Training Design and Channel Estimation in Uplink Cloud Radio Access Networks

Xinqian Xie, Mugen Peng, Wenbo Wang, and H. Vincent Poor, *Fellow, IEEE*

***Abstract*—To decrease the training overhead and improve the channel estimation accuracy in uplink cloud radio access networks (C-RANs), a superimposed-segment training design is proposed. The core idea of the proposal is that each mobile station superimposes a periodic training sequence on the data signal, and each remote radio head prepends a separate pilot to the received signal before forwarding it to the centralized base band unit pool. Moreover, a complex-exponential basis-expansion-model based channel estimation algorithm to maximize a posteriori probability is developed. Simulation results show that the proposed channel estimation algorithm can effectively decrease the estimation mean square error and increase the average effective signal-to-noise ratio (AESNR) in C-RANs.**

***Index Terms*—Channel estimation, cloud radio access networks.**

I. Introduction

IN CLOUD RADIO ACCESS NETWORKS (C-RANs), a large number of remote radio heads (RRHs) are deployed to forward received signals from mobile stations (MSs) to a centralized base band unit (BBU) pool through wired/wireless backhaul links for uplink transmission [1]. To suppress the inter-RRH interference by using cooperative processing techniques at the BBU pool, channel state information (CSI) of both the radio access links (ALs) and wireless backhaul links (BLs) is required [2]. The superimposed-training scheme can significantly reduce the overhead and is valid to perform channel estimation for time-varying environments using the complex-exponential basis-expansion-model (CE-BEM) [3]. 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 the effective signal-to-noise ratio (SNR) [4].

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 AL while segment-training is applied for the wireless BL. Moreover, a CE-BEM based maximum a posteriori probability (MAP) channel estimation algorithm is developed, in which the basis-expansion-model (BEM) coefficients of the time-varying AL and the channel fading of the quasi-static wireless BL are first obtained, after which the time-domain channel samples of the AL are restored in terms of maximizing the average effective SNR.

Manuscript received September 16, 2014; revised November 04, 2014; accepted December 07, 2014. Date of publication December 18, 2014; date of current version January 07, 2015. The work of X. Xie, M. Peng, and W. Wang was supported in part by the National Natural Science Foundation of China under Grant 61222103, the State Major Science and Technology Special Projects under Grant 2013ZX03001001, and by the Beijing Natural Science Foundation under Grant 4131003. The work of H. V. Poor was supported in part by the National Science Foundation under Grant ECCS-1343210. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ron Dabora.

X. Xie, M. Peng, and W. Wang are with Key Laboratory of Universal Wireless Communications (Ministry of Education), Beijing University of Posts & Telecommunications, Beijing, China (e-mail: xxmbupt@gmail.com; pmg@bupt.edu.cn; wbwang@bupt.edu.cn).

H. V. Poor is with the School of Engineering and Applied Science, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2014.2380776

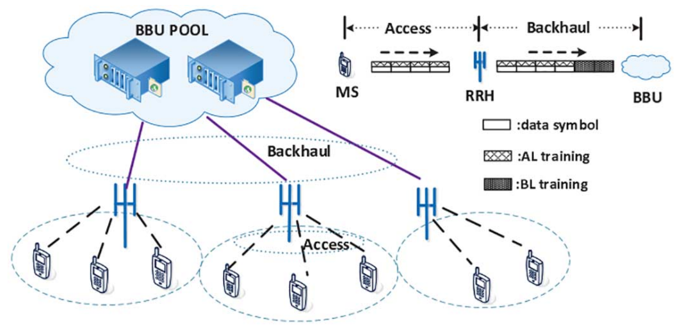

Fig. 1. System model and the superimposed-segment training design.

II. System Model and Training Design

A C-RAN consisting of one BBU pool and multiple RRHs is depicted in Fig. 1, where the RRHs operate in half-duplex mode, and different MSs associated with 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 radio channels of 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 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 fading gain of the time-varying radio AL is denoted by $h(n)$ with mean zero and variance v_h, while the channel fading gain of the quasi-static flat BL is denoted by g with the complex Gaussian distribution of mean zero and variance v_g. The transmit power of the MSs and RRHs are denoted by P_s and P_r, respectively. The noise variances at the RRHs and the BBU pool are denoted by σ_R^2 and σ_B^2, respectively. 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, σ_R^2 and σ_B^2.

