Control Theory and Technology (2021) 19:249–259
https://doi.org/10.1007/s11768-021-00044-0

RESEARCH ARTICLE

Distributed solving Sylvester equations with fractional order dynamics

Songsong Cheng[1] · Shu Liang[2] · Yuan Fan[1]

Received: 20 August 2020 / Revised: 5 November 2020 / Accepted: 30 November 2020 / Published online: 17 May 2021

Abstract
This paper addresses distributed computation Sylvester equations of the form $AX + XB = C$ with fractional order dynamics. By partitioning parameter matrices A, B and C, we transfer the problem of distributed solving Sylvester equations as two distributed optimization models and design two fractional order continuous-time algorithms, which have more design freedom and have potential to obtain better convergence performance than that of the existing first order algorithms. Then, rewriting distributed algorithms as corresponding frequency distributed models, we design Lyapunov functions and prove that the proposed algorithms asymptotically converge to an exact or least squares solution. Finally, we validate the effectiveness of the proposed algorithms by providing a numerical example

Keywords Fractional order calculus · Distributed optimization · Sylvester equation · Frequency distributed model · Convergence

1 Introduction

Linear matrix equations computation is an important and attractive problem in many fields, such as image processing, control theory, and machine learning [1–3]. Sylvester equations, as a special class of linear matrix equations, received many research attention [4–6]. Many methods were designed to obtain exact or least squares solutions of Sylvester equations in centralized schemes [6–8].

However, in some big data or complex systems circumstances, the dimension of Sylvester equations is very large. Conventional centralized algorithms are difficult to solve high-dimensional Sylvester equations. Thanks to the favourable abilities of computation and communication, distributed computation schemes attracted many research attention [9–11]. Deng et al. [12] proposed a distributed algorithm to obtain an exact solution of Sylvester equations. Considering that Stein equations in the form of $X + AXB = C$ has no solution, Chen et al. [13] developed a distributed algorithm to achieve a least squares solution for a special case that A, B and C have standard Row–Column–Column structure. For linear matrix equations $AXB = C$, Zeng et al. [14] considered eight standard structures of A, B and C and developed four distributed algorithms to attain a least squares solution.

Fractional order theory was developed 300 years ago and attracted many scholars' attention in systems theory [15–17], signal processing [18–20] because of the advantages on system modelling and more design freedom. As far as we know, almost all of the centralized and distributed computations of Sylvester equations are designed based on the integer order iteration law. The objective of this paper is to design distributed algorithms with fractional order dynamics to solve Sylvester equations. Because of more design freedom of the fractional order calculus, we try to obtain that the fractional order algorithm can achieve better convergence performance than that of the integer order counterpart. The main contributions of this paper are listed as follows.

✉ Yuan Fan
yuanf@ahu.edu.cn

Songsong Cheng
sscheng@amss.ac.cn

Shu Liang
sliangResearch@163.com

1 Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education, School of Electrical Engineering and Automation, Anhui University, Hefei 230601, Anhui, China

2 Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

1. A universal distributed optimization model is introduced to handle any type of standard partitioning of matrices A, B and C.
2. In comparison with [12–14], we design two distributed algorithms with fractional order dynamics and prove that the proposed algorithms converge to an exact or least squares solution.

The rest of this paper is organized as follows. Section 2 presents some preliminaries in matrices, fractional order calculus, graph theory, and formulate the problem of distributed solving Sylvester equations, while Sect. 3 models the problem of solving Sylvester equations as two distributed optimization problems, designs two algorithms with fractional order dynamics and shows that proposed algorithms asymptotically converge to an exact or least squares solution. Section 4 provides a numerical example to illustrate the effectiveness of the proposed algorithm and Sect. 5 concludes this paper.

2 Preliminaries and problem formulations

2.1 Notations

The real number set, n-dimensional real column vector set, and $n \times m$ real matrix set are denoted as $\mathbb{R}$, $\mathbb{R}^n$, and $\mathbb{R}^{n\times m}$, respectively. $\mathbf{1}_n \in \mathbb{R}^n$ ($\mathbf{0}_n \in \mathbb{R}^n$, $\mathbf{0}_{n\times m} \in \mathbb{R}^{n\times m}$) is a vector (vector, matrix) with all of the elements are one (zero, zero). $\mathrm{diag}\{\cdot\}$ constructs a diagonal matrix with a vector's entries on the diagonal and $\mathrm{blkdiag}\{\cdot\}$ constructs an augmented diagonal matrix with matrix arguments on the diagonal. $I_n \in \mathbb{R}^{n\times n}$ is an identity matrix. For any $A \in \mathbb{R}^{n\times m}$, A^{T} means the transpose of A, a_{ij} means the ith row and jth column entry of A, and $A_{i\cdot}$ ($A_{\cdot i}$) denotes the ith row or row sub-block (column or column sub-block) of the matrix A with proper dimensions. $\|A\| = (\sum_{i=1}^n \sum_{j=1}^m a_{ij}^2)^{\frac{1}{2}}$ is the Frobenius norm of A. Similarly, for any two real matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{n\times m}$, the Frobenius inner product can be calculated as $\langle A, B\rangle = \sum_{i=1}^n \sum_{j=1}^m a_{ij}b_{ij}$. $[m_i]$ and $[r_i]$ are two sequences with $\sum_{i=1}^n m_i = m$, $\sum_{i=1}^n r_i = r$, $m_{[i]} = \sum_{j=1}^i m_j$, and $r_{[i]} = \sum_{j=1}^i r_j$. Then two sub-block matrices $A_{i\cdot} \in \mathbb{R}^{m_i\times m}$ and $B_{\cdot i} \in \mathbb{R}^{r\times r_i}$ can be augmented as $\bar{A}_i := [\mathbf{0}^{\mathrm{T}}_{m_{[i-1]}\times m}, A^{\mathrm{T}}_{i\cdot}, \mathbf{0}^{\mathrm{T}}_{(m-m_{[i]})\times m}]^{\mathrm{T}} \in \mathbb{R}^{m\times m}$ and $\bar{B}_i := [\mathbf{0}_{r\times r_{[i-1]}}, B_{\cdot i}, \mathbf{0}_{r\times(r-r_{[i]})}] \in \mathbb{R}^{r\times r}$, respectively.

2.2 Fractional order calculus

Definition 1 (*See* [21]) For a given function $f(x)$, the αth order Caputo derivative is

$$\mathscr{D}^{\alpha} f(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-\tau)^{n-1-\alpha} f^{(n)}(\tau)\mathrm{d}\tau, \tag{1}$$

where $\alpha \in (n-1, n)$, $n \in \mathbb{N}$, $\Gamma(t) = \int_0^{+\infty} \tau^{t-1}\mathrm{e}^{-\tau}\mathrm{d}\tau$ is Gamma function, and $f^{(n)}(t)$ is the nth order derivative of $f(t)$.

Based on Definition 1, we introduce the frequency distributed model of a fractional order system in the following lemma.

Lemma 1 (See [22]) *For a fractional order system $\mathscr{D}^{\alpha} x(t) = u(t)$ where $\alpha \in (0, 1)$, we provide its frequency distributed model as follows*:

$$\begin{cases} \dfrac{\partial z(\omega, t)}{\partial t} = -\omega z(\omega, t) + u(t), \\ y(t) = \displaystyle\int_0^{+\infty} \mu_{\alpha}(\omega) z(\omega, t)\mathrm{d}\omega, \end{cases} \tag{2}$$

where $\mu_{\alpha} = \dfrac{\sin(\alpha\pi)}{\omega^{\alpha}\pi}$.

2.3 Graph theory

For a network with nodes set $\mathcal{V} = \{1, 2, \ldots, n\}$ and edges set $\mathcal{E} \subseteq \mathcal{V}\times V$, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denotes the corresponding undirected communication graph. Node j is the neighbor of node i if $\{j, i\} \in \mathcal{E}$. $A = [a_{ij}] \in \mathbb{R}^{n\times n}$ is the adjacency matrix of graph $\mathcal{G}$ with $a_{ij} = a_{ji} > 0$ if $\{j, i\} \in \mathcal{E}$ and $a_{ij} = 0$ otherwise. Based on the adjacency matrix we calculate the degree matrix D and Laplacian matrix L as $D = \mathrm{diag}\{\sum_{j=1}^n a_{1j}, \ldots, \sum_{j=1}^n a_{nj}\}$ and $L = D - A$, respectively. Specifically, if an undirected graph $\mathcal{G}$ is connected, the Laplacian matrix L is symmetric positive semi-definite, and the rank and kernel of L are $n-1$ and $k1_n$ with $k \in \mathbb{R}$, respectively.

Assumption 1 The undirected graph $\mathcal{G}$ is connected.

