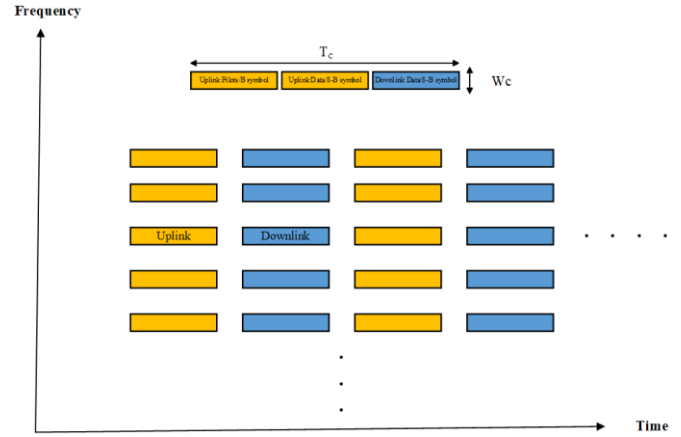


Figure 1. Frame structure

The transmission is divided into blocks of  $S = T_C \times W_C$  symbols.  $B$  symbols,  $S-B$  symbols were used in pilot transmission and data loads, respectively. In Massive MIMO systems, all activities such as uplink pilot transmission, uplink load transport transmission and downlink transmission take place within the coherence time. Some representative values related to this coherence time are included in the study [23].

In a cellular system like in Figure 2,  $\mathbf{h}_{jlk} \in C^{N \times 1}$  shows the channel response. Here,  $l, j$  and  $k$  represent the cell, base station (BS), and user, respectively [22].

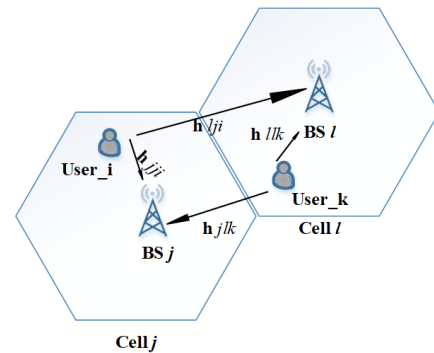


Figure 2. Cellular network

Channel impulse responses are complex Gauss vectors with zero means and are expressed as in Eq. (1).

$$\mathbf{h}_{jlk} \sim CN(\mathbf{0}, d_j(\mathbf{z}_{lk})\mathbf{I}_M) \quad (1)$$

here,  $d_j(\mathbf{z})$  is expressed as channel variance. This variance is the losses between BS  $j$  and any user in the  $\mathbf{z}$  position. Therefore,  $d_j(\mathbf{z}_{lk})$  value is assumed to be known by the base station. However, unlike other studies,  $\mathbf{z}_{lk}$  positions are unknown. The  $M \times M$  matrix is expressed by  $\mathbf{I}_M$  represents complex numbers.

Time Division Duplex (TDD) operation is considered in Massive MIMO system. One of the most important reasons for this is channel reciprocity in TDD [24]. The channel reciprocity is that the inverse channel matrix is equal to the transposition of the forward channel matrix [25]. Thus, the base station is assumed to obtain downlink channel status information by estimating uplink pilot signals. Only uplink channel measurements and up and downlink signals can be processed. Due to these features, TDD scenario is at the forefront in Massive MIMO systems.

## 2.1 Uplink

Downlink channels are estimated with the use of uplink pilots at the base stations. During these operations, all users send uplink data to the base stations. Therefore, the uplink signal received at the base station  $j$  is expressed as in Eq. (2) similar to the work in [10, 26, 27].

$$\mathbf{y}_j = \sum_{l=1}^L \sum_{k=1}^K \sqrt{p_{lk}} \mathbf{h}_{jlk} x_{lk} + \mathbf{n}_j \quad (2)$$

here,  $\mathbf{h}_{jlk}$ ,  $x_{lk}$ ,  $p_{lk}$  and  $\mathbf{n}_j$  represent the channel response, the symbol transmitted by the user  $k$  in cell  $l$ , the uplink transmit power, and the noise vector.

## 2.2 Downlink

The signal received at user  $k$  in cell  $j$  is expressed as in Eq. (3) similar to the works [28, 29].

$$z_{jk} = \sum_{l=1}^L \sum_{m=1}^K \mathbf{h}_{ljk}^T \mathbf{w}_{lm} s_{lm} + \mathbf{n}_j \quad (3)$$

Here,  $s_{lm}$  is the symbol of user  $m$  in cell  $l$ .  $\mathbf{w}_{lm}$  is a precoding vector.  $(\cdot)^T$  represents transpose.

## 3. CHANNEL ESTIMATION

It can achieve various gains with uplink and downlink operations through its numerous antennas in the base station. Because of that, the base station needs to know some information about the channels.  $\sqrt{p_{lk}} \mathbf{h}_{jlk}$  is one of these information which is called channel state information (CSI). Obtaining CSI for multi-cell systems is not easy due to intercellular interference. This interference is referred to as pilot contamination.

It was understood that the pilot contamination problem was caused by using the same orthogonal pilot sequence in all cells [28]. In this study, it is thought that pilots will use different subsets instead of using the same pilot sequences of each cell. As shown in Figure 1, the pilot signals will be considered to be as much as the  $B$  symbol. Pilot signals are represented by the  $\mathbf{v}$  vector. Thus, Eq. (4) was obtained by rewriting Eq. (2).

$$\mathbf{Y}_j = \sum_{l=1}^L \sum_{k=1}^K \sqrt{p_{lk}} \mathbf{h}_{jlk} \mathbf{v}_{ik}^T + \mathbf{N}_j \quad (4)$$

here  $i_{lk}$  is expressed as  $\{1, \dots, B\}$ .

Minimum mean square error (MMSE) is widely used for

channel estimation in Massive MIMO systems [28]. In this study, channel estimation will be performed with this method. Channel estimation with MMSE is considered through Eq. (4). Channel estimation can be obtained as follows, similar to the study [30].

$$\hat{\mathbf{h}}_{jlk}^{MMSE} = \frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} \mathbf{Y}_j(\boldsymbol{\varphi}_j)^{-1} \mathbf{v}_{i_{lk}} \quad (5)$$

here,  $\boldsymbol{\varphi}_j$  is defined as the covariance matrix and is expressed as follows.

$$\boldsymbol{\varphi}_j = \sum_{l=1}^L \sum_{k=1}^K \frac{d_j(\mathbf{z}_{lm})}{d_l(\mathbf{z}_{lm})} \mathbf{v}_{i_{lm}} \mathbf{v}_{i_{lm}}^H + \sigma^2 \mathbf{I}_B \quad (6)$$

here  $(\cdot)^H$  stands for conjugate transpose (Hermitian). Channel estimation error is given as Eq. (7) as in the study [31].

$$\tilde{\mathbf{h}}_{jlk}^{MMSE} = \mathbf{h}_{jlk} - \hat{\mathbf{h}}_{jlk}^{MMSE} \quad (7)$$

The expression of the covariance matrix of the channel estimation error is as follows:

$$\begin{aligned} \mathbf{C}_{jlk}^{MMSE} &= \mathbb{E} \left\{ (\tilde{\mathbf{h}}_{jlk}^{MMSE}) (\tilde{\mathbf{h}}_{jlk}^{MMSE})^H \right\} \\ &= \mathbb{E} \left\{ (\mathbf{h}_{jlk} \right. \\ &\quad \left. - \hat{\mathbf{h}}_{jlk}^{MMSE}) (\mathbf{h}_{jlk} - \hat{\mathbf{h}}_{jlk}^{MMSE})^H \right\} \\ &= \mathbb{E} \left\{ (\mathbf{h}_{jlk} \mathbf{h}_{jlk}^H) \right\} \\ &\quad - \mathbb{E} \left\{ \hat{\mathbf{h}}_{jlk}^{MMSE} (\hat{\mathbf{h}}_{jlk}^{MMSE})^H \right\} \end{aligned} \quad (8)$$

here,  $\mathbb{E} \left\{ (\mathbf{h}_{jlk} \mathbf{h}_{jlk}^H) \right\} = \frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} \mathbf{I}_M$ .

