GRADIENT-BASED SOLUTION FOR HYBRID PRECODING IN MIMO SYSTEMS

Negin Bakhshi Zanjani, Seyran Khademi, Geert Leus

Faculty of Electrical Engineering, Mathematics and Computer Science, TU Delft, The Netherlands
E-mails: {n.bakhshizanjani, s.khademi, g.j.t.leus}@tudelft.nl.

ABSTRACT

The combination of baseband and analog precoding for multiple-input multiple-output (MIMO) systems is considered in this paper which is referred to as hybrid precoding. The system capacity, as a design criterion, is maximized subject to unit modulus constraints on the elements of the analog precoder (phase shifters), and a total power constraint. This is a non-convex problem due to the product of the analog and baseband precoder variables. The proposed technique suggests computing non-trivial complex derivatives of the objective and constraints, analytically, in order to develop an iterative gradient-based sequential optimization algorithm to solve the non-convex problem. Promising simulation results show that the solution of the proposed algorithm is sufficiently close to the optimal (full-baseband) precoder solution, regardless of the channel characteristics.

Index Terms— multiple-input multiple-output (MIMO), hybrid precoding, nonlinear optimization, sequential quadratic programming.

1. INTRODUCTION

Modern multiple-input multiple-output (MIMO) communication networks aim at reducing the number of radio frequency (RF) chains due to energy consumption and high fabrication cost of the RF components. Hybrid precoding which is also referred to as soft antenna selection is an attractive approach to balance the precoding performance and the underlying hardware complexity. Hybrid processing is initially introduced in [1] by using a network of analog phase shifters in the RF domain and a low dimensional baseband (BB) precoder for receive beamforming. Similarly, in hybrid precoding the transmitter communicates by passing the multiplexed data through a BB precoder $\mathbf{P}_{\text{BB}} \in \mathbb{C}^{L_t \times N_s}$, followed by an RF precoder $\mathbf{P}_{\text{RF}} \in \mathbb{C}^{N_t \times L_t}$, where N_t and N_s denote the number of transmit antennas and streams, respectively, and L_t is the number of available transmit RF chains. The formulation of the precoding problem leads to a non-convex objective function for the product of a BB and RF precoder and a non-convex feasible set due to the unit modulus constraints on the RF precoder elements. A common approach in the literature is to perform alternate programming by solving for one parameter at a time in order to overcome the nonlinearity issue, and further restraining the hybrid processor to a unitary matrix for simplification of the objective.

In [1], unitary hybrid beamforming is considered separately for multiplexing (capacity maximization) and diversity combining (SNR maximization), to match the eigen directions of the channel matrix. The matrix of the phase shifters is chosen as the conjugate phase of the L_r (number of available receive RF chains) eigenvectors corresponding to the L_r largest eigenvalues of the channel gain matrix. Given the RF precoder, the optimization problem is solved only with respect to the BB precoder. In [2], the sum-rate capacity maximization is considered with a hybrid structure at both the transmitter and receiver for a large multi-user MIMO system. Exploiting the properties of large MIMO channels, the BB precoder is determined, and then the RF precoder is found by an iterative column update approach. Phased zero forcing is considered in [3], where the RF precoder is designed to align the phase of the channel matrix and then the equivalent low dimensional real channel matrix is used to solve for the BB precoder which serves as a power allocator.

Hybrid unitary processing for both the transmitter and receiver is specialized for millimeter wave systems in [4–12]. In [4], a low-complexity spatially sparse hybrid precoder is proposed using the concept of orthogonal matching pursuit (OMP) where the precoder vectors are chosen iteratively as a linear combination of the steering vectors for a known array geometry. In [5], an adaptive algorithm is proposed to estimate the millimeter wave complex channel gains and then the steering vectors are chosen based on the quantized beamsteering directions of the estimated channel using the same OMP algorithm. In [6], a similar technique is used, however instead of correlation matching over a dictionary to find each column of the RF precoder, an element-wise normalization of the first singular vector of the error matrix (between the optimal precoder and hybrid precoder) is used at each iteration. Reduced complexity codebook-based precoding algorithms are proposed in [7]. In [8–11], sparse approximation problems are formulated to design hybrid precoders. An alternating minimization algorithm is considered in [12] in which digital and analog precoders are separately calculated per iteration.

