

Training Design and Channel Estimation in Uplink Cloud Radio Access Networks

Xinqian Xie, Mugen Peng, H. Vincent Poor

Abstract

To decrease the training overhead and improve the channel estimation accuracy under time-varying environments in uplink cloud radio access networks (C-RANs), a superimposed-segment training design is proposed whose core idea is that each mobile station puts a periodic training sequence on the top of the data signal, and remote radio heads (RRHs) insert a separate pilot prior to the received signal before forwarding to the centralized base band unit (BBU) pool. Moreover, a complex-exponential basis-expansion-model (CE-BEM) based maximum a posteriori probability (MAP) channel estimation algorithm is developed, where the BEM coefficients of access links (ALs) and the channel fading of wireless backhaul links (BLs) are first obtained, after which the time-domain channel samples of ALs are restored in terms of maximizing the average effective signal-to-noise ratio (AESNR). Simulation results show that the proposed channel estimation algorithm can effectively decrease the estimation mean square error and increase the AESNR in C-RANs, thus significantly outperforming the existing solutions.

Index Terms

Channel estimation, cloud radio access networks, time-varying environment.

I. INTRODUCTION

Cloud radio access networks (C-RANs) have received considerable research interest as one of the most promising solutions to mitigate interference, fulfil energy efficiency demands and support high-rate transmission in the fifth generation cellular network [1]. In C-RANs, a large number of remote radio heads (RRHs) are deployed, which operate as non-regenerative relays to forward received signals from mobile stations (MSs) to the centralized base band unit (BBU) pool through wire/wireless backhaul

M. Peng and H. Vincent Poor are with the School of Engineering and Applied Science, Princeton University, Princeton, NJ, USA. X. Xie is with Wireless Signal Processing and Network Lab, Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts & Telecommunications, Beijing, China.

links for uplink transmission [2]. To suppress the inter-RRH interference by using cooperative processing techniques at the BBU pool, the channel state information (CSI) of both the radio access links (ALs) and wireless backhaul links (BLs) are required [3]. In [4], though a segment-training scheme was proposed to estimate the individual channel coefficients for two-hop scenario under flat fading environments, this proposal results in high overhead consumption for backhaul transmission since the RRHs need to forward both AL and BL training sequences to the BBU pool [5]. On the other hand, the superimposed-training scheme [6], where a training sequence is superimposed on the data signal, can significantly reduce the overhead and is valid to perform channel estimation for time-varying environments using complex-exponential basis expansion model (CE-BEM) [7]. However, straightforward implementation of superimposed training in C-RANs would degrade transmission quality due to the fact that superimposing both AL and BL training sequences on the data signal declines the effective signal-to-noise ratio (SNR) [8].

Motivated to reduce the training overhead and enhance the channel estimation performance at the BBU pool, a superimposed-segment training design is proposed in this letter, where superimposed-training is implemented for the radio AL while the segment-training is applied for the wireless BL. Moreover, based on the training design, a CE-BEM based maximum a posteriori probability (MAP) channel estimation algorithm is developed, where the BEM coefficients of the time-varying radio AL and the channel fading of the quasi-static wireless BL are first obtained, after which the time-domain channel samples of the radio AL are restored in terms of maximizing the average effective SNR.

Notation: The transpose, Hermitian and inverse of a matrix are denoted by $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^{-1}$, respectively; $\|\cdot\|$ represents the two-norm of a vector; $|\cdot|$ defines the magnitude of a complex argument; \otimes is the Kronecker product; $\text{diag}(a_1, a_2, \dots, a_M)$ denotes the diagonal matrix with a_m being the m th diagonal element; $\text{tr}(\cdot)$ represents the trace of a matrix; \mathbf{I}_N and $\mathbf{0}_N$ are the $N \times N$ unit diagonal matrix and zero matrix, respectively; $\mathbf{1}_N$ stands for the $N \times 1$ vector with each entry being unit value; $\mathcal{E}\{\cdot\}$ denotes the expectation of a random variable, and \hat{x} represents its estimate.

II. SYSTEM MODEL AND TRAINING DESIGN

Consider a C-RAN consisting of one BBU pool and multiple RRHs depicted in Fig. 1. The RRHs operate in half-duplex modes, and different MSs served by the same RRU are allocated with a single subcarrier through the orthogonal frequency division multiplexing access (OFDMA) technique. It is assumed that MSs move continuously, while RRUs remain fixed. Thus, the channels of radio ALs would vary during one transmission block, while those of wireless BLs undergo quasi-static flat fading.

