RIS Design to Optimize the CRB for Source Localization

Yuhua Jiang, Feifei Gao, and Wanmai Yuan

Abstract—Reconfigurable Intelligent Surface (RIS) plays a pivotal role in enhancing source localization accuracy. Based on the information inequality of Fisher information analyses, the Cramér-Rao Bound (CRB) of the localization error can be evaluated to localize the accuracy for a given set of RIS coefficients. However, there is a lack of research in optimizing these RIS coefficients to decrease the CRB under the constraint imposed by the RIS hardware. In this paper, we adopt the manifold optimization method to derive the locally optimal CRB of the localization error, where the RIS coefficients are restricted to lie on the complex circle manifold. Specifically, the Wirtinger derivatives are calculated in the gradient descent part, and the Riemannian nonlinear acceleration technique is employed to speed up the convergence rate. Simulation results show that the proposed method can yield the locally optimal RIS coefficients and can significantly decrease the CRB of localization error. Moreover, the iteration number can be reduced by the acceleration technique.

Index Terms—source localization, CRB optimization, RIS Design, manifold optimization, Riemannian nonlinear acceleration.

I. INTRODUCTION

Source localization is required in a variety of applications such as beamforming [1], object detection [2], smart environments [3], and Internet-of-Things (IoT) [4]. Recently, the reconfigurable intelligent surface (RIS) has been used in the sixth generation (6G) communications systems to enhance the accuracy of source localization. RIS plays an essential role in source localization when the line-of-sight (LOS) path between the transmitter and the receiver is blocked, hence making localization feasible when conventional systems fail. Moreover, RIS is beneficial to timely improve localization accuracy when the LoS path is present. Such flexibility makes RIS a pivotal localizing technology.

As for RISs-based localization, joint RIS selection and directional reflection beam design have been considered with prior knowledge of the user equipment (UE) location, which aims to concentrate the reflected power towards the UE and increase signal-to-noise-ratio (SNR) [5]. Alternatively, the authors in [6] use simpler random RIS phase profiles for asynchronous positioning in downlink single-input-single-output (SISO) transmission. This scheme does not require any prior information, but is not optimal under a priori UE localization information. In [7], the localization in narrow-band systems is extended to wideband systems with an unfavorable assumption that the phase responses of RIS elements must be constants over the considered frequency band. Another scheme for wideband localization with RIS is designed by dividing the frequency band into a sum of narrow frequency bands [8].

Based on the information inequality in Fisher information analyses, the Cramér-Rao Bound (CRB) can be used to evaluate the localization accuracy for a given set of RIS coefficients. In [9], the electromagnetic (EM) models are built to compute CRB for both discrete and continuous RISs in localization systems. In [10], the RIS coefficients are designed to focus energy on one of the anchors, which yields lower CRB than that with randomly designed RIS coefficients. In [11], CRB is derived for both position and orientation error of localizing a rotated user equipment.

However, the CRB minimization problems in [9]-[11] are nonconvex and are challenging to be solved directly by conventional optimization techniques. To our best knowledge, no optimization method has been proposed to directly minimize CRB by designing the RIS coefficients that take values in the set of practically feasible complex values (e.g., unit-modulus values with quantized phases).

In this paper, we apply the manifold optimization method to derive the locally optimal CRB of the localization error under practical RIS hardware limitations. In order to find the search direction, the Wirtinger derivatives are calculated in the gradient descent part of the optimization method. Moreover, the Riemannian nonlinear acceleration technique is employed to speed up the convergence rate. Simulation results show that the proposed method can significantly decrease the CRB of the localization error, and the iteration number can be reduced by the acceleration technique.

II. SYSTEM MODEL

Consider a three-dimensional (3D) scenario with one RIS, one agent that acts as the source, and $M$ anchors that act as receivers, as shown in Fig. 1. Denote the coordinate of the agent and the $m$th anchor by $q = [q_1, q_2, q_3]^T$ and $p_m = [x_m, y_m, z_m]^T$, respectively. The position of the agent is unknown, while the positions of the anchors are known. We consider the RIS as a rectangular plate with length $a$ in $y$-axis and length $b$ in $x$-axis, located in the horizontal plane. Suppose the RIS is equipped with $N$ passive elements, and each element is $l_1$ long and $l_2$ wide. Suppose the wave number is $k_0 = 2\pi/\lambda_0$, where $\lambda_0$ is the wavelength of the transmitted electromagnetic waves.
Denote $r_m$ as the distance from the $m$th anchor to the agent, $\rho_n$ as the distance from the $n$th element of the RIS to the agent, and $d_{mn}$ as the distance from the $m$th anchor to the $n$th element of the RIS. The overall channel vector from the transmitter to RIS and RIS to the $m$th anchor is denoted as $h_{RIS(m)} (m = 1, \cdots, M)$. The direct path channel from the transmitter to the $m$th anchor is denoted as $h_{DP(m)} (m = 1, \cdots, M)$. The two channels can be respectively written as [10], [9]

