Spectral and Energy Efficient Low-Overhead Uplink and Downlink Channel Estimation for 5G Massive MIMO Systems

Imran Khan [1] [ID], **Mohammad Haseeb Zafar** [1], **Mohammad Tariq Jan** [2], **Jaime Lloret** [3] [ID],
Mohammed Basheri [4] and **Dhananjay Singh** [5,*] [ID]

[1] Department of Electrical Engineering, University of Engineering and Technology, Peshawar 814, Pakistan;
imran_khan@uetpeshawar.edu.pk (I.K.); haseeb@uetpeshawar.edu.pk (M.H.Z.)
[2] Department of Physics, Kohat University of Science and Technology (KUST), Kohat 26000, Pakistan;
tariqjan@kust.edu.pk
[3] Instituto de Investigación para la Gestión Integrada de Zonas Costeras, Universitat Politècnica de València,
46022 Camino de Vera, Spain; jlloret@dcom.upv.es
[4] Department of Information Technology, Faculty of Computing and Information Technology,
King Abdulaziz University, Jeddah 21589, Saudi Arabia; mbasheri@kau.edu.sa
[5] Department of Electronics Engineering, Hankuk University of Foreign Studies, Yongin 449-791, Korea
* Correspondence: dsingh@hufs.ac.kr; Tel.: +82-105-844-9415

Abstract: Uplink and Downlink channel estimation in massive Multiple Input Multiple Output (MIMO) systems is an intricate issue because of the increasing channel matrix dimensions. The channel feedback overhead using traditional codebook schemes is very large, which consumes more bandwidth and decreases the overall system efficiency. The purpose of this paper is to decrease the channel estimation overhead by taking the advantage of sparse attributes and also to optimize the Energy Efficiency (EE) of the system. To cope with this issue, we propose a novel approach by using Compressed-Sensing (CS), Block Iterative-Support-Detection (Block-ISD), Angle-of-Departure (AoD) and Structured Compressive Sampling Matching Pursuit (S-CoSaMP) algorithms to reduce the channel estimation overhead and compare them with the traditional algorithms. The CS uses temporal-correlation of time-varying channels to produce Differential-Channel Impulse Response (DCIR) among two CIRs that are adjacent in time-slots. DCIR has greater sparsity than the conventional CIRs as it can be easily compressed. The Block-ISD uses spatial-correlation of the channels to obtain the block-sparsity which results in lower pilot-overhead. AoD quantizes the channels whose path-AoDs variation is slower than path-gains and such information is utilized for reducing the overhead. S-CoSaMP deploys structured-sparsity to obtain reliable Channel-State-Information (CSI). MATLAB simulation results show that the proposed CS based algorithms reduce the feedback and pilot-overhead by a significant percentage and also improve the system capacity as compared with the traditional algorithms. Moreover, the EE level increases with increasing Base Station (BS) density, UE density and lowering hardware impairments level.

Keywords: 5G; CS; sparsity; feedback; pilot

1. Introduction

Massive MIMO is an upcoming technology for the future 5G Wireless Communication Systems [1]. It is a physical layer (PHY-L) envisioned technology that will resolve various physical layer issues in modern cellular networks. It has attracted substantial research interest from various communities for assessing its various merits. It has the ability to provide improved Spectral Efficiency (SE) and

Energy Efficiency (EE) by utilizing the space-resources. It is considered to be the future enabler of bandwidth-efficient, energy-efficient, secure and robust broadband networks. It will be deployed on the Internet of Things (IoT), Smart City and many other emerging infrastructures such as e-health [2]. It has high spatial-resolution, simple hardware-design, and less inter-cell interference. It uses hundreds or even thousands of low-power, physically-compact, and independently controllable antennas at the Base Stations (BS) in order to serve multiple-users Mobile Stations (MSs) or User Equipment (UE) simultaneously using the same time-frequency resources [3]. It allows adding enhancements such as congestion control systems [4] and intelligent handover algorithms [5]. The data rate capacity is increased due to the spatial-multiplexing which transmit data to the desired UE. The energy efficiency is boosted due to the sharp and more focused beamforms for the desired user location [6].

Massive MIMO provides a favorable propagation channel thanks to a large number of deployed antennas. Such a channel has the important property that the channel-vectors and UEs become orthogonal in pairs, and this can be easily processed using simple linear signal processing techniques such as Matched Filtering (MF). Massive MIMO can also be deployed together with mmWave to efficiently utilize its bandwidth which is also a key technology in 5G Ultra-Dense Networks (UDNs) [7].

To obtain both SE and EE, it requires accurate Channel State Information (CSI) at the Transmitter (Tx) Side [8]. The allocated channel resources are limited and the Massive MIMO channel matrix size grows immensely; therefore, channel feedback at the Uplink (UL) turns out to be contesting problem. The feedback and the data signals are transmitted together, so if the feedback signal overhead is large then it will consume extra bits from the scarce spectrum, which is the problem that is addressed in this work. The conventional codebook channel feedback technique is not appropriate as its size should be greatly extended to ensure adequate CSI accuracy. The non-codebook channel estimation methods have high feedback overhead. Therefore, Compressive-Sensing (CS) and AoD techniques are employed here to decrease its overhead. For Downlink (DL) channel estimation, accurate CSI is also mandatory for improved SE and EE. The conventional DL channel estimation techniques such as Minimum Mean Square Error (MMSE) and Least Square (LS) are not appropriate because, the number of orthogonal-pilots increases directly with the number of BS antennas, and it causes excessively large-overhead. To decrease such overhead, we deploy an enhanced Block-ISD and S-CoSaMP algorithms.

In this paper, we propose a novel approach of deploying various renowned CS-based algorithms for the efficient UL and DL channels estimation with low-overhead. We also analyze the EE of the system in terms of BS density, UE density and level of hardware impairment by deploying small cells in dense network scenario with the specific constraints. We deploy Compressed Sensing (CS), Block-ISD, AoD and S-CoSaMP algorithms to reduce the channel feedback and pilot overhead for UL and DL scenarios. Theoretical formulations are also provided for the better understanding and comparison. Such methods significantly decrease the channel overheads, while enhances the system performance. CS is a new signal-processing concept that can recover high-scale sparse-signals from low-scale measurements with an adequate probability [9]. The wireless channels have limited number of significant-paths which exhibit sparse response and that is considered as the basic principle by CS techniques. The Block-ISD fully utilizes the sparsity inside the block-sparse-equivalent Channel Impulse Response (CIR) which is obtained from the spatial-correlations of the considered CIRs. It is a better approach because it does not depend on prior knowledge of sparsity-level and it has more probability of practical implementations as compared to other schemes.

The AoD adaptive-subspace codebook utilizes the channel characteristics that the variation in path-AoDs is much slower than the variations in path-gains. It uses the phenomenon of channel-quantization to quantize the channel-vectors. The path-AoDs remains constant within the angle-coherence time. In the N-dimensional space, where N is the number of BS-antennas, the channel-vector is only distributed in the sub-space of it. It is also considered as the channel-subspace. The quantization-vectors are focused and mapped accurately on the normalized-channel subspace whose size is very smaller than the complete space. Thus, the number of feedback overhead bits is linearly proportional to the number of path-AoDs rather than the whole number of BS antennas, at that is the basis for reduced overhead. It also assists in keeping a constant data-rate gap without affecting the other parameters. Furthermore, our AoD codebook method is

better than the conventional channel-statistics based method as it can track the channel-vectors more efficient. The S-CoSaMP scheme also provides a clue to reduce the channel overhead efficiently. Dense antenna-arrays provides more correlated signals at the BS. As the BS-antennas are closely-spaced which causes same path-delays, so the CIRs of each antenna has a common support vector which is deployed by S-CoSaMP to decrease the undesired overhead. S-CoSaMP recovers the CIR from the reduced pilots. This technique can further decrease the UL/DL feedback overhead. This scheme uses SCS-joint UL/DL mechanism for channel estimation. It also eliminates the need for channel estimation at each UE and then feedback to BS. UEs directly feedback their pilots to BS where CS-algorithms recover the corresponding CIR. It uses a joint mechanism which improves the system performance as compared to the conventional CS-algorithms that performs channel training and feedback separately.

The rest of the paper is organized as follows: Section 2 describes the literature review of the paper which discusses the related background work. Section 3 discusses the proposed system model with all necessary theoretical as well as mathematical explanation. Section 4 provides the simulation results which show a comparison between the proposed and the conventional algorithms in terms of various parameters. Finally, Section 5 provides our conclusions and future work.

Notations: Lowercase and upper-case letters denote vectors and matrices, respectively; $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$ denote the transpose, conjugate transpose, and inverse of a matrix, respectively; ϕ represents the sensing matrix, γ denotes the feedback overhead, I_K denotes the identity matrix of $K \times K$ dimension; ρ is the correlation coefficient, α_{h_i} is the support element of CIR h_i and μ is the channel sparsity.

2. Literature Review

The recent research activities indicate that the main goals of the 5G mobile communications networks are to accomplish $1000\times$ times the system-capacity and $10\times$ times the Spectral Efficiency (SE), Energy Efficiency (EE) and data-rate, and $25\times$ times the average-cell-throughput [9]. The demand for high data-rate and link-reliability is increasing exponentially due to the increasing number of new generation of devices such as smart-phones, tablets and notebooks, etc. The electromagnetic spectrum is a precious but limited commodity, so, to utilize such spectrum efficiently and to cope with the demands, the high-level perspective technology of massive MIMO is deployed. Massive MIMO is defined as a system that deploys a large number of smaller array antennas at the BS which significantly improve the beamforming and system capacity and can simultaneously provide service to a large number of users (UEs) [10]. It has various merits over the conventional MIMO. First, it uses a large number of antennas at the BS due to which the simplest coherent-combiner and linear-precoder can be used for signal processing such MF or ZF. Second, by increasing the number of antennas increases the system capacity substantially using the channel-reciprocity features and without increasing feedback-overhead. Third, the reduced power benefits in the UL/DL provide the feasibility to shrink the cell-size which can be used in micro and pico-cells. Massive MIMO Technology is used by various wireless communication standards and companies such as the Long-Term-Evolution-Advanced (LTE-A), due to its substantial potential both in channel-capacity and route-reliability. It uses hundreds of antennas at the BS in various geometrical shapes in the form of arrays. Increasing the number of antennas assist to direct the radiated energy towards much smaller-regions, which improves the throughput and EE. It also provides aggressive-spatial multiplexing. It has the ability to decrease the transmitter (Tx) power and also enhances the Spectrum Efficiency (SE) by 10 times and more. Also, low-power devices are deployed for service purpose, which is also a plus point for massive MIMO as it is low-cost and easily manufactured from solid-state devices.

