Article

Massive-MIMO Sparse Uplink Channel Estimation Using Implicit Training and Compressed Sensing

Babar Mansoor *, Syed Junaid Nawaz and Sardar Muhammad Gulfam

Department of Electrical Engineering, COMSATS Institute of Information Technology, Islamabad 44000, Pakistan; junaidnawaz@ieee.org.com (S.J.N.); sardar_muhammad@comsats.edu.pk (S.M.G.)
* Correspondence: babar_mansoor@comsats.edu.pk; Tel.: +92-51-9049-122

Abstract: Massive multiple-input multiple-output (massive-MIMO) is foreseen as a potential technology for future 5G cellular communication networks due to its substantial benefits in terms of increased spectral and energy efficiency. These advantages of massive-MIMO are a consequence of equipping the base station (BS) with quite a large number of antenna elements, thus resulting in an aggressive spatial multiplexing. In order to effectively reap the benefits of massive-MIMO, an adequate estimate of the channel impulse response (CIR) between each transmit–receive link is of utmost importance. It has been established in the literature that certain specific multipath propagation environments lead to a sparse structured CIR in spatial and/or delay domains. In this paper, implicit training and compressed sensing based CIR estimation techniques are proposed for the case of massive-MIMO sparse uplink channels. In the proposed superimposed training (SiT) based techniques, a periodic and low power training sequence is superimposed (arithmetically added) over the information sequence, thus avoiding any dedicated time/frequency slots for the training sequence. For the estimation of such massive-MIMO sparse uplink channels, two greedy pursuits based compressed sensing approaches are proposed, viz: SiT based stage-wise orthogonal matching pursuit (SiT-StOMP) and gradient pursuit (SiT-GP). In order to demonstrate the validity of proposed techniques, a performance comparison in terms of normalized mean square error (NCMSE) and bit error rate (BER) is performed with a notable SiT based least squares (SiT-LS) channel estimation technique. The effect of channels' sparsity, training-to-information power ratio (TIR) and signal-to-noise ratio (SNR) on BER and NCMSE performance of proposed schemes is thoroughly studied. For a simulation scenario of: 4×64 massive-MIMO with a channel sparsity level of 80% and signal-to-noise ratio (SNR) of 10 dB, a performance gain of 18 dB and 13 dB in terms of NCMSE over SiT-LS is observed for the proposed SiT-StOMP and SiT-GP techniques, respectively. Moreover, a performance gain of about 3 dB and 2.5 dB in SNR is achieved by the proposed SiT-StOMP and SiT-GP, respectively, for a BER of 10^{-2}, as compared to SiT-LS. This performance gain NCME and BER is observed to further increase with an increase in channels' sparsity.

Keywords: massive MIMO; superimposed training; compressed sensing; estimation; sparse channel; 5G communications

1. Introduction

Today's modern cellular communication networks are witnessing an ever increasing demand for higher data rates and link reliability. A major hindrance in meeting these demands is scarcity of the electromagnetic spectrum. One of the key solutions to effectively deal with these challenges is to spatially reuse the available spectrum. This is achieved by employing multiple antennas at the access point and/or user equipment (UE), thus enabling the propagation channel to provide more degrees of freedom. Such multiple-input multiple-output (MIMO) communication systems are

generally categorized into point-to-point MIMO or single-user MIMO [1,2] and multi-user MIMO (MU-MIMO) [3,4]. MIMO systems make use of spatial multiplexing to provide an order of increase in the capacity while achieving link reliability by employing space-time coding techniques. MU-MIMO has been implemented in many wireless standards such as 802.11 (WiFi), 802.16 (WiMax), and long-term evolution (LTE), e.g., LTE-Advanced allows up to eight antenna ports at the base station (BS). Moreover, it is envisioned that the future 5G cellular networks will increase the system capacity by a thousand folds and energy efficiency by hundred folds, in addition to decreasing the system latency by ten folds [5].

Massive-MIMO, however, in contrast to the conventional MIMO, is a MU-MIMO technology that employs a large number (on the order of 100 or more) of physically small, low power and independently-controlled antennas at the BS in order to serve a number of single-antenna terminals using the same time and frequency resources [6]. Equipping the BS with a large number of antennas results in manifolds increase in spectral and energy efficiency as compared to conventional MIMO systems. The increase in spectral efficiency is a result of aggressive spatial multiplexing to transmit data streams for a desired user [7]. Whereas the increase in energy efficiency is a consequence of producing sharp beams with the help large number of antennas at the BS, such energy is focused into smaller spatial regions i.e., for intended users only [7]. Furthermore, employing a large number of antennas at the BS results in a favorable propagation channel because the channel vectors and users become pair-wise orthogonal; thus, linear processing becomes optimal [8]. Moreover, in order to benefit from a large amount of under-utilized millimeter wave (mmWave) spectra, an mmWave based massive-MIMO system has also been proposed for the backhaul of future 5G ultra-dense networks (UDN) [5]. Therefore, massive-MIMO is foreseen as a potential technology for future 5G cellular communication networks.

It has been established in literature that the channel impulse response CIR in several wireless communication scenarios tends to exhibit a sparse structure that gets pronounced with growing signal dimensions [9,10]. For example, a sparse structured CIR is exhibited by underwater acoustic communication channels [11]; wideband high frequency channels [12]; high-definition television (HDTV) channels [13], and cellular communication channels, where spatially spread distant scatterers correspond to the arrival of signals [14]. It has been demonstrated in [15] that in propagation environments with an insufficient number of scatterers, the physical MIMO channels tend to exhibit a sparse structure in CIR. Furthermore, in [16], it has been shown that in MIMO systems the time of arrival (ToA) at different antennas is similar. Therefore, the different uplink channels exhibit a CIR that tends to possess a common support. Moreover, MIMO communication channels exhibit joint sparsity across the channel components because of the smaller antenna spacings as compared to the propagation paths' lengths [17]. In massive-MIMO, several experimental studies have established the fact that CIR tends to be jointly sparse due to the shared common local scattering clusters [18]. In [19], authors have explored the joint sparsity of massive-MIMO channels in virtual angular domains.

To reap the benefits of massive-MIMO, adequate knowledge of the CIR between each transmit–receive link is required. In a typical massive-MIMO system, the BS is equipped with a large number of antennas. Therefore, a large number of channels needs to be estimated. Thus, CIR estimation in massive-MIMO systems is quite a challenging task due to high dimensionality of massive-MIMO channels. Several CIR estimation techniques have been proposed in literature for the case of MIMO systems that can be broadly categorized into blind, e.g., [20,21], semiblind, e.g., [22,23], training/pilot based, e.g., [24,25], and superimposed training (SiT), e.g., [14,26], based techniques.

For the training based channel estimation techniques, a known training sequence, also known as a pilot sequence, is transmitted along with a block of transmit symbols. At the receiver side, the channel is estimated by using the received training sequence and the known training sequence. The drawback of training based channel estimation techniques is that a significant portion of the channel capacity is consumed by the training sequence. On the other hand, blind channel estimation schemes use the statistical parameters of the received signal to estimate the CIR. However, blind estimation techniques

are computationally complex and have slow convergence rate. Semi-blind methods use some training symbols along with the statistics of received signal to estimate the CIR. Semi-blind channel estimation techniques share the benefits and trade-offs of the blind and training based estimation techniques. Recently, SiT based techniques have gained significant attention for the purpose of channel estimation. In an SiT based technique, a low power and periodic training sequence is superimposed over the information sequence and transmitted. Consequently, SiT based techniques avoid the allocation of any dedicated time/frequency slots for the training sequence.

Over the past few years, compressed sensing (CS) has emerged as a new paradigm for the recovery of sparse signals. The authors in [27,28] have established the fact that a finite-dimensional sparse signal can be exactly reconstructed from fewer, linear and nonadaptive measurements. The CS approach has been established as an efficient solution to estimate sparse multipath channels—see e.g., [14,29]. Computing the sparse solution requires solving a ℓ_0-minimization problem, which is computationally non-deterministic polynomial-time (*NP*) hard. An alternative approach is to relax the ℓ_0-minimization to ℓ_1-minimization and solve the problem for a sparse solution. This approach is termed as Basis Pursuits (BP) [30]. A BP based approach achieves more accurate solutions but requires higher computational complexity. Another approach used to find the sparse solution of an underdetermined system of linear equations is that of greedy algorithms, which iteratively approximate the sparse signal by suitably choosing the columns from the sparse measurement matrix. In literature, a wide variety of the greedy algorithms have been proposed to solve the CS problem such as matching pursuit (MP) [31], orthogonal matching pursuit (OMP) [32], regularized orthogonal matching pursuit (ROMP) [33], stagewise orthogonal matching pursuit (StOMP) [34], orthogonal complementary matching pursuit (OCMP) [35], compressive sampling matching pursuit (CoSaMP) [36], and gradient pursuit (GP) [37].