This paper is distinguished from the existing literature mainly because it offers a generic approach for solving the non-convex hybrid precoding problem. This is summarized below as

1. There are no prior assumptions on the channel statistical properties or matrix structure such as sparsity or rank constraints, which makes the proposed algorithm nonexclusive and widely applicable.
2. The mutual information can be maximized directly instead of minimizing the Euclidian distance between the optimal and the hybrid precoder (error matrix), hence the full baseband precoder ($\mathbf{P}_{\text{opt}}$) is not required, a priori.
3. The proposed algorithm, based on the analytical gradient derivations, leverages a simple function evaluation in iterative updates, making the algorithm scalable and extensions to alternative precoding schemes can be readily envisioned.
4. Both the RF and BB precoder are updated simultaneously in the algorithm thus no alternate optimization is used.

The notation is defined as follows: bold upper case and bold lower case symbols indicate matrices and vectors, respectively. The conjugate transpose, conjugate, transpose and the inverse of a matrix $\mathbf{A}$ are denoted as $\mathbf{A}^H$, $\mathbf{A}^*$, $\mathbf{A}^T$ and $\mathbf{A}^{-1}$. $\mathbf{I}_N$ denotes an identity matrix of size N and $\mathbf{0}_{MN}$ is the zero matrix of size $M \times N$. $|\mathbf{A}(i,j)|$ denotes the absolute value of the (i,j)th element of $\mathbf{A}$ and $\|\mathbf{A}\|_F$ is the Frobenius norm of $\mathbf{A}$. The operation diag$(\mathbf{a})$ forms a diagonal matrix with the entries of $\mathbf{a}$ on its diagonal, and vec$(\mathbf{A})$ (vec$^T(\mathbf{A})$) lists the columns of $\mathbf{A}$ in a column (row) vector. The Kronecker product is denoted by $\otimes$.

2. PROBLEM FORMULATION

2.1. Signal Model

Consider a multiplexing MIMO system where N_t and N_r denote the number of transmit and receive antennas, respectively. The discrete-time transmit vector, at each time instance, is given by

$$\mathbf{x} = \mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}\mathbf{s}, \quad (1)$$

where $\mathbf{P}_{\text{RF}} \in \mathbb{C}^{N_t \times L_t}$ and $\mathbf{P}_{\text{BB}} \in \mathbb{C}^{L_t \times N_s}$ are the RF and BB precoders, respectively. The number of transmit streams can be $N_s \leq L_t$. we consider maximum rate so $N_s = L_t$. The transmit data sequence is a vector $\mathbf{s} \in \mathbb{C}^{N_s \times 1}$ with a scaled identity covariance matrix $E[\mathbf{s}\mathbf{s}^H] = \frac{1}{N_s}\mathbf{I}_{N_s}$. The elements of the RF precoder are constant modulus so, $\mathbf{P}_{\text{RF}}$ can be implemented using analog phase shifters. In turn, the received noisy signal is given by

$$\mathbf{y} = \sqrt{p}\,\mathbf{H}\,\mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}\,\mathbf{s} + \mathbf{n}, \quad (2)$$

where p is the average received power, the complex channel matrix is denoted by $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$, and $\mathbf{n}$ is assumed to be a spatially white, zero mean Gaussian noise vector with covariance matrix $\sigma_n^2\mathbf{I}_{N_r}$. We also assume that the instantaneous channel matrix is known at the transmitter so $\mathbf{H}$ is given. The product of $\mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}$ is referred to as the full linear precoder. We introduce two precoding problems in the following of which the solutions are discussed in the next section.