The performance characteristics of the wireless system are not solely based on the transmitter (Tx) and receiver (Rx), but are also dependent upon the wireless channel under consideration. The wireless channel is very complicated and unpredictable. It is the key element to understand the development of high-performance, improved Quality-of-Service (QoS) and bandwidth-efficient wireless communication technology. In Massive MIMO Systems, accurate UL and DL CSI are necessary for channel estimation [5]. In [3], it is illustrated that in propagation-environments, the MIMO-channels shows a sparse-structure in Channel Impulse Response (CIR). To reap the advantages of massive MIMO, accurate CSI is mandatory

for each Tx-Rx pair. As there is a large number of antennas used by massive MIMO, therefore the channels estimations associated with such antennas will also be large which is a tedious task because of the enormous increasing dimensions of the channel matrix. Many channel estimation techniques for CIR are proposed in the related research literature which can be classified into superimposed [11], training [12], pilot [13], blind [14], and semi-blind [15] techniques, respectively.

The CS paradigm is getting much research attention as it has the ability to recover the unknown signals from just a small number of measurements. It uses very few samples following the Nyquist sampling rate. It deploys the non-zero elements of a sparse-signal for the recovery operation. It also performs accurate estimation of system-parameters with fewer-pilots, therefore it also increases the bandwidth (BW) efficiency. The classical conventional CS-paradigms need pre-known information of channel-sparsity, which is practically unknown in system analysis. Furthermore, for reliable and accurate CIR retrieving, the CS-sensing or sampling matrix must satisfy the isometry property.

The previous research performed for channel estimation did not perform UL/DL channel estimation jointly using CS [16–19], differential CIR [20], Block-ISD [21], AoD [22] and S-CoSaMP [23]. Therefore, we propose an efficient, low-overhead UL/DL joint channel feedback and pilot-based channel estimation which gives an optimum comparison with the conventional schemes such as direct-CS, classical Block-ISD, BP, LS, CoSaMP, Kalman filter, J-OMP and conventional codebook mechanisms respectively. The simulation results provide a clear understanding of the improvement caused by the proposed algorithms in terms of the mean square error (MSE), data-rate, per-user rate and feedback bits.

3. System Model

The proposed system is modeled using N_t transmitter antennas, *K* receiving antennas such that ($N_t >> K$), SNR, channel length, channel sparsity, slot-interval, and various relevant parameters to obtain the required results.

3.1. CS-Differential Channel Feedback

The proposed CS algorithm for Uplink (UL) differential channel feedback is depicted in Figure 1. The CIR channel feedback is performed at particular time-slots. The direct compressed sensing (DCS) method does not perform the differential operation and directly compresses the CIR with sensing-matrix. But the CS differential algorithm applies differential operation for generating more strong and sparse Differential CIR (DCIR) than the DCS CIR. DCIR is then compressed to Compressed CIR (CCIR) using the sensing-matrix and it is known to base-station (BS) and User (UE). Such CCIR is fed back to BS. The CS Recovery algorithm gets back the DCIR from the CCIR accurately.
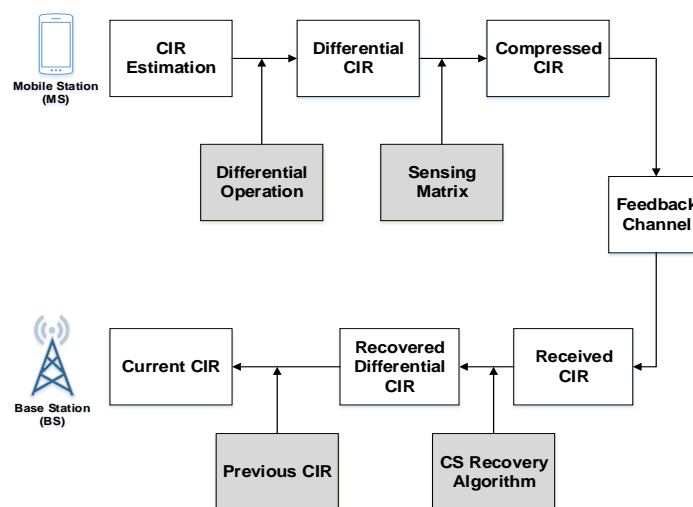

Figure 1. Proposed CS-differential UL channel-feedback.

3.1.1. Temporal-Correlation of Massive MIMO Channels

Consider *k*-time slots for the time-varying channels between *n*-number of BS antennas and single UE antenna. Mathematically, the Channel Impulse Response (CIR) at *k*th time-slot between the *n*th BS antenna and single UE is represented by:

$$h_n^{(k)} = [h_n^{(k)}(0),\ h_n^{(k)}(1),\ h_n^{(k)}(2),\ \ldots,\ h_n^{(k)}(L-1)]^T \text{ for } 1 \le n \le N \tag{1}$$

where: *N*: Number of BS-antennas and L: Maximum channel Length.

The Sparse CIR is represented by a supportive-vector $p_n^{(k)}$ and an amplitude-vector $a_n^{(k)}$:

$$h_n^{(k)} = p_n^{(k)} o\ a_n^{(k)} \tag{2}$$

where $p_n^{(k)}(l) \in \{0, 1\}$, $a_n^{(k)} \in \mathbb{C}$ and "*o*" indicates the Hadamard-product.

3.1.2. Support-Vector

$p_n^{(k)}$ having *l*th elements and *T*-vectors can be represented by a first-order Markov process that represent it by two transition-probabilities:

$$p_{10} \triangleq P_r\{p_n^{(k+1)}(l) = 1 | p_n^{(k)}(l) = 0\} \tag{3}$$

and:

$$p_{01} \triangleq P_r\{p_n^{(k+1)}(l) = 0 | p_n^{(k)}(l) = 1\} \tag{4}$$

where, p_{10} is the transition probability from 0 (at time slot *k*) to 1 (at time slot *k* + 1). The corresponding distribution for initial time-slot (*k* = 1) is:

$$\mu_n^{(1)} \triangleq P_r\{p_n^{(1)}(l) = 1\} \tag{5}$$

For the steady-state Markov process:

$$P_r\{p_n^{(k)}(l) = 1\} = u \tag{6}$$

For all *k*, *n*. Where, μ denotes the channel sparsity. From the above equations, only p_{01} and μ are important to represent the Markov process as p_{10} can be obtained from p_{01} and μ respectively.

3.1.3. Amplitude-Vector

It is represented by a first-order auto-regressive process as [24]:

$$a_n^{(k)} = \rho\ a_n^{(k-1)} + \sqrt{1 - \rho^2}\omega^{(k)} \tag{7}$$

where $\omega^{(k)}$: is the noise-vector whose elements are independent identically distributed (i.i.d) with zero mean Gaussian variables which follow $\mathbb{C}N(0, \sigma_\omega^2)$ distribution and does not depend on $a_n^{(k-1)}$, ρ: is the correlation-coefficient which is obtained by the Bessel-function of zero order as follows:

$$\rho = J_o(2\pi f_d \tau) \tag{8}$$

f_d : Maximal Doppler-frequency and τ : Time-interval of slot.

3.1.4. Differential CIR (DCIR)

It is determined by taking the difference among two CIRs at the previous at current time-slots, which is the required sparse-signal. Mathematically:

$$\Delta h_n^{(k)} = h_n^{(k)} - h_n^{(k-1)} \tag{9}$$

By putting the corresponding values from (2) and (7) in (9) we get:

$$\begin{aligned}
&= p_n^{(k)} \text{o} \left(a_n^{(k)} - a_n^{(k-1)} \right) + \left(p_n^{(k)} - p_n^{(k-1)} \right) \text{o}\, a_n^{(k-1)} \\
&= p_n^{(k)} \text{o} \left[\sqrt{1 - -\rho^2}\,\omega - (1-\rho)a_n^{(k-1)} \right] + \left(p_n^{(k)} - p_n^{(k-1)} \right) \text{o}\, a_n^{(k-1)}
\end{aligned} \tag{10}$$

The above process is graphically illustrated in Figure 2, in which the previous CIR, current CIR and the differential CIR (DCIR) are shown. It is clear from the figure that the DCIR has more sparsity and can be more compressed than the other two CIRs.
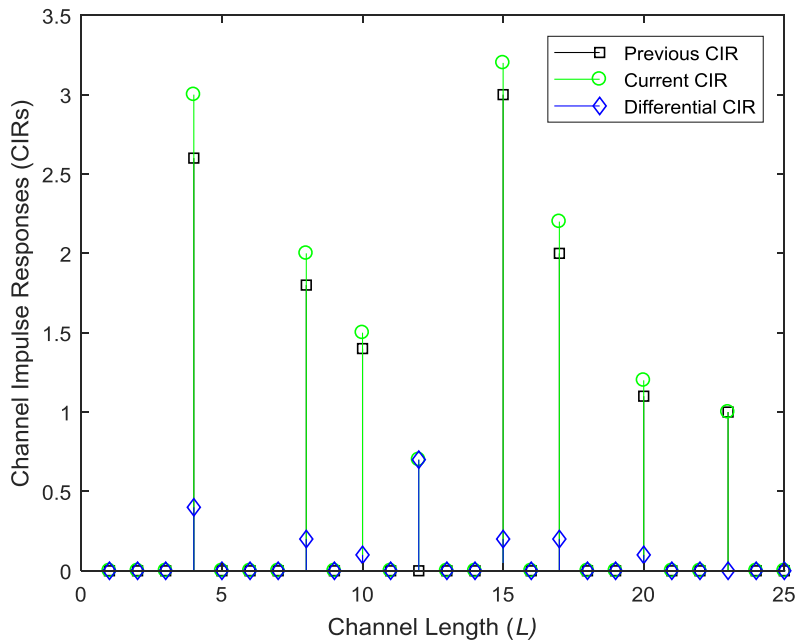

Figure 2. Illustration of the previous, current and differential channel impulse responses (CIRs) in different time slots with same channel length.