During each transmission block, the MS transmits a signal $\mathbf{s}$ with symbol length of N_s 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 value of ϵ is within $0 < \epsilon < 1$. Without loss of generality, we further assume that $K = \frac{N_s}{N_p}$ is an integer [3]. The n-th observation at the RRH can be written as

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

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

$$\mathbf{x}_s = \alpha g \cdot \mathrm{diag}(\mathbf{h}) \cdot \mathbf{s} + \alpha g \mathbf{w}_R + \mathbf{w}_{Ds}, \qquad (3)$$

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

where $\mathbf{w}_R = [w_R(1) \ldots w_R(N_s)]^T$, $\mathbf{w}_{Ds} = [w_{Ds}(1) \ldots w_{Ds}(N_s)]^T$ and $\mathbf{w}_{Dr} = [w_{Dr}(1), \ldots, w_{Dr}(N_r)]^T$ are AWGN vectors. 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.

III. CE-BEM Based MAP Channel Estimation Algorithm

The CE-BEM for time-selective but frequency-flat fading channels, e.g., Rayleigh fading, given in [5] and [6] is chosen to model the time-varying AL as

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

where the λ_q's are assumed to satisfy independent complex Gaussian distributions with λ_q having mean zero and variance v_q. Moreover, the λ_q's are assumed to remain invariant within one transmission block, and $\sum_{q=-Q}^{Q} v_q = v_h$ is satisfied. Substituting (5) into (3), $\mathbf{x}_r$ can be rewritten as

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

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

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

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

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

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

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

We have

$$\mathcal{E}\{\mathbf{w}_{q_1}\mathbf{w}_{q_1}^H\} = v_n \mathbf{I}_{N_p}, \quad \mathcal{E}\{\mathbf{w}_{q_1}\mathbf{w}_{q_2}^H\} = \mathbf{0}_{N_p} \qquad (10)$$

for $q_1 \neq q_2$ with $v_n = \frac{\alpha^2 v_g v_h (1-\epsilon) P_s + (\alpha^2 v_g \sigma_R^2 + \sigma_B^2)}{K}$. In [14, pp. 7], it is mentioned that if the probability density function (PDF) of a noise process is not mathematically tractable, a reasonable model to choose is white Gaussian noise. Since the PDF of $\mathbf{w}_q$ is rather complicated, we thus choose the complex Gaussian density to be the nominal likelihood function of $\mathbf{w}_q$. In this case, the likelihood function of $\mathbf{r}$ can be written as

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

A. Estimation for λ_q's and g

Defining $\boldsymbol{\theta} = [\lambda_{-Q}, \ldots, \lambda_Q, g]^T$, the MAP estimation of $\boldsymbol{\theta}$ gives (12) shown at the bottom of this page, where $p(\mathbf{x}_r|g)$, $p(\boldsymbol{\lambda})$ and $p(g)$ are Gaussian density functions. With a given g, the estimate of λ_q can be obtained as

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

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

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

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

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

The estimate of $|g|$ must be either a local minimum 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 in $|g|$, and thus a numerical method such as one dimensional search is needed to compute the value of $\widehat{|g|}$. To reduce the complexity

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

of such an approach, an iterative approach is developed whereby $\hat{g}$ is initialized from

$$\hat{g} = \arg\max_{g} \{p(\mathbf{x}_r|g)\, p(g)\} = \frac{\mathbf{t}_r^H \mathbf{x}_r}{\|\mathbf{t}_r\|^2 + \frac{\sigma_B^2}{\upsilon_g}}. \tag{16}$$

With the obtained $\hat{g}$, the $\hat{\lambda}_q$'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

$$\hat{g} = \frac{\sum_{q=-Q}^{Q} \frac{\alpha \hat{\lambda}_q^H \tilde{\mathbf{t}}_s^H \mathbf{r}_q}{\upsilon_n} + \frac{\mathbf{t}_r^H \mathbf{x}_r}{\sigma_B^2}}{\sum_{q=-Q}^{Q} \frac{\alpha^2 |\hat{\lambda}_q|^2 \|\tilde{\mathbf{t}}_s\|^2}{\upsilon_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_B^2} + \frac{1}{\upsilon_g}}. \tag{17}$$

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

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

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

where $\boldsymbol{\eta}_q$'s are real factors. The vector $\boldsymbol{\eta}$ that maximizes the average effective SNR (AESNR) [7] denoted by $\boldsymbol{\eta}^*$ is obtained from