2.2. Unitary Precoding

The optimal joint precoder and receiver for a point to point MIMO system is known to decompose the interference MIMO channel to $R = \text{rank}\{\mathbf{H}\}$ independent single-input single-output (SISO) channels. Focusing on the precoding problem, the full unitary precoding matrix includes the right eigen vectors of the channel. By taking the singular value decomposition of the channel, $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$, we introduce $\mathbf{\Sigma}_1 \in \mathbb{C}^{L_t \times L_t}$ and $\mathbf{V}_1 \in \mathbb{C}^{N_t \times L_t}$ such that

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_2 \end{pmatrix}, \mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{pmatrix}. \quad (3)$$

The full unitary precoder in this setting is $\mathbf{V}_1$, (excluding the diagonal power allocation matrix). Once we know the full precoder, the hybrid precoder is the solution of the following problem

$$\mathcal{P}_1: \quad \begin{array}{ll} \underset{\mathbf{P}_{\text{RF}}, \mathbf{P}_{\text{BB}}}{\text{minimize}} & \|\mathbf{P}_{opt} - \mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}\|_F^2 \\ \text{subject to} & |\mathbf{P}_{\text{RF}}(i,j)| = 1;\ \forall i,j \\ & i = 1,...,N_t,\ j = 1,...,L_t \\ & \|\mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}\|_F^2 = L_t \end{array}$$

where $\mathbf{P}_{opt}$ is $\mathbf{V}_1$ for unitary precoding. The first (constant modulus) constraint concerns the analog phase shifters and the last constraint is imposed to limit the total transmit power.

2.3. Non-Unitary Precoding

For a MIMO system with a linear precoding matrix $\mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}$ at the transmitter, the mutual information between the two ends of the system using Gaussian signaling is given by

$$I(\mathbf{x};\mathbf{y}) = \log\left(\det\left[\mathbf{I}_{N_r} + \frac{p}{L_t\sigma_n^2}\mathbf{H}\mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}\mathbf{P}_{\text{BB}}^H\mathbf{P}_{\text{RF}}^H\mathbf{H}^H\right]\right),$$

in unit of bits per second per Hz when the basis of the logarithm is two. The full precoder is designed to maximize the mutual information between $\mathbf{x}$ and $\mathbf{y}$ which gives the capacity of the system [13]. The corresponding hybrid precoding optimization problem is then expressed as

$$\mathcal{P}_2: \quad \begin{array}{ll} \underset{\mathbf{P}_{\text{RF}}, \mathbf{P}_{\text{BB}}}{\text{maximize}} & I(\mathbf{x};\mathbf{y}) \\ \text{subject to} & |\mathbf{P}_{\text{RF}}(i,j)| = 1;\ \forall i,j \\ & i = 1,...,N_t,\ j = 1,...,L_t \\ & \|\mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}}\|_F^2 = L_t \end{array}$$

Both optimization problems of $\mathcal{P}_1$ and $\mathcal{P}_2$ are non-convex for two reasons: 1) The objectives are non-convex functions of the optimization variables (product of $\mathbf{P}_{\text{RF}}$ and $\mathbf{P}_{\text{BB}}$). 2) The feasible set is not convex due to the nonlinear equality constraints. The proposed algorithm to solve these two non-convex problems is given next.

3. PROPOSED APPROACH

The proposed solution for the non-convex problems of $\mathcal{P}_1$ and $\mathcal{P}_2$ is based on gradient descent minimization which requires at least the first-order differentiability of the objective and constraints. First, we use a simple variable change, i.e., $\mathbf{P}_{\text{RF}} = \exp(j\mathbf{\Phi})$,which is an element-wise exponential function, to lift the first constraint and simplify the problem. Accordingly, $\mathcal{P}_1$ and $\mathcal{P}_2$ can be written in terms of two variables $\{\mathbf{\Phi} \in \mathbb{R}^{N_t \times L_t}, \mathbf{P}_{\text{BB}} \in \mathbb{C}^{L_t \times L_t}\}$, so the constant modulus constraints are dropped. Further, we invoke the augmented complex-valued derivation (ACVD) technique [14] to obtain the gradient of the objective and constraint with respect to the variables. This is the key step to approach the nonlinear optimization problem. The elegance of the ACVD technique is evident here since the common variable decomposition to real and imaginary components yields an extremely tedious and impractical representation of the problem, while the ACVD preserves the compact formulation.

