Analysis of Channel Estimation Performance in MPC-RAN: Improved MMSE and Compressed Data Techniques

Emmanuel Mukubwa
Durban University of Technology
Durban, South Africa
mukubwaemmanuel@yahoo.com

Oludare Sokoya
Durban University of Technology
Durban, South Africa
darryso@gmail.com

Abstract—This paper considers channel estimation problem in massive MIMO partially centralized cloud-RAN. By noting that the user activities in massive MIMO partially centralized cloud-RAN are sparse, the channel estimation issue is solved by use of compressed data method to minimize the huge pilot overhead. Comparison and analysis of improved MMSE, via-Q and compressed data methods are done for massive MIMO partially centralized cloud-RAN. The achievable spectral efficiency (SE) and normalized mean square error (NMSE) were investigated. The RNA-based estimator gave the best performance for spectral efficiency than the MR for the multicell massive MIMO partially centralized cloud-RAN system. The performance is also evaluated for normalized mean square error for the three estimators with the RNA-MMSE giving the lowest normalized mean square error. The performance between the compressed CSI and the via-Q method show that the two methods are comparable, and this vindicates compressed data as a method to be utilized in channel state information covariance matrix estimation since it compresses the massive MIMO channel information hence mitigating the fronthaul finite capacity problem.

I. INTRODUCTION

Efficient utilization of the constrained amount of accessible spectrum to consider the exponentially growing interest for throughput has been the focal point in communication and signal processing for a couple of decades. The sporadic rise in technologies has galvanized the once predominantly offline appliances and devices to data generation points through the use of sensors and therefore pushing the demand for throughput higher [1], [2]. The current 5G and future communication systems are being enhanced to cater for this as well as conventional mobile users.

The key enabling technologies for 5G networks has been singled out to be the cloud-RAN and the massive MIMO as they promise to lower operational cost and enhance performance. When massive MIMO is utilized in the remote radio heads (RRH), fronthaul becomes the limiting factor because of its inherent finite capacity [3]. One of the foreseen solutions to fronthaul finite capacity is to split functions so that some are performed at the RRH and others at the baseband unit (BBU). Considering this suggested architecture, the RRH are charged with performing basic functions like beamforming and the BBU is left to carry out the digital functions including channel estimation. This then renders the fronthaul traffic to be mainly dependent on UT data rates and not on the number of antennas [4], [5]. This leads to the massive MIMO partially centralized C-RAN (MPC-RAN) [6].

When the partial centralization is combined with distributed cooperation for the case where RRHs are inter-connected it greatly mitigates capacity constraint and time latency on MPC-RANs fronthaul. Thus the common notion is to configure the topology to be adaptive in a way to strike a common balance between the fronthaul constraints and the distributed cooperative processing complexity [7].

Pilot contamination mitigation is needed to facilitate the approximation of covariance matrix considering one channel vector to a single user terminal (UT). If we assume particular communication setups having particular channel models, there is a possibility of mitigating pilot contamination provided certain separability conditions are satisfied [8]. From [9] It was argued that the elimination of limits on the uplink (UL) and downlink (DL) throughput as a result of pilot contamination can be realized if the covariance matrix are considered to be under certain loose conditions. Realization of this method dictates the estimation of covariance matrix at the base station (BS) and again they are acquired by virtue of observations which are subjected to pilot contamination.

It was demonstrated in [10] that the throughput of massive MIMO grows boundlessly as antenna number goes to infinity provided there is no linear dependence among covariance matrices of co-channel users. In [11] it was demonstrated that the coherence interval of channel vectors is less than that of the covariance matrix hence allowing room for accurate covariance matrix estimation. According to [12] covariance matrix corresponding to a particular UT is obtained through computation of estimated channel from sample cross-correlation of two pilot sequences.

The investigation of imperfect statistical data for the UL system is discussed in [13]. The SE is evaluated considering imperfect covariance information. Low complexity covariance matrix is presented in [14] where it shown that based on this estimated covariance matrix, both UL and DL spectral efficiencies (SEs) increases with the increase in the number of antennas. Estimation of covariance matrix from compressed data using unbiased estimator is discussed in [15].

We use compressed data to approximate the covariance matrix for MPC-RAN system following the method in [16]. But our work is different from this in that we employ this method to calculate the covariance matrix in MPC-RAN for channel estimation something that was not investigated. Individual data points are processed by multiplying them with

ISSN 2305-7254

a single projection matrix $\boldsymbol{S} \in \mathbb{C}^{M\times Z}$ with a Gaussian distribution. For data vectors K, we assume specific projection matrices equal to K [16], [17]. Based on this, we approximate covariance matrix that is consistent and enhance efficiency and accuracy by construction of a distinct sampling matrix. We seek to compress the channel data from UL pilots and extract the covariance matrix in an efficient and accurate way to estimate the channel.

We approximate the covariance matrix using compressed data using a weighted sampling structure. This strategy is data aware with most significant entries being explored allowing for good approximation accuracy with fewer entries. Then the validation of the method is done on simulated data in comparison with the conventional methods.

We begin by modeling the optimal multicell linear received signal and then tailor it to each of the covariance matrix approximation schemes. The imperfect covariance channel is modeled before evaluating the spectral efficiency, and normalized MSE compressed data channel covariance information is used to analyze the behavior of the MPC-RAN network system over the modeled channels in multi-RRH scenario.

Notation: lower-case and upper-case boldface letters denote vectors and matrices, respectively; $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$, and $tr(\cdot)$ denote the transpose, conjugate transpose, matrix inversion, and trace, respectively; $\mathbb{C}$ denotes the set of complex numbers, $\boldsymbol{I}_N$ is the $N \times N$ identity matrix. We let $\{X_t\}_{t=1}^{K}$ to represent $\{X_1, X_2, \dots, X_K\}$ which is a set of matrices and $x_{ji,t}$ to stand for the (j, i) $^{\text{th}}$ element of X_t. And then $\|X\|_2$ and $\|X\|_F$ represents the Spectral and Frobenius norms respectively. $\|\boldsymbol{X}\|_q = \left(\sum_{j=1}^{M}|x_j|^q\right)^{1/q}$ where $q \geq 1$ stands for the l_q-norm of $\mathbf{X} \epsilon \mathbb{C}^M$. We also take $\mathbb{D}(x)$ to represent a square diagonal matrix with the main diagonal having the elements of $\mathbf{X}$. $\mathbb{D}(\boldsymbol{X})$ is a square diagonal matrix with its main diagonal having only the diagonal elements of $\boldsymbol{X}$.

