Robust Distributed Compression for Cloud Radio Access Networks

1Seok-Hwan Park, 1Osvaldo Simeone, 2Onur Sahin and 3Shlomo Shamai (Shitz)
1CWCSPR, New Jersey Institute of Technology, 07102 Newark, New Jersey, USA
2InterDigital Inc., Melville, New York, 11747, USA
3Department of Electrical Engineering, Technion, Haifa, 32000, Israel
Email: {seok-hwan.park, osvaldo.simeone}@njit.edu, Onur.Sahin@interdigital.com, sshlomo@ee.technion.ac.il

Abstract—This work studies distributed compression for the uplink of a cloud radio access network, where multiple multi-antenna base stations (BSs) communicate with a central unit, also referred to as cloud decoder, via capacity-constrained backhaul links. Distributed source coding strategies are potentially beneficial since the signals received at different BSs are correlated. However, they require each BS to have information about the joint statistics of the received signals across the BSs, and are generally sensitive to uncertainties regarding such information. Motivated by this observation, a robust compression method is proposed to cope with uncertainties on the correlation of the received signals. The problem is formulated using a deterministic worst-case approach, and an algorithm is proposed that achieves a stationary point for the problem. From numerical results, it is observed that the proposed robust compression scheme compensates for a large fraction of the performance loss induced by the imperfect statistical information.

I. INTRODUCTION

On the uplink of a cloud radio access network, the base stations (BSs) operate as terminals that relay “soft” information to the cloud decoder regarding the received baseband signals (see, e.g., [1]). For the communication from the BSs to the cloud decoder, distributed source coding strategies are generally beneficial since the signals received at different BSs are correlated. This was first demonstrated in [2] and [3] for single-antenna and multi-antenna BSs, respectively (see also [4]-[7] for related works).

This work is motivated by the observation that the performance of distributed source coding is sensitive to errors in the knowledge of the joint statistics of the received signals at the BSs. This is due to the potential inability of the cloud decoder to decompress the signal received by a BS in case the statistical correlation is less pronounced than envisaged in the code design phase (see, e.g., [8]). As a solution to this problem, here we propose a distributed compression scheme that is robust to uncertainties on the statistics of interest.

The proposed robust distributed compression strategy assumes that the knowledge of the joint statistics available at each BS is imperfect. To model the uncertainty, we adopt a deterministic additive error model with bounds on eigenvalues of the error matrix similar to [9][10] (see also [11]). The problem is formulated following a deterministic worst-case approach and a solution that achieves a stationary point of this problem is provided by solving Karush-Kuhn-Tucker (KKT) conditions [12], which are also shown to be necessary for optimality. We conclude the paper with numerical results in Sec. IV.

Notation: The probability density functions (pdfs) is denoted by \( p(x) \) of a random variable \( X \) and similar notations are used for joint and conditional distributions. Given a vector \( x = [x_1, x_2, ..., x_n]^T \), we define \( x_S \) for a subset \( S \subseteq \{1, 2, ..., n\} \) as the vector including, in arbitrary order, the entries \( x_i \) with \( i \in S \). Notation \( \Sigma_x \) is used for the correlation matrix of random vector \( x \); \( \Sigma_{xy} \) represents the cross-correlation matrix of \( x \) and \( y \); and \( \Sigma_{xy} \) represents the “conditional” correlation matrix of \( x \) given \( y \), namely \( \Sigma_{xy} = \Sigma_x - \Sigma_{xy} \Sigma_y \Sigma_{yx} \). Notation \( \mathcal{H}^n \) represents the set of all \( n \times n \) Hermitian matrices.

II. SYSTEM MODEL

We consider a cluster of cells, which includes a total number \( N_B \) of BSs, each being either a Macro BS (MBS) or a Home BS (HBS), and \( N_M \) active MSs (see Fig. 1). We denote the set of all BSs as \( \mathcal{N}_B = \{1, ..., N_B\} \). Each \( i \)th BS is connected to the cloud decoder via a finite-capacity backhaul link of capacity \( C_i \) and has \( n_{B,i} \) antennas, while each MS has \( n_{M,i} \) antennas. Throughout the paper, we focus on the uplink.

