IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 16, 2022, accepted May 31, 2022, date of publication June 6, 2022, date of current version June 15, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3180366

A Neurodynamic Approach to Nonsmooth Quaternion Distributed Convex Optimization With Inequality and Affine Equality Constraints

GUOCHENG LI[1], LINHUA LUAN[2], SITIAN QIN[2], AND JIQIANG FENG[3]

[1]Department of Mathematics, Beijing Information Science and Technology University, Beijing 100192, China
[2]Department of Mathematics, Harbin Institute of Technology, Weihai 264209, China
[3]Institute of Intelligent Computing Science, Shenzhen University, Shenzhen 518060, China

Corresponding author: Jiqiang Feng (fengjq@szu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62176073 and Grant 11871178 and in part by the Natural Scientific Research Innovation Foundation in Harbin Institute of Technology.

ABSTRACT Quaternions have appeared in many practical fields, such as image processing and data mining, and so on. This paper focuses on designing an efficient quaternion-valued neurodynamic approach (QNA) based on multi-agent systems to solve nonsmooth convex quaternion distributed optimization problems (QDOPs) with inequality and affine equality constraints. Each agent in the system cooperatively solves the minimum of the global objective function through the information of itself and its neighbors. At first, the related convex analysis of the quaternion field is given, which provides a theoretical basis for solving the nonsmooth QDOP. Then the considered QDOP is equivalently transformed by using the connectivity of the communication topology. After that, a distributed QNA is presented, where the adaptive controller is introduced to ensure that the penalty terms with respect to the inequality constraints can self-adjust according to the local states. It is shown that all agents reach a consensus while obtaining the optimal solution to the related QDOP. Finally, a numerical example and an application in dictionary sparse representation of color images based on quaternion are realized to intuitively describe the effectiveness and practical significance of the proposed QNA.

INDEX TERMS Adaptive controller, multi-agent system, nonsmooth, quaternion distributed optimization problem (QDOP), quaternion-valued neurodynamic approach (QNA).

I. INTRODUCTION

Since W.R.Hamilton introduced the well-known quaternion in [1], the quaternion field has been regarded as the field of generalized complex numbers with three independent imaginary parts. With the development of 3D technology and the exploration of high-dimensional problems, the importance of quaternion distributed optimization problems (QDOPs) has gradually emerged in fields of science and technology, such as color image processing [2], [3], satellite tracking [4], bearings-only tracking [5] and so on. In particular, in the field of color image processing [2], the color image is abstracted into a quaternion array by using three imaginary parts of quaternion to represent the gravity of the three primary colors, respectively, thus keeping each signal channel from interfering with others.

The associate editor coordinating the review of this manuscript and approving it for publication was Azwirman Gusrialdi.

Generally, the form of a distributed optimization problem with quaternion decision variables (hereafter referred to as QDOP) can be represented as follows

$$\begin{aligned} &\min g(p) \triangleq \sum_{i=1}^{N} g_i(p) \\ &\text{s.t. } h_i(p) \leq 0, \quad D_i p = d_i, \ i = 1, 2, \cdots, N \end{aligned} \tag{1}$$

where $p = (p_1, p_2, \cdots, p_n)^{\mathrm{T}} \in \mathbb{H}^n$ is the decision variable,$\mathbb{H}^n$ is the set of n-dimensional quaternion vectors, $g_i : \mathbb{H}^n \to \mathbb{R}$ and $h_i : \mathbb{H}^n \to \mathbb{R}^{m_i}$ with $h_{ij} : \mathbb{H}^n \to \mathbb{R}$ ($j = 1, 2, \cdots, m_i$) are quaternion-variable functions, $D_i \in \mathbb{H}^{s_i \times n}$ and $d_i \in \mathbb{H}^{s_i}$. The optimization objective is to find the quaternion p satisfying the constraint conditions so as to minimize the value of the objective function g, which is different from the real-valued optimization problem because of the different domain of decision variables. As the world has entered the information age, multi-agent systems that can protect the security of private information have been applied in various aspects, such as optimal control [6]–[8] and network function virtualization [9]. Thus proposing an efficient algorithm to solve QDOP (1) with a multi-agent system has become a hot topic for many researchers. Inspired by the neural networks proposed by Tank and Hopfield in [10] and [11], various real-valued neurodynamic approaches have emerged to solve real-valued distributed optimization problems in [12]–[16] and references therein until now. However, when studying QDOP (1) and the affine transformation of high-dimensional sets similar to [17] and [18], it is necessary to propose QNAs for searching the minimum of such problems.

As the extension of real-valued and complex-valued neurodynamic approaches, QNAs have the following advantages:

- The structure of quaternions enables QNAs to maintain the integrity of the original structure and the coupling of internal structure in data modeling and processing of three-dimensional or four-dimensional space.
- Quaternions can improve the compactness of the model, preserve the structural framework of the system, and reflect the physical properties of the original problem better.

Therefore, much research in the quaternion field appears, such as [19]–[25]. However, the special properties of quaternion bring many difficulties to the study of QNAs. As we all know, whether in the real or complex field, the design of neurodynamic approaches for solving distributed optimization problems are inseparable from gradients or sub-gradients of the objective function and constraint functions. To design efficient neurodynamic approaches in the quaternion field, [20] and [21] proposed Hamilton-real calculus, generalized Hamilton-real calculus, and the definitions of quaternion derivative and quaternion Hessian matrix, which provided the theoretical support for solving the nonsmooth QDOP later. In [22], the stability of QNAs with continuous-time and discrete-time were analyzed by the complex factorization of quaternions, and [23] adopted quaternion matrix theory to solve the state estimation problem of QNA with delay. As for the existing QNAs used to solved quaternion optimization, [24] promoted the real-valued neurodynamic approach in [12] to solve the quaternion optimization problem with affine equality constraints and set constraints by the penalty method, and [25] further proposed a QNA for the nonsmooth QDOP with inequality constraints.

However, as far as we know, there are still few quaternion-valued neurodynamic approaches for solving nonsmooth convex QDOPs with inequality constraints and affine equality constraints. Inspired by the QNAs in [24] and [25], a continuous-time QNA is constructed in this paper and other main contributions are listed below.

- In comparison with the existing QNAs to solving QDOPs in [24], [25], this paper further studies the sub-differential properties of the global objective function and the optimality condition for QDOPs and takes the inequality and affine equality constraints into account by using multi-agent system.
- Instead of introducing penalty parameters that need to estimate the lower bounds in advance like [16], [24]–[27], this paper designs a novel QNA by introducing the adaptive controller to ensure that the state solution finally enters the feasible region of the QDOP with less computation amount.
- The QNA proposed in this paper can be further simplified as a real-valued neurodynamic approach for solving real-valued distribution optimization problems with inequality and affine equality constraints. Compared with the real-valued neurodynamic approaches in [12]–[16], the approach herein can avoid using global information of the optimization with the help of the adaptive controller.

The remaining part consists of the introduction of basic knowledge and quaternion-related theories in Section II, the equivalent transformation of QDOP (1) in Section III, the convergence analysis of the proposed QNA in Section IV, the corresponding numerical simulations in Section V, and the conclusion in Section VI.

II. PRELIMINARIES

A. NOTATIONS AND GRAPH THEORY

Throughout this paper, we adopt $\mathbb{R}$ as the set of real numbers, $\mathbb{H}$ denotes the set of quaternions. For any vector $v \in \mathbb{R}^n$ (or $\mathbb{H}^n$), matrix $A \in \mathbb{R}^{m\times n}$ (or $\mathbb{H}^{m\times n}$), v_i is the ith component of the vector v, a_{ij} represents the entry located in the ith row and jth column of matrix $A = (a_{ij})_{m\times n}$. I_n represents the n-dimensional identity matrix. $(\cdot)^{\mathrm{T}}$ means the transpose and $\|\cdot\|$ means the Euclidean norm. The notations $\frac{\mathrm{d}}{\mathrm{d}v}$ and $\frac{\partial}{\partial v}$ denote the differential operator and partial differential operator with respect to v. As for any set S, $\mathrm{int}(S)$, $\mathrm{bd}(S)$, and $\overline{S}$ mean the interior, boundary, and closure of the set S, respectively. $\mathrm{col}(x, y) = (x^{\mathrm{T}}, y^{\mathrm{T}})^{\mathrm{T}}$ for any vectors x, y; $\mathrm{diag}(A, B)$ denotes the partitioned diagonal matrix with A and B as diagonal elements and $A \otimes B$ stands for the Kronecker product of A and B.

Let $\mathcal{G}(\mathcal{V}, \mathcal{E})$ be a graph consisting of nodes $i \in \mathcal{V} = \{1, 2, \cdots, N\}$ and corresponding edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ where $(i, j) \in \mathcal{E}$ if and only if nodes i and j are connected. The neighbor set of the node i is defined as $\mathcal{N}^i = \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$. $A = (a_{ij})_{N\times N}$ is the adjacency matrix of $\mathcal{G}$ and the connection weight $a_{ij} = 1$ when $j \in \mathcal{N}^i$, $a_{ij} = 0$ otherwise. In particular, the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is an undirected graph when $a_{ij} = a_{ji}$ for any $i, j \in \mathcal{V}$. If there further exists a path $\{(i, i_1), (i_1, i_2), \cdots, (i_k, j)\}$ from node i to node j for any $i, j \in \mathcal{V}$, then the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is connected. The Laplacian matrix is defined as $L = D - A$ where $D = \mathrm{diag}\{D_{11}, D_{22}, \cdots, D_{NN}\}$ and $D_{ii} = \sum_{j=1}^{N} a_{ij}$. It is well known that L is positive semidefinite when $\mathcal{G}$ is connected, and $x^{\mathrm{T}}Lx = 0$ if and only if $Lx = 0$.

B. QUATERNION ALGEBRA ANALYSIS

The quaternion concept proposed by W.R.Hamilton in [1] is as follows:

$$\begin{aligned}\mathbb{H} :&= \mathrm{span}\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}\\ &= \{x + y\mathbf{i} + z\mathbf{j} + w\mathbf{k} : x, y, z, w \in \mathbb{R}\}\end{aligned}$$

where the imaginary units $\mathbf{i}, \mathbf{j}, \mathbf{k}$ satisfy the following Hamilton rule

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \ \mathbf{ijk} = -1, \\ \mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \ \mathbf{ki} = -\mathbf{ik} = \mathbf{j}. \tag{2}$$

Based on the definition of the quaternion field $\mathbb{H}$, for any $p = p_x + p_y\mathbf{i} + p_z\mathbf{j} + p_w\mathbf{k} \in \mathbb{H}$,

$$\mathcal{R}(p) = p_x, \\ \mathcal{I}(p) = p_y\mathbf{i} + p_z\mathbf{j} + p_w\mathbf{k},$$

are the real part and imaginary part of p, respectively. In addition, take

$$\mathcal{I}^{\mathbf{i}}(p) = p_y, \quad \mathcal{I}^{\mathbf{j}}(p) = p_z, \ \text{and } \mathcal{I}^{\mathbf{k}}(p) = p_w,$$

the quaternion conjugate of p is

$$p^* = p_x - p_y\mathbf{i} - p_z\mathbf{j} - p_w\mathbf{k},$$

and the norm of p is

$$\|p\| = \sqrt{pp^*} = \sqrt{p_x^2 + p_y^2 + p_z^2 + p_w^2}.$$

For any $p \in \mathbb{H}^n$, p^{T} and p^{H} denote the transpose and transpose conjugate of p, respectively. Define $B(p, r) = \{q : \|p - q\| < r\}$, projection operator $\mathcal{P}_S(p) = \arg\min_{q \in S} \|p - q\|$, the distance from p to S as $d(p, S) = \inf\{\|p-q\| : q \in S\}$, and the normal cone of S at p as $\mathbf{N}_S(p) = \{o \in \mathbb{H}^n : \mathcal{R}(o^{\mathrm{H}}(p - q)) \geq 0, \forall q \in S\}$.

C. NECESSARY DEFINITIONS AND LEMMAS

Through the Hamilton rule (2), it can be concluded that quaternion does not satisfy the general commutative law of multiplication, that is, for any $p_1, p_2 \in \mathbb{H}$, the equation $p_1p_2 = p_2p_1$ is not necessarily hold. In addition, the convex analysis theory in the quaternion field is not complete yet, which makes it difficult to solve QDOPs.

In this case, to solve the nonsmooth QDOP, the gradient or subgradient of quaternion function needs to be considered first. According to [24] and [25], some necessary definitions and lemmas are given. In addition, the subdifferential properties of quaternion-variable functions and the optimality condition for QDOPs are studied as shown in the following.

Definition 1 ([24]): The linear invertible mapping $\Psi : X \to Y$ is defined as follows

1) When $X = \mathbb{H}^n$ and $Y = \mathbb{R}^{4n}$, for any $p \in X$,

$$\Psi(p) = \begin{pmatrix} \mathcal{R}(p) \\ \mathcal{I}^{\mathbf{i}}(p) \\ \mathcal{I}^{\mathbf{j}}(p) \\ \mathcal{I}^{\mathbf{k}}(p) \end{pmatrix} \in Y; \tag{3}$$

2) When $X = \mathbb{H}^{m \times n}$ and $Y = \mathbb{R}^{4m \times 4n}$, for any $P \in X$,

$$\Psi(P) = \begin{pmatrix} \mathcal{R}(P) & -\mathcal{I}^{\mathbf{i}}(P) & -\mathcal{I}^{\mathbf{j}}(P) & -\mathcal{I}^{\mathbf{k}}(P) \\ \mathcal{I}^{\mathbf{i}}(P) & \mathcal{R}(P) & -\mathcal{I}^{\mathbf{k}}(P) & \mathcal{I}^{\mathbf{j}}(P) \\ \mathcal{I}^{\mathbf{j}}(P) & \mathcal{I}^{\mathbf{k}}(P) & \mathcal{R}(P) & -\mathcal{I}^{\mathbf{i}}(P) \\ \mathcal{I}^{\mathbf{k}}(P) & -\mathcal{I}^{\mathbf{j}}(P) & \mathcal{I}^{\mathbf{i}}(P) & \mathcal{R}(P) \end{pmatrix}.$$

It is easy to obtain that Ψ is a bijective mapping and establishes a one-to-one correspondence between quaternion field and real field, so the Definition 1 lays the foundation for the convex analysis theory on quaternion field. Therefore, there are some useful propositions about Ψ.