Due to the orthogonality characteristics of OFDMA for the accessing of multiple MSs, we can focus on the transmission of only a single MS. Let \mathbf{b} and \mathbf{t}_s denote the data vector and cyclical training sequence transmitted from the MS, respectively. The training sequence from the RRH is denoted by \mathbf{t}_r . The n -th channel sample of the time-varying radio AL is denoted by $h(n)$ with mean zero and variance v_h , while the channel fading of the quasi-static flat BL is denoted by g which has the complex Gaussian distribution of mean zero and variance v_g . The transmit power of the MS and RRH are denoted by P_s and P_r , respectively. The noise variance at the RRH and BBU pool is denoted by σ_n^2 . It is assumed that the BBU pool acquires the knowledge of \mathbf{t}_s , \mathbf{t}_r , v_h , v_g , P_s , P_r , and σ_n^2 .

During each transmission block, the MS transmits a signal \mathbf{s} of N_s symbol length to the RRU, in which the n -th entry of \mathbf{s} is given by

$$s(n) = \sqrt{1-\epsilon}b(n) + \sqrt{\epsilon}t_s(n), \quad 0 \leq n \leq (N_s-1), \quad (1)$$

where $b(n)$ denotes the n -th entry of \mathbf{b} with M-ary phase shift keying (MPSK) modulation constrained by $\mathcal{E}\{|b(n)|^2\} = P_s$, and $t_s(n)$ represents the n -th entry of \mathbf{t}_s with $|t_s(n)|^2 = P_s$ whose period is denoted by N_p . The ϵ is within $0 < \epsilon < 1$. Without loss of generality, we further assume that $K = \frac{N_s}{N_p}$ is an integer [7]. The n -th observation at the RRH is written as

$$x_R(n) = h(n)s(n) + w_R(n), \quad 0 \leq n \leq (N_s-1), \quad (2)$$

where $w_R(n)$ is additive white Gaussian noise (AWGN) at the RRH. Then the RRH scales the received signal by $\alpha = \sqrt{\frac{P_r}{v_h P_s + \sigma_n^2}}$, and inserts \mathbf{t}_r prior to the received signal. The sequence \mathbf{t}_r is of N_r length and its n -th entry satisfies $|t_r(n)|^2 = P_r$. The BBU pool receives two separate signals as

$$\mathbf{x}_s = \alpha g \cdot \text{diag}(\mathbf{h}) \cdot \mathbf{s} + \alpha g \mathbf{w}_R + \mathbf{w}_{D_s}, \quad (3)$$

$$\mathbf{x}_r = g \mathbf{t}_r + \mathbf{w}_{D_r}, \quad (4)$$

where $\mathbf{w}_R = [w_R(1) \dots w_R(N_s)]^T$, $\mathbf{w}_{D_s} = [w_{D_s}(1) \dots w_{D_s}(N_s)]^T$ and $\mathbf{w}_{D_r} = [w_{D_r}(1), \dots, w_{D_r}(N_r)]^T$ are AWGN vectors with each entry having variance σ_n^2 . In order to perform coherent reception and adopt cooperative processing techniques at the BBU pool, the knowledge of $h(n)$'s and g should be obtained, which is explained in the following section.

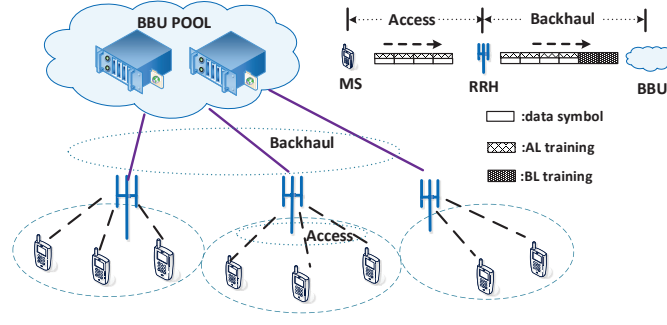


Fig. 1. System model and the superimposed-segment training design.

III. CE-BEM BASED MAP CHANNEL ESTIMATION ALGORITHM

The CE-BEM for time-selective but frequency-flat fading channels in [9] [10] is chosen to model the time-varying AL as

$$h(n) = \sum_{q=-Q}^Q \lambda_q e^{j \frac{2\pi q n}{N_s}}, \quad 0 \leq n \leq (N_s - 1), \quad (5)$$

where λ_q 's are the BEM coefficients that remain invariant within one transmission block, and λ_q has the complex Gaussian distribution with mean zero and variance v_q constrained to $\sum_{q=-Q}^Q v_q = v_h$. Substituting (5) into (3), \mathbf{x}_r can be rewritten as

$$\mathbf{x}_s = \alpha g \mathbf{D} [\mathbf{I}_{2Q+1} \otimes \mathbf{s}] \boldsymbol{\lambda} + \alpha g \mathbf{w}_R + \mathbf{w}_{D_s}, \quad (6)$$

where $\mathbf{D} = [\mathbf{D}_{-Q}, \dots, \mathbf{D}_Q]$ is an $N_s \times (2Q+1) N_s$ dimensional matrix with $\mathbf{D}_q = \text{diag}\left(1, e^{j \frac{2\pi q}{N_s}}, \dots, e^{j \frac{2\pi q(N_s-1)}{N_s}}\right)$.

