Data-Driven Distributed Spectrum Estimation for Linear Time-Invariant Systems

Shenyu Liu Jorge Cortés Sonia Martínez

Abstract—**This paper tackles spectrum estimation of a linear time-invariant system by a multi-agent network using data. We consider a group of agents that communicate over a strongly connected, aperiodic graph and do not have any knowledge of the system dynamics. Each agent only measures some signals that are linear functions of the system states or inputs, and does not know the functional form of this dependence. The proposed distributed algorithm consists of two steps that rely on the collected data: (i) the identification of an unforced trajectory of the system and (ii) the estimation of the coefficients of the characteristic polynomial of the system matrix using this unforced trajectory. We show that each step can be formulated as a problem of finding a common solution to a set of linear algebraic equations which are amenable to distributed algorithmic solutions. We prove that, under mild assumptions on the collected data, when the initial condition of the system is random, the proposed distributed algorithm accurately estimates the spectrum with probability 1.**

I. Introduction

System identification and spectral analysis find numerous applications, including signal processing [1], power systems [2], process control [3], and structural engineering [4], [5]. Despite the fact that algorithms for spectrum estimation of linear time-invariant (LTI) systems are theoretically well-established [6], the problem becomes practically challenging in the case when the global network topology or the overall dynamics are not accessible to a central agent performing the analysis, either due to bandwidth limitations, geographical constraints, privacy considerations, or other specifications. As a result, the spectral analysis cannot be performed by a centralized method and distributed algorithms become more suitable. In addition, distributed algorithms can offer added benefits, such as scalability with the size of the system and robustness against single points of failure. These considerations lead us to the problem of collaborative estimation of the spectrum by a group of agents, where any individual agent only has partial measurements of the data generated by the system and can only communicate locally with others.

Literature review: Spectral analysis approaches for LTI systems have been available for a long time. The classical Ho-Kalman algorithm [7] employs the so-called system Markov parameters to estimate an equivalent representation of the system matrices. The work [8] builds on this to develop the Eigensystem Realization Algorithm (ERA), which can deal with faulty sensors or incomplete unit-pulse-response data. The main limitation of the Ho-Kalman algorithm and ERA is that their inputs must be unit impulses. Instead, subspace-based state-space system identification (4SID) methods [9], [10] can be used to estimate the eigenvalues of an LTI system when inputs are arbitrary. All the aforementioned methods are centralized, relying on singular value decomposition (SVD) and matrix multiplications. In addition, 4SID methods rely on oblique projections, which represent a challenge for distributed implementation. Additional centralized system identification techniques can be found in [6].

Building on frequency domain analysis techniques [11] for system identification, [12], [13] propose distributed algorithms based on Fast Fourier Transforms (FFT) to perform spectral clustering of the nodes of networks and estimate the eigenvalues of their Laplacian matrices. The work [14] presents a decentralized algorithm for computing eigenvectors of a symmetric matrix building on the Power Method [15, Chapter 4.1.1]. The Power Method has also been used for distributed estimation of the algebraic connectivity of undirect [16] and direct [17] graphs. The work [18] provides distributed tests for stability of large-scale interconnected systems. Particularly relevant to our work are [19], which proposes a discrete-time, distributed algorithm for spectrum estimation and [20], which finds eigenvalues and eigenvectors at the same time. The recent work [21] further extends these methods to determine eigenvectors of matrices on spatially distributed networks. Nevertheless, all these works have in common the constraint that the matrix whose spectral properties are to be analyzed needs to be the weighted adjacency matrix of the network and the fact that the algorithms rely on state information generated by autonomous systems. Our problem formulation is different and, in particular, we allow the system to be subject to arbitrary inputs. Our spectrum estimation algorithm is data-driven, in the sense that the analysis is based on both input and output data. Our approach relies on the body of work that designs distributed algorithms to solve linear algebraic equations (LAEs) [22], [23], [24], [25], [26] and distributed optimization problems [27], [28], [29].

