FAST OPTIMIZATION ALGORITHM ON COMPLEX OBLIQUE MANIFOLD FOR HYBRID PRECODING IN MILLIMETER WAVE MIMO SYSTEMS

Hiroyuki Kasai

The University of Electro-Communications, Tokyo, Japan

ABSTRACT

This paper considers a hybrid precoder design for millimeter wave systems. The conventional works suffer from performance loss in spectral efficiency, and also suffer from high complexities of optimization algorithms in terms of iteration as well as processing time. This paper proposes a new fast optimization algorithm that improves the calculation algorithm of the gradient and avoids the nested-loop architecture of the state-of-the-art algorithm, MO-AltMin. Exhaustive evaluations demonstrate that the proposed algorithm yields comparable or faster processing speed than the state-of-the-art algorithms while keeping spectral efficiency near-optimal.

Index Terms— Hybrid precoding, millimeter wave MIMO, manifold optimization, complex oblique manifold

1. INTRODUCTION

A *millimeter wave* (mmWave) system, of which bounds ranges from 30 GHz to 300 GHz, provides high data rates over cellular networks [1, 2, 3, 4, 5]. The short wavelength of the mmWave system allows its transceivers to facilitate a large number of antenna arrays. The mmWave system leverages the multiple-input and multiple-output (MIMO) through such large antenna arrays to provide sufficient received signal power. However, the traditional MIMO systems deploy either fully analog beamforming at the radio frequency (RF) or fully digital precoding at the baseband circuits which control both phase shifts and amplitudes of incoming signals [6]. Subsequently, this makes it hard to devote itself to a separate RF chain[1] with its components for each antenna, resulting in high hardware costs and high power consumptions in large-scale MIMO systems [4, 3, 7, 8]. *Hybrid precoding (-beamforming)*, originally proposed in [9, 10], is the most promising approach to solve this problem, which combines *large*-dimensional *analog* precoding (or beamforming) via *phase shifters* with *lower*-dimensional *digital* baseband precoding. Hybrid precoding drastically reduces the number of RF chains from the number of antenna elements, which is needed in the fully digital scheme, into, at least, the number of data streams [11, 6].

The design of precoder and the decoder of the mmWave system is handled as two sub-problems, i.e., the precoding and decoding problems [11, 12, 13, 14]. The maximization of spectral efficiency approximately boils down to a minimization problem of the Euclidean distance between the *fully digital precoder* $\mathbf{F}_{\rm opt}$ and the hybrid precoder, where $\mathbf{F}_{\rm opt}$ is the unconstrained singular value decomposition (SVD)-based precoding matrix. This problem is further formulated as a *matrix factorization* problem of $\mathbf{F}_{\rm opt}$ with a product of the digital *baseband precoder* matrix $\mathbf{F}_{\rm BB}$ and the analog *RF precoder* (or beamforming) matrix $\mathbf{F}_{\rm RF}$. The noteworthy point is that the phase shifters impose an additional *element-wise unit modulus* constraints on $\mathbf{F}_{\rm RF}$. For this particular problem, addressing the spatial structure of mmWave systems with large antenna arrays, Ayach et al. translated it into a *sparse reconstruction* problem, and solves it by the orthogonal matching pursuit (OMP) algorithm [11, 12]. However, this restriction on the space of feasible analog precoding solutions causes some degradation of spectral efficiency. Yu et al. proposed a superior alternative minimization algorithm, called MO-AltMin, based on a *manifold optimization* technique by embedding the unit modulus constituents on $\mathbf{F}_{\rm RF}$ into a product of multiple circles in the complex plane [14, Section III]. Although MO-AltMin provides near-optimal spectral efficiency, it suffers from not only slow down of the convergence speed of the procedure due to its *nested-loop* architecture between $\mathbf{F}_{\rm BB}$ and $\mathbf{F}_{\rm RF}$, but also an extremely high complexity to handle big matrices due to the Kronecker product. Therefore, Yu et al. proposed, in parallel, an alternative low-complexity algorithm, called PE-AltMin [14, Section IV]. It is, however, hindered by the decrease of spectral efficiency when the number of transmit chains increases, and also hindered by a high complexity originating in the SVD calculation per iteration for a big $\mathbf{F}_{\rm opt}^H\mathbf{F}_{\rm RF}$ matrix. Finally, Zanjani et al. proposed a gradient-based approach to maximize directly the mutual information instead of the approximated Euclidean distance problem, but it requires a high complexity [15].