Defining \( H_{ij} \) as the \( n_{B,i} \times n_{M,j} \) channel matrix between the \( j \)th MS and the \( i \)th BS, the overall channel matrix toward
BS $i$ is given as the $n_{B,i} \times n_M$ matrix

$$H_i = [H_{i1} \cdots H_{iN_M}],$$

with $n_M = \sum_{i=1}^{N_M} n_{M,i}$. Assuming that all the $N_M$ MSs in a cluster are asynchronous, at any discrete-time channel use of a given time-slot, the signal received by the $i$th BS is given by

$$y_i = H_i x + z_i.$$  

In (2), vector $x = [x_1^T \cdots x_{N_M}^T]^T$ is the $n_M \times 1$ vector of symbols transmitted by all the MSs in the cluster at hand. The noise vectors $z_i$ are independent over $i$ and are distributed as $z_i \sim \mathcal{CN}(0, \Sigma_i)$ for $i \in \{1, \ldots, N_B\}$.

Using standard random coding arguments, the coding strategies employed by the MSs in each time-slot entail a distribution $p(x)$ on the transmitted signals that factorizes as $p(x) = \prod_{i=1}^{N_M} p(x_i)$, since the signals sent by different MSs are independent. We will assume throughout that the distribution $p(x_i)$ of the signal transmitted by the $i$th MS is given as $x_i \sim \mathcal{CN}(0, \Sigma_{x_i})$ for a given covariance matrix $\Sigma_{x_i}$.

The BSs communicate with the cloud by providing the latter with soft information derived from the received signal. We consider compression strategies that do not require the BSs to know the codebooks employed by the MSs. Using a conventional rate-distortion theory arguments, a compression strategy for the $i$th BS is described by a test channel $p(\tilde{y}_i|x_i)$ that describes the relationship between the signal to be compressed, namely $y_i$, and its description $\tilde{y}_i$ to be communicated to the cloud (see, e.g., [13]). It is clear that the outcome of such compression is limited to $C_i$ bits per received symbol. The cloud decodes jointly the signals $x$ of all MSs based on all the descriptions $\tilde{y}_i$ for $i \in N_B$, so that, from standard information-theoretic considerations, the achievable sum-rate is given by

$$R_{\text{sum}} = I(x; \tilde{y}_{N_B}).$$

Since the signals $y_i$ measured by different BSs are correlated, distributed source coding techniques have the potential to improve the quality of the descriptions $\tilde{y}_i$ [2]. Specifically, given compression test channels $p(\tilde{y}_i|y_i)$, the descriptions $\tilde{y}_i$ can be recovered at the cloud as long as the capacities $C_i$ satisfy the following conditions (see, e.g., [5][13])

$$C_{\pi(i)} \geq I(y_{\pi(i)}; \tilde{y}_{\pi(i)}|\tilde{y}_{\pi(1)},\ldots,\pi(i-1))$$

for all $i = 1, \ldots, N_B$ given an arbitrary permutation $\pi(i)$ of the indices $i \in \{1, \ldots, N_B\}$.

A. Problem Definition and Greedy Solution

In this section, we aim at maximizing the sum-rate (3) under the constraints (4) over the test channels $\{p(\tilde{y}_i|y_i)\}_{i \in N_B}$ and the BS permutation $\pi$. Since this optimization problem is generally complex, we propose a greedy approach in Algorithm 1 to the selection of the permutation $\pi$, while optimizing the test channels $p(\tilde{y}_i|y_i)$ at each step of the greedy algorithm. As a result of the algorithm, we obtain a permutation $\pi^*$ and feasible test channels $p^*(\tilde{y}_i|y_i)$. The key step of Algorithm 1 is the solution of problem (5) at BS $i$. 

Algorithm 1: Greedy approach to the selection of the ordering $\pi$ and the test channels $p(\tilde{y}_i|y_i)$.

