Resource Allocation for Outdoor-to-Indoor Compress-and-Forward SUDAS with Independent Relay Processing

(Invited Paper)

Aravindh Krishnamoorthy*†, Robert Schober†, and Marco Breiling*
*Fraunhofer Institute for Integrated Circuits (IIS), Erlangen, Germany.
†Institute for Digital Communications, Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Germany.

***Abstract*—In this paper, we consider resource allocation for an outdoor-to-indoor shared user-equipment (UE)-side distributed antenna system (SUDAS) employing multiple independently operating compress-and-forward (CF) relays which utilize both licensed and unlicensed frequency bands to enhance indoor data throughput. First, a non-convex matrix-valued resource allocation problem for maximization of the weighted sum rate is formulated. Next, the non-convex problem is simplified to obtain a low-complexity suboptimal resource allocation algorithm based on sequential quadratic programming (SQP). The proposed algorithm is shown to provide excellent performance in practical scenarios. Furthermore, the algorithm has a low channel state information (CSI) feedback requirement and can accommodate arbitrary communication bands and technologies for indoor relaying. Therefore, the proposed CF-SUDAS scheme can help achieve high outdoor-to-indoor data throughput at low complexity, a crucial requirement for next generation wireless communication systems.**

I. Introduction

Outdoor-to-indoor communication is an important requirement for the next generation communication systems, such as those based on the upcoming 5th generation (5G) communication standard, owing to the increasing indoor usage of mobile video and multimedia applications [1]. A promising technique for improving the outdoor-to-indoor data throughput is spatial multiplexing enabled by multiple-input multiple-output (MIMO) systems. However, in practice, the spatial multiplexing gain is limited due to the low number of antennas that can be accommodated at mobile devices. To overcome this limitation while still enjoying the benefits of MIMO, the use of virtual antenna arrays (VAAs) [2] has been proposed. VAAs consist of adjacent single-antenna receivers which realize a virtual MIMO system by exchanging received signals, thereby emulating multi-antenna reception.

On the other hand, base station (BS) side distributed antenna systems (DASs) are relays which aid transmission of signals from a BS baseband unit to user equipments (UEs). DAS based schemes have been investigated to mitigate shadowing and penetration losses in indoor and outdoor-to-indoor communication scenarios [3], [4]. Furthermore, optimal resource allocation algorithms for DASs [5] and the system performance with imperfect channel state information (CSI) [6] have been studied. Traditionally, a BS baseband unit is connected to a DAS using a wired connection such as optical fibre or cable [4]. However, more recently, wired and wireless connections from BS baseband units to DASs have been investigated in the context of cloud radio access networks (C-RANs) [7], [8].

Similar to a conventional BS-side DAS, a UE-side DAS consists of relays, operating as a dedicated VAA, which forward the signals received from a BS to the UEs. Therefore, UE-side DAS achieves the virtual MIMO gain in single antenna UEs without the need for UE cooperation. However, such a system necessitates the use of wireless resources for both the BS-relay and relay-UE links. Therefore, the gains of UE-side DASs are limited by the system bandwidth which is a scarce resource in the licensed frequency bands.

The shared UE-side distributed antenna system (SUDAS) proposed in [9] overcomes this limitation by using a second unlicensed frequency band for indoor relaying. The availability of a large unlicensed band of about 7 GHz around 60 GHz, and the applicability of millimetre wave (mmWave) to short distance indoor communication provides the necessary bandwidth and technology to relay signals from a DAS to the UEs. In SUDAS, multiple shared indoor relays receive a MIMO signal from an outdoor BS and forward it, after amplification and frequency translation, over orthogonal mmWave frequency bands to indoor UEs. Hence, SUDAS converts the high outdoor spatial multiplexing gain obtained by virtual MIMO to a high indoor frequency multiplexing gain, effectively utilizing the large bandwidth in the unlicensed band to achieve a high system throughput.

However, the baseline SUDAS scheme described in [9] suffers from several practical problems. First, the scheme requires joint processing of the signals received at all relays, which is not practical because of the need for high speed interconnects between the relays. Second, resource allocation for this system requires channel state information (CSI) feedback for both the BS-relay and the relay-UE channels, i.e., twice the amount of feedback of conventional MIMO systems. Third, the use of amplify-and-forward (AF) relays in [9] limits flexibility in mmWave bandwidth usage.

In this paper, we address the shortcomings of the baseline SUDAS scheme by considering SUDAS with independently operating compress-and-forward (CF) relays. The contributions of this paper can be summarized as follows.

- We present a SUDAS scheme with independently oper-

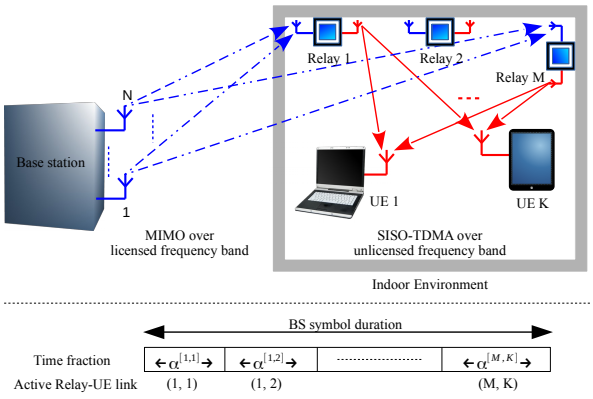

Fig. 1. The upper half of the figure shows the CF-SUDAS model. M single-antenna CF relays aid communication from an outdoor BS to K single-antenna indoor UEs. The lower half of the figure shows the time division multiple access (TDMA) protocol used for indoor relay-UE communication.

ating CF relays for high throughput outdoor-to-indoor communication. The scheme has a low CSI feedback requirement, owing to independent resource allocation on the BS-relay and relay-UE links, and offers flexibility with respect to indoor bandwidth usage.