Statement of contributions: We design a data-driven distributed algorithm for spectrum estimation of unknown linear time-invariant systems by a multi-agent network. Our setup does not assume that agents can measure partial states of the system or have knowledge of the plant's parameters, and in that sense is more general than those previously considered in the literature. Instead, each agent has access to a signal that is a linear transformation of the system's state and/or input. However, agents have no knowledge of the functional form of

Shenyu Liu is with the School of Automation, Beijing Institute of Technology, shenyuliu@bit.edu.cn. Jorge Cortés and Sonia Martínez are with the Department of Aerospace and Mechanical Engineering, University of California, San Diego, {cortes, soniamd}@ucsd.edu. This work was partially supported by AFOSR Award FA9550-19-1-0235 and ARO Award W911NF-23-1-0138.

these linear transformations or the system dynamics.

Our solution strategy proceeds in two steps: first, agents compute an unforced output trajectory of the system based on their measured data. Using this unforced output trajectory, agents then compute the characteristic polynomial of the system matrix, which in turn allows them to estimate the spectrum. We show how each of these steps can be cast as solving a system of linear-algebraic equations, for which employ a distributed algorithm. To study the convergence properties of our desigm we assume mild assumptions on joint system observability and joint input reconstrutability by the multi-agent network. When the collected data is generated by arbitrary inputs and the algorithm is randomly initialized according to a generic distribution, we provide sufficient conditions to show that the proposed distributed algorithmic procedure accurately estimates the spectrum with probability 1. We also particularize our treatment to the case when inputs are generated by another LTI system and provide conditions for accurate estimation in terms of the spectrum of the system matrices. Simulations on a mass/spring/damper system illustrate our results.

Organization: The paper is organized as follows. Section II provides prelimaries on linear algebra and distributed algorithmic solutions to systems of LAEs. Section III introduces our assumptions on the multi-agent network and formulates the distributed spectrum estimation problem. Section IV presents our two-phase algorithm design and Section V establishes its correctness, paying attention to the identification of sufficient conditions that guarantee the accurate estimation of the system spectrum. We analyze two cases: when the inputs to the system are arbitrary and when they are generated by another LTI system. Finally, Section VI illustrates the performance of the proposed distributed algorithm on a mass/spring/damper system and Section VII gathers our conclusions and ideas for future work.

II. Preliminaries

This section describes the notation and basic concepts from linear algebra, graph theory, and distributed algorithms.

Notation: We denote by $\mathbb{N} := \{0, 1, 2, \cdots\}$ the set of non-negative integers, $\mathbb{R}$ the set of real numbers, $\mathbb{R}^n$ the n-dimensional real space, and $\mathbb{R}^{n\times m}$ the space of $n \times m$ real matrices. In particular, $0_{n\times m}$, (resp. $1_{n\times m}$) denotes the $n \times m$-dimensional zero matrix (resp. all-ones matrix), while I_n represents the $n \times n$ identity matrix. When the dimensions are clear from the context, we remove the subindices. The transpose of a matrix is denoted by the superscript $^\top$ and the Kronecker product by $\otimes$. For a given discrete-time signal $x : \mathbb{N} \to \mathbb{R}^n$ and $i, j \in \mathbb{N}$ with $i \le j$, we use the shorthand notation $x_{i:j} := \begin{bmatrix} x(i)^\top & x(i+1)^\top & \cdots & x(j)^\top \end{bmatrix}^\top$. Moreover, with $i, j, k \in \mathbb{N}$, $j, k \ge 1$, we denote the Hankel matrix $H_{i,j,k}(x) \in \mathbb{R}^{nj\times k}$ by

$$H_{i,j,k}(x) := \begin{bmatrix} x(i) & x(i+1) & \cdots & x(i+k-1) \\ x(i+1) & x(i+2) & \cdots & x(i+k) \\ \vdots & \vdots & \ddots & \vdots \\ x(i+j-1) & x(i+j) & \cdots & x(i+j+k-2) \end{bmatrix}.$$

We have $x_{i:j} = H_{i,j-i+1,1}(x)$. Note that one needs $j + k - 1$ consecutive samples of x (from $x(i)$ to $x(i + j + k - 2)$) to form the Hankel matrix $H_{i,j,k}(x)$. A random variable X has a *generic* probability distribution over an n-dimensional vector space $\mathcal{X}$ if, for any proper subspace or affine space $\mathcal{Y} \subset \mathcal{X}$ with $\dim(\mathcal{Y}) < \dim(\mathcal{X})$, it holds that $\Pr(X \in \mathcal{Y}) = 0$. Note that the multivariate Gaussian distribution over $\mathbb{R}^n$ and uniform distribution over an n-dimensional ball in $\mathbb{R}^n$ are generic.