II. SYSTEM MODEL

We assume a MPC-RAN system with L RRHs, each of which has M transmitting antennas and K user terminals (*UTs)* having single antenna. We consider that the time division duplex (*TDD*) protocols are synchronized across RRHS to simultaneously transmit pilot signals and data to and from all BBUs. Pilots initially transmitted in ℓth RRH by *UT*s are same and given by $\boldsymbol{\psi}_K = [\boldsymbol{\varphi}_{j,1}^T, \boldsymbol{\varphi}_{j,2}^T, \dots, \boldsymbol{\varphi}_{j,K}^T]$, where $\boldsymbol{\varphi}_{j,k}$ corresponds to a pilot used by every kth user terminal (*UT*) in each RRH and $\|\boldsymbol{\varphi}_{j,k}\|^2 = 1$. Then a channel from the kth *UT* within the jth RRH is given as $\mathbf{h}_{j,k} \epsilon \mathbb{C}^M$. The channel vectors are assumed to be Rayleigh fading and modelled as in

$$\boldsymbol{h}_{j,k} \sim \mathcal{CN}(0, \boldsymbol{R}_{j,k}) \tag{1}$$

where $\boldsymbol{R}_{j,k}$ represents the covariance matrix corresponding from the jth *RRH* to the kth *UT*. If we assume Rayleigh fading with no correlation between the *UT*s, then $\mathbf{R}_{j,k} = \beta_{j,k}\mathbf{I}_M$. From [18], it is suggested that $\mathbf{R}_{j,k}$ varies slowly over time compared $\boldsymbol{h}_{j,k}$. For this work we assume that $\boldsymbol{R}_{j,k}$ is constant across the bandwidth of transmission and change slowly over time. Hence the training sequences received $\boldsymbol{Y}_j \epsilon \mathbb{C}^M$, is computed as

$$\boldsymbol{Y}_j = \mathbf{H}_{j,k}\boldsymbol{\psi}_k + \boldsymbol{Z}_j \tag{2}$$

where $\boldsymbol{\psi}_k \epsilon \mathbb{C}^K$ is the pilot matrix representing total transmitted sequences by K UTs and $\boldsymbol{Z}_j \epsilon \mathbb{C}^M$ represents the AWGN noise matrix.

III. IMPROVED MMSE CHANNEL ESTIMATION

A. Conventional MMSE Channel Estimation

The MMSE channel approximation is an improvement on the least square (LS) channel approximation. The MMSE relies on channel statistics in approximating the CSI. The MMSE approximation procedure is as follows. In MMSE approximation of the channel, the MSE obtained as the difference between the real channel, $\mathbf{h}_{j,k}$ and the estimated channel $\widehat{\boldsymbol{h}}_{j,\mathrm{k}}^{MMSE}$, as

$$\tilde{\mathbf{h}}_{j,\mathrm{k}} = \mathbf{h}_{j,\mathrm{k}} - \widehat{\boldsymbol{h}}_{j,\mathrm{k}}^{MMSE} \tag{3}$$

The basis of MMSE estimation is to minimize the mean square error (MSE) in (3) as follows

$$\tilde{\mathbf{h}}_{j,\mathrm{k}}^{MMSE} = \mathrm{E}\left\{\left\|\mathbf{h}_{j,\mathrm{k}} - \widehat{\boldsymbol{h}}_{j,\mathrm{k}}^{MMSE}\right\|_F^2\right\} \tag{4}$$

The BBU performs minimum mean-squared error (MMSE) channel estimation for each RRH of which can be written as [13], [19],

$$\widehat{\boldsymbol{h}}_{j,\mathrm{k}}^{MMSE} = \mathbf{R}_{j,\mathrm{k}}\boldsymbol{\phi}_{j,\mathrm{k}}^{-1}\boldsymbol{y}_{j,\mathrm{k}}^{p} \tag{5}$$

where

$$\boldsymbol{y}_{j,\mathrm{k}}^{p} = \mathbf{h}_{j,\mathrm{k}} + \sum_{l=1, l\neq j}^{L} \mathbf{h}_{\ell,k} + \frac{1}{\sqrt{\rho^{tr}}}\boldsymbol{N}_j^p \boldsymbol{\phi}_{j,\mathrm{k}}^{\star} \tag{6}$$

with ρ^{tr} being the normalized total pilot transmission power for each *UT* and $\boldsymbol{\Phi}_{j,k} = \mathbb{E}\left[\boldsymbol{y}_{j,k}^{p}\left(\boldsymbol{y}_{j,k}^{p}\right)^H\right]$ which is re-written as

$$\begin{aligned} \boldsymbol{\Phi}_{j,\mathrm{k}} &= \sum_{\mathrm{j}=1}^{L}\sum_{k=1}^{K}\hat{\mathbf{h}}_{j,k}\hat{\mathbf{h}}_{j,k}^{\mathbf{H}} + \frac{1}{\rho^{tr}}\boldsymbol{I}_M \\ &= \sum_{j=1}^{L}\mathbf{R}_{j,k} + \frac{1}{\rho^{tr}}\boldsymbol{I}_M \end{aligned} \tag{7}$$