1. Initialize set $S$ to an empty set, i.e., $S^{(0)} = \emptyset$. 
2. For $j = 1, \ldots, N_B$, perform the following steps. 
   i) Each $i$th BS with $i \in N_B - S$ evaluates the test channel $p(\tilde{y}_i|y_i)$ by solving the problem

$$\max_{p(\tilde{y}_i|y_i)} I(x;\tilde{y}_i|S) \text{ s.t. } I(y_i;\tilde{y}_i|S) \leq C_i.$$  

   Denote the optimal value of this problem as $\phi_i^*$ and an optimal test channel as $p^*(\tilde{y}_i|y_i)$.
   ii) Choose the BS $i \in N_B - S$ with the largest optimal value $\phi_i^*$ and add it to the set $S$, i.e., $S^{(j)} = S^{(j-1)} \cup \{i\}$ and set $\pi^*(j) = i$.

This problem selects the test channel $p(\tilde{y}_i|y_i)$ at BS $i$ so as to maximize the capacity increase due to the reception of the compressed signal $\tilde{y}_i$ under the backhaul constraint and under the assumption that the cloud decoder has side information $\tilde{y}_S$ with $S = \{\pi(1), \ldots, \pi(i-1)\}$. From now on, we refer to the compression based on (5) in Algorithm 1 as Max-Rate compression.

B. Max-Rate Compression

We now review the solution of problem (5) given in [3]. To describe the optimal solution to problem (5), we define the covariance matrix

$$\Sigma_{xS} = \Sigma_x - \Sigma_x H_S^\dagger \left( H_S \Sigma_x H_S^\dagger + \Sigma_t \right)^{-1} H_S \Sigma_x,$$

where $H_i = A_i H_j$ and $\Sigma_{tj} = A_i A_j^\dagger + I$. A matrix $A_j$ will be defined in Proposition 1. We then have the following result.

Proposition 1. [3] An optimal solution $p(\tilde{y}_i|y_i)$ to problem (5) is given by

$$\tilde{y}_i = A_i y_i + q_i,$$

where $q_i \sim \mathcal{CN}(0, I)$ is the compression noise, which is independent of $x$ and $z_i$, and matrix $A_i$ is such that $\Omega_i = A_i^\dagger A_i$, with

$$\Omega_i = U \text{diag}(\alpha_1, \ldots, \alpha_{n_{B,i}}) U^\dagger.$$  

In (8), we have used the eigenvalue decomposition $H_i \Sigma_{xS} H_i^\dagger + I = U \text{diag}(\lambda_1, \ldots, \lambda_{n_{B,i}}) U^\dagger$ with unitary $U$ and ordered eigenvalues $\lambda_1 \geq \cdots \geq \lambda_{n_{B,i}}$. The diagonal elements $\alpha_1, \ldots, \alpha_{n_{B,i}}$ are computed as

$$\alpha_l = \left[ \frac{1}{\mu} \left( 1 - \frac{1}{\lambda_l} \right) - 1 \right]^+, \quad l = 1, \ldots, n_{B,i},$$

where $\mu$ is such that the condition $\sum_{l=1}^{n_{B,i}} \log (1 + \alpha_l \lambda_l) = C_i$ is satisfied.

III. ROBUST OPTIMAL COMPRESSION

As seen in the previous section, the optimal compression resulting from the solution of problem (5) at the $i$th BS depends on the covariance matrix $\Sigma_{xS}$ in (6) of the vector of transmitted signals conditioned on the compressed version
the uncertainty set \( \log \det(\tilde{\Delta}) \) on the eigenvalues of \( H \) given as (8), where matrix \( \Sigma \) models the estimation error. We assume that the error matrix \( \Delta \) is only known to belong to a set \( U_{\Delta} \subseteq \mathcal{H}^{n_{M}} \), which models the uncertainty at the \( i \)th BS regarding matrix \( \Sigma_{x|y,s} \).