Basic notions from linear algebra: Following [30], for any matrix $A \in \mathbb{R}^{n\times n}$, denote the set of all eigenvalues of A as $\mathrm{spec}(A)$. Given any matrix $M \in \mathbb{R}^{n\times m}$, we denote its image and kernel as $\mathrm{im}(M) := \{y \in \mathbb{R}^n : y = Mx \text{ for some } x \in \mathbb{R}^m\}$ and $\ker(M) := \{x \in \mathbb{R}^m : Mx = 0\}$. The rank and nullity of M are denoted as $\mathrm{rank}(M) := \dim(\mathrm{im}(M))$ and $\mathrm{nullity}(M) := \dim(\ker(M))$. We have $\mathrm{rank}(M) + \mathrm{nullity}(M) = m$. For any vector space $\mathcal{X} \subseteq \mathbb{R}^n$, denote $\mathcal{X}^\perp := \{y \in \mathbb{R}^n : x^\top y = 0, \forall x \in \mathcal{X}\}$. The orthogonal projection matrix onto the vector space $\mathcal{X}$ is given by a matrix $P_\mathcal{X} \in \mathbb{R}^{n\times n}$ such that $P_\mathcal{X} x = x$, for all $x \in \mathcal{X}$, and $P_\mathcal{X} x = 0$, for all $x \in \mathcal{X}^\perp$. In particular, for any $M \in \mathbb{R}^{n\times m}$, $P_{\mathrm{im}(M)} = MM^\dagger$, where $M^\dagger$ is the Moore-Penrose pseudo-inverse of M.

Basic notions from graph theory: Following [31], a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a vertex set $\mathcal{V} := \{1, 2, \cdots, p\}$ and an edge set $\mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$. The adjacency matrix A of $\mathcal{G}$ is a $p \times p$ matrix such that $\mathsf{A}_{ij} = 1$ if $(i, j) \in \mathcal{E}$ and $\mathsf{A}_{ij} = 0$ otherwise. A path is a sequence of vertices connected by edges, and a cycle is a path whose first and last vertices are the same. A directed graph is *strongly connected* if there is a path between any pair of vertices. The graph is *aperiodic* if there is no integer other than 1 that divides the number of edges in every cycle. Given $i \in \mathcal{V}$, the set of in-neighbors of i is $N_i^{\text{in}} := \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}$ and the set of its out-neighbors is $N_i^{\text{out}} := \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$. The (out-)degree d_i of an vertex $i \in \mathcal{V}$ is defined as the cardinality of N_i^{out}.

Distributed algorithmic solution to linear equations: Consider p linear algebraic equations (LAEs)

$$M_i z_i = l_i, \quad i \in \mathcal{V} = \{1, \cdots, p\}, \tag{1}$$

where $M_i \in \mathbb{R}^{n_i\times n}$, $l_i \in \mathbb{R}^{n_i}$ are the system data and $z_i \in \mathbb{R}^n$ are the variables. A common solution $z = z_1 = \cdots = z_p \in \mathbb{R}^n$ to (1) corresponds to a solution z of

$$Mz = l, \tag{2}$$

where

$$M := \begin{bmatrix} M_1 \\ \vdots \\ M_p \end{bmatrix}, \quad l := \begin{bmatrix} l_1 \\ \vdots \\ l_p \end{bmatrix}.$$

A group of p agents communicating over a graph $\mathcal{G}$ aims to solve (2). We assume that each agent $i \in \mathcal{V}$

- knows the value of M_i, l_i and can find a solution to the i-th equation in (1);
- can share information with its in-neighbors N_i^{in} (excluding the values of M_i, l_i for privacy considerations).

