Precoder Design for Distributed Antenna Systems (DAS) with Limited Channel State Information

Hieu Duy Nguyen, Jingon Joung, and Sumei Sun
Institute for Inforcomm Research (I^2R)
Agency for Science, Technology, and Research (A*STAR), Singapore
Email(s): {nguyendh, jgjoung, sunsm}@i2r.a-star.edu.sg

***Abstract*—A distributed antenna system (DAS) consists of multiple baseband units (BBUs) connecting to distributed antennas (DAs) via dedicated access links. In this study, we investigate a DAS with limited channel state information (CSI) and consider an average rate of users as an objective, where the expectation is taken over the channel uncertainty. We propose two distributed precoder designs that are based on a rate lower bound and a rate upper bound, respectively. As a benchmark, coordinated precoder and cooperative dirty paper coding (DPC)-based precoder with full CSI are compared with our proposed algorithms. Numerical results verifies that the rate performance of our upper-bound based scheme with limited CSI approaches tightly the maximum rates of the full CSI schemes, while that of lower-bound based scheme is relatively worse.**

I. Introduction

Various centralized/distributed precoders have been extensively studied for centralized antenna systems (CASs) whose all antennas are co-located at a base station. For example, precoder design for single-cell multiple-input multiple-output (MIMO) system was considered in [1], [2], in which the transmitter receives only an estimated channel state information (CSI) with errors, using which the worst case rate/error performance is optimized. Distributed designs have also been proposed to support multiuser in the cooperative systems, where only local CSI, i.e., knowledge of channel vectors between transmitters and its corresponding receivers, is available at the transmitters [3], [4].

Successful commercialization of the fourth generation (4G) cellular wireless network has encouraged the development of the next generation communication systems. Demands for high data rate and low energy consumption in the future communication systems has motivated the dense cell and/or antenna deployment, namely, small-cell networks and distributed antenna systems (DASs), due to its potential of high capacity and energy efficiency. Different from the co-located antennas in CASs, the antennas in DASs are distributed over the service region and connected to a baseband unit (BBU) via, e.g., optical fibre. Due to short distance between the distributed antennas (DAs) and users, DASs increase the desired signal strength and system capacity while reducing the transmit power and, consequently, the co-channel interferences.

The potential of DAS has attracted much research attention in the literature [5]–[11]. Investigations have been dedicated to information-theoretic rate analysis [5], [6], and precoder/receiver design [7]. In [8] and [9], energy efficiency of DAS is illuminated with antenna selection and user scheduling. The works [5]–[9] are based on the assumption of *perfect/full* CSI at the transmitter (CSIT). Due to the distribution of DAs, a prohibitively large fronthaul data exchange is required to obtain the perfect CSIT. A few recent studies have considered imperfect and/or limited CSIT for DAS with central processing [10], [11]. However, most of the existing work regarding DASs consider only a single BBU without interferences, which is an impractical assumption as dense antenna deployment with multiple BBUs will incur strong interference signals.

In this paper, we investigate DASs with *multiple* BBUs serving *multiple* users, in which distributed precoder is designed with *limited* CSIT. To our best knowledge, this is the first work considering multiple BBUs with limited CSIT in DASs. We adopt an average rate of users as our objective function, where the expectation is over the channel uncertainty, and propose two distributed schemes based on lower and upper bounds of the achievable rate. The proposed schemes can be implemented at each BBU in a distributed manner. For comparison, we also evaluate the performance of DASs using a state-of-the-art coordinated precoder and a cooperative dirty paper coding (DPC)-based precoder with full CSIT. Through simulations, we show that comparable rate performance is achieved by our proposed distributed schemes, but with significantly lower overhead gained from the distributed BBU processing.

Notations: Scalars and vectors/matrices are denoted by lower-case and bold-face lower-case/upper-case letters, respectively. The conjugate, transpose, and conjugate transpose operators are denoted as $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$, respectively. $[\boldsymbol{A}]_{i,j}$ stands for the (i,j)th element of the matrix $\boldsymbol{A}$. $\mathbb{E}_X[\cdot]$ denotes the statistical expectation over a random variable X. $Tr(\cdot)$ and $\det(\cdot)$ represent the trace and determinant of a matrix, respectively. $\boldsymbol{A} \succeq \boldsymbol{0}$ means that $\boldsymbol{A}$ is a positive semi-definite matrix. $\mathbb{R}^{x\times y}$ and $\mathbb{C}^{x\times y}$ denote the space of x-by-y real and complex matrices, respectively.

II. System Model

Consider a DAS with N active users as shown in Fig. 1. User n is assigned to B_n BBUs which, for example, have the DA with the highest signal-to-interference plus noise (SINR) to user n (we do not consider the user and DA association in this paper). In the data transmission phase, the B_n BBUs receive the data from their corresponding control unit (CU) and coordinately transmit the signal to user n via the selected DAs. For ease of analysis, we assume that each BBU transmits to one user in each resource slot; the number of DAs selected by each BBU is identical and equal to N_{DA}; and the number of antennas at each user is N_R. The proposed precoder designs can be readily extended to general cases with multiple users served by each BBU with different numbers of DAs. The received baseband signal at user n is expressed as

$$\boldsymbol{y}^{(n)} = \sum_{b=1}^{B_n} \boldsymbol{H}_b^{(n,n)} \left(\boldsymbol{\Gamma}_b^{(n,n)}\right)^{1/2} \boldsymbol{T}_b^{(n)} \boldsymbol{s}^{(n)} + \sum_{l=1,\ l\neq n}^{N} \sum_{b=1}^{B_l} \boldsymbol{H}_b^{(l,n)} \left(\boldsymbol{\Gamma}_b^{(l,n)}\right)^{1/2} \boldsymbol{T}_b^{(l)} \boldsymbol{s}^{(l)} + \boldsymbol{n}^{(n)}, \quad (1)$$