Proposition 2 ([24]): For any vectors $p, q \in \mathbb{H}^n$, matrix $P \in \mathbb{H}^{m \times n}$, invertible matrix $O \in \mathbb{H}^{n \times n}$ and $\sigma \in \mathbb{R}$, one has Ψ is additive and homogeneous. In addition, it has the following properties

1) $\Psi(P^{\mathrm{H}}) = \Psi(P)^{\mathrm{T}}$, $\mathcal{R}(p^{\mathrm{H}}q) = \Psi(p)^{\mathrm{T}}\Psi(q)$;
2) $\Psi(PO) = \Psi(A)\Psi(O)$, $\Psi(Pp) = \Psi(A)\Psi(p)$;
3) $\Psi(O)^{-1} = \Psi(O^{-1})$;
4) $\|\Psi(p)\| = \|p\|$, $\|\Psi(p)\|_1 = \|p\|_1$;
5) $-\|p\|\|q\| \leq \mathcal{R}(p^{\mathrm{H}}q) \leq \|p\|\|q\|$.

Generally, the objective function (or cost function) of the considered QDOP is usually quaternion-variable (QR) function, that is, its domain is a subset of $\mathbb{H}^n$ and its range is a subset in $\mathbb{R}$. On the basis of the existing work on real-valued distributed optimization problems, considering the quaternion gradient and subgradient (for nonsmooth case) of QR functions is instrumental for practical applications. Thus, the definition of the auxiliary function is shown next for convenience.

Definition 3: For the QR function $g : \mathbb{H}^n \to \mathbb{R}$, the real-valued function $f : \mathbb{R}^{4n} \to \mathbb{R}$ is called the auxiliary function of g, if it satisfies $f(\Psi(p)) = g(p)$ for any $p \in \mathbb{H}^n$.

Lemma 4 ([24]): For the differentiable QR function $g(p) : \mathbb{H} \to \mathbb{R}$ and $p = x + y\mathbf{i} + z\mathbf{j} + w\mathbf{k}$, it follows

1) $\mathrm{d}g(p) = 4\mathcal{R}\left(\frac{\partial g(p)}{\partial p}\mathrm{d}p\right)$;
2) the quaternion gradient of $g(p)$ is

$$\nabla_{\mathbb{H}} g(p) = \partial_x f(x, y, z, w) + \partial_y f(x, y, z, w)\mathbf{i} \\ + \partial_z f(x, y, z, w)\mathbf{j} + \partial_w f(x, y, z, w)\mathbf{k},$$

where f is the auxiliary function of g and $\partial_\iota f(x, y, z, w)$ is the partial derivative of f with respect to $\iota \in \{x, y, z, w\}$.

By introducing the mapping Ψ in Definition 1 and the existing theoretical results of convex analysis of real-valued functions in [28], it is easy to prove that the convexity of the QR function $g : \mathbb{H}^n \to \mathbb{R}$ and its auxiliary function $f : \mathbb{R}^{4n} \to \mathbb{R}$ is equivalent by using the properties of the g and f. For convenience, let

$$\widetilde{S} = \{x \in \mathbb{R}^{4n} : x = \Psi(p), p \in S\} \tag{4}$$

be the auxiliary set of $S \subseteq \mathbb{H}^n$, it also can be obtained that the bounded closed convexity of S is equivalent to that of $\widetilde{S}$, and for any $p \in \mathbb{H}^n$, it follows $d(p, S) = d(\Psi(p), \widetilde{S})$ and $p \in S$ if only if $\Psi(p) \in \widetilde{S}$.

Lemma 5 ([28]): Suppose $f(x) : \mathbb{R}^n \to \mathbb{R}$ is convex and Lipschitz on the convex set $K \subseteq \mathbb{R}^n$, the subdifferential of f at x is $\partial f(x) = \{\xi \in \mathbb{R}^n : f(x + v) - f(x) \geq \xi^{\mathrm{T}}v, \forall v \in \mathbb{R}^n\}$.

Based on Lemma 5, define the partial subdifferential of the convex function $f(x, y)$ at x as

$$\partial_x f(x, y) \\ = \{\xi \in \mathbb{R}^n : f(x + v, y) - f(x, y) \geq \xi^{\mathrm{T}}v, \forall v \in \mathbb{R}^n\}. \tag{5}$$

Definition 6 ([29]): Suppose that $g(p) : \mathbb{H}^n \to \mathbb{R}$ is a convex function, for any $p = x + y\mathbf{i} + z\mathbf{j} + w\mathbf{k} \in \mathbb{H}^n$, the quaternion subdifferential of $g(p)$ is defined as

$$\partial_{\mathbb{H}} g(p) = \partial_x f(x, y, z, w) + \partial_y f(x, y, z, w)\mathbf{i} + \partial_z f(x, y, z, w)\mathbf{j} + \partial_w f(x, y, z, w)\mathbf{k} \quad (6)$$

where $\partial_x f(\cdot)$, $\partial_y f(\cdot)$, $\partial_z f(\cdot)$ and $\partial_w f(\cdot)$ are the partial subdifferential of f with regard to x, y, z and w, respectively.

Remark 7: From equation (6), it is worth pointing out that when the convex QR function g is differentiable, $\partial_{\mathbb{H}} g = \{\nabla_{\mathbb{H}} g\}$ and $\nabla_{\mathbb{H}} g$ is defined in Lemma 4. Therefore, the above result can be regarded as a generalization of the subdifferential of real-valued functions over the quaternion field. Therefore, it is obvious that $\partial_{\mathbb{H}} g(p)$ is a nonempty compact convex set of $\mathbb{H}^n$.

Lemma 8 ([28]): Let $f(x)$ be a convex function at $x = (x_1, x_2, x_3, x_4)$, then

$$\partial f(x) \subseteq \partial_{x_1} f(x_1, x_2, x_3, x_4) \times \partial_{x_2} f(x_1, x_2, x_3, x_4) \times \partial_{x_3} f(x_1, x_2, x_3, x_4) \times \partial_{x_4} f(x_1, x_2, x_3, x_4). \quad (7)$$

Proposition 9: For any $p \in \mathbb{H}^n$, let $g(p)$ be convex QR function and f be the auxiliary function of g, then

$$\Psi(\partial_{\mathbb{H}} g(p)) = \partial f(\Psi(p)).$$

Proposition 10: For the convex QR function $g : \mathbb{H}^n \to \mathbb{R}$, $\forall\, p, q \in \mathbb{H}^n$, we have

$$\partial_{\mathbb{H}} g(p) = \{\eta \in \mathbb{H}^n : g(q) - g(p) \geq \mathcal{R}(\eta^{\mathrm{H}}(q - p))\},$$

and

$$\mathcal{R}((\eta - \varsigma)^{\mathrm{H}}(p - q)) \geq 0$$

where $\eta \in \partial_{\mathbb{H}} g(p)$, $\varsigma \in \partial_{\mathbb{H}} g(q)$.

Proposition 11: Suppose that $S, Q \subseteq \mathbb{H}^n$ are nonempty closed convex sets, $g : \mathbb{H}^n \to \mathbb{R}$,

1) if $g(p)$ is local Lipschitz and attains a minimum over S at $p \in S$, then $0 \in \partial_{\mathbb{H}} g(p) + \mathbf{N}_S(p)$;
2) if $0 \in \mathrm{int}(S - Q)$, then for any $p \in S \cap Q$, $\mathbf{N}_{S \cap Q}(p) = \mathbf{N}_S(p) + \mathbf{N}_Q(p)$.

The proofs of Proposition 9, 10, and 11 can be seen in Appendix- A, B, and C, respectively.

As described in [25], the differential inclusion is expressed as

$$\dot{p}(t) \in F(p(t)), \quad t \geq 0 \quad (8)$$

where F is a set-valued map that for any $p \in \mathbb{H}^n$, there is always a corresponding set satisfying $F(p) \subseteq \mathbb{H}^n$. If there is a measurable function $\eta(t) \in F(p(t))$ such that $\dot{p}(t) = \eta(t)$ for a.e. $t \in [0, T]$ where $T > 0$, then $p(t)$ is a local solution of (8). Moreover, if $V(p) : \mathbb{H}^n \to \mathbb{R}$ is a locally Lipschitz function, $p(t) : [0, +\infty) \to \mathbb{H}^n$ is absolutely continuous on any compact interval of $[0, +\infty)$, then for *a.e.* $t \in [0, +\infty)$, the derivative of $V(p)$ with respect to t along (8) is $\dot{V}(p(t)) = \{\xi \in \mathbb{R} : \exists \nu \in F(p(t))$ such that $4\mathcal{R}(\eta^{\mathrm{H}}\nu) = \xi,\ \forall\, \eta \in \partial_{\mathbb{H}} V(p)\}$ from Lemma 4. Then, the following is the invariance principle in the quaternion field for nonsmooth functions.

Some important conclusions in this section are derived from works of literature [20], [21], [24], and [25], which provide a theoretical basis for us to establish the QNA for solving QDOPs of the form (1) and prove its convergence. In addition, the convex QDOP (1) is meaningful for various practice areas, like color image processing, satellite tracking, and bearings-only tracking.

III. PROBLEM DESCRIPTION

Consider a network consisting of N agents, and the communication topology between agents is described as $\mathcal{G}(\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, 2, \cdots, N\}$. When graph $\mathcal{G}$ is a connected graph and p_i is used to represent the state of agent i, the convex QDOP (1) is equivalent to the following QDOP

$$\begin{aligned} \min\ & \boldsymbol{g}(\boldsymbol{p}) \triangleq \sum_{i=1}^{N} g_i(p_i) \\ \text{s.t. } & \boldsymbol{h}(\boldsymbol{p}) \leq \mathbf{0} \\ & \boldsymbol{D}\boldsymbol{p} = \boldsymbol{d} \\ & \boldsymbol{L}\boldsymbol{p} = \mathbf{0} \end{aligned} \quad (9)$$

where $\boldsymbol{p} = \mathrm{col}\{p_1, p_2, \cdots, p_N\} \in \mathbb{H}^{Nn}$, the global objective function $\boldsymbol{g}(\boldsymbol{p}) : \mathbb{H}^{Nn} \to \mathbb{R}$ is the sum of the local objective functions $g_i(p_i)$ $(i \in \mathcal{V})$, $\boldsymbol{h}(\boldsymbol{p}) = \mathrm{col}\{h_1(p_1), h_2(p_2), \cdots, h_N(p_N)\} : \mathbb{H}^{Nn} \to \mathbb{R}^m$ with $m = \sum_{i=1}^{N} m_i$, $\boldsymbol{D} = \mathrm{diag}(D_1, D_2, \cdots, D_N)$, $\boldsymbol{d} = \mathrm{col}(d_1, d_2, \cdots, d_N)$, $\boldsymbol{L} = L \otimes I_n$ and L is the Laplacian matrix of the communication topology $\mathcal{G}$.

For convenience, note

$$X_1 = \{\boldsymbol{p} \in \mathbb{H}^{Nn} : \boldsymbol{h}(\boldsymbol{p}) \leq \mathbf{0}\} = \prod_{i=1}^{N} X_1^i,$$

$$X_2 = \{\boldsymbol{p} \in \mathbb{H}^{Nn} : \boldsymbol{D}\boldsymbol{p} = \boldsymbol{d}\} = \prod_{i=1}^{N} X_2^i,$$

$$X_3 = \{\boldsymbol{p} = \mathbf{1}_N \otimes p \in \mathbb{H}^{Nn} : p \in \mathbb{H}^n\},$$

where $X_1^i = \{p \in \mathbb{H}^n : h_i(p) \leq \mathbf{0}\}$ and $X_2^i = \{p \in \mathbb{H}^n : D_i p = d_i\}$. Then the feasible region of QDOP (9) is denoted as

$$X = X_1 \cap X_2 \cap X_3.$$

In this paper, we suppose that $X \neq \emptyset$ and the optimal solution set $O \subseteq \mathbb{H}^n$ of QDOP (9) is nonempty.

Assumption 12:

1) There are $\hat{p} \in \mathrm{int}(X_1^i) \cap X_2^i$ and $r_i > 0$ such that $X_1^i \subseteq B(\hat{p}, r_i)$ for every $i \in \mathcal{V}$.
2) For every $i \in \mathcal{V}$, $h_i = \mathrm{col}\{h_{i1}, h_{i2}, \cdots, h_{im_i}\}$, $h_{ij} : \mathbb{H}^n \to \mathbb{R}$ $(j = 1, 2, \cdots, m_i)$ are convex QR functions but may be nonsmooth, $D_i \in \mathbb{H}^{s_i \times n}$ is full row-rank.
3) For every $i \in \mathcal{V}$, $g_i : \mathbb{H}^n \to \mathbb{R}$ is convex but may be nonsmooth, and g_i is Lipschitz.

Assumption 13: The communication topology $\mathcal{G}$ is an undirected connected graph.

Lemma 14: Under Assumption 12, for the global function $\boldsymbol{g}(\boldsymbol{p}) = \sum_{i=1}^{N} g_i(p_i)$ defined as (9), we have

$$\partial_{\mathbb{H}}\boldsymbol{g}(\boldsymbol{p}) = \mathrm{col}(\partial_{\mathbb{H}}g_1(p_1), \partial_{\mathbb{H}}g_2(p_2), \cdots, \partial_{\mathbb{H}}g_N(p_N)). \tag{10}$$

The proof of Lemma 14 can be seen in Appendix-D.