By defining $\mathbf{J} = \frac{1}{K} \mathbf{1}_K^T \otimes \mathbf{I}_{N_p}$ of $N_p \times N_s$ dimension, we can obtain

$$\mathbf{J} \mathbf{D}_q \mathbf{t}_s = \begin{cases} \tilde{\mathbf{t}}_s, & q = 0, \\ 0, & q \neq 0, \end{cases} \quad (7a)$$

$$(7b)$$

where $\tilde{\mathbf{t}}_s$ is an $N_p \times 1$ dimensional vector. Left multiplying \mathbf{x}_s by $(\mathbf{I}_{Q+1} \otimes \mathbf{J}) \mathbf{D}^H$ yields

$$\mathbf{r} = (\mathbf{I}_{Q+1} \otimes \mathbf{J}) \mathbf{D}^H \mathbf{x}_s, \quad (8)$$

whose $[(q+Q)N_p+1]$ -th to $[(q+Q+1)N_p]$ -th entries denoted by \mathbf{r}_q can be expressed as

$$\mathbf{r}_q = \alpha g \tilde{\mathbf{t}}_s \lambda_q + \underbrace{\alpha g \sum_{l=-Q}^Q \mathbf{J} \mathbf{D}_{l-q} \mathbf{b} \lambda_l + \mathbf{J} \mathbf{D}_q^H (\alpha g \mathbf{w}_R + \mathbf{w}_{D_s})}_{\mathbf{w}_q}. \quad (9)$$

It is testified that

$$\mathcal{E}\{\mathbf{w}_{q1}\mathbf{w}_{q1}^H\} = v_n\mathbf{I}_{N_p}, \quad \mathcal{E}\{\mathbf{w}_{q1}\mathbf{w}_{q2}^H\} = \mathbf{0}_{N_p} \quad (10)$$

for $q_1 \neq q_2$ with $v_n = \frac{\alpha^2 v_g v_h (1-\epsilon)^{P_s} + (\alpha^2 v_g + 1)\sigma_n^2}{K}$. Although \mathbf{w}_q is not a Gaussian vector, it is effective to use the Gaussian distribution to model the noise behavior for estimation problems [14]. Thus, we choose the following function to be the nominal likelihood function of \mathbf{r} as

$$p(\mathbf{r}|\boldsymbol{\lambda}, g) = \prod_{q=1}^Q \left(\frac{1}{\pi v_n} \right)^{N_p} e^{-\frac{\|\mathbf{r}_q - \alpha g \lambda_q \tilde{\mathbf{t}}_s\|^2}{v_n}}. \quad (11)$$

A. Estimation for λ_q 's and g

Defining $\boldsymbol{\theta} = [\lambda_{-Q}, \dots, \lambda_Q, g]^T$, the MAP estimation for $\boldsymbol{\theta}$ gives

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= \arg \max_{\boldsymbol{\theta}} \left\{ p(\mathbf{r}|\boldsymbol{\lambda}, g) p(\mathbf{x}_r|g) p(\boldsymbol{\lambda}) p(g) \right\} \\ &= \arg \min_{\boldsymbol{\theta}} \left\{ \underbrace{\sum_{q=-Q}^Q \left\{ \frac{\|\mathbf{r}_q - \alpha g \lambda_q \tilde{\mathbf{t}}_s\|^2}{v_n} + \frac{|\lambda_q|^2}{v_q} \right\}}_{\mathcal{L}(\boldsymbol{\theta})} + \frac{\|\mathbf{x}_r - g \mathbf{t}_r\|^2}{\sigma_n^2} + \frac{|g|^2}{v_g} \right\}, \end{aligned} \quad (12)$$

where $p(\mathbf{x}_r|g)$, $p(\boldsymbol{\lambda})$ and $p(g)$ are Gaussian distribution functions. With a given g , the estimate of λ_q can be obtained as

$$\hat{\lambda}_q = \frac{\alpha g^H \tilde{\mathbf{t}}_s^H \mathbf{r}_q}{\alpha^2 |g|^2 \|\tilde{\mathbf{t}}_s\|^2 + \frac{v_n}{v_q}}, \quad -Q \leq q \leq Q. \quad (13)$$