This paper proposes a new fast optimization algorithm on *complex oblique manifold* in terms of iteration and processing time without rendering the decrease of spectral efficiency. *Manifold optimization* has shown superior perfor-

H. Kasai was partially supported by JSPS KAKENHI Grant Numbers JP16K00031 and JP17H01732.

[1]Up converter, down converter, digital-to-analog converter (DAC), analog-to-digital converter (ADC), mixers, and power amplifiers [6].

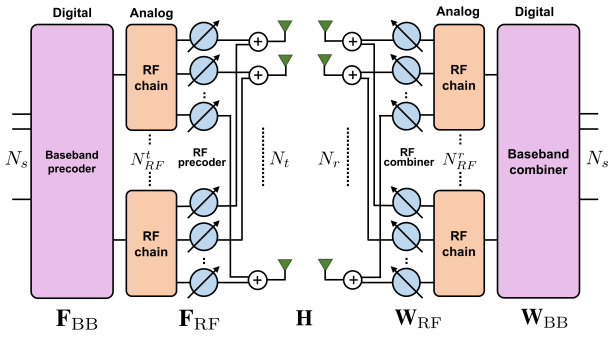


Fig. 1. Hybrid precoder in mmWave MIMO system [11, 14].

mances on many applications that include, to name a few, the low-rank matrix completion and tensor completion problems on the manifold of fixed-rank matrices and tensors [16, 17], and the independent component analysis (ICA) is another example defined on the *oblique* manifold [18]. Although following the same line of the strategy of MO-AltMin, the calculation algorithm of the gradient avoids big matrix calculations originating in the Kronecker product. Furthermore, a *single-loop* architecture of the alternating strategy between $\mathbf{F}_{\mathrm{BB}}$ and $\mathbf{F}_{\mathrm{RF}}$ is also proposed to alleviate the slow convergence caused by the nested-loop architecture in MO-AltMin. These techniques enable the proposed algorithm to achieve comparable or faster processing speed than not only MO-AltMin but also PE-AltMin while keeping spectral efficiency higher than the state-of-the-art algorithms. The codes of the proposed algorithm are implemented in the MATLAB toolbox Manopt [19] and are available at `https://github.com/hiroyuki-kasai/HybridPrecodingOpt`.

2. PRELIMINARIES

2.1. Notations

The scalars are represented with lower-case letters $(a, b, \ldots)$, vectors are with bold lower-case letters $(\boldsymbol{a}, \boldsymbol{b}, \ldots)$, and matrices are with bold-face capitals $(\mathbf{A}, \mathbf{B}, \ldots)$. $(\mathbf{A})_{i,j}$ represents the element of $\mathbf{A}$ at (i, j). $\mathbf{A}^T$. $\mathbf{A}^*$ and $\mathbf{A}^H$ are designated for the transpose, conjugate, and conjugate transpose of $\mathbf{A}$, respectively. The Moore-Penrose pseudo inverse of $\mathbf{A}$ is denoted as $\mathbf{A}^\dagger$. $\mathrm{vec}(\mathbf{A})$ and $\mathrm{vec}^{-1}(\mathbf{A})$ are vectorization and inverse-vectorization operators, respectively. $\mathbb{E}[\cdot]$ represents the expectation, and $\mathrm{Real}[\cdot]$ extracts the real part of a complex variable. $\|\cdot\|_F$ denotes the Frobenius norm. $\otimes$ and $\circ$ denote the Kronecker and Hadamard products of two matrices.