where $\boldsymbol{H}_b^{(l,n)} \in \mathbb{C}^{N_R \times N_{DA}}$ denotes the channel from between user n and BBU b who serves user l; each element of $\boldsymbol{H}_b^{(l,n)}$ is assumed to be independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian with zero mean and unit variance, denoted as $\mathcal{CN}(0,1)$; $\boldsymbol{\Gamma}_b^{(l,n)} = \operatorname{diag}\left(\gamma_{b,1}^{(l,n)}, \ldots, \gamma_{b,N_{DA}}^{(l,n)}\right)$ is a real diagonal matrix that represents the signal attenuation from BBU b serving user l to user n; $\boldsymbol{s}^{(n)} \in \mathbb{C}^{N_R \times 1}$ is the transmitted data signal from the BBUs serving user n, normalized to be $\mathbb{E}_{\boldsymbol{s}}\left[\boldsymbol{s}^{(n)}\left(\boldsymbol{s}^{(n)}\right)^H\right] = \boldsymbol{I}$; $\boldsymbol{T}_b^{(n)} \in \mathbb{C}^{N_{DA} \times N_R}$ is a precoding matrix of BBU b serving user n satisfying the power constraint:

$$Tr\left(\boldsymbol{T}_b^{(n)} \left(\boldsymbol{T}_b^{(n)}\right)^H\right) \leq P_b^{(n)};$$

and $\boldsymbol{n}^{(n)} \in \mathbb{C}^{N_R \times 1}$ is the additive white Gaussian noise (AWGN) at user n with i.i.d. elements, each of which is distributed as $\mathcal{CN}(0, \sigma^2)$. Here, the first and second terms of the right hand side in (1) represent the desired and interfering signals, respectively.

Denote covariance matrices $\boldsymbol{A}^{(n)}$ and $\boldsymbol{W}^{(n)}$ of the desired and interference-plus-noise signals, respectively, as

$$\boldsymbol{A}^{(n)} = \left(\sum_{b=1}^{B_n} \boldsymbol{G}_b^{(n,n)} \boldsymbol{T}_b^{(n)}\right) \left(\sum_{b=1}^{B_n} \boldsymbol{G}_b^{(n,n)} \boldsymbol{T}_b^{(n)}\right)^H, \quad (2)$$

$$\boldsymbol{W}^{(n)} = \sigma^2 \boldsymbol{I} + \sum_{l\neq n}^{N} \left(\sum_{b=1}^{B_l} \boldsymbol{G}_b^{(l,n)} \boldsymbol{T}_b^{(l)}\right) \left(\sum_{b=1}^{B_l} \boldsymbol{G}_b^{(l,n)} \boldsymbol{T}_b^{(l)}\right)^H, \quad (3)$$

where $\boldsymbol{G}_b^{(l,n)} \triangleq \boldsymbol{H}_b^{(l,n)} \left(\boldsymbol{\Gamma}_b^{(l,n)}\right)^{1/2}$, $b \in \{1, \ldots, B_n\}$ and $l, n \in \{1, \ldots, N\}$. The instantaneous spectral efficiency of the n-th user $R_{ins}^{(n)}$ is then given by

$$R_{ins}^{(n)} = \log_2 \det\left(\boldsymbol{A}^{(n)} + \boldsymbol{W}^{(n)}\right) - \log_2 \det \boldsymbol{W}^{(n)}. \quad (4)$$

In practical communications systems, the training for CSI estimation between a BBU and an user, who is served by another BBU, is difficult to be achieved. In this study, we therefore assume that only the direct link CSI, i.e., $\boldsymbol{G}_k^{(n,n)}$, is available at BBU k which serves user n [3] [4]. Since each BBU knows only direct link channel $\boldsymbol{G}_k^{(n,n)}$, the problem needs to be solved in a distributed manner at each BBU, i.e., *distributed* scheme. Defining user n's desired signal matrix from its corresponding BBUs except BBU k as

$$\boldsymbol{Z}_{-k}^{(n)} \triangleq \sum_{b=1, b\neq k}^{B_n} \boldsymbol{G}_b^{(n,n)} \boldsymbol{T}_b^{(n)}, \quad (5)$$

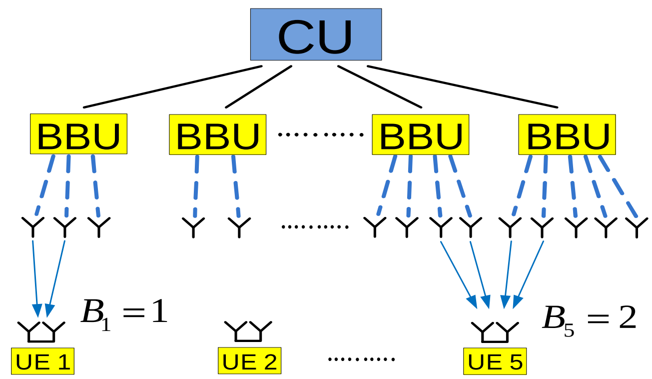

Fig. 1: A DAS with $N = 5$ users each having $N_R = 2$. Here user 1 and 5 are served by $B_1 = 1$ and $B_5 = 2$ BBUs, respectively.

the *average* achievable rate of user n is derived as

$$R_{avg}^{(n)} = \mathbb{E}_{\boldsymbol{W},\boldsymbol{Z}}\left[\log_2 \det\left(\left(\boldsymbol{G}_k^{(n,n)} \boldsymbol{T}_k^{(n)} + \boldsymbol{Z}_{-k}^{(n)}\right)\left(\boldsymbol{G}_k^{(n,n)} \boldsymbol{T}_k^{(n)} + \boldsymbol{Z}_{-k}^{(n)}\right)^H + \boldsymbol{W}^{(n)}\right) \middle| \boldsymbol{G}_k^{(n,n)}\right] - \mathbb{E}_{\boldsymbol{W}}\left[\log_2 \det \boldsymbol{W}^{(n)}\right], \quad (6)$$

where the expectation is over the uncertainty, i.e., $\boldsymbol{W}^{(n)}$ and $\boldsymbol{Z}_{-k}^{(n)}$ given in (3) and (5), respectively. Under the limited CSI constraint, we aim to maximize the average achievable rate (6) for each user by designing the precoder $\boldsymbol{T}_k^{(n)}$, $n = 1, \ldots, N$, $k = 1, \ldots, B_n$. The optimization problem is formulated for BBU k serving user n as follows:

$$\underset{\boldsymbol{T}_k^{(n)}}{\texttt{maximize}} \quad R_{avg}^{(n)}$$
$$\texttt{s.t.} \quad Tr\left(\boldsymbol{T}_k^{(n)} \left(\boldsymbol{T}_k^{(n)}\right)^H\right) \leq P_k^{(n)}, \quad (7)$$

The constraint in (7) is the power limitation of BBU k serving user n by maximum power $P_k^{(n)}$. Finally, since the precoder $\boldsymbol{T}_k^{(n)}$ is designed based on the direct link $\boldsymbol{G}_k^{(n,n)}$ only, it is reasonable and intuitive to assume that $\boldsymbol{T}_k^{(n)}$ is independent of other links, namely, $\boldsymbol{G}_b^{(l,m)}$, $b = 1, \ldots, B_l$, and $l, m \in \{1, \ldots, N\}$, in which $(b, l, m) \neq (k, n, n)$.