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

Define $\phi_{q,n} = e^{j\frac{2\pi q n}{N_s}}$, and denote by Δg and $\Delta\lambda_q$ the estimation error of g and λ_q, respectively. We can obtain (20), shown at the bottom of the page. Note that $\delta_{\lambda_q}, \delta_g \sim \mathcal{O}(\frac{\upsilon_n}{P_s})$ while the last term in (20) has the order of $\mathcal{O}[(\frac{\upsilon_n}{P_s})^2]$, and thus we remove the last term for the high SNR approximation, i.e.,

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

Similarly,

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

Due to the non-linearity of the MAP estimators (13) and (17), it is difficult to derive the expressions for the corresponding mean-square errors (MSEs) δ_{λ_q} and δ_g, respectively, in closed form. Thus, we approximate δ_{λ_q} and δ_g as [12, eq. 27, pp. 454]:

$$\delta_{\lambda_q} = \frac{\upsilon_n}{\alpha^2 |g|^2 \|\tilde{\mathbf{t}}_s\|^2}, \delta_g = \frac{1}{\frac{\alpha^2\|\lambda\|^2\|\tilde{\mathbf{t}}_s\|^2}{\upsilon_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_B^2}}, \tag{23}$$

respectively. It is stressed that both δ_{λ_q} and δ_g in (23) are the corresponding MSEs of the asymptotic MAP estimators when N_s and N_r are sufficiently large.

Substituting (23) into $\bar{\gamma}$ and taking the expectation with respect to g, the λ_q's and the noise terms, we can obtain

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

where (see (25)–(27) at the bottom of the page), and $\mathbf{R}_\lambda = \mathrm{diag}(\upsilon_{-Q} \ldots \upsilon_Q)$. On $\varphi^* = (\boldsymbol{\eta}^*)^T \boldsymbol{\Xi} \boldsymbol{\eta}^*$, the optimization of (19) transforms to

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

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

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

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

Substituting (30) back into (24), the optimization problem is transformed as

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

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

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

$$\boldsymbol{\Upsilon} = \frac{\alpha^2 \upsilon_g \|\tilde{\mathbf{t}}_s\|^2}{\upsilon_n} \mathbf{R}_\lambda^2 + \frac{\alpha^2 \upsilon_g \|\tilde{\mathbf{t}}_s\|^2}{\upsilon_n} \mathrm{tr}\{\mathbf{R}_\lambda\} \mathbf{R}_\lambda + \frac{\upsilon_g \|\mathbf{t}_r\|^2}{\sigma_B^2} \mathbf{R}_\lambda, \tag{25}$$

$$\Xi = \boldsymbol{\Upsilon} + \mathbf{R}_\lambda + \left(\upsilon_h + \frac{\upsilon_n \|\mathbf{t}_r\|^2}{\sigma_B^2 \alpha^2 \|\tilde{\mathbf{t}}_s\|^2}\right) \mathbf{I}_{2Q+1}, \tag{26}$$

$$C = \mathbf{1}_{2Q+1}^T \boldsymbol{\Upsilon} \mathbf{1}_{2Q+1} + \left(\upsilon_g \sigma_R^2 + \frac{\sigma_B^2}{\alpha^2}\right) \left[\frac{\upsilon_h \alpha^2 \|\tilde{\mathbf{t}}_s\|^2}{\upsilon_n P_s} + \frac{\|\mathbf{t}_r\|^2}{\sigma_B^2 P_s}\right], \tag{27}$$

TABLE I
ITERATIVE CHANNEL ESTIMATION ALGORITHM

- **Initialize** $\hat{g}$ from (16), and obtain $\hat{\lambda}_q$'s by substituting $\hat{g}$ into (13).
- **Repeat**
  - $g_{\text{temp}} = \hat{g}$; $\lambda_{q,\text{temp}} = \hat{\lambda}_q$.
  - Update $\hat{g}$ by substituting $\lambda_{q,\text{temp}}$'s into (17).
  - For each q, update $\hat{\lambda}_q$ by substituting g_{temp} into (13).
- **Until** $|\hat{g} - g_{\text{temp}}|^2 + \sum_q |\hat{\lambda} - \lambda_{q,\text{temp}}|^2 \leq 0.001$ 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 $\hat{g}$.

whose solution $\varphi^* = \frac{C^2}{\mathbf{1}_{2Q+1}^T \boldsymbol{\Upsilon} \boldsymbol{\Xi}^{-1} \boldsymbol{\Upsilon} \mathbf{1}_{2Q+1}}$, which leads to

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

The proposed channel estimation algorithm is summarized in Table I, and 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 a lower MSE than the maximum likelihood (ML) method does.