Substituting (13) into (12), the estimate of g can be obtained from

$$\hat{g} = \arg \min_g \left\{ \underbrace{\frac{\|\mathbf{x}_r - g \mathbf{t}_r\|^2}{\sigma_n^2} + \frac{|g|^2}{v_g} + \frac{\alpha^2 |g|^2 \sum_{q=-Q}^Q |\mathbf{t}_s^H \mathbf{r}_q|^2}{\alpha^2 v_n \|\tilde{\mathbf{t}}_s\|^2 |g|^2 + \frac{v_n^2}{v_q}}}_{\mathcal{L}(g)} \right\}. \quad (14)$$

Note that, only the first term in $\mathcal{L}(g)$ relates to the phase of g denoted by $\angle g$, and thus the estimate of $\angle g$ can be directly estimated from minimizing $\|\mathbf{x}_r - g \mathbf{t}_r\|^2$ as

$$\widehat{\angle g} = \arg \min \left\{ \|\mathbf{x}_r - g \mathbf{t}_r\|^2 \right\} = \angle(\mathbf{t}_r^H \mathbf{x}_r). \quad (15)$$

The estimate of $|g|$ must be either a local minima of $\mathcal{L}(|g|, \angle g)$ or at the boundary $|g| = 0$, which can be obtained from solving $\frac{\partial \mathcal{L}(|g|, \angle g)}{\partial |g|} = 0$. Unfortunately, a closed-form expression for \hat{g} is hard to derive since $\frac{\partial \mathcal{L}(|g|, \angle g)}{\partial |g|}$ is an $m (\geq 4)$ th-order polynomial of $|g|$, and thus numerical methods such

as one dimensional search are needed to compute the value of $|\widehat{g}|$. To reduce the complexity of such approaches, an iterative approach is developed whereby \hat{g} is initialized from

$$\hat{g} = \arg \max_g \left\{ p(\mathbf{x}_r | g) p(g) \right\} = \frac{\mathbf{t}_r^H \mathbf{x}_r}{\|\mathbf{t}_r\|^2 + \frac{\sigma_n^2}{v_g}}. \quad (16)$$

With \hat{g} obtained, $\hat{\lambda}_q$'s can be estimated according to (13) with \hat{g} in place of g . Then, \hat{g} can be further updated by substituting $\hat{\lambda}_q$'s into (15) as

$$\hat{g} = \frac{\sum_{q=-Q}^Q \frac{\alpha \hat{\lambda}_q^H \tilde{\mathbf{t}}_s^H \mathbf{r}_q}{v_n} + \frac{\mathbf{t}_r^H \mathbf{x}_r}{\sigma_n^2}}{\sum_{q=-Q}^Q \frac{\alpha^2 |\hat{\lambda}_q|^2 \|\tilde{\mathbf{t}}_s\|^2}{v_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_n^2} + \frac{1}{v_g}}. \quad (17)$$

B. Restoration for Channel Samples $h(n)$'s

With $\hat{\lambda}_q$'s obtained, $\hat{h}(n)$ is restored as

$$\hat{h}(n) = \sum_{q=-Q}^Q \left\{ \eta_q \cdot \hat{\lambda}_q e^{j \frac{2\pi q n}{N_s}} \right\}, \quad 0 \leq n \leq (N_s - 1), \quad (18)$$

where η_q 's are real factors. The vector $\boldsymbol{\eta}$ that maximizes the average effective SNR (AESNR) [11] denoted by $\boldsymbol{\eta}^*$ is obtained from

$$\boldsymbol{\eta}^* = \arg \max_{\boldsymbol{\eta}} \frac{\mathcal{E} \left\{ \mathcal{E} \left\{ |\hat{g} \hat{h}(n)|^2 \mid \boldsymbol{\lambda}, g \right\} \right\} (1 - \epsilon)}{\underbrace{\mathcal{E} \left\{ \mathcal{E} \left\{ |\hat{g} \hat{h}(n) - g h(n)|^2 \mid \boldsymbol{\lambda}, g \right\} + (|g|^2 + \frac{1}{\alpha^2}) \frac{\sigma_n^2}{P_s} \right\}}_{\tilde{\gamma}}}. \quad (19)$$

Define $\phi_{q,n} = e^{j \frac{2\pi q n}{N_s}}$, and denote Δg and $\Delta \lambda_q$ to be the estimation error of g and λ_q , respectively. It is obtained that

$$\mathcal{E} \left\{ |\hat{g} \hat{h}(n)|^2 \mid \boldsymbol{\lambda}, g \right\} = |g|^2 \sum_{q=-Q}^Q \underbrace{\eta_q^2 \mathcal{E} \left\{ |\Delta \lambda_q|^2 \right\}}_{\delta_{\lambda_q}} + \underbrace{\left| \sum_{q=-Q}^Q \eta_q \lambda_q \phi_{q,n} \right|^2}_{\tilde{h}(n)} \underbrace{\mathcal{E} \left\{ |\Delta g|^2 \right\}}_{\delta_g} + |g|^2 \left| \sum_{q=-Q}^Q \eta_q \lambda_q \phi_{q,n} \right|^2 + \mathcal{E} \left\{ \left| \Delta g \sum_{q=-Q}^Q \eta_q \Delta \lambda_q \phi_{q,n} \right|^2 \right\}. \quad (20)$$