In general, in order to define the uncertainty set \( U_{\Delta} \), one can impose some bounds on given measures of the eigenvalues and/or eigenvectors of matrix \( \Delta_{x|y,s} \). Based on the observation that the mutual information \( I(x;\hat{y}|y,s) \) can be written as

\[
I(x;\hat{y}|y,s) = f_{i}(\Omega_{i},\tilde{\Delta}_{x|y,s}) - \log \det (I + \Omega_{i}),
\]

where we have defined for this section \( f_{i}(\Omega_{i},\tilde{\Delta}_{x|y,s}) = \log \det(I + \Omega_{i}(H_{i}\Sigma_{x|y,s}H_{i}^{\dagger} + \tilde{\Delta}_{x|y,s} + I)) \) and \( \tilde{\Delta}_{x|y,s} = H_{i}\Sigma_{x|y,s}H_{i}^{\dagger} \), we take the approach of bounding the uncertainty on the eigenvalues of \( \tilde{\Delta}_{x|y,s} \). Specifically, we define the uncertainty set \( U_{\Delta} \) as the set of Hermitian matrices \( \Delta_{x|y,s} \) such that conditions

\[
\lambda_{\min}(\tilde{\Delta}_{x|y,s}) \geq \lambda_{LB} \text{ and } \lambda_{\max}(\tilde{\Delta}_{x|y,s}) \leq \lambda_{UB}
\]

hold for given lower and upper bounds \( \lambda_{LB}, \lambda_{UB} \) on the eigenvalues of matrix \( \Delta_{x|y,s} \). Note that when using the additive model (10) a non-trivial lower bound \( \lambda_{LB} \) must satisfy \( \lambda_{LB} \geq \min(H_{i}\Sigma_{x|y,s}H_{i}^{\dagger}) \) in order to guarantee that matrix \( H_{i}\Sigma_{x|y,s}H_{i}^{\dagger} \) is positive semidefinite.

Under this model, the problem of deriving the optimal robust compression strategy can be formulated as

\[
\max_{\mu_{1},\ldots,\mu_{n_{B,i}}} \min_{\tilde{\Delta}_{x|y,s} \in \mathcal{H}^{n_{M}}} f_{i}(\Omega_{i},\tilde{\Delta}_{x|y,s}) - \log \det (I + \Omega_{i})
\]

s.t.

\[
\begin{align*}
\lambda_{\min}(\tilde{\Delta}_{x|y,s}) & \geq \lambda_{LB} \\
\lambda_{\max}(\tilde{\Delta}_{x|y,s}) & \leq \lambda_{UB}
\end{align*}
\]

Theorem 1. A stationary point for problem (13) can be found as (7) with matrix \( A_{i} \) such that \( \Omega_{i} = A_{i}^{\dagger}A_{i} \) is given as (8), where matrix \( U \) is obtained from the eigenvalue decomposition \( H_{i}\Sigma_{x|y,s}H_{i}^{\dagger} + I = \text{Udiag}(\lambda_{1},\ldots,\lambda_{n_{B,i}})U^{\dagger} \) and the diagonal elements \( \alpha_{1},\ldots,\alpha_{n_{B,i}} \) are calculated by solving the following mixed integer-continuous problem:

\[
\begin{align*}
\max_{\mu,\alpha_{1},\ldots,\alpha_{n_{B,i}}} & \; \sum_{l=1}^{n_{B,i}} \log \det(1 + \alpha_{l}) \\
\text{s.t.} & \; 0 < \mu < 1 \\
& \; \alpha_{l} \in \mathcal{P}(\mu), \quad l = 1,\ldots,n_{B,i} \\
& \; g_{l}^{U}(\alpha_{1},\ldots,\alpha_{n_{B,i}}) = C_{l}
\end{align*}
\]

where the functions \( g_{l}^{U} \) and \( g_{l}^{U} \) are defined as

\[
\begin{align*}
g_{l}^{U}(\alpha_{1},\ldots,\alpha_{n_{B,i}}) & = \sum_{l=1}^{n_{B,i}} \log \det(1 + \alpha_{l}c_{l}^{U}), \\
g_{l}^{U}(\alpha_{1},\ldots,\alpha_{n_{B,i}}) & = \sum_{l=1}^{n_{B,i}} \log \det(1 + \alpha_{l}c_{l}^{L}),
\end{align*}
\]