III. Distributed Precoder Design for DAS

To solve (7), we consider two suboptimal approaches that design two distributed precoders based on lower and upper bounds of the achievable rate $R_{avg}^{(n)}$ in (6), respectively. To this end, we first obtain Proposition 1.

Proposition 1: Assuming that BBU k who serves user n knows *only* direct link $\boldsymbol{G}_k^{(n,n)}$, the optimal precoder structure for (7) is as follows:

$$\widehat{\boldsymbol{T}}_k^{(n)} = \boldsymbol{V}_k^{(n,n)} \boldsymbol{\Lambda}_k^{(n)} \left(\boldsymbol{U}_k^{(n,n)}\right)^H, \quad (8)$$

where $\boldsymbol{\Lambda}_k^{(n)}$ is a real-valued diagonal matrix[1] whose ith element is α_i such that $Tr\left(\left(\boldsymbol{\Lambda}_k^{(n)}\right)^2\right) \leq P_k^{(n)}$, $k = 1, \ldots, B_n$, and $n = 1, \ldots, N$; $\boldsymbol{V}_k^{(n,n)}$ and $\boldsymbol{U}_k^{(n,n)}$ are the right and left singular matrices of $\boldsymbol{G}_k^{(n,n)}$, which is obtained from the

[1] An $N \times M$ matrix $\boldsymbol{X}$ is called diagonal if $[\boldsymbol{X}]_{i,j} = 0$, $\forall i \neq j$, i.e, the only possibly non-zero elements are located at the (i,i)th position where $i = 1, \ldots, \min(N, M)$.

singular value decomposition (SVD) such that $\boldsymbol{G}_k^{(n,n)} = \boldsymbol{U}_k^{(n,n)}\boldsymbol{\Sigma}_k^{(n,n)}\left(\boldsymbol{V}_k^{(n,n)}\right)^H$; and $\boldsymbol{\Sigma}_k^{(n,n)}$ is a diagonal matrix whose ith element s_i is a singular value of $\boldsymbol{G}_k^{(n,n)}$.

Proof: The proof is based on Jensen's inequality and the invariance of $R_{avg}^{(n)}$ under unitary matrix multiplication. Detailed derivation is omitted due the page limit. ∎

Using the optimal structure in (8), the precoding design problem (7) is reformulated to the following optimization

$$\begin{aligned} &\underset{\boldsymbol{\Lambda}_k^{(n)}}{\texttt{maximize}} \ \ \widehat{R}_{avg}^{(n)} \\ &\texttt{s.t.} \ \ Tr\Big(\left(\boldsymbol{\Lambda}_k^{(n)}\right)^2\Big) \le P_k^{(n)}, \end{aligned} \tag{9}$$

where

$$\begin{aligned} \widehat{R}_{avg}^{(n)} = \mathbb{E}_{\widehat{\boldsymbol{W}},\widehat{\boldsymbol{Z}}}\Bigg[\log_2\det\bigg(\Big(\boldsymbol{\Sigma}_k^{(n,n)}\boldsymbol{\Lambda}_k^{(n)} + \widehat{\boldsymbol{Z}}_{-k}^{(n)}\Big)\Big(\boldsymbol{\Sigma}_k^{(n,n)}\boldsymbol{\Lambda}_k^{(n)} \\ + \widehat{\boldsymbol{Z}}_{-k}^{(n)}\Big)^H + \widehat{\boldsymbol{W}}^{(n)}\bigg)\bigg|\boldsymbol{\Sigma}_k^{(n,n)}\Bigg] - \mathbb{E}_{\hat{W}}\left[\log_2\det\widehat{\boldsymbol{W}}^{(n)}\right]; \end{aligned} \tag{10}$$

and $\boldsymbol{\Lambda}_k^{(n)}$ is a real diagonal matrix; $\widehat{\boldsymbol{Z}}_{-k}^{(n)}$ and $\widehat{\boldsymbol{W}}^{(n)}$ that are derived from (3), (5), and (8) as

$$\widehat{\boldsymbol{Z}}_{-k}^{(n)} = \sum_{b\neq k}^{B_n}\boldsymbol{U}_b^{(n,n)}\boldsymbol{\Sigma}_b^{(n,n)}\boldsymbol{\Lambda}_b^{(n)}\left(\boldsymbol{U}_b^{(n,n)}\right)^H, \tag{11}$$

$$\begin{aligned} \widehat{\boldsymbol{W}}^{(n)} = \sum_{l\neq n}^{N}\left(\sum_{b=1}^{B_l}\boldsymbol{G}_b^{(l,n)}\boldsymbol{V}_b^{(l,l)}\boldsymbol{\Lambda}_b^{(l)}\left(\boldsymbol{U}_b^{(l,l)}\right)^H\right) \\ \times\left(\sum_{b=1}^{B_l}\boldsymbol{G}_b^{(l,n)}\boldsymbol{V}_b^{(l,l)}\boldsymbol{\Lambda}_b^{(l)}\left(\boldsymbol{U}_b^{(l,l)}\right)^H\right)^H + \sigma^2\boldsymbol{I}. \end{aligned} \tag{12}$$