Several massive-MIMO channel estimation techniques have been proposed in the literature—see e.g., [23,38] and references therein. In order to estimate CIR in massive-MIMO systems, linear minimum mean-squared error (LMMSE) estimation techniques are usually used. Several CS based techniques have also been proposed to estimate the CIR in massive-MIMO (see e.g., [39,40]). In [41], authors have proposed a low complexity polynomial channel estimation based on Bayesian channel estimators to estimate CIR in massive-MIMO. An estimation technique for an uplink channel of massive-MIMO has been proposed in [42] that exploits the joint sparsity of the channels in a massive-MIMO system. In [43], authors have devised a Gaussian-mixture Bayesian learning based channel estimation technique for massive-MIMO. In [44], authors have devised a channel estimation technique that requires a small number of pilot sequences and exploits the sparsity and common support property of massive-MIMO communication channels. In [45], authors have exploited CS along with random linear network coding (RLNC) to devise an energy efficient scheme for vital signal telemonitoring in wireless body area networks (WBAN).

By exploiting the sparsity of wireless multipath channels, SiT sequence based compressive channel sensing methods have been studied in various contexts such as single-input single-output (SISO) systems [14,46], sparse MIMO channels [47,48], and underwater acoustic channels [49]. In [46], a genetic algorithm (GA) based channel estimation method is proposed using an SiT sequence for SISO systems. In [14], a Dantzig selector (DS) algorithm based method is proposed for estimation of SISO sparse multipath channels using a known SiT sequence. This study is further extended in [47,48] for the case of multiuser MIMO systems, where SiT based DS and MP algorithms are proposed.

To the best of the authors' knowledge, no study has considered the use of CS based greedy algorithms in conjunction with an SiT sequence for estimation of massive-MIMO sparse uplink channels. The prime motivation for using a superimposed training based channel estimation approach is its improved spectral efficiency as compared to the conventional training based approaches. In this regard, the main contributions of this paper is to devise an SiT sequence based estimation of massive-MIMO sparse uplink channels by exploiting StOMP and GP as CS greedy algorithms. For large-scale sparse reconstruction problems, it has been established in [34] that the StOMP algorithm

achieves better performance as compared to MP and OMP. Similarly, the GP algorithm enhances the performance (in terms of faster computations) of the OMP algorithm by adopting a directional gradient pursuit based approach [37]. As the CIR estimation of massive-MIMO sparse uplink channels involves a large number of channel coefficients, the choice of StOMP and GP as sparse reconstruction algorithms is more favorable for this purpose.

The rest of the paper is structured as follows. In Section 2, the system model under consideration for multiuser massive-MIMO uplink communications is presented. Section 3 presents first-order statistics of the received signals along with the proposed channel estimation techniques. In Section 3.1, an SiT least squares (SiT-LS) based channel estimation technique available in literature [26] is presented. Section 3.2 presents the proposed SiT-StOMP and SiT-GP channel estimation techniques for the case of sparse uplink channels in massive-MIMO. Section 4 discusses a minimum mean square error (MMSE) equalizer for a massive-MIMO uplink communication scenario. Section 5 presents the simulation results and a performance analysis for the proposed channel estimation techniques. In Section 6, conclusions are presented.

Notations: Matrices are denoted by boldface uppercase letters, e.g., $\mathbf{X}$, while boldface lowercase letters are used to represent vectors, e.g., $\mathbf{x}$. Small case letters, e.g., x are used to denote scalar quantities. Transpose and Hermitian transpose are represented by the superscripts T and H, respectively. Γ^i represents the set of indices of the elements selected up to and including iteration i. The matrix $\mathbf{X}_{\Gamma^i}$ represents a sub-matrix of $\mathbf{X}$ consisting of only those columns of $\mathbf{X}$ whose indices belong to the set Γ^i. The same convention is also followed for the case of vectors.

2. Massive-MIMO System Model for Uplink Communications

The uplink communications system model for massive-MIMO is assumed to consist of N_t single antenna UEs communicating with N_r antennas at the BS, as depicted in Figure 1. The information sequence to be transmitted by n_t-th UE is represented by $\mathbf{b}_{n_t} = [b_{n_t}(0), b_{n_t}(1), \cdots, b_{n_t}(M-1)]^T$, such that $\mathbf{b}_{n_t}$ is zero-mean, statistically independent of other UEs, and with $E\{|b_{n_t}(k)|^2\} = 1$. Let $\mathbf{c}_{n_t} = [c_{n_t}(0), c_{n_t}(1), \cdots, c_{n_t}(M-1)]^T$ represent a low power and periodic training sequence for n_t-th UE. The training sequence $\mathbf{c}_{n_t}$ is periodic with period P such that $c_{n_t}(k) = c_{n_t}(k+aP)$, for k and a being any integers, and is assumed to also be known at the BS side. After superimposing the training sequence $\mathbf{c}_{n_t}$ over the information sequence, $\mathbf{b}_{n_t}$, the resultant transmit signal for n_t-th UE is given as

$$\mathbf{x}_{n_t} = \mathbf{b}_{n_t} + \mathbf{c}_{n_t}. \tag{1}$$

The sequence $\mathbf{x}_{n_t} = [x_{n_t}(0), x_{n_t}(1), \cdots, x_{n_t}(M-1)]^T$ is then transmitted by the n_t-th UE over the massive-MIMO sparse uplink channel. Between n_t-th UE and n_r-th receive antenna at the BS, a frequency selective and time invariant channel is assumed with CIR denoted by a sparse vector $\mathbf{h}_{n_r n_t} = [h_{n_r n_t}^0, h_{n_r n_t}^1, \cdots, h_{n_r n_t}^{L-1}]^T$. The number of resolvable mutipaths is represented by L. Moreover, the CIR vector $\mathbf{h}_{n_r n_t}$ is assumed to be Q sparse i.e., $\{Q = \|\mathbf{h}_{nm}\|_{\ell_0}\} \ll L$, and has support $\check{\mathbf{p}} = [\check{p}_0, \check{p}_1, \cdots, \check{p}_{Q-1}]$ such that

$$h_{n_r n_t}^{\ell} = \begin{cases} \neq 0 & ; \ell \in \check{\mathbf{p}}; \\ = 0 & ; \text{otherwise.} \end{cases} \tag{2}$$

The signal received at time instant k by the n_r-th antenna element of receiver array at the BS is given below,

$$y_{n_r}(k) = \sum_{n_t=1}^{N_t} \sum_{\ell=0}^{L-1} h_{n_r n_t}^{\ell} x_{n_t}(k-\ell) + z_{n_r}(k), \tag{3}$$

where $z_{n_r}(k)$ denotes kth sample of zero-mean, additive white Gaussian noise (AWGN) with variance σ_z^2. At time instant k, the overall combined received signal by all antennas at the BS is denoted by $\mathbf{y}(k) = [y_1(k), y_2(k), \cdots, y_{N_t}(k)]^T$ and given by,

$$\mathbf{y}(k) = \sum_{\ell=0}^{L-1} \mathbf{H}^\ell \mathbf{x}(k-\ell) + \mathbf{z}(k), \tag{4}$$

where $\mathbf{x}(k-\ell) = [x_1(k-\ell), x_2(k-\ell), \cdots, x_{N_t}(k-\ell)]^T$, $\mathbf{z}(k) = [z_1(k), z_2(k), \cdots, z_{N_r}(k)]^T$ and the $N_r \times N_t$ channel matrix $\mathbf{H}^\ell$ for ℓth tap delay is given by,