3.1. Complex-Valued Gradient

The crucial step towards the calculation of the gradient of the objective and the constraint is to write them as a function of the optimization variables and their conjugates, i.e., $\{\mathbf{P}_{\text{BB}}, \mathbf{P}^*_{\text{BB}}, \mathbf{\Phi}\}$. For this purpose, a new variable is introduced as $\mathbf{Z} := \mathbf{P}_{\text{RF}}\mathbf{P}_{\text{BB}} = \exp(j\mathbf{\Phi})\mathbf{P}_{\text{BB}}$, accordingly

$$\begin{aligned} \mathbf{z} \in \mathbb{C}^{N_t L_t \times 1} &= \text{vec}\left(\mathbf{P}_{\text{RF}}\, \mathbf{P}_{\text{BB}}\right) = \text{vec}\left(\exp(j\mathbf{\Phi})\mathbf{P}_{\text{BB}}\right) \\ &= \left(\, \mathbf{I}_{L_t} \otimes \exp(j\mathbf{\Phi})\,\right) \underbrace{\text{vec}\left(\mathbf{P}_{\text{BB}}\right)}_{\mathbf{p}_{\text{BB}}} \\ &= \underbrace{\left(\, {\mathbf{P}_{\text{BB}}}^T \otimes \mathbf{I}_{N_t}\,\right)}_{\mathbf{I}_{\text{BB}}} \; \underbrace{\text{vec}\left(\exp(j\mathbf{\Phi})\right)}_{\exp(j\text{vec}\,(\mathbf{\Phi}))=\exp(j\boldsymbol{\phi})} \end{aligned} \quad (4)$$

where the optimal precoder is vectorized as $\mathbf{p}_{\text{opt}} = \text{vec}(\mathbf{P}_{\text{opt}})$.

Let $f, g \;:\; \{\mathbf{Z} \in \mathbb{C}^{N_t \times L_t}, \mathbf{Z}^* \in \mathbb{C}^{N_t \times L_t}\} \to \mathbb{R}$ be real-valued and differentiable mappings of the objective and the constraint, respectively. We follow the same notation and formulation for ACVD as in [14] to calculate the analytical expression for the gradient. The gradient vector of the objective with respect to two optimization variables $\nabla f(\mathbf{p}_{\text{BB}}, \mathbf{p}^*_{\text{BB}}, \boldsymbol{\phi}) \in \mathbb{R}^{(2L_t L_t + L_t N_t)\times 1}$ is given by

$$\nabla f(\mathbf{p}_{\text{BB}}, \mathbf{p}^*_{\text{BB}}, \boldsymbol{\phi})^T = \begin{pmatrix} \frac{\partial f}{\partial \mathbf{z}} & \frac{\partial f}{\partial \mathbf{z}^*} \end{pmatrix} \begin{pmatrix} \frac{\partial \mathbf{z}}{\partial \mathbf{p}_{BB}} & \frac{\partial \mathbf{z}}{\partial \mathbf{p}^*_{BB}} & \frac{\partial \mathbf{z}}{\partial \boldsymbol{\phi}} \\ \frac{\partial \mathbf{z}^*}{\partial \mathbf{p}_{BB}} & \frac{\partial \mathbf{z}^*}{\partial \mathbf{p}^*_{BB}} & \frac{\partial \mathbf{z}^*}{\partial \boldsymbol{\phi}} \end{pmatrix}, \quad (5)$$

using the chain rule properties. The gradient of the constraint (∇g) is derived in a similar way, by substituting f with g in the first term of (5).