3.1.5. Compression CIR (CCIR)

The $\Delta h_n^{(k)}$ is compressed by the sensing matrix $\phi \in \mathbb{C}^{MxL}$ with $M \ll L$ and convert it into a lower level measurement-vector as follows:

$$y = \phi \Delta h_n^{(k)} \tag{11}$$

y: compressed channel-feedback to BS.

3.1.6. Feedback-Overhead

It is expressed as a ratio of channel compression w.r.t channel length. This ratio should be small enough for efficient channel-feedback and recovery process. Mathematically:

$$\gamma = \frac{M}{L} \tag{12}$$

where $M \ll L$ so the ratio is less than 1 (100%) and it is added as extra overhead bits to the feedback channel attributes.

3.1.7. CS-Recovery Algorithm

The desired $\Delta h_n^{(k)}$ is recovered by the BS from the noise accumulated signal if ϕ fulfills the isometric-property [25]. Mathematically:

$$y = \phi \Delta h_n^{(k)} + n \tag{13}$$

where n represents the feedback-channel noise having the same attributes as the noise-vector ω in (7) and follows $\mathbb{CN}\left(0, \sigma_n^2\right)$.

3.1.8. Flowchart

The flowchart in Figure 3 describes the major steps involved in the CS-differential channel estimation and feedback process. First, the system parameters are initialized which are mentioned in Table 1. Then the differential operation is performed to produce DCIR. The sensing matrix then performs compressing operation for CCIR and that is feedback to BS for recovering. The condition checks whether φ obey isometry, if yes then the required CIR is obtained to complete the desired channel-estimation process.
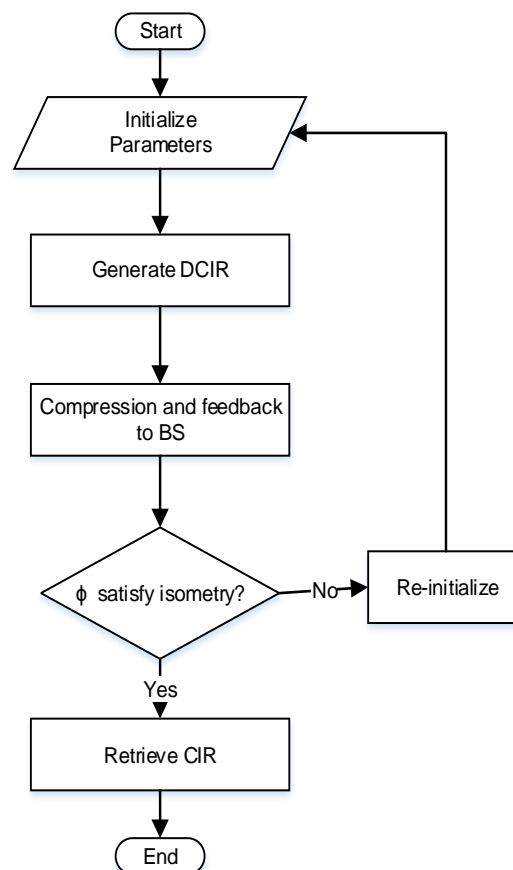

Figure 3. CS-UL channel-feedback flowchart.

3.2. Block-ISD Pilot Feedback

Figure 4 shows the proposed block-ISD based channel-feedback schemes. The same number of antennas are deployed for downlink channel estimation as in uplink channel estimation process.

Orthogonal Frequency Division Multiplexing (OFDM) scheme is employed to determine the number of pilots easily in frequency-domain [26–28]. Mathematically, the received pilots-signals at the MS or UE is expressed as:

$$y_j = \sum_{i=1}^{N_T} P_i(F_L)_j h_i + n_j \tag{14}$$

where $j = [1, 2, \ldots, N]$: Is the index of subcarrier-set that are chosen randomly and allocate to pilots-signals, y_j: Measurements performed for CIR recovery; $P_i = \text{diag}(p_i)$: Is the diagonal matrix of the pilot-vector p_i at the corresponding ith Tx antenna, $p_i \in \mathbb{C}^{px1}$ here p represents the number of pilot-signals; $F_L \in \mathbb{C}^{MxL}$ is the sub-matrix which contains the Discrete Fourier Transform (DFT) column matrix of L columns of order M × M; M is the symbol-length of OFDM; $(F_L)_j$ is the sub-matrix containing rows of F_L sub-matrix with index of j; $h_i = [h_i(1), h_i(2), \ldots, h_i(L)]^T$ is the CIR between the ith BS antenna and single-user Rx antenna; L is the maximum channel-length; $n_j = [n_1, n_2, \ldots, n_p]^T$ is the channel noise-vector, having the same characteristics as discussed in Equation (13).
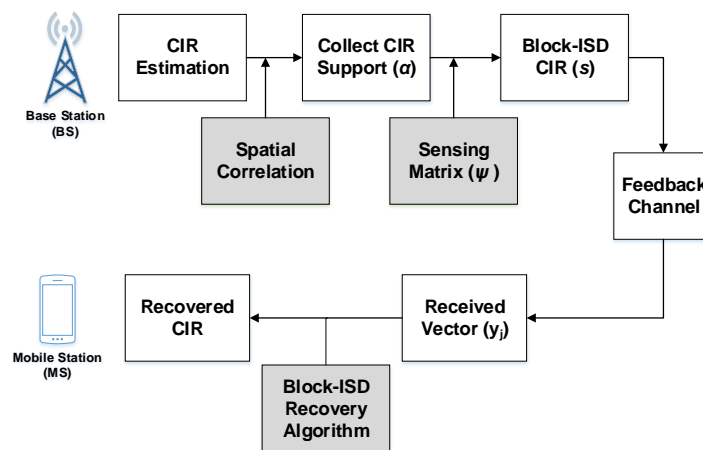

Figure 4. Proposed Block-ISD channel-feedback.

For the sake of convenience, (14) can also be written as:

$$y_j = \phi \times h + n_j \tag{15}$$

where $\phi = [P_1(F_L)_j, P_2(F_L)_j, \ldots, P_{N_T}(F_L)_j]$ is the sensing-matrix and $h = [h_1^{(T)}, h_2^{(T)}, \ldots, h_{N_T}^{(T)}]^T$ is the cumulative CIR from the N_T number of antennas.

Creating Block-ISD CIR

For the recovery process, the CIR (h) is required which is extracted from the noise-accumulated signal. The spatial-correlation operation is performed on various CIRs of the considered antennas. The CIRs between the BS and UE of each antenna exhibit analogous properties such as the Time of Arrival (ToA), so they share a common-supporting element as follows [29]:

$$\alpha_{h_1} = \alpha_{h_2} = \ldots = \alpha_{h_{N_T}} \tag{16}$$

where $\alpha_{h_i} = \{q : h_i(q) \neq 0\}$ is the support of h_i.

Figure 5 illustrates the production of block-sparse equivalent CIR (s) from the considered CIRs which has more sparsity. The green, yellow and orange colors of CIRs (h) represents the non-zero elements and that is also the common support elements among them from which the desired sparse-CIR is constructed.
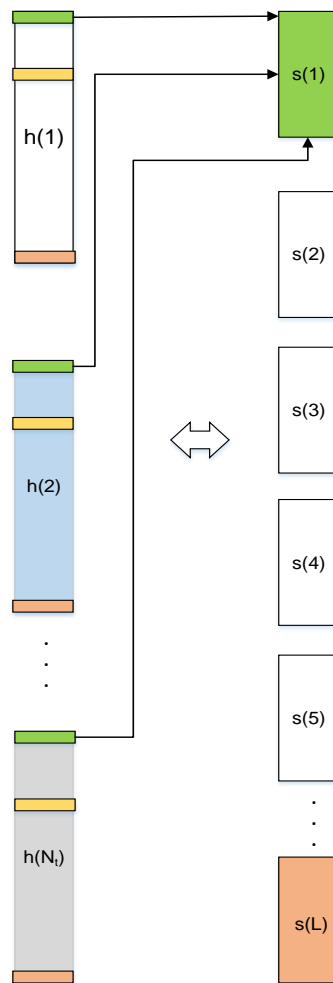
Figure 5. Creation of block-sparse equivalent-CIR.

The equivalent CIR (*s*) is given as:

$$s = \left[s_1, \; s_2, \; \ldots, \; s_L \right]^T \tag{17}$$

which is obtained from the following equation:

$$s((l-1)N_T) + n_t) = h((n_t - 1)L + l) \tag{18}$$

where $l = 1, 2, \ldots, L; n_t = 1, 2, \ldots, N_T$.

From (18), the zero and non-zero elements can be grouped separately for feasibility if s is decomposed into L-blocks where each block contains N_T-elements. Therefore, the supportive elements of s (i.e., α_s) can be updated collectively using the basic pursuit (BP) technique. To represent the new sensing-matrix ψ with more-sparse attributes in terms of ϕ in (15) and (18) it is written as:

$$\psi(:, \; (l-1)N_T + n_t) = \phi(:, \; (n_t - 1)L + l) \tag{19}$$

and the corresponding estimated channel from (15) can be written as:

$$y_j = \psi \times s + n_j \tag{20}$$

where y_j Is the received vector of order $p \times 1$; *s* is the require sparse CIR with the order $N_T L \times 1$.

3.3. AoD-Adaptive Subspace Codebook Algorithm

We consider the same number of BS and UE antennas as in CS and Block-ISD techniques. It is also based on CS algorithms. To obtain the DL channel-vector, we employ classical narrowband ray-based mmWave channel model which is given by [30,31]:

$$h_j = \sum_{i=1}^{P_j} G_{j,i} a(\theta_{j,i}) \tag{21}$$

where h_j is the DL channel-vector for jth user, that is also required by the BS for precoding and power allocation; P_j is the number of resolvable paths from BS to jth UE; $G_{j,i}$ is the i.i.d complex-gain of the i-th propagation-path of the j-th UE having zero mean and unit variance; $\theta_{j,i}$ is the i-th path-AoD of the j-th UE; $a(\theta_{j,i}) \in \mathbb{C}^{Nx1}$ is the steering-vector representing the antenna-response of the i-th path of the j-th UE.

To obtain $a(\theta_{j,i})$, let us suppose that the BS antennas are uniform linear-arrays (ULA), therefore:

$$a(\theta_{j,i}) = \left[1, e^{-j2\pi\frac{d}{\lambda}\sin(\theta_{j,i})}, \ldots, e^{-j2\pi\frac{d}{\lambda}(M-1)\sin(\theta_{j,i})}\right]^H \tag{22}$$

where d is the spacing or separation between antenna array elements at the BS; λ Is the wavelength of the carrier frequency.