B. RNA-based MMSE Channel Estimation

The MMSE approximation still requires inversion of the matrix and therefore we replace it with the rapid numerical algorithm (RNA) method. RNA-based approximation totally avoids the matrix inversion and uses multiplication and addition instead. Let Ŧ be a $M \times M$ nonsingular matrix that we are tasked to compute the inverse. Again, let $\boldsymbol{\mathcal{R}}_{\mathrm{k}}$ represent the estimated inverse in the k^{th} iteration. The residual matrix represents the divergence of the computed inverse from the real inverse of the matrix $Ғ$. The residual matrix $\mathcal{E}_k$ is obtained as

$$\mathcal{E}_k = \boldsymbol{I} - Ғ\mathcal{R}_k \tag{8}$$

This is the residual matrix in the computation of the first inverse $\mathcal{R}_k$. From which,

$$\begin{aligned} Ғ\mathcal{R}_k &= \boldsymbol{I} - \mathcal{E}_k \\ (Ғ\mathcal{R}_k)^{-1} &= (\boldsymbol{I} - \mathcal{E}_k)^{-1} \\ Ғ^{-1} &= \mathcal{R}_k(\boldsymbol{I} - \mathcal{E}_k)^{-1} \end{aligned} \tag{9}$$

This expression is a power series in $\mathcal{E}_k$. Thus,

$$\begin{aligned} Ғ^{-1} = \mathcal{R}_k(\boldsymbol{I} - \mathcal{E}_k)^{-1} &= \mathcal{R}_k \sum_{k=0}^{\infty} \left({\mathcal{E}_k}^k\right) \\ &= \mathcal{R}_k(\boldsymbol{I} + \mathcal{E}_k + \mathcal{E}_k^2 + \cdots) \end{aligned} \tag{10}$$

The first two terms can be used to represent the inverse of an infinite series as in [20], [21]. Limiting the infinite series to the first two terms, we obtain,

$$\begin{aligned} Ғ^{-1} = \mathcal{R}_k(\boldsymbol{I} + \mathcal{E}_k) &= \mathcal{R}_k(\boldsymbol{I} + \boldsymbol{I} - Ғ\mathcal{R}_k) \\ &= \mathcal{R}_k(2\boldsymbol{I} - Ғ\mathcal{R}_k) \end{aligned} \tag{11}$$

This can also be written as,

$$\mathcal{R}_k = \mathcal{R}_{k-1}(2\boldsymbol{I} - Ғ\mathcal{R}_{k-1}) \tag{12}$$

where, $\mathcal{R}_k$ represent the inverse in the next iteration. This expression is known as the Schulz iterative method for inverting a matrix [22], [23]. It was pointed out in [24] that the consideration of the initial three terms gives the quickest convergence of the iterative process for finding the inverse. Thus,

$$\begin{aligned} \mathcal{R}_k &= \mathcal{R}_{k-1}(\boldsymbol{I} + \mathcal{E}_{k-1} + \mathcal{E}_{k-1}^2) \\ &= \mathcal{R}_{k-1}\big(\boldsymbol{I} + \mathcal{E}_{k-1}(\boldsymbol{I} + \mathcal{E}_{k-1})\big) \\ &= \mathcal{R}_{k-1}\Big(\boldsymbol{I} + (\boldsymbol{I} - Ғ\mathcal{R}_{k-1})\big(\boldsymbol{I} + (\boldsymbol{I} - Ғ\mathcal{R}_{k-1})\big)\Big) \\ &= \mathcal{R}_{k-1}(\boldsymbol{I} + 2\boldsymbol{I} - \boldsymbol{I}Ғ\mathcal{R}_{k-1} - 2\boldsymbol{I}Ғ\mathcal{R}_{k-1} + (Ғ\mathcal{R}_{k-1})^2) \\ &= \mathcal{R}_{k-1}\big(3\boldsymbol{I} - Ғ\mathcal{R}_{k-1}(3\boldsymbol{I} - Ғ\mathcal{R}_{k-1})\big) \end{aligned} \tag{13}$$

This method was proposed by Amat in [25], and this sequence converges to $Ғ^{-1}$.

We assume that $\boldsymbol{\Phi}_{j,k}^{RNA}$ is a positive definite Hermitian matrix. Applying Cholesky decomposition to $\boldsymbol{\Phi}_{j,k}^{RNA}$ [26]

$$\boldsymbol{\Phi}_{j,k}^{RNA} = \boldsymbol{L}_{j,k}\boldsymbol{L}_{j,k}^{H} \tag{14}$$

where $\boldsymbol{L}_{j,k}$ represents the lower triangular matrix. This implies that

$$\left(\boldsymbol{\Phi}_{j,k}^{RNA}\right)^{-1} = \left(\boldsymbol{L}_{j,k}^{H}\right)^{-1}\boldsymbol{L}_{j,k}^{-1} \tag{15}$$

Here the computation of the inversion of matrix $\boldsymbol{\Phi}_{j,k}^{RNA}$ can be changed into computation of the inversion of matrix $\boldsymbol{L}_{j,k}$.

From (15), we let $\mathcal{R}_0 = \boldsymbol{L}_{j,k}^{-1}$ and then the residual matrix can be obtained as

$$\mathcal{E}_0 = \boldsymbol{I_M} - \boldsymbol{L}_{j,k}\mathcal{R}_0 \tag{16}$$

from which the first iteration can be expressed as

$$\mathcal{R}_1 = \mathcal{R}_0\Big(3\boldsymbol{I_M} - \boldsymbol{L}_{j,k}\mathcal{R}_0\big(3\boldsymbol{I_M} - \boldsymbol{L}_{j,k}\mathcal{R}_0\big)\Big) \tag{17}$$

and hence the k^{th} iteration is obtained in the same way as follows

$$\mathcal{R}_k = \mathcal{R}_{k-1}\Big(3\boldsymbol{I_M} - \boldsymbol{L}_{j,k}\mathcal{R}_{k-1}\big(3\boldsymbol{I_M} - \boldsymbol{L}_{j,k}\mathcal{R}_{k-1}\big)\Big) \tag{18}$$

Once again, the inversion of $\boldsymbol{L}_{j,k}$ can be computed by iterating $K - 1$ times. From which $\boldsymbol{L}_{j,k}^{-1} = \mathcal{R}_k$ and the inverse of $\boldsymbol{\Phi}_{j,k}^{RNA} = \boldsymbol{\Phi}_{j,k}$ is computed as per equation (15). From which the approximated channel is expressed as

$$\widehat{\boldsymbol{h}}_{j,k}^{RNA} = \mathbf{R}_{j,k}\left(\boldsymbol{\Phi}_{j,k}^{RNA}\right)^{-1}\boldsymbol{y}_{j,k}^{p} \tag{19}$$

The RNA-based channel estimation may have the same computational complexity as the conventional MMSE channel estimation of $O(\mathrm{M}^3)$. However, according to [27] a method that renders itself to parallelization like RNA-based channel estimation, if implemented on machine with more cores it becomes superior. And this is the case in MPC-RAN where multi-cores are utilized at the baseband unit (BBU) to enhance computational process.