Compared with Theorem 2 in [25], it is obvious that Lemma 14 above relaxes the assumptions of the strong convexity of the local objective functions and $\partial_{\mathbb{H}}\boldsymbol{g}(\boldsymbol{p}) \subset \mathrm{col}(\partial_{\mathbb{H}}g_1(p_1), \partial_{\mathbb{H}}g_2(p_2), \cdots, \partial_{\mathbb{H}}g_N(p_N))$.

Lemma 15 ([24]): Under Assumption 12, let $\mathcal{I}_i \triangleq \{1, 2, \cdots, m_i\}$, then $U_i(p) \triangleq \sum_{k=1}^{m_i} \max\{0, h_{ik}(p)\}$ is convex and

$$\partial_{\mathbb{H}}U_i(p) = \begin{cases} \sum\limits_{k\in\mathcal{I}_i^+(p)} \partial_{\mathbb{H}}h_{ik}(p) \\ \quad + \sum\limits_{k\in\mathcal{I}_i^0(p)} [0,1]\partial_{\mathbb{H}}h_{ik}(p), & p \in \mathbb{H}^n\backslash X_1^i \\ \sum\limits_{k\in\mathcal{I}_i^0(p)} [0,1]\partial_{\mathbb{H}}h_{ik}(p), & p \in \mathrm{bd}(X_1^i) \\ \{\mathbf{0}\}, & p \in \mathrm{int}(X_1^i) \end{cases} \tag{11}$$

where $\mathcal{I}_i^+(p) = \{k \in \mathcal{I}_i : h_{ik}(p) > 0\}$ and $\mathcal{I}_i^0(p) = \{k \in \mathcal{I}_i : h_{ik}(p) = 0\}$.

By means of the exact penalty method, the convex QDOP (9) can be equivalently transformed as follows.

Theorem 16: Under Assumptions 12 and 13, $\boldsymbol{p}^* \in \mathbb{H}^{Nn}$ is an optimal solution of QDOP (9) if and only if there is $a_0 > 0$ such that $\boldsymbol{p}^* \in \mathbb{H}^{Nn}$ is an optimal solution of the following QDOP with $u_i > a_0$ $(i \in \mathcal{V})$

$$\begin{aligned} \min \boldsymbol{G}(\boldsymbol{p}) &\triangleq \sum_{i=1}^{N} g_i(p_i) + \sum_{i=1}^{N} u_iU_i(p_i) \\ \text{s.t. } \boldsymbol{D}\boldsymbol{p} &= \boldsymbol{d} \\ \boldsymbol{L}\boldsymbol{p} &= \mathbf{0} \end{aligned} \tag{12}$$

where $U_i(p_i)$ is from Lemma 15.

Proof: We prove this theorem from two aspects.

Necessity: If $\boldsymbol{p}^*$ is an optimal solution to the QDOP (9), then there exists $p^* \in \mathbb{H}^n$ such that $\boldsymbol{p}^* = I_N \otimes p^*$ and p^* is an optimal solution of the QDOP (1), that is $p^* \in X_1^i \cap X_2^i$. Based on Assumption 12, $\mathbf{N}_{X_1^i\cap X_2^i}(p^*) = \mathbf{N}_{X_1^i}(p^*) + \mathbf{N}_{X_2^i}(p^*)$. Due to Proposition 11, we can find $\theta_{ik} \in [0, +\infty)$ satisfying

$$0 \in \sum_{i=1}^{N} \partial_{\mathbb{H}}g_i(p^*) + \sum_{i=1}^{N} \sum_{k\in\mathcal{I}_i^0(p^*)} \theta_{ik}\partial_{\mathbb{H}}h_{ik}(p^*) + \sum_{i=1}^{N} \mathbf{N}_{X_2^i}(p^*).$$

Then there exist $-\xi^* \in \sum_{i=1}^{N} \partial_{\mathbb{H}}g_i(p^*)$, $\eta_{ik}^* \in \partial_{\mathbb{H}}h_{ik}(p^*)$, and $\gamma^* \in \sum_{i=1}^{N} \mathbf{N}_{X_2^i}(p^*)$, such that $\xi^* = \sum_{i=1}^{N} \sum_{k\in\mathcal{I}_i^0(p^*)} \theta_{ik}\eta_{ik}^* + \gamma^*$. We claim that for any $k \in \mathcal{I}_i^0(p^*)$ $(i \in \mathcal{V})$,

$$\theta_{ik} \le Nlr/\hat{h} \triangleq a_0 \tag{13}$$

where $l = \max_{i\in\mathcal{V}} l_i$, l_i is the Lipschitz constant of g_i, $r = \max_{i\in\mathcal{V}} r_i$, $\hat{h} = -\max\{\boldsymbol{h}_1(\hat{\boldsymbol{p}}), \boldsymbol{h}_2(\hat{\boldsymbol{p}}), \boldsymbol{h}_m(\hat{\boldsymbol{p}})\}$, $\boldsymbol{h}_k(\boldsymbol{p})$ is the kth component of $\boldsymbol{h}(\boldsymbol{p}) = \mathrm{col}\,\{h_1(p_1), \cdots, h_N(p_N)\} : \mathbb{H}^{Nn} \to \mathbb{R}^m$ and $\hat{\boldsymbol{p}} = \mathbf{1} \otimes \hat{p}$.

When $\mathcal{I}_i^0(p^*) = \emptyset$ $(i \in \mathcal{V})$, it is obvious that $\theta_{ik} = 0$ satisfies (13). On the other hand, when there is $\tilde{i} \in \mathcal{V}$ such that $\mathcal{I}_{\tilde{i}}^0(p^*) \neq \emptyset$, we next prove (13) still holds. If (13) does not hold, there is at least one $\tilde{k} \in \mathcal{I}_{\tilde{i}}^0(p^*)$ such that $\theta_{\tilde{i}\tilde{k}} > a_0$. Thus, by the convexity of $h_{ik}(p)$, it has

$$\begin{aligned} &\langle -\xi^*, p^* - \hat{p}_i\rangle \\ &= \sum_{i=1}^{N} \sum_{k\in\mathcal{I}_i^0(p^*)} \theta_{ik}\langle\eta_{ik}^*, p^* - \hat{p}_i\rangle + \langle\gamma^*, p^* - \hat{p}_i\rangle \\ &\ge \sum_{i=1}^{N} \sum_{k\in I^0(p^*)} \theta_{ik}(h_{ik}(p^*) - h_{ik}(\hat{p}_i)) \ge \theta_{\tilde{i}\tilde{k}}\hat{h} > Nlr, \end{aligned}$$

which implies that $Nl < \|\xi^*\|$ by Assumption 12 and contradicts $\|\xi^*\| < Nl$. Therefore, when $u_i \ge a_0$,

$$\begin{aligned} \sum_{k\in\mathcal{I}_i^0(p^*)} \theta_{ik}\partial_{\mathbb{H}}h_{ik}(p^*) &= u_i \sum_{k\in\mathcal{I}_i^0(p^*)} \frac{\theta_{ik}}{u_i}\partial_{\mathbb{H}}h_{ik}(p^*) \\ &\subseteq u_i \sum_{k\in\mathcal{I}_i^0(p^*)} [0,1]\partial_{\mathbb{H}}h_{ik}(p^*). \end{aligned}$$

Therefore, from (11), it is easy to obtain

$$\sum_{k\in\mathcal{I}_i^0(p^*)} \theta_{ik}\partial_{\mathbb{H}}h_{ik}(p^*) \subseteq u_i\partial_{\mathbb{H}}U_i(p^*)$$

for any $u_i \ge a_0$. As a result, we have

$$\mathbf{0} \in \sum_{i=1}^{N} \partial_{\mathbb{H}}g_i(p^*) + \sum_{i=1}^{N} \sum_{k\in\mathcal{I}_i^0(p^*)} u_i\partial_{\mathbb{H}}U_i(p^*) + \sum_{i=1}^{N} \mathbf{N}_{X_2^i}(p^*)$$

for all $u_i \ge a_0$. That is $\boldsymbol{p}^* = \mathrm{col}(p^*, p^*, \cdots, p^*) \in \mathbb{H}^{Nn}$ is an optimal solution of QDOP (12) from Proposition 11, which means that the optimal solution to (9) must be the optimal solution to (12) with $u_i > a_0$.

Sufficiency: If $\boldsymbol{p}^*$ is an optimal solution to QDOP (12). Without loss of generality, assume that $\tilde{\boldsymbol{p}}$ is an optimal solution to QDOP (9). Then $\boldsymbol{p}^* \in X_2 \cap X_3$, and $\tilde{\boldsymbol{p}} = \mathrm{col}(\tilde{p}, \tilde{p}, \cdots, \tilde{p}) \in \mathbb{H}^{Nn}$ is also an optimal solution to the following QDOP

$$\begin{aligned} &\min \boldsymbol{g}(\boldsymbol{p}) + a_0U(\boldsymbol{p}) \\ &\text{s.t. } \boldsymbol{p} \in X_2 \cap X_3. \end{aligned}$$

where $a_0 = lr/\hat{h}$ and $U(\boldsymbol{p}) = \sum_{i=1}^{N} U_i(p_i)$. Thereby,

$$\sum_{i=1}^{N} g_i(\tilde{p}) + a_0\sum_{i=1}^{N} U_i(\tilde{p}) \le \sum_{i=1}^{N} g_i(p^*) + a_0\sum_{i=1}^{N} U_i(p^*). \tag{14}$$

In addition, by the choice of $\boldsymbol{p}^*$, there is a $p^* \in \mathbb{H}^n$ such that $\boldsymbol{p}^* = \mathrm{col}(p^*, p^*, \cdots, p^*) \in \mathbb{H}^{Nn}$ satisfies

$$\sum_{i=1}^{N} g_i(p^*) + \sum_{i=1}^{N} u_i U_i(p^*) \leq \sum_{i=1}^{N} g_i(\tilde{p}) + \sum_{i=1}^{N} u_i U_i(\tilde{p}) \tag{15}$$

when $u_i > a_0$. By combining (14) and (15), we get that

$$\sum_{i=1}^{N} (u_i - a_0) U_i(p^*) \leq \sum_{i=1}^{N} (u_i - a_0) U_i(\tilde{p}) = 0.$$

From the definition of $U_i(p)$, $p \in X_1^i$ if and only if $U_i(p) = 0$. As a result, it is clear that $p^* \in X_1^i$ for all $i \in \mathcal{V}$. Then it implies that $\boldsymbol{g}(\boldsymbol{p}^*) = \sum_{i=1}^{N} g_i(p^*) = \sum_{i=1}^{N} g_i(p^*) + \sum_{i=1}^{N} u_i U_i(p^*) = \boldsymbol{G}(\boldsymbol{p}^*) \leq \boldsymbol{G}(\boldsymbol{p}) = \boldsymbol{g}(\boldsymbol{p})$, $\forall \boldsymbol{p} \in X_1 \cap X_2 \cap X_3$. Thus $\boldsymbol{p}^*$ is an optimal solution to the QDOP (9), which means that the optimal solution to (12) must be the optimal solution to (9). ∎

Remark 17: It can be known from Theorem 16 that when establishing the equivalence between QDOP (9) and QDOP (12), the penalty parameters need to be larger than a uniform lower bound. However, calculating the lower bound of the penalty parameters will involve global information and cause additional computational burden. Therefore, introducing adaptive controllers that can be self-adjusted has the following advantages.

1) Drive the state solution into the constraint set;
2) Make the adaptive control terms remain constant after entering the constraint set to avoid penalty parameters becoming ill-conditioned;
3) Reduce the amount of calculation, and avoid using the global information of the multi-agent system.

IV. MAIN RESULTS

In this section, a QDOP (1) with affine equality and inequality constraints is solved. The proposed QNA is based on a multi-agent system with adaptive controllers under undirected and connected topology, and it is theoretically proved that the state solution exists globally and converges to an optimal solution to the QDOP.

Next, benefiting from the Theorem 16, we propose the following QNA for agent $i \in \mathcal{V}$ in the multi-agent system in order to solve the QDOP (9),

$$\begin{cases} \dot{p}_i(t) \in -\alpha(t)(I - Q_i)\big(\partial_{\mathrm{H}} g_i(p_i(t)) + u_i(t)\partial_{\mathrm{H}} U_i(p_i(t)) \\ \qquad + \sum_{j=1}^{N} a_{ij}(z_i(t) - z_j(t))\big) - \sum_{j=1}^{N} a_{ij}(p_i(t) - p_j(t)) \\ \qquad - \partial_{\mathrm{H}} \|D_i p_i(t) - d_i\|_1 \\ \dot{u}_i(t) = U_i(p_i(t)), \ u_i(0) > 0 \\ \dot{z}_i(t) = \alpha(t) \sum_{j=1}^{N} a_{ij}(p_i(t) - p_j(t)), \end{cases} \tag{16}$$

where p_i is the estimate of the optimal solution for agent i, u_i is the adaptive penalty variable of agent i and z_i is auxiliary variable of agent i, I is the identity matrix and $Q_i = D_i^{\mathrm{H}}(D_i D_i^{\mathrm{H}})^{-1} D_i$ is the projection matrix and $K_i(p_i) = \|\Psi(D_i)\Psi(p_i) - \Psi(d_i)\|_1 \triangleq B_i(\Psi(p_i))$. Let t_{X_2} be the latest time to enter constraint set X_2 and define $\alpha(t)$ as

$$\alpha(t) = \begin{cases} 0, & t < t_{X_2} \\ 1, & t \geq t_{X_2}. \end{cases} \tag{17}$$

From the definition of $B_i(\cdot)$, it holds

$$\partial B_i(\Psi(p_i)) = \Psi(D_i)^{\mathrm{T}} h_{[-1,1]}(\Psi(D_i)\Psi(p_i) - \Psi(d_i))$$

where $h_{[-1,1]}(x) = (h_{[-1,1]}(x_1), \cdots, h_{[-1,1]}(x_n))$ for $x \in \mathbb{R}^n$ and

$$h_{[-1,1]}(x_k) = \begin{cases} 1, & x_k > 0 \\ [-1, 1], & x_k = 0 \\ -1, & x_k < 0 \end{cases} \tag{18}$$

for $k = 1, 2, \cdots, n$. Let

$$\psi_{[-1,1]}(p) = h_{[-1,1]}(\mathcal{R}(p)) + h_{[-1,1]}(\mathcal{I}^{\mathbf{i}}(p))\mathbf{i} + h_{[-1,1]}(\mathcal{I}^{\mathbf{j}}(p))\mathbf{j} + h_{[-1,1]}(\mathcal{I}^{\mathbf{k}}(p))\mathbf{k},$$

then it follows from Proposition 9 that $\partial_{\mathrm{H}} K_i(p_i) = D_i^{\mathrm{H}} \psi_{[-1,1]}(D_i p_i - d_i)$.