For the sake of convenience, we write (21) in matrix form as:

$$h_j = A_j G_j \tag{23}$$

where $A_j = \left[a(\theta_{j,1}), a(\theta_{j,2}), \ldots, a(\theta_{j,P_j})\right] \in \mathbb{C}^{NxP_j}$; $G_j = \left[g_{j,1}, g_{j,2}, g_{j,3}, \ldots, g_{j,P_j}\right]^H \in \mathbb{C}^{P_j x1}$.

The channel-vectors of all K-users are concatenated and represented as follows:

$$H = [h_1, h_2, h_3, \ldots, h_K] \in \mathbb{C}^{NxK} \tag{24}$$

The channel quantization is performed by the codebook:

$$\mathbb{C}_j = [c_{j,1}, c_{j,2}, c_{j,3}, \ldots, c_{j,2^b,}] \tag{25}$$

where 2^b is the N-dimensional different column-vectors of unit-magnitude, and b is the number of bits required for feedback.

The quantization of the jth user channel-vector h_j to a quantization vector is given by:

$$h_j \rightarrow c_{j,F_j} \in \mathbb{C}^{N\times1} \tag{26}$$

where F_j is the quantization index. It is calculated by the following equation:

$$F_j = \underset{i\in\{1,\,2^b\}}{\arg\min}\left[\sin^2\left[\angle(h_j, c_{j,i,})\right]\right] = \underset{i\in\{1,\,2^b\}}{\arg\max}|\tilde{h}_j^H c_{j,i}|^2 \tag{27}$$

where $\tilde{h}_j$ is the direction of the concerned channel which can be obtained from the following relationship:

$$\tilde{h}_j = \frac{h_j}{||h_j||} \tag{28}$$

The quantization vector is transmitted and F_j can be feedback from the jth UE to BS by employing the feedback bits b. When the BS receive such b-bits and its associated F_j, it can produce the feedback channel-vector which is given by:

$$\check{h}_j = ||h_j|| c_{F_j} \tag{29}$$

The feedback channel-vectors of all *K*-users are concatenated and represented as follows:

$$\check{H} = [\check{h}_1, \check{h}_2, \check{h}_3, \ldots, \check{h}_K] \in \mathbb{C}^{N \times K} \tag{30}$$

The per-user rate is obtained from (30) using zero-forcing (ZF) linear precoding technique. ZF has the ability to attain near-optimal performance with lower complexity when the BS antennas approach infinity [32]. The signal that is transmitted after ZF precoding is determined by:

$$X = \sqrt{\frac{P_t}{K}} \, \mathbb{\Gamma} s \tag{31}$$

where P_t is the transmit power; $s = [s_1, s_2, s_3, \ldots, s_K] \in \mathbb{C}^{K \times 1}$ is a set of signals dedicated to the *K*-users with unit normalized-power; $\mathbb{\Gamma} = [\mathbb{\Gamma}_1, \mathbb{\Gamma}_2, \mathbb{\Gamma}_3, \ldots, \mathbb{\Gamma}_K] \in \mathbb{C}^{N \times K}$ is the ZF precoding-matrix which contains *K*-different unit-magnitude precoding-vectors $\mathbb{\Gamma}_i$ of order $N \times 1$. To obtain the precoding vectors, let us represent the aggregated channel vector as follows:

$$\eta = \check{H} \, (\check{H}^H \check{H})^{-1} \tag{32}$$

Therefore, the precoding-vectors can be obtained as the normalized *i*th column of η as given below:

$$\mathbb{\Gamma}_i = \frac{\eta(:,i)}{||\eta(:,i)||} \tag{33}$$

The received-signal for the *k*th user can be expressed as:

$$y_k = h_k^H X + \mathrm{n}_k \tag{34}$$

By putting (31) in (34) we get:

$$= \sqrt{\frac{P_t}{K}} h_k^H \, \mathbb{\Gamma}_k \mathrm{s}_k + \sqrt{\frac{P_t}{K}} \sum_{i=1,\, i \neq k}^{K} h_k^H \, \mathbb{\Gamma}_i \mathrm{s}_i + n_k \tag{35}$$

where n_k is the Gaussian noise-vector at the *k*th user having zero-mean and unit-variance. Therefore, the signal-to-interference noise ratio (*SINR*) at the *k*th user is given by:

$$SINR_k = \frac{\frac{P_t}{K} \left| h_k^H \, \mathbb{\Gamma}_k \right|^2}{1 + \frac{P_t}{K} \sum_{i=1,\, i \neq k}^{K} \left| h_k^H \, \mathbb{\Gamma}_i \right|^2} \tag{36}$$

The per-user data rate $\mathcal{R}$ is then calculated by:

$$\mathcal{R} = E\{\log_2(1 + SINR_k)\} \tag{37}$$

Putting (35) in (36), we get:

$$\mathcal{R} = E\{\log_2 (1 + \frac{\frac{P_t}{K} \left| h_k^H \, \mathbb{\Gamma}_k \right|^2}{1 + \frac{P_t}{K} \sum_{i=1,\, i \neq k}^{K} \left| h_k^H \, \mathbb{\Gamma}_i \right|^2})\} \tag{38}$$

In (38), $\mathcal{R}$ is dependent upon $\mathbb{\Gamma}$-matrix which in turn depends on the $\check{H}$-matrix quality. The quantization-vectors are accurately placed on the normalized-channel subspace which is pictorially depicted in Figure 6.
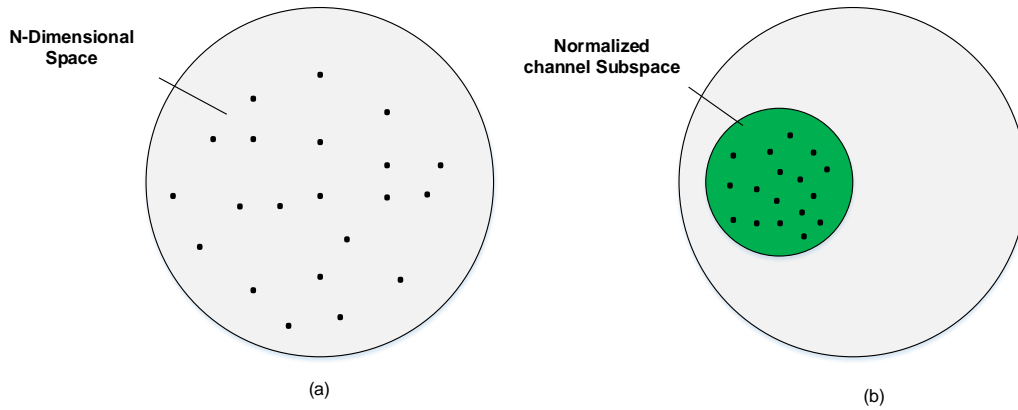
Figure 6. Comparison of the codebooks: (**a**) classical Grassmannian N-dimension; (**b**) Proposed AoD based codebook.

The path-AoD ($\theta_{j,i}$) in (29) is mainly dependent on the nearby hurdles which are around the BS and may not change physically their place in a lot of time as compare to the coherence-time. On the other hand, for the path-gain $(G_{j,i})$ a resolvable-path is created by a cluster comprised of scatters which is nearby the kth user and it contains a number of unresolvable-paths. Therefore, the kth user sees a path-gain that depends on a number of unresolvable-paths. It is clear that the path-gain has much quicker variation than the path-AoDs which vary much slowly. The coherence-time of the angle in which the path-AoDs are considered static and the channel-vector is present in the channel-subspace, is much longer as compare to the channel coherence-time.

Due to joint UL/DL channel estimation benefit, the quantized AoDs can be produced at the BS as well as users. They can also create the steering-matrix. Therefore, the proposed AoD-subspace based codebook is written as:

$$c_{j,i} \;=\; \frac{1}{\sqrt{N}}\check{A}_j\mathbb{Z}_i \tag{39}$$

where, $\mathbb{Z}_i \in \mathbb{C}^{k\times 1}$ is the traditional-codebook quantization-vector. The feedback bits b is determined as follows:

$$b \;\geq\; \frac{P-1}{3}\,\text{SNR} + (P-1)\log_2\left(\frac{K-1}{c-1}\right) \tag{40}$$

where SNR is computed at the Rx side which is expressed by:

$$\text{SNR} = 10\log_{10}\frac{P_t}{K}E[||h_k||^2] \tag{41}$$

It is clear from (41) that the slope of the b-bits $(Q-1)$ is directly proportional to SNR and the rate-gap remains constant. It is also clear that as Q is very much smaller than N, so the proposed AoD-codebook algorithm can substantially reduce the size of codebook and the overhead.

3.4. S-CoSaMP Algorithm

The conventional UL/DL schemes is a two-step process. In the first step, DL CSI is estimated at the UE and in the second step, the CSI is feedback in UL to BS. These two steps processes are individually optimized which is shown in Figure 7a. To point out the loopholes associated with such technique, we propose a novel approach that performs channel UL and DL feedback jointly. The proposed scheme process is depicted in Figure 7b. In this algorithm, the user feedbacks the pilots to BS without channel-estimation. The BS then extracts the required CIR using the CS algorithm. This is the main advantage of this algorithm as it provides freedom to users from channel-estimation which is a tedious task. Also, the performance of channel-feedback is also enhanced.
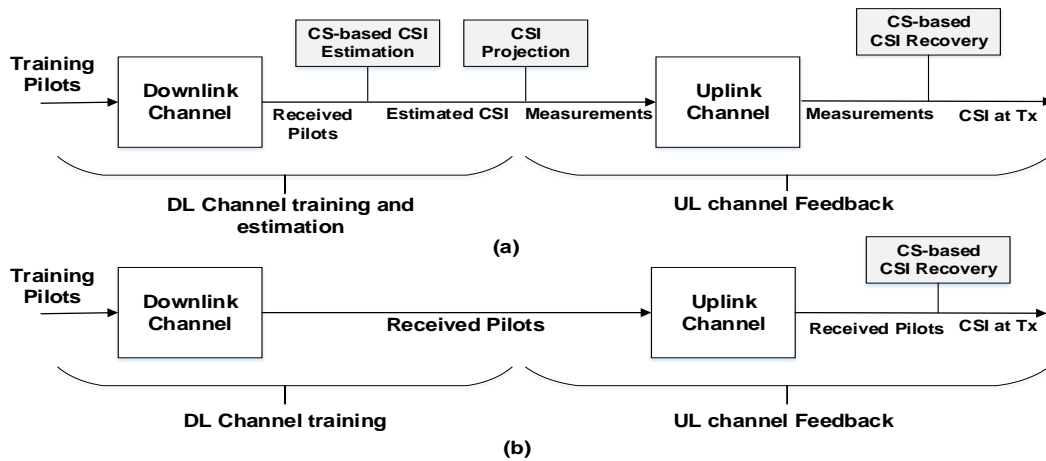
Figure 7. Comparison of the: (**a**) Conventional CS based separate DL and UL channel feedback; (**b**) Proposed CS based joint algorithm for channel estimation.