$$
\begin{align}
    h_{RIS(m)} &= \frac{j k_0 l_1 l_2 \left[ \gamma_{mN} e^{j k_0 (\rho_n + d_{mn})} \right]}{4\pi \sqrt{R_L}}, \\
    h_{DP(m)} &= \frac{s_m DP \sqrt{F_m}}{2\sqrt{R_L}} e^{-j k_0 r_m},
\end{align}
$$

where $R_L$ is the load resistance of the anchor antenna; $\gamma_{mN}$ is the loss due to the polarization difference between the antenna at the $m$th anchor and the transmitted EM waves via the $n$th element of the RIS; $F_{m}^{RIS}$ is the directive gain of the antenna of the $m$th anchor and the reflected EM waves via the $n$th element of the RIS; $\gamma_{DP}$ is the loss due to the polarization difference between the antenna at the $m$th anchor and the transmitted EM waves via the direct path; $F_m$ is the directive gain of the antenna at the $m$th anchor and the transmitted EM waves via the direct path; $s_m$ is the signal transmitted by the agent. Then, the signal received by the $m$th anchor is

$$
y_m = (h_{RIS(m)} w + h_{DP(m)}) x + n,
$$

where $x$ denotes the signal transmitted by the agent, $s_m$ denotes the received signal of the $m$th anchor without noise, and $n$ denotes zero-mean additive white Gaussian noise (AWGN) at the anchor with variance $\sigma^2$.

After receiving signals from the agent, the anchors can jointly estimate the location of the agent by methods such as the maximum likelihood estimator (MLE) [11].

### III. The Cramér-Rao Bound of Localization Error

Denote the Fisher information matrix (FIM) as $J(q)$. According to [12], $J(q)$ can be expressed by the FIM of the $m$th anchor $J_m(q), (m = 1, \cdots, M)$ as

$$
J(q) = \sum_{m=1}^{M} J_m(q).
$$

The elements of $J_m(q)$ are given by

$$
[J_m(q)]_{i,k} = \frac{2}{\sigma^2} \Re \left\{ \frac{\partial s_m}{\partial q_i} \frac{\partial s_m^*}{\partial q_k} \right\},
$$

where $\Re\{\cdot\}$ denotes the real part of the element. By substituting (1) and (2) into (3) and using the chain rule of derivative, the first derivative in (7) can be computed as

$$
\frac{\partial s_m}{\partial q_i} = \frac{j k_0 l_1 l_2^2}{4\pi \sqrt{R_L}} \sum_{n=1}^{N} \frac{\gamma_{mN} e^{j k_0 (\rho_n + d_{mn})}}{\rho_n d_{mn}} \frac{\partial \rho_n}{\partial q_i} \times (-\frac{1}{\rho_n - j k_0}) + \frac{\gamma_{DP}}{2\sqrt{R_L}} \frac{F_m}{r_m} e^{-j k_0 r_m} \frac{\partial r_m}{\partial q_i},
$$

where $A_{mn}$ and $c_{mi}$ are constants when $x$, $q_i$, and $p_m$ are fixed. Thus, (7) and (8) are only related to $w$. Then, the mean square error (MSE) of any unbiased estimator $\hat{q}$ of $q$ is lower bounded by CRB computed as [13]

$$
\text{CRB} = \text{tr}[J^{-1}(q)] = f(w).
$$

To enhance the positioning accuracy under the constraint on the RIS reflection coefficients, one needs to solve

$$
P_1 : \min_{w} f(w)
$$

s.t. $|w_n| = 1, n = 1, \cdots, N$.

Since both the objective function and the constraint in $P_1$ are nonconvex, $P_1$ is challenging to be solved by convex optimization techniques. In the following, an efficient algorithm is developed to derive a high quality locally optimal solution.