- We formulate a non-convex resource allocation problem for weighted sum-rate maximization and propose a low-complexity suboptimal resource allocation algorithm based on sequential quadratic programming (SQP).
- We study the convergence of the resource allocation algorithm based on computer simulations, and compare its performance with that of the baseline SUDAS scheme [9].

The remainder of this paper is organized as follows. In Section II, we present the system and signal model for the proposed CF-SUDAS. In Section III, we formulate the non-convex resource allocation problem for weighted sum-rate maximization. In Section IV, we solve the resource allocation problem and develop a low-complexity, suboptimal resource allocation algorithm. The performance of the proposed algorithm is evaluated via computer simulations in Section V, and concluding remarks are presented in Section VI.

Notation: Boldface capital $\boldsymbol{X}$ and lower case $\boldsymbol{x}$ letters denote matrices and vectors, respectively. $\boldsymbol{X}^{\mathrm{H}}$, $\mathrm{Tr}(\boldsymbol{X})$, and $\det(\boldsymbol{X})$ denote the Hermitian transpose, trace, and determinant of matrix $\boldsymbol{X}$, respectively. $\mathbb{C}^{M\times N}$ and $\mathbb{R}^{M\times N}$ denote the set of all $M\times N$ matrices with complex-valued and real-valued entries, respectively. $\mathrm{diag}(d_1,\cdots,d_M)$ denotes a diagonal matrix with the diagonal elements given by $\{d_1,\cdots,d_M\}$. The circularly symmetric complex Gaussian (CSCG) distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is denoted by $\mathcal{CN}(\boldsymbol{\mu},\boldsymbol{\Sigma})$; $\sim$ stands for "distributed as." $\mathrm{E}\{\cdot\}$ denotes statistical expectation.

II. System Model

In this section, we present the system and signal models for CF-SUDAS.

A. SUDAS Downlink System Model

We consider a system comprising an outdoor BS equipped with N antennas, a set of M independently operating indoor CF relays (the CF-SUDAS), and K indoor UEs as shown in the upper half of Figure 1. The relays are equipped with one antenna respectively in the licensed and unlicensed frequency bands. The UEs are equipped with a single antenna operating in the unlicensed frequency band.

The BS employs multi-carrier downlink transmission with N_F subcarriers and symbol duration T operating in the sub-6 GHz licensed band. The relays receive a sub-6 GHz MIMO signal from the BS, quantize the received signal amplitudes, encode and modulate the quantized bits, and transmit the resulting *indoor signal* over the unlicensed band to the UEs. To simplify resource allocation and signal processing, interference between the relay-UE links is avoided by using an orthogonal time division multiple access (TDMA) protocol as shown in the lower half of Figure 1. In this paper, we assume that the relay-UE links operate in the mmWave band around 60 GHz. Furthermore, for simplicity of implementation, we restrict the quantization scheme to scalar quantization although the resource allocation algorithm described herein can be extended to vector quantization and distributed quantization using Wyner-Ziv binning [10] in a straightforward manner.

A given UE receives indoor signals from the M relays in different time slots and decodes them to recover the quantized version of the MIMO signal received by the relays. Hence, the relays act as shared virtual antennas for the UEs. The recovered quantized MIMO signal at the UEs is then used for MIMO decoding of the BS transmit signal.

B. BS-Relay Links

Let $\boldsymbol{x}^{[i,k]} \in \mathbb{C}^{N_{\mathrm{S}}\times 1}$, $N_{\mathrm{S}} \leq N, M$, denote the data symbol vector, where $\mathrm{E}\{\boldsymbol{x}^{[i,k]}(\boldsymbol{x}^{[i,k]})^{\mathrm{H}}\} = \boldsymbol{I}_{N_S}$, intended for UE k on subcarrier i. It is assumed that each subcarrier is allocated to a single UE. Let $\boldsymbol{P}^{[i,k]} \in \mathbb{C}^{N\times N_{\mathrm{S}}}$ denote the precoder matrix used to map the symbol vector of UE k to the N transmit antennas, and let $\boldsymbol{H}^{[i]} \in \mathbb{C}^{M\times N}$ denote the MIMO channel matrix between the BS and the M relays. The received signal at the M relays on subcarrier i allocated to UE k can be expressed as

$$\boldsymbol{y}^{[i,k]} = \boldsymbol{H}^{[i]}\boldsymbol{P}^{[i,k]}\boldsymbol{x}^{[i,k]} + \boldsymbol{n}^{[i]}, \tag{1}$$

where $\boldsymbol{n}^{[i]} \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_M)$ is the additive white Gaussian noise (AWGN) and $\boldsymbol{y}^{[i,k]} = [y_1^{[i,k]},\cdots,y_M^{[i,k]}]^{\mathrm{T}}$. The m-th element, $y_m^{[i,k]}$, of $\boldsymbol{y}^{[i,k]}$ denotes the received signal at relay m.

The M relays *independently* perform scalar quantization of the signal amplitudes received on the N_F subcarriers. At relay m, the quantization level for subcarrier i allocated to UE k is adjusted such that the quantization noise variance $q_m^{[i,k]}$ has a desired value. To aid analysis, the scalar quantization error at the relays is approximated as AWGN[1] [11]. Therefore, the

[1]The AWGN approximation is also valid for distributed [10] and vector [11] quantization.