The procedure of obtaining the DCIR for UL is the same as explained in subsection A of CS differential CIR. To obtain the received pilots at the BS, we modify (15) as follows:

$$y_j^{(k)} = \phi \cdot h^{(k)} + n_j^{(k)} \tag{42}$$

We try to use the temporal-correlation of channels by calculating the differential-pilots (D-Pilots) at the BS in two consecutive time-slots $k-1$ and k respectively as [33]:

$$\Delta y_j^{(k)} = y_j^{(k)} - y_j^{(k-1)} \tag{43}$$

By putting (42) in (43) we get:

$$
\begin{aligned}
&= \phi(h^{(k)} - h^{(k-1)}) + n_j^{(k)} - n_j^{(k-1)} \\
&= \phi \Delta h^{(k)} + \Delta n_j^{(k)}
\end{aligned}
\tag{44}
$$

where $\Delta h^{(k)} = h^{(k)} - h^{(k-1)}$, which is the same as (9).

We know that $h^{(k)}$ and $h^{(k-1)}$ have sparse structure so the $\Delta h^{(k)}$ also have sparse nature. So, to reap these sparse benefits, we employ the S-CoSaMP algorithm which has low-complexity and robustness to recover the CIR for LS technique which is given by [34]:

$$\check{h} = \phi y_j \tag{45}$$

This holds true when the CIR has sparse nature. The CoSaMP technique recovers the CIR in the initial time-slot from the pilots received at the BS. It also recovers the DCIR from the D-pilots in the succeeding time-slots. It is also more beneficial than the conventional CoSaMP algorithm as it updates each support of CIR. The conventional CoSaMP does not take into account the sparse structure of CIR. Whereas, the proposed S-CoSaMP utilizes the sparsity inherent in the CIR.

3.5. Energy Efficiency Maximization

In this section, we analyze the Uplink energy efficiency (EE) in terms of the number of BSs, the number of users per cell, the density of BS antennas and the pilot reuse factor respectively. The closed form of analytical expressions is derived to provide insights into the energy optimization

of small cell based massive MIMO systems. The transmit power of user equipment *i* in *j* cell can be expressed by:

$$P_{ji} = \rho \omega d_{jji}^{\alpha} \tag{46}$$

where ρ is the coefficient of power control. The average transmit power for such UE can then be represented by:

$$E\{P_{ji}\} = \rho \omega E\{d_{jji}^{\alpha}\} = \rho \omega \, \frac{\mathbb{I}\left(\frac{\alpha}{2} + 1\right)}{(\pi \lambda)^{\frac{\alpha}{2}}} \tag{47}$$

Therefore, by using (47), we can calculate the optimized EE for the proposed massive MIMO system by through the statistical power control policy [35].

4. Simulation Results

The simulation results are provided to verify our theoretical analysis. Table 1 provides the simulation parameters for the proposed algorithms which are utilized to obtain the simulation results in MATLAB. The sub-space pursuit (SP) technique is used for the recovery process which is more robust to noise and have low complexity [36].

Table 1. Proposed System Simulation Parameters.

S. No	Parameter	Symbol	Value
1	Max. Channel Length	L	32
2	Pilot length	p	640
3	OFDM symbol-length	$\mathcal{S}$	2048
4	Signal-to-Noise-Ratio	SNR	25 dB
5	Block Elements	B	8
6	Noise Variance	σ_n	1
7	Transition-Probability	P_{01}	0.05 (5%)
8	Channel-Sparsity	μ	0.1 (10%)
9	Doppler-frequency	f_d	10 Hz
10	Slot-interval	τ	1 ms
11	Channel-variance	σ_w	1
12	Correlation coefficient	ρ	0.0628
13	Number of BS antennas	N_t	130
14	Number of User antennas	K	10
15	Number of Paths	P	5
16	AoD Feedback bits	b	5
17	Sensing-Matrix	φ	Random Gaussian
18	Static Power Consumption	P_s	10 W
19	Circuit power per antenna	P_{ca}	0.1 W
20	Path Loss Exponent	α	3.76
21	Coherence Block Length	S	400
22	Coherence Time	T_c	2.16 ms
23	Bandwidth	BW	20 MHz
24	Macrocell Radius	r	250 m

Figure 8 shows the comparison results for the uplink (UL) channel-feedback for massive MIMO system using the direct compressed-sensing (DCS) and the proposed CS differential techniques. The simulations are performed using two compression ratios (γ) for CIRs which are 25% and 50% respectively. The results clearly reveal that the DCS scheme fails for 25% compression ratio as its overhead is lesser than required and the gap between this and the proposed CS differential algorithm gets larger and its NMSE does not reduce as required, while the proposed scheme performs much better than former scheme. For the second compression ratio (50%), the proposed scheme still show better performance as compared to DCS. It also got 6-dB gain over DCS. The proposed CS differential channel feedback algorithm using 25% overhead have similar performance as the DCS algorithm that uses 50% overhead. Its mean that the proposed algorithm decreases the channel-feedback overhead from 50% to 25%, and the net reduction is 25%.
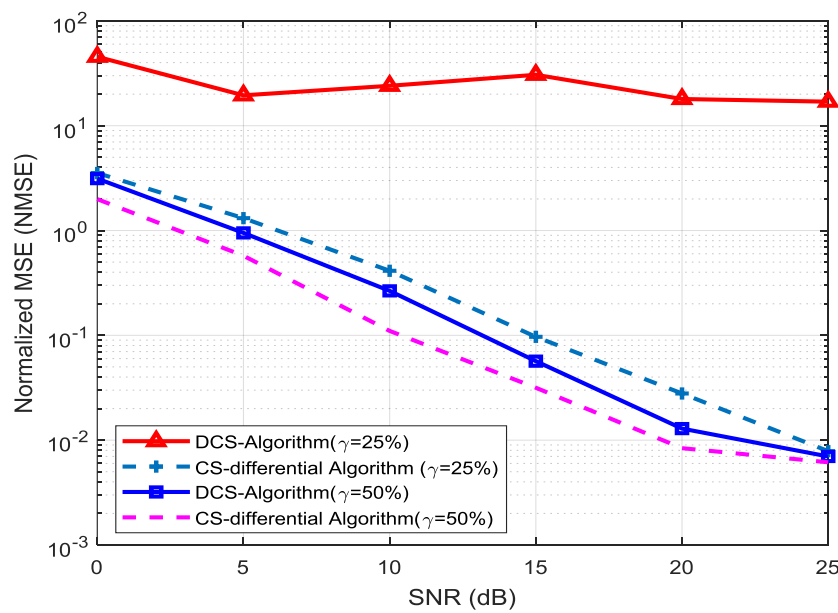
Figure 8. Efficiency analysis of channel-feedback between DCS and CS-differential algorithms.

Figure 9 shows the NMSE performance comparison of the classical ISD and the proposed Block-ISD algorithms. The results clearly show that the proposed Block-ISD performs better than the classical-ISD and BP algorithms due to more inherent sparsity. The exact LS is a reference for comparison of our results. The number of pilots for LS and other scheme is $p = LN_t = 32 \times 130 = 4160$ which is far more than the proposed pilots of 640. By taking the percentage between these two pilots signals $\left\lceil \frac{(4160-640)}{4160} \right\rceil \times 100 = 84.6\%$, it is evident that a significant reduction in the channel-overhead is obtained by using Block-ISD algorithm for channel-estimation of massive MIMO systems.
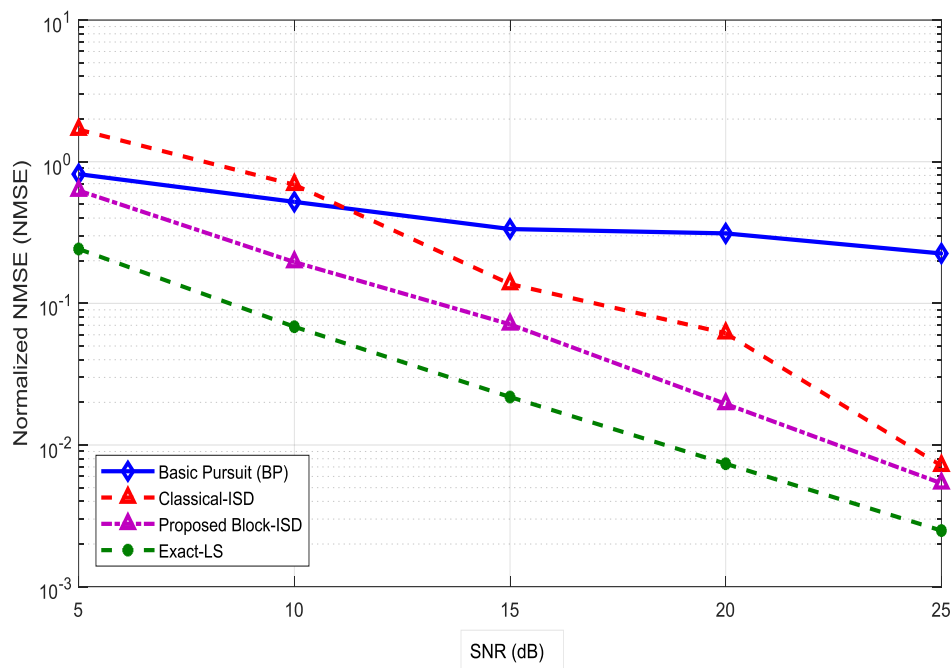
Figure 9. Efficiency analysis of channel-feedback between Block-ISD and classic-ISD algorithms.