It is worth pointing out that penalty parameters u_i in Theorem 16 are local information for agents, which can guarantee the complete distribution properties of the QNA (16). Actually, the role of the term $\partial_{\mathrm{H}} g_i(p_i(t))$ is to minimize the objective function $g_i(p_i(t))$, $u_i(t)\partial_{\mathrm{H}} U_i(p_i(t))$ is to ensure that the state solution enters inequality constraint set X_1^i and $\alpha(t)(I - Q_i)$ as well as $-\partial_{\mathrm{H}}\|D_i p_i(t) - d_i\|_1$ are to ensure the state solution enters equality constraint set X_2^i, while $-\sum_{j=1}^{N} a_{ij}(z_i(t) - z_j(t))$ and $z_i(t)$ play the role in driving all agents to reach the consensus.

Denote $\boldsymbol{I} = I \otimes I_N$, $\boldsymbol{Q} = \mathrm{diag}(Q_1, Q_2, \cdots, Q_N)$, $\boldsymbol{u} = \mathrm{col}(u_1, u_2, \cdots, u_N)$, $\boldsymbol{z} = \mathrm{col}(z_1, z_2, \cdots, z_N)$, $\mathrm{diag}(u) = \mathrm{diag}(u_1, u_2, \cdots, u_N)$, $\mathrm{diag}(\boldsymbol{u}) = \mathrm{diag}(u) \otimes I_n$, and $W(\boldsymbol{p}) = \mathrm{col}(U_1(p_1), U_2(p_2), \cdots, U_N(p_N))$. Let $K(\boldsymbol{p}) = \sum_{i=1}^{N} K_i(p_i)$ and $U(\boldsymbol{p}) = \sum_{i=1}^{N} U_i(p_i)$, then from (10), $\partial_{\mathrm{H}} U(\boldsymbol{p}(t))$ and $\partial_{\mathrm{H}} K(\boldsymbol{p}(t))$ can be computed similarly. Therefore, the QNA (16) can be rewritten in a compact form

$$\begin{cases} \dot{\boldsymbol{p}}(t) \in -\alpha(t)(\boldsymbol{I} - \boldsymbol{Q})\big(\partial_{\mathrm{H}} \boldsymbol{g}(\boldsymbol{p}(t)) + \mathrm{diag}(\boldsymbol{u}(t)) \times \\ \qquad \partial_{\mathrm{H}} U(\boldsymbol{p}(t)) + \boldsymbol{L}\boldsymbol{z}(t)\big) - \boldsymbol{L}\boldsymbol{p}(t) - \partial_{\mathrm{H}} K(\boldsymbol{p}(t)) \\ \dot{\boldsymbol{u}}(t) = W(\boldsymbol{p}(t)), \ \boldsymbol{u}(0) > 0 \\ \dot{\boldsymbol{z}}(t) = \alpha(t)\boldsymbol{L}\boldsymbol{p}(t). \end{cases} \tag{19}$$

Remark 18: As we all know, the Karush-Kuhn-Tucker (KKT) condition and negative gradient systems play an important role in designing continuous-time neurodynamic approaches such as [12]–[18], [24]–[27], and the Lyapunov stability theory is always used to study their convergence performance. As the importance of quaternions gradually emerging in real life, this paper is no longer limited to solving real-valued convex optimization problems, but proposes a novel QNA for solving QDOPs with inequality and affine equality constraints by means of the adaptive controller $\boldsymbol{u}(t)$. From the dynamic properties of $\boldsymbol{u}(t)$, it can be seen that when the state solution doesn't enter the constraint set, $\boldsymbol{u}(t)$ will increase to intensify the penalty for the points outside the constraint set, and $\boldsymbol{u}(t)$ stops increasing until the state solution enters the constraint set. This not only ensures that the state solution advances in the direction of the feasible region, but also avoids the occurrence of infinite variables that may cause the system to collapse.

In addition, when the imaginary part of the quaternion takes zero, the QNA proposed in this paper can be simplified to a real-valued neurodynamic approach. According to the dynamic behavior of the adaptive controller that can be adjusted with its own state, it is capable of avoiding the calculation of the lower bound of the penalty parameter while solving the optimization problem with constraints, which makes the QNA in this paper also have the competitive advantage of reducing the amount of calculation than ones in [16], [24]–[27] when solving both convex QDOPs and convex real-valued distributed optimization problems.

Theorem 19: Under Assumption 12, there is $T_{X_2} > 0$ such that the state solution $(\boldsymbol{p}(t), \boldsymbol{u}(t), \boldsymbol{z}(t))$ of QNA (19) from the initial solution $(\boldsymbol{p}(0), \boldsymbol{u}(0), \boldsymbol{z}(0)) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$, fulfills $\boldsymbol{p}(t) \in X_2$ when $t \geq T_{X_2} > 0$.

Proof: For $K(\boldsymbol{p}) = B(\Psi(\boldsymbol{p})) = \|\Psi(\boldsymbol{D})\Psi(\boldsymbol{p}) - \Psi(\boldsymbol{d})\|_1$, it is easy to obtain that there are $\xi(t) \in \partial_{\mathrm{H}} K(\boldsymbol{p}(t))$ and $\eta(t) \in \psi_{[-1,1]}(\boldsymbol{D}\boldsymbol{p}(t) - \boldsymbol{d})$ satisfying $\xi(t) = \boldsymbol{D}^{\mathrm{H}}\eta(t)$, thus $\Psi(\xi(t)) \in \partial B(\Psi(\boldsymbol{p}(t)))$ and $\Psi(\xi(t)) = \Psi(\boldsymbol{D})^{\mathrm{T}}\Psi(\eta(t)) \in \partial B(\Psi(\boldsymbol{p}(t)))$. From Proposition 9, it has

$$\frac{\mathrm{d}}{\mathrm{d}t} B(\Psi(\boldsymbol{p}(t))) = \langle \Psi(\boldsymbol{D})^{\mathrm{T}}\Psi(\eta(t)), \Psi(\dot{\boldsymbol{p}}(t))\rangle.$$

Since $\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{Q}) = \boldsymbol{0}$, if $\boldsymbol{p}(t) \in \mathbb{H}^{Nn}\backslash X_2$ for any $t \geq 0$ and $\boldsymbol{p}_0 = \boldsymbol{p}(0) \in \mathbb{H}^{Nn}$, we have $\Psi(\boldsymbol{p}(t)) \in \mathbb{R}^{4Nn}\backslash\widetilde{X}_2$, where $\widetilde{X}_2$ is defined as same as $\widetilde{S}$ in (4). Therefore, it follows from the definition of Ψ and QNA (19) that

$$\Psi(\boldsymbol{D})\Psi(\dot{\boldsymbol{p}}(t)) = -\Psi(\boldsymbol{D})\partial B(\Psi(\boldsymbol{p}(t))).$$

Obviously, from $\Psi(\boldsymbol{p}(t)) \in \mathbb{R}^{4Nn}\backslash\widetilde{X}_2$, $\Psi(\boldsymbol{D})\Psi(\boldsymbol{p}) - \Psi(\boldsymbol{d}) \neq \boldsymbol{0}$ holds. Then it obtains $\|\Psi(\eta(t))\| \geq 1$ and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} &B(\Psi(\boldsymbol{p}(t)))\\ &= -\Psi(\eta(t))^{\mathrm{T}}\Psi(\boldsymbol{D})\Psi(\boldsymbol{D})^{\mathrm{T}}\Psi(\eta(t))\\ &\leq -\lambda_{min}\|\Psi(\eta(t))\|^2 \leq -\lambda_{min}\end{aligned}$$

where $\lambda_{min} > 0$ is the minimum eigenvalue of $\Psi(\boldsymbol{D})\Psi(\boldsymbol{D})^{\mathrm{T}}$. Through integrating the above inequality from $t_0 = 0$ to t, we have

$$B(\Psi(\boldsymbol{p}(t))) \leq B(\Psi(\boldsymbol{p}(0))) - \lambda_{min}t.$$

Hence, when $t = \frac{\|\Psi(\boldsymbol{D}\Psi(\boldsymbol{p}(0))) - \Psi(\boldsymbol{d})\|_1}{\lambda_{min}} \triangleq T_{X_2}$, $B(\Psi(\boldsymbol{p}(t))) = 0$, namely, $\Psi(\boldsymbol{p}(t)) \in \widetilde{X}_2$ if and only if $\boldsymbol{p}(t) \in X_2$. If there is $t_1 \geq T_{X_2}$ such that $\boldsymbol{p}(t_1) \in \mathrm{bd}(X_2)$ and $\boldsymbol{p}(t) \notin X_2$ for $t \in (t_1, t_2)$, then $K(\boldsymbol{p}(t_1)) = B(\Psi(\boldsymbol{p}(t_1))) = 0$. It can be known from the previous proof that $B(\Psi(\boldsymbol{p}(t))) < 0$ when $t \in (t_1, t_2)$, which is a contradiction with $B(\Psi(\boldsymbol{p})) \geq 0$. Therefore, when $t \geq T_{X_2}$, one has $\boldsymbol{p}(t) \in X_2$. ■

Before revealing the correspondence between (9) and (19), we assume $(\boldsymbol{p}^*, \boldsymbol{u}^*, \boldsymbol{z}^*) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$ is an equilibrium solution of QNA (19), that is,

$$\begin{cases}\boldsymbol{0} \in -(\boldsymbol{I} - \boldsymbol{Q})\big(\partial_{\mathrm{H}}\boldsymbol{g}(\boldsymbol{p}^*) + \mathrm{diag}(\boldsymbol{u}^*)\partial_{\mathrm{H}}U(\boldsymbol{p}^*) + \boldsymbol{L}\boldsymbol{z}^*\big)\\ \qquad -\partial_{\mathrm{H}}K(\boldsymbol{p}^*)\\ \boldsymbol{0} = W(\boldsymbol{p}^*)\\ \boldsymbol{0} = \boldsymbol{L}\boldsymbol{p}^*.\end{cases} \tag{20}$$

Theorem 20: Under Assumptions 12 and 13, if $(\boldsymbol{p}^*, \boldsymbol{u}^*, \boldsymbol{z}^*) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$ is an equilibrium solution of QNA (19), then $\boldsymbol{p}^* \in \mathbb{H}^{Nn}$ is an optimal solution to the QDOP (9). Conversely, if $\boldsymbol{p}^* \in \mathbb{H}^{Nn}$ is an optimal solution to the QDOP (9), then there are $\boldsymbol{u}^* \in \mathbb{R}^N_{>0}$ and $\boldsymbol{z}^* \in \mathbb{H}^{Nn}$ such that $(\boldsymbol{p}^*, \boldsymbol{u}^*, \boldsymbol{z}^*) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$ is an equilibrium solution of QNA (19) when $\boldsymbol{u}^* > a_0 \otimes \boldsymbol{1}_N$.

Proof: First, assume $(\boldsymbol{p}^*, \boldsymbol{u}^*, \boldsymbol{z}^*) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$ is an equilibrium solution of QNA (19), then from (20) that $\boldsymbol{p}^* \in X_3$ and $U_i(p^*) = 0$ for all $i \in \mathcal{V}$. Thus $\boldsymbol{p}^* = \mathrm{col}(p^*, p^*, \cdots, p^*) \in X_1 \cap X_3$. From the conclusion of Theorem 19, it is obvious that $\boldsymbol{p}^* \in X_1 \cap X_2 \cap X_3$ and $\alpha^* = 1$.

Based on (20) and Theorem 19, it follows that there is $\xi^* \in \partial_{\mathrm{H}}\boldsymbol{g}(\boldsymbol{p}^*)$, $\zeta^* \in \partial_{\mathrm{H}}U(\boldsymbol{p}^*)$, $\eta^*_{ik} \in \psi_{[-1,1]}(\boldsymbol{D}\boldsymbol{p}^* - \boldsymbol{d})$ such that

$$-(\boldsymbol{I} - \boldsymbol{Q})[\xi^* + (\mathrm{diag}(\boldsymbol{u}^*))\zeta^* + \boldsymbol{L}\boldsymbol{z}^*] - \boldsymbol{D}^{\mathrm{H}}\eta^*_{ik} = \boldsymbol{0}. \tag{21}$$

Multiplying (21) by $\boldsymbol{D}$ from the left, it deduces that $\boldsymbol{D}\boldsymbol{D}^{\mathrm{H}}\eta^*_{ik} = \boldsymbol{0}$ due to $\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{Q}) = \boldsymbol{0}$. Therefore, we have $\eta^*_{ik} = \boldsymbol{0}$ from the Assumption 12. Then (21) degenerates as follows

$$-(\boldsymbol{I} - \boldsymbol{Q})[\xi^* + (\mathrm{diag}(\boldsymbol{u}^*) \otimes I_n)\zeta^* + \boldsymbol{L}\boldsymbol{z}^*] = \boldsymbol{0}.$$

According to the definition of Ψ and Proposition 9, we have

$$\begin{aligned}&-(\Psi(\boldsymbol{I}) - \Psi(\boldsymbol{Q}))\\ &\quad\times[\Psi(\xi^*) + (\mathrm{diag}(\boldsymbol{u}^*) \otimes I_n)\Psi(\zeta^*) + \Psi(\boldsymbol{L})\Psi(\boldsymbol{z}^*)] = \boldsymbol{0}\end{aligned}$$

where $\Psi(\xi^*)$ and $\Psi(\zeta^*)$ is a subgradient of the corresponding auxiliary function $\boldsymbol{f}(\Psi(\boldsymbol{p})) = \boldsymbol{g}(\boldsymbol{p})$ and $J(\Psi(\boldsymbol{p})) = U(\boldsymbol{p})$, respectively.