### IV. Manifold Optimization

#### A. Wirtinger Gradient

Since $f(w)$ is a non-trivial (not constant) real-valued function with complex arguments $w$, $f(w)$ is non-analytic and therefore is not complex differentiable. Thus, the steepest descent direction of $f(w)$ is $-\nabla_{w^*} f(w)$ given by Wirtinger
gradient [14], where \( w^* \) denotes the conjugate of \( w \). The \( g \)th element of \( \nabla_{w^*} f(w) \) is
\[
\frac{\partial f(w)}{\partial w_i^*} = -\text{tr}[J^{-1}(q) \frac{\partial J(q)}{\partial w_i} J^{-1}(q)],
\]
\[
= -\text{tr}[J^{-1}(q) \sum_{m=1}^M \frac{\partial J_m(q)}{\partial w_i} J^{-1}(q)].
\]
(11)

According to [14], \( w^* \) and \( w \) can be viewed as two independent variables in Wirtinger’s derivative rule. Hence, the elements of \( \frac{\partial J_m(q)}{\partial w_i} \) are computed by utilizing \( \Re(z) = \frac{z + z^*}{2} \) and substituting (8) into (7) as
\[
\left[ \frac{\partial J_m(q)}{\partial w_i} \right]_{i,k} = \frac{1}{\sigma^2} A_{mgk} \left( \sum_{n=1}^N w_n A_{mn} + c_{mi} \right) + \frac{1}{\sigma^2} A_{mg} \left( \sum_{n=1}^N w_n A_{mnk} + c_{mk} \right). \]
(12)

B. Riemannian Gradient Descent

The unit modulus constraint in \( P_1 \) can be geometrically interpreted as restricting \( w \) to lie on the complex circle manifold that is defined as
\[
\mathcal{M} = \{ w \in \mathbb{C}^N : |w_1| = |w_2| = \cdots = |w_N| = 1 \}. \]
(13)
The tangent space at \( w_i \) on the complex circle manifold is denoted as \( T_{w_i} \mathcal{M} \), which is the space of tangent vectors passing through \( w_i \) and is given by
\[
T_{w_i} \mathcal{M} = \{ v \in \mathbb{C}^N : \Re(v \odot w_i^*) = 0 \}, \]
(14)
where \( \odot \) denotes the Hadamard element-wise multiplication, and \( 0 \) denotes a vector with all zero elements. Among all tangent vectors, the one that yields the fastest increase of the objective function is defined as the Riemannian gradient, \( \text{grad} f(w_i) \). Numerically, \( \text{grad} f(w_i) \) is computed by first calculating the steepest ascent direction in the Euclidean space, i.e., \( \nabla_{w^*} f(w_i) \), and then projecting it onto the tangent space via a projection operator. The projection operator from the Euclidean space onto the tangent space \( T_{w_i} \mathcal{M} \) is given by [15]
\[
P_{T_{w_i} \mathcal{M}}(v) = v - \Re\{v \odot w_i^*\} \odot w_i.
\]
(15)
Hence, the Riemannian gradient of \( f(w_i) \) is expressed as
\[
\text{grad} f(w_i) = P_{T_{w_i} \mathcal{M}}(\nabla_{w^*} f(w_i)) = \nabla_{w^*} f(w_i) - \Re\{\nabla_{w^*} f(w_i) \odot w_i^*\} \odot w_i.
\]
(16)

We then employ the conjugate gradient (CG) method [15]. The update rule of the search direction in the Euclidean space is given by
\[
\mu_{i+1} = -\nabla_{w^*} f + \alpha_i \mu_i,
\]
(17)
where \( \mu_i \) denotes the search direction at \( w_i \) and \( \alpha_i \) is chosen as the Polak-Ribiere parameter to achieve fast convergence [15]. However, since \( \mu_i \) and \( \mu_{i+1} \) in (17) lie in \( T_{w_i} \mathcal{M} \) and \( T_{w_{i+1}} \mathcal{M} \), respectively, they cannot be integrated directly over different tangent spaces. Thus, we need to project \( \mu_i \) from tangent space \( T_{w_i} \mathcal{M} \) to tangent space \( T_{w_{i+1}} \mathcal{M} \) [16]. Similar to (17), the search direction based on the Riemannian gradient can be updated as:
\[
\mu_{i+1} = -\text{grad} f(w_i) + \alpha_i P_{T_{w_{i+1}} \mathcal{M}}(\mu_i),
\]
(18)
where \( \mu_{i+1} \) is the Polak-Ribiere parameter \( \alpha_i \) is computed as [15]
\[
\alpha_i = \frac{\text{grad} f(w_i) \cdot (\text{grad} f(w_i) - \text{grad} f(w_{i-1}))}{\text{grad} f(w_{i-1}) \cdot \text{grad} f(w_{i-1})}.
\]
(19)