The equivalent formulation of $\mathcal{P}_1$ in terms of $\mathbf{z}$ is given by

$$\mathcal{P}_1: \quad \begin{aligned} &\underset{\mathbf{z} = \text{vec}(\mathbf{Z})}{\text{minimize}} && f_1(\mathbf{Z}, \mathbf{Z}^*) = (\mathbf{p}_{\text{opt}} - \mathbf{z})^H(\mathbf{p}_{\text{opt}} - \mathbf{z}) \\ &\text{s.t.} && g(\mathbf{Z}, \mathbf{Z}^*) = \mathbf{z}^H\mathbf{z} - L_t = 0 \end{aligned}$$

In turn, $\mathcal{P}_2$ can be reformulated as

$$\mathcal{P}_2: \quad \begin{aligned} &\underset{\mathbf{z} = \text{vec}(\mathbf{Z})}{\text{minimize}} && f_2(\mathbf{Z}, \mathbf{Z}^*) = \log\left(\det\left[\mathbf{I}_{N_r} + \tfrac{p}{L_t\sigma_n^2}\mathbf{H}\mathbf{Z}\mathbf{Z}^H\mathbf{H}^H\,\right]\right) \\ &\text{s.t.} && g(\mathbf{Z}, \mathbf{Z}^*) = \mathbf{z}^H\mathbf{z} - L_t = 0 \end{aligned}$$

The gradient of the objective functions and the constraint for both problems $\mathcal{P}_1$ and $\mathcal{P}_2$ can be explicitly derived using (5). Note that the first term in (5) differs for $\mathcal{P}_1$ and $\mathcal{P}_2$ while the second term is common and is given by

$$\begin{pmatrix} \mathbf{I}_{L_t} \otimes \exp(j\mathbf{\Phi}) & \mathbf{0}_{N_t L_t \times L_t L_t} & j\mathbf{I}_{\text{BB}}\, \text{diag}\left(\exp(j\boldsymbol{\phi})\right) \\ \mathbf{0}_{N_t L_t \times L_t L_t} & \mathbf{I}_{L_t} \otimes \exp(-j\mathbf{\Phi}) & -j\mathbf{I}_{\text{BB}}\, \text{diag}\left(\exp(-j\boldsymbol{\phi})\right) \end{pmatrix}.$$

For f_1 in $\mathcal{P}_1$, the first term is readily calculated as

$$\begin{pmatrix} \frac{\partial f_1}{\partial \mathbf{z}} & \frac{\partial f_1}{\partial \mathbf{z}^*} \end{pmatrix} = \begin{pmatrix} \mathbf{p}^H_{\text{opt}} - \mathbf{z}^H & \mathbf{z}^T - \mathbf{p}^T_{\text{opt}} \end{pmatrix}. \quad (6)$$

For f_2 in $\mathcal{P}_2$, taking $\mathbf{R} = (\mathbf{I}_{N_r} + \frac{p}{L_t\sigma_n^2}\mathbf{H}\mathbf{Z}\mathbf{Z}^H\mathbf{H}^H)$ yields

$$\begin{pmatrix} \frac{\partial f_2}{\partial \mathbf{z}} & \frac{\partial f_2}{\partial \mathbf{z}^*} \end{pmatrix} = \left(\text{vec}^T\left(\tfrac{p}{L_t\sigma_n^2}\mathbf{H}^T\mathbf{R}^{-T}\mathbf{H}^*\mathbf{Z}^*\right) \quad \text{vec}^T\left(\tfrac{p}{L_t\sigma_n^2}\mathbf{H}^H\mathbf{R}^{-1}\mathbf{H}\mathbf{Z}\right) \right).$$

Using the analytical gradient vectors that are derived in this section, we can take the next step to find a proper optimization algorithm that provides the required accuracy. Note that problems $\mathcal{P}_1$ and $\mathcal{P}_2$ are non-convex with respect to the original variables $(\mathbf{P}_{\text{BB}}, \mathbf{\Phi})$ so the optimal global solution can not be found, theoretically. Therefore, to evaluate the performance of the proposed algorithms we use statistical measures, in Section 4, which verify the accuracy (convergence) of the solution in an average sense.