2.2. System model

A single-user (point-to-point) multi-stream mmWave MIMO system is shown in Figure 1, in which a transmitter equipped with N_t antennas and N_{RF}^t RF chains and a receiver with N_r antennas and N_{RF}^r RF chains communicate via N_s streams. N_s is constrained to be bounded as $N_s \leq N_{RF}^t \leq N_t$ and $N_s \leq N_{RF}^r \leq N_r$. This hardware architecture enables the transmitter to apply $N_{RF}^t \times N_s$ digital baseband precoder $\mathbf{F}_{\mathrm{BB}} \in \mathbb{R}^{N_{RF}^t \times N_s}$ using its N_{RF}^t transmit chains, followed by an $N_t \times N_{RF}^t$ analog RF precoder $\mathbf{F}_{\mathrm{RF}} \in \mathbb{R}^{N_t \times N_{RF}^t}$. Thus, the transmitted signal can be represented as $\boldsymbol{y} = \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\boldsymbol{s}$ $(\in \mathbb{R}^{N_t})$, where $\boldsymbol{s} \in \mathbb{R}^{N_s}$ is the the symbol vector with $\mathbb{E}[\boldsymbol{s}\boldsymbol{s}^H] = \frac{1}{N_s}\mathbf{I}_{N_s}$. For this system, the normalized tranmit power constraint, i.e., $\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 = N_s$, is imposed. More importantly, the analog precoders are implemented with phase shifters to control the phases of the signals. Hence, all the nonzero elements of $\mathbf{F}_{\mathrm{RF}}$ should satisfy the element-wise unit modulus constrains, i.e., $|(\mathbf{F}_{\mathrm{RF}})_{i,j}| = 1$.

3. PROPOSED OPTIMIZATION ALGORITHM

This section defines the problem formulations, and then the proposed algorithm and its complexity are described in detail.

3.1. Problem formulation [11, 14]

The following problem (1) has been shown as an equivalent formulation to maximize spectral efficiency [11, 13, 14]. This can be intuitively understandable since the optimal hybrid precoders is suffciently close to the unconstrained optimal fully digital precoder. Here, because the optimal fully digital precoder matrix $\mathbf{F}_{\mathrm{opt}} \in \mathbb{R}^{N_t \times N_s}$ includes the eigen vectors of the channel matrix $\mathbf{H}$, $\mathbf{F}_{\mathrm{opt}}$ equivalently becomes the first N_s columns of $\mathbf{V}$, where $\mathbf{V}$ is derived from the SVD of the channel $\mathbf{H}$, i.e., $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. Consequently, based on this $\mathbf{F}_{\mathrm{opt}}$, the problem is formulated as

$$\begin{aligned} \min_{\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}} \quad & \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F \\ \text{subject to} \quad & |(\mathbf{F}_{\mathrm{RF}})_{i,j}| = 1, \quad \|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 = N_s. \end{aligned} \tag{1}$$

Furthermore, the extended problem formulation for multi-carrier techniques such as the OFDM system to alleviate the effect of the multi-path fading can be defined as

$$\begin{aligned} \min_{\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}[l]} \quad & \sum_{l=1}^{L-1} \|\mathbf{F}_{\mathrm{opt}}[l] - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}[l]\|_F \\ \text{subject to} \quad & |(\mathbf{F}_{\mathrm{RF}})_{i,j}| = 1, \quad \|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}[l]\|_F^2 = N_s, \end{aligned} \tag{2}$$

where L is the total number of sub-carriers, and $l \in [0, L-1]$ is the sub-carrier index. $\mathbf{F}_{\mathrm{opt}}[l]$ and $\mathbf{F}_{\mathrm{BB}}[l]$ are the optimal digital precoder and the baseband precoder for the l-th sub-carrier, respectively.