Figure 10 illustrates the number of feedback bits b required to maintain a constant rate-gap between the ideal CSI and the proposed practical AoD-adaptive subspace codebook. We can see that

the rate gap is limited to within 0.12 bps/Hz. The graph is linear, therefore, the number of feedback bits required is linearly proportional to the number of paths. By comparing this simulation result with our theoretical formulation in (39), it is clearly validating it.
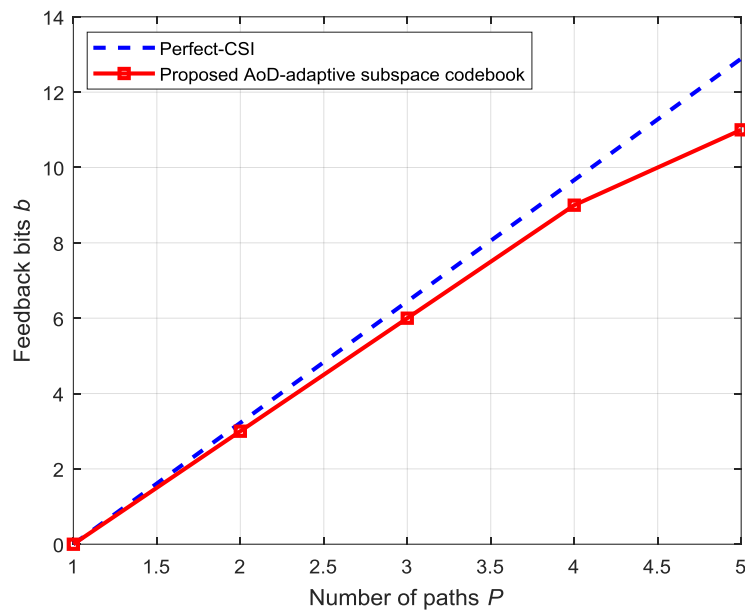

Figure 10. Feedback-bits comparison for maintaining a data-rate gap between the perfect-CSI and proposed AoD-codebook.

Figure 11 compares the required feedback bits versus the SNR between the ideal scenario of perfect-CSI and the proposed scheme with limited channel-feedback. We observe that the required number of feedback bits vary linearly with the SNR under consideration.
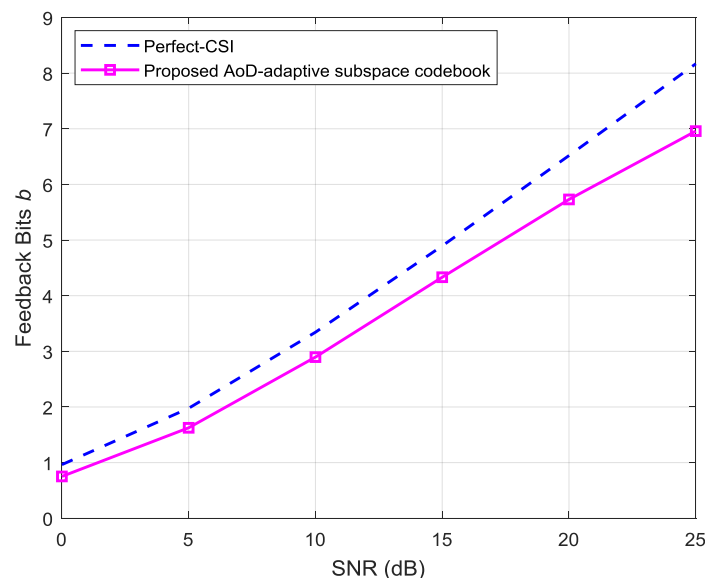

Figure 11. Comparison of the required number of feedback-bits with increasing the SNR between the ideal CSI and the proposed AoD-codebook schemes.

Figure 12 shows the per-user data rate comparison of the ideal condition, conventional channel statistics-based codebook and the proposed AoD-adaptive subspace codebook with perfect path-AoDs. The feedback bits *b* is constant for the quantized channel feedback algorithms. It is clear from the

results that the per-user rate gap between the ideal perfect-CSI and the proposed AoD-codebook is at a constant level when the Rx SNR increases. From the Figure, the rate gap at 0 dB SNR is approximately 0.18 bps/Hz, whereas, at 15 dB SNR, the proposed AoD-codebook overlaps on the ideal case. This validates our theoretical analysis. The rate gap between the ideal CSI and the conventional-codebook increases with increasing the SNR, which is undesired in practice. Also, the proposed AoD-codebook performs much better than the conventional-codebook.
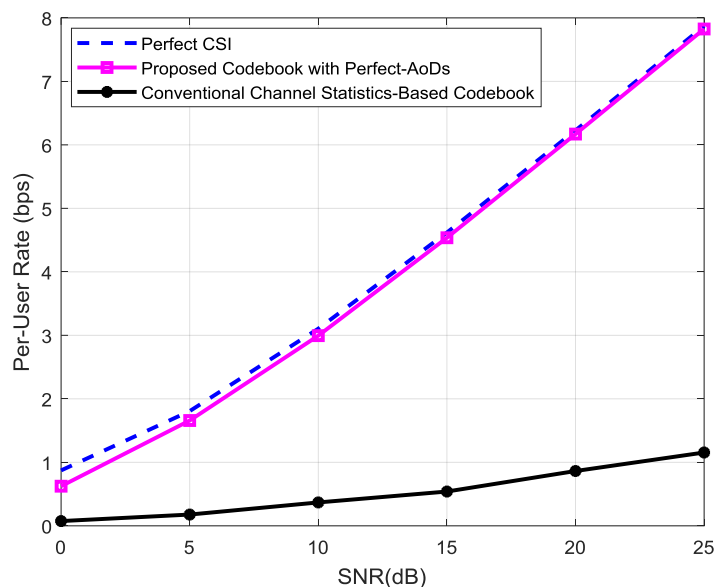

Figure 12. Comparison of the per-user rate of the ideal perfect-CSI and the proposed scheme with quantized-channel feedback in terms of SNR.

Figure 13 shows the NMSE comparison of the proposed S-CoSaMP and the conventional algorithms. The conventional algorithms are represented by dotted-lines while the joint algorithms are represented by solid lines. It is clear from the Figure that the S-CoSaMP based differential joint scheme performs best as compare to other algorithms as it deploys temporal-correlation and sparsity. The proposed algorithm has the least NMSE which means reduced feedback-overhead as compared to the conventional schemes. On the second best is the S-CoSaMP joint algorithm which performs better than the conventional schemes. The JOMP and CoSaMP joint schemes have the same performance as their common support vectors are not satisfied. The Kalman Filter response is not better as it requires the autocorrelation information of the channel. We also observe that the NMSE gap between the conventional algorithms and the proposed algorithms increases with increasing SNR. Thus, it also indicates the dominance of our proposed algorithms.

Figure 14 compares the sum-rate between the ideal perfect-CSI, S-CoSaMP, and the conventional CoSaMP algorithms. We consider two compression ratios (γ) for comparing their performance evaluation. The perfect-CSI case is used for reference purpose with which we compare our proposed results. It is clear from the Figure that our proposed S-CoSaMP based joint differential feedback shows better performance and it is operating near the ideal case of perfect-CSI at $\gamma = 70\%$. The proposed S-CoSaMP at $\gamma = 40\%$ perform better than the conventional CoSaMP which is operating with the overhead of $\gamma = 70\%$ and they have approximately the same data-rate. The conventional CoSaMP fails at the overhead of $\gamma = 40\%$ because the rate gap between this and the ideal CSI gets much larger which is undesired in practice. It shows that our proposed algorithm provides 30% reduction in overhead and improves the channel feedback efficiency as compared to the conventional algorithm in attaining the same sum-rate.
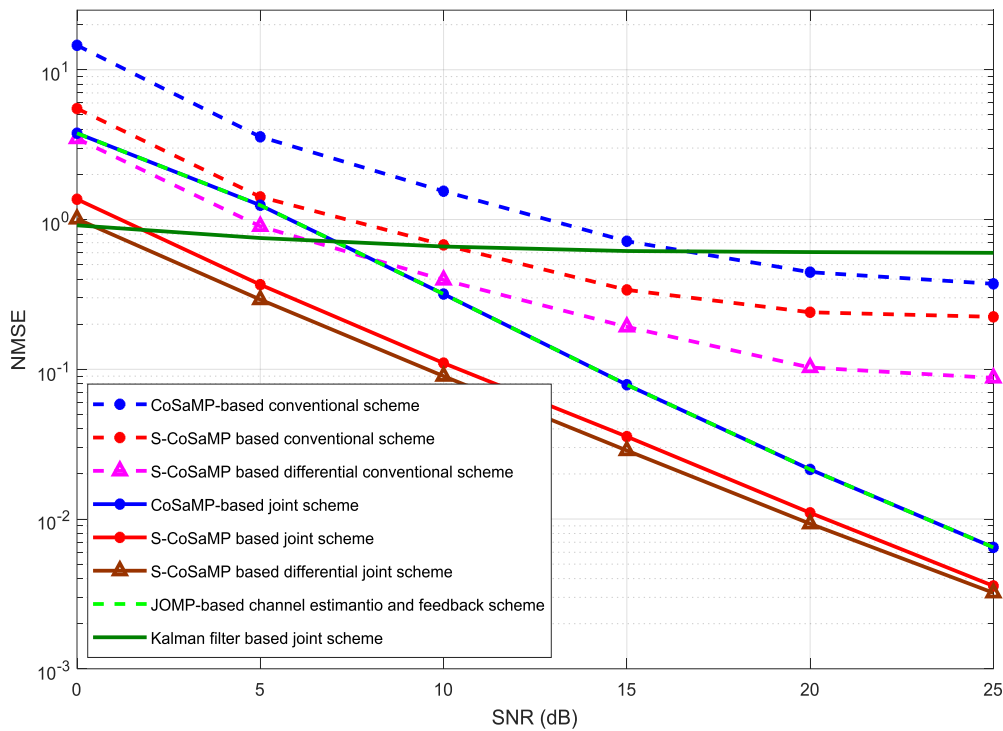
Figure 13. Comparison of the proposed and the conventional channel feedback algorithms in terms of NMSE versus SNR.
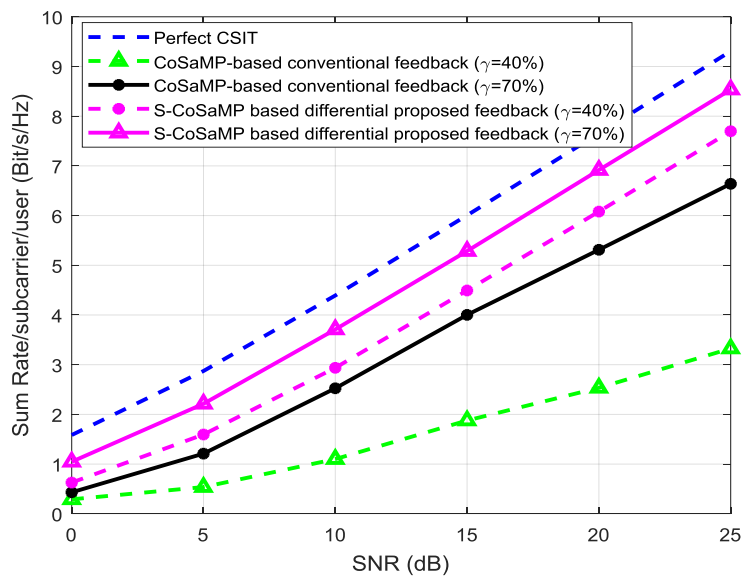

Figure 14. Comparison of sum-rate between the ideal CSI; proposed S-CoSaMP and the conventional CoSaMP algorithms under two-compression ratios.