$$\mathbf{H}^\ell = \begin{bmatrix} h_{11}^\ell & h_{12}^\ell & \cdots & h_{1N_t}^\ell \\ h_{21}^\ell & h_{22}^\ell & \cdots & h_{2N_t}^\ell \\ \vdots & \cdots & \ddots & \vdots \\ h_{N_r1}^\ell & h_{N_r2}^\ell & \cdots & h_{N_rN_t}^\ell \end{bmatrix}. \tag{5}$$
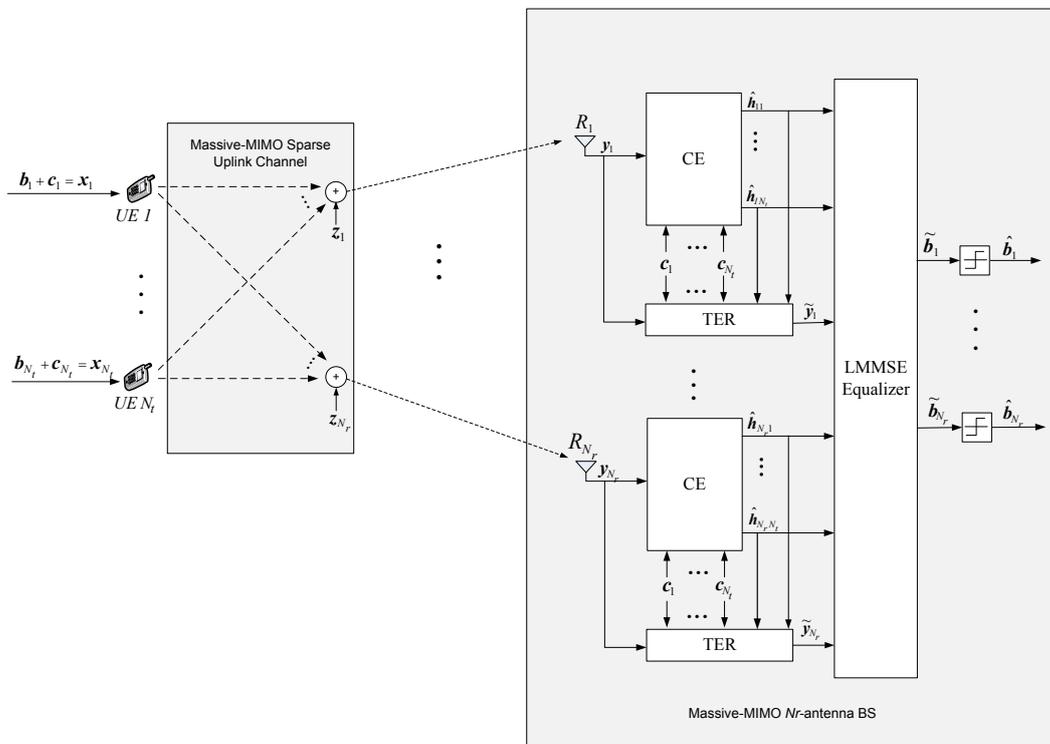
Figure 1. Proposed system model for massive multiple-input multiple-output (massive-MIMO) uplink communications.

The channel estimator (CE) block, shown in Figure 1, is implemented by using the SiT and first-order statistics based least squares (SiT-LS) technique presented in [50] and the proposed SiT-StOMP and SiT-GP techniques. Once the CIR is estimated, the effect of superimposed training sequence is removed from the information sequence at the receiver side by the training effect remover (TER) block, as shown in Figure 1. After eliminating the superimposed training sequence effect, the output of TER is fed as an input to the LMMSE equalizer.

3. Massive-MIMO Sparse Uplink Channel Estimation Using a First-Order Statistics Based Approach

For the uplink scenario of a massive-MIMO communication system, if each mobile user is assigned with a specific pilot sequence that is superimposed on the information sequence, a first order statistics of the received signal can be used to estimate the CIR as outlined in [50]. In this section, an SiT based channel estimation technique of [26] is extended for the estimation of massive-MIMO sparse uplink channels. For n_t-th transmitter, the training sequence $c_{n_t}(k)$ is periodic with period $P = \tilde{P}N_t$, where $\tilde{P}$ is a positive integer. The training sequence $c_{n_t}(k)$ is given as,

$$c_{n_t}(k) = \sum_{i=0}^{P-1} c_{i,n_t} \, e^{j(2\pi i/P)k}, \qquad \forall k, \tag{6}$$

where $j = \sqrt{-1}$ and

$$c_{i,n_t} = \frac{1}{P} \sum_{k=0}^{P-1} c_{n_t} \, e^{-j(2\pi i/P)k}. \tag{7}$$

The pilot sequence $c_{n_t}(k)$ is selected such that only $\tilde{P}$ coefficients among P are non-zero. Hence, the training sequence $c_{n_t}(k)$ for n_t-th UE is given by,

$$c_{n_t}(k) = \sum_{i=0}^{\tilde{P}-1} c'_{i,n_t} \, e^{j\alpha_{i,n_t}k}, \qquad \forall k, \tag{8}$$

where $\alpha_{i,n_t} = 2\pi(iN_t + n_t - 1)/P$, and c'_{i,n_t} are suitably chosen coefficients for $1 \leq n_t \leq N_t$ and $0 \leq i \leq \tilde{P} - 1$. The training sequence $c_{n_t}(k)$ can be designed by first choosing a periodic base sequence $\bar{c}_o(k)$ with a period of $\tilde{P}$ [26] such that,

$$\bar{c}_{i,o} = \frac{1}{\tilde{P}} \sum_{k=0}^{\tilde{P}-1} \bar{c}_o(k) \, e^{-j(2\pi i/\tilde{P})k}. \tag{9}$$

The training sequence $\bar{c}_1(n_t)$ with period P is then generated by replicating $\bar{c}_o(k)$ for N_t times. Therefore, for n_t-th UE, the training sequence is obtained as [26]

$$c_{n_t}(k) = \sigma_{cn_t} \, \bar{c}_1(k) \, e^{j(2\pi/P)(n_t-1)k} \text{ for } n_t = 1, 2, \cdots, N_t. \tag{10}$$

Taking the expected value of the received signal $y_{n_r}(k)$ at n_r-th receive antenna gives

$$E\{y_{n_r}(k)\} = \sum_{n_t=1}^{N_t} \sum_{i=0}^{\tilde{P}-1} \left[\sum_{\ell=0}^{L} c'_{i,n_t} h^\ell_{n_r n_t} e^{-j\alpha_{i,n_t}\ell} \right] e^{j\alpha_{i,n_t}k}. \tag{11}$$

For $n_1 \neq n_2$, we have $\alpha_{i_1,n_1} \neq \alpha_{i_2,n_2}$ for any $\{i_1, i_2\} \in 0, 1, ..., \tilde{P} - 1$. Let $\mathbf{d}_{n_r n_t} = [d_{n_r n_t,0}, d_{n_r n_t,1}, \cdots, d_{n_r n_t,(\tilde{P}-1)}]^T$, where $d_{n_r n_t,i}$ is given by

$$d_{n_r n_t,i} = \sum_{\ell=0}^{L} c'_{i,n_t} h^\ell_{n_r n_t} e^{-j\alpha_{i,n_t}\ell}. \tag{12}$$

The mean square consistent estimate $\hat{\mathbf{d}}_{n_r n_t} = [\hat{d}_{n_r n_t,0}, \hat{d}_{n_r n_t,1}, \cdots, \hat{d}_{n_r n_t,(\tilde{P}-1)}]^T$ of $\mathbf{d}_{n_r n_t}$ is computed as in [26] and is given by

$$\hat{d}_{n_r n_t,i} = \frac{1}{M} \sum_{k=0}^{M-1} y_{n_r}(k) \, e^{-j\alpha_{i,n_t}k}, \tag{13}$$

where M is the total number of received symbols, as $M \to \infty$, $\hat{d}_{n_r n_t, i} \to d_{n_r n_t, i}$. The vector form of (13) is given below,

$$\hat{\mathbf{d}}_{n_r n_t} = \mathbf{C}_{n_t} \mathbf{h}_{n_r n_t}, \tag{14}$$

where the matrix $\mathbf{C}_{n_t}$ is computed as

$$\mathbf{C}_{n_t} = diag\,\{c'_{0,n_t}, c'_{1,n_t}, \cdots, c'_{(\tilde{P}-1),n_t}\} \mathbf{V}_{n_t}, \tag{15}$$

where $\mathbf{V}_{n_t}$ is the Vandermonde matrix given by

$$\mathbf{V}_{n_t} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-j\alpha_{1,n_t}} & \cdots & e^{-j\alpha_{1,n_t}L} \\ 1 & e^{-j\alpha_{2,n_t}} & \cdots & e^{-j\alpha_{2,n_t}L} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & e^{-j\alpha_{(\tilde{P}-1),n_t}} & \cdots & e^{-j\alpha_{(\tilde{P}-1),n_t}L} \end{bmatrix}. \tag{16}$$

For distinct α_{i,n_t}, the rank of matrix $\mathbf{V}_{n_t}$ is L if $\tilde{P} \geq L + 1$ ([51], p. 257). Since all α_{i,n_t} are distinct and $c_{i,n_t} \neq 0 \quad \forall i$, $\mathrm{rank}(\mathbf{C}_{n_t}) = \mathrm{rank}(diag\,\{c'_{0,n_t}, c'_{1,n_t}, \cdots, c'_{(\tilde{P}-1),n_t}\} \mathbf{V}_{n_t}) = L$, $\mathbf{h}_{n_r n_t}$ can be uniquely determined from Equation (14).

