

Joint Antenna Detection and Channel Estimation for Non-Coherent User Terminals

Ema Becirovic, Emil Björnson and Erik G Larsson

The self-archived postprint version of this journal article is available at Linköping University Institutional Repository (DiVA):

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-168236>

N.B.: When citing this work, cite the original publication.

Becirovic, E., Björnson, E., Larsson, E. G, (2019), Joint Antenna Detection and Channel Estimation for Non-Coherent User Terminals, *2019 IEEE 20TH INTERNATIONAL WORKSHOP ON SIGNAL PROCESSING ADVANCES IN WIRELESS COMMUNICATIONS (SPAWC 2019)*.
<https://doi.org/10.1109/SPAWC.2019.8815500>

Original publication available at:

<https://doi.org/10.1109/SPAWC.2019.8815500>

Copyright: IEEE

<http://www.ieee.org/>

© 2019 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.



Joint Antenna Detection and Channel Estimation for Non-Coherent User Terminals

Enma Becirovic, Emil Björnson, and Erik G. Larsson
 Dept. of Electrical Engineering (ISY), Linköping University, Linköping, Sweden
 Email: {ema.becirovic, emil.bjornson, erik.g.larsson}@liu.se

Abstract—In this paper we propose a method of improving channel estimates for non-coherent terminals with channels that can be considered constant over multiple time slots. The terminals have multiple antennas and are free to choose whichever antenna they want to use in each time slot. An unknown phase shift is introduced in each time slot as we cannot guarantee that the terminals are phase coherent across time slots. The proposed methods of improving channel estimates are a combination of clustering and heuristic methods. With our proposed methods we can have an improvement of 1.5 dB at 1 bit/s/Hz.

Index Terms—Channel estimation, non-coherent terminal.

I. INTRODUCTION

Channel estimation is crucial in multiple antenna wireless communication systems. In reciprocity-based precoding, which is superior to feedback-based precoding, uplink channel estimation is especially important for it to work effectively, and also achieve accurate demodulation of uplink data [1]. Since the uplink power can be two orders-of-magnitude lower than the downlink power, the uplink estimation quality can be poor even when the downlink signal-to-noise ratio (SNR) is good. The estimates can be improved by averaging over multiple slots. The slot duration is equal to the minimum coherence time that the system supports. The vast majority of users will have a substantially longer coherence time, and the majority of traffic is generated by stationary indoor users that are the ones with the longest coherence time. Filtering of channel estimates is well known and used in state-of-the-art systems, see [2], [3], [4] for examples in OFDM-systems.

In this paper, we study the class of non-coherent 5G terminals [5, Ch. 7], which only use one of its antennas (or, mathematically equivalent, one of the pre-defined beams in the codebook) simultaneously. Moreover, in each time slot a phase shift might occur which is due to the fact that the terminals are not required to be phase coherent across time slots which facilitates building energy efficient hardware, e.g., hardware that is allowed to turn components on and off. The phase shift may also be due to power amplifiers that change the range of operation.

Technical contributions: We propose methods of improving the channel estimates by jointly detecting the antenna that the terminal used, and estimating the phase shifts and obtaining

our channel estimates. The problem statement in this paper is to the authors knowledge new.

II. SYSTEM MODEL

We consider a single-cell system with M antennas at the base station and one user with K antennas. The channel between the user and the base station is considered constant over many time slots. The base station wants to exploit this and use the received pilot signal from many time slots to estimate the channel to the user. However, the user is free to use any of its antennas but can only use one at a time. Here, the word antenna is not necessarily referring to a physical antenna but can refer to a virtual antenna defined by a per-defined beam.¹ The choice of antenna is unknown to the base station. We denote the selected antenna at time t with $S(t) \in \{1, \dots, K\}$.

Let \mathcal{T}_k be the set of time slots where the user selected to use antenna k : $\mathcal{T}_k = \{t : S(t) = k\}$. The union of the sets is equal to the set of all time slots, $\bigcup_{k=1}^K \mathcal{T}_k = \{1, \dots, T\}$, where T is the number of slots where the channels can be considered to be constant. T is assumed to be known at the base station. The set of time slots of the selected antennas are all disjoint, $\mathcal{T}_k \cap \mathcal{T}_{k'} = \emptyset, \forall k \neq k'$, meaning that the user can only activate one antenna at a time.