Therefore, the covariance of the channel estimation error is written as

$$\mathbf{C}_{jlk}^{MMSE} = \frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} \left( 1 - \frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} \mathbf{v}_{i_{lm}} (\boldsymbol{\varphi}_j)^{-1} \mathbf{v}_{i_{lm}}^H \right) \mathbf{I}_M. \quad (9)$$

The simplification of this expression is carried out with the help of Ref. [30, 32, 33] as follows:

$$\begin{aligned} &\mathbf{v}_{i_{lm}} (\boldsymbol{\varphi}_j)^{-1} \\ &= \frac{1}{\sum_{l=1}^L \sum_{k=1}^K \frac{d_j(\mathbf{z}_{lm})}{d_l(\mathbf{z}_{lm})} \mathbf{v}_{i_{lm}} \mathbf{v}_{i_{lm}}^H + \sigma^2} \mathbf{v}_{i_{lm}} \end{aligned} \quad (10)$$

$$\begin{aligned} &\mathbf{C}_{jlk}^{MMSE} \\ &= \rho \frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} \left( 1 \right. \\ &\quad \left. - \frac{\frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} B}{\sum_{l=1}^L \sum_{k=1}^K \frac{d_j(\mathbf{z}_{lm})}{d_l(\mathbf{z}_{lm})} \mathbf{v}_{i_{lm}} \mathbf{v}_{i_{lm}}^H + \frac{\rho}{\sigma^2}} \right) \mathbf{I}_M \end{aligned} \quad (11)$$

The covariance given in Eq. (11) shows that the errors depend on the SNR and the user terminals using the same pilot signal. Here, SNR is defined as  $\frac{\rho}{\sigma^2}$ . The expression of users using the same pilot signal means that  $\mathbf{v}_{i_{lm}} \mathbf{v}_{i_{lk}}^H$  is not zero. The

expression  $\frac{d_j(z_{lm})}{d_l(z_{lm})}$  describes the interference received from user  $m$  in cell  $l$  at the base station  $j$ . Considering this interference as a value between 0-1, it is the highest at the cell edges. Considering as cell  $l$ , the further away this cell is from the base station, the lower the interference [34].

#### 4. SPECTRUM EFFICIENCY

In this section, closed form expressions of signal-to-interference-and-noise ratio (SINR) are obtained and a general spectrum efficiency expression is derived accordingly. The point to be considered when derived a general spectrum efficiency expression is capacity. Here, the achievable capacity of spectrum efficiency for user  $k$  in cell  $j$  is expressed as follows, similar to the work [35].

$$R_{jk} = \left(1 - \frac{B}{S}\right) \{\log_2(1 + SINR_{jk})\} [\text{bit/s/Hz}] \quad (12)$$

The two situations to be mentioned in this section are uplink and downlink. Uplink is a state where  $K$  users send signals to BS. With the received signal vector CSI (Channel State Information) information, BS will consistently detect the signals transmitted from  $K$  users. Uplink payload data transmission is when the BS in the  $l$ th cell uses only the received signal and aims only to detect the signals sent by its  $K$  users. Signals from users in other cells are perceived as intracellular interference and ultimately evaluated as additional noise. Downlink refers to the situation where the BS sends a signal to all  $K$  users.

In the MIMO system, BS replicates one data stream per user downlink and receives one stream per user on the uplink. That is, BS antennas use each signal to direct each signal to the desired receiver on the downlink and separate multiple received signals on the uplink. Each consistency range operates in TDD mode and can include both downlink and uplink load transmissions. To enable channel estimation in the BS, symbols in each coherence range are allocated in the uplink transmission of the pilot sequences, while the remaining symbols can be arbitrarily allocated between the uplink and downlink load. To summarize, the channel estimates received for the uplink are also used for downlink linear precoding. Equations to be used in uplink and downlink spectral efficiencies with MIMO optimize downlink power allocation and BS user for MIMO cellular systems.

##### 4.1 Spectrum efficiency for uplink

Channel estimates are carried out with the signals from users are properly detected in BS. While performing these operations, some detection techniques are used in BS. Thanks to these, the signals of the user are separated from the interference from other users. Here, MRC and ZF, previously mentioned, will be emphasized. SINR expressions will be calculated according to MRC and ZF and spectrum efficiency equations will be derived accordingly.

The uplink spectrum efficiency for each cell is expressed as follows.

$$SE_{jk}^{ul} = K \left(1 - \frac{B}{S}\right) \{\log_2(1 + SINR_{jk})\} \left[\frac{\text{bit}}{\text{Hz}} / \text{cell}\right] \quad (13)$$

After the channel estimation is performed, the detection of the transmitted signal in Eq. (2) with a suitable  $\mathbf{v}$  detection vector is performed as follows.

$$\begin{aligned} \hat{x}_{jk} &= \mathbf{v}_{jk}^H \mathbf{y}_j = \sqrt{p_{jk}} \mathbf{v}_{jk}^H \mathbf{h}_{jk} x_{jk} \\ &\quad + \sqrt{p_{jk}} \sum_{m=1, m \neq k}^K \mathbf{v}_{jk}^H \mathbf{h}_{jk} x_{jk} \\ &\quad + \sqrt{p_{lm}} \sum_{l=1, l \neq j}^L \sum_{m=1}^K \mathbf{v}_{jk}^H \mathbf{h}_{lm} x_{lm} \\ &\quad + \mathbf{v}_{jk}^H \mathbf{n}_j \end{aligned} \quad (14)$$

here, the first term represents the desired signal, the second and third terms represent intracellular signals and intercellular interference respectively, the last term is noise. The uplink signal vector specified in Eq. (2) can be rewritten as the uplink symbol received using the inference [36] as follows.

$$\begin{aligned} y_{jk} &= \mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}] \sqrt{p_{jk}} x_{jk} \\ &\quad + (\mathbf{v}_{jk}^H \mathbf{h}_{jk} - \mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}]) \sqrt{p_{jk}} x_{jk} \\ &\quad + \sqrt{p_{jk}} \sum_{m \neq k} \mathbf{v}_{jk}^H \mathbf{h}_{jk} x_{jk} \\ &\quad + \underbrace{\sqrt{p_{lm}} \sum_{(l,m) \neq (j,k)} \mathbf{v}_{jk}^H \mathbf{h}_{lm} x_{lm}}_{(l,m) \neq (j,k)} + \mathbf{v}_{jk}^H \mathbf{n}_j \end{aligned} \quad (15)$$

here,  $\mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}]^2$  channel gain. Variance of this channel gain is calculated as  $\text{var}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}] = \mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk} - \mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}]]^2 = \mathbb{E}[|\mathbf{v}_{jk}^H \mathbf{h}_{jk}|^2] - |\mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}]|^2$ . Finally, interference power is in the form of  $\sum_{(l,m) \neq (j,k)} \mathbb{E}[|\mathbf{v}_{jk}^H \mathbf{h}_{lm}|^2] = \sum_{l,k} \mathbb{E}[|\mathbf{v}_{jk}^H \mathbf{h}_{lm}|^2] - \mathbb{E}[|\mathbf{v}_{jk}^H \mathbf{h}_{jk}|^2]$ . It is also assumed that all these values are known. SINR statement It is known that receive signal power is expressed as the ratio of interference power to noise total [37]. So the SINR statement is written as follows:

$$SINR_{jk}^{ul} = \frac{\sqrt{p_{jk}} |\mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}]|^2}{\sum_{j=1}^L \sum_{m=1}^K \sqrt{p_{lm}} \mathbb{E}[|\mathbf{v}_{jk}^H \mathbf{h}_{jm}|^2] - (\sqrt{p_{jk}} |\mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}]|^2 + \sigma^2 \mathbb{E}[\|\mathbf{v}_{jk}\|^2])} \quad (16)$$

here  $\sqrt{p_{jk}} |\mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}]|^2$  is power of desired signal,  $\sum_{j=1}^L \sum_{m=1}^K \sqrt{p_{lm}} \mathbb{E}[|\mathbf{v}_{jk}^H \mathbf{h}_{jm}|^2]$  represents multiple user interference,  $\sqrt{p_{jk}} |\mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}]|^2$  intracellular interference and finally  $\sigma^2 \mathbb{E}[\|\mathbf{v}_{jk}\|^2]$  is the total noise power. Therefore, it is understood that the spectrum efficiency given in Eq. (13) depends on detection techniques.