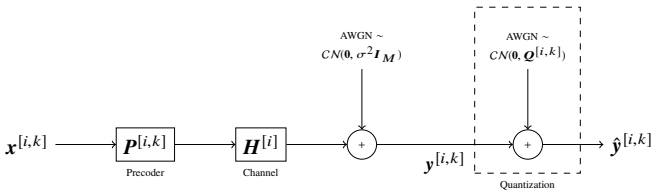

Fig. 2. Schematic illustration of the CF-SUDAS signal model in (2), where the quantization process is modelled as AWGN $\mathcal{CN}(\mathbf{0}, \boldsymbol{Q}^{[i,k]})$.

quantized MIMO signal $\hat{\boldsymbol{y}}^{[i,k]}$ on subcarrier i for user k at the M relays can be expressed as

$$\hat{\boldsymbol{y}}^{[i,k]} = \boldsymbol{H}^{[i]}\boldsymbol{P}^{[i,k]}\boldsymbol{x}^{[i,k]} + \boldsymbol{n}^{[i]} + \boldsymbol{n}_q^{[i,k]}, \tag{2}$$

where $\boldsymbol{n}_q^{[i,k]} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{Q}^{[i,k]})$ denotes the quantization noise modelled as AWGN and $\hat{\boldsymbol{y}}^{[i,k]} = [\hat{y}_1^{[i,k]}, \cdots, \hat{y}_M^{[i,k]}]^{\mathrm{T}}$. Due to the independent quantization at the relays, the quantization noise covariance matrix is diagonal, i.e., $\boldsymbol{Q}^{[i,k]} = \mathrm{diag}(q_1^{[i,k]}, \cdots, q_M^{[i,k]})$. The signal model described above is schematically illustrated in Figure 2.

If the quantized signal can be communicated error-free from the M relays to UE k, the UE can perform MIMO decoding on $\hat{\boldsymbol{y}}^{[i,k]}$ given in (2).

C. Relay-UE Links

After appropriate coding and modulation of the quantized signal, the relays transmit the resulting *indoor signal* to the UEs. In order to simplify the processing at the UEs, interference amongst relay-UE links is avoided using an orthogonal TDMA protocol as illustrated in the lower half of Figure 1. Thereby, relay m transmits to UE k using a time fraction $\alpha^{[m,k]}$ of the BS symbol duration T. For sustained communication, the time fractions must satisfy the condition $\sum_{k=1}^{K}\sum_{m=1}^{M}\alpha^{[m,k]} \leq 1$.

The resource allocation for each indoor relay-UE link is performed independently by the corresponding relay and is not considered here. Instead, we assume that the relay-UE link from relay m to UE k has an achievable rate of $C^{[m,k]}$ for the BS symbol duration. Hence, in the relaying phase, relay m transmits to UE k with an effective rate of $\alpha^{[m,k]}C^{[m,k]}$.

For error-free communication from relay m to UE k, the entropy of the quantized signal at the subcarriers allocated to UE k must be less than the effective relaying rate. This leads to the condition

$$\sum_{i=1}^{N_F} s^{[i,k]}\mathrm{I}(\hat{y}_m^{[i,k]}; y_m^{[i,k]}) \leq \alpha^{[m,k]}C^{[m,k]}, \forall\, m, k, \tag{3}$$

which is henceforth referred to as the *relay-UE rate constraint*. Here, $\mathrm{I}(\cdot;\cdot)$ denotes mutual information, and $s^{[i,k]} \in \{0,1\}\,\forall\, i,k$ are Boolean variables specifying whether or not subcarrier i is allocated to UE k.

For scalar quantization, the mutual information of the test channel at relay m for UE k can be expressed as [11]

$$\mathrm{I}(\hat{y}_m^{[i,k]}; y_m^{[i,k]}) = \log_2\left(1 + \frac{\boldsymbol{h}_m^{[i]}\boldsymbol{P}^{[i,k]}(\boldsymbol{P}^{[i,k]})^{\mathrm{H}}(\boldsymbol{h}_m^{[i]})^{\mathrm{H}} + \sigma^2}{q_m^{[i,k]}}\right), \tag{4}$$

where $\boldsymbol{h}_m^{[i]}$ is the m-th row of the MIMO channel matrix $\boldsymbol{H}^{[i]}$, i.e., $\boldsymbol{H}^{[i]} = [(\boldsymbol{h}_1^{[i]})^{\mathrm{T}}, \cdots, (\boldsymbol{h}_M^{[i]})^{\mathrm{T}}]^{\mathrm{T}}$.

Furthermore, for the Schur-concave function $\mathrm{I}(\cdot;\cdot)$, we have $\sum_{m=1}^{M}\mathrm{I}(\hat{y}_m^{[i,k]}; y_m^{[i,k]}) \geq \mathrm{I}(\hat{\boldsymbol{y}}^{[i,k]}; \boldsymbol{y}^{[i,k]})$, i.e., independent scalar quantization at the relays results in redundancy in the indoor signals. However, due to the large communication bandwidth in the unlicensed mmWave band used for indoor relaying and the short indoor communication distance, the achievable rates $C^{[m,k]}$ on the relay-UE links are typically much higher than the end-to-end MIMO rates. Therefore, the additional performance gains possible with more advanced quantization schemes are expected to be small.

In addition, due to our modelling of the relay-UE communication channel as finite-rate links with independent resource allocation, we can accommodate arbitrary unlicensed mmWave bandwidths and transmission formats for indoor relaying.

III. Problem Formulation

In this section, we present the expression for the weighted sum rate of the CF-SUDAS system and formulate a resource allocation problem for maximization of the weighted sum rate.

A. Weighted Sum Rate

Assuming error-free communication on the relay-UE links, UE k can perform MIMO decoding of the quantized signal given in (2). Hence, the achievable rate for UE k can be expressed as

$$R_k = \sum_{i=1}^{N_F} s^{[i,k]}\log_2\det\left(\boldsymbol{I}_M + \boldsymbol{H}^{[i]}\boldsymbol{P}^{[i,k]}(\boldsymbol{H}^{[i]}\boldsymbol{P}^{[i,k]})^{\mathrm{H}}\boldsymbol{T}^{[i,k]}\right), \tag{5}$$

where $\boldsymbol{T}^{[i,k]} = (\sigma^2\boldsymbol{I}_M + \boldsymbol{Q}^{[i,k]})^{-1} = \mathrm{diag}(t_1^{[i,k]}, \cdots, t_M^{[i,k]})$.