At each time slot t , the user first sends a pilot signal $\phi(t) \in \mathbb{C}^{\tau_p}$ of length τ_p , $\|\phi(t)\| = 1$, which is known to the base station. The base station receives

$$\mathbf{Y}^p(t) = \sqrt{p\tau_p} e^{i\psi(t)} \mathbf{g}_{S(t)} \phi(t)^H + \mathbf{W}^p(t), \quad t = 1, \dots, T, \quad (1)$$

where p is the power, $\mathbf{g}_k \in \mathbb{C}^M$ is the channel between user antenna k and the base station and $\psi(t)$ is an unknown phase shift that occurs at time t . Finally, $\mathbf{W}^p(t) \in \mathbb{C}^{M \times \tau_p}$ is AWGN with i.i.d. $\mathcal{CN}(0, 1)$ elements. The base station despreads the received signal with the pilot signal as

$$\mathbf{Y}^p(t) \phi(t) = \mathbf{y}^p(t) = \sqrt{p\tau_p} e^{i\psi(t)} \mathbf{g}_{S(t)} + \mathbf{w}^p(t), \quad \forall t, \quad (2)$$

where $\mathbf{w}^p(t) = \mathbf{W}^p(t) \phi(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. $\mathbf{y}^p(t)$ are the signals that the base station uses to estimate the channel between itself and the user.

Furthermore, in each time slot, there is an uplink data phase where we consider the symbol $x(t)$, $\mathbb{E}\{|x(t)|^2\} \leq 1$, in an

¹If \mathbf{G} is the $M \times K$ matrix consisting of the physical channel between the base station and the user antennas, detecting one of the antennas, $\mathbf{g}_k = \mathbf{G} \mathbf{e}_k$ is the same as detecting one of the pre-defined beams, $\mathbf{G} \mathbf{w}_k$, where \mathbf{w}_k is the k :th beamforming vector.

TABLE I
THE FOUR CASES OF CHANNEL ESTIMATION OVER MANY SLOTS.

	One antenna	Multiple antennas
Without phase shift	MMSE	Section III-A
With phase shift	Section III-B	Section III-C

arbitrary time slot t . When the user sends $x(t)$, the base station receives

$$\mathbf{y}^d(t) = \sqrt{p}e^{i\psi(t)}\mathbf{g}_{S(t)}x(t) + \mathbf{w}^d(t), \quad (3)$$

where $\mathbf{w}^d(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is AWGN.

III. CHANNEL ESTIMATION

There are four different cases of the channel estimation, see Table I. The simplest case where only one antenna is active and there is no phase shift ($\psi(t)$ absent from the above formulas) is trivial and can be solved with an MMSE estimate (assuming that the distribution of \mathbf{g} is known). The other cases are described below.

A. With Multiple Active Antennas, Without Phase Shifts

In the case without phase shifts, i.e., $\psi(t) = 0, t = 1, \dots, T$, the received despread pilot signal (2) simplifies to

$$\mathbf{y}^p(t) = \sqrt{p\tau_p}\mathbf{g}_{S(t)} + \mathbf{w}^p(t), \quad t = 1, \dots, T. \quad (4)$$

We first cluster the points to detect the time slots where the user selected antenna k , $\hat{\mathcal{T}}_k$. As with the true sets, the detected sets are also disjoint. Furthermore, we can express the detected selected antenna with $\hat{S}(t)$, defined analogously to $S(t)$. The clustering algorithms are described in Section IV. The channel estimates are formed from the despread pilot signal and the detected clusters:

$$\hat{\mathbf{g}}_k = \frac{1}{|\hat{\mathcal{T}}_k|\sqrt{p\tau_p}} \sum_{t \in \hat{\mathcal{T}}_k} \mathbf{y}^p(t) \quad (5)$$

$$= \frac{|\hat{\mathcal{T}}_k \cap \mathcal{T}_k|}{|\hat{\mathcal{T}}_k|} \mathbf{g}_k + \sum_{\substack{k'=1 \\ k' \neq k}}^K \frac{|\hat{\mathcal{T}}_k \cap \mathcal{T}_{k'}|}{|\hat{\mathcal{T}}_k|} \mathbf{g}_{k'} + \frac{1}{|\hat{\mathcal{T}}_k|} \sum_{t \in \hat{\mathcal{T}}_k} \frac{\mathbf{w}^p(t)}{\sqrt{p\tau_p}}. \quad (6)$$