Figure 15 shows the energy efficiency of the proposed massive MIMO system against the density of BSs. It is clear from the Figure that the EE gap between the upper and lower bounds for all three signal-to-interference ratios (SINRs) ζ is very small and it also decreases as SINR increases. The analytical expression of the SINR is given by:

$$\zeta = \frac{\left(1 - \epsilon^2\right)^2}{\left(1 - \epsilon^2\right)\left(\frac{1}{\beta(\partial - 1)} + \epsilon^2\right)} \tag{48}$$

where ϵ is the level of hardware impairment, β is the optimization variable and ∂ is the power level coefficient respectively. Moreover, the EE is significantly improved with increasing the BS density. This means that by deploying smalls cells, the EE greatly improves which makes it an attractive solution to such issue. Furthermore, the EE decreases as the SINR increases which make a constraint on the EE and for this purpose, a target SINR should be specified from a practical perspective to obtain the required EE.
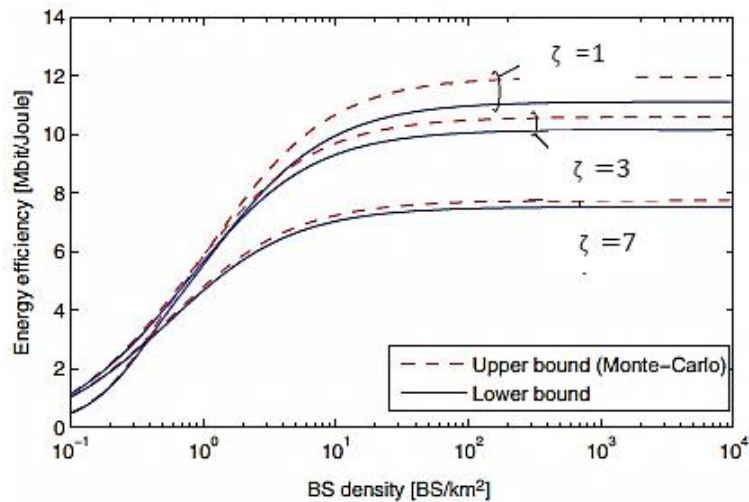

Figure 15. Energy Efficiency versus the BS density comparison of massive MIMO system for different SINRs (ζ).

Figure 16 illustrates the EE against the UE density for the average SINR (ζ) of 3 dB respectively. It is compared with two references usage cases, such as single-input multiple outputs (SIMO) and MIMO for performance evaluation. It is clear from the results that the EE becomes independent as the UE density increases to a large level. We can see that the EE saturation levels occurs at UE density of 100 and above respectively for the optimized and MIMO cases. The saturation for SIMO case occurs at UE of 2 and above. Moreover, we can observe that the proposed system shows optimized performance than the other two competing cases which makes it suitable for practical usage scenarios of small cells dense networks.
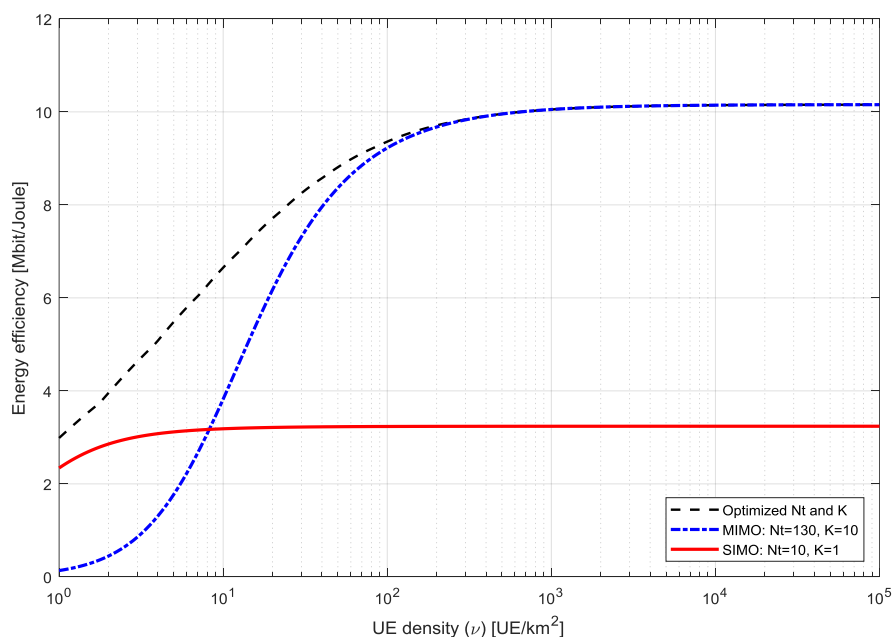

Figure 16. Energy Efficiency versus the UE density comparison of massive MIMO system.

Figure 17 shows the EE levels against the level of hardware impairments for three different SINR levels. It is clear from the results that the EE decreases with increasing hardware impairments. It is due to the decay of the desired signal power.
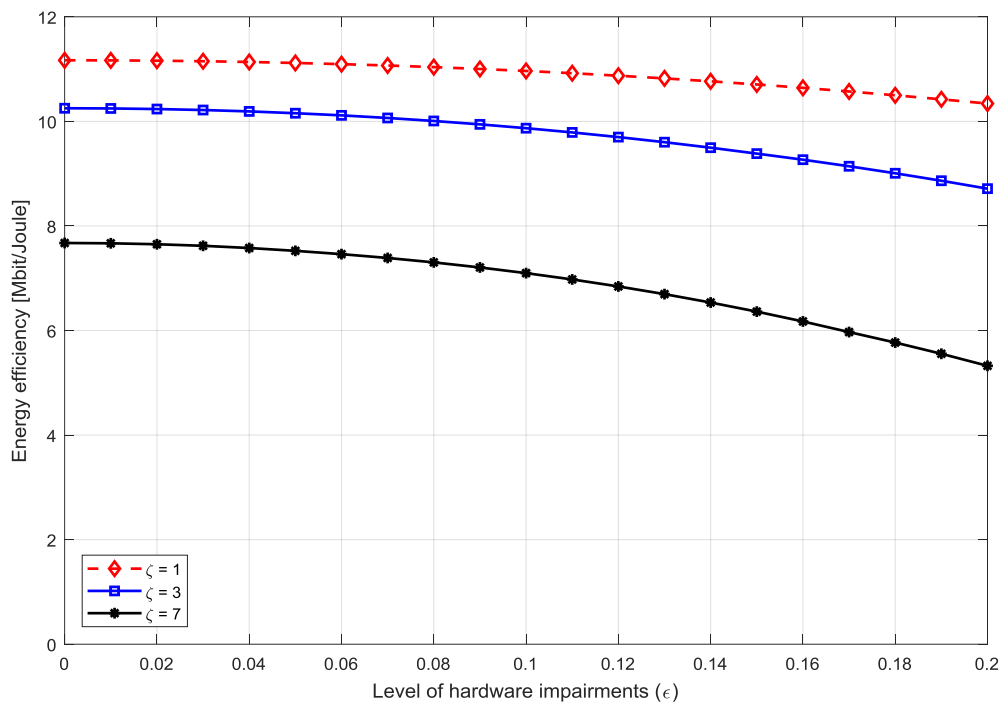

Figure 17. Energy Efficiency versus the level of hardware impairment of massive MIMO system for different SINRs.

The loss can be high when the SINR increases to a large level which means that the hardware impairment substantially affects the EE in high SINR levels. Moreover, the EE loss is negligible for hardware impairment of level of 0.1 and below. Therefore, it can be concluded that for the moderate hardware impairment level, the effect on the system under consideration is negligible.

5. Conclusions

This paper scrutinizes the feedback and pilot overhead issues of uplink and downlink channel estimation in massive MIMO systems. It also investigates the EE evaluation in terms of BS density, UE density, and hardware impairment levels respectively. For the UL-channel estimation, it is concluded that by using temporal-correlation of massive MIMO time-varying channels which produces the compression-efficient differential CIR has 25% less overhead requirement or in other words, 25% more effective as compare to the direct CS (DCS) algorithm. It is also shown that the proposed CS algorithm has improved NMSE than the DCS algorithm. For the DL-channel estimation, the proposed Block-ISD scheme shows enhanced performance as compare to classical-ISD and BP schemes. It utilizes spatial-correlation of massive MIMO channels which produces block-sparse CIR that have more compression-efficiency and due to such mechanism, the pilot-overhead is decreased by 84.6%. In other words, the efficiency of the Block-ISD is 84.6% more than the classical-ISD or BP techniques. The AoD-adaptive subspace codebook provides a substantial reduction in the training and feedback overhead as well as the codebook size which results in improved data-rate and bandwidth-efficiency. It uses the idea that the time-based variation in channels path-AoDs is much slower than path-gains. The proposed results are so close to the theoretical analysis in terms of per-user data-rate and feedback bits. As we have proved that the feedback bits vary linearly with the number of paths P rather than the whole N-dimensional space which provides freedom from the large BS antenna-based feedback

bits scaling. The S-CoSaMP algorithm is used for joint channel training and feedback which is based on CS that further reduces the channel overhead. It has reduced the joint UL/DL overhead by 30%. It also provides the benefit of getting rid of the burden on the UE which estimate their channel and then feedback to BS. The UE directly feedback their pilots to BS and then the CS algorithm obtain the required CSI at the BS. The EE levels of the proposed system outperform the SIMO and MIMO usage cases which makes it suitable for small cells dense networks operation. Also, the EE levels are optimum in upper and lower bounds which are defined using three different SINR level of 1, 3 and 7, respectively. The EE gap between both bounds is close enough which decreases further with increasing SINR levels. We also conclude that the EE level decreases as the SINR levels increases which provide information to specify a target SINR for each considered usage case in practical perspective. Therefore, the EE increases with increasing BS density, UE density, and lower hardware impairment levels respectively. Our future work is focused on reducing the channel overhead for which we will utilize the spatial-correlation of CSI from various UEs at the BS. Also, the quantization error of AoD in terms of feedback bits will be evaluated and analyzed.