Given the achievable rates of the individual UEs, the weighted sum rate can be expressed as

$$R_{\mathrm{sum}} = \sum_{k=1}^{K}\mu_k R_k, \tag{6}$$

where the μ_k's are fixed weights which can be chosen to ensure user fairness [12, Sec. 4].

B. Resource Allocation Problem

In this subsection, we define the resource allocation problem for the system model under consideration. The optimization variables include: the BS precoder matrix $\boldsymbol{P}^{[i,k]}$, the subcarrier allocation variables $s^{[i,k]}$, the inverse noise covariance matrix $\boldsymbol{T}^{[i,k]}$, and the TDMA time fractions $\alpha^{[m,k]}$. Let $\mathcal{V} = \{\boldsymbol{P}^{[i,k]}, s^{[i,k]}, \boldsymbol{T}^{[i,k]}, \alpha^{[m,k]}\,\forall\, i,m,k\}$ denote the set of optimization variables. The optimal resource allocation $\mathcal{V}^*$ is the solution to the following problem:

$$\text{(P)} \quad \underset{\mathcal{V}}{\text{maximize}} \quad R_{\mathrm{sum}}$$

$$
\begin{aligned}
\text{s.t.}\quad & \text{C1:}\quad \sum_{i=1}^{N_F}\sum_{k=1}^{K} s^{[i,k]}\mathrm{Tr}((\boldsymbol{P}^{[i,k]})^{\mathrm{H}}\boldsymbol{P}^{[i,k]}) \leq P_B,\\
& \text{C2:}\quad \sum_{k=1}^{K}\sum_{m=1}^{M} \alpha^{[m,k]} \leq 1, \quad \text{C3:}\quad \sum_{k=1}^{K} s^{[i,k]} \leq 1\,\forall\, i,\\
& \text{C4:}\quad \epsilon_1 \leq t_m^{[i,k]} \leq \left(\tfrac{1}{\sigma^2} - \epsilon_1\right)\,\forall\, i,m,k,\\
& \text{C5:}\quad \alpha^{[m,k]} \geq 0\,\forall\, m,k, \qquad \text{C6:}\quad s^{[i,k]} \in \{0,1\}\,\forall\, i,k,\\
& \text{C7:}\quad c_{mk} \leq 0\,\forall\, m,k. \qquad (7)
\end{aligned}
$$

Here, constraint C1 is the BS power constraint that ensures a maximum transmission power of P_B. C2 is the TDMA constraint on the relays as discussed earlier. C3 ensures that each subcarrier is allocated at most to one UE. C4 ensures that the quantization noise variance is positive. In a practical system, zero and infinite quantization noise variance cannot be achieved. Therefore, a small positive constant $\epsilon_1 > 0$, which is chosen based on the computer precision, is used to ensure that the quantization noise is bounded[2]. C5 is the boundary constraint for the time fractions. C6 enforces that the subcarrier allocation variables are Boolean. C7 is the relay-UE rate constraint, with c_{mk} defined based on (3), (4) as

$$
\begin{aligned}
c_{mk} = & \sum_{i=1}^{N_F} s^{[i,k]} \log_2\left(1 + \frac{\boldsymbol{h}_m^{[i]}\boldsymbol{P}^{[i,k]}(\boldsymbol{P}^{[i,k]})^{\mathrm{H}}\boldsymbol{h}_m^{[i]\mathrm{H}} + \sigma^2}{\frac{1}{t_m^{[i,k]}} - \sigma^2}\right)\\
& - \alpha^{[m,k]} C^{[m,k]}. \qquad (8)
\end{aligned}
$$

We note that solving resource allocation problem (P) for CF-SUDAS requires only the knowledge of the indoor achievable rates $C^{[m,k]}$ instead of the full CSI on the relay-UE channel as in [9]. Therefore, the amount of CSI feedback to the BS, where the resource allocation is performed for the BS-relay links, is greatly reduced compared to [9].

However, finding a solution to (P) is difficult for three reasons. First, the objective function R_{sum} is non-convex. Second, the combinatorial constraint C6 and the relay-UE rate constraint C7 are non-convex. Third, there are matrix-valued optimization variables which increase the overall complexity of the solution. Therefore, an optimal solution to the problem may not be possible. Hence, in the following section, we aim to develop an efficient suboptimal solution to (P).

IV. Resource Allocation

In this section, we present a low complexity suboptimal solution to (P) and describe the associated resource allocation algorithm.

A. Problem Simplification

We begin by making the problem tractable by introducing three simplifications.

1) Enforcing equal quantization noise variance: We restrict the quantization noise variance at all the relays for user k on subcarrier i to be identical. This implies $t_m^{[i,k]} = t^{[i,k]}$, $\forall\, m$ and $\boldsymbol{T}^{[i,k]}$ is a scaled identity matrix, i.e., $\boldsymbol{T}^{[i,k]} = t^{[i,k]}\boldsymbol{I}_M$. The motivation behind this restriction is twofold. First, in small indoor environments, all relay-UE links typically have a line of sight, and thus achieve similar data rates which suggests that using equal quantization noise variances is near optimal. Second, in the absence of non-convex constraint C7, the use of equal quantization noise variances is optimal [13, Corollary 2.1], [14].