A. Lower-Bound Based Design

It is still difficult to solve (9) due to the expectation over $\widehat{\boldsymbol{Z}}_{-k}^{(n)}$ and $\widehat{\boldsymbol{W}}^{(n)}$. As the first suboptimal approach, we first consider maximizing an approximated *lower bound* (ALB) of (10) instead. The ALB $R_{ALB}^{(n)}$ of (10), i.e., $R_{ALB}^{(n)} \le \widehat{R}_{avg}^{(n)}$, is derived under the assumption that the power of interference links is as large as that of the direct link as[2]

$$\begin{aligned} R_{ALB}^{(n)} = \sum_{m=1}^{N_R}\log_2\left(s_m^2\alpha_m^2 + \frac{2(B_n-1)}{N_R\sqrt{N_R}}s_m\alpha_m\sum_{i=1}^{N_R}s_i\sqrt{\sum_{i=1}^{N_R}\alpha_i^2}\right. \\ \left. + \left[\frac{B_n-1}{N_R^2} + \frac{\sum_{l\neq n}^{N}B_l}{N_{DA}N_R}\right]\sum_{i=1}^{N_R}s_i^2\sum_{i=1}^{N_R}\alpha_i^2 + \sigma^2\right) \\ - N_R\log_2\left(\frac{\sum_{l\neq n}^{N}B_l}{N_{DA}N_R}\sum_{i=1}^{N_R}s_i^2\sum_{i=1}^{N_R}\alpha_i^2 + \sigma^2\right). \end{aligned} \tag{13}$$

[2]Certainly, this assumption is very strict and conservative, yet we can obtain a performance lower bound, which is be always guaranteed.

The detailed derivation for $R_{ALB}^{(n)}$ is omitted due to the page limit. The optimization problem is rewritten as

$$\begin{aligned} &\underset{\{\alpha_m\}_{m=1}^{N_R}}{\texttt{maximize}} \ \ R_{ALB}^{(n)} \\ &\texttt{s.t.} \ \ \sum_{i=1}^{N_R}\alpha_i^2 \le P_k^{(n)}, \end{aligned} \tag{14}$$

Next, we define

$$\boldsymbol{q} = \left[\alpha_1,\ldots,\alpha_{N_R},\sqrt{\sum_{i=1}^{N_R}\alpha_i^2}\right]^T, \tag{15}$$

$$\begin{aligned} \boldsymbol{C}_m = \boldsymbol{D}_m^T \\ \times\begin{bmatrix} s_m^2 & \frac{B_n-1}{N_R\sqrt{N_R}}s_m\sum_{i=1}^{N_R}s_i \\ \frac{B_n-1}{N_R\sqrt{N_R}}s_m\sum_{i=1}^{N_R}s_i & \left[\frac{B_n-1}{N_R^2} + \frac{\sum_{l\neq n}^{N}B_l}{N_{DA}N_R}\right]\sum_{i=1}^{N_R}s_i^2 \end{bmatrix}\boldsymbol{D}_m, \end{aligned} \tag{16}$$

where $\boldsymbol{D}_m \in \mathbb{R}^{2\times(N_R+1)}$ with the (only non-zero) elements $[\boldsymbol{D}_m]_{1,m} = [\boldsymbol{D}_m]_{2,N_R+1} = 1$, $m = 1,\ldots,N_R$, and $\boldsymbol{C}_{N_R+1} \in \mathbb{R}^{(N_R+1)\times(N_R+1)}$ with the only non-zero $[\boldsymbol{C}_{N_R+1}]_{N_R+1,N_R+1} = \frac{\sum_{l\neq n}^{N}B_l}{N_{DA}N_R}\left(\sum_{i=1}^{N_R}s_i^2\right)$. Problem (14) is then reformulated as

$$\begin{aligned} \underset{\boldsymbol{q}}{\texttt{maximize}} \ \ &\sum_{m=1}^{N_R}\log_2\left(\boldsymbol{q}^T\boldsymbol{C}_m\boldsymbol{q} + \sigma^2\right) \\ &\quad - N_R\log_2\left(\boldsymbol{q}^T\boldsymbol{C}_{N_R+1}\boldsymbol{q} + \sigma^2\right) \\ \texttt{s.t.} \ \ &||\boldsymbol{q}||_2 = \sqrt{2}q_{N_R+1}, \\ &q_{N_R+1} \le \sqrt{P_k^{(n)}}. \end{aligned} \tag{17}$$

To solve (17), we first note the following fact.

Fact 1: For every $x > 0$, we have : $-\log(x) = \max_{y>0}\log(y) - xy + 1$. The optimal value of y to achieve that is $y^* = 1/x$.

Using Fact 1 and denoting $\boldsymbol{Q} = \boldsymbol{q}\boldsymbol{q}^T$, problem (17) is further reformulated as

$$\begin{aligned} \underset{\boldsymbol{Q},\delta}{\texttt{maximize}} \ \ &\sum_{m=1}^{N_R}\log_2\Big(Tr(\boldsymbol{C}_m\boldsymbol{Q}) + \sigma^2\Big) + \log_2\delta \\ &\quad - N_R\delta\Big(Tr(\boldsymbol{C}_{N_R+1}\boldsymbol{Q}) + \sigma^2\Big) \\ \texttt{s.t.} \ \ &\boldsymbol{Q} \succeq \boldsymbol{0}, \quad \delta > 0, \\ &Tr(\boldsymbol{Q}) = 2Q_{N_R+1,N_R+1}, \\ &Q_{N_R+1,N_R+1} \le P_k^{(n)}, \\ &rank(\boldsymbol{Q}) = 1. \end{aligned} \tag{18}$$

We consider to optimize (18) alternatively with respect to $\boldsymbol{Q}$ and δ. Concretely, we obtain δ optimally as $\delta^* = \Big(Tr(\boldsymbol{C}_{N_R+1}\boldsymbol{Q}) + \sigma^2\Big)^{-1}$ for given $\boldsymbol{Q}$ from Fact 1, and find the optimal $\boldsymbol{Q}^*$ for given δ. However, it is a difficult to find the optimal $\boldsymbol{Q}^*$ even for (18) due to the rank-1 constraint, i.e., $rank(\boldsymbol{Q}) = 1$. To tackle the problem, we consider two different cases separately for the number of coordinated BBUs B_n.