3.2. Proposed algorithm

The discussions here base on those of [14]. The narrowband mmWave system is first addressed for simplicity. Due to the element-wise unit modulus constrains on $\mathbf{F}_{\mathrm{RF}}$, i.e., $|(\mathbf{F}_{\mathrm{RF}})_{i,j}| = 1$, the vectorized representation of $\mathbf{F}_{\mathrm{RF}}$, i.e., $\mathrm{vec}(\mathbf{F}_{\mathrm{RF}}) \in \mathbb{R}^{N_t N_{RF}^t}$, forms a *complex oblique manifold* $\mathcal{OB}(m, \mathbb{C})$, where $\mathcal{OB}(m, \mathbb{C}) = \{\mathbf{X} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m] \in$

$\mathbb{C}^{n\times m} : \|\boldsymbol{x}_1\| = \cdots = \|\boldsymbol{x}_m\| = 1\}$ with $m = N_t N_{RF}^t$ [14, 18]. Note that $\mathcal{OB}(m, \mathbb{C})$ is equivalent to the product of m complex circles. Thus, the problem (1) becomes a problem on a complex oblique manifold and a Euclidean space for which systematic procedures are proposed in [20, 21]. Consequently, the problem (1) is conceptually translated into unconstrained optimization problem over the search space $\mathcal{OB}(m, \mathbb{C}) \times \mathbb{R}^{N_{RF}^t \times N_s}$. Hence, ignoring $\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 = N_s$, the problem (1) is reformulated as

$$\min_{\mathrm{vec}(\mathbf{F}_{\mathrm{RF}})\in\mathcal{OB}(N_t N_{RF}^t,\mathbb{C})} \min_{\mathbf{F}_{\mathrm{BB}}\in\mathbb{R}^{N_{RF}^t\times N_s}} \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2. \quad (3)$$

The inner least-squares optimization problem in (3) is solved in closed form solution $\hat{\mathbf{F}}_{\mathrm{BB}} = \mathbf{F}_{\mathrm{RF}}^{\dagger}\mathbf{F}_{\mathrm{opt}}$ exploiting the least squares structure of the cost function. Hence, the final problem formulation is presented below.

$$\min_{\mathrm{vec}(\mathbf{F}_{\mathrm{RF}})\in\mathcal{OB}(m,\mathbb{C})} \|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}\|_F^2. \quad (4)$$

Denoting the Euclidean gradient of $\|\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}\|_F^2$ with respect to $\mathbf{F}_{\mathrm{RF}}$ as $\mathbf{G}(\mathbf{F}_{\mathrm{RF}})$, the Euclidean gradient $\nabla f(\boldsymbol{x}) \in \mathbb{R}^{N_t N_{RF}^t}$ of (4) is calculated as

$$\nabla f(\boldsymbol{x}) = \mathrm{vec}(\mathbf{G}(\mathbf{F}_{\mathrm{RF}})) = \mathrm{vec}(-2(\mathbf{F}_{\mathrm{opt}} - \mathbf{F}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}})\hat{\mathbf{F}}_{\mathrm{BB}}^H). \quad (5)$$

It should be emphasized that this *matrix-based* calculation of $\mathbf{G}(\mathbf{F}_{\mathrm{RF}})$ differs from MO-AltMin, which relies on the vectorization of $\mathbf{F}_{\mathrm{opt}}$ and the Kronecker product of $\mathbf{F}_{\mathrm{BB}}$ as

$$\nabla f(\boldsymbol{x}) = -2(\hat{\mathbf{F}}_{\mathrm{BB}}^* \otimes \mathbf{I}_{N_t})[\mathrm{vec}(\mathbf{F}_{\mathrm{opt}}) - (\hat{\mathbf{F}}_{\mathrm{BB}}^T \otimes \mathbf{I}_{N_t})\boldsymbol{x}]. \quad (6)$$