2) Scalarization of the problem: In order to tackle the matrix-valued optimization variable $\boldsymbol{P}^{[i,k]}$, we adopt a singular value decomposition (SVD) based matrix structure which leads to end-to-end diagonalization, i.e., the MIMO channel is decomposed into parallel scalar channels thereby simplifying the objective function. In general, the SVD-based structure is suboptimal except when constraint C7 is inactive in which case optimality is ensured by [13, Theorem 3.1]. Let $\boldsymbol{H}^{[i]} = \boldsymbol{U}^{[i]}\boldsymbol{\Lambda}^{[i]}(\boldsymbol{V}^{[i]})^{\mathrm{H}}$ denote the SVD of $\boldsymbol{H}^{[i]}$, where $\boldsymbol{U}^{[i]} \in \mathbb{C}^{M\times M}$ and $\boldsymbol{V}^{[i]} \in \mathbb{C}^{N\times N}$ are Unitary matrices containing the left and right singular vectors, respectively. Matrix $\boldsymbol{\Lambda}^{[i]} \in \mathbb{R}^{M\times N}$ is a diagonal matrix with main diagonal entries $\lambda_1^{[i]}, \cdots, \lambda_R^{[i]}$, where $R = \min(M,N)$ and the $\lambda_r^{[i]}$'s, $r = 1, \cdots, R$, are the singular values of $\boldsymbol{H}^{[i]}$ ordered as $\lambda_1^{[i]} \geq \lambda_2^{[i]} \geq \cdots \geq \lambda_R^{[i]}$. We choose the precoder matrix as $\boldsymbol{P}^{[i,k]} = \tilde{\boldsymbol{V}}^{[i]}(\boldsymbol{W}^{[i,k]})^{\frac{1}{2}}$, where $\tilde{\boldsymbol{V}}^{[i]} \in \mathbb{C}^{N\times N_S}$ is a matrix containing the N_S right singular vectors corresponding to the N_S largest singular values of $\boldsymbol{H}^{[i]}$, and $\boldsymbol{W}^{[i,k]} \in \mathbb{R}^{N_S\times N_S} = \mathrm{diag}(w_1^{[i,k]}, \cdots, w_{N_S}^{[i,k]})$ is a diagonal matrix. It can be verified that the combination of enforcing equal quantization noise variance and the chosen SVD based matrix structure leads to end-to-end diagonalization. Based on the above simplifications, constraint C1 can be rewritten as

$$
\text{C1}'\text{:}\ \sum_{k=1}^{K}\sum_{i=1}^{N_F}\sum_{n=1}^{N_S} s^{[i,k]} w_n^{[i,k]} \leq P_B. \qquad (9)
$$

3) Time sharing relaxation: In order to tackle the combinatorial constraints C3, C6, we use the time-sharing relaxation from [12], [15] where the Boolean allocation variables are relaxed to real-valued variables. In the presence of convex constraints, the time-sharing relaxation has zero duality gap resulting in an optimal subcarrier allocation. However, in the presence of non-convex constraint C7, strong duality cannot be guaranteed. Therefore, the time-sharing relaxation may result in a suboptimal solution. With the time-sharing relaxation, constraint C3 remains unchanged, however constraint C6 is replaced by

$$
\text{C6}'\text{:}\ \epsilon_2 \leq s^{[i,k]} \leq 1\,\forall\, i,k, \qquad (10)
$$

where $\epsilon_2 > 0$ is a small design constant, which is chosen according to computer precision to avoid the discontinuity of the objective function given below at $s^{[i,k]} = 0$.

B. Problem Reformulation

Using the simplifications outlined above, the weighted sum rate (objective function) can be rewritten as

$$
R'_{\text{sum}} = \sum_{k=1}^{K} \mu_k R'_k, \qquad (11)
$$

[2]Note that the use of ϵ_1 necessitates a minimum achievable rate $C_{\min}^{[m,k]}$, $\forall\, m,k$, for (P) to be feasible, which tends to zero as $\epsilon_1 \to 0$.

where

$$R'_k = \sum_{i=1}^{N_F} \sum_{n=1}^{N_S} s^{[i,k]} \log_2(1 + (\lambda_n^{[i]})^2 \frac{\hat{w}_n^{[i,k]}}{s^{[i,k]} t^{[i,k]}}). \quad (12)$$

Here, $\hat{w}_n^{[i,k]} = s^{[i,k]} w_n^{[i,k]}$ is an auxiliary variable introduced for time-sharing. Furthermore, constraint C7 is replaced by

$$\text{C7}': \quad c'_{mk} \leq 0 \, \forall \, m, k \quad (13)$$

with

$$c'_{mk} = \sum_{i=1}^{N_F} s^{[i,k]} \log_2 \left(1 + \frac{\sum_{n=1}^{N_S} |u_{m,n}^{[i]}|^2 (\lambda_n^{[i]})^2 \frac{\hat{w}_n^{[i,k]}}{s^{[i,k]}} + \sigma^2}{\frac{1}{t_m^{[i,k]}} - \sigma^2} \right) - \alpha^{[m,k]} C^{[m,k]}, \quad (14)$$

where $u_{m,n}^{[i]}$ is the element in the m-th row and n-th column of matrix $\boldsymbol{U}^{[i]}$. Next, constraint C1′ can be expressed in terms of the auxiliary variable as

$$\text{C1}'': \sum_{k=1}^{K} \sum_{i=1}^{N_F} \sum_{n=1}^{N_S} \hat{w}_n^{[i,k]} \leq P_B. \quad (15)$$

The updated set of optimization variables

$$\mathcal{W} = \{\hat{w}_1^{[i,k]}, \cdots, \hat{w}_{N_S}^{[i,k]}, s^{[i,k]}, t^{[i,k]}, \alpha^{[1,k]}, \cdots, \alpha^{[M,k]} \, \forall \, i, k\}$$

now consists only of scalar variables. With the above changes, problem (P) simplifies to problem (SP) given below.

$$\text{(SP)} \quad \underset{\mathcal{W}}{\text{maximize}} \quad R'_{\text{sum}}$$
$$\text{subject to} \quad \text{C1}'', \text{C2}, \text{C3}, \text{C4}, \text{C5}, \text{C6}', \text{C7}'. \quad (16)$$

C. Sequential Quadratic Programming

Despite the above simplifications, the objective function and constraint C7′ are still non-convex. Therefore, we use SQP [16] to obtain a locally optimum solution of problem (SP). In SQP, an associated quadratic program (QP) is iteratively solved until convergence. The associated QP is constructed as follows: (a) the objective function R'_{sum} is replaced by its quadratic approximation, and (b) constraint C7′ is replaced by its first order linear approximation.