with \( c_{l}^{U} = \lambda_{l} + \lambda_{LB} \) and \( c_{l}^{U} = \lambda_{l} + \lambda_{UB} \); and the discrete set \( \mathcal{P}(\mu) \) is defined as

\[
\begin{align*}
\mathcal{P}(\mu) & = \begin{cases}
\{0\}, & \text{if } Q_{l} \geq 0, S_{l} \geq 0 \\
\{-Q_{l}+\sqrt{Q_{l}^{2}-4S_{l}}\}, & \text{if } Q_{l} \geq 0, S_{l} < 0 \\
\{-Q_{l}+\sqrt{Q_{l}^{2}-4S_{l}}\}, & \text{if } Q_{l} < 0, S_{l} \geq 0 \\
\{0\}, & \text{if } Q_{l} < 0, S_{l} < 0
\end{cases}
\end{align*}
\]

with \( Q_{l} \) and \( S_{l} \) given as

\[
Q_{l} = \frac{c_{l}^{U}(1 + \mu + (\mu - 1)c_{l}^{U})}{\mu c_{l}^{U} c_{l}^{L}}, \quad S_{l} = \frac{1}{2} + c_{l}^{L} - \frac{1}{\mu c_{l}^{U} c_{l}^{L}}
\]

Proof: See Appendix A.

The following corollary shows that, in some special case, the search over parameters \( \alpha_{l} \) is not necessary, since the sets \( \mathcal{P}(\mu) \) only contain one element.

Corollary 1. If \( \lambda_{UB} - \lambda_{LB} < 1 \), a stationary point for problem (13) is given by \( \alpha_{l} = \frac{-Q_{l}+\sqrt{Q_{l}^{2}-4S_{l}}}{2} \) for \( l = 1,\ldots,n_{B,i} \) with \( \mu \) such that the constraint \( g_{l}^{U} = C_{l} \) is satisfied.

Proof: If \( \lambda_{UB} - \lambda_{LB} < 1 \), \( \alpha_{l} \) is computed from (17) as \( \alpha_{l} = \frac{-Q_{l}+\sqrt{Q_{l}^{2}-4S_{l}}}{2} \) if \( \mu < \frac{(c_{l}^{U} - 1)}{c_{l}^{L}} \) and \( \alpha_{l} = 0 \) otherwise, which entails the claimed result by direct calculation.

IV. Numerical Results

In this section, we present numerical results for a single-cell scenarios. It is assumed that the considered cell has radius \( R_{cell} \), a single MBS and multiple HBSs. The MBS is located at cell center while the HBS and MS are randomly dropped within a circular cell according to uniform distribution. All channel elements of \( H_{i,j} \) are i.i.d. \( \mathcal{CN}(0, (D_{0}/d_{i,j})^{\nu}) \), where
the path-loss exponent $\nu$ is chosen as 3.5 and $d_{i,j}$ is the distance from MS $j$ to BS $i$. The reference distance $D_0$ is set to half of cell-radius, i.e., $D_0 = R_{cell}/2$. For simplicity, each MS uses single antenna, i.e., $n_{M,i} = 1$ with transmit power $P_{tx}$ such that the aggregated transmit vector $x$ has a covariance of $\Sigma_x = P_{tx}I$. The eigenvectors of the uncertainty matrix $\Delta_x \Sigma_x$ are selected randomly according to isotropic distribution on the column space of $H_i$ and the eigenvalues uniformly in the set (12) where $\lambda_{UB} = \lambda_{\min}(H_i \Sigma_x \Sigma_y \Sigma_y^H H_i^H)$ and $\lambda_{LB} = -\lambda_{\min}(H_i \Sigma_x \Sigma_y \Sigma_y^H H_i^H)$, respectively, to guarantee the positive semi-definiteness of matrix $H_i \Sigma_x \Sigma_y \Sigma_y^H H_i^H$ (see Sec. III). Moreover, it is assumed that the MBS is connected to the cloud via a backhaul link of capacity $C$ while the HBSs' backhaul is of capacity equal to a fraction of $C$, namely $\omega C$ with $0 < \omega \leq 1$.