##### 4.1.1 Maximum ratio combining (MRC)

MRC is defined as  $\mathbf{V} = \hat{\mathbf{H}}$  as stated in the study [11]. Here is the  $\mathbf{V}$  detector matrix and  $\hat{\mathbf{H}}$  is the matrix of the channel estimation. Given this definition, Eq. (16) is rewritten as Eq. (17). As follows

$$SINR_{jk}^{ul} = \frac{\sqrt{p_{jk}} |\mathbb{E}[\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{jk}]|^2}{\sum_{j=1}^L \sum_{m=1}^K \sqrt{p_{lm}} \mathbb{E} [|\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{jm}|^2] - (\sqrt{p_{jk}} |\mathbb{E}[\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{jk}]|^2 + \sigma^2 \mathbb{E} [\|\hat{\mathbf{h}}_{jk}\|^2])} \quad (17)$$

The calculation of the expected values in the above statement is as follows.  $V\{\cdot\}$  stands for variance.

$$\sqrt{p_{jk}} |\mathbb{E}[\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{jk}]|^2 = M^2 \sqrt{p_{jk}} (V\{\hat{\mathbf{h}}_{jkm}\})^2 \quad (18)$$

$$\begin{aligned} & \sum_{j=1}^L \sum_{m=1}^K \sqrt{p_{lm}} \mathbb{E} [|\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{jm}|^2] \\ & - (\sqrt{p_{jk}} |\mathbb{E}[\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{jk}]|^2 \\ & + \sigma^2 \mathbb{E} [\|\hat{\mathbf{h}}_{jk}\|^2]) \\ & = \sum_j \sqrt{p_{lk}} \mathbb{E} [|\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{lk}|^2] \\ & + \sum_j \sum_{m=1, m \neq k} \sqrt{p_{lm}} \mathbb{E} [|\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{lm}|^2] \\ & + \sum_{m=k} \sqrt{p_{lm}} \mathbb{E} [|\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{lm}|^2] \\ & - \sqrt{p_{jk}} |\mathbb{E}[\hat{\mathbf{h}}_{jk}^H \mathbf{h}_{jk}]|^2 \\ & + \sigma^2 \mathbb{E} [\|\hat{\mathbf{h}}_{jk}\|^2] \\ & = M^2 V\{\hat{\mathbf{h}}_{jkm}\} \sum_j \sqrt{p_{lm}} V\{\hat{\mathbf{h}}_{lkm}\} \\ & + MV\{\hat{\mathbf{h}}_{jkm}\} \sum_{m=1}^j \sqrt{p_{lm}} V\{\hat{\mathbf{h}}_{lkm}\} \\ & + MV\{\hat{\mathbf{h}}_{jkm}\} \sum_{m=1}^j \sqrt{p_{lm}} V\{\hat{\mathbf{h}}_{lkm}\} \\ & + MV\{\hat{\mathbf{h}}_{jkm}\} \sigma^2 \end{aligned} \quad (19)$$

Therefore, the SINR statement can be rewritten by simplifying it as in Eq. (20) with MRC detection.

$$SINR_{jk}^{MRC,ul} = \frac{M \sqrt{p_{jk}} V\{\hat{\mathbf{h}}_{jkm}\}}{\sum_j \sqrt{p_{lm}} V\{\hat{\mathbf{h}}_{lkm}\} + \sum_{m=1} \sqrt{p_{lm}} V\{\hat{\mathbf{h}}_{lkm}\} + \sigma^2} \quad (20)$$

#### 4.1.2 Zero forcing (ZF)

ZF is defined as  $\mathbf{V} = \hat{\mathbf{H}}(\hat{\mathbf{H}}^H \hat{\mathbf{H}})^{-1}$  as stated in the study [11]. The main purpose here is to try to minimize intracellular interferences.

In the case of ZF,  $\mathbb{E}[\mathbf{v}_{jk}^H \mathbf{h}_{jk}] = 1$  due to its reverse channel structure. The estimation of the noise term is as follows.

$$\sigma^2 \mathbb{E} [\|\mathbf{v}_{jk}\|^2] = \frac{\sigma^2}{(M-K)V\{\hat{\mathbf{h}}_{jkm}\}} \quad (21)$$

Following these estimations, the SINR statement for ZF is derived as in Eq. (22).

$$SINR_{jk}^{ZF,ul} = \frac{\sqrt{p_{jk}}}{\sum_{j=1}^L \sum_{m=1}^K \sqrt{p_{lm}} \mathbb{E} [|\hat{\mathbf{v}}_{jk}^H \mathbf{h}_{jm}|^2] - \sqrt{p_{jk}} + \frac{\sigma^2}{(M-K)V\{\hat{\mathbf{h}}_{jkm}\}}} \quad (22)$$

## 4.2 Spectrum efficiency for downlink

In this section, the statements obtained from the estimations in uplink are taken considering the Massive MIMO downlink system. The signal vector transmitted in  $\mathbf{x}_l \in \mathbb{C}^M$  BS  $l$  can be calculated with the  $\mathbf{w}_{lm} \in \mathbb{C}^M$  precoding vector as follows.

$$\mathbf{x}_l = \sum_{m=1}^K \sqrt{p_{lm}} \mathbf{w}_{lm} s_{lm} \quad (23)$$

here,  $s_{l,m}$  denotes the load symbol for user  $m$  in cell  $l$  and  $\mathbb{E}[|s_{lm}|^2] = 1$ .  $p_{lm}$  is the transmitter power of user  $m$ . The  $y_{jm}$  signal received in user  $m$  in the cell  $j$  is expressed as follows.

$$y_{jm} = \sum_{l=1}^L \mathbf{h}_{ljm}^H \mathbf{x}_l + n_{jm} \quad (24)$$

here,  $\mathbf{h}_{ljm}$  is the channel vector,  $n_{jm}$  is the noise vector. Eq. (24) is developed as follows using Eq. (23).

$$y_{jm} = \sqrt{p_{lk}} \sum_{l=1}^L \sum_{k=1}^K \mathbf{h}_{ljm}^H \mathbf{w}_{lk} s_{lk} + n_{jm} \quad (25)$$

Similar to the study [38], this statement is rewritten as follows.

$$\begin{aligned} y_{jm} & = \sqrt{p_{jm}} s_{jm} \mathbb{E}[\mathbf{h}_{jjm}^H \mathbf{w}_{jm}] \\ & + (\mathbf{h}_{jjm}^H \mathbf{w}_{jm} \\ & - \mathbb{E}[\mathbf{h}_{jjm}^H \mathbf{w}_{jm}]) \sqrt{p_{jm}} s_{jm} \\ & + \sqrt{p_{lk}} \sum_{l,k \neq j,m} \mathbf{h}_{ljm}^H \mathbf{w}_{lk} s_{lk} + n_{jm} \end{aligned} \quad (26)$$

here, it is assumed that channel gain  $\mathbb{E} [|\mathbf{h}_{jjm}^H \mathbf{w}_{jm}|^2]$ , its variance  $\text{var}[\mathbf{h}_{jjm}^H \mathbf{w}_{jm}]$  and interference power  $\sum_{l,k \neq j,m} \mathbb{E} [|\mathbf{h}_{ljm}^H \mathbf{w}_{lk}|^2]$  are known. The SINR definition is specified in the study [39] as follows. Based on this definition, it is obtained as in SINR Eq. (29) for downlink. as follows.

$$SINR = \frac{\mathbb{E}[|Useful Signal|^2]}{\text{var}[Pilot Cont. Interf.] + \text{var}[Interf. plus Noise Terms]} \quad (27)$$

$$SINR_{jm} = \frac{\sqrt{p_{jm}} |\mathbb{E}[\mathbf{h}_{jjm}^H \mathbf{w}_{jm}]|^2}{\sqrt{p_{jm}} \text{var}[\mathbf{h}_{jjm}^H \mathbf{w}_{jm}] + \sqrt{p_{lk}} \sum_{l,k \neq j,m} \mathbb{E} [|\mathbf{h}_{ljm}^H \mathbf{w}_{lk}|^2] + \sigma^2} \quad (28)$$