B. With One Active Antenna, With Phase Shifts

In the case with only one active user antenna, i.e., $S(t) = 1, t = 1, \dots, T$, the received despread pilot signal (2) simplifies to (with removed antenna subscript)

$$\mathbf{y}^p(t) = \sqrt{p\tau_p}e^{i\psi(t)}\mathbf{g} + \mathbf{w}^p(t), \quad t = 1, \dots, T. \quad (7)$$

Here, we formulate the following least-squares problem:

$$\min_{\mathbf{g}, \{\psi(t)\}} \sum_{t=1}^T \|\mathbf{y}^p(t) - e^{i\psi(t)}\mathbf{g}\|^2 \quad (8)$$

to form estimates of the channel and the phase shifts. We experiment with two approaches of solving this problem:

- **Pairwise:** The first approach is called ‘‘Pairwise’’ and works by adapting the phase shift pairwise between time slots. We create the corrected observations:

$$\mathbf{y}^{p,\text{corr}}(1) = \mathbf{y}^p(1), \quad (9)$$

$$\mathbf{y}^{p,\text{corr}}(t) = e^{-i \arg(\mathbf{y}^{p,\text{corr}}(t-1)^H \mathbf{y}^p(t))} \mathbf{y}^p(t), \quad t = 2, \dots, T. \quad (10)$$

After we have formed the corrected observations, we can form our channel estimate to be

$$\hat{\mathbf{g}} = \frac{1}{T\sqrt{p\tau_p}} \sum_{t=1}^T \mathbf{y}^{p,\text{corr}}(t) \quad (11)$$

and the estimated phase shifts to be

$$\hat{\psi}(t) = \arg(\hat{\mathbf{g}}^H \mathbf{y}^p(t)), \quad t = 1, \dots, T. \quad (12)$$

- **Iterative ML:** The next approach is based on the maximum-likelihood (ML) estimate and is a cyclic optimization algorithm. We call it ‘‘iterative ML’’ here. We decouple the least-squares problem (8) to form two separate problems where we minimize over the channel in one and the phase shifts in the other. The minimizers are, for the channel estimation and phase shift estimation, respectively:

$$\hat{\mathbf{g}} = \frac{1}{T\sqrt{p\tau_p}} \sum_{t=1}^T e^{-i\psi(t)} \mathbf{y}^p(t) \quad (13)$$

and

$$\hat{\psi}(t) = \arg(\hat{\mathbf{g}}^H \mathbf{y}^p(t)), \quad t = 1, \dots, T. \quad (14)$$

In our ‘‘iterative ML’’ method we obtain the solution by iteratively insert (13) into (14) and vice versa.

C. With Multiple Active Antennas, With Phase Shifts

In the case where we have both phase shifts and multiple active antennas, we modify the clustering algorithms to handle the case where there are phase shifts between each time slot to form our detected clusters. This is discussed after introducing the candidates for the clustering algorithms in Section IV-D.

Additionally, we need to modify the channel estimation. This is basically done by using the same methods as in Section III-B for each detected antenna cluster, $\hat{\mathcal{T}}_k$, rather than going over all time slots, $t = 1, \dots, T$, e.g., $\frac{1}{T} \sum_{t=1}^T$ is replaced by $\frac{1}{|\hat{\mathcal{T}}_k|} \sum_{t \in \hat{\mathcal{T}}_k}$. That is, after obtaining the detected antenna clusters we apply the methods in the previous section on every cluster.

IV. CLUSTERING ALGORITHMS

Below, we present three candidates for clustering the active antennas.

A. K-means Clustering

The first clustering algorithm is the K-means clustering algorithm and the general form of it is presented in Algorithm 1. The general form can be directly applied to our problem without phase shifts. K-means iteratively finds the K centroids, $\mathbf{c}_1, \dots, \mathbf{c}_K$, of the T input points, $\{\mathbf{y}(t)\}$, while trying to minimize the Euclidian distance between the points in the cluster and the centroids, $\min_{\hat{\mathcal{T}}_k} \sum_{k=1}^K \sum_{t \in \hat{\mathcal{T}}_k} \|\mathbf{y}(t) - \mathbf{c}_k\|^2$ by taking the mean of all the points in each cluster [6]. The centroids

Algorithm 1 The K-means algorithm.