IV. COMPRESSED DATA CHANNEL ESTIMATION

Practically we need to estimate the covariance matrices based on the pilot samples received at the RRH. We set out to investigate the approximation of the needed covariance information by the *BBU* and the impact of these estimates. The computation of the MMSE approximation of $\mathbf{h}_{j,k}$ at the *j*th *RRH* from (5) requires the knowledge of $\mathbf{R}_{j,k} = \mathbb{E}\left[\mathbf{h}_{j,k}\mathbf{h}_{j,k}^{H}\right]$ and $\boldsymbol{\Phi}_{j,k} = \mathbb{E}\left[\boldsymbol{y}_{j,k}^{p}\left(\boldsymbol{y}_{j,k}^{p}\right)^{H}\right]$. Bearing in mind that these are $M \times M$ (quite large) matrices, and regularization of the estimates might be needed [28], [29].

To compute these covariance matrices, we require huge communication and storage resources since the use of MPC-RAN results in high-dimensional data. Enormous bandwidth and power resources are critically needed [16] to transmit the CSI information from the RRHs to the BBU. To mitigate this problem, we specify our C-RAN system to be partially centralized with the massive MIMO RRHs interconnected and cooperating (MPC-RAN) [7]. As pointed out earlier this renders the fronthaul traffic to be mainly dependent on UT data rates and not on the number of antennas Then we leverage compressed data to approximate the covariance matrix. We follow the via-Q method to compute the covariance matrices [13] but using the compressed data as in [16].

A. Approximation of $\phi_{j,k}$

We assume that the pilot signal $\boldsymbol{y}_{j,k}^{p}$ arrives at the *j*th *RHH* over N_ϕ coherence blocks. These N_ϕ observations can be denoted by $\boldsymbol{y}_{j,k}^{p}[1], \ldots, \boldsymbol{y}_{j,k}^{p}[N_\phi]$. Then the sample observation is formulated as follows [13]

$$\boldsymbol{y}_{j,k}^{(Sample)} = \frac{1}{N_\phi}\sum_{n=1}^{N_\phi} \boldsymbol{y}_{j,k}^{p}[n] \tag{20}$$

Equation (20) for the case of an antenna index k almost surely (a. s.) tends to true $\boldsymbol{y}_{j,k}^{p}$ as $N_\phi \to \infty$.

$$\frac{1}{N_\phi}\sum_{n=1}^{N_\phi} \left[\boldsymbol{y}_{j,k}^{p}[n]\right]_{m,k} \xrightarrow{a.s.} \left[\boldsymbol{y}_{j,m}^{p}\right]_{m,k} \tag{21}$$

This is a derivative of the large numbers law and channels ergodicity. To find a good $\boldsymbol{y}_{j,k}^{p}$ approximation we need just a few observations as the standard deviation decays by $1/\sqrt{N_\phi}$ for a given sample $\boldsymbol{y}_{j,k}^{p}$. The elements within $\mathbf{y}_{j,k}^{(Sample)}$ will individually tend to the corresponding elements in $\boldsymbol{y}_{j,k}^{p}$. Then we follow the method presented in [16] to mitigate this problem.

Weighted sampling matrices $\{\boldsymbol{S}_{j,k}\}_{k=1}^{K} \in \mathbb{C}^{M\times Z}$ are adopted in compression of data through $\boldsymbol{S}_{j,k}^{T}\mathbf{y}_{j,k}^{(Sample)}$ and the data is projected back into original space via $\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{T}\mathbf{y}_{j,k}^{(Sample)}$. The data obtained is then used in approximation of the covariance matrix. The weighted sampling matrix $\boldsymbol{S}_{j,k}$ removes at least M-Z elements from the *k*th vector, the remaining ones are retained as they can be most informative. When the sampling probabilities are designed carefully, the unbiased estimator $\widehat{\boldsymbol{\Phi}}_{j,k}$ would perform accurately in relation to the spectral norm of the matrix $\|\widehat{\boldsymbol{\Phi}}_{j,k} - \boldsymbol{\Phi}_{j,k}\|_2$ [15], [30], [31].

The weighted sampling evoked is good enough to explore the most important entries to reduce the estimation error $\|\widehat{\boldsymbol{\Phi}}_{j,k} - \boldsymbol{\Phi}_{j,k}\|_2$. We begin of by setting up the necessary variables and then carry out the approximation process.

We know that $\mathbf{y}_{j,k}^{(Sample)} \in \mathbb{C}^{M\times K}$ and we set $\alpha\in[0,1]$ as our regularizing factor. Then the uplink received information is compressed as follows with $l = [1,2,\dots,M]$

$$\boldsymbol{v}_{j,k} = \left\|\boldsymbol{y}_{j,k}^{(Sample)}\right\|_1 = \sum_{l=1}^{M}\left|\boldsymbol{y}_{j,l,k}^{(Sample)}\right| \tag{22}$$

and

$$\boldsymbol{\omega}_{j,k} = \left\|\boldsymbol{y}_{j,k}^{(Sample)}\right\|_2^2 = \sum_{l=1}^{M}\left(\boldsymbol{y}_{j,l,k}^{(Sample)}\right)^2 \tag{23}$$