In here,  $\text{var}[\mathbf{h}_{jjm}^H \mathbf{w}_{jm}] = \mathbb{E} [|\mathbf{h}_{jjm}^H \mathbf{w}_{jm}|^2] - |\mathbb{E}[\mathbf{h}_{jjm}^H \mathbf{w}_{jm}]|^2$ ,  $\sum_{l,k \neq j,m} \mathbb{E} [|\mathbf{h}_{ljm}^H \mathbf{w}_{lk}|^2] = \sum_{l,k} \mathbb{E} [|\mathbf{h}_{ljm}^H \mathbf{w}_{lk}|^2] - \mathbb{E} [|\mathbf{h}_{jjm}^H \mathbf{w}_{jm}|^2]$ . Therefore SINR is;

$$SINR_{jm} = \frac{\sqrt{p_{jm}} |\mathbb{E}[\mathbf{h}_{jjm}^H \mathbf{w}_{jm}]|^2}{\sqrt{p_{jm}} \sum_{l,k} \mathbb{E} [|\mathbf{h}_{ljm}^H \mathbf{w}_{lk}|^2] - \sqrt{p_{lk}} |\mathbb{E}[\mathbf{h}_{jjm}^H \mathbf{w}_{jm}]|^2 + \sigma^2} \quad (29)$$

It is seen that the available rate depends on the channel vector and the precoding vector. Therefore, the precoding vector should be chosen well for high intracellular and intercellular interference with low signal gains.

## 5. ACHIEVING SPECTRUM EFFICIENCY EQUATIONS

In this section, the derivation of spectrum efficiency equations is discussed. Contrary to the general belief in the literature, closed form expressions are derived from different distributions of users. Since different distributions of users are taken into account, three different interference levels that are taken into account when deriving closed form statements of spectrum efficiency are described as follows. These environments generally emerge with the classical pathloss model method. Studies on pathloss models are also available in the researches [35, 38]. In addition, the statements of the processing schemes for spectrum efficiency equations that are considered according to these three different interference levels are given below.

### 5.1 First environment

This situation actually symbolizes the concept of an ideal environment. So, given a particular cell in a multi-cell system, users in other cells are far from BS. Therefore, there is no interference. But this situation is almost impossible. But it is theoretically stated in terms of comparison.

### 5.2 Second environment

It is an interference level where the averages of the users in all cells are taken into account, assuming that they have similar, close positions. It is estimated that this will be the most suitable level for the system under consideration.

### 5.3 Third environment

This situation represents the worst attempt to encounter similar to the logic in the definition of a very good environment. It symbolizes the situations in which users in other cells are closest to BS. It is aimed to make an objective comparison by showing the most ideal and worst case.

An interference level equation must be created in order to express the above interference levels that will occur due to a multi-cell structure. The point to be considered here is the relationship with the base stations in other cells. The equation below indicates the level of interference.

$$\alpha_{total} = 1 + parts \left( \sum \alpha_{jl\_case} - 1 \right) \quad (30)$$

here, for any user in cell  $l$ , the ratio of channel variance in base station  $j$  to channel variance in base station  $l$  is considered as  $\alpha_{jl}$ . Therefore, this statement is expressed as follows.

$$\alpha_{jl\_case} = \mathbb{E} \left\{ \left( \frac{|d_j(\mathbf{z})|}{|d_l(\mathbf{z}) + \mathbf{b}_j|} \right)^\kappa \right\} \quad (31)$$

here, the positions of users with  $\mathbf{z}$ , and the positions of base stations are represented with  $\mathbf{b}$ . Detailed explanations regarding the determination of these locations can be found in

[40, 41].  $\kappa$  is pathloss exponent. The value of  $\alpha_{jl}$  varies according to the environments mentioned above. The basic reason in Eq. (30); since one center cell is accepted, 1 is subtracted from the sum of each element in the interference vector and then multiplied by the number of cells considered, that is parts. The expression -1 here indicates that the central cell has been removed. Then this center cell was added as +1. In the assumption below, the estimation of the interference level for the first, second and third environment conditions is specified.

$$\alpha_{jl\_FirstEnv.} = \min \mathbb{E} \left\{ \left( \frac{|d_j(\mathbf{z})|}{|d_l(\mathbf{z}) + \mathbf{b}_j|} \right)^\kappa \right\} \quad (32)$$

$$\alpha_{jl\_SecondEnv.} = \text{mean} \mathbb{E} \left\{ \left( \frac{|d_j(\mathbf{z})|}{|d_l(\mathbf{z}) + \mathbf{b}_j|} \right)^\kappa \right\} \quad (33)$$

$$\alpha_{jl\_ThirdEnv.} = \max \mathbb{E} \left\{ \left( \frac{|d_j(\mathbf{z})|}{|d_l(\mathbf{z}) + \mathbf{b}_j|} \right)^\kappa \right\} \quad (34)$$

As can be seen from these equations, considering the particular cell, the absence of any interference (less than negligible) for the first environment where the users in other cells are far from BS is determined by finding the minimum value of the expression specified in Eq. (31). Similarly, for the third environment, that is, when the interferences are too high, the maximum value of the expression specified in Eq. (31) is taken. The average value of this expression is used for the second environment. Thus, SE equations calculated according to SINR values in the previous section will be rewritten according to these interference levels considering MRC and ZF. While these processes are being carried out, the uplink system has been considered. By making similar approaches, SE expressions can be performed for downlink.

The SINR expression given in the equation Eq. (16) for the uplink system is rearranged taking into account the reasons stated in the study [15] and the status of the above mentioned interference levels with MRC and ZF approaches [11] as follows.

$$\begin{aligned} SINR_{jk}^{MRC,ul} &= \frac{B}{\sum_l \left( \alpha_{jl\_case}^{(2)} + \frac{(\alpha_{jl\_case}^{(2)} - (\alpha_{jl\_case}^{(1)})^2)}{M} \right) + \frac{(\sum_l \alpha_{jl\_case}^{(1)} K + \frac{\sigma^2}{\rho}) (\sum_l \alpha_{jl\_case}^{(1)} + \frac{\sigma^2}{B\rho})}{M}} \end{aligned} \quad (35)$$

$$\begin{aligned} SINR_{jk}^{ZF,ul} &= \frac{B}{\left( \alpha_{jl\_case}^{(2)} + \frac{(\alpha_{jl\_case}^{(2)} - (\alpha_{jl\_case}^{(1)})^2)}{M-K} \right) + \frac{\sum_l \left( \sum_l \alpha_{jl\_case}^{(1)} K \left( 1 - \frac{\alpha_{jl\_case}^{(1)}}{\sum_l \alpha_{jl\_case}^{(1)} + \frac{\sigma^2}{B\rho}} \right) + \frac{\sigma^2}{\rho} \right) (\sum_l \alpha_{jl\_case}^{(1)} + \frac{\sigma^2}{B\rho})}{M}} \end{aligned} \quad (36)$$

These statements are substituted in the spectrum efficiency equation given below.

$$SE_{jk}^{ul} = K \left(1 - \frac{B}{S}\right) \left\{ \log_2(1 + SINR_{jk}) \right\} \text{ [bit/s/Hz /cell]} \quad (37)$$

#### 5.4 Determining user locations

In this study, uniform distribution is thought for user terminals in a cell as in Figure 3. The radius of the cell is determined as  $r_{max}$  and the minimum distance between users and BS is  $d_{min}$ . The location of any  $k$  user is shown as  $x_k \in \mathbb{R}^2$ . This is determined by the BS, which is supposed to be at the center. User locations are selected from the distribution of users, expressed as  $f(x)$ , as in the study [42].

$$f(x) = \begin{cases} \frac{1}{\pi(r_{max}^2 - d_{min}^2)}, & d_{min} \leq \|x\| \leq r_{max} \\ 0, & otherwise \end{cases} \quad (38)$$

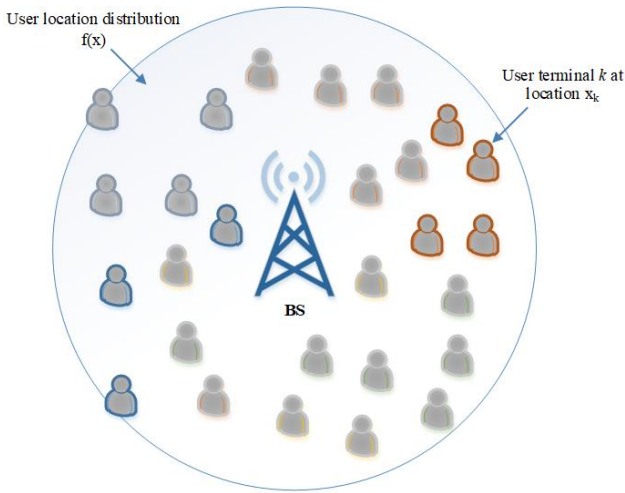


Figure 3. User locations distribution

#### 5.5 Upper limit of spectrum efficiency

Since the first studies that Massive MIMO was put forward, its asymptotic analysis has been carried out continuously. An unlimited number of antenna arrays are mentioned in Marzetta's work [43]. In subsequent studies, a limited number of antennas was considered. However, upper limit analyzes are also performed, which include the situation where the number of antennas in the base station  $M$  goes to infinity.