2.4 Problem formulation

Consider a problem of solving the following Sylvester equation:

$$AX + XB = C, \tag{3}$$

where $A \in \mathbb{R}^{m\times m}$, $B \in \mathbb{R}^{r\times r}$, and $C \in \mathbb{R}^{m\times r}$ are given matrices, and $X \in \mathbb{R}^{m\times r}$ is the solution to be achieved.

Definition 2 X^* is an exact solution to (3), if $AX^* + X^*B = C$; and X^* is a least squares solution to (3), if $X^* = \underset{X}{\operatorname{argmin}}\|AX + XB - C\|^2$.

In the scheme of distributedly solving (3) with the aid of multi-agent systems, the ith agent only holds the ith sub-blocks parameter matrices A, B, and C, and exchanges information with their neighbors. For instance, matrices A, B, and C are partitioned as the Row–Column–Column structure, namely,

$$\begin{cases} A = [A_{1\cdot}^{\mathrm{T}}, A_{2\cdot}^{\mathrm{T}}, \dots, A_{n\cdot}^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{m\times m}, \\ B = [B_{\cdot 1}, B_{\cdot 2}, \dots, B_{\cdot n}] \in \mathbb{R}^{r\times r}, \\ C = [C_{\cdot 1}, C_{\cdot 2}, \dots, C_{\cdot n}] \in \mathbb{R}^{m\times r}, \end{cases} \tag{4}$$

where $A_{i\cdot} \in \mathbb{R}^{m_i\times m}$, $\sum_{i=1}^{n} m_i = m$, $B_{\cdot i} \in \mathbb{R}^{r\times r_i}$, $\sum_{i=1}^{n} r_i = r$, $C_{\cdot i} \in \mathbb{R}^{m\times r_i}$, and the ith agent has access to $A_{i\cdot}$, $B_{\cdot i}$, and $C_{\cdot i}$. Referring to the augmented technique in Notations, we augment $A_{i\cdot}$, $B_{\cdot i}$, and $C_{\cdot i}$ as $\bar{A}_i \in \mathbb{R}^{m\times m}$, $\bar{B}_i \in \mathbb{R}^{r\times r}$, and $\bar{C}_i \in \mathbb{R}^{m\times r}$, respectively. Then,

$$\sum_{i=1}^{n} \bar{A}_i X + X \sum_{i=1}^{n} \bar{B}_i - \sum_{i=1}^{n} \bar{C}_i = \mathbf{0}. \tag{5}$$

By introducing local decision variables X_i for agent i with consensus constrains, we have

$$\begin{cases} \sum_{i=1}^{n} \left(\bar{A}_i X_i + X_i \bar{B}_i - \bar{C}_i\right) = \mathbf{0}, \\ X_i = X_j. \end{cases} \tag{6}$$

To decouple (6), we introduce an intermediate matrix $M = [M_1^{\mathrm{T}}, M_2^{\mathrm{T}}, \dots, M_n^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{nm\times r}$ with $M_i \in \mathbb{R}^{m\times r}$, $\forall i \in \mathcal{V}$. Then,

$$\begin{cases} \bar{A}_i X_i + X_i \bar{B}_i - \bar{C}_i + \sum_{j=1}^{n} l_{ij} M_j = \mathbf{0}, \\ X_i = X_j, \end{cases} \tag{7}$$

where the matrix $L := [l_{ij}]$ is the Laplacian matrix of undirected graphs. Rewrite (7) as

$$\begin{cases} \bar{A}_i X_i + X_i \bar{B}_i - \bar{C}_i + \bar{L}_{i\cdot} M = \mathbf{0}, \\ \bar{L} X = \mathbf{0}, \end{cases} \tag{8}$$

where $\bar{L} = L \otimes I_m$ and $\bar{L}_{i\cdot} = L_{i\cdot} \otimes I_m$.

***Remark* 1** The Row–Column–Column partition in (4) is presented as an example to decouple coefficient matrices A, B, and C. Actually, with the aid of augmented technique, the decoupling in (8) can handle any partition of A, B, and C.

3 Main results

In this section, we design two algorithms to obtain an exact solution and a least-squares one, respectively. Firstly, we design an algorithm to obtain the exact solution.

3.1 Obtaining an exact solution

According to (8), we model the following optimization problem.

$$\begin{aligned} \min_{X, M} \quad & \frac{1}{2} \sum_{i=1}^{n} \langle X_i, \sum_{j=1}^{n} l_{ij} X_j \rangle, \\ \text{s.t.} \quad & \bar{A}_i X_i + X_i \bar{B}_i - \bar{C}_i + \bar{L}_{i\cdot} M = \mathbf{0}. \end{aligned} \tag{9}$$

It can be verified that the problems (3) and (9) are equivalent. Based on (9), we construct an augmented Lagrangian function as

$$\begin{aligned} &\mathcal{L}(X, M, \Lambda) \\ &= \frac{1}{2} \sum_{i=1}^{n} \langle X_i, \sum_{j=1}^{n} l_{ij} X_j \rangle \\ &\quad + \sum_{i=1}^{n} \langle \Lambda_i, \bar{A}_i X_i + X_i \bar{B}_i - \bar{C}_i + \bar{L}_{i\cdot} M \rangle \\ &\quad + \frac{1}{2} \sum_{i=1}^{n} \|\bar{A}_i X_i + X_i \bar{B}_i - \bar{C}_i + \bar{L}_{i\cdot} M\|^2, \end{aligned} \tag{10}$$

where $\Lambda = [\Lambda_1^{\mathrm{T}}, \Lambda_2^{\mathrm{T}}, \dots, \Lambda_n^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{nm\times r}$ with $\Lambda_i \in \mathbb{R}^{m\times r}$, $\forall i \in \mathcal{V}$.

According to (10), we design the following fractional order continuous-time optimization algorithm from primal-dual viewpoint, i.e., gradient descend for primal variables X and M and gradient ascent for the dual variable Λ, and both primal and duel variables governed by the fractional order update law.

$$\begin{cases} \mathscr{D}^{\alpha_1} X_i(t) = -\nabla_{X_i} \mathcal{L}(X(t), M(t), \Lambda(t)), \\ \mathscr{D}^{\alpha_2} M_i(t) = -\nabla_{M_i} \mathcal{L}(X(t), M(t), \Lambda(t)), \\ \mathscr{D}^{\alpha_3} \Lambda_i(t) = \nabla_{\Lambda_i} \mathcal{L}(X(t), M(t), \Lambda(t)), \end{cases} \tag{11}$$

where $0 < \alpha_1, \alpha_2 < 2$, $0 < \alpha_3 < 1$, $\nabla_{X_i} \mathcal{L}(\cdot,\cdot,\cdot)$, $\nabla_{M_i} \mathcal{L}(\cdot,\cdot,\cdot)$, and $\nabla_{\Lambda_i} \mathcal{L}(\cdot,\cdot,\cdot)$ are gradients of $\mathcal{L}$ on variables $X_i(t)$, $M_i(t)$, and $\Lambda_i(t)$, respectively. Based on (10), the detailed dynamics of $X_i(t)$, $M_i(t)$, and $\Lambda_i(t)$ are expressed in Algorithm 1.

Algorithm 1

Initialization: For each $i \in \mathcal{V}$,
$X_i(0) \in \mathbb{R}^{m\times r},\ M_i(0) \in \mathbb{R}^{m\times r},\ \Lambda_i(0) \in \mathbb{R}^{m\times r}$.

Update flows: For each $i \in \mathcal{V}$,

$$\begin{aligned}\mathscr{D}^{\alpha_1}X_i(t) =& -\left[\bar{A}_i^{\mathrm{T}}\Lambda_i(t)+\Lambda_i(t)\bar{B}_i^{\mathrm{T}}+\bar{L}_{i\cdot}X(t)\right]-A_i^{\mathrm{T}}\left[\bar{A}_iX_i(t)+X_i(t)\bar{B}_i-\bar{C}_i+\bar{L}_{i\cdot}M(t)\right]\\ &-\left[\bar{A}_iX_i(t)+X_i(t)\bar{B}_i-\bar{C}_i+\bar{L}_{i\cdot}M(t)\right]B_i^{\mathrm{T}},\\ \mathscr{D}^{\alpha_2}M_i(t) =& -\bar{L}_{i\cdot}\Lambda(t)-\bar{L}_{i\cdot}\left[\bar{A}_1X_1(t)+X_1(t)\bar{B}_1-\bar{C}_1+\bar{L}_{1\cdot}M(t),\cdots,\bar{A}_nX_n(t)+X_n(t)\bar{B}_n-\bar{C}_n+\bar{L}_{n\cdot}M(t)\right],\\ \mathscr{D}^{\alpha_3}\Lambda_i(t) =& \ \bar{A}_iX_i(t)+X_i(t)\bar{B}_i-\bar{C}_i+\bar{L}_{i\cdot}M(t).\end{aligned}$$

***Remark* 2** In Algorithm 1, we design a distributed algorithm to solve Sylvester equations with fractional order iteration dynamics. Compared with conventional integer order scheme, the proposed algorithm has more design freedom and has potential to obtain better convergence performance than that of the conventional first order counterpart.