TABLE I
LOWER BOUNDED, DISTRIBUTED ALGORITHM

Given $\boldsymbol{G}_k^{(n,n)} = \boldsymbol{U}_k^{(n,n)} \boldsymbol{\Sigma}_k^{(n,n)} \left(\boldsymbol{V}_k^{(n,n)}\right)^H$

1. Randomly generate $\boldsymbol{Q} \succeq \boldsymbol{0}$ such that $Tr(\boldsymbol{Q}) \leq 2Q_{N_R+1,N_R+1}$ and $Q_{N_R+1,N_R+1} \leq P_k^{(n)}$.
2. Set $\delta = \left(Tr(\boldsymbol{C}_{N_R+1}\boldsymbol{Q}) + \sigma^2\right)^{-1}$.
3. Update $\boldsymbol{Q}$ by solving problem (20).
4. Repeat steps 2 and 3 until both $\boldsymbol{Q}$ and δ converge within the prescribed accuracy.
5. $\boldsymbol{T}_k^{(n)} = \boldsymbol{V}_k^{(n,n)} \boldsymbol{\Lambda}_k^{(n)} \left(\boldsymbol{U}_k^{(n,n)}\right)^H$, where $[\boldsymbol{\Lambda}_k^{(n)}]_{i,i} = \sqrt{Q_{i,i}}$, $i = 1, \ldots, N_R$.

1) Case when $B_n = 1$: In this case, we define $\beta_m = \alpha_m^2$, $\forall m = 1, \ldots, N_R$. Problem (18) given δ is transformed to a convex optimization problem (cf. (14)) as

$$\begin{aligned}\underset{\{\beta_m\}_{m=1}^{N_R}}{\texttt{maximize}} \quad & \sum_{m=1}^{N_R} \log_2\left(s_m^2\beta_m + \frac{\sum_{l\neq n}^N B_l}{N_{DA}N_R}\sum_{i=1}^{N_R} s_i^2 \sum_{i=1}^{N_R}\beta_i + \sigma^2\right) \\ & + N_R\log_2(\delta) - \delta\frac{\sum_{l\neq n}^N B_l}{N_{DA}}\sum_{i=1}^{N_R} s_i^2\sum_{i=1}^{N_R}\beta_i - \delta\sigma^2 \\ \texttt{s.t.} \quad & \sum_{i=1}^{N_R}\beta_i \leq P_k^{(n)}. \end{aligned} \tag{19}$$

The problem (19) is solvable by using, e.g., the primal-dual method [13].

2) Case when $B_n \geq 2$: In this case, we show that the rank-1 constraint in (18) can be omitted from Proposition 2.

Proposition 2: Consider a relaxed problem of (18) with given δ as

$$\begin{aligned}\underset{\boldsymbol{Q}}{\texttt{maximize}} \quad & \sum_{m=1}^{N_R}\log_2\left(Tr(\boldsymbol{C}_m\boldsymbol{Q}) + \sigma^2\right) + \log_2\delta \\ & - N_R\delta\Big(Tr(\boldsymbol{C}_{N_R+1}\boldsymbol{Q}) + \sigma^2\Big) \\ \texttt{s.t.} \quad & \boldsymbol{Q} \succeq \boldsymbol{0}, \\ & Tr(\boldsymbol{Q}) = 2Q_{N_R+1,N_R+1}, \\ & Q_{N_R+1,N_R+1} \leq P_k^{(n)}. \end{aligned} \tag{20}$$

The optimal solution $\boldsymbol{Q}^*$ of (20) always satisfies $rank(\boldsymbol{Q}^*) = 1$. In other words, (20) is equivalent to (18) with given δ.

Proof: The proof is based on the Lagrangian and Karush-Kuhn-Tucker (KKT) conditions [13] associated with (20). Detailed derivation is omitted due to the page limit. ■

Since (20) is convex in $\boldsymbol{Q}$, it is also solvable by using the primal-dual method [13].

Combining the two cases, we propose a lower-bound based, distributed design for BBU k serving user n as in Table I. Note that the proposed schemes to solve problems (19) and (20) are describable in the same framework. However, the solution $\boldsymbol{Q}^*$ is a diagonal matrix when $B_n = 1$, while its rank is always one when $B_n \geq 2$ as stated by Proposition 2.

B. Upper-Bound Based Design

The second suboptimal approach to solve problem (9) is maximizing an approximated *upper bound* (AUB) of (10), i.e., $R_{AUB}^{(n)} \geq \widehat{R}_{avg}^{(n)}$, which is obtained by ignoring the interference $\tilde{\boldsymbol{W}}^{(n)}$ given in (12). The resulting optimization is given as follows

$$\begin{aligned}\underset{\boldsymbol{\Lambda}_k^{(n)}}{\texttt{maximize}} \quad & R_{AUB}^{(n)} \\ \texttt{s.t.} \quad & Tr\Big(\left(\boldsymbol{\Lambda}_k^{(n)}\right)^2\Big) \leq P_k^{(n)}, \end{aligned} \tag{21}$$

where

$$\begin{aligned} R_{AUB}^{(n)} = \log_2\det\Bigg(&\left(\boldsymbol{\Sigma}_k^{(n,n)}\boldsymbol{\Lambda}_k^{(n)}\right)^2 \\ & + \left[\frac{2(B_n-1)}{N_R\sqrt{N_R}}Tr\left(\boldsymbol{\Sigma}_k^{(n,n)}\right)Tr\left(\left(\boldsymbol{\Lambda}_k^{(n,n)}\right)^2\right)\right]\boldsymbol{\Sigma}_k^{(n,n)}\boldsymbol{\Lambda}_k^{(n)} \\ & + \left[\frac{B_n-1}{N_R^2}Tr\left(\left(\boldsymbol{\Lambda}_k^{(n)}\right)^2\right)Tr\left(\left(\boldsymbol{\Sigma}_k^{(n,n)}\right)^2\right)\right]\boldsymbol{I} + \sigma^2\boldsymbol{I}\Bigg). \end{aligned} \tag{22}$$