In order to construct the QP, we first determine the quadratic approximation of the objective function around a point $\mathcal{W}_l$ as follows:

$$R''_{\text{sum}} = \sum_{k=1}^{K} \mu_k R''_k(\mathcal{W}, \mathcal{W}_l), \quad (17)$$

where

$$R''_k(\mathcal{W}, \mathcal{W}_l) = R'_k(\mathcal{W}_l) + \nabla R'_k(\mathcal{W}_l)(\mathcal{W} - \mathcal{W}_l) + (\mathcal{W} - \mathcal{W}_l)^{\text{H}} \boldsymbol{B}_k (\mathcal{W} - \mathcal{W}_l). \quad (18)$$

Here, $\nabla R'_k(\mathcal{W}_l)$ is the gradient of $R'_k(\mathcal{W})$ evaluated at point $\mathcal{W}_l$, and $\boldsymbol{B}_k$ is the Hessian matrix of the Lagrangian function of (SP)[3]. However, in the presence of non-convex constraint C7′, the Hessian matrix is not guaranteed to be negative-definite. Therefore, we use a negative-definite approximation of the Hessian matrix. Multiple methods for approximating the Hessian matrix for non-convex optimization problems have been studied in [16, Sec. 3.2], [17, Chap. 3]. The methods offer a trade-off between convergence speed and computational complexity. In this paper, we use a low-complexity diagonal negative-definite approximation for the Hessian matrix [16, Sec. 3.2] by neglecting the off-diagonal terms.

Furthermore, we construct the first order approximation of constraint C7′ around a point $\mathcal{W}_l$ as follows:

$$\text{C7}'(\mathcal{W}_l): c'_{mk}(\mathcal{W}_l) + \nabla c'_{mk}(\mathcal{W}_l)(\mathcal{W} - \mathcal{W}_l) \leq 0 \, \forall \, m, k, \quad (19)$$

where $\nabla c'_{mk}(\mathcal{W}_l)$ is the gradient of function $c'_{mk}(\mathcal{W})$ evaluated at $\mathcal{W}_l$.

Using the approximations of the objective function and constraint C7′($\mathcal{W}_l$) as given above, the QP, parametrized by point $\mathcal{W}_l$, is given as follows:

$$\text{(QP}(\mathcal{W}_l)\text{)} \quad \underset{\mathcal{W}}{\text{maximize}} \quad R''_{\text{sum}}$$
$$\text{subject to} \quad \text{C1}'', \text{C2}, \text{C3}, \text{C4}, \text{C5}, \text{C6}', \text{C7}'(\mathcal{W}_l). \quad (20)$$

The QP in (20) is solved iteratively with parameter $\mathcal{W}_l$, $l = 1, 2, 3, \cdots$, starting from an initial value $\mathcal{W}_0$. QPs can be solved efficiently using numerical solvers. Algorithms for solving QPs have been studied extensively since the 1950s; see [18, Chap. 4] for details. Before describing the detailed iterative resource allocation algorithm, we first provide the update rule for mapping the relaxed subcarrier allocation variables to their discrete counterparts.

Let $\mathcal{L}$ denote the Lagrangian function of (SP) given by

$$\mathcal{L} = \sum_{k=1}^{K} \mu_k \sum_{i=1}^{N_F} \sum_{n=1}^{N_S} s^{[i,k]} \log_2(1 + (\lambda_n^{[i]})^2 w_n^{[i,k]} t^{[i,k]}) - \sum_{i=1}^{N_F} g_3^{[i]} \left(\sum_{k=1}^{K} s^{[i,k]} - 1 \right) - \sum_{k=1}^{K} \sum_{i=1}^{N_F} g_{6,l}^{[i,k]} \left(-s^{[i,k]} + \epsilon_2 \right) - \sum_{k=1}^{K} \sum_{i=1}^{N_F} g_{6,u}^{[i,k]} \left(s^{[i,k]} - 1 \right) - \sum_{k=1}^{K} \sum_{m=1}^{M} g_{7,m}^{[k]} c'_{mk} + \text{terms independent of } s^{[i,k]}, \quad (21)$$

where $g_3^{[i]}$, $g_{6,l}^{[i,k]}$, $g_{6,u}^{[i,k]}$, $g_{7,m}^{[k]}$, $\forall \, i, k, m$ denote the Lagrangian variables corresponding to constraint C3, the lower bound in C6′, the upper bound in C6′, and C7′($\mathcal{W}_l$), respectively. For computing the Lagrangian, variable $w_n^{[i,k]} = \hat{w}_n^{[i,k]} / s^{[i,k]}$ is used instead of the auxiliary variable $\hat{w}_n^{[i,k]}$. Furthermore, let $\psi(i,k) = \frac{\partial \mathcal{L}}{\partial s^{[i,k]}}$ denote the partial derivative of $\mathcal{L}$ with respect to $s^{[i,k]}$ evaluated at a locally optimal point of (SP). Function $\psi(i,k)$ is referred to as the link quality indicator function [12] and is a measure of the relative merit of assigning subcarrier i to UE k. A subcarrier is allocated to the user with the highest corresponding link quality indicator, i.e.,

$$s^{[i,k]} \leftarrow \begin{cases} 1 & k = \arg\max_k \psi(i,k) \\ 0 & \text{otherwise.} \end{cases} \quad (22)$$

[3]We use the Hessian matrix of the Lagrangian function instead of the Hessian matrix of the objective function as it is well-known that the latter can cause the SQP to be unbounded, see [16, Sec. 2.2] for details.