3.2. Optimization Algorithm

Knowing the gradient and the Hessian of the objective and the equality constraint, problems $\mathcal{P}_1$ and $\mathcal{P}_2$ can be solved using the Lagrange multiplier (LM) method. The Hessian matrix is the Jacobian of the gradient vector. However, the explicit evaluation of the Hessian matrix and its inverse are commonly computationally inefficient and sometimes infeasible due to the singularity of the matrix. Here the Hessian, $\nabla^2 f(\mathbf{p}_{\text{BB}k}, \mathbf{p}^*_{\text{BB}k}, \phi_k)$, is found to be singular because the problem is ill-conditioned due to the exponential relaxation of $\mathbf{P}_{\text{RF}}$. Hence, the inverse of the Hessian matrix is approximated using the Broyden Fletcher Goldfarb Shanno (BFGS) updating method [15].

The main competence of LM method is to transform a constrained optimization problem to an unconstrained root-finding problem. The roots of the gradient of the Lagrangian function associated with the problem can be found using Newton like methods [16]. The simulation results suggest that the performance of LM is not satisfying for the considered hybrid precoding problems.

The sequential quadratic programming (SQP) algorithm [17] is used as an alternative robust technique to solve the non-convex hybrid precoding problems $\mathcal{P}_1$ and $\mathcal{P}_2$. SQP solves a sequence of quadratic programming (QP) optimization subproblems, i.e., by considering the second-order and the first order Taylor expansion of the objective and constraint functions of the original problem, respectively, at each iteration. As a result, the objective function for QP is quadratic and the constraints are linear, so a convex problem with an standard iterative solution. To formulate a real cost and constraint functions for the QP problem, a concatenation of two optimization variables is represented by $\mathbf{p}_k \in \mathbb{R}^{(2L_tL_t+L_tN_t)\times 1} = [\Re\{\mathbf{p}_{\mathrm{BB}k}^T\}, \Im\{\mathbf{p}_{\mathrm{BB}k}^T\}, \phi_k^T]^T$. The gradients with respect to $(\mathbf{p}_{\mathrm{BB}}, \mathbf{p}_{\mathrm{BB}}^*, \phi)$, that are derived in Sec. 3.1, needs to be translated to the ones with respect to the real and imaginary parts of the variables [14]. Accordingly, $\nabla f(\mathbf{p}_k) = (\frac{\partial f}{\partial \Re\{\mathbf{p}_{\mathrm{BB}}\}}, \frac{\partial f}{\partial \Im\{\mathbf{p}_{\mathrm{BB}}\}}, \frac{\partial f}{\partial \phi})$ is calculated using

$$\frac{\partial f}{\partial \Re\{\mathbf{p}_{\mathrm{BB}}\}} = \frac{1}{2}(\frac{\partial f}{\partial \mathbf{p}_{\mathrm{BB}}} + \frac{\partial f}{\partial \mathbf{p}_{\mathrm{BB}}^*}), \quad (7)$$

$$\frac{\partial f}{\partial \Im\{\mathbf{p}_{\mathrm{BB}}\}} = \frac{1}{2j}(\frac{\partial f}{\partial \mathbf{p}_{\mathrm{BB}}} - \frac{\partial f}{\partial \mathbf{p}_{\mathrm{BB}}^*}). \quad (8)$$

The QP optimization problem of the hybrid precoding problem at kth iteration is given by

$$\begin{aligned} \underset{\mathbf{p}_k}{\text{minimize}} \quad & \tfrac{1}{2}\mathbf{p}_k^T \nabla^2 f(\mathbf{p}_k)\, \mathbf{p}_k + \nabla f(\mathbf{p}_k)^T \mathbf{p}_k \\ \text{s.t.} \quad & \nabla g(\mathbf{p}_k)^T \mathbf{p}_k + g(\mathbf{p}_k) = 0 \end{aligned}$$

where the Hessian is approximated at each iteration. The QP subproblem is solved at each iteration using a Matlab built-in function (quadprog) and the solution is used to construct the next approximation of the problem, until the solution converges to the optimum.