Input: $\{\mathbf{y}(t)\}, K$ **Output:** $\hat{\mathcal{T}}_k \forall k = 1, \dots, K$

- 1: Initialize $\mathbf{c}_k \forall k = 1, \dots, K$ with appropriate algorithm
 - 2: $\hat{\mathcal{T}}_k = \{t : \|\mathbf{y}(t) - \mathbf{c}_k\| \leq \|\mathbf{y}(t) - \mathbf{c}_{k'}\|, k' = 1, \dots, K\}$
 - 3: **while** $\{\hat{\mathcal{T}}_k\}$ unchanged **do**
 - 4: $\mathbf{c}_k \leftarrow \frac{1}{|\hat{\mathcal{T}}_k|} \sum_{t \in \hat{\mathcal{T}}_k} \mathbf{y}(t) \forall k = 1, \dots, K$
 - 5: $\hat{\mathcal{T}}_k = \{t : \|\mathbf{y}(t) - \mathbf{c}_k\| \leq \|\mathbf{y}(t) - \mathbf{c}_{k'}\|, k' = 1, \dots, K\}$
 - 6: **end while**
-

Algorithm 2 Modified EM algorithm, where $p(\mathbf{x} | \mathbf{m})$ is the circularly symmetric complex Gaussian probability density function with covariance matrix \mathbf{I} evaluated in point $\mathbf{x} - \mathbf{m}$.**Input:** $\{\mathbf{y}(t)\}, K$ **Output:** $\hat{\mathcal{T}}_k \forall k = 1, \dots, K$

- 1: Initialize $\mathbf{c}_k \forall k = 1, \dots, K$ with appropriate algorithm
 - 2: $\eta_k \leftarrow \frac{1}{K} \forall k = 1, \dots, K$
 - 3: **while** $\{\gamma_{t,k}\}$ unchanged **do**
 - 4: $\gamma_{t,k} \leftarrow \frac{\eta_k p(\mathbf{y}(t) | \mathbf{c}_k)}{\sum_{k'=1}^K \eta_{k'} p(\mathbf{y}(t) | \mathbf{c}_{k'})} \forall t, k$
 - 5: $\eta_k \leftarrow \frac{1}{T} \sum_{t=1}^T \gamma_{t,k}$
 - 6: $\mathbf{c}_k \leftarrow \frac{\sum_{t=1}^T \gamma_{t,k} \mathbf{y}(t)}{\sum_{t=1}^T \gamma_{t,k}}$
 - 7: **end while**
 - 8: $\hat{\mathcal{T}}_k = \{t : \gamma_{t,k} \geq \gamma_{t,k'}, k' = 1, \dots, K\}$
-

need to be initialized which can be done by randomly choosing them from the set of the points to be clustered or by a more sophisticated algorithm like the K-means++ algorithm [7].

B. Expectation-Maximization (EM) Algorithm

The clustering problem can be formulated as a Gaussian mixture model fit,

$$\max_{\{\mathbf{g}_k\}, \{\eta_k\}} \prod_{t=1}^T \sum_{k=1}^K \eta_k p(\mathbf{y}^p(t) | \sqrt{p\tau_p} \mathbf{g}_k) \quad (15)$$

where $p(\mathbf{y}^p(t) | \sqrt{p\tau_p} \mathbf{g}_k) = \frac{\exp(-\|\mathbf{y}^p(t) - \sqrt{p\tau_p} \mathbf{g}_k\|^2)}{\pi^M}$ is the circularly symmetric complex Gaussian probability density function with covariance matrix \mathbf{I} evaluated in point $\mathbf{y}^p(t) - \sqrt{p\tau_p} \mathbf{g}_k$ and η_k denotes the probability of any point belonging cluster k . (15) produces the ML estimate of the mixture parameters, \mathbf{g}_k and η_k . Therefore, a possibility is to use the EM algorithm which iteratively finds the ML estimate of the mixture parameters [6]. The algorithm is presented in Algorithm 2, where a slight modification is made to the original algorithm as we utilize the knowledge of the noise covariance matrix. The $\gamma_{t,k}$ variable denotes the probability of point $\mathbf{y}^p(t)$ belonging to cluster k and can be interpreted as “fuzzy”, or soft, clusters. In this work, we do not consider soft clusters and instead assign point $\mathbf{y}^p(t)$ to the cluster k with the highest probability, $\hat{S}(t) = \text{argmax}_{k'} \gamma_{t,k'}$.

Algorithm 3 The spectral clustering algorithm [8].