For Algorithm 1, we provide the following convergence results.

Theorem 1 *Under Assumption 1, let $X_i(t)$, $M_i(t)$, and $\Lambda_i(t)$ be generated by Algorithm 1. If $0<\alpha_1,\alpha_2,\alpha_3<1$, then $X_i(t)$ asymptotically converges to exact solutions of Sylvester equations.*

Proof By introducing Kronecker product, we rewrite the dynamic of $X_i(t)$ in Algorithm 1 in the following column form

$$\begin{aligned}\mathscr{D}^{\alpha_1}\boldsymbol{x}_i(t) = & -R_i^{\mathrm{T}}\left[R_i\boldsymbol{x}_i(t)-\boldsymbol{c}_i+\hat{L}_{i\cdot}\boldsymbol{m}(t)\right]\\ & -R_i^{\mathrm{T}}\lambda_i(t)-\hat{L}_{i\cdot}\boldsymbol{x}(t),\end{aligned} \tag{12}$$

where $\hat{L}_{i\cdot}=L_{i\cdot}\otimes I_{mr}\in\mathbb{R}^{mr\times nmr}$, $R_i=I_r\otimes\bar{A}_i+\bar{B}_i^{\mathrm{T}}\otimes I_m\in\mathbb{R}^{mr\times mr}$, $\boldsymbol{x}(t)=\mathrm{col}\{\boldsymbol{x}_1(t),\boldsymbol{x}_2(t),\ldots,\boldsymbol{x}_n(t)\}\in\mathbb{R}^{nmr}$, and $\boldsymbol{m}(t)=\mathrm{col}\{\boldsymbol{m}_1(t),\boldsymbol{m}_2(t),\ldots,\boldsymbol{m}_n(t)\}\in\mathbb{R}^{nmr}$, where $\boldsymbol{x}_i(t)\in\mathbb{R}^{mr}$, $\boldsymbol{m}_i(t)\in\mathbb{R}^{mr}$, and $\boldsymbol{c}_i(t)\in\mathbb{R}^{mr}$ are augmented column vectors by accumulating the column of $X_i(t)$, $M_i(t)$, and $\bar{C}_i$, respectively.

Then, we rewrite (12) as

$$\begin{aligned}\mathscr{D}^{\alpha_1}\boldsymbol{x}(t) = & -R^{\mathrm{T}}\left[R\boldsymbol{x}(t)-\boldsymbol{c}+\hat{L}\boldsymbol{m}(t)\right]\\ & -R^{\mathrm{T}}\lambda(t)-\hat{L}\boldsymbol{x}(t),\end{aligned} \tag{13}$$

where $R=\mathrm{blkdiag}\{R_1,R_2,\ldots,R_n\}$, $\boldsymbol{c}=\mathrm{col}\{\boldsymbol{c}_1,\boldsymbol{c}_2,\ldots,\boldsymbol{c}_n\}\in\mathbb{R}^{nmr}$, and $\lambda(t)=\mathrm{col}\{\lambda_1(t),\lambda_2(t),\ldots,\lambda_n(t)\}\in\mathbb{R}^{nmr}$ where $\lambda_i(t)\in\mathbb{R}^{mr}$ is an augmented column vector by accumulating the column of $\Lambda_i(t)$. Considering the equilibrium of (13)

$$\mathbf{0}=-R^{\mathrm{T}}\left[R\boldsymbol{x}^*-\boldsymbol{c}+\hat{L}\boldsymbol{m}^*\right]-R^{\mathrm{T}}\lambda^*-\hat{L}\boldsymbol{x}^*. \tag{14}$$

Then, combining (14) and (13),

$$\mathscr{D}^{\alpha_1}\tilde{\boldsymbol{x}}(t)=-R^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)+\tilde{\lambda}(t)\right]-\hat{L}\tilde{\boldsymbol{x}}(t). \tag{15}$$

According to Lemma 1, we rewrite (15) as

$$\begin{cases}\dfrac{\partial z_1(\omega,t)}{\partial t} = -\omega z_1(\omega,t)-R^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)\right.\\ \qquad\qquad\qquad \left.+\tilde{\lambda}(t)\right]-\hat{L}\tilde{\boldsymbol{x}}(t),\\ \tilde{\boldsymbol{x}}(t)=\displaystyle\int_0^{+\infty}\mu_{\alpha_1}(\omega)z_1(\omega,t)\mathrm{d}\omega,\end{cases} \tag{16}$$

where $z_1(\omega,t)\in\mathbb{R}^{nmr}$.

Similarly, for $M_i(t)$ and $\Lambda_i(t)$ of Algorithm 1, we have

$$\begin{cases}\dfrac{\partial z_2(\omega,t)}{\partial t}=-\omega z_2(\omega,t)-\hat{L}^{\mathrm{T}}\left[\tilde{\lambda}(t)+R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)\right],\\ \tilde{\boldsymbol{m}}(t)=\displaystyle\int_0^{+\infty}\mu_{\alpha_2}(\omega)z_2(\omega,t)\mathrm{d}\omega,\\ \dfrac{\partial z_3(\omega,t)}{\partial t}=-\omega z_3(\omega,t)+R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t),\\ \tilde{\lambda}(t)=\displaystyle\int_0^{+\infty}\mu_{\alpha_3}(\omega)z_3(\omega,t)\mathrm{d}\omega,\end{cases} \tag{17}$$

where $z_2(\omega,t)\in\mathbb{R}^{nmr}$ and $z_3(\omega,t)\in\mathbb{R}^{nmr}$.

Consider the following Lyapunov function

$$V_1=\frac{1}{2}\int_0^{+\infty}\sum_{i=1}^{3}\mu_{\alpha_i}(\omega)\|z_i(\omega,t)\|^2\mathrm{d}\omega. \tag{18}$$

Note that

$$
\begin{aligned}
\dot V_1 &= \int_0^{+\infty} \textstyle\sum_{i=1}^{3} \mu_{\alpha_i}(\omega)\langle z_i(\omega,t), \dfrac{\partial z_i(\omega,t)}{\partial t}\rangle \mathrm{d}\omega \\
&= \int_0^{+\infty} \mu_{\alpha_1}(\omega)\langle z_1(\omega,t), -\omega z_1(\omega,t) - R^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) \\
&\quad + \hat{L}\tilde{\boldsymbol{m}}(t) + \tilde{\lambda}(t)\big] - \hat{L}\tilde{\boldsymbol{x}}(t)\rangle \mathrm{d}\omega \\
&\quad + \int_0^{+\infty} \mu_{\alpha_2}(\omega)\langle z_2(\omega,t), -\omega z_2(\omega,t) \\
&\quad - \hat{L}^{\mathrm{T}}\big[\tilde{\lambda}(t) + R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big]\rangle \mathrm{d}\omega \\
&\quad + \int_0^{+\infty} \mu_{\alpha_3}(\omega)\langle z_3(\omega,t), -\omega z_3(\omega,t) + R\tilde{\boldsymbol{x}}(t) \\
&\quad + \hat{L}\tilde{\boldsymbol{m}}(t)\rangle \mathrm{d}\omega \\
&= - \int_0^{+\infty} \omega \textstyle\sum_{i=1}^{3} \mu_{\alpha_i}(\omega)\|z_i(\omega,t)\|^2 \mathrm{d}\omega \\
&\quad + \langle \tilde{\boldsymbol{x}}(t), -R^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t) + \tilde{\lambda}(t)\big] - \hat{L}\tilde{\boldsymbol{x}}(t)\rangle \\
&\quad + \langle \tilde{\boldsymbol{m}}(t), -\hat{L}^{\mathrm{T}}\big[\tilde{\lambda}(t) + R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big]\rangle \\
&\quad + \langle \tilde{\lambda}(t), R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\rangle \\
&= - \int_0^{+\infty} \omega \textstyle\sum_{i=1}^{3} \mu_{\alpha_i}(\omega)\|z_i(\omega,t)\|^2 \mathrm{d}\omega \\
&\quad - \langle \tilde{\boldsymbol{x}}(t), \hat{L}\tilde{\boldsymbol{x}}(t)\rangle - \|R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\|^2 \\
&\leqslant \mathbf{0}.
\end{aligned} \tag{19}
$$

According to the Lasalle invariance principle, this completes the proof. □

Next, we show a convergence result of Algorithm 1 by extending α_1 and α_2 from interval (0, 1) to interval (1, 2).