The main purpose in upper limit analysis is not to examine their behavior, but to have knowledge about performance evaluations in practice. Therefore, an upper limit analysis was carried out in this study. It is aimed to have information about the results to be obtained for the situation where  $M$  goes to infinity. Because this result is an indicator that can never be achieved. But how close we come to this goal matters.

When Eq. (35) ve Eq. (36) are examined, it is seen that the SINR value must be maximized in order to perform the upper limit analysis. Since  $M$  is assumed to go to infinity, SINR in both equations will express the same limit. This expression is given as Eq. (39).

The spectrum efficiency equations, the above analyzes of which are obtained for MRC and ZF, can be performed for a certain number of antennas and number of user terminals. This section shows that the number of antennas in the base station, that is to say  $M$ , goes to infinity, is the upper limit for spectrum

efficiency. In the SINR in Eq. (37), the spectrum efficiency expression is rewritten as follows if  $M$  goes to infinity.

$$SE_{jk} = K \left(1 - \frac{B}{S}\right) \left\{ \log_2 \left(1 + \frac{B}{\sum_l (\alpha_{jl}^{(2)})}\right) \right\} \text{ [bit/s/Hz/cell]} \quad (39)$$

The result obtained here is one of the main contributions of this study. Because, thanks to this upper limit equation, the ideal number of users will be realized in the next section. As a matter of fact, the ideal number of users in the specified system model will be determined when the number of antennas in the base station goes to infinity. In the following section, necessary expressions and results are shown to determine the ideal number of users. transpose.

### 6. NUMERICAL RESULTS FOR DETERMINATION OF IDEAL USERS

In this section, simulation of the spectrum efficiency equation analyzed was performed. Here, MRC and ZF linear processing states are taken into account. Basically, considering all the values for the  $M$  antenna number, the ideal user number  $U$  was tried to be found. In addition, the pilot reuse factor  $f$  was taken into consideration while performing these operations, thus maximizing the spectrum efficiency. Table 1 lists some of the parameters that were used in the simulation. These some parameters have been tested with Matlab and Monte-Carlo simulation. There are various codes related to these. Some of them are mentioned in [4]. These parameters are generally used in Massive MIMO systems in studies such as [4, 41, 42].

Table 1. Some parameters used in simulation

| Parameter                     | Value   |
|-------------------------------|---------|
| Coherence Block Length, $S$   | 400     |
| Coherence Time                | 2 ms    |
| Coherence Bandwidth           | 200 kHz |
| SINR, $\frac{\sigma^2}{\rho}$ | 5 dB    |
| Path Loss Exponent, $\kappa$  | 3.8     |
| Parts (Number of Cells)       | 6       |

For some positive integers  $f$  (eg 1, 3, 4, 7, ...), which are called pilot reuse factors,  $B = fK$  equation is accepted [44]. When  $B = fU$  is replaced with this equation Eq. (39), the following statement is obtained.

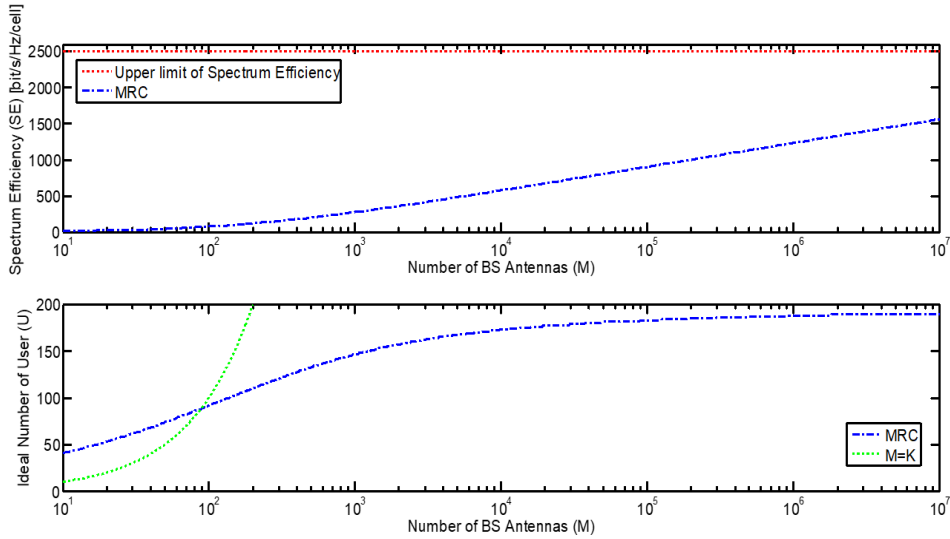
$$SE_{jk} = K \left(1 - \frac{fU}{S}\right) \left\{ \log_2 \left(1 + \frac{fU}{\sum_l (\alpha_{jl}^{(2)})}\right) \right\} \text{ [bit/s /Hz/cell]} \quad (40)$$

If analytical operations are carried out to find the point at which the spectrum efficiency will be maximum, the ideal number of users will be  $U = \frac{S}{2f}$ . Therefore, it will be understood for which pilot reuse factor the spectrum efficiency in the ideal number of users to be achieved is realized. This point is another major contribution of this study.

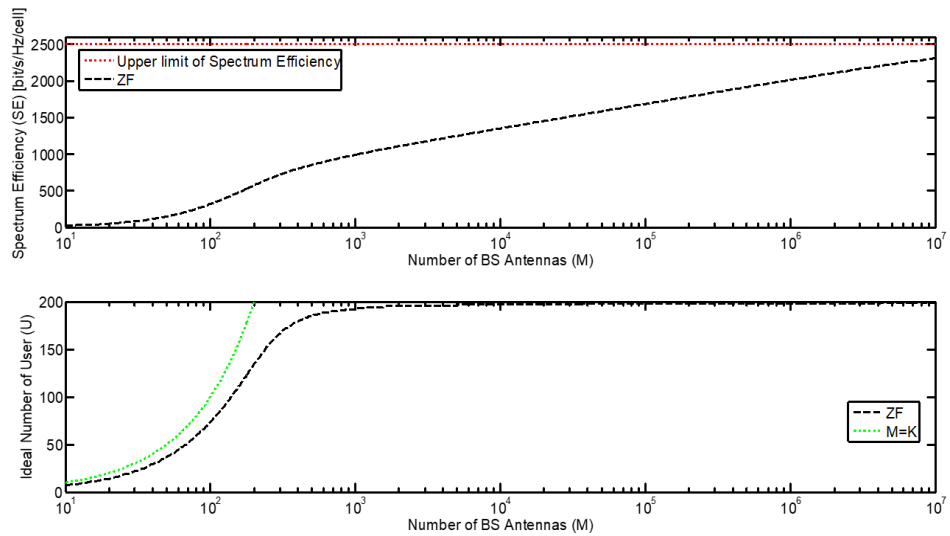
The figures below show the results obtained according to the situations determined as first environment, second environment and third environment respectively.