3.1. SiT Based Least Squares (SiT-LS) Channel Estimation Approach

The least squares estimate of CIR between n_t-th transmitter and n_r-th receiver can be obtained from the linear model in Equation (14), as proposed in [50], and is given below

$$\hat{\mathbf{h}}_{n_r n_t}^{\mathrm{SiT-LS}} = \arg \min_{\tilde{\mathbf{h}}_{n_r n_t}} \|\hat{\mathbf{d}}_{n_r n_t} - \mathbf{C}_{n_t} \tilde{\mathbf{h}}_{n_r n_t}\|_2^2. \tag{17}$$

The above estimate can also be obtained as

$$\hat{\mathbf{h}}_{n_r n_t}^{\mathrm{SiT-LS}} = (\mathbf{C}_{n_t}^H \mathbf{C}_{n_t})^{-1} \mathbf{C}_{n_t}^H \hat{\mathbf{d}}_{n_r n_t}. \tag{18}$$

To obtain the channel estimate for non-zero mean noise, set $\tilde{P} \geq L + 1$, omit the first row from $\mathbf{C}_{n_t}$ and $\hat{d}_{n_r n_t, 0}$ from $\hat{\mathbf{d}}_{n_r n_t}$.

3.2. Proposed SiT Based Massive-MIMO Sparse Uplink Channel Estimation Techniques

The estimation error between $\hat{d}_{n_r n_t, i}$ and $d_{n_r n_t, i}$ can be computed by using Equation (3) in Equation (13), and is given by

$$\hat{d}_{n_r n_t, i} = d_{n_r n_t, i} + \varepsilon_{n_r n_t, i}, \tag{19}$$

where $\varepsilon_{n_r n_t, i}$ denotes the estimation error of $d_{n_r n_t, i}$. This estimation error constitutes contributions from interference due to superimposed information sequence of all UEs ($\tilde{b}_{n_r n_t, i}$), additive noise ($\tilde{z}_{n_r n_t, i}$), and interference due to training sequences of cross channels ($\tilde{c}_{n_r \tilde{n}_t, i}$). The estimation error is thus given by, $\varepsilon_{n_r n_t, i} = \tilde{c}_{n_r \tilde{n}_t, i} + \tilde{b}_{n_r n_t, i} + \tilde{z}_{n_r n_t, i}$, where

$$\tilde{c}_{n_r \tilde{n}_t, i} = \frac{1}{M} \sum_{k=0}^{M-1} \left[\sum_{\substack{\tilde{n}_t = 1 \\ \tilde{n}_t \neq n_t}}^{N_t - 1} \sum_{\ell=0}^{L} h_{n_r \tilde{n}_t}^\ell c_{\tilde{n}_t}(k - \ell) \right] e^{-j\alpha_{i,\tilde{n}_t} k}, \tag{20}$$

$$\tilde{b}_{n_r n_t, i} = \frac{1}{M} \sum_{k=0}^{M-1} \left[\sum_{n_t=1}^{N_t - 1} \sum_{\ell=0}^{L} h_{n_r n_t}^\ell b_{n_t}(k - \ell) \right] e^{-j\alpha_{i,n_t} k}, \tag{21}$$

$$\tilde{z}_{n_r n_t, i} = \frac{1}{M} \sum_{k=0}^{M-1} z_{n_r}(k) e^{-j\alpha_{i,n_t}k}. \tag{22}$$

Taking into account this inherent error $\varepsilon_{n_r n_t, i}$ results in more adequate estimate of the CIR. Moreover, the first-order statistics based technique discussed previously does not take into consideration the sparse nature of uplink massive-MIMO channels. This section, thus, presents two extensions of the first-order statistics based channel estimation techniques for sparse uplink massive-MIMO channels by using the model in Equation (14).

3.2.1. SiT Based Stage-Wise Orthogonal Matching Pursuit (SiT-StOMP)

In [34], StOMP was proposed with an objective of performance enhancement in the reconstruction of sparse signals for large-scale CS problems while keeping the computational cost low. This performance enhancement in StOMP is achieved by allowing for selecting multiple columns per iteration as opposed to single column selection based strategy of MP and OMP algorithms. It has been established in [34] that for the reconstruction of large-scale sparse signals, StOMP performs faster than MP and OMP.

At each iteration of the StOMP algorithm, several columns from the measurement matrix are added to the active set as compared to the single column selection of the OMP algorithm. In StOMP, the selection of columns takes place according to a certain pre-determined threshold value. Only those columns are selected whose absolute correlations with the current residual exceed the threshold value. After selection of columns, StOMP then solves for a least squares problem to update the residual vector. Thus, StOMP converges faster than OMP since it requires less number of iterations to reconstruct the sparse solution.

Due to faster convergence and better performance for large-scale systems, StOMP has been considered as a viable solution for the estimation of massive-MIMO sparse uplink channels. Therefore, in order to reconstruct the sparse channel vector $\mathbf{h}_{n_r n_t}$ from the model presented in Equation (14), we incorporate the StOMP algorithm. The proposed SiT-StOMP algorithm for the estimation of channel vector between n_r-th receiver and n_t-th transmitter is outlined as below:

Input: Matrix $\mathbf{C}_{n_t}$, vector $\hat{\mathbf{d}}_{n_r n_t}$, and threshold ϵ_i.
Output: Channel estimate vector $\hat{\mathbf{h}}_{n_r n_t}^{\text{StOMP}}$.

1. Initialize residual $\mathbf{r}_0 = \hat{\mathbf{d}}_{n_r n_t}$, index set $\Phi_0 = \varnothing$, and iteration counter $i = 1$.
2. Create a set $\boldsymbol{\Omega}_i$ consisting of the indices of all elements in the vector, $\boldsymbol{\psi}_i = \mathbf{C}_{n_t}^H \mathbf{r}_{i-1}$, which are above the threshold ϵ_i

$$\boldsymbol{\Omega}_i = \{j : \boldsymbol{\psi}_i(j) \geq \epsilon_i\}.$$

3. Update the index set by $\Phi_i = \Phi_{i-1} \cup \Omega_i$ and residual by

$$\check{\mathbf{h}}_{n_r n_t} = \arg \min_{\check{\mathbf{h}}_{n_r n_t} \in \mathbb{R}^{\Phi_i}} \|\hat{\mathbf{d}}_{n_r n_t} - \mathbf{C}_{n_t} \check{\mathbf{h}}_{n_r n_t}\|_2^2,$$

$$\mathbf{r}_i = \hat{\mathbf{d}}_{n_r n_t} - \mathbf{C}_{n_t} \check{\mathbf{h}}_{n_r n_t}.$$

4. Check stopping criteria; if it is not met then update index $i = i + 1$, and go to step 2; if stopping criteria is met, set the final output vector as $\hat{\mathbf{h}}_{n_r n_t}^{\text{StOMP}} = \check{\mathbf{h}}_{n_r n_t}$.

The stopping criteria in this case is a fixed number of maximum iterations, Λ. In order to avoid false alarms and missed detection, as proposed in [34], the threshold is set as $\epsilon_i = t_i \|\mathbf{r}_i\|_2 / \sqrt{P}$, where $2 \leq t_i \leq 3$. When the channel is sufficiently sparse, after the algorithm exits, Φ_i is expected to have no more than P entries and all the non-zeros in $\hat{\mathbf{h}}_{n_r n_t}^{\text{StOMP}}$ are selected in Φ_i.

3.2.2. SiT Based Gradient Pursuit (SiT-GP)

The GP [37] is a greedy algorithm for the reconstruction of sparse signals. It utilizes the steepest descent methodology to compute the step-size for each iteration, i.e., the sparse solution vector is updated at each iteration with a directional update computed on the basis of gradient or conjugate gradient. The only additional computational cost compared to MP is that of the evaluation of the step size. The SiT-GP algorithm is given below:

1. Initialize the residual vector $\mathbf{r}^0 = \hat{\mathbf{d}}_{n_r n_t}$, the estimate of the channel coefficients vector $\hat{\mathbf{h}}^0_{n_r n_t} = \mathbf{0}$, and $\Gamma^0 = \varnothing$;

2. for $i = 1; i = i + 1$ until stopping criteria is met, do

   (a) $\mathbf{g}^i = \mathbf{C}^H_{n_t} \mathbf{r}^{i-1}$;

   (b) $m^i = \arg_m \max | g^i_m |$;

   (c) $\Gamma^i = \Gamma^{i-1} \cup m^i$;

   (d) Compute the update direction $\mathbf{u}_{\Gamma^i} = \mathbf{C}^H_{n_{t_{\Gamma^i}}} (\hat{\mathbf{d}}_{n_r n_t} - \mathbf{C}_{n_{t_{\Gamma^i}}} \hat{\mathbf{h}}^{i-1}_{n_r n_{t_{\Gamma^i}}})$;

   (e) $\mathbf{v}^i = \mathbf{C}_{n_{t_{\Gamma^i}}} \mathbf{u}_{\Gamma^i}$;

   (f) $\alpha^i = \langle \mathbf{r}^i, \mathbf{v}^i \rangle / \|\mathbf{v}^i\|^2_2$;

   (g) $\hat{\mathbf{h}}^i_{n_r n_{t_{\Gamma^i}}} := \hat{\mathbf{h}}^{i-1}_{n_r n_{t_{\Gamma^i}}} + \alpha^i \mathbf{u}_{\Gamma^i}$;

   (h) $\mathbf{r}^i = \mathbf{r}^{i-1} - \alpha^i \mathbf{v}^i$;

3. Output $\mathbf{r}^i, \hat{\mathbf{h}}^i_{n_r n_t}$.

The adoption of directional gradient pursuit in the GP algorithm results in fast approximations to OMP but with the same computational complexity as that of MP algorithm.