To compress this matrix, we sample Z rows of $\mathbf{y}_{j,k}^{(Sample)}$ instead of all the M rows. We let $z \in [Z]$, and then pick $t_{z,k} \in [M]$ with

$$\begin{aligned}\boldsymbol{p}_{j,l,k} &\equiv \mathbb{P}(t_{z,k} = l)\\ &= \alpha\frac{\left|\boldsymbol{y}_{j,l,k}^{(Sample)}\right|}{\boldsymbol{v}_{j,k}}\\ &+ (1-\alpha)\frac{\left(\boldsymbol{y}_{j,l,k}^{(Sample)}\right)^2}{\boldsymbol{\omega}_{j,k}}\end{aligned} \tag{24}$$

and we let

$$\boldsymbol{x}_{j,l,k} = \boldsymbol{y}_{j,t_{z,k},k}^{(Sample)} \tag{25}$$

Then the compressed data $\mathbf{X}$, the indices used in sampling $\mathbf{T}, \boldsymbol{V}, \boldsymbol{W}$ *and* α are transmitted to the BBU from the RRH and used to construct the unbiased covariance matrix estimator from the compressed data as follows

$$\boldsymbol{p}_{j,t_{z,k},k} = \alpha\frac{\left|\boldsymbol{x}_{j,z,k}^{(Sample)}\right|}{\boldsymbol{v}_{j,k}} + (1-\alpha)\frac{\left(\boldsymbol{x}_{j,z,k}^{(Sample)}\right)^2}{\boldsymbol{\omega}_{j,k}} \tag{26}$$

and

$$\boldsymbol{s}_{j,t_{z,k},k} = \frac{1}{\sqrt{Z\boldsymbol{p}_{j,t_{z,k},k}}} \tag{27}$$

Due to imperfection in correlation matrix knowledge, we realize robust approximation through experimental optimization of the parameter α. With advances in computing, the manipulation of vectors with length $O(M)$ in the memory is possible. Thus, the compression of data through weighted sampling will need a single pass in moving data to memory from the RRH to the BBU. This makes the algorithm to render itself to streaming data and therefore is well suited for use in MPC-RAN systems.

The estimator is unbiased and is expressed through $\{\boldsymbol{S}_{j,k}\}_{k=1}^{K}$ and $\{\boldsymbol{S}_{j,k}^{T}\mathbf{y}_{j,k}^{(Sample)}\}_{k=1}^{K}$. We have that $\mathbf{y}_{j,k}^{(Sample)} \in \mathbb{C}^{M\times K}$ and we let our sampling window be $2 \le Z < M$. Thus, we take Z entries for every $\mathbf{y}_{j,k}^{(Sample)} \in \mathbb{C}^{M}$. The sampling probabilities are taken to be $\{\mathrm{p}_{j,l,k}\}_{l=1}^{M}$ and the sampling matrix is expressed as $\boldsymbol{S}_{j,k} \in \mathbb{C}^{M\times Z}$. To recover the unbiased estimator corresponding to $\boldsymbol{\Phi}_{\mathrm{j,k}} = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{y}_{j,k}^{p}\left(\boldsymbol{y}_{j,k}^{p}\right)^H = \frac{1}{K}\boldsymbol{Y}^p(\boldsymbol{Y}^p)^H$ which is our covariance matrix target, we follow

$$\widehat{\boldsymbol{\Phi}}_{j,k}^{Compressed} = \widehat{\boldsymbol{\Phi}}_{j,k}^{1} - \widehat{\boldsymbol{\Phi}}_{j,k}^{2} \tag{28}$$

with $\mathbb{E}[\widehat{\boldsymbol{\Phi}}_{j,k}^{Compressed}] = \boldsymbol{\Phi}_{\mathrm{j,k}}$

$$\widehat{\boldsymbol{\Phi}}_{j,k}^{1} = \frac{Z}{KZ-K}\sum_{k=1}^{K}\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}\boldsymbol{y}_{j,k}^{(Sample)}\left(\boldsymbol{y}_{j,k}^{(Sample)}\right)^H\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H} \tag{29}$$

$$\widehat{\boldsymbol{\Phi}}_{j,k}^{2} = \frac{Z}{KZ-K}\sum_{k=1}^{K}\mathbb{D}\left(\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}\boldsymbol{y}_{j,k}^{(Sample)}\left(\boldsymbol{y}_{j,k}^{(Sample)}\right)^H\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}\right)\mathbb{D}(\boldsymbol{b}_{j,k}) \tag{30}$$

with $\boldsymbol{b}_{j,l,k} = \frac{1}{1+(Z-1)\mathrm{p}_{j,l,k}}$. A maximum of Z entries have to be calculated for each $\mathbf{b}_{\mathrm{j,l,k}}$ since for a given $\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}\boldsymbol{y}_{j,k}^{p}\left(\boldsymbol{y}_{j,k}^{p}\right)^H\boldsymbol{S}_{j,k}\boldsymbol{S}_{j,k}^{H}$ the maximum none zero elements on the diagonal are Z.

B. Approximation of $R_{j,k}$

To estimate the $\mathbf{R}_{j,k} \in \mathbb{C}^{M\times M}$ we follow a similar method used for $\boldsymbol{\phi}_{\mathrm{j,k}}$. The task is to acquire the $\mathbf{h}_{j,k}$ observations with minimal interference from other UTs. From [13], [32] it has been pointed out that the UT can employ a set of unique orthogonal pilots to realize a training phase to learn $\mathbf{R}_{j,k}$. We assume that the *j*th RHH has N_R observations of the noisy $\mathbf{h}_{j,k}$ which lays the basis of constructing the approximate covariance matrix $\widehat{\mathbf{R}}_{j,k}$.

This will simply mean that more data transmission from RRH to BBU over the fronthaul and increased computations for that matter. We adopt the via-Q method presented in [13] to evaluate the covariance matrix $\widehat{\mathbf{R}}_{j,k}$. This permits the estimation of $\widehat{\mathbf{\Phi}}_{j,-k}^{(Sample)} = \widehat{\mathbf{\Phi}}_{j,k} - \mathbf{h}_{j,k}\mathbf{h}_{j,k}^{H}$ which combines all interfering UTs. We set the $\widehat{\mathbf{\Phi}}_{j,k}^{(Sample)} = \widehat{\mathbf{\Phi}}_{j,k}^{1}$. Thus, the covariance matrix $\widehat{\mathbf{R}}_{j,k}^{(Sample)}$ can be computed as

$$\widehat{\mathbf{R}}_{j,k}^{(Sample)} = \widehat{\mathbf{\Phi}}_{j,k}^{(Sample)} - \widehat{\mathbf{\Phi}}_{j,-k}^{(Sample)} \tag{31}$$

Then we compute the approximate covariance matrix $\widehat{\mathbf{R}}_{j,k}^{Compressed}$ as follows

$$\widehat{\mathbf{R}}_{j,k}^{Compressed} = \beta\widehat{\mathbf{R}}_{j,k}^{(Sample)} - (1-\beta)\widehat{\mathbf{R}}_{j,k}^{(Sample)} \tag{32}$$

where $\beta \in [0,1]$ is the regularizing parameter employed in approximation of $\widehat{\mathbf{R}}_{j,k}$.