However, given the search direction \( \mu_i \), the solution cannot be simply updated via \( w_{i+1} = w_i + \eta_i \mu_i \), where \( \eta_i \) denotes the searching step size and \( \eta_i \mu_i \) is defined as search vectors. The reason is that \( w_i + \eta_i \mu_i \) would lie in the tangent space \( T_{w_i} \mathcal{M} \) but not on the surface of the manifold. Hence, a retraction function is needed from the tangent space to the surface of the manifold such that the RIS hardware limitation is not violated. For the complex circle manifold, the retraction function that maps \( v \in T_{w_i} \mathcal{M} \) onto \( \mathcal{M} \) can be defined as [15]
\[
R_{w_i}(v) = (w_i + v) \odot |w_i + v|, \]
(20)
where \( | \cdot | \) and \( \odot \) denote element-wise absolute value and element-wise division, respectively. We then adopt retraction to find the next iterate \( w_{i+1} \) on the manifold as
\[
w_{i+1} = R_{w_i}(\eta_i \mu_i),
\]
(21)
where the Armijo backtracking line search algorithm [15] can be used to choose the step size \( \eta_i \), which ensures that the objective function is decreasing at each iteration, i.e., \( f(w_{i+1}) < f(w_i) \).

A locally optimal solution can be found by repeating the above steps until \( \|\text{grad} f(w_k)\|_2 < \epsilon \), where \( \epsilon \) is the convergence tolerance, and \( k \) is the maximum number of iteration until convergence. This procedure is defined as Riemannian gradient descent (RGD).

C. Riemannian Nonlinear Acceleration

Although the monotonic decrease in the objective function is guaranteed after each iteration, the convergence is reached with maximum number of iteration \( k = \mathcal{O}(\epsilon^{-2}) \) [17]. When small convergence tolerance is demanded, \( k \) may be unacceptably large, and the RGD algorithm may be slow in practice. Hence, we apply the regularized nonlinear acceleration gradient algorithm [18] to accelerate the RGD algorithm in parallel.

Let us consider a generalization of nonlinear acceleration for Riemannian optimization via a weighted Riemannian average on the manifold. Define \( M_d \) as the memory depth that controls the interval between two acceleration operations. After obtaining a sequence of iterates from the RGD process denoted as \( \{w_i\}_{i=0}^{k} \), we define the residuals as the projection of all search vectors onto the tangent space at \( w_{M_d-1} \), which is expressed as
\[
r_i = P_{T_{w_{M_d-1}} \mathcal{M}}(\eta_i \text{grad} f(w_i)).
\]
(22)
The weights for the acceleration is defined as \( c \), and can be obtained by minimizing the sum of a weighted combination of
the residuals and a regularization term in the weights [18]. The solution of \( c \) is given by the following optimization problem:

\[
\begin{align*}
\hat{c} &= \arg\min_{c \in \mathbb{R}^{M_d}} \left\{ \sum_{i=0}^{k} c r_i \right\}^2 + \lambda \|c\|^2 \right\},
\end{align*}
\]

where \( I \) denotes a vector with all elements equal to 1, and \( \lambda \) is the regularization parameter. We then show that the optimal weight \( c \) has a closed-form solution. Define the residual matrix as \( \mathbf{R} = \left\{ r_i^T r_j \right\}_{j=0}^{M_d} \in \mathbb{C}^{M_d \times M_d} \), which collects all pairwise inner products of \( r_i \). Note that (23) is a linearly constrained quadratic program that can be solved by introducing \( \mu \in \mathbb{R} \) as the dual variable. Then (23) indicates that \( c \) and \( \mu \) satisfy the Karush-Kuhn-Tucker (KKT) system:

\[
\begin{bmatrix}
2(\mathbf{R} + \lambda I) & 1
\end{bmatrix}^{-1} \begin{bmatrix}
c \\
\mu
\end{bmatrix} = \begin{bmatrix}
0 \\
1
\end{bmatrix}.
\]