4. Minimum Mean Square Error (MMSE) Based Equalizer

Once an adequate estimate of the channel vector $\mathbf{h}_{n_r n_t}$ has been obtained, the effect of the training sequence must be removed from the received signal before equalization. This is achieved by using TER block, as shown in Figure 1. Since the training sequence for each UE is also known at the BS, the effect of this superimposed training sequence can be removed as below

$$\tilde{y}_{n_r}(k) = y_{n_r}(k) - \sum_{n_t=1}^{N_t} \sum_{\ell=0}^{L} \hat{h}^\ell_{n_r n_t} c_{n_t}(k - \ell), \tag{23}$$

where $\hat{h}^\ell_{n_r n_t}$ is the estimate of ℓth tap CIR from n_t-th UE to n_r-th receive antenna. The CIR estimate $\hat{\mathbf{h}}_{n_r n_t}$ may be taken from any of the estimation techniques discussed previously, i.e., $\hat{\mathbf{h}}^{\text{SiT-LS}}_{n_r n_t}$, $\hat{\mathbf{h}}^{\text{SiT-StOMP}}_{n_r n_t}$ or $\hat{\mathbf{h}}^{\text{SiT-GP}}_{n_r n_t}$.

The optimal weights, $\mathbf{w}_{n_r}$, for the equalizer at n_r-th receive antenna, can be computed as in [52], and are given below

$$\mathbf{w}_{n_r n_t} = (\hat{\mathbf{H}}\hat{\mathbf{H}}^* + 2\sigma^2_{n_r}\mathbf{I})^{-1} \hat{\mathbf{H}} \,|_{(n_r-1)(L_e+L-1)+(\tau_d+1)}, \tag{24}$$

where n_r is the receiver index such that $1 \leq n_r \leq N_r$, L_e is the length of equalizer, τ_d represents the symbol mappers' decision delay at equalizer's output, $\mathbf{I}$ denotes the $(N_t \times L_e) \times (N_t \times L_e)$ identity

matrix and $\hat{\mathbf{H}}|_i$ is the ith column of $\hat{\mathbf{H}}$. The estimate of the noise variance $\sigma_{n_r}^2$ is obtained as in [26]. The $L_e \times (L_e + L - 1)$ convolutional matrix $\hat{\mathbf{H}}$ is given by

$$
\hat{\mathbf{H}}_{n_r n_t} = \begin{bmatrix}
\hat{h}_{n_r n_t}^0 & \hat{h}_{n_r n_t}^1 & \cdots & \hat{h}_{n_r n_t}^{L-1} & 0 & \cdots & 0 \\
0 & \hat{h}_{n_r n_t}^0 & \hat{h}_{n_r n_t}^1 & \cdots & \hat{h}_{n_r n_t}^{L-1} & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \hat{h}_{n_r n_t}^0 & \hat{h}_{n_r n_t}^1 & \cdots & \hat{h}_{n_r n_t}^{L-1}
\end{bmatrix}.
\tag{25}
$$

The overall convolutional matrix $\hat{\mathbf{H}}$ is given below

$$
\hat{\mathbf{H}} = \begin{bmatrix}
\hat{\mathbf{H}}_{1,1} & \hat{\mathbf{H}}_{1,2} & \cdots & \hat{\mathbf{H}}_{1,N_t} \\
\hat{\mathbf{H}}_{2,1} & \mathbf{H}_{2,2} & \cdots & \hat{\mathbf{H}}_{2,N_t} \\
\vdots & \vdots & \cdots & \vdots \\
\hat{\mathbf{H}}_{N_r,1} & \hat{\mathbf{H}}_{N_r,2} & \cdots & \hat{\mathbf{H}}_{N_r,N_t}
\end{bmatrix}.
\tag{26}
$$

The information symbols transmitted from the n_t-th transmitter can thus be estimated as below, followed by a symbol mapper, as shown in Figure 1:

$$
\tilde{b}_{n_t}(k) = \sum_{n_r=1}^{N_r} \sum_{i=0}^{L_e-1} w_{n_r n_t}^i \tilde{y}_{n_r}(k-i).
\tag{27}
$$

The estimated symbol vector, $\tilde{\mathbf{b}}_n$, is then fed as an input to a symbol mapper, as shown in Figure 1. The symbol mapper performs mapping of the estimated symbols as per used modulation scheme at the UEs. The output of the symbol mapper is the decoded symbol vector $\hat{\mathbf{b}}_n$.

5. Results and Discussion

This section presents the computer based simulation results of the proposed techniques and analysis of the obtained results. The normalized channel mean square error (NCMSE) and bit error rate (BER) have been used as the performance metrics for this purpose. The NCMSE for the proposed estimation techniques is defined as

$$
\text{NCMSE} = \frac{\sum_{n_r=1}^{N_r} \sum_{n_t=1}^{N_t} \sum_{\ell=0}^{L-1} \left| \hat{h}_{n_r n_t}^\ell - h_{n_r n_t}^\ell \right|^2}{\sum_{n_r=1}^{N_r} \sum_{n_t=1}^{N_t} \sum_{\ell=0}^{L-1} \left| h_{n_r n_t}^\ell \right|^2}.
\tag{28}
$$

A massive-MIMO system consisting of 64-antenna BS serving four UEs i.e., $N_r = 64$ and $N_t = 4$, is considered for the purpose of simulations. The underlying massive-MIMO sparse uplink channels are assumed to be time-invariant and frequency-selective. The channel vectors $\mathbf{h}_{n_r n_t}$, having a fixed sparsity level Q/L, are independently generated for each Monte Carlo run. The non-zero channel coefficients of $h_{n_r n_t}$ are drawn from a zero-mean Gaussian distribution with variance $1/(N_r(L+1))$. Moreover, a fixed channel length of $L = 14$ is used for all of the channels. The positions of non-zero channel taps of all the channels from a certain UE to all of the receive antennas at the BS is taken as the same because of the small separation distance of the antennas as compared to the line-of-sight (LoS) path length. The periodic and low power training sequence for each UE is generated by using the $m-$ sequence based approach, as presented in [26]. For this purpose, a base sequence $\{-1,-1,-1,1,1,1,1,-1,1,-1,1,1,-1,-1,1\}$ with period $\tilde{P} = 15$ is used for all of the simulation results. AWGN is independently generated at each receiver for a specific signal-to-noise (SNR) ratio. The SNR at n_r-th receiver is defined as the ratio of power of received signal $\sigma_{y_{n_r}}^2$ to noise power $\sigma_{n_r}^2$, i.e., $\text{SNR}_{n_r} = \sigma_{y_{n_r}}^2 / \sigma_{n_r}^2$. The zero mean binary phase-shift keying (BPSK) modulated information sequences ($\mathbf{b}_{n_t} \in \{1, -1\}$) are generated mutually independent for each transmitter.

Performance comparison of the proposed SiT-StOMP and SiT-GP with that of the SiT-LS is presented in Figure 2 for NCMSE and BER against SNR, respectively. For this purpose, length of the information sequence is set as $M = 1500$ bits, sparsity level of the channel vectors as $Q/L = 3/14$, and training-to-information ratio (TIR) as $\sigma^2_{c_{n_t}}/\sigma^2_{b_{n_t}} = 0.2$, respectively.