In Fig. 2, per-MS sum-rate performance is evaluated for a single-cell with $N_B = 4$, namely one MBS and three randomly placed HBSs, $N_M = 12$, $n_{B,i} = 4$, $\omega = 0.5$ and $P_{tx} = 10\text{dB}$ versus the MBS backhaul capacity $C$. For reference, we plot the performance attained by a variation of the Max-Rate scheme that ignores side information and that of a scheme that operates by assuming that $\hat{\Sigma}_x \hat{\Sigma}_y$ is the true covariance matrix. Note that, in this case, the backhaul constraint (5) can be violated, which implies that the cloud decoder cannot recover the corresponding compressed signal $\hat{y}_i$ (labeled as “imperfect Side Information (SI)” in the figure).

The figure shows that the incorrect matrix $\hat{\Sigma}_x \hat{\Sigma}_y$ as being the actual one can result in severe performance degradation. However, this performance loss can be overcome by adopting the proposed robust algorithm, which shows intermediate performance between the ideal setting with perfect side information and that with no side information. We also observe the more pronounced performance gain of the proposed robust solution for a larger backhaul link capacity.

In a similar vein, in Fig. 3, we investigate the effect of the number $N_B - 1$ of HBSs. It is seen that, as the number of BSs grows, leveraging side information provides more relevant gains, so that, even assuming imperfect side information can be useful.

V. Conclusions

We have studied distributed compression schemes for the uplink of cloud radio access networks. We proposed a robust compression scheme for a practical scenario with inaccurate statistical information. The scheme is based on a deterministic worst-case problem formulation. Via numerical results, we have demonstrated that, while errors in the statistical model of the side information make distributed source coding strategy virtually useless, the proposed robust compression scheme allows to tolerate sizable errors without drastic performance degradation.

APPENDIX A

PROOF OF THEOREM 1

Since the problem (13) involves infinitely many inequality constraints, we first convert it into a problem with finite number of inequalities [14].

**Lemma 1.** Problem (13) is equivalent to the problem

$$\max_{\Omega_i \succeq 0} f_i(\Omega_i, \lambda_{LB} I) - \log \det (\Omega_i + I)$$  \hspace{1cm} (19)

$$\text{s.t.} \quad f_i(\Omega_i, \lambda_{UB} I) - C_i \leq 0.$$ 

**Proof:** First, we observe that problem (13) is equivalent
to the following problem with one inequality constraint:
\[
\max_{\Omega \geq 0, t} \quad t - \log \det (I + \Omega_t)
\]
\[
\text{s.t.} \quad \max_{\Delta x_{S|\Theta}} \max_{\Omega_i} \left\{ t - f_i \left( \Omega_i, \Delta x_{S|\Theta} \right), \quad f_i \left( \Omega_i, \Delta x_{S|\Theta} \right) - C_t \right\} \leq 0.
\]
In order to proceed, we need to maximize and minimize the function \( f_i \) with respect to \( \Delta x_{S|\Theta} \) for given \( \Omega_i \) under constraint (12). To this end, note that function \( f_i \) can be written as log det \( (I + K_i \Delta x_{S|\Theta}) = \sum_{l=1}^{n_{B,i}} \log \left( 1 + \lambda_l \left( K_i \Delta x_{S|\Theta} \right) \right) \), where \( K_i = \left( I + \Omega_i \left( H_i \Sigma_{S|\Theta} H_i^H + I \right) \right)^{-1} \Omega_i \) and \( \lambda_l(X) \) represents the \( l \)-th largest eigenvalue of \( X \). Finally, using the following eigenvalue inequalities [15],
\[
\lambda_l \left( K_i \right) \lambda_{\min} \left( \Delta x_{S|\Theta} \right) \leq \lambda_l \left( K_i \Delta x_{S|\Theta} \right) \leq \lambda_l \left( K_i \right) \lambda_{\max} \left( \Delta x_{S|\Theta} \right),
\]
for \( l = 1, \ldots, n_{B,i} \), the optimal values for the maximization and minimization of \( f_i \) are obtained by setting \( \Delta x_{S|\Theta} = \lambda_{UB} I \) and \( \Delta x_{S|\Theta} = \lambda_{LB} I \), respectively. This leads to problem (19).