Figure 4 and Figure 5 show the spectrum efficiency and corresponding ideal user numbers under MRC and ZF for the first environment. As can be seen from the figures, spectrum efficiencies are similar for MRC and ZF in smaller ranges of antenna numbers in the base station. But it requires a different

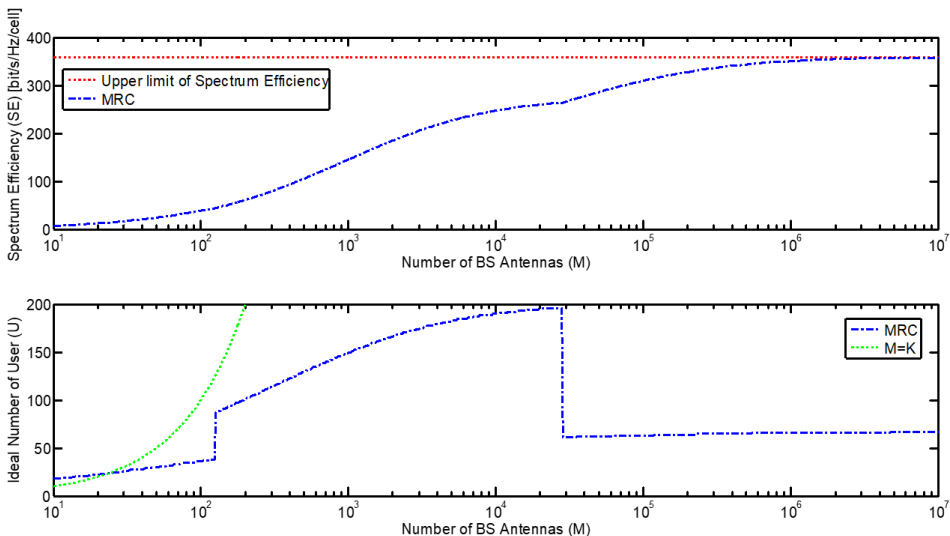
number of user terminals. Spectrum efficiency is much better than others, due to the model of the first environment. Because it is considered as the maximum point that can be obtained. It is an important comparative inference.



**Figure 4.** Ideal number of users with maximizing spectrum efficiency for MRC (for first environment)

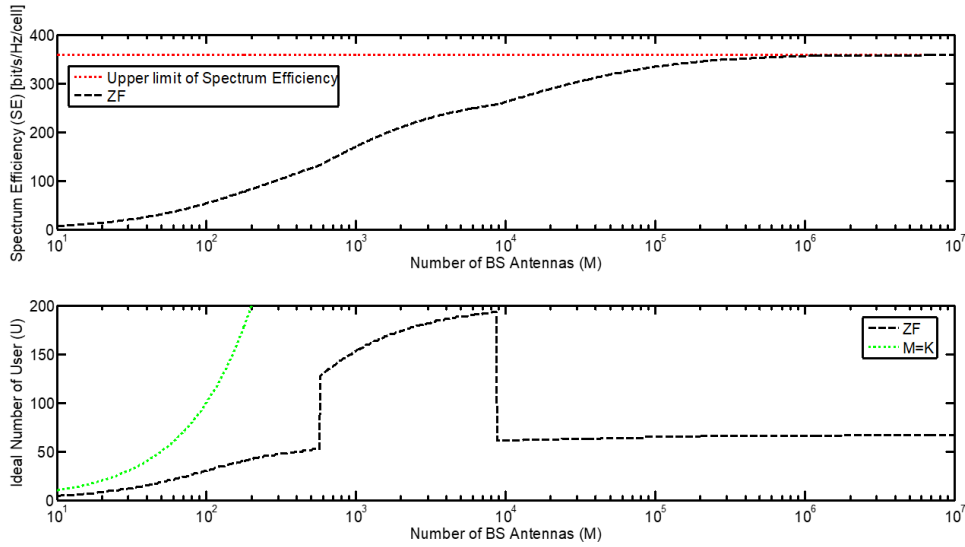


**Figure 5.** Ideal number of users with maximizing spectrum efficiency for ZF (for first environment)

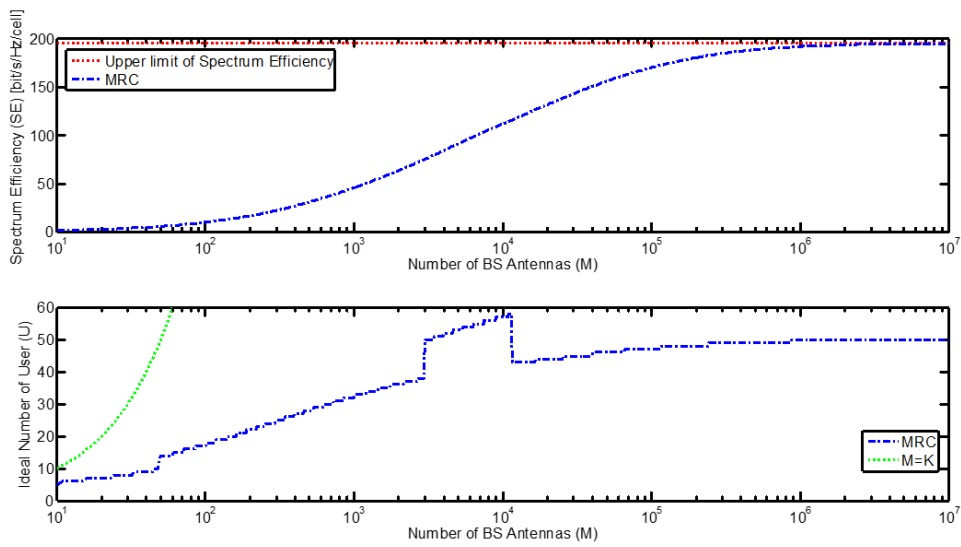


**Figure 6.** Ideal number of users with maximizing spectrum efficiency for MRC (for second environment)

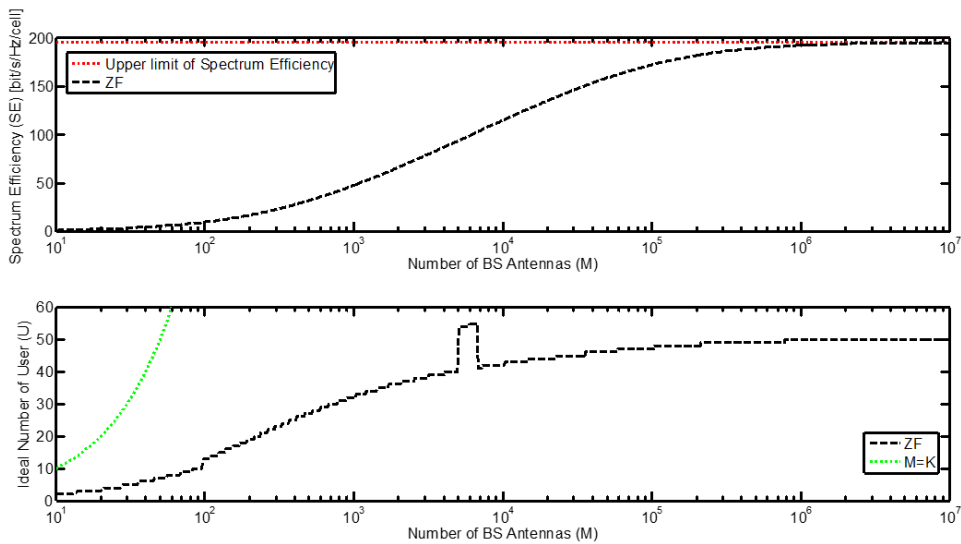




**Figure 7.** Ideal number of users with maximizing spectrum efficiency for ZF (for second environment)



**Figure 8.** Ideal number of users with maximizing spectrum efficiency for MRC (for third environment)



**Figure 9.** Ideal number of users with maximizing spectrum efficiency for ZF (for third environment)

Figure 6 and Figure 7 were performed under MRC and ZF for second environment. This is an environment that includes

situations that may be encountered in practice. It is modeled accordingly and its expressions are derived. Therefore,

different results are seen depending on the number of antennas in base stations. Some jumping points are shown in the figures. These points show the most suitable pilot reuse factor selection of the system. In that, according to the formula  $U = \frac{S}{2f}$ , it is the determination that the appropriate pilot reuse factor  $f$  is used according to the ideal number of users.

In this environment, there are both intracellular and intercellular interferences. It is a situation that can be encountered practically. MRC and ZF are important in increasing spectrum efficiency by eliminating these interferences. In addition, different new processing schemes may be tried in the future.