Algorithm 1 presents a distributed algorithm from [24] that solves this problem. The following result extends the asymp-

Algorithm 1 Distributed LAE solver

Input: M_i, l_i for each agent $i \in \mathcal{V}$, K

Output: $z_i(K) \quad \forall i \in \mathcal{V}$

for $i \in \mathcal{V}$ **do**

  Agent i computes $z_i(0)$ such that

$$M_i z_i(0) = l_i \tag{3}$$

end for

for $k = 0, 1, \cdots, K-1$ **do**

  for $i \in \mathcal{V}$ **do**

    Agent i broadcasts $z_i(k)$ to in-neighbors $j \in N_i^{\text{in}}$

    Agent i receives $z_j(k)$ from out-neighbors $j \in N_i^{\text{out}}$

    Agent i updates its value of z_i by

$$z_i(k+1) = z_i(k) + \frac{1}{d_i} P_{\ker(M_i)} \sum_{j \in N_i^{\text{out}}} (z_j(k) - z_i(k)) \tag{4}$$

  end for

end for

totic convergence result of [24, Theorem 1] by providing an explicit formula for the computed solution.

Theorem II.1 (Convergence of the distributed LAE solver)**.** *Suppose there exists a solution to* (2)*. Let $\mathcal{G}$ be strongly connected and aperiodic graph, and denote by* A *its adjacency matrix,* $\mathsf{D} := \operatorname{diag}(d_1, \cdots, d_p)$*, and $w \in \mathbb{R}^n$ the left Perron-Frobenius eigenvector of* $\mathsf{D}^{-1}\mathsf{A}$*. Then, there exists $\lambda \in (0,1)$, and $c \geq 1$ such that the output $z(k) := \begin{bmatrix} z_1(k)^\top & \cdots & z_p(k)^\top \end{bmatrix}^\top$ of Algorithm 1 satisfies*

$$|z(k) - z_{\text{c}}| \leq c\lambda^k \quad \forall k \in \mathbb{N}, \tag{5}$$

where

$$z_{\text{c}} := 1_{p \times 1} \otimes (P_{\ker(M)} Z(0) w + z_{\text{mn}}), \tag{6}$$

with $Z(0) := \begin{bmatrix} z_1(0) & \cdots & z_p(0) \end{bmatrix}$ and $z_{\text{mn}} \in \mathbb{R}^n$ the minimum-norm solution of (2)*.*

We provide the proof of Theorem II.1 in the Appendix.

Remark II.2 (Expression for the computed solution of linear equations). The explicit formula (6) of the computed solution z_{c} provided in Theorem II.1 consists of two parts. The first part $P_{\ker(M)} Z(0) w$ is a weighted sum of the projections of $z_i(0)$'s onto the kernel of M. This depends on the agents' initial guesses of a common solution and is independent of the true solutions of (2). The second part z_{mn} is independent of the agents' initial guesses. Moreover, the Kronecker product with 1 in the expression of z_{c} implies that $z_i(k)$ will all converge to the same vector $P_{\ker(M)} Z(0) w + z_{\text{mn}}$, which is a solution to (2). When (2) has a unique solution, that is, when $\ker(M) = \{0\}$, then $P_{\ker(M)} \equiv 0$ and z_{mn} is the unique solution. •

III. Problem formulation

Here, we formally state the distributed spectrum estimation problem. Consider a discrete-time LTI system

$$x(t+1) = Ax(t) + Bu(t), \tag{7}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the input and A, B are the system and input matrices of compatible dimensions. Our setup is as follows: p agents aim to collaboratively estimate $\operatorname{spec}(A)$ via a distributed algorithm over a strongly connected, aperiodic communication graph $\mathcal{G}$. In addition, each agent $i \in \mathcal{V}$

- has no knowledge of the matrices A, B;
- can observe a signal $y_i(t) = C_i x(t) \in \mathbb{R}^{n_{y,i}}$ and a signal $v_i(t) = E_i u(t) \in \mathbb{R}^{n_{v,i}}$, but does not know the matrices C_i, E_i either;
- can send information to its in-neighbors (excluding the values of $y_i(t), v_i(t)$ for privacy considerations).

Note that the input u is an external signal and cannot be customized by the agents. This means that u is arbitrary and may not be particularly suited for spectrum estimation. Our goal is to design a distributed spectrum estimation algorithm that relies solely on the collected data $y_i(t), v_i(t)$ from the system (7). Figure 1 shows a graphical illustration of the problem. We make the following assumptions.

Assumption 1 (Joint observability)**.** The matrix pair (A, C) is observable, where $C := \begin{bmatrix} C_1^\top & C_2^\top & \cdots C_p^\top \end{bmatrix}^\top$.