Input: $\{\mathbf{y}(t)\}, K, \rho$ **Output:** $\hat{\mathcal{T}}_k \forall k = 1, \dots, K$

- 1: Create $\mathbf{A} \in \mathbb{R}^{T \times T}$
 - 2: $A_{i,j} \leftarrow \exp\left(\frac{\|\mathbf{y}(i) - \mathbf{y}(j)\|^2}{\rho}\right) \forall i = 1, \dots, T, j = 1, \dots, T, i \neq j$
 - 3: $A_{i,i} \leftarrow 0 \forall i = 1, \dots, T$
 - 4: Create $\mathbf{D} \in \mathbb{R}^{T \times T}$
 - 5: $D_{i,j} \leftarrow 0 \forall i = 1, \dots, T, j = 1, \dots, T, i \neq j$
 - 6: $D_{i,i} \leftarrow \sum_{j=1}^T A_{i,j} \forall i = 1, \dots, T$
 - 7: Create $\mathbf{L} \in \mathbb{R}^{T \times T}$
 - 8: $\mathbf{L} \leftarrow \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}$
 - 9: Find the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_K \in \mathbb{R}^T$ corresponding to the K largest eigenvalues of \mathbf{L} and put them as columns of $\mathbf{V} \in \mathbb{R}^{T \times K}$
 - 10: Create $\mathbf{U} \in \mathbb{R}^{T \times K}$
 - 11: $U_{i,j} \leftarrow \frac{V_{i,j}}{\left(\sum_{j'=1}^K V_{i,j'}^2\right)^{\frac{1}{2}}}$
 - 12: Cluster the rows of \mathbf{U} with Algorithm 1 to obtain $\hat{\mathcal{T}}_k$
-

C. Spectral Clustering

The spectral clustering is done with the Ng-Jordan-Weiss algorithm [8] and is presented in Algorithm 3. The affinity matrix \mathbf{A} is describing the similarity between the points and ρ is deciding how fast the similarity should decrease with increased distance between the points.

D. Modifying the Clustering Algorithms to Consider Phase Shifts

When introducing phase shifts to our model, we need to take these into consideration when performing the clustering as well. This is achieved by slightly modifying the clustering algorithms.

We consider the pseudo metric $d(\mathbf{x}, \mathbf{y}) = \min_{\theta} \|\mathbf{x} - e^{i\theta} \mathbf{y}\| = \|\mathbf{x} - e^{-i \arg(\mathbf{x}^H \mathbf{y})} \mathbf{y}\|$ as a substitute for the Euclidean distance when assigning points to clusters in K-means, when calculating probabilities in the EM algorithm and when determining the similarity between two points in spectral clustering. The pseudo metric, $d(\mathbf{x}, \mathbf{y})$, has all the properties that a metric has except for $d(\mathbf{x}, \mathbf{y}) = 0 \not\Rightarrow \mathbf{x} = \mathbf{y}$ [9, Ch. 4], i.e., a pseudo metric is non-negative, symmetric and fulfills the triangle inequality.

Additionally, in K-means and the EM algorithm, the cluster centers are calculated with either the “Pairwise” or “Iterative ML” method instead of taking the (weighted) mean.

V. ACHIEVABLE RATE

To evaluate the performance of our channel estimation methods we use the uplink achievable rate, which is a good measure of performance given that we study long code words and that suitable channel coding techniques are used.

To calculate the achievable rate, we consider the uplink received signal (3). We multiply the received signal with the conjugate transpose of the receive combining vector $\mathbf{v}(t)$ to form

$$\mathbf{v}(t)^H \mathbf{y}^d(t) = \sqrt{p} e^{i\psi(t)} \mathbf{v}(t)^H \mathbf{g}_{S(t)} x(t) + \mathbf{v}(t)^H \mathbf{w}^d(t). \quad (16)$$

We study three different capacity bounds; the use-and-then-forget bound [1, Sec. 2.3.4] [10, Thm. 4.4] with two different receive combining vectors and the estimation bound (EB) in [11, Lemma 3]: The two receive combining vectors that we consider for the use-and-then-forget bound are the channel estimate scaled by a deterministic constant (UaF) $\mathbf{v}(t) = \frac{e^{i\psi(t)} \hat{\mathbf{g}}_{S(t)}}{M}$ and the the channel estimate divided by its squared norm (UaF N) $\mathbf{v}(t) = \frac{e^{i\psi(t)} \hat{\mathbf{g}}_{S(t)}}{\|\hat{\mathbf{g}}_{S(t)}\|^2}$. To adhere to the system model for EB in [11, Lemma 3] we scale the channel estimate by its norm $\mathbf{v}(t) = \frac{e^{i\psi(t)} \hat{\mathbf{g}}_{S(t)}}{\|\hat{\mathbf{g}}_{S(t)}\|}$. The system model in [11] is a downlink multiuser MIMO system with linearly precoded transmission to the K single-antenna users. To apply the EB bound from [11] to our system model, we state the conjugate of (16) with the notation from [11] under the braces (without the block indices);