Algorithm 1 Resource allocation algorithm for CF SUDAS.

1: Choose a feasible initial value for the variables $\mathcal{W}_0$ e.g. as $\hat{w}_n^{[i,k]} = \frac{P_B}{KN_F N_S}, s^{[i,k]} = 1/K, t_m^{[i,k]} = \epsilon_1, \alpha^{[m,k]} = \frac{1}{KM}, \forall\, i, m, k.$
2: **for** $l = 1, 2, 3, \cdots$ **do**
3: Update the gradient and Hessian of the objective function and the gradient of the non-convex constraint C7′ based on the point $\mathcal{W}_{l-1}$.
4: Solve the quadratic program QP($\mathcal{W}_{l-1}$) in (20).
5: **if** $|\mathcal{W}_l - \mathcal{W}_{l-1}| < N_{\text{tol}}$, the numerical tolerance, **then**
6: Exit l loop and goto 11.
7: **end if**
8: Calculate the maximum step size $0 < \alpha \leq 1$ using bisection search such that the updated solution satisfies non-convex constraint C7′.
9: Update the solution as $\mathcal{W}_l = \alpha\mathcal{W}_l + (1-\alpha)\mathcal{W}_{l-1}$.
10: **end for**
11: Update the subcarrier allocation variables according to (22).

D. Resource Allocation Algorithm

The proposed resource allocation algorithm for the CF-SUDAS system is summarized in Algorithm 1. First, the algorithm is initialized with a feasible initial value. Within the iteration loop, in Line 3, the gradient and the Hessian of the objective function R'_{sum} and the gradient of the non-convex constraint C7′ are updated according to the solution from the previous iteration. Next, in Line 4, the QP is solved using a numerical solver [18] to obtain a solution for the current iteration. If the previous solution and the current solution are equal, upto a numerical tolerance N_{tol}, the algorithm is deemed to have converged and the iterations are stopped. Otherwise, in Line 8, a step size α, $0 < \alpha \leq 1$, is determined based on bisection search such that the updated solution satisfies non-convex constraint C7′. Next, in Line 9, the current solution is updated using the determined step size. A stepwise update of the solution is necessary for convergence due to the presence of non-convex constraint C7′ [16, Sec. 2.3].

The solution found by Algorithm 1 is a locally optimum solution of problem (SP) [16], [19, Lemma 1]. For a detailed discussion of the convergence properties of SQPs, see [16, Chap. 3,4], [17, Sec. 5.6].

V. Simulation Results

In this section, we evaluate the performance of CF-SUDAS for the given resource allocation algorithm using computer simulations.

As simulation setup, we consider an outdoor BS with N transmit antennas, M indoor CF-SUDAS relays, and K indoor UEs. The outdoor-to-indoor channel is simulated using the QuaDRiGa framework [20] with non-line-of-sight (NLOS) paths. For the indoor channel, the mmWave 60 GHz channel model from [21] with the 'conference room' setting is used. The relays are uniformly distributed on the ceiling of the room, and the UEs are located randomly within the indoor environment. The achievable rates on the relay-UE links are calculated using single-user waterfilling on the indoor mmWave channel. The system parameters are chosen to conform to the 16x16 Long Term Evolution (LTE) Advanced Pro and Wireless Gigabit Alliance (WiGiG) standards for the BS-relay and relay-UE links, respectively. The detailed simulation parameters are given in Table I.

TABLE I
Table of simulation parameters.

BS-Relay, Sub 6 GHz	
N	16
Antenna Gain	18 dBi
N_F	1200
Subcarrier bandwidth	15 kHz
Bandwidth	18 MHz
M	16
K	2
BS-room distance	50 m
Relay-UE, 60 GHz	
Number of subcarriers	336
Subcarrier bandwidth	5.15625 MHz
Effective bandwidth	1.7325 GHz
Max. transmit power	23 dBm
Conference room size	10 m x 10 m

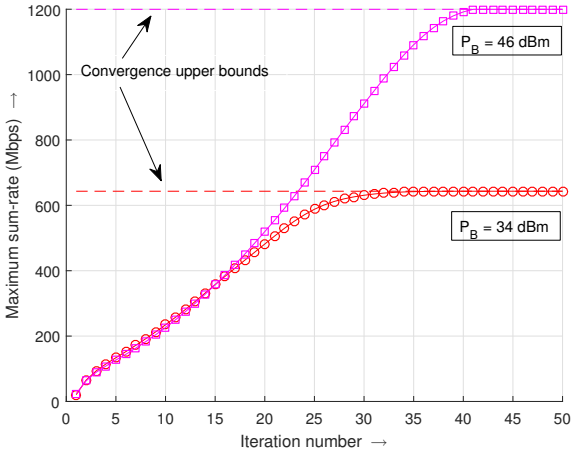

Fig. 3. Convergence of the proposed resource allocation algorithm (Algorithm 1) as a function of the number of iterations.

Figure 3 shows the convergence behaviour of Algorithm 1. At low-to-moderate transmit powers, e.g., P_B = 34 dBm, the relay-UE rate constraint is not active, which has a positive effect on the speed of convergence. The convergence is slower at higher transmit powers, e.g., P_B = 46 dBm, when constraint C7 is active. Nevertheless, in both cases, we observe that the algorithm converges within 50 iterations to the maximum sum rate value. Therefore, in the following, we set the number of iterations to 50.