Assumption 2 (Joint rank condition on observed inputs)**.** The matrix $E := \begin{bmatrix} E_1^\top & E_2^\top & \cdots E_p^\top \end{bmatrix}^\top$ has full column rank.

In other words, while each individual agent only has access to partial information about the inputs and states (which, individually, is insufficient for the spectrum estimation), the system's inputs are collectively known and the system is observable to the group. These assumptions are justified by the fact that, without them, the agents would not be able to estimate the spectrum of A. A distributed solution to this problem is challenging since the agents do not directly share their collected data with neighbors.

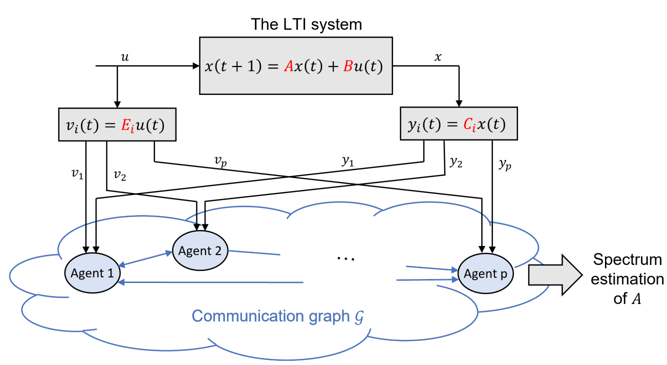


Fig. 1: Illustration of the distributed spectrum estimation problem. The matrices marked in red are unknown to the agents.

Remark III.1 (Agents with no measurements of inputs or outputs). The setup described above does not rule out the case $C_i = 0$ and $E_i \neq 0$, and vice versa, indicating cases where the i-th agent is capable of either measuring inputs or outputs. It is also possible that both $C_i = 0$ and $E_i = 0$ for some $i \in \mathcal{V}$, representing agents which only participate in sharing information with others but do not measure any signal about the system. We point out that the total number of agents measuring states is not critical and it can be as small as 1, provided that the observability condition of Assumption 1 is satisfied (the spectrum of an observable linear system with no inputs can be estimated based on a single output via the Ho-Kalman method [7] or other techniques). On the other hand, we do require the agents to collectively know the full input set. This is because if some inputs are not measured at all, the agents will be unable to distinguish if the "patterns" observed in the states are caused by the spectral properties of the system or by the unmeasured inputs. •

IV. Two-phase distributed spectrum estimation

We now present our distributed method for spectrum estimation, which consists of two phases. First, agents compute an unforced output trajectory in a distributed way based on their measured data. Second, using this unforced output trajectory, agents compute in a distributed way the characteristic polynomial of the matrix A. This allows them to estimate the spectrum as the roots of the polynomial.

A. Gathering information about an unforced trajectory

In the first phase, the system (7) evolves from time 0 to a sufficiently large time $T-1$, with $T \geq 2n$. Each agent collects data pairs $y_i(t), v_i(t)$ of length T. Using the linearity of the system (7) (see [32, Section 2]), we have for any $g \in \mathbb{R}^{T-2n+1}$, the linear combination

$$\begin{bmatrix} \bar{u}_{0:2n-1} \\ \bar{x}_{0:2n-1} \end{bmatrix} := \begin{bmatrix} H_{0,2n,T-2n+1}(u) \\ H_{0,2n,T-2n+1}(x) \end{bmatrix} g, \tag{8}$$

is an input/state trajectory of length $2n$ of the system (7). If this corresponds to an unforced trajectory, $\bar{u}_{0:2n-1} = 0$ must hold. Therefore,

$$H_{0,2n,T-2n+1}(u) g = 0. \tag{9}$$

Left multiplying this equation by $I_{2n} \otimes E$ and using the identity $v_i(t) = E_i u(t)$, we obtain

$$H_{0,2n,T-2n+1}(v) g = 0, \tag{10}$$

where $v := \begin{bmatrix} v_1^\top & \cdots & v_p^\top \end{bmatrix}^\top$. Moreover, since E is full column rank, cf. Assumption 2, equation (9) holds if and only if (10) holds. Now, observe that the LAE (10) can be split into p subsets of LAEs, one per agent, as follows

$$H_{0,2n,T-2n+1}(v_i) g = 0, \quad \forall i \in \mathcal{V}. \tag{11}$$

Now because agent i has access to v_i, this agent can set up the corresponding LAE in (11). Thus, the problem of finding g takes exactly the form of (1) with $M_i = H_{0,2n,T-2n+1}(v_i)$ and $l_i = 0$, which can be solved in a distributed way via Algorithm 1.