Note that, $\delta_{\lambda_q}, \delta_g \sim \mathcal{O} \left(\frac{\sigma_n^2}{P_s} \right)$ while the last term in (20) has the order of $\mathcal{O} \left[\left(\frac{\sigma_n^2}{P_s} \right)^2 \right]$, thus we remove the last term for high SNR approximation, i.e.,

$$\mathcal{E} \left\{ |\hat{g} \hat{h}(n)|^2 \mid \boldsymbol{\lambda}, g \right\} = |g|^2 \sum_{q=-Q}^Q \eta_q^2 \delta_{\lambda_q} + |\tilde{h}(n)|^2 \delta_g + |g|^2 |\tilde{h}(n)|^2. \quad (21)$$

Similarly,

$$\mathcal{E}\left\{\left|\hat{g}h(n)-gh(n)\right|^2\middle|\boldsymbol{\lambda},g\right\}\approx|g|^2\sum_{q=-Q}^Q\eta_q^2\delta_{\lambda_q}+|\tilde{h}(n)|^2\delta_g+|g|^2\left|\sum_{q=-Q}^Q(\eta_q-1)\lambda_q\phi_{n,q}\right|^2. \quad (22)$$

Due to the non-linearity of the MAP estimation, it is hard to derive the corresponding MSE expressions in closed forms. Moreover, it is known that the MAP estimation MSEs converge to the Cramér-Rao bound (CRB) [12] when the training length is sufficiently large, and thus it is effective to use the CRBs in the computation of the AESNR as

$$\delta_{\lambda_q}=\text{CRB}_{\lambda_q}=\frac{v_n}{\alpha^2|g|^2\|\tilde{\mathbf{t}}_s\|^2}, \quad \delta_g=\text{CRB}_g=\frac{1}{\frac{\alpha^2\|\boldsymbol{\lambda}\|^2\|\tilde{\mathbf{t}}_s\|^2+\|\mathbf{t}_r\|^2}{v_n}+\frac{\sigma_n^2}{\sigma_n^2}}. \quad (23)$$

Substituting the above expressions for δ_{λ_q} and δ_g into $\bar{\gamma}$ yields

$$\bar{\gamma}=\frac{\boldsymbol{\eta}^T\boldsymbol{\Xi}\boldsymbol{\eta}(1-\epsilon)}{\boldsymbol{\eta}^T\boldsymbol{\Xi}\boldsymbol{\eta}-2\mathbf{1}_{2Q+1}^T\boldsymbol{\Upsilon}\boldsymbol{\eta}+C}, \quad (24)$$

where

$$\boldsymbol{\Upsilon}=\frac{\alpha^2v_g\|\tilde{\mathbf{t}}_s\|^2}{v_n}\mathbf{R}_\lambda^2+\frac{\alpha^2v_g\|\tilde{\mathbf{t}}_s\|^2}{v_n}\text{tr}\{\mathbf{R}_\lambda\}\mathbf{R}_\lambda+\frac{v_g\|\mathbf{t}_r\|^2}{\sigma_n^2}\mathbf{R}_\lambda, \quad (25)$$

$$\boldsymbol{\Xi}=\boldsymbol{\Upsilon}+\mathbf{R}_\lambda+\left(v_h+\frac{v_n\|\mathbf{t}_r\|^2}{\sigma_n^2\alpha^2\|\tilde{\mathbf{t}}_s\|^2}\right)\mathbf{I}_{2Q+1}, \quad (26)$$

$$C=\mathbf{1}_{2Q+1}^T\boldsymbol{\Upsilon}\mathbf{1}_{2Q+1}+\left(v_g+\frac{1}{\alpha^2}\right)\left[\frac{v_h\alpha^2\|\tilde{\mathbf{t}}_s\|^2}{v_n}+\frac{\|\mathbf{t}_r\|^2}{\sigma_n^2}\right]\frac{\sigma_n^2}{P_s}, \quad (27)$$

and $\mathbf{R}_\lambda=\text{diag}(v_{-Q}\dots v_Q)$. On setting $\varphi^*=(\boldsymbol{\eta}^*)^T\boldsymbol{\Xi}\boldsymbol{\eta}^*$, the optimization for $\boldsymbol{\eta}$ transforms to

$$\max_{\boldsymbol{\eta}} \quad 2\mathbf{1}_{2Q+1}^T\boldsymbol{\Upsilon}\boldsymbol{\eta}, \quad (28)$$

$$\text{s.t.} \quad \boldsymbol{\eta}^T\boldsymbol{\Xi}\boldsymbol{\eta}=\varphi^*. \quad (29)$$