From Figure 2a, it can be seen that the proposed schemes perform much better as compared to SiT-LS. In terms of NCMSE at $SNR = 10$ dB, the proposed techniques SiT-StOMP and SiT-GP give a performance gain of 18 dB and 13 dB, respectively, over the SiT-LS. The BER based performance comparison is shown in Figure 2b. For a BER of 10^{-2}, a performance gain of about 3 dB and 2.5 dB in SNR is achieved by proposed SiT-StOMP and SiT-GP, respectively, over SiT-LS.
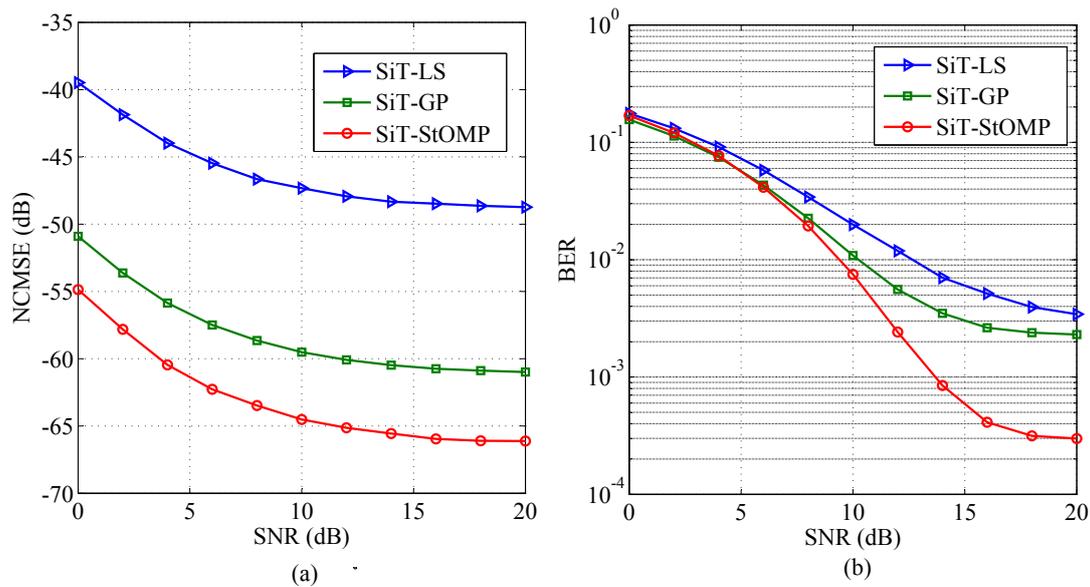

Figure 2. Normalized channel mean square error (NCMSE) and bit error rate (BER) based comparison of proposed SiT-StOMP and SiT-GP techniques with SiT-LS for massive-MIMO uplink communications, $M = 1500$ bits, $Q/L = 3/14$, and $\sigma^2_{c_{n_t}}/\sigma^2_{b_{n_t}} = 0.2$, (**a**) MSE based performance comparison; (**b**) BER based performance comparison.

In order to demonstrate the performance of proposed schemes for different variants of channel's sparsity parameter, i.e., Q/L, the NCMSE and BER are plotted in Figures 3 and 4, respectively, for several values of Q/L (i.e., $Q/L = 3/14, 5/14$ and $7/14$) by keeping the value of TIR equal to 0.2.

It can be observed from Figures 3 and 4 that NCMSE and BER decrease with increasing channel sparsity and vice versa. However, SiT-StOMP performs better than SiT-GP, in terms of NCMSE, even if the channel is less sparse.
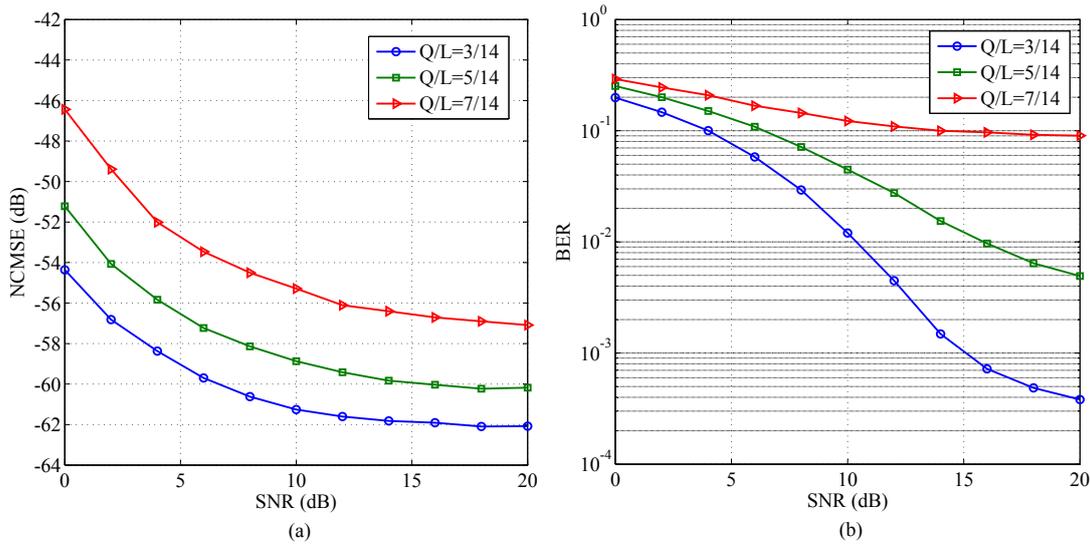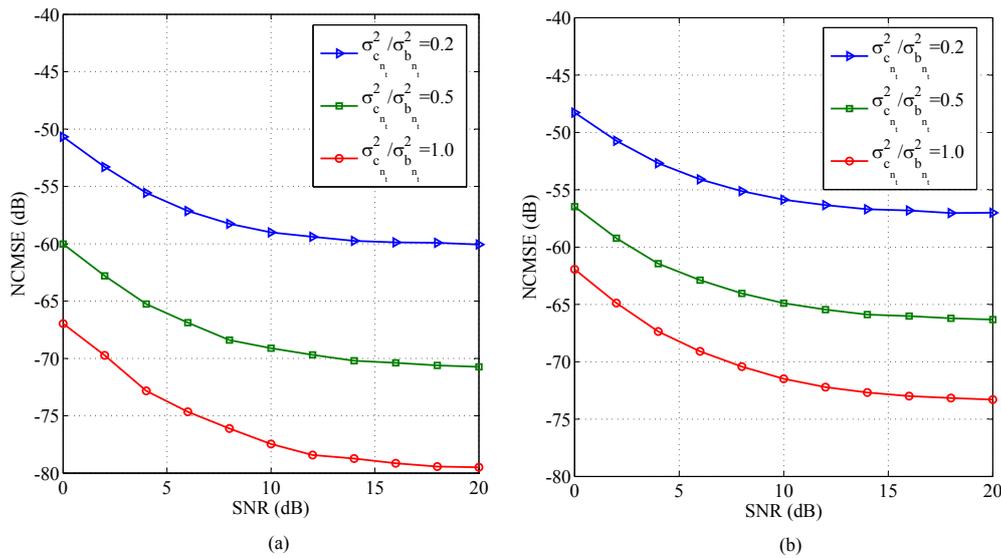
Figure 3. Effect of channel sparsity variation on superimposed training stagewise orthogonal matching pursuit (SiT-StOMP), $M = 1500$ bits, and $\sigma^2_{c_{n_t}}/\sigma^2_{b_{n_t}} = 0.2$, (**a**) NCMSE based performance; (**b**) BER based performance.
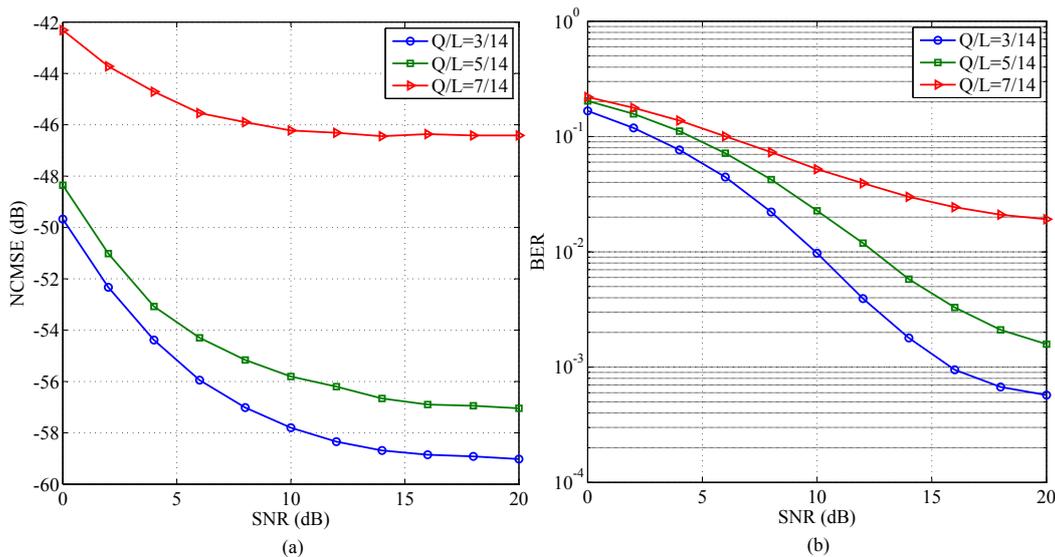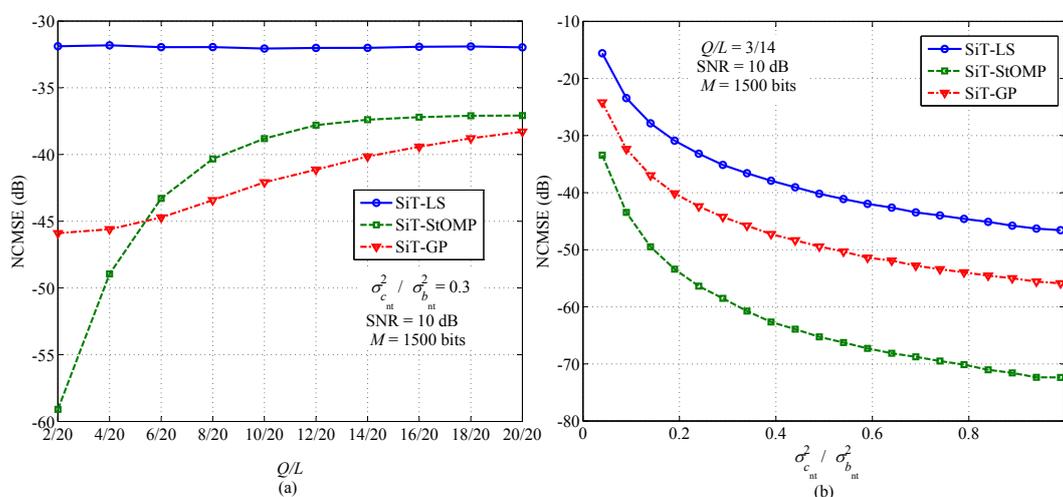