$$\underbrace{\mathbf{y}^d(t)^H \mathbf{v}(t)}_{=y_k} = \underbrace{e^{-i\psi(t)} \mathbf{g}_{S(t)}^H}_{=\mathbf{h}_k^H} \underbrace{\sqrt{p} \mathbf{v}(t) x(t)^*}_{=\sqrt{\mathcal{E}_k} = \mathbf{v}_k = \mathbf{s}_k} + \underbrace{\mathbf{w}^d(t)^H \mathbf{v}(t)}_{=z_k}.$$

Unit precoding vectors are assumed in [11], i.e., $\|\mathbf{v}_k\|^2 = 1$, which is why we use the normalized receive combining vector in our model. Also, z_k should be independent of \mathbf{v}_k which is also fulfilled in our model with the normalized $\mathbf{v}(t)$. This can be verified by noting that the probability distribution of $\mathbf{w}^d(t)^H \mathbf{v}(t)$ conditioned $\mathbf{v}(t)$ does not depend on $\mathbf{v}(t)$ for any unit-norm $\mathbf{v}(t)$.

The achievable rates are as follows; the use-and-then-forget bound

$$C \geq \log_2 \left(1 + \frac{p \mathbb{E} \left\{ \left| e^{i\psi(t)} \mathbf{v}(t)^H \mathbf{g}_{S(t)} \right|^2 \right\}}{\mathbb{V} \left\{ \mathbf{v}(t)^H \mathbf{w}^d(t) \right\} + p \mathbb{V} \left\{ e^{i\psi(t)} \mathbf{v}(t)^H \mathbf{g}_{S(t)} \right\}} \right), \quad (17)$$

where UaF and UaF N are obtained by inserting the corresponding receive combining vector and the expectations are across channel realizations and t , and finally EB with normalized receive combining vector

$$C^{\text{EB}} \geq \mathbb{E} \left\{ \log_2 \left(1 + p \left| \frac{e^{i(\psi(t) - \hat{\psi}(t))} \hat{\mathbf{g}}_{S(t)}^H \mathbf{g}_{S(t)}}{\|\hat{\mathbf{g}}_{S(t)}\|} \right|^2 \right) \right\} - \frac{1}{\tau_c} \log_2 \left(1 + \tau_c p \mathbb{V} \left\{ \frac{e^{i(\psi(t) - \hat{\psi}(t))} \hat{\mathbf{g}}_{S(t)}^H \mathbf{g}_{S(t)}}{\|\hat{\mathbf{g}}_{S(t)}\|} \right\} \right). \quad (18)$$

VI. NUMERICAL SIMULATIONS

The channels are in the simulations assumed to be uncorrelated Rayleigh fading and independent between the user antennas, $\mathbf{g}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. The large scale fading is normalized and included in p . The phase shift $\psi(t)$ is assumed to

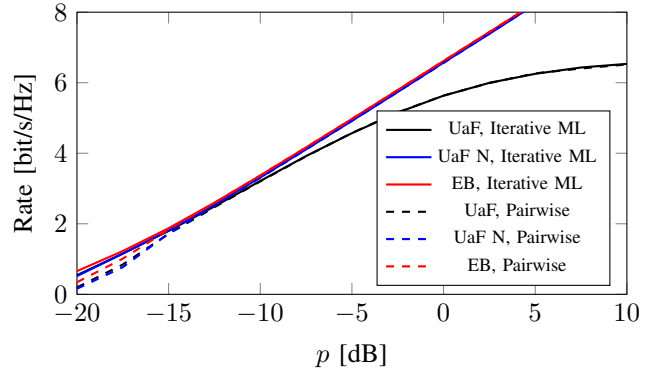


Fig. 1. Simulation of achievable rates with one active antenna, with phase shifts as in Section III-B, $M = 100$, $T = 20$, $\tau_p = 10$.