Theorem 2 *Under Assumption 1, let $X_i(t)$, $M_i(t)$, and $\Lambda_i(t)$ be generated by Algorithm 1. If $1 < \alpha_1 = \alpha_2 < 2$ and $\alpha_1 + \alpha_3 = 2$, then $X_i(t)$ asymptotically converges to an exact solution of Sylvester equations.*

Proof Similar to the proof of Theorem 1 and note that $\alpha_1 = 1 + \bar{\alpha}_1$ and $\alpha_2 = 1 + \bar{\alpha}_2$, we rewrite the dynamics of Algorithm 1 as follows:

$$
\begin{cases}
\dot{\tilde{\boldsymbol{x}}}_{\mathrm{a}}(t) = -R^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t) + \tilde{\lambda}(t)\big] - \hat{L}\tilde{\boldsymbol{x}}(t), \\
\mathscr{D}^{\bar{\alpha}_1}\tilde{\boldsymbol{x}}(t) = \tilde{\boldsymbol{x}}_{\mathrm{a}}(t), \\
\dot{\tilde{\boldsymbol{m}}}_{\mathrm{a}}(t) = -\hat{L}^{\mathrm{T}}\big[\tilde{\lambda}(t) + R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big], \\
\mathscr{D}^{\bar{\alpha}_2}\tilde{\boldsymbol{m}}(t) = \tilde{\boldsymbol{m}}_{\mathrm{a}}(t), \\
\mathscr{D}^{\alpha_3}\tilde{\lambda}(t) = R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t).
\end{cases} \tag{20}
$$

Apparently, $\alpha_1 = 1 + \bar{\alpha}_1$ and $\alpha_1 + \alpha_3 = 2$ implies $\bar{\alpha}_1 + \alpha_3 = 1$. As a result,

$$
\begin{aligned}
\dot{\tilde{\lambda}}(t) &= \mathscr{D}^{\bar{\alpha}_1}\mathscr{D}^{\alpha_3}\tilde{\lambda}(t) \\
&= \mathscr{D}^{\bar{\alpha}_1}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big] \\
&= R\tilde{\boldsymbol{x}}_{\mathrm{a}}(t) + \hat{L}\tilde{\boldsymbol{m}}_{\mathrm{a}}(t).
\end{aligned} \tag{21}
$$

Based on Lemma 1, we have the following frequency distributed model of (20):

$$
\begin{cases}
\dot{\tilde{\boldsymbol{x}}}_{\mathrm{a}}(t) = -R^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t) + \tilde{\lambda}(t)\big] - \hat{L}\tilde{\boldsymbol{x}}(t), \\
\dfrac{\partial z_1(\omega,t)}{\partial t} = -\omega z_1(\omega,t) + \tilde{\boldsymbol{x}}_{\mathrm{a}}(t), \\
\tilde{\boldsymbol{x}}(t) = \displaystyle\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega) z_1(\omega,t)\mathrm{d}\omega, \\
\dot{\tilde{\boldsymbol{m}}}_{\mathrm{a}}(t) = -\hat{L}^{\mathrm{T}}\big[\tilde{\lambda}(t) + R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big], \\
\dfrac{\partial z_2(\omega,t)}{\partial t} = -\omega z_2(\omega,t) + \tilde{\boldsymbol{m}}_{\mathrm{a}}(t), \\
\tilde{\boldsymbol{m}}(t) = \displaystyle\int_0^{+\infty} \mu_{\bar{\alpha}_2}(\omega) z_2(\omega,t)\mathrm{d}\omega, \\
\dot{\tilde{\lambda}}(t) = R\tilde{\boldsymbol{x}}_{\mathrm{a}}(t) + \hat{L}\tilde{\boldsymbol{m}}_{\mathrm{a}}(t).
\end{cases} \tag{22}
$$

Based on (22), we design the following Lyapunov function

$$
V_2 = V_{21} + V_{22}, \tag{23}
$$

where

$$
\begin{cases}
V_{21} = \dfrac{1}{2}\displaystyle\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\|Rz_1(\omega,t) + \hat{L}z_2(\omega,t)\|^2 \mathrm{d}\omega \\
\qquad + \dfrac{1}{2}\displaystyle\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\big\langle z_1(\omega,t), \hat{L}z_1(\omega,t)\big\rangle \mathrm{d}\omega, \\
V_{22} = \dfrac{1}{2}\|\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)\|^2 + \dfrac{1}{2}\|\tilde{\boldsymbol{m}}_{\mathrm{a}}(t)\|^2 + \dfrac{1}{2}\|\tilde{\lambda}(t)\|^2.
\end{cases} \tag{24}
$$

Note that

$$
\begin{aligned}
\dot V_{21} &= \int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\left\langle Rz_1(\omega,t) + \hat{L}z_2(\omega,t), R\frac{\partial z_1(\omega,t)}{\partial t} + \hat{L}\frac{\partial z_2(\omega,t)}{\partial t}\right\rangle \mathrm{d}\omega \\
&\quad + \int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\left\langle z_1(\omega,t), \hat{L}\frac{\partial z_1(\omega,t)}{\partial t}\right\rangle \mathrm{d}\omega.
\end{aligned} \tag{25}
$$

Substituting (22) into (25),

$$\begin{aligned}\dot{V}_{21}=&-\frac{1}{2}\int_0^{+\infty}\mu_{\bar{\alpha}_1}(\omega)\omega\|Rz_1(\omega,t)+\hat{L}z_2(\omega,t)\|^2\mathrm{d}\omega\\&-\frac{1}{2}\int_0^{+\infty}\mu_{\bar{\alpha}_1}(\omega)\omega\langle z_1(\omega,t),\hat{L}z_1(\omega,t)\rangle\mathrm{d}\omega\\&+\int_0^{+\infty}\mu_{\bar{\alpha}_1}(\omega)\langle Rz_1(\omega,t)+\hat{L}z_2(\omega,t),R\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)+\hat{L}\tilde{\boldsymbol{m}}_{\mathrm{a}}(t)\rangle\mathrm{d}\omega\\&+\int_0^{+\infty}\mu_{\bar{\alpha}_1}(\omega)\langle z_1(\omega,t),\hat{L}\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)\rangle\mathrm{d}\omega\\=&-\frac{1}{2}\int_0^{+\infty}\mu_{\bar{\alpha}_1}(\omega)\omega\|Rz_1(\omega,t)+\hat{L}z_2(\omega,t)\|^2\mathrm{d}\omega\\&-\frac{1}{2}\int_0^{+\infty}\mu_{\bar{\alpha}_1}(\omega)\omega\langle z_1(\omega,t),\hat{L}z_1(\omega,t)\rangle\mathrm{d}\omega\\&+\langle R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t),R\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)+\hat{L}\tilde{\boldsymbol{m}}_{\mathrm{a}}(t)\rangle\\&+\langle\tilde{\boldsymbol{x}}(t),\hat{L}\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)\rangle.\end{aligned}\tag{26}$$

Similarly,

$$\begin{aligned}\dot{V}_{22}=&\langle\tilde{\boldsymbol{x}}_{\mathrm{a}}(t),\dot{\tilde{\boldsymbol{x}}}_{\mathrm{a}}(t)\rangle+\langle\tilde{\boldsymbol{m}}_{\mathrm{a}}(t),\dot{\tilde{\boldsymbol{m}}}_{\mathrm{a}}(t)\rangle+\langle\tilde{\lambda}(t),\dot{\tilde{\lambda}}(t)\rangle\\=&\langle\tilde{\boldsymbol{x}}_{\mathrm{a}}(t),-R^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)+\tilde{\lambda}(t)\right]-\hat{L}\tilde{\boldsymbol{x}}(t)\rangle\\&+\langle\tilde{\boldsymbol{m}}_{\mathrm{a}}(t),-\hat{L}^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)+\tilde{\lambda}(t)\right]\rangle\\&+\langle\tilde{\lambda}(t),R\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)+\hat{L}\tilde{\boldsymbol{m}}_{\mathrm{a}}(t)\rangle\\=&-\langle R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t),R\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)+\hat{L}\tilde{\boldsymbol{m}}_{\mathrm{a}}(t)\rangle\\&-\langle\tilde{\boldsymbol{x}}(t),\hat{L}\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)\rangle.\end{aligned}\tag{27}$$

Combining (28) with (27) yields

$$\begin{aligned}\dot{V}_2=&-\frac{1}{2}\int_0^{+\infty}\mu_{\bar{\alpha}_1}(\omega)\omega\|Rz_1(\omega,t)+\hat{L}z_2(\omega,t)\|^2\mathrm{d}\omega\\&-\frac{1}{2}\int_0^{+\infty}\mu_{\bar{\alpha}_1}(\omega)\omega\langle z_1(\omega,t),\hat{L}z_1(\omega,t)\rangle\mathrm{d}\omega\\\leqslant&\,\mathbf{0}.\end{aligned}\tag{28}$$

According to the Lasalle invariance principle, this completes the proof. □

***Remark* 3** In [12], the authors utilized derivative feedback in their algorithm and then obtained the convergence result. In this paper, we design Algorithm 1 based on augmented Lagrangian function and does not need derivative feedback.