Let $\varrho^* = \xi^* + (\mathrm{diag}(\boldsymbol{u}^*) \otimes I_n)\zeta^* + \boldsymbol{L}\boldsymbol{z}^*$ and $y^* = -(\boldsymbol{D}\boldsymbol{D}^{\mathrm{H}})^{-1}\boldsymbol{D}\varrho^*$, then

$$\varrho^* = \boldsymbol{Q}\varrho^* = \boldsymbol{D}^{\mathrm{H}}(\boldsymbol{D}\boldsymbol{D}^{\mathrm{H}})^{-1}\boldsymbol{D}\varrho^* = -\boldsymbol{D}^{\mathrm{H}}y^*.$$

Since $\boldsymbol{Z}(\boldsymbol{p}) = \boldsymbol{g}(\boldsymbol{p}) + \mathrm{diag}(\boldsymbol{u}^*)U(\boldsymbol{p})$ is a convex quaternion-valued function and the graph $\mathcal{G}$ is undirected, so for any $\boldsymbol{p} \in X_1 \cap X_2 \cap X_3$, it holds

$$\begin{aligned}\boldsymbol{Z}(\boldsymbol{p}) - \boldsymbol{Z}(\boldsymbol{p}^*) &\geq \mathcal{R}((\boldsymbol{p} - \boldsymbol{p}^*)^{\mathrm{H}}(\varrho^* - \boldsymbol{L}\boldsymbol{z}^*)\\ &= -\mathcal{R}((\boldsymbol{p} - \boldsymbol{p}^*)^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}y^* + (\boldsymbol{p} - \boldsymbol{p}^*)^{\mathrm{H}}\boldsymbol{L}^{\mathrm{H}}\boldsymbol{z}^*) = 0.\end{aligned}$$

Thus,

$$\boldsymbol{p}^* = \arg \min_{\boldsymbol{p} \in X_1 \cap X_2 \cap X_3} \boldsymbol{Z}(\boldsymbol{p})$$

$$= \arg \min_{p \in X_1 \cap X_2 \cap X_3} g(p),$$

that is, p^* is an optimal solution to the QDOP (9).

Conversely, if p^* is an optimal solution to the QDOP (9), then $U_i(p_i^*) = 0$, $D_i p_i^* - d_i = \mathbf{0}$ and $p_i^* = p_j^*$ for all $i, j \in \mathcal{V}$, i.e. $W(p^*) = \mathbf{0}$ and $Lp^* = \mathbf{0}$. From Theorem 16, there is $a_0 > 0$ such that

$$p^* = \arg \min_{p \in X_2 \cap X_3} \{\sum_{i=1}^{N} g_i(p_i) + \sum_{i=1}^{N} u_i U_i(p_i)\}.$$

Thus, take $u^* = \mathrm{col}(u_1^*, u_2^*, \cdots, u_N^*)$ with $u_i^* > a_0$ $(i \in \mathcal{V})$, it holds that

$$p^* = \arg \min_{p \in X_2 \cap X_3} \{g(p) + \mathrm{diag}(u^*) U(p)\},$$

then from Proposition 11, we have $\mathbf{0} \in \partial_{\mathrm{H}} Z(p^*) + \mathbf{N}_{X_2 \cap X_3}(p^*)$, which implies that

$$\mathbf{0} \in (I - Q) \partial_{\mathrm{H}} Z(p^*) + (I - Q) \mathbf{N}_{X_2 \cap X_3}(p^*).$$

Since there is $z^* \in \mathbb{H}^{Nn}$ such that $Lz^* = \mathbf{N}_{X_2 \cap X_3}(p^*)$, then $\mathbf{0} \in (I - Q)(\partial_{\mathrm{H}} Z(p^*) + Lz^*)$. Because $p^* \in X_2$, it obtains $\mathbf{0} \in D^{\mathrm{H}} \psi(Dp^* - d) = \partial_{\mathrm{H}} K(p^*)$. As a result,

$$\mathbf{0} \in -(I - Q)\big(\partial_{\mathrm{H}} g(p^*) + \mathrm{diag}(u^*) \partial_{\mathrm{H}} U(p^*) + Lz^*\big) - \partial_{\mathrm{H}} \|Dp^* - d\|_1,$$

which means $(p^*, u^*, z^*) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$ is an equilibrium solution of QNA (19). ■

Theorem 21: Under Assumptions 12 and 13, the state solution $(p(t), u(t), z(t))$ of QNA (19) from any initial solution $(p(0), u(0), z(0)) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$ is bounded and exists globally.

Proof: Since the right hand of QNA (19) is a set-valued mapping with nonempty compact convex values, then there is at least a local solution $(p(t), u(t), z(t))$ for any initial solution $(p(0), u(0), z(0)) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$ for $t \in [0, T)$ with $T > 0$.

In particular, for the local solution $(p(t), u(t), z(t))$, there exist measurable quaternion-valued functions $\xi(t) = \mathrm{col}(\xi_1(t), \xi_2(t), \cdots, \xi_N(t)) \in \partial_{\mathrm{H}} g(p(t))$ with $\xi_i(t) \in \partial_{\mathrm{H}} g_i(p_i(t))$, $\zeta(t) \in \partial_{\mathrm{H}} U(p(t))$ and $\eta(t) \in \psi_{[-1,1]}(Dp(t) - d)$ such that

$$\begin{cases} \dot{p}(t) = -\alpha(t)(I - Q)\big(\xi(t) + \mathrm{diag}(u(t))\zeta(t) + Lz(t)\big) \\ \qquad -Lp(t) - D^{\mathrm{H}}\eta(t) \\ \dot{u}(t) = W(p(t)) \\ \dot{z}(t) = \alpha(t) Lp(t) \end{cases} \tag{22}$$

for a.e. $t \in [0, T)$. Let p^* be an optimal solution to the QDOP (9), then from Theorem 20, there is $u^* = \mathrm{col}(u_1^*, u_2^*, \cdots, u_N^*) \in \mathbb{R}^N_{>0}$ with $u_i > a_0$ $(i \in \mathcal{V})$ and $z^* \in \mathbb{H}^{Nn}$ such that $(p^*, u^*, z^*) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$ is an equilibrium solution of QNA (19). Based on Theorem 19, it has $p^* \in X_2$, so

$$\begin{cases} \mathbf{0} = -(I - Q)\big(\xi^* + \mathrm{diag}(u^*)\zeta^* + Lz^*\big) - D^{\mathrm{H}}\eta_{ik}^* \\ \mathbf{0} = W(p^*) \\ \mathbf{0} = Lp^* \end{cases} \tag{23}$$

where $\xi^* = \mathrm{col}(\xi_1^*, \xi_2^*, \cdots, \xi_N^*) \in \partial_{\mathrm{H}} g(p^*)$ with $\xi_i^* \in \partial_{\mathrm{H}} g_i(p_i^*)$ for $i \in \mathcal{V}$, $\zeta^* \in \partial_{\mathrm{H}} U(p^*)$ and $\eta_{ik}^* \in \psi_{[-1,1]}(Dp^* - d)$. Due to Assumption 12, it has $\eta_{ik}^* = \mathbf{0}$, so $\mathbf{0} = -(I - Q)\big(\xi^* + \mathrm{diag}(u^*)\zeta^* + Lz^*\big)$.

Consider the candidate Lyapunov function

$$V_1(p(t), u(t), z(t)) = \frac{1}{2}(\|p(t) - p^*\|^2 + \|u(t) - u^*\|^2 + \|z(t) - z^*\|^2). \tag{24}$$

Then the trajectory of V_1 along QNA (19) satisfies

$$\begin{aligned} \dot{V}_1 = &\ 4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}}[-\alpha(t)(\mathbf{I} - Q)\big(\xi(t) \\ &+ \mathrm{diag}(u(t))\zeta(t) + Lz(t)\big) - Lp(t) - D^{\mathrm{H}}\eta(t)]\} \\ &+ 4\mathcal{R}\{(u(t) - u^*)^{\mathrm{H}} W(p(t))\} \\ &+ 4\mathcal{R}\{(z(t) - z^*)^{\mathrm{H}} \alpha(t) Lp(t)\}. \end{aligned}$$

If $p(0) \notin X_2$, by Theorem 19, there is $T_{X_2} > 0$ such that $p(t) \notin X_2$ when $t \in [0, T_{X_2})$ while $p(t) \in X_2$ when $t \geq T_{X_2}$. When $t \in [0, T_{X_2})$, according to (17), it has $\alpha(t) = 0$ and

$$\begin{aligned} \dot{V}_1 &= 4\mathcal{R}\{-(p(t) - p^*)^{\mathrm{H}} D^{\mathrm{H}} \eta(t)\} \\ &\quad - 4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}} L(p(t) - p^*)\} \\ &\quad + 4\mathcal{R}\{(u(t) - u^*)^{\mathrm{H}} W(p(t))\} \\ &\leq K(p^*) - K(p(t)) \\ &\quad - 4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}} L(p(t) - p^*)\} \\ &\leq -4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}} L(p(t) - p^*)\} \end{aligned}$$

because the convexity of $K(p)$ and the nonnegativity of $K(p)$ and $W(p)$. When $t \geq T_{X_2}$, $\alpha(t) = 1$ and $Dp = Dp^* = d$, then $Qp = Qp^* = q \triangleq D^{\mathrm{H}}(DD^{\mathrm{H}})^{-1} d$. Therefore,

$$\begin{aligned} \dot{V}_1 &= 4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}}[-(I - Q)\big(\xi(t) + \mathrm{diag}(u(t))\zeta(t) \\ &\quad + Lz(t)\big) - Lp(t) - D^{\mathrm{H}}\eta(t)]\} \\ &\quad + 4\mathcal{R}\{(u(t) - u^*)^{\mathrm{H}} W(p(t))\} \\ &\quad + 4\mathcal{R}\{(z(t) - z^*)^{\mathrm{H}} Lp(t)\} \\ &= 4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}}[-\xi(t) - \mathrm{diag}(u(t))\zeta(t) - Lz(t) \\ &\quad + \big(\xi^* + \mathrm{diag}(u^*)\zeta^* + Lz^*\big) - D^{\mathrm{H}}\eta(t)]\} \\ &\quad + 4\mathcal{R}\{(u(t) - u^*)^{\mathrm{H}} W(p(t))\} \\ &\quad + 4\mathcal{R}\{(z(t) - z^*)^{\mathrm{H}} Lp(t)\} \\ &\quad - 4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}} L(p(t) - p^*)\} \\ &= -4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}}(\xi(t) - \xi^*)\} \\ &\quad - 4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}}[\mathrm{diag}(u(t))\zeta(t) - \mathrm{diag}(u^*)\zeta^*]\} \\ &\quad - 4\mathcal{R}\{(p(t) - p^*)^{\mathrm{H}} D^{\mathrm{H}} \eta(t)]\} \\ &\quad + 4\mathcal{R}\{(u(t) - u^*)^{\mathrm{H}} W(p(t))\} \end{aligned}$$

$$
\begin{aligned}
&-4\mathcal{R}\{(\boldsymbol{p}(t)-\boldsymbol{p}^*)^{\mathrm{H}}\boldsymbol{L}(\boldsymbol{p}(t)-\boldsymbol{p}^*)\}\\
\leq& \sum_{i=1}^{N}[(p_i(t)-p_i^*)^{\mathrm{H}}(u_i^*\xi_i-u_i(t)\xi_i(t))\\
&+(u_i(t)-u_i^*)^{\mathrm{H}}U_i(p_i(t))]\\
&-4\mathcal{R}\{(\boldsymbol{p}(t)-\boldsymbol{p}^*)^{\mathrm{H}}\boldsymbol{L}(\boldsymbol{p}(t)-\boldsymbol{p}^*)\}.
\end{aligned}
$$

Since $U_i(\cdot)$ is convex, it follows that

$$
\begin{aligned}
&\sum_{i=1}^{N}[(p_i(t)-p_i^*)^{\mathrm{H}}(u_i^*\xi_i-u_i(t)\xi_i(t))+(u_i(t)-u_i^*)^{\mathrm{H}}U_i(p_i(t))]\\
\leq& \sum_{i=1}^{N}u_i^{*\mathrm{H}}(U_i(p_i(t))-U_i(p_i^*))+\sum_{i=1}^{N}u_i(t)^{\mathrm{H}}(U_i(p_i^*)-U_i(p_i(t)))\\
&+\sum_{i=1}^{N}(u_i(t)-u_i^*)^{\mathrm{H}}U_i(p_i(t))\\
=& \sum_{i=1}^{N}(u_i(t)-u_i^*)^{\mathrm{H}}U_i(p_i)=0.
\end{aligned}
$$

Thus, we have $\dot{V}_1 \leq -4\mathcal{R}\{(\boldsymbol{p}(t)-\boldsymbol{p}^*)^{\mathrm{H}}\boldsymbol{L}(\boldsymbol{p}(t)-\boldsymbol{p}^*)\} \leq 0$. Combined with the above conclusions, it holds $V_1(t) \leq V_1(0)$ for $t \in [0, T)$, which means $(\boldsymbol{p}(t), \boldsymbol{u}(t), \boldsymbol{z}(t))$ is bounded when $t \in [0, T)$. Consequently, it follows that from any initial solution $(\boldsymbol{p}(0), \boldsymbol{u}(0), \boldsymbol{z}(0)) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$, the state solution $(\boldsymbol{p}(t), \boldsymbol{u}(t), \boldsymbol{z}(t))$ of QNA (19) is bounded and exists globally. ■

Theorem 22: Under Assumptions 12 and 13, from any initial solution $(\boldsymbol{p}(0), \boldsymbol{u}(0), \boldsymbol{z}(0)) \in \mathbb{H}^{Nn}\times\mathbb{R}^N_{>0}\times\mathbb{H}^{Nn}$, the state solution $(\boldsymbol{p}(t), \boldsymbol{u}(t), \boldsymbol{z}(t))$ of QNA (19) converges to an equilibrium solution of QNA(19). In particular, $\boldsymbol{p}(t)$ converges to an optimal solution to the QDOP (9).