4. SIMULATION RESULTS

In this section, we present simulation results to evaluate the performance of different algorithms. Each algorithm is evaluated using 1000 random channel realizations to evaluate the performance statistically. Next, a complementary cumulative density function (CCDF) is derived to show the probability that the maximum mutual information (capacity) is greater than a specified value ($C > C_0$). To model the propagation environment, we consider the geometric channel model with L paths from the transmitter to the receiver from [5]. The simulations compare the proposed algorithm to the optimal full linear precoder which is obtained using CVX [18] for $\mathcal{P}_2$.

In Fig. 1, problems $\mathcal{P}_1$ and $\mathcal{P}_2$ are solved using SQP for 8×8 and a 128×64 MIMO system using $N_s = L_t = 3$. The results are compared with the full unitary precoder ($\mathbf{V}_1$ in (3)) for two different numbers of channel paths ($L = 3, 100$). The higher the L, the richer the propagation environment which deviates from the sparse channel assumption in millimeter-wave MIMO systems. The parameters of the model in [5] are selected as $M = 2$ and $N = 128$ where M and N are respectively the number of beamforming vectors per stage and the desired resolution parameter. The most left CCDF curve (worse) in Fig. 1(a) (8×8 MIMO) corresponds to the capacity of the unitary precoder ($\mathcal{P}_1$) using the LM algorithm. The SQP solution for $\mathcal{P}_1$ coincides with the optimal solution and

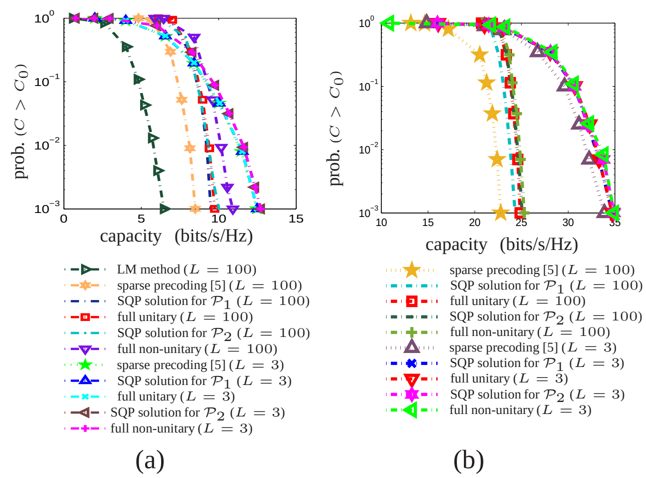

Fig. 1: CCDF of capacity for $\mathcal{P}_1$ and $\mathcal{P}_2$ using $L = 3$ and $L = 100$ received paths (a) 8×8 MIMO (b) 128×64 MIMO.

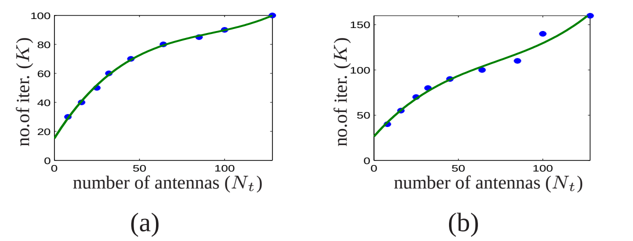

Fig. 2: Number of iterations in average for SQP, versus number of transmit antennas for (a) $\mathcal{P}_1$ (b) $\mathcal{P}_2$. Solid line and dots represent the fitted curve and experimental data, respectively.