At the same time, from (8), we deduce that $\bar{x}_{0:2n-1} = H_{0,2n,T-2n+1}(x) g$. Left multiplying by $I_{2n} \otimes C_i$ and using the identity $y_i(t) = C_i x(t)$ yields

$$(\bar{y}_i)_{0:2n-1} = H_{0,2n,T-2n+1}(y_i) g, \tag{12}$$

where $\bar{y}_i(t) := C_i \bar{x}(t)$. In other words, after obtaining $g \in \mathbb{R}^{T-2n+1}$ which satisfies (9), each agent can employ (12) to compute the corresponding unforced output trajectory $\bar{y}_i$ using its own original output observations y_i. The second phase of the algorithm employs these unforced output trajectories to determine the characteristic polynomial of the system matrix A.

B. Calculation of the characteristic polynomial

We next show how to estimate the spectrum of A. Recall that the eigenvalues of A are the roots of its characteristic polynomial

$$p_A(s) := \det(sI - A) = s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0. \tag{13}$$

The roots of this polynomial can be estimated numerically, e.g., using the Aberth method [33]. Thus, each agent only needs to determine the coefficients of p_A to find the spectrum of A. To this end, notice that it follows from the Cayley-Hamilton theorem [34, Theorem 2.1] that

$$p_A(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = 0. \tag{14}$$

For each $i \in \mathcal{V}$, left-multiplying this equation by C_i and right-multiplying by $\bar{x}(t)$ yield the expression

$$\bar{y}_i(t+n) + a_{n-1}\bar{y}_i(t+n-1) + \cdots + a_1 \bar{y}_i(t+1) + a_0 \bar{y}_i(t) = 0, \tag{15}$$

where we have used $\bar{y}_i(t+r) = C_i \bar{x}_i(t+r) = C_i A^r \bar{x}(t)$. Since the LAE (15) holds for any $t \in \mathbb{N}$ and $i \in \mathcal{V}$, we subtract the term $\bar{y}_i(t+n)$ on both sides and stack the LAEs from $t=0$ to $t=n-1$ together to conclude

$$H_{0,n,n}(\bar{y}_i) a = -(\bar{y}_i)_{n:2n-1}, \quad \forall i \in \mathcal{V}, \tag{16}$$

where $a := \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \end{bmatrix}^\top \in \mathbb{R}^n$ is the vector of coefficients of $p_A(s)$ in reverse order. Note that the problem of finding the common solution a to the LAEs (16) is in the form of (1) with $M_i = H_{0,n,n}(\bar{y}_i)$ and $l_i = -(\bar{y}_i)_{n:2n-1}$. Hence, a can be found by the agents using the distributed LAE solver.

Algorithm 2 presents the two-phase distributed spectrum estimation algorithm. Because of the exponential rate of the distributed LAE solver, cf. Theorem II.1, the convergence of Algorithm 2 is fast and hence a relatively small number of K timesteps is sufficient for achieving consensus on the values of g_i and a_i, $i \in \mathcal{V}$.

Algorithm 2 Distributed spectrum estimation algorithm

Input: $T, K \in \mathbb{N}$, each agent $i \in \mathcal{V}$ with data $\{(v_i)_{0:T-1}, (y_i)_{0:T-1}\}$

Output: Estimated eigenvalues $\{\lambda_j^i\}_{j=1}^n$ for each agent $i \in \mathcal{V}$