Proof: Let $(\boldsymbol{p}(t), \boldsymbol{u}(t), \boldsymbol{z}(t))$ be the state solution of QNA (19) and $(\boldsymbol{p}^*, \boldsymbol{u}^*, \boldsymbol{z}^*)$ be the equilibrium solution of QNA (19), then from Theorem 20, $\boldsymbol{p}^*$ is an optimal solution to the QDOP (9). By Theorem 19, we have $\boldsymbol{p}(t) \in X_2$ for all $t \geq 0$.

Define the function

$$
\begin{aligned}
V_2(\boldsymbol{p}(t), \boldsymbol{u}(t), \boldsymbol{z}(t)) =& \boldsymbol{g}(\boldsymbol{p}(t))+\boldsymbol{u}(t)^{\mathrm{H}}W(\boldsymbol{p}(t))-\mathbf{1}_N^{\mathrm{T}}\boldsymbol{u}(t)\\
&+\tilde{\boldsymbol{p}}(t)^{\mathrm{H}}\tilde{\boldsymbol{L}}\tilde{z}(t)-\frac{1}{2}\boldsymbol{p}(t)^{\mathrm{H}}\boldsymbol{L}\boldsymbol{p}(t)+K(\boldsymbol{p}(t))
\end{aligned}
$$

where $\tilde{\boldsymbol{p}} = \Psi(\boldsymbol{p})$, $\tilde{z} = \Psi(z)$ and $\tilde{\boldsymbol{L}} = \Psi(\boldsymbol{L})$. Then there are $\xi(t) \in \partial_{\mathrm{H}}\boldsymbol{g}(\boldsymbol{p}(t))$, $\zeta(t) \in \partial_{\mathrm{H}}U(\boldsymbol{p}(t))$ and $\eta(t) \in \psi_{[-1,1]}(\boldsymbol{D}\boldsymbol{p}(t)-\boldsymbol{d})$ such that

$$
\begin{aligned}
\dot{V}_2 =& 4\mathcal{R}\{\big(\xi(t)+\mathrm{diag}(\boldsymbol{u}(t))\zeta(t)+\boldsymbol{L}\boldsymbol{z}(t)-\boldsymbol{L}\boldsymbol{p}(t)\\
&+\boldsymbol{D}^{\mathrm{H}}\eta(t)\big)^{\mathrm{H}}\dot{\boldsymbol{p}}(t)\}-\|\boldsymbol{L}\boldsymbol{p}(t)\|^2\\
&+4\mathcal{R}\{\big(W(\boldsymbol{p}(t)-\mathbf{1}_N)\big)^{\mathrm{H}}W(\boldsymbol{p}(t))\}\\
\leq& 4\mathcal{R}\{\big(\xi(t)+\mathrm{diag}(\boldsymbol{u}(t))\zeta(t)+\boldsymbol{L}\boldsymbol{z}(t)-\boldsymbol{L}\boldsymbol{p}(t)\\
&+\boldsymbol{D}^{\mathrm{H}}\eta(t)\big)^{\mathrm{H}}\dot{\boldsymbol{p}}(t)\}\\
&+4\mathcal{R}\{\big(W(\boldsymbol{p}(t)-\mathbf{1}_N)\big)^{\mathrm{H}}W(\boldsymbol{p}(t))\}\\
=& -4\|(\boldsymbol{I}-\boldsymbol{Q})[\xi(t)+\mathrm{diag}(\boldsymbol{u}(t))\zeta(t)+\boldsymbol{L}\boldsymbol{z}(t)]\\
&-\boldsymbol{L}\boldsymbol{p}(t)+\boldsymbol{D}^{\mathrm{H}}\eta(t)\|^2\\
&+\sum_{i=1}^{N}U_i^2(p_i(t))-\sum_{i=1}^{N}U_i(p_i(t)).
\end{aligned}
$$

where the last equality holds due to QNA (19) and $(\boldsymbol{I}-\boldsymbol{Q})^2 = \boldsymbol{I}-\boldsymbol{Q}$.

Moreover, based on Theorem 21, the state solution $(\boldsymbol{p}(t), \boldsymbol{u}(t), \boldsymbol{z}(t))$ of QNA (19) from any initial solution $(\boldsymbol{p}(0), \boldsymbol{u}(0), \boldsymbol{z}(0)) \in \mathbb{H}^{Nn} \times \mathbb{R}^N_{>0} \times \mathbb{H}^{Nn}$ is bounded. Thus, it holds $U_i(p_i(t)) \to 0$ $(i \in \mathcal{V})$ as $t \to +\infty$ by the dynamic of $u_i(t)$. As a result, there is $T_0 = T_0(\boldsymbol{p}(0), \boldsymbol{u}(0), \boldsymbol{z}(0)) > 0$ satisfies

$$
U_i^2(p_i(t)) - U_i(p_i(t)) \leq -\frac{1}{2}U_i(p_i(t))
$$

when $t \geq T_0$, which implies

$$
\begin{aligned}
\dot{V}_2 \leq& -\|(\boldsymbol{I}-\boldsymbol{Q})[\xi(t)+\mathrm{diag}(\boldsymbol{u}(t))\zeta(t)+\boldsymbol{L}\boldsymbol{z}(t)]\\
&-\boldsymbol{L}\boldsymbol{p}(t)+\boldsymbol{D}^{\mathrm{H}}\eta(t)\|^2-\frac{1}{2}\sum_{i=1}^{N}U_i(p_i(t)).
\end{aligned}
$$

Define $V = V_1 + V_2$ where V_1 is defined in (24), by the proof of Theorem 21, we have

$$
\begin{aligned}
\dot{V} \leq& -\|(\boldsymbol{I}-\boldsymbol{Q})[\xi(t)+\mathrm{diag}(\boldsymbol{u}(t))\zeta(t)+\boldsymbol{L}\boldsymbol{z}(t)]\\
&-\boldsymbol{L}\boldsymbol{p}(t)+\boldsymbol{D}^{\mathrm{H}}\eta(t)\|^2-\boldsymbol{p}(t)^{\mathrm{H}}\boldsymbol{L}\boldsymbol{p}(t)-\frac{1}{2}U(\boldsymbol{p}(t)) \qquad (25)
\end{aligned}
$$

for $t \geq T_0$. According to results in [25], it can obtain that $(\boldsymbol{p}(t), \boldsymbol{u}(t), \boldsymbol{z}(t))$ converges to M which is the largest weakly invariant subset of $\Upsilon \cap L_l$, $\Upsilon = \{(\boldsymbol{p}, \boldsymbol{u}, z) : 0 \in \dot{V}(\boldsymbol{p}, \boldsymbol{u}, z)\}$ and $L_l = \{(\boldsymbol{p}, \boldsymbol{u}, z) : V((\boldsymbol{p}, \boldsymbol{u}, z)) \leq V((\boldsymbol{p}(0), \boldsymbol{u}(0), z(0)))\}$. By the definition of M and (25), if $(\boldsymbol{p}(t), \boldsymbol{u}(t), z(t)) \in M$, it must satisfy $0 = \|(\boldsymbol{I}-\boldsymbol{Q})[\xi(t)+\mathrm{diag}(\boldsymbol{u}(t))\zeta(t)+\boldsymbol{L}\boldsymbol{z}(t)]-\boldsymbol{L}\boldsymbol{p}(t)+\boldsymbol{D}^{\mathrm{H}}\eta(t)\|^2$, $\boldsymbol{p}(t)^{\mathrm{H}}\boldsymbol{L}\boldsymbol{p}(t) = 0$ for any $t \geq 0$ and $U(\boldsymbol{p}(t)) = 0$. Therefore, $(\boldsymbol{p}(t), \boldsymbol{u}(t), z(t))$ is an equilibrium solution of QNA (19).

Further, since $\phi(t) \triangleq (\boldsymbol{p}(t), \boldsymbol{u}(t), z(t))$ is bounded, there exist $\tilde{\phi} \triangleq (\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{u}}, \tilde{z})$ and a increasing sequence $\{t_k\}$ such that $\phi(t_k) \to \tilde{\phi}$ as $k \to +\infty$. Thus, $\tilde{\phi} \in M$, i.e. $(\tilde{\boldsymbol{p}}, \tilde{\boldsymbol{u}}, \tilde{z})$ is an equilibrium solution of QNA (19). On the other hand, the definition of V_1 implies the fact that each equilibrium solution of QNA (19) is Lyapunov stable. Hence, for any $\varepsilon > 0$, there are $\delta > 0$ and $T_1 > 0$, $\phi(t) \in B(\tilde{\phi}, \varepsilon)$ when $t \geq T_1$. Because $\phi(t_k) \to \tilde{\phi}$ $(k \to +\infty)$, there is $m > 0$ such that $\phi(t_m) \in B(\tilde{\phi}, \varepsilon)$ for all $t \geq t_m$. Consequently, $(\boldsymbol{p}(t), \boldsymbol{u}(t), z(t))$ converges to an equilibrium solution of QNA (19), and then $\boldsymbol{p}(t)$ converges to an optimal solution to the QDOP (9). ■

Remark 23: Compared to the centralized QNA in [24], the proposed QNA can effectively protect the security of privacy by means of the multi-agent system. Because one of the most significant advantages of distributed optimization is that agents store the information related to the optimization problem in a distributed way, the private information of each agent will be protected.

Specifically, the private information of agent i involved in the QNA (16) includes the local objective function g_i, local constraint information h_i, D_i and d_i and its estimation of the optimal solution p_i of the QDOP (1). Although p_i and z_i can be obtained by some agents of the multi-agent system, it is only available to those agents with permissions, i.e. neighbors of agent i. In addition, the QNA (16) does not require agents to exchange their more important private information, such as gradient, objective function, local constraints, etc.

Remark 24: The QNA (19) for solving QDOPs can also be used to solve real-valued or complex-valued distributed convex optimization problems with inequality and affine equality constraints. For example, when the imaginary parts of $p \in \mathbb{H}^n$ reduce to zero, the QDOP (1) becomes the following real-valued distributed optimization problems

$$\begin{aligned} &\min f(x) \triangleq \sum_{i=1}^{N} f_i(x) \\ &\text{s.t. } h_i(x) \le 0, \quad D_i x = d_i, \ i = 1, 2, \cdots, N. \end{aligned} \tag{26}$$

Correspondingly, the QNA (16) is simplified as

$$\begin{cases} \dot{x}_i(t) \in -\alpha(t)(I - Q_i)\big(\partial f_i(x_i(t)) + u_i(t)\partial U_i(x_i(t)) \\ \qquad + \sum_{j=1}^{N} a_{ij}(z_i(t) - z_j(t))\big) - \sum_{j=1}^{N} a_{ij}(x_i(t) - x_j(t)) \\ \qquad - \partial_{\mathrm{H}} \|D_i x_i(t) - d_i\|_1 \\ \dot{u}_i(t) = U_i(x_i(t)), \ u_i(0) > 0 \\ \dot{z}_i(t) = \alpha(t) \sum_{j=1}^{N} a_{ij}(x_i(t) - x_j(t)), \end{cases} \tag{27}$$

where $x_i \in \mathbb{R}^n, f_i : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$ and $h_i : \mathbb{R}^n \to \mathbb{R}^{m_i}$ with $h_{ij} : \mathbb{R}^n \to \mathbb{R}$ $(j = 1, 2, \cdots, m_i)$, $D_i \in \mathbb{R}^{s_i \times n}$ and $d_i \in \mathbb{R}^{s_i}$. u_i is the adaptive control variable and z_i is the auxiliary variable. The remaining symbols are as defined in the QNA (16). Compared with most of the existing real-valued neurodynamic approaches for solving real-valued distributed optimization problems such as [12], [15], and [16], neurodynamic approach (27) not only avoids calculating the lower bound of the penalty parameters by introducing the adaptive controller, but also has lower dimension and looser initial conditions.

V. NUMERICAL SIMULATION

Example 25: Consider a network consisting of ten agents interacting over an undirected communication graph shown in FIGURE 1 to collaboratively solve the following optimization problem:

$$\begin{aligned} &\min g(\boldsymbol{p}) = \sum_{i=1}^{10} \|p_i\|^2 \\ &\text{s.t. } \|p_i - a\|^2 \le 4; \ \mathcal{R}(p_{i1}) + i\mathcal{R}(p_{i2}) \ge 0; \\ &\quad (-2 + \mathbf{i} + \mathbf{j} + \mathbf{k})p_{i1} + (1 - 2\mathbf{i} - 2\mathbf{j} - 2\mathbf{k})p_{i2} \\ &\quad = -0.14\mathbf{i} + 0.15\mathbf{j} - 0.91\mathbf{k} \end{aligned} \tag{28}$$

where $p_i = (p_{i1}, p_{i2}) \in \mathbb{H}^2$ for $i = 1, 2, \ldots, 10$, $\boldsymbol{p} = \text{col}\{p_1, p_2, \ldots, p_{10}\} \in \mathbb{H}^{20}$ and

$$a = \begin{pmatrix} 1 + \mathbf{i} + \mathbf{j} + \mathbf{k} \\ 1 + \mathbf{i} + \mathbf{j} + \mathbf{k} \end{pmatrix}.$$

Obviously, problem (28) is a convex QDOP with inequality constraints and affine equality constraints. Here, the considered QNA (19) demonstrates that the state solution $\boldsymbol{p}(t)$ from a random initial point converges to an optimal solution $p^* = (0.3+0.41\mathbf{i}+0.1\mathbf{j}+0.25\mathbf{k}, 0.15+0.18\mathbf{i}+0.3\mathbf{j}+0.125\mathbf{k})$ to the QDOP (28). FIGURE 2 displays that the trajectories of all agents reached a consensus while the optimal solution is obtained, which illustrates the effectiveness of the QNA (19).