C. Channel Estimate Approximation

The approximation of MMSE estimate corresponding to $\widehat{\boldsymbol{h}}_{j,k}$ is computed based on $\widehat{\mathbf{R}}_{j,k}^{Compressed}$ and $\mathbf{\Phi}_{j,k}^{Compressed}$, that are assumed to be the correct covariance matrices, as

$$\hat{\mathbf{h}}_{j,k}^{\text{Compressed}} = \mathrm{W}_{j,k}^{Compressed}\boldsymbol{y}_{j,k}^{p} \tag{33}$$

where $\boldsymbol{W}_{j,k} = \widehat{\mathbf{R}}_{j,k}^{Compressed}\left(\mathbf{\Phi}_{j,k}^{Compressed}\right)^{-1}$.

D. Spectral Efficiency Estimate Approximation

A lower bound of the capacity that is independent of MMSE estimates is required to facilitate the quantification of SE in the presence of imperfect covariance information. From [33], the kth UT in the jth RRH has a channel capacity that is lower bounded by

$$\boldsymbol{SE}_{j,k} = \left(1-\frac{K}{\tau_c}\right)log_2\big(1 + \boldsymbol{SINR}_{j,k}\big)\ [bit/s/Hz] \tag{34}$$

where

$$\boldsymbol{SINR}_{j,k} = \frac{\left|\mathbb{E}[\mathbf{h}_{j,k}^{H}\mathbf{f}_{j,k}]\right|^2}{\sigma^2 + \mathbb{E}\left[\left|\mathbf{h}_{j,k}^{H}\mathbf{f}_{j,k}\right|^2\right] - \left|\mathbb{E}[\mathbf{h}_{j,k}^{H}\mathbf{f}_{j,k}]\right|^2 + \sum_{\ell,m}\mathbb{E}\left[\left|\mathbf{h}_{\ell,k}^{H}\mathbf{f}_{\ell,m}\right|^2\right] - \mathbf{\Upsilon}_{cp}} \tag{35}$$

and the term $\mathbf{\Upsilon}_{cp} = \sum_{\ell,m}\mathbb{E}\left[\left|\mathbf{h}_{j,k}^{H}\frac{diag\{\hat{\mathbf{h}}_{\ell,m}\}}{diag\{\mathbf{h}_{j,k}\}}\mathbf{f}_{\ell,m}\right|^2\right]$ represents the BSs cooperation component and with the expectation taken relative to the channel realization. We assumed partially centralized C-RAN with interconnected RRHs. Here it is clear to see that the capacity bound is independent of the channel estimation method and combining scheme used.

V. NUMERICAL RESULTS AND ANALYSIS

In this section we look at the performance parameters NMSE, SNR and M for all the channel estimation techniques. Tradeoffs amongst these parameters are analyzed for the massive MIMO uplink with all the four channel estimation schemes discussed above. The NMSE can be computed as,

$$MSE = \frac{E\{trace(\tilde{\mathbf{h}}_{j,k}\tilde{\mathbf{h}}_{j,k}^{H})\}}{trace(\mathbf{R}_{j,k})} \tag{36}$$

The performance of the four different channel estimation techniques is analyzed in view of the NMSE, SNR and the respective corresponding M. At first, the performance of the four corresponding channel estimation techniques is compared in with different values of SNR and transmit antennas.

We provides the comparison and analysis of NMSE, SE, and M for the RNA-MMSE channel estimation technique, via-Q method in [13] and the channel estimation using compressed channel data in MPC-RAN. We compare these channel estimation techniques in multicell MPC-RAN. This comparison is carried out for M varying from $16\ to\ 160$, with a step of 16 and $K = 10$ MPC-RAN system and the SNR varies from 0dB to 20dB. Again these parameters are different from those used in [13] and putting in mind the architecture of the network we expect its values to differ from those specified. We also assume that $\mathbf{y}_{j,k}^{(Sample)}$ is already known at the BBU. We then average the SE and NMSE over this SNR range to get the average SE and NMSE.

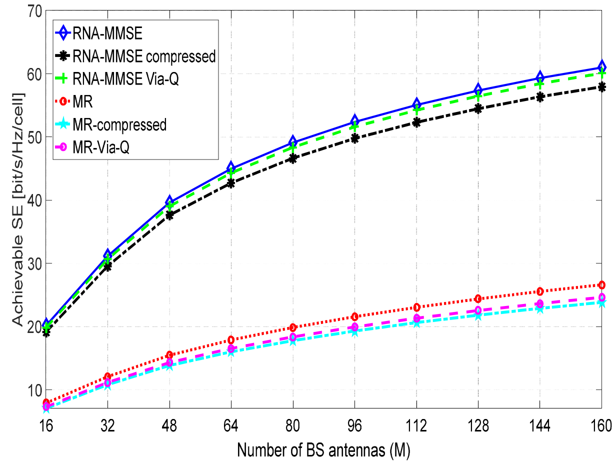

Fig. 1. Achievable SE per RRH vs. number of RRH antennas with a reuse factor of 1

Fig. 1 compares the achievable SE per cell between MR and the RNA-MMSE precoding techniques in multicell massive MIMO with a reuse factor of 1 for RNA-MMSE channel estimation, the via-Q channel estimation and the compressed data channel estimation. Based on this figure, several observations can be made. The RNA-MMSE based SE has the best achievable SE per RRH as expected from previous analyses, but it can be observed that the SE computed using the compressed data channel estimation is lower than the RNA-MMSE channel estimation and the via-Q channel estimation for both MR and RNA-MMSE. This corresponds well with the expectation since the compressed data channel estimation approximates the RNA-MMSE channel estimation. But it can also be noted that as the number of the RRH antennas grow large, the SE performance improves for both MR and RNA-MMSE based precoding.