3.2 Obtaining the least squares solution

Based on (9), Algorithm 1 can achieve exact solutions of (3). However, Algorithm 1 cannot handle the circumstance that (3) has no exact solution. Then, we design a novel optimization model to achieve least squares solution next.

According to (8), we develop the following distributed optimization problem

$$\begin{aligned}\min_{X,M}\ &\frac{1}{2}\textstyle\sum_{i=1}^n\|\bar{A}_iX_i+X_i\bar{B}_i-\bar{C}_i+\bar{L}_{i\cdot}M\|^2,\\\text{s.t.}\ &\bar{L}X=0.\end{aligned}\tag{29}$$

It can be verified that the problems (3) and (29) are equivalent in the sense of least square. Based on (29), we construct an augmented Lagrangian function as

$$\begin{aligned}\mathcal{L}(X,M,\Lambda)=&\frac{1}{2}\textstyle\sum_{i=1}^n\|\bar{A}_iX_i+X_i\bar{B}_i-\bar{C}_i+\bar{L}_{i\cdot}M\|^2\\&+\langle\Lambda,\bar{L}X\rangle+\frac{1}{2}\langle X,\bar{L}X\rangle,\end{aligned}\tag{30}$$

where $\Lambda=[\Lambda_1^{\mathrm{T}},\Lambda_2^{\mathrm{T}},\ldots,\Lambda_n^{\mathrm{T}}]^{\mathrm{T}}\in\mathbb{R}^{nm\times r}$ with $\Lambda_i\in\mathbb{R}^{m\times r}$, $\forall i\in\mathcal{V}$.

According to (30), we design the following fractional order continuous-time optimization algorithm from primal-dual viewpoint, i.e., gradient descend for primal variables X and M and gradient ascent for the dual variable Λ, and both primal and duel variables governed by the fractional order update law.

$$\begin{cases}\mathscr{D}^{\alpha_1}X_i(t)=-\nabla_{X_i}\mathcal{L}(X(t),M(t),\Lambda(t)),\\\mathscr{D}^{\alpha_2}M_i(t)=-\nabla_{M_i}\mathcal{L}(X(t),M(t),\Lambda(t)),\\\mathscr{D}^{\alpha_3}\Lambda_i(t)=\nabla_{\Lambda_i}\mathcal{L}(X(t),M(t),\Lambda(t)),\end{cases}\tag{31}$$

where $0<\alpha_1,\ \alpha_2<2,\ 0<\alpha_3<1$. Based on (31), we express the detailed dynamics of $X_i(t)$, $M_i(t)$, and $\Lambda_i(t)$ in Algorithm 2.

For Algorithm 2, we provide the following results.

Theorem 3 *Under Assumption 1, let $X_i(t)$, $M_i(t)$, and $\Lambda_i(t)$ be generated by Algorithm 2. If $0<\alpha_1,\alpha_2,\alpha_3<1$, then $X_i(t)$ asymptotically converges to a least squares solution of Sylvester equations.*

Algorithm 2

Initialization: For each $i\in\mathcal{V}$,
$X_i(0)\in\mathbb{R}^{m\times r},\ M_i(0)\in\mathbb{R}^{m\times r},\ \Lambda_i(0)\in\mathbb{R}^{m\times r}$.

Update flows: For each $i\in\mathcal{V}$,

$\mathscr{D}^{\alpha_1}X_i(t)=-\bar{L}_{i\cdot}[X(t)+\Lambda(t)]-\bar{A}_i^{\mathrm{T}}[\bar{A}_iX_i(t)+X_i(t)\bar{B}_i-\bar{C}_i+\bar{L}_{i\cdot}M(t)]+[\bar{A}_iX_i(t)+X_i(t)\bar{B}_i-\bar{C}_i+\bar{L}_{i\cdot}M(t)]B_i^{\mathrm{T}}$,

$\mathscr{D}^{\alpha_2}M_i(t)=-\bar{L}_{i\cdot}[\bar{A}_1X_1(t)+X_1(t)\bar{B}_1-\bar{C}_1+\bar{L}_{1\cdot}M(t),\cdots,\bar{A}_nX_n(t)+X_n(t)\bar{B}_n-\bar{C}_n+\bar{L}_{n\cdot}M(t)]$,

$\mathscr{D}^{\alpha_3}\Lambda_i(t)=\bar{L}_{i\cdot}X(t)$.

Proof By introducing Kronecker product, we rewrite the dynamic of $X_i(t)$ in Algorithm 2 in the following column form:

$$\begin{aligned}\mathscr{D}^{\alpha_1}\boldsymbol{x}_i(t) = & -R_i^{\mathrm{T}}\left[\boldsymbol{x}_i(t)-\hat{L}_{i\cdot}\boldsymbol{m}(t)-\boldsymbol{c}_i\right]\\ & -\hat{L}_{i\cdot}(\lambda(t)+\boldsymbol{x}(t)),\end{aligned} \tag{32}$$

where $\hat{L}_{i\cdot}=L_{i\cdot}\otimes I_{mr}\in\mathbb{R}^{mr\times nmr}$ and $R_i=I_r\otimes\bar{A}_i+\bar{B}_i^{\mathrm{T}}\otimes I_m\in\mathbb{R}^{mr\times mr}$.

$$\boldsymbol{x}(t)=\mathrm{col}\{\boldsymbol{x}_1(t),\boldsymbol{x}_2(t),\dots,\boldsymbol{x}_n(t)\}\in\mathbb{R}^{nmr},$$
$$\boldsymbol{m}(t)=\mathrm{col}\{\boldsymbol{m}_1(t),\boldsymbol{m}_2(t),\dots,\boldsymbol{m}_n(t)\}\in\mathbb{R}^{nmr},$$
$$\lambda(t)=\mathrm{col}\{\lambda_1(t),\lambda_2(t),\dots,\lambda_n(t)\}\in\mathbb{R}^{nmr},$$

where $\boldsymbol{x}_i(t)\in\mathbb{R}^{mr}$, $\boldsymbol{m}_i(t)\in\mathbb{R}^{mr}$, $\lambda_i(t)\in\mathbb{R}^{mr}$, and $\boldsymbol{c}_i(t)\in\mathbb{R}^{mr}$ are augmented column vectors by accumulating the column of $X_i(t)$, $M_i(t)$, $\Lambda_i(t)$, and $\bar{C}_i$, respectively.

Then, we rewrite (32) as follows:

$$\begin{aligned}\mathscr{D}^{\alpha_1}\boldsymbol{x}(t) = & -R^{\mathrm{T}}\left[R\boldsymbol{x}(t)+\hat{L}\boldsymbol{m}(t)-\boldsymbol{c}\right]\\ & -\hat{L}\left[\boldsymbol{x}(t)+\lambda(t)\right].\end{aligned} \tag{33}$$

Considering the equilibrium of (33)

$$R^{\mathrm{T}}\boldsymbol{c}=-R^{\mathrm{T}}\left(R\boldsymbol{x}^*+\hat{L}\boldsymbol{m}^*\right)-\hat{L}\left(\boldsymbol{x}^*+\lambda^*\right). \tag{34}$$

Then, substituting (34) into (33) and defining $\tilde{\boldsymbol{x}}(t)=\boldsymbol{x}(t)-\boldsymbol{x}^*$, $\tilde{\boldsymbol{m}}(t)=\boldsymbol{m}(t)-\boldsymbol{m}^*$, $\tilde{\lambda}(t)=\lambda(t)-\lambda^*$,

$$\begin{aligned}\mathscr{D}^{\alpha_1}\tilde{\boldsymbol{x}}(t) = & -R^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)\right]\\ & -\hat{L}\left[\tilde{\boldsymbol{x}}(t)+\tilde{\lambda}(t)\right].\end{aligned} \tag{35}$$

According to Lemma 1, we rewrite (35) as

$$\begin{cases}\dfrac{\partial z_1(\omega,t)}{\partial t} = -\omega z_1(\omega,t)-R^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)\right]\\ \qquad\qquad -\hat{L}\left[\tilde{\boldsymbol{x}}(t)+\tilde{\lambda}(t)\right],\\ \tilde{\boldsymbol{x}}(t)=\displaystyle\int_0^{+\infty}\mu_{\alpha_1}(\omega)z_1(\omega,t)\mathrm{d}\omega,\end{cases} \tag{36}$$

where $z_1(\omega,t)\in\mathbb{R}^{nmr}$.