Following a similar procedure as in Section III-A, we first reformulate (21) to

$$\begin{aligned}\underset{\boldsymbol{Q}}{\texttt{maximize}} \quad & \sum_{m=1}^{N_R}\log_2\left(Tr(\tilde{\boldsymbol{C}}_m\boldsymbol{Q}) + \sigma^2\right) \\ \texttt{s.t.} \quad & \boldsymbol{Q} \succeq \boldsymbol{0}, \\ & Tr(\boldsymbol{Q}) = 2Q_{N_R+1,N_R+1}, \\ & Q_{N_R+1,N_R+1} \leq P_k^{(n)}, \\ & rank(\boldsymbol{Q}) = 1, \end{aligned} \tag{23}$$

where

$$\tilde{\boldsymbol{C}}_m = \boldsymbol{D}_m^T\begin{bmatrix} s_m^2 & \frac{B_n-1}{N_R\sqrt{N_R}}s_m\sum_{i=1}^{N_R}s_i \\ \frac{B_n-1}{N_R\sqrt{N_R}}s_m\sum_{i=1}^{N_R}s_i & \frac{B_n-1}{N_R^2}\sum_{i=1}^{N_R}s_i^2\end{bmatrix}\boldsymbol{D}_m, \tag{24}$$

and $m = 1, \ldots, N_R$. Next, we consider two cases when $B_n = 1$ and $B_n \geq 2$, separately, and state problems (25) and (26), respectively, as

$$\begin{aligned}\underset{\{\beta_m\}_{m=1}^{N_R}}{\texttt{maximize}} \quad & \sum_{m=1}^{N_R}\log_2\left(s_m^2\beta_m + \sigma^2\right) \\ \texttt{s.t.} \quad & \sum_{i=1}^{N_R}\beta_i \leq P_k^{(n)}, \end{aligned} \tag{25}$$

$$\begin{aligned}\underset{\boldsymbol{Q}}{\texttt{maximize}} \quad & \sum_{m=1}^{N_R}\log_2\left(Tr(\tilde{\boldsymbol{C}}_m\boldsymbol{Q}) + \sigma^2\right) \\ \texttt{s.t.} \quad & \boldsymbol{Q} \succeq \boldsymbol{0}, \\ & Tr(\boldsymbol{Q}) = 2Q_{N_R+1,N_R+1}, \\ & Q_{N_R+1,N_R+1} \leq P_k^{(n)}. \end{aligned} \tag{26}$$

Herein, we show that (26) is equivalent to (23) from Proposition 3.

Proposition 3: Given that $B_n \geq 2$, the optimal solution $\boldsymbol{Q}^*$ of (26) always satisfies $rank(\boldsymbol{Q}^*) = 1$.

Proof: The proof is similar to the that of Proposition 2 and is thus omitted. ■

Based on these analyses, an upper-bound based, distributed algorithm is proposed in Table II. Again, note that the schemes

TABLE II
UPPER-BOUND, DISTRIBUTED ALGORITHM

Given $\boldsymbol{G}_k^{(n,n)} = \boldsymbol{U}_k^{(n,n)}\boldsymbol{\Sigma}_k^{(n,n)}\left(\boldsymbol{V}_k^{(n,n)}\right)^H$
1. Obtain $\boldsymbol{Q}$ by solving problem (26).
2. $\boldsymbol{T}_k^{(n)} = \boldsymbol{V}_k^{(n,n)}\,\boldsymbol{\Lambda}_k^{(n)}\left(\boldsymbol{U}_k^{(n,n)}\right)^H$, where $\left[\boldsymbol{\Lambda}_k^{(n)}\right]_{i,i} = \sqrt{Q_{i,i}}$, $i = 1,\ldots,N_R$.

to solve (25) and (26) can be given in the same framework. The solution $\boldsymbol{Q}^*$ when $B_n = 1$ is a diagonal matrix, while that when $B_n \geq 2$ always has rank one, as stated by Proposition 3. Finally, it is worth noting that problem (25) is solvable by using a well-known water-filling algorithm.

IV. COORDINATE DAS DESIGN

We will compare our distributed algorithm with a *coordinate* DAS where each BBU has full CSI, yet no information sharing among BBUs serving different users. Since there is no channel uncertainty, the objective is to maximize the minimum *instantaneous* rate of each user, instead of the average rate as in (9). The coordinate DAS precoder design is expressed as

$$\begin{aligned}
\underset{\left\{\boldsymbol{T}_k^{(n)}\right\}_{\substack{k=1,\ldots,B_n\\ n=1,\ldots,N}}}{\texttt{maximize}} \quad & \min_{n=1,\ldots,N} R_{ins}^{(n)} \\
\texttt{s.t.} \quad & Tr\left(\boldsymbol{T}_k^{(n)}\left(\boldsymbol{T}_k^{(n)}\right)^H\right) \leq P_k^{(n)}, \\
& k = 1,\ldots,B_n, \quad n = 1,\ldots,N, \qquad (27)
\end{aligned}$$

where $R_{ins}^{(n)}$ is defined in (4). Note that this scheme requires prohibitively huge fronthaul and backhaul data transmission for exchanging the CSI. Particularly, the exchange of CSI for *each* BBU grows as $O\left(NB_nN_{DA}N_R\right)$.

Defining the following aggregate matrices

$$\boldsymbol{H}_n^{(l)} \triangleq \left[\boldsymbol{H}_1^{(l,n)} \quad \ldots \quad \boldsymbol{H}_{B_l}^{(l,n)}\right], \qquad (28)$$

$$\boldsymbol{\Gamma}_n^{(l)} \triangleq diag\left(\boldsymbol{\Gamma}_1^{(l,n)},\ldots,\boldsymbol{\Gamma}_{B_l}^{(l,n)}\right), \qquad (29)$$

$$\boldsymbol{T}_n \triangleq \left[\boldsymbol{T}_1^{(n,n)} \quad \ldots \quad \boldsymbol{T}_{B_n}^{(n,n)}\right]^T, \qquad (30)$$

we rewrite (4) as

$$\begin{aligned}
R_{ins}^{(n)} = \log_2\det&\left(\sum_{l=1}^{N}\boldsymbol{H}_n^{(l)}\boldsymbol{\Gamma}_n^{(l)}\boldsymbol{\Psi}_l\boldsymbol{\Gamma}_n^{(l)}\left(\boldsymbol{H}_n^{(l)}\right)^H + \sigma^2\boldsymbol{I}\right) \\
&- \log_2\det\left(\sum_{l\neq n}^{N}\boldsymbol{H}_n^{(l)}\boldsymbol{\Gamma}_n^{(l)}\boldsymbol{\Psi}_l\boldsymbol{\Gamma}_n^{(l)}\left(\boldsymbol{H}_n^{(l)}\right)^H + \sigma^2\boldsymbol{I}\right), \qquad (31)
\end{aligned}$$