be uniformly distributed between $-\pi$ and π and independent between time slots. When having multiple antennas, the terminal is assumed to have four antennas, $K = 4$. The selected antenna is uniformly chosen at each time slot. Throughout the simulations we use $\tau_c = 200$ as in [10]. Furthermore, the simulations depicted are with $M = 100$, $T = 20$, and $\tau_p = 10$. We insert $\mathbf{y}^p(t)$ in the clustering algorithms. Both the K-means and EM algorithm are initialized by the K-means++ algorithm [7]. The scaling parameter, ρ , in the spectral clustering is set to $M p \tau_p$ to not make the distances explode for high SNR or large number of antennas. We limit the number of iterations in K-means and EM and set the maximum number of iterations to 1000. Because of some numerical issues we re-initialize the EM algorithm if $\sum_{k=1}^K \eta_k P(\mathbf{y}(t) | \mathbf{c}_k) = 0$ (to machine precision) for any t , that is the probability of a point belonging to any cluster is zero to machine precision in the simulation. The K-means algorithm, EM algorithm and spectral clustering algorithm are denoted by KM, EM and SC, respectively in the figure legends. We run the ‘‘Iterative ML’’ for 5 iterations, as it converges fairly quickly.

In the first simulation, we compare the different methods of attending to the phase shifts without dealing with multiple antennas, i.e., the case in Section III-B. The result is depicted in Fig. 1. The difference between the iterative ML and the pairwise method is most noticeable at low SNR where the iterative ML method is outperforming the pairwise method. This is achieved at a cost of increasing complexity. Although not depicted here, when having fewer antennas at the base station, the iterative ML and the pairwise method get closer in performance.

Next, we study the case with multiple antennas but without phase shifts, i.e., Section III-A, to compare the different clustering algorithms. The result is depicted in Fig. 2. Here we can see that the EM algorithm is clearly the best of the three methods at high SNR, and that spectral clustering is slightly better at low SNR. In fact, the spectral clustering algorithm gains in performance when there are more points to cluster, i.e., T is larger. In this figure, we can also more clearly see the difference between the different achievable rates. The UaF bound is the worst of the three as it is very pessimistic. This bound can be calculated in closed form in some system

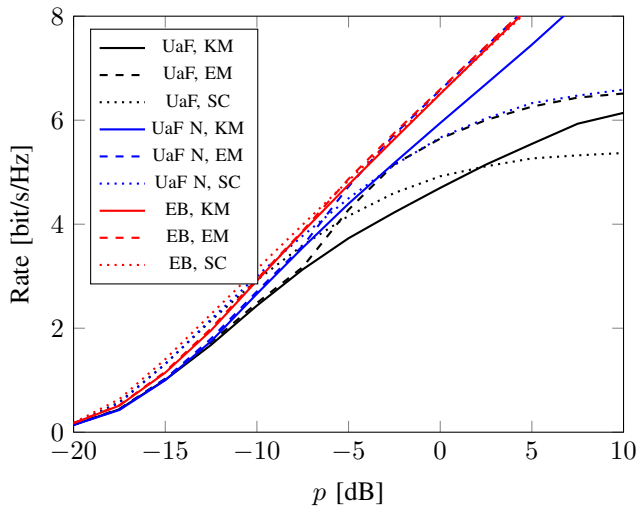


Fig. 2. Simulation of achievable rates with multiple active antennas, without phase shifts as in Section III-A, $M = 100$, $T = 20$, $\tau_p = 10$.

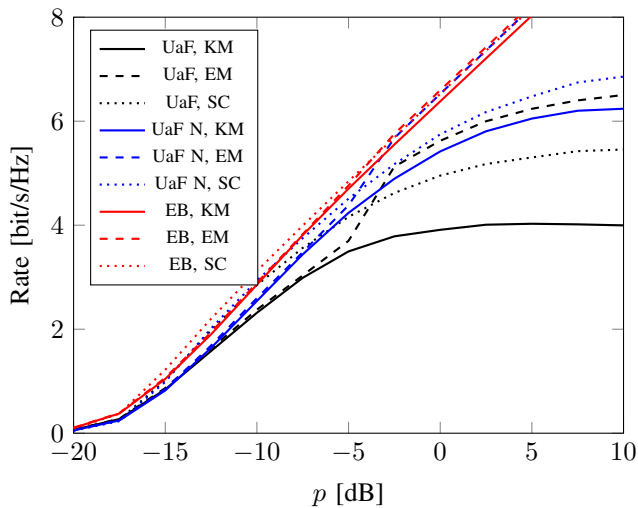


Fig. 3. Simulation of achievable rates with multiple active antennas, with phase shifts as in Section III-C with the iterative ML method used to deal with the phase shifts, $M = 100$, $T = 20$, $\tau_p = 10$.