Similarly, for the dynamic of $M_i(t)$ and $\Lambda_i(t)$ of Algorithm 2, we have

$$\begin{cases}\dfrac{\partial z_2(\omega,t)}{\partial t}=-\omega z_2(\omega,t)-\hat{L}^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)\right],\\ \tilde{\boldsymbol{m}}(t)=\displaystyle\int_0^{+\infty}\mu_{\alpha_2}(\omega)z_2(\omega,t)\mathrm{d}\omega,\\ \dfrac{\partial z_3(\omega,t)}{\partial t}=-\omega z_3(\omega,t)+\hat{L}\tilde{\boldsymbol{x}}(t),\\ \tilde{\lambda}(t)=\displaystyle\int_0^{+\infty}\mu_{\alpha_3}(\omega)z_3(\omega,t)\mathrm{d}\omega.\end{cases} \tag{37}$$

where $z_2(\omega,t)\in\mathbb{R}^{nmr}$ and $z_2(\omega,t)\in\mathbb{R}^{nmr}$.

Consider the following Lyapunov function

$$V_3=\frac{1}{2}\int_0^{+\infty}\sum_{i=1}^{3}\mu_{\alpha_i}(\omega)\|z_i(\omega,t)\|^2\mathrm{d}\omega. \tag{38}$$

Then,

$$\dot{V}_3=\int_0^{+\infty}\sum_{i=1}^{3}\mu_{\alpha_i}(\omega)\left\langle z_i(\omega,t),\frac{\partial z_i(\omega,t)}{\partial t}\right\rangle\mathrm{d}\omega. \tag{39}$$

Substituting (36) and (37) into (39),

$$\begin{aligned}\dot{V}_3 = & \int_0^{+\infty}\mu_{\alpha_1}(\omega)z_1^{\mathrm{T}}(\omega,t)\big(-\omega z_1(\omega,t)-R^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)\right.\\ & \left.+\hat{L}\tilde{\boldsymbol{m}}(t)\right]-\hat{L}\left[\tilde{\boldsymbol{x}}(t)+\tilde{\lambda}(t)\right]\big)\mathrm{d}\omega\\ & +\int_0^{+\infty}\mu_{\alpha_2}(\omega)z_2^{\mathrm{T}}(\omega,t)\big(-\omega z_2(\omega,t)-\hat{L}^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)\right.\\ & \left.+\hat{L}\tilde{\boldsymbol{m}}(t)\right]\big)\mathrm{d}\omega\\ & +\int_0^{+\infty}\mu_{\alpha_3}(\omega)z_3^{\mathrm{T}}(\omega,t)\big(-\omega z_3(\omega,t)+\hat{L}\tilde{\boldsymbol{x}}(t)\big)\mathrm{d}\omega\\ = & -\int_0^{+\infty}\omega\textstyle\sum_{i=1}^{3}\mu_{\alpha_i}(\omega)\|z_i(\omega,t)\|^2\mathrm{d}\omega\\ & -\tilde{\boldsymbol{x}}^{\mathrm{T}}(t)R^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)\right]-\tilde{\boldsymbol{x}}^{\mathrm{T}}(t)\hat{L}\left[\tilde{\boldsymbol{x}}(t)+\tilde{\lambda}(t)\right]\\ & -\tilde{\boldsymbol{m}}^{\mathrm{T}}(t)\hat{L}^{\mathrm{T}}\left[R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)\right]+\tilde{\lambda}^{\mathrm{T}}\hat{L}\tilde{\boldsymbol{x}}(t)\\ = & -\int_0^{+\infty}\omega\textstyle\sum_{i=1}^{3}\mu_{\alpha_i}(\omega)\|z_i(\omega,t)\|^2\mathrm{d}\omega\\ & -\|R\tilde{\boldsymbol{x}}(t)+\hat{L}\tilde{\boldsymbol{m}}(t)\|^2-\tilde{\boldsymbol{x}}^{\mathrm{T}}(t)\hat{L}\tilde{\boldsymbol{x}}(t)\\ \leqslant & \ \boldsymbol{0}.\end{aligned} \tag{40}$$

According to the Lasalle invariance principle, this completes the proof. □

Next, we show a result by extending α_1 and α_2 from interval (0, 1) to interval (1, 2).

Theorem 4 *Under Assumption 1, let $X_i(t)$, $M_i(t)$, and $\Lambda_i(t)$ be generated by Algorithm 2. If* $1<\alpha_1=\alpha_2<2$ and $\alpha_1+\alpha_3=2$, *then $X_i(t)$ asymptotically converges to the least squares solution of Sylvester equations.*

Proof Similar to the proof of Theorem 1 and note that $\alpha_1=1+\bar{\alpha}_1$ and $\alpha_2=1+\bar{\alpha}_2$, we rewrite the dynamics of Algorithm 2 as follows:

$$\begin{cases} \dot{\tilde{\boldsymbol{x}}}_{\mathrm{a}}(t) = -R^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big] - \hat{L}\big[\tilde{\boldsymbol{x}}(t) + \tilde{\lambda}(t)\big], \\ \mathscr{D}^{\bar{\alpha}_1}\tilde{\boldsymbol{x}}(t) = \tilde{\boldsymbol{x}}_{\mathrm{a}}(t), \\ \dot{\tilde{\boldsymbol{m}}}_{\mathrm{a}}(t) = -\hat{L}^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big], \\ \mathscr{D}^{\bar{\alpha}_2}\tilde{\boldsymbol{m}}(t) = \tilde{\boldsymbol{m}}_{\mathrm{a}}(t), \\ \mathscr{D}^{\alpha_3}\tilde{\lambda}(t) = \hat{L}\tilde{\boldsymbol{x}}(t). \end{cases} \tag{41}$$

Since $\alpha_1 = 1 + \bar{\alpha}_1$ and $\alpha_1 + \alpha_3 = 2$ implies $\bar{\alpha}_1 + \alpha_3 = 1$. As a result,

$$\begin{aligned} \dot{\tilde{\lambda}}(t) &= \mathscr{D}^{\bar{\alpha}_1}\mathscr{D}^{\alpha_3}\tilde{\lambda}(t) \\ &= \mathscr{D}^{\bar{\alpha}_1}\hat{L}\tilde{\boldsymbol{x}}(t) \\ &= \hat{L}\tilde{\boldsymbol{x}}_{\mathrm{a}}(t). \end{aligned} \tag{42}$$

According to Lemma 1, we have the following frequency distributed model of (41)

$$\begin{cases} \dot{\tilde{\boldsymbol{x}}}_{\mathrm{a}}(t) = -R^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big] - \hat{L}\big[\tilde{\boldsymbol{x}}(t) + \tilde{\lambda}(t)\big], \\ \dfrac{\partial z_1(\omega, t)}{\partial t} = -\omega z_1(\omega, t) + \tilde{\boldsymbol{x}}_{\mathrm{a}}(t), \\ \tilde{\boldsymbol{x}}(t) = \displaystyle\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega) z_1(\omega, t)\mathrm{d}\omega, \\ \dot{\tilde{\boldsymbol{m}}}_{\mathrm{a}}(t) = -\hat{L}^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big], \\ \dfrac{\partial z_2(\omega, t)}{\partial t} = -\omega z_2(\omega, t) + \tilde{\boldsymbol{m}}_{\mathrm{a}}(t), \\ \tilde{\boldsymbol{m}}(t) = \displaystyle\int_0^{+\infty} \mu_{\bar{\alpha}_2}(\omega) z_2(\omega, t)\mathrm{d}\omega, \\ \dot{\tilde{\lambda}}(t) = \hat{L}\tilde{\boldsymbol{x}}_{\mathrm{a}}(t). \end{cases} \tag{43}$$

Based on (43), we design the following Lyapunov function

$$V_4 = V_{41} + V_{42}, \tag{44}$$

where,

$$\begin{cases} V_{41} = \dfrac{1}{2}\displaystyle\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\|Rz_1(\omega, t) + \hat{L}z_2(\omega, t)\|^2\mathrm{d}\omega \\ \qquad + \dfrac{1}{2}\displaystyle\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\langle z_1(\omega, t), \hat{L}z_1(\omega, t)\rangle\mathrm{d}\omega, \\ V_{42} = \dfrac{1}{2}\|\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)\|^2 + \dfrac{1}{2}\|\tilde{\boldsymbol{m}}_{\mathrm{a}}(t)\|^2 + \dfrac{1}{2}\|\tilde{\lambda}(t)\|^2. \end{cases} \tag{45}$$