Spectrum efficiency results are not the point here. The basic expression emphasized is the determination of the ideal number of user terminals in obtaining spectrum efficiency results and the selection of the pilot reuse factor used.

Figure 8 and Figure 9 are performed under MRC and ZF for the third environment. The third environment is the worst of all in terms of interferences. When evaluated in terms of pilot reuse factor, it is observed that the value of  $f$  is needed to be at a high level in the worst environment. This environment is a situation with the highest levels of interference, which is not likely to be encountered in practice. However, the first environment and third environment situation is an important result in terms of evaluating the second environment situation. These results allow comparisons to be made more efficiently.

The figures above show the spectrum efficiency is maximized for all three cases and the result of the corresponding ideal number of users. While the upper graphic in the figures shows the spectrum efficiency in response to the number of antennas in the base station, the lower graphic shows the corresponding ideal number of users. The results were carried out with MRC and ZF in each case.

As can be seen from the figures, when spectrum efficiencies are examined, it is understood that the first environment is very good and the third environment is very bad. Indeed, this confirms the expressions presented as Eq. (30) and Eq. (31). Also, for the first case, it is understood that ZF is better than MRC in terms of spectrum efficiency. Because while defining the first case, it is thought to be an ideal environment without intercellular interference. But there is intracellular interference. Thus, intracellular interference reduction, which is one of the aims of ZF, is achieved and a much better spectrum efficiency is provided compared to MRC.

If we recall Eq. (40), which is the main contribution of the study, considering the number of antennas in the base station as infinite, the ideal number of users can be expressed as  $U = \frac{S}{2f}$ , the ideal number of users for first environment is  $U = 200$ ,  $f = 1$ ;  $U = 67$ ,  $f = 3$  for the second environment and  $U = 50$ ,  $f = 4$  for the third environment.

From the results, an improvement of about 50-100 times was observed in spectrum efficiency according to the requirement of 2.25 bit/s/Hz/cell of IMT-Advanced. In addition, the results we obtained seem to be more efficient in terms of ideal number of users-spectrum efficiency metric than the study [11].

The main purpose in this study is not to explain how spectrum efficiency is achieved. The main goal is spectrum efficiency is achieved in which ideal number of users  $U$  and in which pilot reuse factor. One of the desired situations here is to achieve the highest value of  $U$ . The jumping points seen when the figures are examined show the pilot reuse factors that MRC and ZF have changed to ensure the ideal number of users

and to maximize spectrum efficiency.

## 7. CONCLUSIONS

This study investigated the answer to what the ideal number of users should be, along with the number of antennas in the base station to maximize spectrum efficiency in multi-user Massive MIMO systems. Spectrum efficiency equations depend on the locations of the user terminals. Because this situation actually makes it difficult to determine the ideal number of users. Therefore, in this study, new spectrum efficiency expressions were created in which we can determine the ideal number of users according to the status of the user locations. These statements, which will help determine the ideal number of users while maximizing spectrum efficiency, have been verified by simulation under MRC and ZF processing.

Analytically obtained results were performed under MRC and ZF processes. The analysis brings the ideal number of users closer to  $\frac{S}{2f}$ , regardless of linear processing. However, the effect of the pilot reuse factor, which is changed momentarily to maximize spectrum efficiency, should not be forgotten. In addition, this study has been examined except for the effect of any hardware failure or other parameters. Among the future studies, it is considered to examine the effects such as hardware failure.

## REFERENCES

- [1] Bhang, D., Deth, C. (2020). Performance optimization of LS/LMMSE using swarm intelligence in 3D MIMO-OFDM systems. *Traitement du Signal*, 37(1): 107-112. <https://doi.org/10.18280/ts.370114>
- [2] Marzetta, T.L. (2006). How much training is required for multiuser MIMO? 2006 Fortieth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, pp. 359-363. <https://doi.org/10.1109/ACSSC.2006.354768>
- [3] Marzetta, T.L. (1999). BLAST training: Estimating channel characteristics for high capacity space-time wireless. In *Proceedings of the Annual Allerton Conference on Communication Control and Computing*, 37: 958-966.
- [4] Björnson, E., Larsson, E.G., Debbah, M. (2016). Maximal spectral efficiency: How many users and pilots should be allocated? *IEEE Transactions on Wireless Communications*, 15(2): 1293-1308. <https://doi.org/10.1109/TWC.2015.2488634>
- [5] Tse, D., Viswanath, P. (2005). *Fundamentals of wireless communication*. Cambridge University Press.
- [6] Rusek, F., Persson, D., Lau, B.K., Larsson, E.G., Marzetta, T.L., Edfors, O., Tufvesson, F. (2013). Scaling Up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Processing Magazine*, 30(1): 40-60. <https://doi.org/10.1109/MSP.2011.2178495>
- [7] Robin, C., Robert, A. (2020). Massive MIMO systems for 5G and beyond networks—overview, recent trends, challenges, and future research direction. *Sensors*, 20(10): 2753. <https://doi.org/10.3390/s20102753>
- [8] Akyildiz, I.F., Kak, A., Nie, S. (2020). 6G and Beyond: The Future of Wireless Communications Systems. *IEEE Access*, 8: 133995-134030.