Figure 4. Effect of channel sparsity variation for SiT gradient pursuit (SiT-GP), $M = 1500$ bits, and $\sigma^2_{c_{n_t}}/\sigma^2_{b_{n_t}} = 0.2$, (**a**) NCMSE based performance; (**b**) BER based performance.

To determine the effect of variations in training power sequence, NCMSE is plotted in Figure 5 for the proposed SiT-StOMP and SiT-GP for several values of TIR i.e., $\sigma^2_{c_{n_t}}/\sigma^2_{b_{n_t}} = 0.2, 0.5$ and 1. For this purpose, the channel sparsity parameter Q/L is set equal to 3/14. It can be observed from Figure 5 that, for a fixed information sequence power, $\sigma^2_{b_{n_t}}$, increasing the training sequence power $\sigma^2_{c_{n_t}}$ results in an improved NCMSE performance for both SiT-StOMP and SiT-GP. However, increasing the TIR leads to a poor BER performance, as this increased training power could have been utilized for the information sequence. Therefore, a suitable value of TIR must be chosen in order to relieve the NCMSE and BER trade-off.

Figure 5. Effect of training-to-information ratio (TIR) variation on NCMSE performance of proposed methods, $M = 1500$ bits, and $Q/L = 3/14$, (**a**) SiT-StOMP; (**b**) SiT-GP.

A detailed performance comparison (in terms of NCMSE versus channel sparsity and TIR) of SiT-StOMP, SiT-GP, and SiT-LS algorithms is presented in Figure 6a,b, respectively. In Figure 6, it can be observed that, for all levels of TIR and channel sparsity, both SiT-StOMP and SiT-GP outperform SiT-LS. Moreover, it is evident from Figure 6a that for a TIR of 0.3, SiT-StOMP outperforms SiT-GP for high levels of channel sparsity; however, for a less sparse channel, a converse behaviour is observed. This performance gain (for highly sparse channels) of SiT-StOMP compared to SiT-GP is also influenced by the TIR level (which is also depicted in Figure 5). Therefore, the effect of TIR levels on NCMSE performance of the proposed algorithms is demonstrated in Figure 6b. It can be observed that SiT-StOMP performs equally well both for lower as well as higher values of TIR in comparison with SiT-GP and SiT-LS (plotted for a highly sparse channel). However, there is a performance trade-off (as discussed earlier) between the accuracy of channel estimates and retrieved information sequence for different values of TIR. Therefore, it is desirable to keep TIR at a certain low value (e.g., TIR = 0.2) that promises an adequate BER performance. With such setting of TIR for an environment exhibiting highly sparse CIR, it is realized that SiT-StOMP is a favorable choice for obtaining channel estimates.

Figure 6. Performance comparison of SiT-StOMP, SiT-GP, and SiT least squares (SiT-LS) algorithms. (**a**) effect of channel's sparsity; (**b**) effect of TIR.

6. Conclusions

Two channel estimation techniques based on implicit training and compressed sensing have been proposed for massive-MIMO sparse uplink channels. A comprehensive analysis based on the results obtained from computer based simulations of these techniques has been presented. Performance of the proposed schemes is evaluated on the basis of NCMSE and BER as the performance criteria. In order to prove the validity of simulation results, NCMSE and BER based performance comparison of the proposed schemes with that of a notable SiT-LS scheme has been presented. Moreover, effect of variations in channel sparsity parameter and training-to-information ratio has also been presented for the proposed techniques. It has been established that the proposed SiT-StOMP and SiT-GP techniques outperform the first-order statistics based SiT-LS in terms of NCMSE and BER for the case of sparse multipath channels. It has been shown that the proposed SiT-StOMP and SiT-GP can provide a performance gain of 18 dB and 13 dB, respectively, in terms of NCMSE at an SNR of 12 dB and channel sparsity of 80%, over SiT-LS. Similarly, for a BER of 10^{-2}, a gain of about 3 dB and 2.5 dB in SNR is achieved by SiT-StOMP and SiT-GP, respectively, over SiT-LS, for a channel sparsity level of 80%. Furthermore, in order to compare the performance of proposed SiT-StOMP and SiT-GP techniques, the NCMSE has been plotted against variations in channel sparsity and TIR. It has been demonstrated that, for sparser channels, SiT-StOMP achieves better performance in terms of NCMSE as compared to SiT-GP. However, as the channel becomes less sparse, the NCMSE for SiT-StOMP increases in comparison to that of SiT-GP. Furthermore, both of the proposed techniques perform better than SiT-LS even if the channel is less sparse. Moreover, it has been shown that SiT-StOMP performs better as compared to SiT-GP for lower as well as higher values of TIR.

Author Contributions: All authors discussed and agreed on the idea and scientific contribution. Syed Junaid Nawaz and Babar Mansoor contributed to mathematical modeling, computer simulations, and writing of the manuscript. Sardar Muhammad Gulfam contributed to the discussion on results.

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

References

1. Foschini, G.J.; Gans, M.J. On limits of wireless communications in a fading environment when using multiple antennas. *Wirel. Pers. Commun.* **1998**, *6*, 311–335.
2. Telatar, E. Capacity of Multi-antenna Gaussian Channels. *Eur. Trans. Telecommun.* **1999**, *10*, 585–595.
3. Caire, G.; Shamai, S. On the achievable throughput of a multiantenna Gaussian broadcast channel. *IEEE Trans. Inf. Theory* **2003**, *49*, 1691–1706.
4. Gesbert, D.; Kountouris, M.; Heath, R.W., Jr.; Chae, C.B.; Salzer, T. Shifting the MIMO Paradigm. *IEEE Signal Process. Mag.* **2007**, *24*, 36–46.
5. 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.
6. Marzetta, T.L. Massive MIMO: An Introduction. *Bell Labs Tech. J.* **2015**, *20*, 11–22.
7. Larsson, E.; Edfors, O.; Tufvesson, F.; Marzetta, T. Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **2014**, *52*, 186–195.
8. Ngo, H.Q.; Larsson, E.G.; Marzetta, T.L. Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems. *IEEE Trans. Commun.* **2013**, *61*, 1436–1449.
9. Cotter, S.F.; Rao, B.D. Sparse channel estimation via matching pursuit with application to equalization. *IEEE Trans. Commun.* **2002**, *50*, 374–377.
10. Raghavan, V.; Hariharan, G.; Sayeed, A.M. Capacity of Sparse Multipath Channels in the Ultra-Wideband Regime. *IEEE J. Sel. Top. Signal Process.* **2007**, *1*, 357–371.
11. Kocic, M.; Brady, D.; Stojanovic, M. Sparse equalization for real-time digital underwater acoustic communications. In Proceedings of the MTS/IEEE Challenges of Our Changing Global Environment (OCEANS'95), San Diego, CA, USA, 9–12 October 1995.

12. Ying, J.; Zhong, J.; Zhao, M.; Cai, Y. Turbo equalization based on compressive sensing channel estimation in wideband HF systems. In Proceedings of the International Conference on Wireless Communications & Signal Processing, Hangzhou, China, 24–26 October 2013; pp. 1–5.