where $\boldsymbol{\Psi}_l = \boldsymbol{T}_l\,\boldsymbol{T}_l^H$ is the aggregate transmit covariance matrix of the BBUs serving user n. Denoting $\boldsymbol{\Psi}_n(k)$ as the kth block-diagonal matrix of $\boldsymbol{\Psi}_n$, we reformulate (27) as

$$\begin{aligned}
\underset{\left\{\boldsymbol{\Psi}_n\right\}_{n=1}^{N}}{\texttt{maximize}} \quad & \min_{n=1,\ldots,N} R_{ins}^{(n)} \\
\texttt{s.t.} \quad & \boldsymbol{\Psi}_n \succeq \boldsymbol{0}, \\
& Tr\left(\boldsymbol{\Psi}_n(k)\right) \leq P_k^{(n)}, \\
& k = 1,\ldots,B_n, \quad n = 1,\ldots,N, \qquad (32)
\end{aligned}$$

TABLE III
COORDINATE DAS ALGORITHM

1. Randomly generate $\boldsymbol{\Psi}_n$, $n = 1,\ldots,N$.
2. Set $\boldsymbol{S}_n$, $n = 1,\ldots,N$, as in (35).
3. Given $\left\{\boldsymbol{S}_n\right\}_{n=1}^N$, update $\left\{\boldsymbol{\Psi}_n\right\}_{n=1}^N$ by solving problem (34).
4. Repeat steps 2 and 3 until both $\left\{\boldsymbol{\Psi}_n\right\}_{n=1}^N$ and $\left\{\boldsymbol{S}_n\right\}_{n=1}^N$ converge within the prescribed accuracy.

where the second constraint is the power limitation on BBU k that serves user n.

In general, problem (32) is NP-hard [12], and thus, it is highly possible that no global optimum is achievable in polynomial time. Hence, we consider a suboptimal algorithm based on alternating optimization [12] as a benchmark. We first note the following fact.

Fact 2: Let $\boldsymbol{X} \in \mathbb{C}^{N\times N}$, $\boldsymbol{X} \succ \boldsymbol{0}$. We have

$$-\log\det\boldsymbol{X} = \max_{\boldsymbol{Z}\succeq\boldsymbol{0}} \quad -Tr(\boldsymbol{Z}\boldsymbol{X}) + \log\det\boldsymbol{Z} + N, \qquad (33)$$

and the equality is achieved if and only if $\boldsymbol{Z}^* = \boldsymbol{X}^{-1}$.

Applying Fact 2 to (32), we obtain an equivalent problem.

$$\begin{aligned}
\underset{\left\{\boldsymbol{S}_n,\boldsymbol{\Psi}_n\right\}_{n=1}^{N}}{\texttt{maximize}} \quad & t \\
\texttt{s.t.} \quad & \boldsymbol{\Psi}_n \succeq \boldsymbol{0}, \\
& Tr\left(\boldsymbol{\Psi}_n(k)\right) \leq P_b^{(k)}, \\
& \log_2\det\left(\sum_{l=1}^{N}\boldsymbol{H}_n^{(l)}\boldsymbol{\Gamma}_n^{(l)}\boldsymbol{\Psi}_l\boldsymbol{\Gamma}_n^{(l)}\left(\boldsymbol{H}_n^{(l)}\right)^H + \sigma^2\boldsymbol{I}\right) \\
& - Tr\left(\boldsymbol{S}_n\sum_{l\neq n}^{N}\boldsymbol{H}_n^{(l)}\boldsymbol{\Gamma}_n^{(l)}\boldsymbol{\Psi}_l\boldsymbol{\Gamma}_n^{(l)}\left(\boldsymbol{H}_n^{(l)}\right)^H + \sigma^2\boldsymbol{S}_n\right) \\
& + \log_2\det\boldsymbol{S}_n \geq t, \\
& k = 1,\ldots,B_n, \quad n = 1,\ldots,N, \qquad (34)
\end{aligned}$$

Since (34) is convex in each of $\left\{\boldsymbol{\Psi}_n\right\}_{n=1}^N$ or $\left\{\boldsymbol{S}_n\right\}_{n=1}^N$ but not jointly, we use an alternating optimization algorithm to solve it. Concretely, given $\left\{\boldsymbol{\Psi}_n\right\}_{n=1}^N$, we update (cf. Fact 2)

$$\boldsymbol{S}_n = \left(\sum_{l\neq n}^{N}\boldsymbol{H}_n^{(l)}\boldsymbol{\Gamma}_n^{(l)}\boldsymbol{\Psi}_l\boldsymbol{\Gamma}_n^{(l)}\left(\boldsymbol{H}_n^{(l)}\right)^H + \sigma^2\boldsymbol{I}\right)^{-1}. \qquad (35)$$

Also, given $\left\{\boldsymbol{S}_n\right\}_{n=1}^N$, $\left\{\boldsymbol{\Psi}_n\right\}_{n=1}^N$ is obtained by solving (34). The detail algorithm is given in Table III.

V. NUMERICAL RESULTS AND DISCUSSIONS

In this section, we evaluate the user rate of the distributed and coordinate algorithms given in Section III and IV, respectively. For the sake of comparison, we also evaluate the performance of a DPC-based precoder design for *cooperative* systems, i.e., each BBU has full CSI and shares the transmitted signal data. We consider nine BBUs each with 36 DAs and identical transmit power P_T. The detailed network parameters is given in Table IV. Here, the users are uniformly generated up to 9 inside the operating region as illustrated in Fig. 2. The closest DAs are allocated to each user. The channel links are

TABLE IV
CELLULAR SYSTEM PARAMETERS

Network Parameters	
BBU model	square grid (1.2 km^2)
# of BBUs	9
# of DAs/BBU	36
# of user antennas	2
# of serving DAs/user	4
Channel Parameters [14]	
Path-loss model	$g - 128 + 10\log_{10}(d^{-\mu})$
Path-loss exponent	$\mu = 3.7$
Feeder loss and antenna gain	$g = 15$ dB
Small-scale fading	$\mathcal{CN}(0,1)$
Bandwidth	Ω = 10 MHz
Transmit power	0 – 30 dBm
AWGN power	σ_k^2 = -174 dBm/Hz