Note that,

$$\begin{aligned} \dot{V}_{41} = &\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\langle Rz_1(\omega, t) + \hat{L}z_2(\omega, t), \\ &R\frac{\partial z_1(\omega, t)}{\partial t} + \hat{L}\frac{\partial z_2(\omega, t)}{\partial t}\rangle\mathrm{d}\omega \\ &+ \int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\langle z_1(\omega, t), \hat{L}\frac{\partial z_1(\omega, t)}{\partial t}\rangle\mathrm{d}\omega. \end{aligned} \tag{46}$$

Substituting (43) into (46),

$$\begin{aligned} \dot{V}_{41} = &-\frac{1}{2}\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\omega\|Rz_1(\omega, t) + \hat{L}z_2(\omega, t)\|^2\mathrm{d}\omega \\ &- \frac{1}{2}\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\omega\langle z_1(\omega, t), \hat{L}z_1(\omega, t)\rangle\mathrm{d}\omega \\ &+ \int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\langle Rz_1(\omega, t) + \hat{L}z_2(\omega, t), R\tilde{\boldsymbol{x}}_{\mathrm{a}}(t) + \hat{L}\tilde{\boldsymbol{m}}_{\mathrm{a}}(t)\rangle\mathrm{d}\omega \\ &+ \int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\langle z_1(\omega, t), \hat{L}\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)\rangle\mathrm{d}\omega \\ = &-\frac{1}{2}\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\omega\|Rz_1(\omega, t) + \hat{L}z_2(\omega, t)\|^2\mathrm{d}\omega \\ &- \frac{1}{2}\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\omega\langle z_1(\omega, t), \hat{L}z_1(\omega, t)\rangle\mathrm{d}\omega \\ &+ \langle R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t), R\tilde{\boldsymbol{x}}_{\mathrm{a}}(t) + \hat{L}\tilde{\boldsymbol{m}}_{\mathrm{a}}(t)\rangle \\ &+ \langle \tilde{\boldsymbol{x}}(t), \hat{L}\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)\rangle. \end{aligned} \tag{47}$$

Similarly,

$$\begin{aligned} \dot{V}_{42} = &\langle \tilde{\boldsymbol{x}}_{\mathrm{a}}(t), \dot{\tilde{\boldsymbol{x}}}_{\mathrm{a}}(t)\rangle + \langle \tilde{\boldsymbol{m}}_{\mathrm{a}}(t), \dot{\tilde{\boldsymbol{m}}}_{\mathrm{a}}(t)\rangle + \langle \tilde{\lambda}(t), \dot{\tilde{\lambda}}(t)\rangle \\ = &\langle \tilde{\boldsymbol{x}}_{\mathrm{a}}(t), -R^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big] - \hat{L}\big[\tilde{\boldsymbol{x}}(t) + \tilde{\lambda}(t)\big]\rangle \\ &+ \langle \tilde{\boldsymbol{m}}_{\mathrm{a}}(t), -\hat{L}^{\mathrm{T}}\big[R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t)\big]\rangle \\ &+ \langle \tilde{\lambda}(t), \hat{L}\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)\rangle \\ = &-\langle R\tilde{\boldsymbol{x}}(t) + \hat{L}\tilde{\boldsymbol{m}}(t), R\tilde{\boldsymbol{x}}_{\mathrm{a}}(t) + \hat{L}\tilde{\boldsymbol{m}}_{\mathrm{a}}(t)\rangle \\ &- \langle \tilde{\boldsymbol{x}}(t), \hat{L}\tilde{\boldsymbol{x}}_{\mathrm{a}}(t)\rangle. \end{aligned} \tag{48}$$

Combining (49) with (48),

$$\begin{aligned} \dot{V}_4 = &-\frac{1}{2}\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\omega\|Rz_1(\omega, t) + \hat{L}z_2(\omega, t)\|^2\mathrm{d}\omega \\ &- \frac{1}{2}\int_0^{+\infty} \mu_{\bar{\alpha}_1}(\omega)\omega\langle z_1(\omega, t), \hat{L}z_1(\omega, t)\rangle\mathrm{d}\omega \\ \leqslant &\,\mathbf{0}. \end{aligned} \tag{49}$$

Following from the Lasalle invariance principle, this completes the proof. □

***Remark* 4** For distributed solving Sylvester equation (3), we decoupled it as (8). Then, one utilized consensus constrain as an objective function and the Sylvester equation as an equation constrain in [12] and Sect. 3.1 of this paper. However, the corresponding optimization problem (9) is only effective for exact solutions. Moreover, we developed a novel distributed optimization model (29), which is effective for least squares solutions.

***Remark* 5** Both in Algorithms 1 and 2, the dynamics can be viewed as a class of stable linear time-invariant fractional order systems. Consider a stable linear time-invariant fractional order system $\mathscr{D}^{\alpha}\boldsymbol{x}(t) = A\boldsymbol{x}(t)$, which has a fast response but chattering trajectory if $\alpha \in (1, 2)$, while

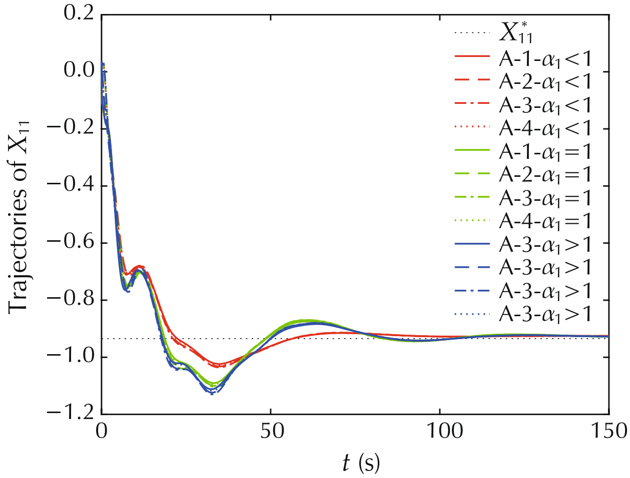

Fig. 1 The trajectories of estimation of X_{11}

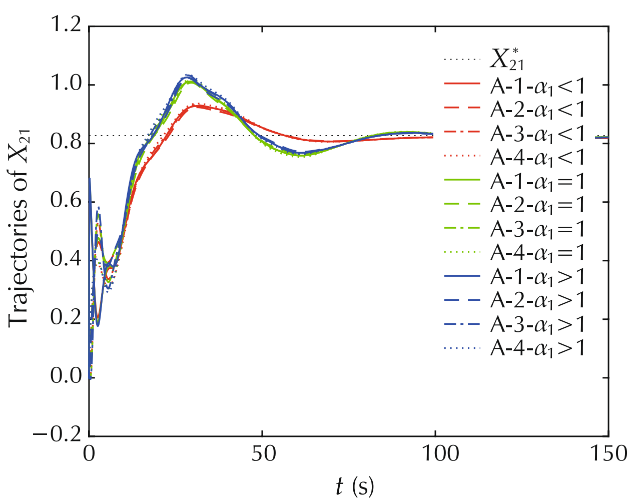

Fig. 2 The trajectories of estimation of X_{21}

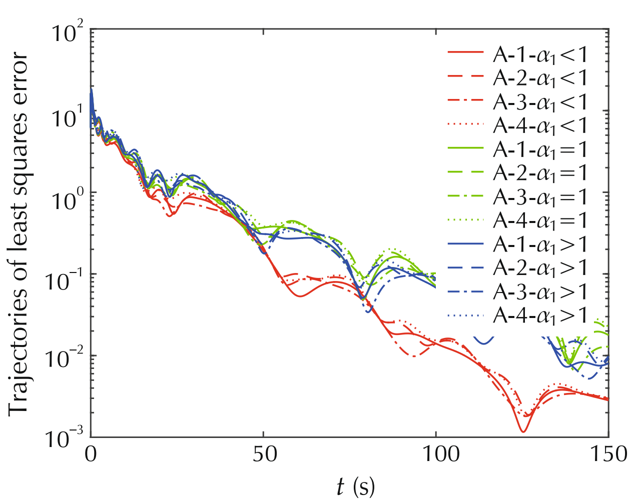

Fig. 3 The trajectories of exact solution error $\|AX + XB - C\|$

it has a slow response but smooth trajectory if $\alpha \in (0, 1)$. In comparison with [12–14], the proposed algorithms with fractional order dynamics have more design freedom. Therefore, we can find a proper update order to adjust the trade-off between the response rate and the smooth trajectory to obtain better convergence performance than that of the first order algorithms.