For each agent $i \in \mathcal{V}$,

1: let g_i be the output of running Algorithm 1 with $M_i = H_{0,2n,T-2n+1}(v_i)$ and $l_i = 0$ for K timesteps
2: compute $(\bar{y}_i)_{0:2n-2} = H_{0,2n-1,T-2n+1}(y_i) g_i$
3: let $a_i = \begin{bmatrix} (a_i)_0 & \cdots & (a_i)_{n-1} \end{bmatrix}^\top$ be the output of running Algorithm 1 with $M_i = H_{0,n,n}(\bar{y}_i)$ and $l_i = -(\bar{y}_i)_{n:2n-1}$ for K timesteps
4: compute the roots $\{\lambda_j^i\}_{j=1}^n$ of the polynomial $s^n + (a_i)_{n-1}s^{n-1} + \cdots + (a_i)_1 s + (a_i)_0$

Remark IV.1 (Preservation of local data privacy). We discuss here the extent to which Algorithm 2 preserves the privacy of the local data. In the first execution of the distributed LAE solver, cf. Algorithm 1, agents aim to agree on a common solution $g \in \mathbb{R}^{T-2n+1}$ to (11). For the k-th timestep, the data broadcasted by agent i is only a vector $g_i(k) \in \ker(H_{0,2n,T-2n+1}(v_i))$. As a result, the in-neighbors of agent i can at most find the kernel space $\ker(H_{0,2n,T-2n+1}(v_i))$ based on the received vectors $g_i(k), k = 0, 1, \cdots K-1$, which are not enough to uniquely determine v_i. During the second execution of Algorithm 1, agents aim to agree on a common solution $a \in \mathbb{R}^n$ to the LAEs (16). By a similar argument, we see that neither the measurement y_i is directly transmitted to the neighboring agents, nor can it be determined by them based on the received data. •

V. Sufficient conditions for successful spectrum estimation

Here we characterize when the proposed distributed spectrum estimation algorithm correctly estimates the spectrum. In fact, for accurate spectrum estimation, we need the LAE (10) in the first stage of Algorithm 2 to be solved with a non-trivial solution, and the LAE (16) in the second stage of Algorithm 2 to have a unique solution. We provide conditions on the data initially collected by the agents that ensure these properties. In our treatment, we consider two scenarios for generating the data: when the inputs to the system (7) are arbitrary and when the inputs are generated by another LTI system.

A. Non-trivial LAE solution in first stage of distributed spectrum estimation

Note that since the LAE (10) is homogeneous, the trivial $g = 0$ is always a solution. Such solution is not useful for estimating $\mathrm{spec}(A)$, since it corresponds to the 0-state trajectory that is common to every linear system. The next result shows that, when non-trivial solution exists (i.e., $\mathrm{nullity}(H_{0,n,T-n+1}(v)) \neq 0$) and the initial conditions $z_i(0)$ in Algorithm 1 are randomly chosen, then the LAE solver almost surely obtains a non-trivial solution.

Proposition V.1 (Almost sure non-zero solution to the homogeneous LAE)**.** *Consider the LAE* (2) *with* $l = 0$ *and let* $S_1 := \prod_{i\in\mathcal{V}} \ker(M_i)$*. Let the initial condition* $Z(0) := (z_1(0), z_2(0), \cdots, z_p(0))$ *be a random variable with a generic probability distribution over* S_1 *and* z_{c} *be computed according to* (6)*. If* $\mathrm{nullity}(M) \neq 0$*, then* $\Pr(z_{\mathrm{c}} = 0) = 0$.

Proof. Since $l = 0$, the minimum-norm solution to (2) is $z_{\mathrm{mn}} = 0$. From (6), it holds that $z_{\mathrm{c}} = 0$ if and only if $\sum_{i\in\mathcal{V}} w_i z_i(0) \in \ker(M)^\perp$. Define $S_2 := \{(z_1, \cdots, z_p) : \sum_{i\in\mathcal{V}} w_i z_i \in \ker(M)^\perp\}$. Note that both S_1, S_2 are vector spaces. Consider the vector $z^* := (z_1^*, \cdots, z_p^*)$ with $z_1^* \in \ker(M)\backslash\{0\}$ and $z_i^* = 0$, for all $i \in \mathcal{V}\backslash\{1\}$. Because $\ker(M) \subseteq \ker(M_1)$, we have $z^* \in S_1$. In addition, since $\sum_{i\in\mathcal{V}} w_i z_i^* = w_1 z_1^* \notin \ker(M)^\perp$, then $z^* \notin S_2$. This means that $S_1 \cap S_2$ is a proper subspace of S_1. Finally because $Z(0)$ has a generic probability distribution over S_1, $\Pr(z_{\mathrm{c}} = 0) = \Pr(Z(0) \in S_1 \cap S_2) = 0$. □