is close to optimal for $\mathcal{P}_2$, while the sparse precoder of [5] underperforms in a non-sparse channel ($L = 100$), as expected. For a sparse channel ($L = 3$) both the SQP and sparse precoder reach the optimum. In Fig. 1(b), the results for 128×64 MIMO are presented, where the optimal full precoder is known using water-filling algorithm [19]. Similarly, the SQP algorithm outperforms the sparse precoder for a large MIMO setting, however, it is more complex compared to the sparse precoder. Note, the sparse precoder considers only unitary precoding ($\mathcal{P}_1$) while the proposed approach can be used to design other precoding problems such as $\mathcal{P}_2$.

The complexity of the SQP algorithm is $K \times O(QP)$, where K is the number of SQP iterations. The complexity of the QP problem is almost scale-independent, while K is changing by the size of the MIMO channel. In Fig. 2 the average number of iterations is plotted versus the number of transmit antennas for $\mathcal{P}_1$ and $\mathcal{P}_2$, which shows a cubic growth with respect to N_t.

To conclude, an effective optimization scheme is proposed in this paper for the hybrid precoding problem in MIMO systems. The SQP algorithm shows a satisfying performance compared to the sparse precoder of [5] and LM method, while the complexity is higher. Note that the proposed approach overreaches many types of precoding while the available literature concerns unitary precoding.

5. REFERENCES

[1] Xinying Zhang, A. F. Molisch, and Sun-Yuan Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, pp. 4091–4103, Nov 2005.

[2] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale MIMO systems," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015, pp. 2929–2933.

[3] L. Liang, W. Xu, and X. Dong, "Low-complexity hybrid precoding in massive multiuser MIMO systems," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 653–656, Dec 2014.

[4] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, March 2014.

[5] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, Oct 2014.

[6] R. Mendez-Rial, C. Rusu, N. Gonzalez-Prelcic, and R. W. Heath, "Dictionary-free hybrid precoders and combiners for mmWave MIMO systems," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2015, pp. 151–155.

[7] J. Singh and S. Ramakrishna, "On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 2670–2683, May 2015.

[8] M. S. Rahman and K. Josiam, "Low complexity RF beam search algorithms for millimeter-wave systems," in *2014 IEEE Global Communications Conference*, Dec 2014, pp. 3815–3820.

[9] C. Rusu, R. Mendez-Rial, N. Gonzalez-Prelcicy, and R. W. Heath, "Low complexity hybrid sparse precoding and combining in millimeter wave MIMO systems," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 1340–1345.

[10] M. Kim and Y. H. Lee, "MSE-based hybrid RF/baseband processing for millimeter-wave communication systems in MIMO interference channels," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 6, pp. 2714–2720, June 2015.

[11] H. Ghauch, T. Kim, M. Bengtsson, and M. Skoglund, "Subspace estimation and decomposition for large millimeter-wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 528–542, April 2016.

[12] X. Yu, J. C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 485–500, April 2016.

[13] Andrea Goldsmith, *Wireless Communications*, Cambridge University Press, New York, NY, USA, 2005.

[14] Are Hjorungnes, *Complex-valued matrix derivatives: With applications in signal processing and communications*, Cambridge University Press, 2011.

[15] Mordecai Avriel, *Nonlinear Programming : Analysis and Methods*, Dover Publishing, 2003.

[16] Kris Hauser, "B553 lecture 7: Constrained optimization, Lagrange multipliers, and KKT conditions," 2012.

[17] P. T. Boggs and J. W. Tolle., *Sequential Quadratic Programming*, Acta Numerica, 1995.

[18] Inc. CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0," Aug 2012.

[19] I. Emre Telatar, "Capacity of multi-antenna gaussian channels," *EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS*, vol. 10, pp. 585–595, 1999.