13. Schreiber, W. Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions. *Proc. IEEE* **1995**, *83*, 958–981.

14. Nawaz, S.J.; Ahmed, K.I.; Patwary, M.N.; Khan, N.M. Superimposed training-based compressed sensing of sparse multipath channels. *IET Commun.* **2012**, *6*, 3150–3156.

15. Matthaiou, M.; Sayeed, A.M.; Nossek, J.A. Sparse multipath MIMO channels: Performance implications based on measurement data. In Proceedings of the IEEE 10th Workshop on Signal Processing Advances in Wireless Communications, Perugia, Italy, 21–24 June 2009; pp. 364–368.

16. Barbotin, Y.; Hormati, A.; Rangan, S.; Vetterli, M. Estimation of Sparse MIMO Channels with Common Support. *IEEE Trans. Commun.* **2012**, *60*, 3705–3716.

17. Gui, G.; Wan, Q.; Peng, W.; Adachi, F. Sparse multipath channel estimation using compressive sampling matching pursuit algorithm. In Proceedings of the IEEE APWCS, Kaohsiung, Taiwan, 20–21 May 2010; pp. 1–5.

18. Hoydis, J.; Hoek, C.; Wild, T.; ten Brink, S. Channel measurements for large antenna arrays. In Proceedings of the International Symposium on Wireless Communication Systems, Paris, France, 28–31 August 2012; pp. 811–815.

19. Yu, Y.; Wang, P.; Chen, H.; Li, Y.; Vucetic, B. A Low-Complexity Transceiver Design in Sparse Multipath Massive MIMO Channels. *IEEE Signal Process. Lett.* **2016**, *23*, 1301–1305.

20. Tugnait, J.K. Blind estimation and equalization of MIMO channels via multidelay whitening. *IEEE J. Sel. Areas Commun.* **2001**, *19*, 1507–1519.

21. Shin, C.; Heath, R.W.; Powers, E.J. Blind Channel Estimation for MIMO-OFDM Systems. *IEEE Trans. Veh. Tech.* **2007**, *56*, 670–685.

22. Shahbazpanahi, S.; Gershman, A.B.; Giannakis, G.B. Semiblind multiuser MIMO channel estimation using capon and MUSIC techniques. *IEEE Trans. Signal Process.* **2006**, *54*, 3581–3591.

23. Wan, F.; Zhu, W.P.; Swamy, M.N.S. Semiblind Sparse Channel Estimation for MIMO-OFDM Systems. *IEEE Trans. Veh. Tech.* **2011**, *60*, 2569–2582.

24. Lee, S.J. On the training of MIMO-OFDM channels with least square channel estimation and linear interpolation. *IEEE Commun. Lett.* **2008**, *12*, 100–102.

25. Ma, J.; Orlik, P.; Zhang, J.; Li, G.Y. Pilot Matrix Design for Estimating Cascaded Channels in Two-Hop MIMO Amplify-and-Forward Relay Systems. *IEEE Trans. Wirel. Commun.* **2011**, *10*, 1956–1965.

26. He, S.; Tugnait, J.K.; Meng, X. On superimposed training for MIMO channel estimation and symbol detection. *IEEE Trans. Signal Process.* **2007**, *55*, 3007–3021.

27. Candes, E.J.; Tao, T. Decoding by linear programming. *IEEE Trans. Inf. Theory* **2005**, *51*, 4203–4215.

28. Donoho, D.L. Compressed sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306.

29. Taubock, G.; Hlawatsch, F.; Eiwen, D.; Rauhut, H. Compressive Estimation of Doubly Selective Channels in Multicarrier Systems: Leakage Effects and Sparsity-Enhancing Processing. *IEEE J. Sel. Top. Signal Process.* **2010**, *4*, 255–271.

30. Chen, S.S.; Donoho, D.L.; Saunders, M.A. Atomic Decomposition by Basis Pursuit. *SIAM J. Sci. Comput.* **1998**, *20*, 33–61.

31. Mallat, S.G.; Zhang, Z. Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.* **1993**, *41*, 3397–3415.

32. Davis, G.; Mallat, S.; Avellaneda, M. Adaptive greedy approximations. *J. Constr. Approx.* **1997**, *13*, 57–98.

33. Needell, D.; Vershynin, R. Signal Recovery From Incomplete and Inaccurate Measurements Via Regularized Orthogonal Matching Pursuit. *IEEE J. Sel. Top. Signal Process.* **2010**, *4*, 310–316.

34. Donoho, D.L.; Tsaig, Y.; Drori, I.; Starck, J.L. Sparse Solution of Underdetermined Systems of Linear Equations by Stagewise Orthogonal Matching Pursuit. *IEEE Trans. Inf. Theory* **2012**, *58*, 1094–1121.

35. Rath, G.; Guillemot, C. Sparse approximation with an orthogonal complementary matching pursuit algorithm. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Taipei, Taiwan, 19–24 April 2009; pp. 3325–3328.

36. Needell, D.; Tropp, J. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *J. Appl. Comput. Harmon. Anal.* **2009**, *26*, 301–321.

37. Blumensath, T.; Davies, M.E. Gradient Pursuits. *IEEE Trans. Signal Process.* **2008**, *56*, 2370–2382.

38. Ngo, H.Q.; Larsson, E.G. EVD-based channel estimation in multicell multiuser MIMO systems with very large antenna arrays. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto, Japan, 25–30 March 2012; pp. 3249–3252.

39. Nguyen, S.L.H.; Ghrayeb, A. Compressive sensing-based channel estimation for massive multiuser MIMO systems. In Proceedings of the IEEE Wireless Communication and Networking Conference, Shanghai, China, 7–10 April 2013; pp. 2890–2895.

40. Rao, X.; Lau, V.K.N. Compressive Sensing With Prior Support Quality Information and Application to Massive MIMO Channel Estimation With Temporal Correlation. *IEEE Trans. Signal Process.* **2015**, *63*, 4914–4924.

41. Shariati, N.; Björnson, E.; Bengtsson, M.; Debbah, M. Low-complexity channel estimation in large-scale MIMO using polynomial expansion. In Proceedings of the IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), London, UK, 8–9 September 2013; pp. 1157–1162.

42. Qi, C.; Wu, L. Uplink channel estimation for massive MIMO systems exploring joint channel sparsity. *Electron. Lett.* **2014**, *50*, 1770–1772.

43. Wen, C.K.; Jin, S.; Wong, K.K.; Chen, J.C.; Ting, P. Channel Estimation for Massive MIMO Using Gaussian-Mixture Bayesian Learning. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 1356–1368.

44. Masood, M.; Afify, L.H.; Al-Naffouri, T.Y. Efficient collaborative sparse channel estimation in massive MIMO. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, South Brisbane, Australia, 19–24 April 2015; pp. 2924–2928.

45. Lalos, A.S.; Antonopoulos, A.; Kartsakli, E.; Renzo, M.D.; Tennina, S.; Alonso, L.; Verikoukis, C. RLNC-Aided Cooperative Compressed Sensing for Energy Efficient Vital Signal Telemonitoring. *IEEE Trans. Wirel. Commun.* **2015**, *14*, 3685–3699.

46. Nawaz, S.J.; Tiwana, M.I.; Patwary, M.N.; Khan, N.M.; Tiwana, M.I.; Haseeb, A. GA Based Sensing of Sparse Multipath Channels with Superimposed Training Sequence. *Elektronika Elektrotechnika* **2016**, *22*, 87–91.

47. Amin, B.; Mansoor, B.; Nawaz, S.J.; Sharma, S.K.; Patwary, M.N. Compressed Sensing of Sparse Multipath MIMO Channels with Superimposed Training Sequence. *Wirel. Pers. Commun.* **2016**, doi:10.1007/s11277-016-3778-7.

48. Mansoor, B.; Nawaz, S.J.; Amin, B.; Sharma, S.K.; Patwary, M.N. Superimposed training based estimation of sparse mimo channels for emerging wireless networks. In Proceedings of the IEEE International Conference on Telecommunications, Kuala Lumpur, Malaysia, 22–27 May 2016; pp. 1–6.

49. Zhao, J.; Meng, W.; Jia, S. Sparse underwater acoustic OFDM channel estimation based on superimposed training. *J. Mar. Sci. Appl.* **2009**, *8*, 65–70.

50. Tugnait, J.K.; Luo, W. On channel estimation using superimposed training and first-order statistics. *IEEE Commun. Lett.* **2003**, *7*, 413–415.

51. Stoica, P.; Moses, R. *Introduction to Spectral Analysis*; Prentice Hall: Upper Saddle River, NJ, USA, 1997.

52. Chen, S.; Livingstone, A.; Hanzo, L. Minimum bit-error rate design for space-time equalization-based multiuser detection. *IEEE Trans. Commun.* **2006**, *54*, 824–832.