As a result, this modification drastically reduces the calculation complexity as seen in Section 3.3 and 4.1. Next, the Riemannian gradient at $\boldsymbol{x}$ is obtained by projecting $\nabla f(\boldsymbol{x})$ onto the tangent space $T_{\boldsymbol{x}}\mathcal{OB}(m, \mathbb{C})$ as $\mathrm{grad} f(\boldsymbol{x}) = \mathrm{P}_{\boldsymbol{x}}\nabla f(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) - \mathrm{Real}[\nabla f(\boldsymbol{x}) \circ \boldsymbol{x}^*] \circ \boldsymbol{x}$, where $\mathrm{P}_{\boldsymbol{x}}$ is a projection operator onto $T_{\boldsymbol{x}}\mathcal{OB}(m, \mathbb{C})$ [20]. Finally, the k-th iterate $\boldsymbol{x}^k$ is updated as $\boldsymbol{x}^{k+1} = R_{\boldsymbol{x}^k}(-\alpha^k \xi^k)$, where α^k is a step size, and $\xi^k \in T_{\boldsymbol{x}^k}\mathcal{OB}(m, \mathbb{C})$ is a direction derived from $\mathrm{grad} f(\boldsymbol{x})$. The derivation of ξ^k depends on the algorithm to be used, which is, for example, the Riemannian steepest descent (RSD) and the Riemannian conjugate gradient (RCG) [20, Section 8.3]. $R_{\boldsymbol{x}} : T_{\boldsymbol{x}}\mathcal{M} \to \mathcal{M} : \zeta \mapsto R_{\boldsymbol{x}}(\zeta)$ is called a *retraction* at $\boldsymbol{x}$, and maps tangent space $T_{\boldsymbol{x}}\mathcal{M}$ onto $\mathcal{M}$ with a local rigidity condition that preserves the gradients at $\boldsymbol{x}$. This case uses $R_{\boldsymbol{x}^k}(\alpha^k \xi^k) = [\frac{(\boldsymbol{x}^k + \alpha^k \xi^k)_i}{|(\boldsymbol{x}^k + \alpha^k \xi^k)_i|}]$. After the termination of the loop, $\mathbf{F}_{\mathrm{BB}}$ is obtained by normalizing as $\mathbf{F}_{\mathrm{BB}} = \sqrt{N_s}/\|\mathbf{F}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}^k\|_F$ to satisfy the constraint $\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}\|_F^2 = N_s$ in (1).

Now, the extension to OFDM defined in (2) is derived. The Euclidean gradient of $\sum_{l=1}^{L-1} \|\mathbf{F}_{\mathrm{opt}}[l] - \mathbf{F}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}[l]\|_F^2$ with respect to $\mathbf{F}_{\mathrm{RF}}$ is calculated instead of (5) as

$$\mathbf{G}(\mathbf{F}_{\mathrm{RF}}) = -2(\mathbf{F}_{\mathrm{opt}}[l] - \mathbf{F}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}[l])\hat{\mathbf{F}}_{\mathrm{BB}}^H[l].$$

Algorithm 1 Proposed algorithm with RSD.

Require: $\mathbf{F}_{\mathrm{opt}}$
1: Initialize $\boldsymbol{x}^0$.
2: **for** $k = 0, 1, 2, \ldots$ **do**
3: Inverse-vectorize $\boldsymbol{x}^k$ as $\mathbf{F}_{\mathrm{RF}}^k = \mathrm{vec}^{-1}(\boldsymbol{x}^k)$.
4: Compute $\hat{\mathbf{F}}_{\mathrm{BB}}^k = (\mathbf{F}_{\mathrm{RF}}^k)^{\dagger}\mathbf{F}_{\mathrm{opt}}$.
5: Compute the Euclidean gradient $\nabla f(\boldsymbol{x}^k)$ using (5).
6: Compute the Riemannian gradient $\xi^k = \mathrm{P}_{\boldsymbol{x}^k}\nabla f(\boldsymbol{x}^k)$.
7: Find the step size α^k, and update $\boldsymbol{x}^{k+1} = R_{\boldsymbol{x}^k}(-\alpha^k \xi^k)$.
8: **end for**
9: Output $\mathbf{F}_{\mathrm{RF}} = \mathrm{vec}^{-1}(\boldsymbol{x}^{k+1})$ and $\mathbf{F}_{\mathrm{BB}} = \frac{\sqrt{N_s}}{\|\mathbf{F}_{\mathrm{RF}}\hat{\mathbf{F}}_{\mathrm{BB}}^k\|_F}\hat{\mathbf{F}}_{\mathrm{BB}}^k$.