Clearly, the optimization problem described in (28) constrained by (29) is concave, and $\boldsymbol{\eta}^*$ can be directly obtained from the Lagrange dual function as

$$\boldsymbol{\eta}^*=\frac{\sqrt{\varphi^*}\boldsymbol{\Xi}^{-1}\boldsymbol{\Upsilon}\mathbf{1}_{2Q+1}}{\sqrt{\mathbf{1}_{2Q+1}^T\boldsymbol{\Upsilon}\boldsymbol{\Xi}^{-1}\boldsymbol{\Upsilon}\mathbf{1}_{2Q+1}}}. \quad (30)$$

Substituting (30) back to (24), the optimization problem becomes

$$\max_{\varphi} \quad \frac{\varphi}{\varphi-2\sqrt{\mathbf{1}_{2Q+1}^T\boldsymbol{\Upsilon}\boldsymbol{\Xi}^{-1}\boldsymbol{\Upsilon}\mathbf{1}_{2Q+1}}\sqrt{\varphi}+C}, \quad (31)$$

$$\text{s.t.} \quad \varphi>0, \quad (32)$$

whose solution is $\varphi^* = \frac{C^2}{\mathbf{1}_{2Q+1}^T \mathbf{\Upsilon} \mathbf{\Xi}^{-1} \mathbf{\Upsilon} \mathbf{1}_{2Q+1}}$ leading to

$$\boldsymbol{\eta}^* = \frac{C \mathbf{\Xi}^{-1} \mathbf{\Upsilon} \mathbf{1}_{2Q+1}}{\mathbf{1}_{2Q+1}^T \mathbf{\Upsilon} \mathbf{\Xi}^{-1} \mathbf{\Upsilon} \mathbf{1}_{2Q+1}}. \quad (33)$$

Combining the estimation for $\boldsymbol{\theta}$ and the restoration for $h(n)$'s, the proposed channel estimation algorithm is summarized in Table I. Moreover, the following proposition is given to demonstrate the effectiveness of the proposed algorithm:

Proposition 1: The iterative channel estimation algorithm is convergent, and it achieves lower MSE than that of the maximum likelihood (ML) method.

Proof: Each iteration consists of $(2Q+2)$ steps. Denote the i th entry of $\boldsymbol{\theta}$ by θ_i , and the updated estimate of θ_i denoted by $\hat{\theta}_i^{\text{new}}$ satisfies $\mathcal{L}(\hat{\theta}_i^{\text{new}}) > \mathcal{L}(\hat{\theta}_i)$. This indicates that $\mathcal{L}(\boldsymbol{\theta})$ strictly increases after each step as well as after one round of iteration. It is known that $\mathcal{L}(\boldsymbol{\theta})$ is continuous with respect to θ_i and $\mathcal{L}(\boldsymbol{\theta}) < +\infty$. Thus, it is concluded that the iterative algorithm is convergent.

From (13), the MAP estimate of g with a given $\boldsymbol{\lambda}$ is

$$\hat{g}^{\text{MAP}} = \frac{\frac{\alpha^2 \|\boldsymbol{\lambda}\|^2 \|\tilde{\mathbf{t}}_s\|^2 + \|\mathbf{t}_r\|^2}{v_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_n^2}}{\frac{\alpha^2 \|\boldsymbol{\lambda}\|^2 \|\tilde{\mathbf{t}}_s\|^2 + \|\mathbf{t}_r\|^2}{v_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_n^2} + \frac{1}{v_g}} g + \frac{\sum_{q=-Q}^Q \frac{\alpha \lambda_q^H \tilde{\mathbf{t}}_s^H \mathbf{w}_q + \mathbf{t}_r^H \mathbf{w}_{Dr}}{v_n} + \frac{\mathbf{t}_r^H \mathbf{w}_{Dr}}{\sigma_n^2}}{\frac{\alpha^2 \|\boldsymbol{\lambda}\|^2 \|\tilde{\mathbf{t}}_s\|^2 + \|\mathbf{t}_r\|^2}{v_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_n^2} + \frac{1}{v_g}},$$

whose MSE is calculated as

$$\begin{aligned} \delta_g^{\text{MAP}} &= \mathcal{E}\left\{\|\hat{g}^{\text{MAP}} - g\|^2\right\} \\ &= \frac{1}{\frac{\alpha^2 \|\boldsymbol{\lambda}\|^2 \|\tilde{\mathbf{t}}_s\|^2 + \|\mathbf{t}_r\|^2}{v_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_n^2}} \frac{v_g}{v_g + \frac{1}{\frac{\alpha^2 \|\boldsymbol{\lambda}\|^2 \|\tilde{\mathbf{t}}_s\|^2 + \|\mathbf{t}_r\|^2}{v_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_n^2}}}. \end{aligned} \quad (34)$$