Problem (19) is not convex. In the next two lemmas, we list some necessary conditions for the optimality of problem (19) whose proofs are straightforward.

Lemma 2. The KKT conditions are necessary conditions for optimality in problem (19).

Lemma 3. At any optimal point \( \Omega^*_i \) for problem (19), the backhaul capacity should be fully utilized, i.e., \( f_i (\Omega^*_i, \lambda_{UB} I) = C_t \).

Now, we can consider only the points satisfying the necessary conditions described in Lemmas 2 and 3. To this end, with the choice (8), the KKT conditions can be written as
\[
\frac{e^l_l}{1 + \alpha_i e^l_l} - \frac{1}{1 + \alpha_i} - \frac{\mu e^U}{1 + \alpha_i e^U} = \theta_l = 0, \quad l = 1, \ldots, n_{B,i},
\]
\[
\theta_l \alpha_i = 0, \quad \theta_l \geq 0, \quad l = 1, \ldots, n_{B,i},
\]
and
\[
g^U \left( \alpha_1, \ldots, \alpha_{n_{B,i}} \right) - C_t = 0,
\]
with Lagrange multipliers \( \theta_l \geq 0 \) for \( l = 1, \ldots, n_{B,i} \) and \( \mu \geq 0 \). Similar to Lemmas 2 and 3, the conditions (22a)-(22c) can be shown to be necessary for the optimality of the following problem.
\[
\max_{\alpha_1 \geq 0, \ldots, \alpha_{n_{B,i}} \geq 0} g^U \left( \alpha_1, \ldots, \alpha_{n_{B,i}} \right) - \sum_{l=1}^{n_{B,i}} \log \det (1 + \alpha_i),
\]
\[
\text{s.t.} \quad g^U \left( \alpha_1, \ldots, \alpha_{n_{B,i}} \right) - C_t = 0.
\]
But, according to the Weierstrass Theorem [12], the problem (23) has a solution due to the compact constraint set. Thus, we can find parameters \( \alpha_1, \ldots, \alpha_{n_{B,i}} \) satisfying the KKT conditions (22a)-(22c) with a proper choice of \( \mu \).

The discussion above shows that any solution of problem (23) provides a solution to the KKT conditions (22a)-(22c). Moreover, we show that \( \alpha_l \) must lie in the set \( P_{\lambda} (\mu) \) with \( \mu \in (0, 1) \) in order to satisfy the conditions (22a)-(22c), and thus we can limit the domain of the optimization (23) as done in (14). Firstly, from the following lemma, the search region for \( \mu \) can be restricted to the interval \( \mu \in (0, 1) \).

Lemma 4. For \( \mu = 0 \) and \( \mu \geq 1 \), the conditions (22a)-(22c) cannot be satisfied simultaneously.

Proof: With \( \mu = 0 \), it is impossible to satisfy (22a) and (22b) simultaneously. For \( \mu \geq 1 \), (22a) does not hold with nonnegative \( \alpha_l \).

Lemma 5. A value of \( \alpha_l \) with \( \alpha_l \notin P_{\lambda} \) cannot satisfy the conditions (22a)-(22c).

Proof: In order for (22a) and (22b) to hold together, parameter \( \alpha_l \) should be such that
\[
\alpha_l^2 + Q_l \alpha_l + S_l = 0, \quad \text{if} \ \alpha_l > 0,
\]
and
\[
\alpha_l^2 + Q_l \alpha_l + S_l \geq 0, \quad \text{if} \ \alpha_l = 0.
\]
By direct calculation, it follows that candidate \( \alpha_l \in P_{\lambda} (\mu) \) must hold in order to satisfy both (24) and (25).

REFERENCES