Since the following procedure is analogous to that of the narrowband mmWave system, it is omitted. Lastly, the overall algorithm is summarized in Algorithm 1, where, for simplicity, the RSD is used. It is noteworthy that, comparing with MO-AltMin, the updates of $\hat{\mathbf{F}}_{\mathrm{BB}}$ and $\mathbf{F}_{\mathrm{RF}}$ in the proposed algorithm are mutually integrated into one single-loop structure, achieving faster convergence of the proposed algorithm.

3.3. Complexity analysis

The proposed algorithm needs to calculate the Euclidean gradient defined in (5), which is the dominant calculation cost of the proposed algorithm. Hence, the computational complexity for the gradient calculation is $\mathcal{O}(N_{RF}^t N_t N_s)$. Meanwhile, the calculation cost of (6) in MO-AltMin algorithm is $\mathcal{O}(N_{RF}^t N_t^3 N_s^2)$ [14]. Apparently, the proposed algorithm has a lower complexity than the MO-AltMin algorithm because the latter relies on the Kronecker product calculation. Thus, as pointed out in [14], the MO-AltMin algorithm is numerically infeasible for large-scale systems in terms of the processing time. Furthermore, since PE-AlgMin requires $\min\{N_s^2 N_{RF}^t, N_s(N_{RF}^t)^2\}$ for the SVD calculation per iteration, the proposed algorithm competes the low-complexity PE-AlgMin algorithm. Consequently, as summarized in Table 1, the proposed algorithm is comparable to or much faster processing than the state-of-the-art algorithms whereas it produces near-optimal spectral efficiency as seen in Section 4.

Table 1. Complexity comparison

Algorithm	Complexity
MO-AltMin [14, Section III]	$\mathcal{O}(N_{RF}^t N_t^3 N_s^2)$
PE-AlgMin [14, Section IV]	$\min\{N_s^2 N_{RF}^t, N_s(N_{RF}^t)^2\}$
Proposed	$\mathcal{O}(N_{RF}^t N_t N_s)$

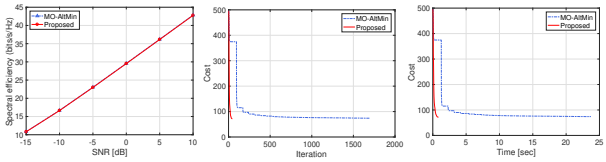
4. NUMERICAL EVALUATIONS

This section evaluates the performance of the proposed algorithm under the same conditions as [14]: The propagation environment is modeled as a $N_{cl} = 8$ cluster environment with $N_{ray} = 10$ rays per cluster with Laplacian distributed azimuth and elevation angles of arrival and departure with uniformly distributed mean angles over $[0, 2\pi)$ and angular spread of 10 degrees [22]. The experiments assume the

multi-carrier situation with $L = 128$, and that N_{RF}^t is equal to N_{RF}^r as $N_{RF}^t = N_{RF}^r = N_{RF}$. The RCG algorithm is used for the optimization algorithm. All experiments are executed in MATLAB on a 4.0 GHz Intel Core i7 PC with 32 GB RAM. The codes of the conventional algorithms are downloaded from the authors' site [2].

4.1. Comparison with near-optimal MO-AltMin

This subsection first presents the fast convergence property of the proposed algorithm against MO-AltMin which produces near-optimal spectral efficiency. The parameters are as follows; $N_s = 4$, $N_{RF}^t = N_{RF}^r = 6$ and $(N_t, N_r) = (144, 36)$. Fig.2(a) shows that the spectral efficiency of the proposed algorithm is almost identical to that of MO-AltMin in all SNRs. Fig.2(b) and (c) also show the degradation behaviors of the objective function, i.e., $\|\mathbf{F}_{\rm opt} - \mathbf{F}_{\rm RF}\hat{\mathbf{F}}_{\rm BB}\|_F^2$, with respect to the iteration number and the processing time, respectively. These reveal that the proposed algorithm is much faster than MO-AltMin for the two metrics. Consequently, the proposed algorithm yields the same spectral efficiency as MO-AltMin with much faster convergence speed than MO-AltMin.