Solving (24) yields the closed form of \( c \) as

\[
\hat{c} = \frac{(\mathbf{R} + \lambda I)^{-1} 1}{1^T (\mathbf{R} + \lambda I)^{-1}} ,
\]

where \( I \) is the identical matrix. The sequence of converging iterates from the RGD process recursively constructs an updated iterate \( \hat{w}_i \) as

\[
\hat{w}_i = \mathcal{R}_{\hat{w}_{i-1}} \left( \frac{c_i}{\sum_{j=0}^{M_d} c_j} \mathcal{R}_{\hat{w}_{i-1}}^{-1}(\hat{w}_i) \right), \quad \hat{w}_0 = w_0.
\]

The inverse of the retraction \( \mathbf{v} = \mathcal{R}_{w_0}^{-1}(w_{i+1}) \) maps \( w_{i+1} \in \mathcal{M} \) to \( \mathbf{v} \in T_{w_0} \mathcal{M} \), and can be derived as follows. Since the angles of all elements in \( \langle \hat{w}_i + \mathbf{v} \rangle \) are unchanged in (21), there is

\[
\arg(\hat{w}_i + \mathbf{v}) = \arg(\hat{w}_{i+1}),
\]

where \( \arg(\cdot) \) denotes the arguments of all elements in a vector. Note that \( \mathbf{v} \in T_{w_0} \mathcal{M} \) and (14) can be rewritten as

\[
\arg(\mathbf{v}) - \arg(\hat{w}_i) = \pm \frac{\pi}{2}.
\]

Based on the geometrical constraints in (27) and (28), the closed form of the inverse of the retraction is given by

\[
\mathcal{R}_{w_0}^{-1}(w_{i+1}) = j \mathbf{w}_i \odot \tan(\arg(w_{i+1}) - \arg(w_i)).
\]

Substituting (29) into (26), the iterate sequence is updated and RGD can restart with \( w_0 = \hat{w}_{M_d-1} \). The nonlinear acceleration is performed in parallel to the RGD process in the complete procedure, which is summarized in Algorithm 1.

\[\textbf{Algorithm 1:} \text{Riemannian gradient descent (RGD) with nonlinear acceleration}\]

\begin{algorithm}
\textbf{Input:} Initialization \( w_0 \), regularization parameter \( \lambda \), convergence tolerance \( \epsilon \), and memory depth \( M_d \).
\begin{algorithmic}[1]
\STATE while \( \|\text{grad} f(w_0)\|_2 \geq \epsilon \) do
\FOR {\( i = 1 \) to \( M_d \)}
\STATE Compute \( \text{grad} f(w_i) \) via (16).
\STATE Compute \( \mu_{i-1} \) via (18) with \( \mu_{-1} = 0 \).
\STATE Update \( w_i = \mathcal{R}_{w_{i-1}}(\eta_{i-1}\mu_{i-1}) \) with \( \eta_{i-1} \) given by Armijo backtracking line search [16, 4.2.2].
\ENDFOR
\STATE Compute \( r_i \) via (22).
\STATE Compute \( \hat{w}_i \) via (25).
\STATE Compute \( \hat{w}_i(0, \ldots, M_d-1) \) via (26), with \( \hat{w}_{-1} = w_0 \).
\STATE Restart with \( w_0 = \hat{w}_{M_d-1} \).
\end{algorithmic}
\end{algorithm}

Fig. 2. Riemannian gradient descent algorithm with and without Riemannian nonlinear acceleration.

\[\text{B. The RCRB versus The Distance from The Agent to The RIS}\]

The position of the agent is set at \( q = (0, 0, 20) \), and the length of the RIS is set as \( a = b = 0.8 \). We perform the Riemannian gradient descent algorithm with and without Riemannian nonlinear acceleration. It is seen from Fig. 2 that, the acceleration technique can reduce the iteration number by more than half.
C. The RCRB versus The Length of The RIS

The RCRB becomes smaller when the size of the RIS becomes larger. This is because with the controlled RIS coefficients, more energy of the scattered wave can be directed toward the main lobe of the pattern of illuminated RIS as the operating frequency increases. The proposed RGD method outperforms the focus-on-\textbf{p}_1 method in all cases, because the latter scheme only ensures a maximum modulus of noise-free signal received by the first anchor while does not consider signals received by the other anchors.