Proof: Each iteration consists of $(2Q+2)$ steps. Denote the i-th entry of $\boldsymbol{\theta}$ by θ_i, and note that 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. Thus, it is concluded that the iterative algorithm is convergent.

By substituting (4) and (9) into (17), the MAP estimate of g with a given $\boldsymbol{\lambda}$ is shown at the bottom of the page, whose MSE is calculated as

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

Similarly, the MSE of ML estimation for g is calculated as

$$\delta_g^{\text{ML}} = \mathcal{E}\{\|\hat{g}^{\text{ML}} - g\|^2\} = \frac{1}{\frac{\alpha^2\|\lambda\|^2\|\tilde{\mathbf{t}}_s\|^2}{v_n} + \frac{\|\mathbf{t}_r\|^2}{\sigma_B^2}}. \tag{35}$$

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

IV. NUMERICAL RESULTS

Numerical results are provided to evaluate the performance of the proposal. The AL channel $\{h(n)\}$ and BL channel g are generated from the spatial channel model (SCM) in *3GPP TR 25.996* [9]. The parameters are set as $N_s = 800$ and $N_r = 4$. We assume binary-phase-shift-keying (BPSK) modulation for $\{b(n)\}$, while $\tilde{\mathbf{t}}_s$ is selected as the 2nd column of the selected $N_p \times N_p$ discrete Fourier transform (DFT) matrix and $\mathbf{t}_r$ is selected as the 3rd column of the $N_r \times N_r$ DFT matrix. The transmit powers P_s and P_r are set to be equal, and the noise variances σ_R^2 and σ_B^2 are set to be of unit value. Thus the SNR is equal to P_s.

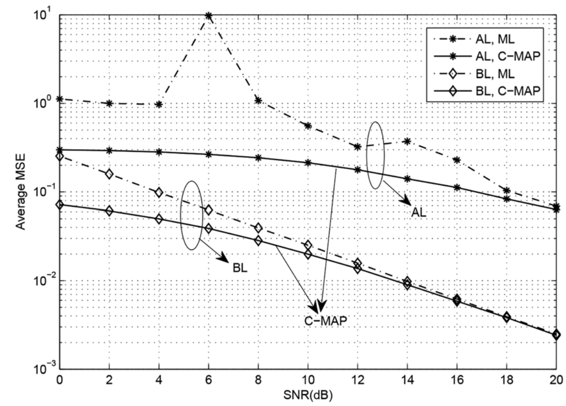

Fig. 2. Average MSE versus SNR for different estimation methods.

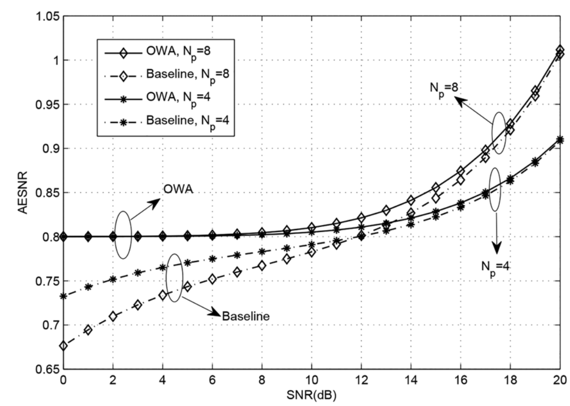

Fig. 3. AESNR versus SNR for different channel restoring methods.

In Fig. 2, the average MSEs 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 over 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 can result in singularity, e.g., $g \to 0$, leading to $\Delta\lambda_q \to \infty$ 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) does, especially in the low SNR region.

V. CONCLUSIONS

A superimposed-segment training design has been proposed to decrease training overhead and enhance channel estimation accuracy in uplink C-RANs. Simulation results have demonstrated that the proposed algorithm lowers the estimation MSE and increases the AESNR.

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

References

[1] 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, Jun. 2014.

[2] 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.

[3] G. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time-varying channels," in *Proc. IEEE*, Nov. 1998, vol. 86, pp. 1969–986.

[4] 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.

[5] X. Ma and G. Giannakis, "Maximum-diversity transmissions over doubly-selective wireless channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1832–1840, Jul. 2003.

[6] 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.

[7] 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.

[8] X. Xie *et al.*, "Maximum a posteriori based channel estimation strategy for two-way relaying channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 450–463, Jan. 2014.

[9] "Spatial channel model for multiple input multiple output simulations," 3GPP TR 25.996, release 9 Dec. 2009.

[10] S. M. Kay, "Fundamentals of statistical signal processing," in *Estimation Theory*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1993, vol. I.

[11] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.