Author Contributions: Imran Khan conceived and designed the presented idea and developed the theory, performed the simulations and wrote the paper; Mohammad Haseeb Zafar provides extensive support in the overall research and verified the analytical methods and encouraged to investigate various relevant aspects of the proposed research; Mohammad Tariq Jan analyzed the data; Jaime Lloret revised the article with critical revision and technical support; Mohammad Basheri verified the experimental results; Dhananjay Singh provide critical feedback and helped shape the research, analysis and manuscript. All authors discussed the results and contributed to the final manuscript.

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

References

1. Krishna, M.B.; Lloret, J. Evolution toward 5G Networks. In *Advances in Mobile Computing and Communications: Perspectives and Emerging Trends in 5G Networks*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2016; ISBN 9781498701136-CAT# K24551.
2. Lloret, J.; Parra, L.; Taha, M.; Tomás, J. An architecture and protocol for smart continuous eHealth monitoring using 5G. *Comput. Netw.* **2017**. [CrossRef]
3. Mansoor, B.; Nawaz, S.J.; Gulfam, S.M. Massive-MIMO Sparse Uplink Channel Estimation Using Implicit Training and Compressed Sensing. *Appl. Sci.* **2017**, *7*, 63. [CrossRef]
4. Najm, I.A.; Ismail, M.; Lloret, J.; Ghafoor, K.Z.; Zaidan, B.B.; Rahem, A.A.T. Improvement of SCTP congestion control in the LTE-A network. *J. Netw. Comput. Appl.* **2015**, *58*, 119–129. [CrossRef]
5. Taha, M.; Parra, L.; Garcia, L.; Lloret, J. An Intelligent handover process algorithm in 5G networks: The use case of mobile cameras for environmental surveillance. In Proceedings of the IEEE International Conference Communications Workshops (ICC Workshops 2017), Paris, France, 21–25 May 2017; pp. 840–844.
6. Larsson, E.; Edfors, O.; Tufvesson, F.; Marzetta, T. Massive MIMO for next-generation wireless systems. *IEEE Commun. Mag.* **2014**, *52*, 186–195. [CrossRef]
7. Gao, Z.; Dai, L.; Mi, D.; Wang, Z.; Imran, M.A.; Shakir, M.Z. MmWave massive-MIMO-based wireless backhaul for the 5G ultra-dense network. *IEEE Wirel. Commun.* **2015**, *22*, 13–21. [CrossRef]
8. Rusek, F.; Persson, D.; Lau, B.; Larsson, E.G.; Marzetta, T.L. Scaling up MIMO: Opportunities and challenges with very large arrays. *IEEE Signal Process. Mag.* **2013**, *30*, 44–60. [CrossRef]
9. Al Salehi, H.; Nakhai, M.R. Efficient Bayesian compressed sensing-based channel estimation techniques for massive MIMO-OFDM systems. *EURASIP J. Wirel. Commun. Netw.* **2017**, *38*, 1–10. [CrossRef]
10. Jungnickel, V.; Manolakis, K.; Zirwas, W.; Panzner, B.; Braun, V.; Lossow, M.; Sternad, M.; Apelfrojd, R.; Svensson, T. The role of small cells, coordinated multipoint, and massive MIMO in 5G. *IEEE Commun. Mag.* **2014**, *52*, 44–51. [CrossRef]
11. Zhang, H.; Gao, S.; Li, D.; Chen, H.; Yang, L. On superimposed pilot for channel estimation in multi-cell multiuser MIMO uplink: Large system analysis. *IEEE Trans. Veh. Technol.* **2015**, *65*, 1492–1505. [CrossRef]
12. Soltanalian, M.; Nagash, M.M.; Shariati, N.; Stocia, P.; Hassibi, B. Training Signal Design for correlated Massive MIMO Channel Estimation. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 1135–1143. [CrossRef]

13.　Noh, S.; Zoltowski, M.D.; Sung, Y.; David, J.L. Pilot beam pattern design for channel estimation in massive MIMO systems. *IEEE J. Sel. Top. Signal Process.* **2014**, *8*, 787–801. [CrossRef]

14.　Peken, T.; Vanhoy, G.; Bose, T. Blind estimation for massive MIMO. *Analog IC Signal Process.* **2017**, *91*, 257–266. [CrossRef]

15.　Mawatwal, K.; Sen, D.; Roy, R. An iterative SAGE-based semi-blind channel estimation for massive MIMO. In Proceedings of the IEEE Conference on GLOBECOM, Washington, DC, USA, 4–8 December 2016.

16.　Fan, D.; Gao, F.; Wang, G.; Zhong, Z.; Nallanathan, A. A practical channel estimation scheme for indoor 60GHz massive MIMO systems via array signal processing. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017; pp. 1–6.

17.　Gonzalez-Prelcic, N.; Truong, K.T.; Rusu, C.; Heath, R.W. Compressive Channel Estimation in FDD Multi-Cell Massive MIMO Systems with Arbitrary Arrays. In Proceedings of the Globecom Workshops (GC Wkshps), Washington, DC, USA, 4–8 December 2016.

18.　Yan, H.; Cabria, D. Compressive sensing based initial beamforming training for massive MIMO millimeter-wave systems. In Proceedings of the 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Washington, DC, USA, 7–9 December 2016.

19.　Lee, J.; Gil, G.-T.; Lee, Y.H. Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications. *IEEE Trans. Commun.* **2014**, *64*, 2370–2386. [CrossRef]

20.　Sure, P.; Bhuma, C.M. On the Comparison of Various Overhead Arrangements for Massive MIMO-OFDM Channel Estimation. *Int. J. Electron. Telecommun.* **2014**, *60*, 173–179.

21.　Shen, W.; Dai, L.; Gao, Z.; Wang, Z. Spatially correlated channel estimation based on block iterative support detection for large-scale MIMO. *Electron. Lett.* **2015**, *51*, 587–588. [CrossRef]

22.　Zhu, D.; Choi, J.; Health, R.W. Auxilliary Beam Pair Enabled AoD and AoA Estimation in Closed-Loop Large-Scale Millimeter-Wave MIMO Systems. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 4770–4785. [CrossRef]

23.　Shen, W.; Wang, B.; Feng, J.; Gao, C.; Ma, J. Differential CSIT Acquisition Based on Compressive Sensing for FDD Massive MIMO Systems. In Proceedings of the 81st Vehicular Technology Conference (VTC Spring), Glasgow, UK, 1–14 May 2015; pp. 1–4.

24.　Ziniel, J.; Schniter, P. Dynamic compressive sensing of time-varying signals via approximate message passing. *IEEE Trans. Signal Process.* **2013**, *61*, 5270–5284. [CrossRef]

25.　Qi, C.; Yue, G.; Wu, L.; Arumugam, N. Pilot design for sparse channel estimation in OFDM-based cognitive radio systems. *IEEE Trans. Veh. Technol.* **2014**, *63*, 982–987. [CrossRef]

26.　Dai, W.; Milenkovic, O. Subspace pursuit for compressive sensing signal reconstruction. *IEEE Trans. Inf. Theory* **2009**, *55*, 2230–2249. [CrossRef]

27.　Dai, L.; Wang, Z.; Yang, Z. Compressive sensing-based time domain synchronous OFDM transmission for vehicular communications. *IEEE J. Sel. Areas Commun.* **2013**, *32*, 460–469. [CrossRef]

28.　Gui, G.; Adachi, F. Stable adaptive sparse filtering algorithms for estimating multiple-input-multiple-output channels. *IET Commun.* **2014**, *8*, 1032–1040. [CrossRef]

29.　Barbotin, Y.; Hormati, A.; Rangan, S.; Vetterli, M. Estimation of sparse MIMO channels with common support. *IEEE Trans. Commun.* **2012**, *60*, 3705–3716. [CrossRef]

30.　Yu, K.; Ottersten, B. Models for MIMO propagation channels: A review. *Wirel. Commun. Mob. Comput.* **2002**, *2*, 653–666. [CrossRef]

31.　Tsai, Y.; Zheng, L.; Wang, X. Millimeter-Wave Beamformed Full-dimension MIMO Channel Estimation Based on Automatic Norm Minimization. *arXiv*, 2017.

32.　Tse, D.; Viswanath, P. The capacity of wireless channels. In *Fundamentals of Wireless Communication*; Cambridge University Press: Cambridge, UK, 2005; pp. 166–227. ISBN 9780521845274.

33.　Abdelrahman, O.M.; Prajapati, P.P. Analysis of precoding algorithms and pilot contamination problem for massive MIMO System-Review. *Int. J. Innov. Res. Comput. Commun. Eng.* **2017**, *5*, 5266–5271. [CrossRef]

34.　Duarte, M.; Eldar, Y. Structured compressed sensing: From theory to applications. *IEEE Trans. Signal Process.* **2011**, *59*, 4053–4085. [CrossRef]

35. Ozdemir, M.; Arslan, H. Channel estimation for wireless OFDM systems. *IEEE Commun. Surv. Tutor.* **2007**, *9*, 18–48. [CrossRef]

36. Bjornson, E.; Sanguinetti, L.; Kountouris, M. Designing wireless broadband access for energy efficiency: Are small cells the only answer? In Proceedings of the IEEE International Conference on Communications (ICC), London, UK, 8–12 June 2015; pp. 136–141.