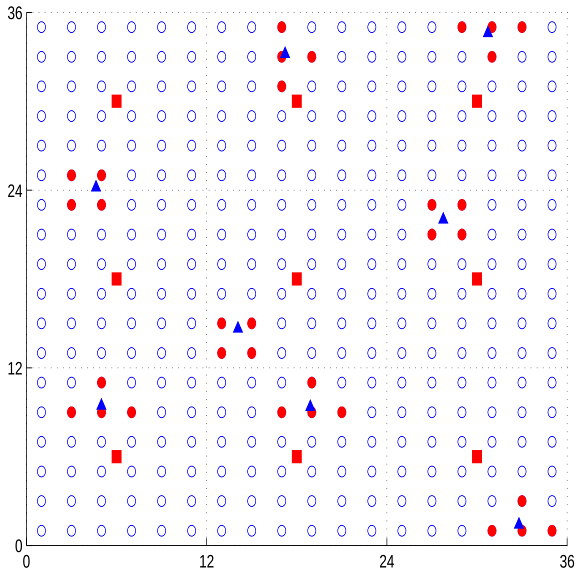

Fig. 2: A channel realization of DAS with nine BBUs. Here, "■", "○", "●", "▲", and "$\cdots$" denote BBUs, DAs, selected DAs, users, and BBU coverage, respectively.

then generated based on the channel parameters in Table IV. The number of channel realizations is 300.

In Fig. 3, we compare the average rate and minimum rate per user for the distributed, coordinate, and cooperative schemes. The lower-bound based, distributed scheme does not perform well since it is a strict lower bound while the interference is small in this setup. In certain cases with severe interference, the lower-bound based scheme would perform better than the upper-bound based design. It is observed that the rate gap between the upper-bound based, distributed scheme and the coordinate precoder design is negligible in the low-SNR regime, and it grows up to 3% at SNR $=$ 30 dB. Further comparisons show that the rate performance of the upper-bound based, distributed scheme is also similar to that of the cooperative DPC scheme at low-SNR regime. The performance difference is only 8% even at a high value of SNR $=$ 30 dB.

VI. Conclusions

In this study, we investigate distributed antenna systems with multiple baseband units and limited CSI. We propose two distributed algorithms based on lower and upper bounds, respectively, for the precoder design. The distributed algorithms are then compared with a state-of-the-art coordinate and DPC-based schemes that utilize full CSI. Simulation results

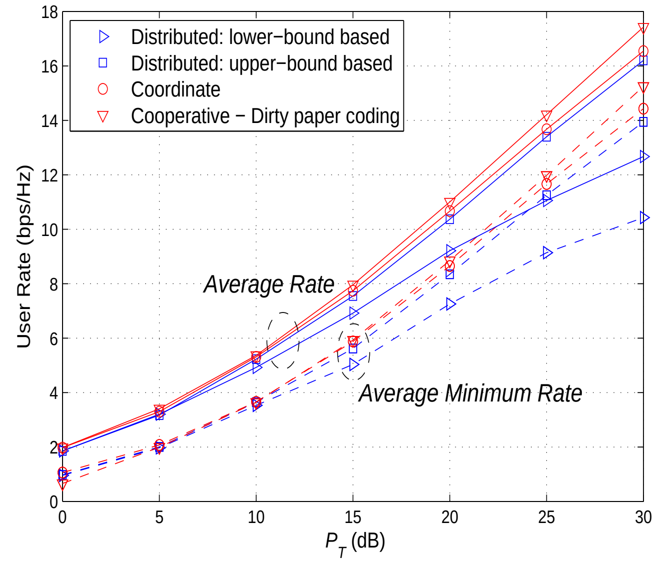

Fig. 3: Comparison of the user rates obtained from distributed, coordinate, and cooperative precoder designs.

show that comparable average and minimum rates are achieved by the proposed upper-bound based, distributed scheme and benchmark precoder designs.

References

[1] N. Vucic, H. Boche, and S. Shi, "Robust transceiver optimization in downlink multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3576–3587, Sep. 2009.
[2] J. Wang, M. Bengtsson, B. Ottersten, and D. P. Palomar, "Robust MIMO precoding for several classes of channel uncertainty," *IEEE Trans. Signal Process.*, vol. 61, no. 12, pp. 3056–3070, June 2013.
[3] E. Bjornson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4298–4310, Aug. 2010.
[4] R. Mochaourab and E. A. Jorswieck, "Optimal beamforming in interference networks with perfect local channel information," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1128–1141, Mar. 2011.
[5] L. Dai, S. Zhou, and Y. Yao, "Capacity analysis in CDMA distributed antenna systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2613–2620, Nov. 2005.
[6] W. Choi and J. G. Andrews, "Downlink performance and capacity of disributed antenna systems in a multicell environment," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 69–73, Jan. 2007.
[7] H. Kim, S.-R. Lee, K.-J. Lee, and I. Lee, "Transmission schemes based on sum rate analysis in distributed antenna systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1201–1209, Mar. 2012.
[8] A. Liu and V. K.N. Lau, "Joint power and antenna selection optimization for energy efficiency in large cloud radio access networks," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1319–1328, Mar. 2014.
[9] J. Joung, Y. K. Chia, and S. Sun, "Energy-efficient, large-scale distributed-antenna system for multiple users," *IEEE J. Sel. Topics Signal Process.*, vol. 8, pp. 954–965, Sep. 2014.
[10] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai (Shitz), "Joint precoding and multivariate backhaul compression for the downlink of cloud radio access networks," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5646–5658, Nov. 2013.
[11] J. Zhang, X. Yuan, and L. Ping, "Hermitian precoding for distributed MIMO systems with individual channel state information," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 241–251, Feb. 2013.
[12] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "Linear transceiver design for a MIMO interfering broadcast channel achieving max-min fairness," *Signal Process., special issue on sensor array processing*, vol. 93, no. 12, pp. 3327–3340, Dec. 2013.
[13] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
[14] "LTE; E-UTRA; RF requirements for LTE pico node B," ETSI, Tech. Rep. 136 931 V9.0.0, 2011. [Online]. Available at http://www.etsi.org/deliver/