Figure 4 shows the maximum achievable sum rate as a function of the BS transmit power. The curve labelled 'Upper Bound' is the *max-flow-min-cut* performance upper bound [22,

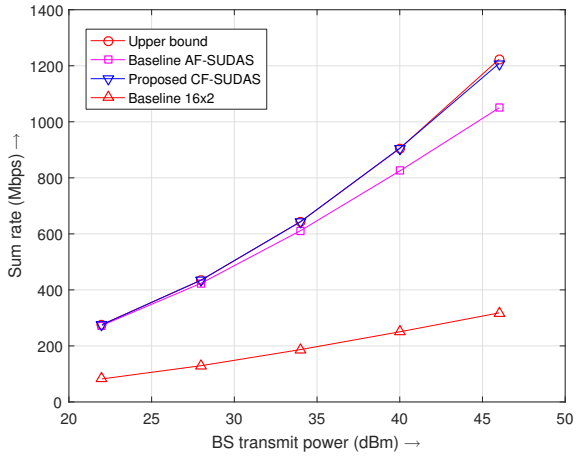

Fig. 4. Performances of CF-SUDAS with the proposed resource allocation algorithm and the baseline AF-SUDAS scheme with joint processing [9].

Theorem 15.10.1] for the SUDAS system. The curve labelled 'Baseline 16x2' is obtained for UEs equipped with two sub-6 GHz antennas and without SUDAS. We observe that the proposed algorithm closely approaches the performance upper bound. Furthermore, the proposed algorithm performs better than the baseline AF-SUDAS with joint relay processing [9] even though the proposed algorithm uses independently operating relays which has significant advantages in terms of complexity and signalling overhead. The improved performance of the proposed scheme is due to the exploitation of the entire indoor capacity for relaying. In comparison, the baseline AF-SUDAS scheme can use only a small amount of the available mmWave bandwidth. We note that the performance of the baseline AF-SUDAS scheme could be improved by adding more relays, thereby increasing the number of relayed MIMO signals. However, this imposes an even larger complexity on the UEs for MIMO decoding. The baseline scheme without SUDAS has a poor performance despite the two receive antennas at the UE as it does not benefit from the virtual MIMO gain.

VI. Conclusion

In this paper, we considered resource allocation for an outdoor-to-indoor SUDAS employing multiple independently operating CF relays which utilize both licensed and unlicensed frequency bands to enhance indoor data throughput. We formulated a non-convex matrix-valued resource allocation problem for maximization of the weighted sum rate. Next, we simplified the non-convex problem to provide a low-complexity suboptimal resource allocation algorithm based on SQP. Computer simulations revealed that the algorithm has good convergence properties. Furthermore, our simulations also confirmed that the proposed algorithm outperforms the baseline AF-SUDAS scheme from [9] as it can take full advantage of the high achievable rates of the indoor relay-UE links.

References

[1] "Ericsson mobility report," Nov. 2013.

[2] M. Dohler, "Virtual antenna arrays," Ph.D. dissertation, King's College London, University of London, Nov. 2003.

[3] A. A. M. Saleh, A. Rustako, and R. Roman, "Distributed antennas for indoor radio communications," *IEEE Trans. Commun.*, vol. 35, no. 12, pp. 1245–1251, Dec. 1987.

[4] R. Heath, S. Peters, Y. Wang, and J. Zhang, "A current perspective on distributed antenna systems for the downlink of cellular systems," *IEEE Commun. Mag.*, vol. 51, no. 4, pp. 161–167, Apr. 2013.

[5] H. Zhu and J. Wang, "Resource allocation in OFDMA-based distributed antenna systems," in *Proc. IEEE/CIC Int. Conf. Communications in China (ICCC)*, Aug. 2013, pp. 565–570.

[6] S. Schwarz, R. W. Heath, and M. Rupp, "Multiuser MIMO in distributed antenna systems with limited feedback," in *IEEE Globecom Workshops*, Dec. 2012, pp. 546–551.

[7] S. H. Park, O. Simeone, O. Sahin, and S. Shamai, "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.

[8] R. G. Stephen and R. Zhang, "Joint millimeter-wave fronthaul and OFDMA resource allocation in ultra-dense CRAN," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1411–1423, Mar. 2017.

[9] D. W. K. Ng, M. Breiling, C. Rohde, F. Burkhardt, and R. Schober, "Energy-efficient 5G outdoor-to-indoor communication: SUDAS over licensed and unlicensed spectrum," *IEEE Trans. Wireless Commun.*, vol. 15, pp. 3170 – 3186, May 2016.

[10] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, 1976.

[11] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[12] X. Wang and G. B. Giannakis, "Resource allocation for wireless multiuser OFDM networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4359–4372, 2011.

[13] D. P. Palomar and Y. Jiang, *MIMO Transceiver Design Via Majorization Theory*. Now Publishers, 2007.

[14] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and its Applications*, 2nd ed. Springer, 2011.

[15] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.

[16] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," *Acta Numerica*, vol. 4, pp. 1–51, 001 1995.

[17] M. J. Goldsmith, "Sequential quadratic programming methods based on indefinite Hessian approximations," Ph.D. dissertation, Stanford, CA, USA, 1999.

[18] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[19] Q. T. Dinh and M. Diehl, "Local convergence of sequential convex programming for nonconvex optimization," in *Recent Advances in Optimization and its Applications in Engineering*. Springer, 2010, pp. 93–102.

[20] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "Quadriga: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, Jun. 2014.

[21] A. Maltsev, R. Maslennikov, A. Sevastyanov, A. Lomayev, and A. Khoryaev, "Statistical channel model for 60 GHz WLAN systems in conference room environment," in *Proc. Fourth European Conf. Antennas and Propagation*, Apr. 2010, pp. 1–5.

[22] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.