(a) Spectral efficiency (b) Cost vs. Iteration (c) Cost vs. Time [sec]

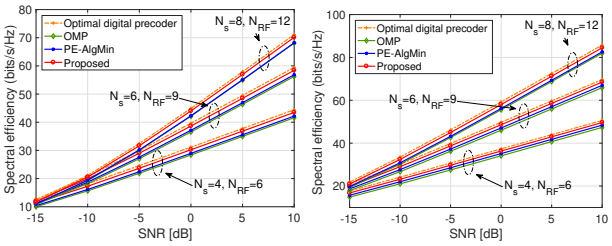
Fig. 2. Spectral efficiency and convergence behaviors.

4.2. Spectral efficiency and processing time evaluations

This subsection compares the proposed algorithm with the OMP algorithm [11], the PE-AlgMin algorithm [14, Section IV] and the optimal digital precoder for a benchmark. The MO-AltMin algorithm in [14, Section III] is omit because the spectral efficiency is the same as the proposed algorithm and its processing time is too high to be compared as shown earlier. Fig.3 first presents the spectral efficiency for each SNR in case of $(N_s, N_{RF}^t) = \{(4, 6), (6, 9), (8, 12)\}$ for $(N_t, N_r) = \{(144, 36), (256, 64)\}$. As the figures show, although the PE-AlgMin algorithm gives higher spectral efficiency than the OMP algorithm when $(N_s, N_{RF}^t) = \{(4, 6), (6, 9)\}$, it gives similar values of the OMP algorithm when $(N_s, N_{RF}^t) = (8, 12)$. On the other hand, the proposed algorithm yields stably higher results than both of them, and very close values as the optimal digital precoder.

The performances are next evaluated when changing N_{RF}^t under $N_s = \{6, 8, 12\}$ at $(N_t, N_r) = \{(144, 36), (256, 64)\}$. The range of $N_s < 2N_{RF}^t$ are addressed because $N_s \geq 2N_{RF}^t$ have been thoroughly investigated in [23]. Fig.4(a) and (b) show the spectral efficiency and processing time, respectively.

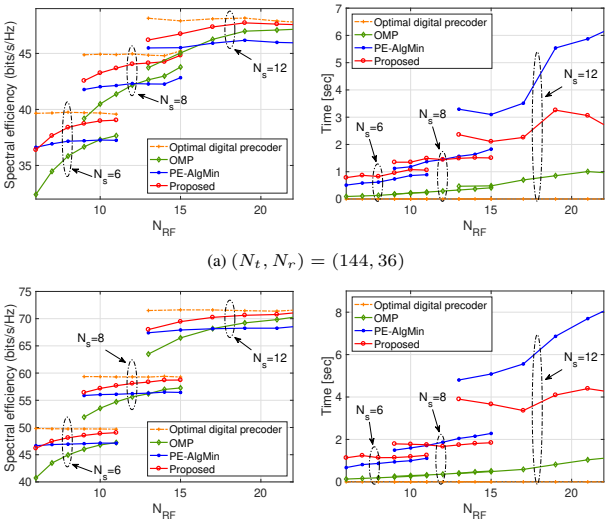
[2] https://github.com/yuxianghao/.



(a) $(N_t, N_r) = (144, 36)$ (b) $(N_t, N_r) = (256, 64)$

Fig. 3. Spectral efficiency $(N_s = \{4, 6, 8\})$.