From Fig. 2, with a reuse factor of 4, the SE computed from RNA-MMSE channel estimation, the via-Q channel estimation and the compressed data channel estimation with MR and RNA precoders improves. The increase in SE per RRH can be attributed to the fact that the pre-log factor reduces with the increased number of pilots. Also, this leads to increased instantaneous SINR as the channel estimates become better with reduced pilot contamination.

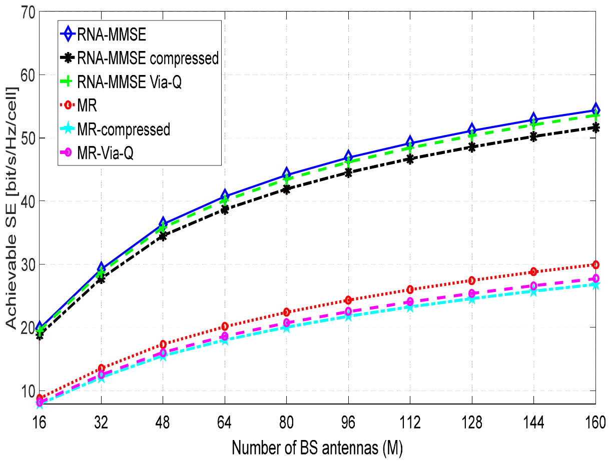

Fig. 2. Achievable SE per RRH vs. number of RRH antennas with a reuse factor of 4

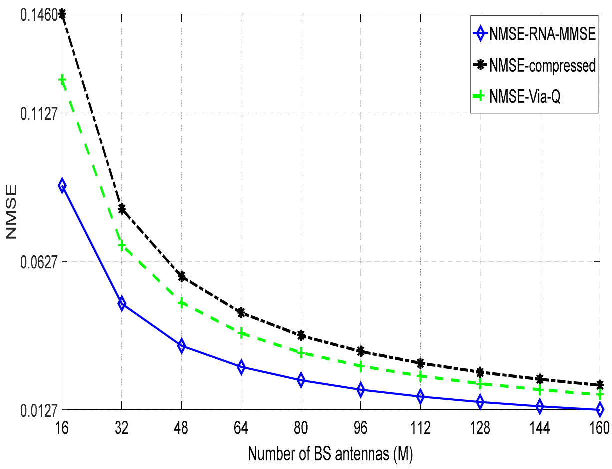

Fig. 3. The normalized MSE vs. number of RRH antennas with a reuse factor of 2

From Fig. 3, the normalized MSE (NMSE) against the number of BS antennas is depicted. As the number of RRH antennas increase the NMSE decreases since the channel estimation improves due channel hardening phenomenon. Again, the RNA-MMSE channel estimation and the via-Q channel estimation have less NMSE compared to the compressed data channel estimation because the compressed data channel estimation approximates the RNA-MMSE channel estimation. But as M increases the compressed data channel estimation NMSE nears that of the RNA-MMSE channel estimation since the approximation improves with the increase in the number of antennas.

In Fig. 4 the reuse factor is set to 4 and the NMSE is reduced as compared to the case when the reuse factor is set to 1. This can be attributed to the fact that as the reuse factor increases the pilot contamination reduces and this enhances the channel estimation process leading to a reduction in NMSE for all the RNA-MMSE channel estimation, the via-Q channel estimation and the compressed data channel estimation.

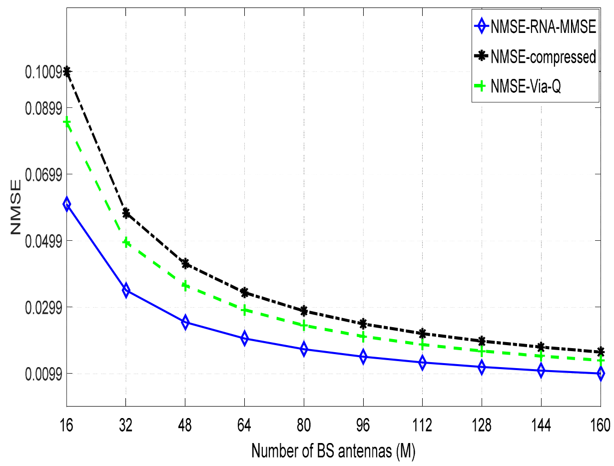

Fig. 4. The normalized MSE vs. number of RRH antennas with a reuse factor of 4

VI. CONCLUSION

The paper gives the performance analysis and comparison of the RNA-MMSE, Via-Q method, and Compressed data-based channel estimators for MPC-RAN system. The performance of the three channel estimation schemes in terms of SE and the NMSE is studied. The SE and NMSE were derived theoretically for each of the covariance matrix estimation schemes under similar assumptions and for the MPC-RAN system. From the simulation and the theoretical results, RNA-based precoding has higher SE than the MR. The NMSE for the compressed data estimator is lower than that of the RNA-MMSE and via-Q method. And it can also be seen that both the compressed CSI and the via-Q method performance is comparable, and this points to better study around the compressed CSI method to enhance its applicability. As the number of the antennas increase the compressed data estimator NMSE performance nears that of the RNA-MMSE and the via-Q method, this is attributed to better approximation as the number of antennas increase. The future work to this study will be to look at combination of the compressed data estimator and sub-space tracking algorithm to realize a semi-blind channel estimator. This will offer better estimation with reduced data size and number of pilots and need to render itself to high parallelization.

REFERENCES

[1] V. Va, T. Shimizu, G. Bansal, and R. W. Heath, “Millimeter wave vehicular communications: A survey,” *Found. Trends Netw.*, vol. 10, no. 1, pp. 1–116, 2016.