The ML estimate of g gives

$$\hat{g}^{\text{ML}} = g + \frac{\sum_{q=-Q}^Q \frac{\alpha \lambda_q^H \tilde{\mathbf{t}}_s^H \mathbf{w}_q + \mathbf{t}_r^H \mathbf{w}_{Dr}}{v_n} + \frac{\mathbf{t}_r^H \mathbf{w}_{Dr}}{\sigma_n^2}}{\frac{\alpha^2 \|\boldsymbol{\lambda}\|^2 \|\tilde{\mathbf{t}}_s\|^2 + \|\mathbf{t}_r\|^2}{v_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_n^2}},$$

whose MSE is calculated as

$$\delta_g^{\text{ML}} = \mathcal{E}\left\{\|\hat{g}^{\text{ML}} - g\|^2\right\} = \frac{1}{\frac{\alpha^2 \|\boldsymbol{\lambda}\|^2 \|\tilde{\mathbf{t}}_s\|^2 + \|\mathbf{t}_r\|^2}{v_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_n^2}}. \quad (35)$$

Clearly, $\delta_g^{\text{MAP}} < \delta_g^{\text{ML}}$ always holds, and it can be also testified that $\delta_{\lambda_q}^{\text{MAP}} < \delta_{\lambda_q}^{\text{ML}}$ is satisfied similarly. ■

Remark: According to (34) and (35), we see that $\delta_g^{\text{MAP}} < \text{CRB}_g$ and $\delta_g^{\text{ML}} = \text{CRB}_g$. This is because the proposed MAP estimation algorithm is biased since $\mathcal{E}\{\hat{g}^{\text{MAP}}\} \neq g$.

TABLE I
ITERATIVE CHANNEL ESTIMATION ALGORITHM

-
- **Initialize** \hat{g} in accordance with (16); $i_Index=i_Times(> 0)$
 - **Repeat**
 - For each q , update $\hat{\lambda}_q$ by substituting \hat{g} into (13).
 - Update \hat{g} by substituting $\hat{\lambda}_q$'s into (17).
 - $i_Index=i_Index-1$
 - **Until** $i_Index \leq 0$ is satisfied.
 - **Calculate** the optimal $\boldsymbol{\eta}^*$ according to (33).
 - **Restore** $\hat{h}(n)$'s according to (18) by using $\hat{\lambda}_q$'s and $\boldsymbol{\eta}^*$.
 - **Return** $\hat{h}(n)$'s and \hat{g} .
-

IV. NUMERICAL RESULTS

Numerical results are provided to evaluate the performance of the CE-BEM based MAP (C-MAP) channel estimation algorithm. The AL channel $\{h(n)\}$ and BL channel g are generated from the spatial channel model (CSM) in *3GPP TR 25.996* [13]. The parameters are set as $N_s = 800$, $N_r = 4$ and $i_Times=10$. We assume binary-phase-shift-keying (BPSK) modulation for $\{b(n)\}$, while $\tilde{\mathbf{t}}_s$ is selected as the 2nd column of the $N_p \times N_p$ discrete Fourier transform (DFT) matrix and \mathbf{t}_r is as the 3rd column of the $N_r \times N_r$ DFT matrix. The transmit power P_s and P_r are set to be equal, and the noise variance is set to be of unit value; thus the SNR is equal to P_s .

In Fig. 2, the average MSEs of both AL and BL channels for C-MAP estimation are compared with that for ML estimation. It is observed that the proposed C-MAP estimation outperforms the traditional ML method since C-MAP achieves lower MSEs of both AL and BL channels than ML. Moreover, it is seen that the MSE of the AL channel for the ML method is not convergent. This is because random generation of g would result in singularity, e.g., $g \rightarrow 0$, leading to $\Delta\lambda_q \rightarrow \infty$ for the ML method, while the proposed C-MAP algorithm is robust at the singularity. In Fig.3, we evaluate the AESNR performance for the optimal weighted approach (OWA) to channel restoration. It is seen that OWA obtains higher AESNRs than the baseline (restoring according to CE-BEM), especially in the low SNR region, and it draws near to the baseline as SNR increases.