- <https://doi.org/10.1109/ACCESS.2020.3010896>
- [9] Jose, J., Ashikhmin, A., Marzetta, T.L., Vishwanath, S. (2011). Pilot contamination and precoding in multi-cell TDD systems. *IEEE Transactions on Wireless Communications*, 10(8): 2640-2651. <https://doi.org/10.1109/TWC.2011.060711.101155>
- [10] Hoydis, J., Brink, S.T., Debbah, M. (2013). Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? *IEEE Journal on Selected Areas in Communications*, 31(2): 160-171. <https://doi.org/10.1109/JSAC.2013.130205>
- [11] Ngo, H.Q., Larsson, E.G., Marzetta, T.L. (2013). Energy and spectral efficiency of very large multiuser MIMO systems. *IEEE Transactions on Communications*, 61(4): 1436-1449. <https://doi.org/10.1109/TCOMM.2013.020413.110848>
- [12] Ha, D., Lee, K., Kang, J. (2013). Energy efficiency analysis with circuit power consumption in massive MIMO systems. 2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), London, UK, pp. 938-942. <https://doi.org/10.1109/PIMRC.2013.6666272>
- [13] Yang, H., Marzetta, T.L. (2013). Total energy efficiency of cellular large scale antenna system multiple access mobile networks. 2013 IEEE Online Conference on Green Communications (OnlineGreenComm), Piscataway, NJ, USA, pp. 27-32. <https://doi.org/10.1109/OnlineGreenCom.2013.6731024>
- [14] Björnson, E., Hoydis, J., Kountouris, M., Debbah, M. (2014). Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits. *IEEE Transactions on Information Theory*, 60(11): 7112-7139. <https://doi.org/10.1109/TIT.2014.2354403>
- [15] Björnson, E., Matthaiou, M., Debbah, M. (2015). Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design. *IEEE Transactions on Wireless Communications*, 14(8): 4353-4368. <https://doi.org/10.1109/TWC.2015.2420095>
- [16] Pitarokoilis, A., Mohammed, S.K., Larsson, E.G. (2015). Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise. *IEEE Transactions on Wireless Communications*, 14(2): 711-723. <https://doi.org/10.1109/TWC.2014.2359018>
- [17] Huh, H., Caire, G., Papadopoulos, H.C., Ramprasad, S.A. (2012). Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas. *IEEE Transactions on Wireless Communications*, 11(9): 3226-3239. <https://doi.org/10.1109/TWC.2012.070912.11138>
- [18] Li, Y., Nam, Y.H., Ng, B.L., Zhang, J. (2012). A non-asymptotic throughput for massive MIMO cellular uplink with pilot reuse. 2012 IEEE Global Communications Conference (GLOBECOM), Anaheim, CA, USA, pp. 4500-4504. <https://doi.org/10.1109/GLOCOM.2012.6503827>
- [19] Li, M., Jin, S., Gao, X. (2013). Spatial orthogonality-based pilot reuse for multi-cell massive MIMO transmission. In 2013 International Conference on Wireless Communications and Signal Processing, pp. 1-6.
- [20] Alemayehu, T.S., Kim, J.H., Yoon, W. (2017). Full-duplex distributed massive MIMO system with optimal beamforming and resource allocation for WPCN. *Journal of Communications Technology and Electronics*, 62: 1383-1387. <https://doi.org/10.1134/S1064226917140030>
- [21] Raghavan, V., Hariharan, G., Sayeed, A.M. (2007). Capacity of sparse multipath channels in the ultra-wideband regime. *IEEE Journal of Selected Topics in Signal Processing*, 1(3): 357-371. <https://doi.org/10.1109/JSTSP.2007.906666>
- [22] Alwakeel, A.S., Mehana, A.M.H. (2017). Achievable rates in uplink massive MIMO systems with pilot hopping. *IEEE Transactions on Communications*, 65(10): 4232-4246. <https://doi.org/10.1109/TCOMM.2017.2719034>
- [23] Karlsson, M., Larsson, E.G. (2014). On the operation of massive MIMO with and without transmitter CSI. 2014 IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Toronto, ON, Canada, pp. 1-5. <https://doi.org/10.1109/SPAWC.2014.6941305>
- [24] Khalilsarai, M.B., Haghightashoar, S., Yi, X., Caire, G. (2018). FDD massive MIMO: Efficient downlink probing and uplink feedback via Active channel sparsification. 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, pp. 1-6. <https://doi.org/10.1109/ICC.2018.8422262>
- [25] Sheikh, T.A., Bora, J., Hussain, A. (2019). Performance analysis of massive multi-input and multi-output with imperfect channel state information. *Traitement du Signal*, 36(4): 361-368. <https://doi.org/10.18280/ts.360409>
- [26] Wu, Y., Liu, T., Cao, M., Li, L., Xu, W. (2018). Pilot contamination reduction in massive MIMO systems based on pilot scheduling. *EURASIP Journal on Wireless Communications and Networking*, 2018(1): 1-9. <https://doi.org/10.1186/s13638-018-1029-1>
- [27] Upadhyay, K., Vorobyov, S.A., Vehkaperä, M. (2016). Downlink performance of superimposed pilots in massive MIMO systems in the presence of pilot contamination. 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Washington, DC, pp. 665-669. <https://doi.org/10.1109/GlobalSIP.2016.7905925>
- [28] Lu, L., Li, G.Y., Swindlehurst, A.L., Ashikhmin, A., Zhang, R. (2014). An overview of massive MIMO: Benefits and challenges. *IEEE Journal of Selected Topics in Signal Processing*, 8(5): 742-758. <https://doi.org/10.1109/JSTSP.2014.2317671>
- [29] Smaili, N., Djeddou, M., Ghanem, K., Azrar, A. (2018). Effect of oblique projection beamforming on the achievable rate in massive MIMO system. *Journal of Communications Technology and Electronic*, 63: 1296-1302. <https://doi.org/10.1134/S1064226918110128>
- [30] Alwakeel, A.S., Mehana, A.H. (2019). Multi-cell MMSE data detection for massive MIMO: New simplified bounds. *IET Communications*, 13(15): 2386-2394. <https://doi.org/10.1049/iet-com.2018.5029>
- [31] Kay, S.M. (1993). *Fundamentals of Statistical Signal Processing: Estimation Theory*. NJ, USA, Prentice-Hall Signal Processing Series.
- [32] Silverstein, J.W., Bai, Z.D. (1995). On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate Analysis*, 54(2): 175-192. <https://doi.org/10.1006/jmva.1995.1051>
- [33] Alwakeel, A.S., Mehana, A.H., Ghoneim, A. (2017). Pilot hopping in multi-cell massive MIMO systems: Pilot assignment and new results. 2017 International

Conference on Computing, Networking and Communications (ICNC), Silicon Valley, CA, USA, pp. 460-464. <https://doi.org/10.1109/ICCNC.2017.7876172>

[34] Zakhour, R., Hanly, S.V. (2012). Base station cooperation on the downlink: Large system analysis. *IEEE Transactions on Information Theory*, 58(4): 2079-2106. <https://doi.org/10.1109/TIT.2011.217777>

[35] Li, X., Bjornson, E., Larsson, E.G., Zhou, S., Wang, J. (2015). A multi-cell MMSE detector for massive MIMO systems and new large system analysis. 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, CA, USA, pp. 1-6. <https://doi.org/10.1109/GLOCOM.2015.7417112>

[36] Fatema, N., Hua, G., Xiang, Y., Peng, D., Natgunanathan, I. (2018). Massive MIMO linear precoding: A survey. *IEEE Systems Journal*, 12(4): 3920-3931. <https://doi.org/10.1109/JSYST.2017.2776401>

[37] Bashar, M., Cumanan, K., Burr, A.G., Debbah, M., Ngo, H.Q. (2018). Enhanced max-min SINR for uplink cell-free massive MIMO systems. 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, pp. 1-6. <https://doi.org/10.1109/ICC.2018.8422577>

[38] Kammoun, A., Müller, A., Björnson, E., Debbah, M. (2014). Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO systems. *IEEE Journal of Selected Topics in Signal Processing*, 8(5): 861-875. <https://doi.org/10.1109/JSTSP.2014.2322582>

[39] Adhikary, A., Ashikhmin, A., Marzetta, T.L. (2017). Uplink interference reduction in large-scale antenna systems. *IEEE Transactions on Communications*, 65(5): 2194-2206. <https://doi.org/10.1109/TCOMM.2017.2662023>

[40] Donald, V.H.M. (1979). *Advanced Mobile Phone Service: The Cellular Concept*. NY, USA, Bell System Technical Journal.

[41] Khan, M.H.A., Chung, J., Lee, M.H. (2016). Downlink performance of cell edge using cooperative BS for multicell cellular network. *EURASIP Journal on Wireless Communications and Networking*, 2016(1): 56. <https://doi.org/10.1186/s13638-016-0537-0>

[42] Björnson, E., Sanguinetti, L., Hoydis, J., Debbah, M. (2015). Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer? *IEEE Transactions on Wireless Communications*, 14(6): 3059-3075. <https://doi.org/10.1109/TWC.2015.2400437>

[43] Marzetta, T.L. (2010). Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9(11): 3590-

3600. <https://doi.org/10.1109/TWC.2010.092810.091092>

[44] Atzeni, I., Arnau, J., Debbah, M. (2015). Fractional pilot reuse in massive MIMO systems. 2015 IEEE International Conference on Communication Workshop (ICCW), London, UK, pp. 1030-1035. <https://doi.org/10.1109/ICCW.2015.7247312>

## NOMENCLATURE

|                         |                                      |
|-------------------------|--------------------------------------|
| $M$                     | Total number of base station antenna |
| $B$                     | bandwidth                            |
| $\mathbf{H}$            | channel matrix                       |
| $\mathbf{R}$            | correlation matrix                   |
| $\hat{\mathbf{H}}$      | estimation channel matrix            |
| $l$                     | cell                                 |
| $j$                     | base station                         |
| $k$                     | user                                 |
| $\mathbf{w}$            | pre-coding vector                    |
| $P$                     | total transmit power                 |
| $x$                     | transmitted data                     |
| $y$                     | received signal                      |
| $d(\mathbf{z})$         | channel variance                     |
| $\mathbf{Im}$           | complex numbers                      |
| $\mathbf{h}$            | channel response                     |
| $\mathbf{n}$            | noise vector                         |
| $slm$                   | symbol of user $m$ in cell $l$       |
| $\mathbf{z}$            | positions of users                   |
| $\mathbf{b}$            | positions of base stations           |
| $\mathbf{V}$            | detector matrix                      |
| $\frac{\rho}{\sigma^2}$ | SNR                                  |

## Greek symbols

|                        |                                 |
|------------------------|---------------------------------|
| $(\cdot)^T$            | transpose                       |
| $(\cdot)^H$            | conjugate transpose (Hermitian) |
| $\kappa$               | pathloss exponent               |
| $CN$                   | complex gaussian distribution   |
| $\boldsymbol{\varphi}$ | covariance matrix               |
| $\mathbb{E}$           | expected value                  |
| $V\{\cdot\}$           | variance                        |

## Subscripts

|     |                    |
|-----|--------------------|
| Hz  | hertz              |
| m   | meter              |
| dBm | decibel milliwatts |