[2] J. Choi, V. Va, N. González-Prelcic, R. Daniels, C. R. Bhat, and R. W. Heath, “Millimeter-Wave Vehicular Communication to Support Massive Automotive Sensing,” *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 160–167, 2016.

[3] J. Francis and G. Fettweis, “Power Allocation for Massive MIMO-based , Fronthaul-constrained Cloud RAN Systems,” *eprint arXiv:1810.04529*, vol. arXiv:1810, pp. 1–10, 2018.

[4] J. Park, D. M. Kim, E. De Carvalho, and C. N. Manch, “Hybrid Precoding for Massive MIMO Systems in Cloud RAN Architecture with Capacity-Limited Fronthauls,” *arXiv.org*, vol. arXiv:1709, pp. 1–13, 2017.

[5] J. Francis and G. Fettweis, “Energy Efficiency Maximization in

Massive MIMO-aided, Fronthaul-constrained C-RAN," *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, vol. 2019-Septe, pp. 1–6, 2019.

[6] S. Park, H. Lee, C. B. Chae, and S. Bahk, "Massive MIMO operation in partially centralized cloud radio access networks," *Comput. Networks*, vol. 115, no. 2017, pp. 54–64, 2017.

[7] M. Peng, C. Wang, V. Lau, and H. V. Poor, "Fronthaul-constrained cloud radio access networks: Insights and challenges," *IEEE Wirel. Commun.*, vol. 22, no. 2, pp. 152–160, 2015.

[8] S. Haghighatshoar and G. Caire, "Massive MIMO pilot decontamination and channel interpolation via wideband sparse channel estimation," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 12, pp. 8316–8332, 2017.

[9] E. Bjornson, J. Hoydis, and L. Sanguinetti, "Massive MIMO Has Unlimited Capacity," *IEEE Trans. Wirel. Commun.*, vol. 17, no. 1, pp. 574–590, 2018.

[10] E. Bjornson, J. Hoydis, and L. Sanguinetti, "Pilot contamination is not a fundamental asymptotic limitation in massive MIMO," in *IEEE International Conference on Communications*, 2017.

[11] D. Neumann, M. Joham, and W. Utschick, "Covariance Matrix Estimation in Massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 863–867, 2018.

[12] K. Upadhya and S. A. Vorobyov, "Covariance Matrix Estimation for Massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 4, pp. 546–550, 2018.

[13] E. Bjornson *et al.*, "Massive MIMO with imperfect channel covariance information," in *50th Asilomar Conference on Signals, Systems and Computers*, 2016.

[14] X. Zhang, S. Wu, S. Yu, and J. Liu, "Spectral efficiency of massive MIMO networks with pilot contamination and channel covariance matrix estimation," *IET Commun.*, vol. 13, no. 1, pp. 59–65, 2018.

[15] F. Pourkamali-Anaraki, "Estimation of the sample covariance matrix from compressive measurements," *IET Signal Process.*, pp. 1089–1095, 2016.

[16] X. Chen, M. R. Lyu, and I. King, "Toward Efficient and Accurate Covariance Matrix Estimation on Compressed Data," in *Proceedings of the 34th International Conference on Machine Learning*, 2017, vol. 70, pp. 767–776.

[17] F. Pourkamali-Anaraki and S. Becker, "Preconditioned Data Sparsification for Big Data with Applications to PCA and K-Means," *IEEE Trans. Inf. Theory*, vol. 63, no. 5, pp. 2954–2974, 2017.

[18] I. Viering, H. Hofstetter, and W. Utschick, "Spatial Long-term Variantions in Urban, Rurla and Indoor Enviroments," in *in the 5th Meeting of COST273*, 2002.

[19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. NJ,USA: Prentice-Hall, Inc., 1993.

[20] V. Pan and J. Reif, "Efficient Parallel Solution of Linear Systems.," *Conf. Proc. Annu. ACM Symp. Theory Comput.*, pp. 143–152, 1985.

[21] V. Pan, "Complexity of Algorithms for Linear Systems of Equations," *Comput. Algorithms Solving Linear Algebr. Equations*, pp. 27–56, 1991.

[22] H. B. Li, T. Z. Huang, Y. Zhang, X. P. Liu, and T. X. Gu, "Chebyshev-type methods and preconditioning techniques," *Appl. Math. Comput.*, vol. 218, no. 2, pp. 260–270, 2011.

[23] A. Ben-Israel, "An iterative method for computing the generalized inverse of a matrix," *Math. Comput.*, vol. 19, no. 91, pp. 452–455, 1965.

[24] E. Isaacson and H. B. Keller, *Analysis of Numerical Methods*. New York: Dover Publications, INC., 1994.

[25] S. Amat, S. Busquier, and J. M. Gutiã Errez, "Geometric constructions of iterative functions to solve nonlinear equations," *J. Comput. Appl. Math.*, vol. 157, no. 1, pp. 197–205, 2003.

[26] A. A. Rontogiannis, V. Kekatos, and K. Berberidis, "A square-root adaptive V-BLAST algorithm for fast time-varying MIMO channels," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 3135–3139, 2006.

[27] W. Ford, *Numerical Linear Algebra with Applications*, 4th ed. London: Elsevier Inc., 2015.

[28] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *J. Multivar. Anal.*, vol. 88, no. 2, pp. 365–411, 2004.

[29] N. Shariati, E. Bjornson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale MIMO with arbitrary statistics," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 815–830, 2014.

[30] A. Gittens, "The spectral norm error of the naive Nystrom extension," *arXiv Prepr. arXiv1110.5305*, pp. 1–9, 2011.

[31] D. Achlioptas, Z. Karnin, and E. Liberty, "Near-Optimal Entrywise Sampling for Data Matrices," in *Neural Information Processing Systems*, 2013, pp. 1565–1573.

[32] H. Yin, D. Gesbert, M. FiliPpou, and Y. Liu, "A Coordinated Approach to Channel Estimation in Large-Scale Multiple-Antenna Systems," *IEEE J. Sel. AREAS COMM UNICATIONS*, vol. 31, no. 2, pp. 264–273, 2013.

[33] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of massive MIMO*. 2016.