4 Numerical example

In this section, we provide an example to validate the effectiveness of the proposed algorithm with the following three cases

1. Fractional order dynamics with $\alpha_1 = \alpha_2 = \alpha_3 = 0.95$;
2. integer order dynamics with $\alpha_1 = \alpha_2 = \alpha_3 = 1.00$; and
3. fractional order dynamics with $\alpha_1 = \alpha_2 = 1.05, \alpha_3 = 0.95$,

where the fractional order operators of this example are implemented via a piecewise numerical approximation algorithm in the frequency domain [23].

Consider a Sylvester equation with the Row–Column–Column structure, whose parameter matrices are randomly chosen as

$$\begin{aligned}
A_{1\cdot} &= [3, 8, 3, 4],\ B_{\cdot 1} = [6, 5, 6, 6]^{\mathrm{T}},\ C_{\cdot 1} = [2, 7, 6, 4]^{\mathrm{T}},\\
A_{2\cdot} &= [2, 8, 4, 8],\ B_{\cdot 2} = [2, 2, 8, 2]^{\mathrm{T}},\ C_{\cdot 2} = [2, 4, 4, 1]^{\mathrm{T}},\\
A_{3\cdot} &= [4, 1, 5, 3],\ B_{\cdot 3} = [1, 5, 8, 6]^{\mathrm{T}},\ C_{\cdot 3} = [5, 2, 4, 5]^{\mathrm{T}},\\
A_{4\cdot} &= [3, 6, 8, 6],\ B_{\cdot 4} = [2, 3, 4, 8]^{\mathrm{T}},\ C_{\cdot 4} = [3, 3, 5, 3]^{\mathrm{T}}.
\end{aligned}$$

It is not hard to verify that the corresponding solution is

$$X^* = \begin{bmatrix} -0.9258 & 0.2154 & 0.2541 & 0.2164 \\ 0.8193 & -0.1204 & 0.0038 & 0.0158 \\ 0.9454 & -0.2641 & 0.1657 & 0.1717 \\ -0.7395 & 0.5161 & 0.0570 & 0.0418 \end{bmatrix}.$$

We solve the Sylvester equation with the aid of a multi-agent system, where four agents are connected by an undirected circular graph and the ith agent has access to sub-blocks $A_{i\cdot}$, $B_{\cdot i}$, and $C_{\cdot i}$.

We present the trajectories of elements $X_{11}(t)$, $X_{21}(t)$ and exact solution error $\|AX_i(t) + X_i(t)B - C\|$ in Figs. 1, 2 and 3, where "A-i" denotes the corresponding curve of agent i. The simulation results illuminate that the proposed algorithm with fractional order between 0 and 2 asymptotically achieves the solution of the Sylvester equation. Moreover, with more design freedom of fractional order, we obtain better convergence performance than that of the first order counterpart.

5 Conclusions

In this paper, we have constructed two distributed optimization models to achieve the exact or least squares solutions of Sylvester equations. We have proposed two distributed algorithm with fractional order dynamics from the primal-dual viewpoint. By transferring the proposed algorithms into corresponding frequency distributed models, we designed Lyapunov functions and then proved the asymptotic convergence of proposed algorithms.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China (Nos. 61903027, 61973002), and in part by the National Postdoctoral Program for Innovative Talents (BX20180346), and in part by the General Financial Grant from the China Postdoctoral Science Foundation (2019M660834), and in part by the Anhui Provincial Natural Science Foundation (No. 2008085J32).

References

1. Gajic, Z., & Qureshi, M. (2008). *Lyapunov matrix equation in system stability and control*. Dover.
2. Li, W. (2018). A recurrent neural network with explicitly definable convergence time for solving time-variant linear matrix equations. *IEEE Transactions on Industrial Informatics*, *14*(12), 5289–5298.
3. Agarwal, A., Peng, J., & Milenkovic, O. (2019). Online convex matrix factorization with representative regions. In *Proceedings of the 33rd conference on neural information processing systems (NeurIPS 2019),* Vancouver, Canada.
4. Duan, G. (2015). *Generalized Sylvester equations: Unified parametric solutions*. CRC Press.
5. Zhang, Z., & Zheng, L. (2018). A complex varying parameter convergent-differential neural-network for solving online time-varying complex Sylvester equation. *IEEE Transactions on Cybernetics*, *49*(10), 3627–3639.
6. Wang, Q., He, Z., & Zhang, Y. (2019). Constrained two-sided coupled Sylvester-type quaternion matrix equations. *Automatica*, *101*, 207–213.
7. Sylvester, J. J. (1884). Sur l'équation en matrices $px = xq$. *Comptes Rendus de l'AcadÃ©des Sciences*, *Paris*, *99*(2), 67–71.
8. Wimmer, H. K. (1996). The generalized Sylvester equation in polynomial matrices. *IEEE Transactions on Automatic Control*, *41*(9), 1372–1376.
9. Yi, P., & Hong, Y. (2016). Distributed cooperative optimization and its applications. *Scientia Sinica Mathematica*, *46*(10), 1547–1564.
10. Liang, S., Zeng, X., & Hong, Y. (2017). Distributed nonsmooth optimization with coupled inequality constraints via modified Lagrangian function. *IEEE Transactions on Automatic Control*, *63*(6), 1753–1759.
11. Liu, S., Qiu, Z., & Xie, L. (2017). Convergence rate analysis of distributed optimization with projected subgradient algorithm. *Automatica*, *83*, 162–169.
12. Deng, W., Zeng, X., & Hong, Y. (2020). Distributed computation for solving the Sylvester equation based on optimization. *IEEE Systems Control Letters*, *4*(2), 414–419.
13. Chen, G., Zeng, X., & Hong, Y. (2019). Distributed optimisation design for solving the Stein equation with constraints. *IET Control Theory & Applications*, *13*(15), 2492–2499.
14. Zeng, X., Liang, S., Hong, Y., & Chen, J. (2018). Distributed computation of linear matrix equations: An optimization perspective. *IEEE Transactions on Automatic Control*, *64*(5), 1858–1873.
15. Lu, J., & Chen, Y. (2010). Robust stability and stabilization of fractional-order interval systems with the fractional order: The case $0 < \alpha < 1$. *IEEE Transactions on Automatic Control*, *55*(1), 152–158.
16. Wei, Y., Liu, D., & Boutat, D. (2019). Innovative fractional derivative estimation of the pseudo-state for a class of fractional order linear systems. *Automatica*, *99*, 157–166.
17. Wei, Y., Chen, Y., Wang, J., & Wang, Y. (2019). Analysis and description of the infinite-dimensional nature for nabla discrete fractional order systems. *Communications in Nonlinear Science and Numerical Simulation*, *72*, 472–492.
18. Cheng, S., Wei, Y., Chen, Y., Li, Y., & Wang, Y. (2017). An innovative fractional order LMS based on variable initial value and gradient order. *Signal Processing*, *133*, 260–269.
19. Sheng, H., Chen, Y., & Qiu, T. (2012). *Fractional processes and fractional-order signal processing*. Springer.
20. Liu, T., Wei, Y., Yin, W., Wang, Y., & Liang, Q. (2019). State estimation for nonlinear discrete-time fractional systems: A Bayesian perspective. *Signal Processing*, *150*, 250–261.
21. Monje, C., Chen, Y., Vinagre, B., Xue, D., & Feliu-Batlle, V. (2010). *Fractional-order systems and controls: Fundamentals and applications*. Springer.
22. Trigeassou, J., Maamri, N., Sabatier, J., & Oustaloup, A. (2012). Transients of fractional-order integrator and derivatives. *Signal, Image Video Process*, *6*(3), 359–372.
23. Wei, Y., Gao, Q., Peng, C., & Wang, Y. (2014). Rational approximate method to fractional order systems. *International Journal of Control Automation and Systems*, *12*(6), 1180–1186.

Songsong Cheng received the Ph.D. degree in Engineering from the University of Science and Technology of China, Hefei, China, in 2018. From 2018 to 2020, he was a postdoctoral fellow with the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China. He is currently a lecturer with the School of Electrical Engineering and Automation, Anhui University, Hefei, China. His current research interests include distributed computation, optimization, and game.

Shu Liang received the B.E. degree in Automatic Control and the Ph.D. degree in Engineering from the University of Science and Technology of China, Hefei, China, in 2010 and 2015, respectively. From 2015 to 2017, he was a postdoctoral fellow with the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China. He is currently an

associate professor with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China. His research interests include non-smooth systems and control, distributed optimizations, game theory, and fractional order systems.

Yuan Fan received the B.E. degree in Automation from the University of Science and Technology of China in 2006, and the Ph.D. degree in Control Science and Engineering from the University of Science and Technology of China and the City University of Hong Kong in 2011. He was a research fellow in the Nanyang Technological University in 2015, and a senior research associate in the City University of Hong Kong in 2018. He has been with Anhui University since 2011, where he is currently a professor. His main research interests include networked dynamic systems, multi-agent coordination, event-triggered control, and robot control.