VI. Conclusion

In this paper, we apply the manifold optimization method to derive the locally optimal CRB of the localization error under practical RIS hardware limitations, where the Wirtinger gradient is calculated to find the steepest descent direction. To solve the problem of slow convergence, the Riemannian nonlinear acceleration technique is employed to speed up the convergence rate. Simulation results show that the optimally configured RIS designed by the proposed method can significantly reduce the CRB. Moreover, the iteration number can be reduced by the acceleration technique.

REFERENCES

[1] Y. Jiang, F. Gao, M. Jian, S. Zhang, and W. Zhang, “Reconfigurable intelligent surface for near field communications: Beamforming and sensing,” 2022.
[2] A. A. Saucan and M. Z. Win, “Information-seeking sensor selection for ocean-of-things,” IEEE Internet Things J., vol. 7, no. 10, pp. 10072–10088, 2020.
[3] G. Pasolini, C. Buratti, L. Feltrin, F. Zabini, C. D. Castro, R. Verdone, and O. Andrisano, “Smart city pilot projects using lora and 802.15.4 technologies,” Sensors (Basel, Switzerland), vol. 18, 2018.
[4] M. Z. Win, F. Meyer, Z. Liu, W. Dai, S. Bartoletti, and A. Conti, “Efficient multisensor localization for the internet of things: Exploring a new class of scalable localization algorithms,” IEEE Signal Process. Mag., vol. 35, no. 5, pp. 153–167, 2018.
[5] H. Wymeersch and B. Denis, “Beyond 5g wireless localization with reconfigurable intelligent surfaces,” in ICC 2020, pp. 1–6, 2020.
[6] K. Keykhoosravi, M. F. Keskin, G. Seco-Granados, and H. Wymeersch, “Ssiso ris-enabled joint 3d downlink localization and synchronization,” in ICC 2021, pp. 1–6, 2021.
[7] T. Ma, Y. Xiao, X. Lei, W. Xiong, and Y. Ding, “Indoor localization with reconfigurable intelligent surfaces,” IEEE Commun. Lett., vol. 25, no. 1, pp. 161–165, 2021.
[8] M. He, W. Xu, H. Shen, G. Xie, C. Zhao, and M. Di Renzo, “Cooperative multi-ris communications for wideband mmwave miso-ofdm systems,” IEEE Wireless Commun. Lett., vol. 10, no. 11, pp. 2360–2364, 2021.
[9] M. Z. Win, Z. Wang, Z. Liu, Y. Shen, and A. Conti, “Location awareness via intelligent surfaces: A path toward holographic NLN,” IEEE Veh. Technol. Mag., vol. 17, no. 2, pp. 37–45, 2022.
[10] Z. Wang, Z. Liu, Y. Shen, A. Conti, and M. Z. Win, “Source localization with intelligent surfaces,” in ICC 2022, pp. 895–900, 2022.
[11] A. Elzanaty, A. Guerra, F. Guidi, and M.-S. Alouini, “Reconfigurable intelligent surfaces for localization: Position and orientation error bounds,” IEEE Trans. Signal Process., vol. 69, pp. 5386–5402, 2021.
[12] M. J. Khojasteh, A. A. Saucan, Z. Liu, A. Conti, and M. Z. Win, “Node deployment under position uncertainty for network localization,” in ICC 2022, pp. 889–894, 2022.
[13] Z. Wang, Z. Liu, Y. Shen, A. Conti, and M. Z. Win, “Wideband localization with reconfigurable intelligent surfaces,” in 2022 IEEE 95th Vehicular Technology Conference (VTC2022-Spring), pp. 1–6, 2022.
[14] R. Burckel and R. Remmert, Theory of Complex Functions. Graduate Texts in Mathematics / Readings in Mathematics, Springer New York, 1999.
[15] P.-A. Absil, R. Mahony, and R. Sepulchre, Optimization Algorithms on Matrix Manifolds. Princeton: Princeton University Press, 2009.
[16] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, “Alternating minimization algorithms for hybrid precoding in millimeter wave mimo systems,” IEEE J. Select. Topics in Signal Process., vol. 10, no. 3, pp. 485–500, 2016.
[17] N. Boumal, P.-A. Absil, and C. Cartis, “Global rates of convergence for nonconvex optimization on manifolds,” IMA Journal of Numerical Analysis, vol. 39, pp. 1–33, Feb 2018.
[18] S. Damien, d. Alexandre, and B. Francis, “Regularized nonlinear acceleration,” Mathematical Programming, vol. 179(1), p. 4783, 2020.