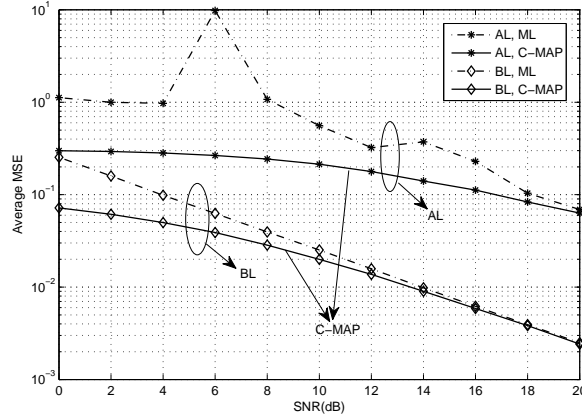


Fig. 2. Average MSE versus SNR for different estimation methods.

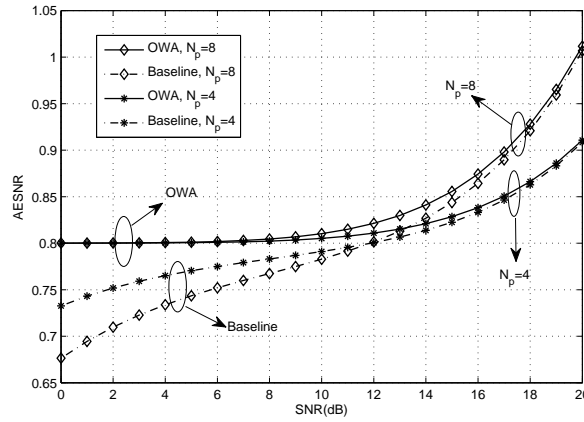


Fig. 3. AESNR versus SNR for different channel restoring methods.

V. CONCLUSIONS

A superimposed-segment training design has been proposed to decrease the training overhead and enhance the channel estimation accuracy in uplink C-RANs. Based on the training design, a CE-BEM based MAP channel estimation algorithm has been developed, where the BEM coefficients of the time-varying radio AL and the channel fading of the quasi-static wireless BL are first obtained, then the time-domain channel samples of AL are restored from maximizing AESNR. Simulation results have demonstrated that the proposed algorithm declines the estimation MSE and increases the AESNR.

REFERENCES

- [1] C. I, *et al.*, “Toward green and soft: A 5G perspective,” *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66–73, Feb. 2014.
- [2] Y. Zhou and W. Yu, “Optimized backhaul compression for uplink cloud radio access network,” *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1295–1307, June 2014.
- [3] S. Park, *et al.*, “Robust layered transmission and compression for distributed uplink reception in cloud radio access networks,” *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 204–216, Jan. 2014.
- [4] S. Zhang, F. Gao, C. Pei and X. He, “Segment training based individual channel estimation in one-way relay network with power allocation,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 3, pp. 1300–1309, Mar. 2013.
- [5] T. Biermann, *et al.*, “How backhaul networks influence the feasibility of coordinated multipoint in cellular networks,” *IEEE Commun. Mag.*, vol. 51, no. 8, pp. 168–176, Aug. 2013.
- [6] J. Tugnait and W. Luo, “On channel estimation using superimposed training and first-order statistics,” *IEEE Commun. Lett.*, vol. 7, no. 9, pp. 413–416, Sep. 2003.
- [7] G. Giannakis and C. Tepedelenlioglu, “Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels,” *Proc. IEEE*, vol. 86, pp. 1969–986, Nov. 1998.
- [8] G. Dou, C. He, C. Li and J. Gao, “A weighted first-order statistical method for time-varying channel and DC-offset estimation using superimposed training,” *IEEE Commun. Lett.*, vol. 17, no. 5, pp. 852–855, May 2013.
- [9] X. Ma and G. Giannakis, “Maximum-diversity transmissions over doubly-selective wireless channels,” *IEEE Trans. Inf. Theory*, vol. 49, pp. 1832–1840, July 2003.
- [10] G. Wang, *et al.*, “Channel estimation and training design for two-way relay networks in time-selective fading environments,” *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2681–2691, Aug. 2011.
- [11] F. Gao, R. Zhang and Y. Liang, “Optimal channel estimation and training design for two-way relay networks,” *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, Oct. 2009.
- [12] E. Carvalho, J. Cioffi and D. Slock, “Cramér-rao bounds for blind multichannel estimation,” in *Proc. IEEE GLOBECOM*, San Francisco, pp. 1036–1040, Nov. 2000.
- [13] “Spatial channel model for multiple input multiple output simulations,” *3GPP TR 25.996*, release 9, Dec. 2009.
- [14] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Prentice Hall PTR, 1993.
- [15] A. Goldsmith, *Wireless Communications*. Cambridge, England: Cambridge University Press, 2005.