The algorithm (29) in [24] is applied to solve the following centralized version of the QDOP (28),

$$\begin{aligned} &\min g(p) = \|p\|^2 \\ &\text{s.t. } \|p - a\|^2 \le 4; \\ &\quad \mathcal{R}(p_1) + i\mathcal{R}(p_2) \ge 0, \quad i = 1, 2, \cdots, 10; \\ &\quad (-2 + \mathbf{i} + \mathbf{j} + \mathbf{k})p_1 + (1 - 2\mathbf{i} - 2\mathbf{j} - 2\mathbf{k})p_2 \\ &\quad = -0.14\mathbf{i} + 0.15\mathbf{j} - 0.91\mathbf{k} \end{aligned} \tag{29}$$

the corresponding state solution trajectory is shown in FIGURE 3. Compared with FIGURE 2 of the result generated by the QNA (19) in this paper, it takes a little longer time to find the optimal solution, which is due to the stronger parallel computing capability of distributed algorithms. In addition, with the self-adjustment ability of adaptive control, QNA (19) also has advantages in algorithm structure and solution space dimension, and thus has less computation.

Example 26: Color image restoration is an important task in imaging science. Many variational models treat a color image as a Euclidean vector or a direct combination of three monochromatic images, which ignores the inherent color structure between channels. To better describe the connection between color channels, we represent the color image as a pure quaternion matrix, that is, each pixel of image u is represented by a quaternion. K-means clustering singular value decomposition (K-*SVD*) based on dictionary learning has the advantage of preserving image texture. If the image $u \in \mathbb{H}^{m\times m}$ satisfies $u = DA$ (or $u \approx DA$), where $A \in \mathbb{H}^{m\times m}$ is the sparse coefficient matrix, that is, there are only a few non-zero numbers, then we call the dictionary sparse representation of u is sparse (or nearly sparse) under dictionary $D \in \mathbb{H}^{m\times m}$, which is an important step in color image restoration. In order to obtain the sparse representation of image u based on fixed dictionary D, the following optimization problem generally need to be solved:

$$\begin{aligned} &\min \|A\|_0 \\ &\text{s.t. } \|u - DA\|_2^2 \le \varepsilon \end{aligned} \tag{30}$$

where $\|\cdot\|_0$ defines ℓ_0-norm, $\varepsilon > 0$ is a parameter related to the noise level. Since ℓ_0-norm is difficult to solve optimally (NP-hard problem), and large-scale matrix calculation in the solving process should be avoided, a network consisting of m agents is introduced. Thus, the problem (30) is deformed to obtain the following QDOP

$$\min \sum_{i=1}^{m} \|Vec(A_i)\|_1$$

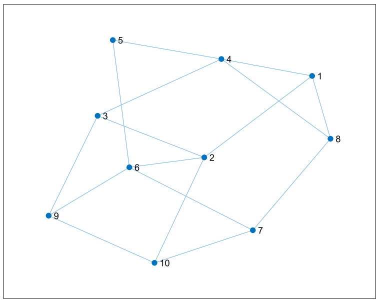


FIGURE 1. Communication topology of the multi-agent system in Example 25.

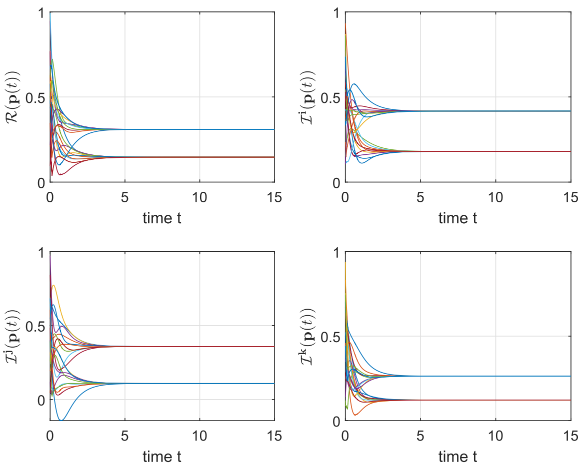


FIGURE 2. Trajectories of the state solution $p(t)$ in Example 25.

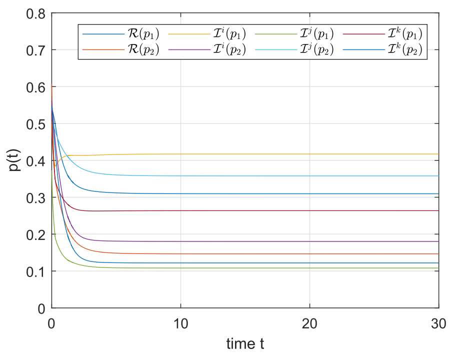


FIGURE 3. Trajectories of the state solution $p(t)$ by Algorithm in [24].

$$\text{s.t. } \|Vec(u - DA_i)\|_2^2 \le \varepsilon_1, \quad A_i = A_j, \\ i, j = 1, 2, \cdots, m \tag{31}$$

where $Vec(A) = (A_{11}, A_{12}, \cdots, A_{1m}, A_{21}, \cdots, A_{mm})^{\mathrm{T}} \in \mathbb{H}^{m^2}$, $\varepsilon_1 > 0$.

For numerical simulation, let $m = 3$ and the communication diagram of three agents is shown in FIGURE 4, $\varepsilon_1 = 0.2$,

$$u = \begin{pmatrix} 0.2\mathbf{i} - 0.6\mathbf{j} + 0.68\mathbf{k} & 0.07\mathbf{i} + 0.1\mathbf{k} & 0.2\mathbf{i} \\ 0.68\mathbf{i} + 0.5\mathbf{k} & -0.1\mathbf{i} - 0.07\mathbf{k} & -0.2\mathbf{k} \\ 0.9\mathbf{i} + 0.1\mathbf{j} + 0.68\mathbf{k} & -0.03\mathbf{i} & 0.2\mathbf{i} \end{pmatrix}$$

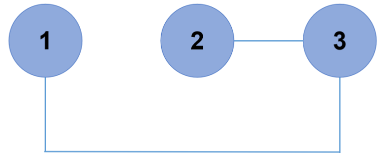


FIGURE 4. Communication topology of the multi-agent system in Example 26.

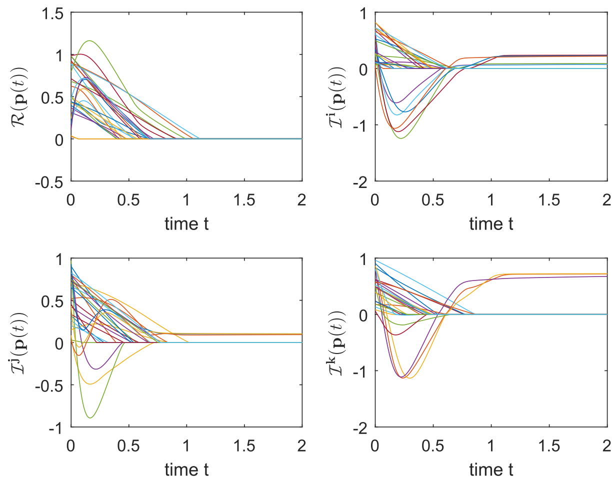


FIGURE 5. Trajectories of the state solution $p(t)$ in Example 26.

and

$$D = \begin{pmatrix} 1 & \mathbf{i} & \mathbf{i} \\ \mathbf{j} & 1 & \mathbf{k} \\ 1 & \mathbf{j} & \mathbf{k} \end{pmatrix}.$$

Therefore, through applying the QNA (19), we get the dictionary sparse representation of u under D is

$$A = \begin{pmatrix} 0.2\mathbf{i} + 0.1\mathbf{j} + 0.71\mathbf{k} & 0.07\mathbf{i} & 0.2\mathbf{i} \\ 0.7\mathbf{k} & 0 & 0 \\ 0 & 0.1\mathbf{j} & 0 \end{pmatrix}$$

and the trajectories of the state solution $\boldsymbol{p}(t)$ of the QNA (19) is showed in FIGURE 5.

VI. CONCLUSION

In this paper, we have designed an effective QNA for nonsmooth convex QDOPs based on multi-agent systems. The QNA has been proven that that it is capable of addressing convex QDOPs with inequality and affine equality constraints. After the convexity analysis of the quaternion-variable functions is given, the equivalent transformation of the QDOP has been carried out by using the connectivity of the communication topology and penalty method, and then a distributed QNA with adaptive controllers has been proposed. Finally, both theoretical results and numerical examples show the validity of the proposed QNA.

In the future, we will focus on QNAs with finite-time or fixed-time convergence rates and QNAs with event-triggered mechanisms for solving QDOPs. In addition, we believe the potential application of QDOPs will be also interesting.

APPENDIX

A. PROOF OF PROPOSITION 9

According to the propositions of Ψ and (6), for $p = x + y\mathbf{i} + z\mathbf{j} + w\mathbf{k}$, it has $\Psi(\partial_{\mathbb{H}} g(p)) = \partial_x f(x, y, z, w) \times \partial_y f(x, y, z, w) \times \partial_z f(x, y, z, w) \times \partial_w f(x, y, z, w)$. Thus, based on Lemma 8, we have $\partial f(\Psi(p)) \subseteq \Psi(\partial_{\mathbb{H}} g(p))$.

Conversely, from Theorem 1 in [24] for any $\eta \in \partial_{\mathbb{H}} g(p)$ and $q \in \mathbb{H}^n$, let $u = \Psi(p)$ and $v = \Psi(q)$, then $f(v) - f(u) = g(q) - g(p) \geq \mathcal{R}(\eta^{\mathrm{H}}(q-p)) = \Psi(\eta)^{\mathrm{T}}\Psi(q-p) = \Psi(\eta)^{\mathrm{T}}(v-u)$, that is, $\Psi(\eta) \in \partial f(u) = \partial f(\Psi(p))$. Thus, $\Psi(\partial_{\mathbb{H}} g(p)) \subseteq \partial f(\Psi(p))$. In a word, $\Psi(\partial_{\mathbb{H}} g(p)) = \partial f(\Psi(p))$ holds.

B. PROOF OF PROPOSITION 10

As the proof of Theorem 1 in [24], it can be obtained that $\partial_{\mathbb{H}} g(p) \subseteq \{\eta \in \mathbb{H}^n : g(q) - g(p) \geq \mathcal{R}(\eta^{\mathrm{H}}(q-p))\} \triangleq Y$. But for any $\eta \in Y$, we have $g(q) - g(p) \geq \mathcal{R}(\eta^{\mathrm{H}}(q-p))$ for any $q \in \mathbb{H}^n$.

According to Definition 1 and Proposition 9, it implies $f(v) - f(u) = g(q) - g(p) \geq \mathcal{R}(\eta^{\mathrm{H}}(q-p)) = \Psi(\eta)^{\mathrm{T}}(v-u)$, where $u = \Psi(p)$ and $v = \Psi(q)$. Thus,

$$\Psi(\eta) \in \partial f(u) = \partial f(\Psi(p)) = \Psi(\partial_{\mathbb{H}} g(p)),$$

that is, $\eta \in \partial_{\mathbb{H}} g(p)$. As a result, $Y \subseteq \partial_{\mathbb{H}} g(p)$. To sum up, $\partial_{\mathbb{H}} g(p) = \{\eta \in \mathbb{H}^n : g(q) - g(p) \geq \mathcal{R}(\eta^{\mathrm{H}}(q-p))\}$.

In addition, for any $\eta \in \partial_{\mathbb{H}} g(p)$, $\varsigma \in \partial_{\mathbb{H}} g(q)$, there is $g(q) - g(p) \geq \mathcal{R}(\eta^{\mathrm{H}}(q-p))$ and $g(p) - g(q) \geq \mathcal{R}(\varsigma^{\mathrm{H}}(p-q))$. After adding the above two inequalities, it comes to $\mathcal{R}((\eta - \varsigma)^{\mathrm{H}}(p-q)) \geq 0$.

C. PROOF OF PROPOSITION 11

1) Let $f : \mathbb{R}^{4n} \to \mathbb{R}$ be the auxiliary function of g that $g(p) = f(\Psi(p))$. Since $g(p)$ is local Lipschitz on S, there exists $l > 0$ such that $|g(p) - g(q)| \leq l\|p - q\|$ for any $p, q \in S$. Therefore, it follows

$$\begin{aligned} |f(\Psi(p)) - f(\Psi(q))| &= |g(p) - g(q)| \\ &\leq l\|p-q\| = l\|\Psi(p) - \Psi(q)\| \end{aligned}$$

for any $p, q \in S$ from Proposition 2, that is, f is local Lipschitz on $\widetilde{S}$ where $\widetilde{S} = \{x \in \mathbb{R}^{4n} : x = \Psi(p), p \in S\}$.