models and is tight when the channel hardens which is why it is commonly used. Here, we do not have much channel hardening due to occasional missed detection of the antenna clusters. By scaling the receive combining vector by its norm squared we obtain a much better bound, UaF N. However, the bound that is the best in this case is EB, which is because the second term in (18) becomes very small when τ_c is much larger than the number of users which is the case here ($200 \gg 1$). EB almost attains the achievable rate when the channel is perfectly known, $\log_2(1 + Mp)$, at high SNR.

Finally, we combine the clustering and phase shifts and consider the case in Section III-C. Fig. 3 shows a simulation where the iterative ML method is used to deal with the phase shifts. We can see that the rates suffer but that the modifications done to the clustering algorithms are fairly adequate to handle the phase shifts.

To answer how much we gain from utilizing multiple time

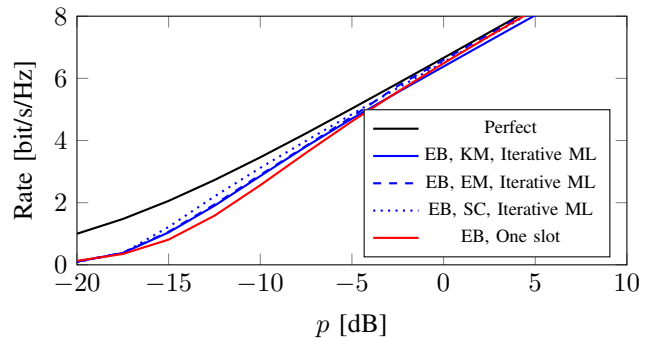


Fig. 4. Comparison of achievable rates when using only a single slot for channel estimation and the proposed method of utilizing multiple slots, $M = 100$, $T = 20$, $\tau_p = 10$. The curve denoted by “Perfect” is the case where the channel is exactly known at the base station.

slots when estimating the channel, we compare the case where the base station estimates the channel separately in each slot (with the least-squares estimator) to our case with phase shifts and multiple antennas. The result is shown in Fig. 4. In this figure the achievable rate when having perfect channel knowledge is also considered. We see that we will improve the rates most at relatively low SNR; there is an improvement of approximately 1.5 dB at 1 bit/s/Hz.

VII. CONCLUSIONS

In this paper we considered the problem of filtering the channel estimates between time slots of a non-coherent 5G user terminal. Our suggested methods detect the selected antenna and estimate the phase shift in each time slot. We show that the uplink data rate can be improved by considering more time slots than only considering one time slot for channel estimation.

REFERENCES

- [1] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
- [2] Y. Li, “Pilot-symbol-aided channel estimation for OFDM in wireless systems,” *IEEE Transactions on Vehicular Technology*, vol. 49, no. 4, pp. 1207–1215, July 2000.
- [3] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, “Pilot-assisted time-varying channel estimation for OFDM systems,” *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 2226–2238, May 2007.
- [4] M. K. Ozdemir and H. Arslan, “Channel estimation for wireless OFDM systems,” *IEEE Communications Surveys Tutorials*, vol. 9, no. 2, pp. 18–48, Second 2007.
- [5] A. Zaidi, F. Athley, J. Medbo, U. Gustavsson, G. Durisi, and X. Chen, *5G Physical Layer*. Academic Press, 2018.
- [6] J. Han, M. Kamber, and J. Pei, *Data Mining Concepts and Techniques*, 3rd ed. Elsevier, 2012.
- [7] D. Arthur and S. Vassilvitskii, “k-means++: The advantages of careful seeding,” in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, 2007, pp. 1027–1035.
- [8] A. Y. Ng, M. I. Jordan, and Y. Weiss, “On spectral clustering: Analysis and an algorithm,” in *Advances in neural information processing systems*, 2002, pp. 849–856.
- [9] J. L. Kelley, *General topology*, ser. The University series in higher mathematics. Van Nostrand, 1955.
- [10] E. Björnson, J. Hoydis, and L. Sanguinetti, “Massive MIMO networks: Spectral, energy, and hardware efficiency,” *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.
- [11] G. Caire, “On the ergodic rate lower bounds with applications to massive MIMO,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3258–3268, May 2018.