Whereas the spectral efficiency of the PE-AlgMin algorithm is not changed, those of the OMP and proposed algorithm become improved as N_{RF} increases. At the higher range of N_{RF}, the proposed algorithm approaches to the optimal digital precoder at all N_s values. For the processing time, the OMP algorithm gives very lower values as expected at the expense of its lower spectral efficiency because it requires only basis searches and no iterations for optimization. Meanwhile, although the PE-AlgMin algorithm gives lower processing time than the propose algorithm for $N_s = 6$, it is hindered by the higher processing time than the proposed algorithm when higher N_{RF} and N_s. These results are coincident with Table 1 in terms of the sizes and orders of N_s and N_{RF}. To conclude, the proposed algorithm achieves near-optimal spectral efficiency with a comparable to or lower complexity than PE-AlgMin for a large-scale system with higher N_s and N_{RF}.



(a) $(N_t, N_r) = (144, 36)$

(b) $(N_t, N_r) = (256, 64)$

Fig. 4. Spectral efficiency and processing time.

5. CONCLUSIONS

The proposed algorithm is superior in terms of the comparable to or higher spectral efficiency than the state-of-the-art algorithms including MO-AltMin algorithm, and is faster processing than the low-complexity PE-AltMin algorithm.

6. REFERENCES

[1] S. Khiong and C.-C. Chong, "An overview of multigigabit wireless through millimeter wave technology: Potentials and technical challenges," *EURASIP J. Wirel. Commun. and Netw.*, vol. 2007, no. 1, 2006.

[2] R. W. Daniels, R. C. Heath, "60 GHz wireless communications: Emerging requirements and design recommendations," *IEEE Veh. Technol. Mag.*, vol. 2, no. 3, pp. 41–50, 2007.

[3] C.H. Doan, S. Emami, D.A. Sobel, A.M. Niknejad, and R.W. Brodersen, "Design considerations for 60 GHz CMOS radios," *IEEE Commun. Mag.*, vol. 42, no. 12, pp. 132–140, 2004.

[4] T. S. Rappaport, R. W. Heath Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*, Prentice Hall, 2014.

[5] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.

[6] H. Seleem, A. S. Sulyman, and A. Alsanie, "Hybrid precoding-beamforming design with hadamard RF codebook for mmWave large-scale MIMO systems," *IEEE Access*, vol. 5, pp. 6813–6823, 2017.

[7] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, 2011.

[8] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, 2013.

[9] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091–4103, 2005.

[10] V. Venkateswaran and Alle-Jan van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4131–4143, 2010.

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

[12] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, 2014.

[13] Y. Lee, C.-H. Wang, and Y.-H. Huang, "A hybrid RF/baseband precoding processor based on parallel-index-selection matrix-inversion-bypass simultaneous orthogonal matching pursuit for millimeter wave MIMO systems," *IEEE Trans. Signal Process.*, vol. 63, no. 2, pp. 305–317, 2015.

[14] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 10, no. 3, pp. 485–500, 2016.

[15] N. B. Zanjani, S. Khademi, and G. Leus, "Gradient-based solution for hybrid precoding in MIMO systems," in *ICASSP*, 2017.

[16] B. Mishra and R. Sepulchre, "R3MC: A Riemannian three-factor algorithm for low-rank matrix completion," in *IEEE CDC*, 2014.

[17] H. Kasai and B. Mishra, "Low-rank tensor completion: a Riemannian manifold preconditioning approach," in *ICML*, 2016.

[18] P.-A. Absil and K. A. Gallivan, "Joint diagonalization on the oblique manifold for independent component analysis," in *ICASSP*, 2006.

[19] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt: a Matlab toolbox for optimization on manifolds," *JMLR*, vol. 15, no. 1, pp. 1455–1459, 2014.

[20] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, 2008.

[21] A. Edelman, T.A. Arias, and S.T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.

[22] H. Xu, V. Kukshya, and T. S. Rappaport, "Spatial and temporal characteristics of 60-GHz indoor channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 620–630, 2001.

[23] E. Zhang and C. Huang, "On achieving optimal rate of digital precoder by RF-baseband codesign for MIMO systems," in *VTC2014-Fall*, 2014.