Moreover, since $\mathbf{N}_S(p) = \{o \in \mathbb{H}^n : \mathcal{R}(o^{\mathrm{H}}(p-q)) \leq 0, \forall\ q \in S\} = \{\Psi(o) \in \mathbb{R}^{4n} : \Psi(o)^{\mathrm{T}}(\Psi(p) - \Psi(q)) \leq 0, \forall\ \Psi(q) \in \widetilde{S}\} = \mathbf{N}_{\widetilde{S}}(\Psi(p))$ and if $g(p)$ attains a minimum over S at $p \in S$, $f(\Psi(p))$ attains a minimum over $\widetilde{S}$ at $\Psi(p)$. Then from [28], it has $0 \in \partial f(\Psi(p)) + \mathbf{N}_{\widetilde{S}}(\Psi(p))$, which means $0 \in \partial_{\mathbb{H}} g(p) + \mathbf{N}_S(p)$.

2) It is obvious that $0 \in \mathrm{int}(S - Q)$ if and only if $0 \in \mathrm{int}(\widetilde{S} - \widetilde{Q})$. Then for any $p \in S \cap Q$, $\Psi(p) \in \widetilde{S} \cap \widetilde{Q}$, so $\mathbf{N}_{\widetilde{S}\cap\widetilde{Q}}(p) = \mathbf{N}_{\widetilde{S}}(\Psi(p)) + \mathbf{N}_{\widetilde{Q}}(\Psi(p))$ [28]. Thus, it has $\mathbf{N}_{S\cap Q}(p) = \mathbf{N}_S(p) + \mathbf{N}_Q(p)$.

D. PROOF OF LEMMA 14

For any $\lambda \in [0, 1]$, $\boldsymbol{p} = \mathrm{col}(p_1, p_2, \cdots, p_N)$ and $\boldsymbol{q} = \mathrm{col}(q_1, q_2, \cdots, q_N)$, there is $\boldsymbol{g}(\lambda\boldsymbol{p} + (1-\lambda)\boldsymbol{q}) = \sum_{i=1}^N g_i(\lambda p_i + (1-\lambda)q_i) \leq \sum_{i=1}^N \lambda g_i(p_i) + (1-\lambda)g_i(q_i) = \lambda\boldsymbol{g}(\boldsymbol{p}) + (1-\lambda)\boldsymbol{g}(\boldsymbol{q})$, which implies $\boldsymbol{g}$ is convex. Moreover, the proof of Theorem 1 in [25] has confirmed

$$\mathrm{col}(\partial_{\mathbb{H}} g_1(p_1), \partial_{\mathbb{H}} g_2(p_2), \cdots, \partial_{\mathbb{H}} g_N(p_N)) \subseteq \partial_{\mathbb{H}} \boldsymbol{g}(\boldsymbol{p}).$$

So we just need to prove that $\partial_{\mathbb{H}} \boldsymbol{g}(\boldsymbol{p}) \subseteq \mathrm{col}(\partial_{\mathbb{H}} g_1(p_1), \partial_{\mathbb{H}} g_2(p_2), \cdots, \partial_{\mathbb{H}} g_N(p_N))$ is true.

In fact, by introducing the auxiliary functions $\boldsymbol{f}$ and f_i $(i \in \mathcal{V})$ such that $\boldsymbol{g}(\boldsymbol{p}) = \boldsymbol{f}(\Psi(\boldsymbol{p}))$ and $g_i(p_i) = f_i(\Psi(p_i))$, where Ψ is defined in Definition 1, there is $\boldsymbol{f}(\Psi(\boldsymbol{p})) = \sum_{i=1}^N f_i(\Psi(p_i))$. Define a bijection $\varphi : \mathbb{R}^{4Nn} \to \mathbb{R}^{4Nn}$ such that $\varphi(\Psi(\boldsymbol{p})) = \mathrm{col}(\Psi(p_1), \Psi(p_2), \cdots, \Psi(p_N))$. For every $\eta \in \partial_{\mathbb{H}} \boldsymbol{g}(\boldsymbol{p})$, let $\widetilde{\Psi} = \varphi \circ \Psi$, then similar to Proposition 9, we have $\zeta = \widetilde{\Psi}(\eta) \in \widetilde{\Psi}(\partial_{\mathbb{H}} \boldsymbol{g}(\boldsymbol{p})) = \partial \boldsymbol{f}(\widetilde{\Psi}(\boldsymbol{p}))$. Since $\boldsymbol{g}$ is convex, $\boldsymbol{f}$ is convex. Thus, $\partial \boldsymbol{f}(\widetilde{\Psi}(\boldsymbol{p})) = \mathrm{col}(\partial f_1(\Psi(p_1)), \partial f_2(\Psi(p_2)), \cdots, \partial f_N(\Psi(p_N)))$ and there exist $\zeta_i \in \partial f_i(\Psi(p_i))$ such that $\zeta = \mathrm{col}(\zeta_1, \zeta_2, \cdots, \zeta_N)$. So it is obvious that $\zeta_i \in \Psi(\partial_{\mathbb{H}} g_i(p_i))$ $(i \in \mathcal{V})$, and $\zeta \in \mathrm{col}(\Psi(\partial_{\mathbb{H}} g_1(p_1)), \Psi(\partial_{\mathbb{H}} g_2(p_2)), \cdots, \Psi(\partial_{\mathbb{H}} g_N(p_N)))$. It follows

$$\begin{aligned} &\Psi(\eta) \\ &\in \varphi^{-1}(\mathrm{col}(\Psi(\partial_{\mathbb{H}} g_1(p_1)), \Psi(\partial_{\mathbb{H}} g_2(p_2)), \cdots, \Psi(\partial_{\mathbb{H}} g_N(p_N)))) \\ &= \Psi(\mathrm{col}(\partial_{\mathbb{H}} g_1(p_1), \partial_{\mathbb{H}} g_2(p_2), \cdots, \partial_{\mathbb{H}} g_N(p_N))) \end{aligned}$$

and $\eta \in \mathrm{col}(\partial_{\mathbb{H}} g_1(p_1), \partial_{\mathbb{H}} g_2(p_2), \cdots, \partial_{\mathbb{H}} g_N(p_N))$ from the bijectivity of φ and Ψ.

Therefore, $\partial_{\mathbb{H}} \boldsymbol{g}(\boldsymbol{p}) \subseteq \mathrm{col}(\partial_{\mathbb{H}} g_1(p_1), \partial_{\mathbb{H}} g_2(p_2), \cdots, \partial_{\mathbb{H}} g_N(p_N))$ holds, which completes the proof.

REFERENCES

[1] W. R. Hamilton, "On a new species of imaginary quantities connected with a theory of quaternions," *Proc. Roy. Irish Acad.*, vol. 2, no. 1, pp. 424–434, Nov. 1844.

[2] N. Matsui, T. Isokawa, H. Kusamichi, F. Peper, and H. Nishimura, "Quaternion neural network with geometrical operators," *J. Intell. Fuzzy Syst.*, vol. 15, nos. 3–4, pp. 149–164, 2004.

[3] S.-C. Pei and C.-M. Cheng, "Color image processing by using binary quaternion-moment-preserving thresholding technique," *IEEE Trans. Image Process.*, vol. 8, no. 5, pp. 614–628, May 1999.

[4] L. Fortuna, G. Muscato, and M. G. Xibilia, "A comparison between HMLP and HRBF for attitude control," *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 318–328, Mar. 2001.

[5] C. Jahanchahi and D. P. Mandic, "A class of quaternion Kalman filters," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 3, pp. 533–544, Mar. 2014.

[6] M. Toyoda and Y. Wu, "Mayer-type optimal control of probabilistic Boolean control network with uncertain selection probabilities," *IEEE Trans. Cybern.*, vol. 51, no. 6, pp. 3079–3092, Jun. 2021.

[7] Y. Wu, Y. Guo, and M. Toyoda, "Policy iteration approach to the infinite horizon average optimal control of probabilistic Boolean networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 7, pp. 2910–2924, Jul. 2021.

[8] Y. Wu, J. Zhang, and T. Shen, "A logical network approximation to optimal control on continuous domain and its application to HEV control," *Sci. China Inf. Sci.*, 2022, doi: 10.1007/s11432-021-3446-8.

[9] S. Le, Y. Wu, Y. Guo, and C. D. Vecchio, "Game theoretic approach for a service function chain routing in NFV with coupled constraints," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 68, no. 12, pp. 3557–3561, Dec. 2021.

[10] J. J. Hopfield and D. W. Tank, "Neural computation of decisions in optimization problems," *Biological*, vol. 52, no. 3, pp. 141–152, 1985.

[11] D. Tank and J. Hopfield, "Simple 'neural' optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit," *IEEE Trans. Circuits Syst.*, vol. CAS-33, no. 5, pp. 533–541, May 1986.

[12] Q. Liu and J. Wang, "A one-layer recurrent neural network for constrained nonsmooth optimization," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 41, no. 5, pp. 1323–1333, Oct. 2011.

[13] X. He, T. Huang, J. Yu, C. Li, and Y. Zhang, "A continuous-time algorithm for distributed optimization based on multiagent networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 12, pp. 2700–2709, Dec. 2019.

[14] H. Zhou, X. Zeng, and Y. Hong, "Adaptive exact penalty design for constrained distributed optimization," *IEEE Trans. Autom. Control*, vol. 64, no. 11, pp. 4661–4667, Nov. 2019.

[15] N. Liu and S. Qin, "A neurodynamic approach to nonlinear optimization problems with affine equality and convex inequality constraints," *Neural Netw.*, vol. 109, pp. 147–158, Jan. 2019.

[16] X. Wen, L. Luan, and S. Qin, "A continuous-time neurodynamic approach and its discretization for distributed convex optimization over multi-agent systems," *Neural Netw.*, vol. 143, pp. 52–65, Nov. 2021.

[17] S. L. Goh, M. Chen, D. H. Popović, K. Aihara, D. Obradovic, and D. P. Mandic, "Complex-valued forecasting of wind profile," *Renew. Energy*, vol. 31, no. 11, pp. 1733–1750, 2006.

[18] S. Qin, J. Feng, J. Song, X. Wen, and C. Xu, "A one-layer recurrent neural network for constrained complex-variable convex optimization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 3, pp. 534–544, Mar. 2018.

[19] J. P. Ward, *Quaternions and Cayley Numbers: Algebra and Applications*. Dordrecht, The Netherlands: Springer, 2001.

[20] D. Xu, C. Jahanchahi, C. C. Took, and D. P. Mandic, "Enabling quaternion derivatives: The generalized HR calculus," *Roy. Soc. Open Sci.*, vol. 2, no. 8, 2015, Art. no. 150255.

[21] D. Xu, Y. Xia, and D. P. Mandic, "Optimization in quaternion dynamic systems: Gradient, hessian, and learning algorithms," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 2, pp. 249–261, Feb. 2016.

[22] X. Chen, Q. Song, Z. Li, Z. Zhao, and Y. Liu, "Stability analysis of continuous-time and discrete-time quaternion-valed nueral networks with linear threshold neurons," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 7, pp. 2769–2781, Jul. 2018.

[23] G. Chen and Q. Yang, "Distributed constrained optimization for multi-agent networks with nonsmooth objective functions," *Syst. Control Lett.*, vol. 124, pp. 60–67, Feb. 2019.

[24] Y. Liu, Y. Zheng, J. Lu, J. Cao, and L. Rutkowski, "Constrained quaternion-variable convex optimization: A quaternion-valued recurrent neural network approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 3, pp. 1022–1035, Mar. 2020.

[25] Z. Xia, Y. Liu, J. Lu, J. Cao, and L. Rutkowski, "Penalty method for constrained distributed quaternion-variable optimization," *IEEE Trans. Cybern.*, vol. 51, no. 11, pp. 5631–5636, Nov. 2021.

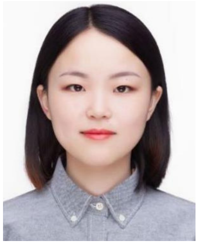
[26] J. Lu and C. Y. Tang, "Zero-gradient-sum algorithms for distributed convex optimization: The continuous-time case," *IEEE Trans. Autom. Control*, vol. 57, no. 9, pp. 2348–2354, Sep. 2012.

[27] X. Wang, G. Wang, and S. Li, "A distributed fixed-time optimization algorithm for multi-agent systems," *Automatica*, vol. 122, Dec. 2020, Art. no. 109289.

[28] F. H. Clarke, *Optimization and Nonsmooth Analysis*. New York, NY, USA: Wiley, 1983.

[29] Y. Zheng, Y. Liu, and J. Cao, "One-layer neural network for nonlinear convex programming with linear constraints," in *Proc. 37th Chin. Control Conf. (CCC)*, Jul. 2018, pp. 2329–2332.

GUOCHENG LI received the B.Sc. degree from Hebei Normal University, in 1986, and the M.Sc. and Ph.D. degrees in basic mathematics from the Harbin Institute of Technology, in 1993 and 2003, respectively. From 2003 to 2005, he was a Postdoctoral Fellow at Tsinghua University. He is currently a Professor at the School of Science, Beijing Information Science and Technology University. His current research interests include neural network optimization and sparse signals.



LINHUA LUAN was born in Shandong, China, in 1998. She received the B.Sc. degree in applied mathematics from Qingdao University, in 2020. She is currently pursuing the master's degree with the School of Science, Harbin Institute of Technology, Weihai. Her current research interests include distributed optimization and multi-agent systems.

SITIAN QIN was born in Shandong, China, in 1981. He received the Ph.D. degree in mathematics from the Harbin Institute of Technology, Harbin, China, in 2010. He was an Assistant Professor, from 2010 to 2013, and an Associate Professor, from 2013 to 2018, with the Harbin Institute of Technology, Weihai, China, where he is currently a Professor. His current research interests include neural network theory and optimization.

JIQIANG FENG received the B.Sc. degree from the Yantai Normal College, Yantai, China, in 2005, and the M.Sc. and Ph.D. degrees from Shenzhen University, Shenzhen, China, in 2008 and 2011, respectively. In 2011, he joined Shenzhen University, where he is currently an Associate Professor. His current research interests include swarm optimization, neural network theory, and optimization.