Recall that in the first stage of the distributed spectrum estimation algorithm, we have $M = H_{0,2n,T-2n+1}(v)$. For a sufficiently large T, the columns of the Hankel matrix $H_{0,2n,T-2n+1}(v)$ become linearly dependent and hence $\mathrm{nullity}(M) \neq 0$. Thus, a random $Z(0)$ and sufficiently large T ensure the existence of a non-trivial solution with probability 1.

B. Unique LAE solution in second stage of distributed spectrum estimation

Finding a non-zero g in the first stage of Algorithm 1 is not enough to guarantee its correctness. The reason is that the satisfaction of the LAEs (16) are only a necessary condition for the vector a to correspond to the coefficients of the characteristic polynomial (13). In other words, if these LAEs have multiple common solutions, our algorithm may only find one of them, which may differ from the intended one. To address this issue, the next result provides a sufficient condition to ensure the uniqueness of the common solution to (16).

Lemma V.2 (Sufficient condition for a unique LAE solution)**.** *Let* $g \in \ker(H_{0,2n,T-2n+1}(v))$ *be such that the pair* $(A, H_{0,1,T-2n+1}(x)g)$ *is controllable and let* $(\bar{y})_{0:2n-1} = (H_{0,2n,T-2n+1}(y))g$*. Then, the coefficients of the characteristic polynomial* (13) *of* A *can be uniquely determined by solving the LAE* (17).

Proof. Firstly, observe that the system of LAEs (16) is equivalent to a single LAE

$$H_{0,n,n}(\bar{y})a = -(\bar{y})_{n:2n-1}, \tag{17}$$

where $\bar{y} = \begin{bmatrix}\bar{y}_1^\top & \bar{y}_2^\top & \cdots y_p^\top\end{bmatrix}^\top$. Therefore the LAEs (16) have a unique common solution if and only if $H_{0,n,n}(\bar{y})$ is full column rank.[1] Using the definition of Hankel matrix and the properties of the unforced trajectory, we have

$$\begin{aligned} H_{0,n,n}(\bar{y}) &= \begin{bmatrix} C\bar{x}(0) & CA\bar{x}(0) & \cdots & CA^{n-1}\bar{x}(0) \\ CA\bar{x}(0) & CA^2\bar{x}(0) & \cdots & CA^n\bar{x}(0) \\ \vdots & \vdots & \ddots & \vdots \\ CA^{n-1}\bar{x}(0) & CA^n\bar{x}(0) & \cdots & CA^{2n-2}\bar{x}(0) \end{bmatrix} \\ &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \begin{bmatrix}\bar{x}(0) & A\bar{x}(0) & \cdots & A^{n-1}\bar{x}(0)\end{bmatrix} \\ &:= \mathcal{O}(A,C)\mathcal{C}(A,\bar{x}(0)). \end{aligned}$$

Therefore, the LAE (17) has a unique solution if the observability matrix $\mathcal{O}(A,C)$ is full column rank and the square controllability matrix $\mathcal{C}(A,\bar{x}(0))$ is full rank. The former condition is guaranteed by the observability of (A,C), cf. Assumption 1, whereas the latter condition requires $(A,\bar{x}(0))$ to be controllable. Note that, from (8), the initial condition can be expressed as $\bar{x}(0) = H_{0,1,T-2n+1}(x)g$, so this condition is satisfied by hypothesis. □

This result provides a sufficient condition under which the proposed distributed algorithm can estimate the coefficients of the characteristic polynomial of A and, hence,

[1] The Hankel matrix $H_{0,n,n}(\bar{y})$ is not square if $\bar{y}$ is not scalar; nevertheless existence of solutions when $H_{0,n,n}(\bar{y})$ is not full row rank is guaranteed since the vector of the coefficients of the characteristic polynomial is always a solution.

compute its eigenvalues. However, the controllability condition in Lemma V.2 cannot be checked in practice since it involves knowledge of the system matrix A and the state x, which are both unknown. Our ensuing discussion addresses this limitation.

C. Sufficient conditions for correct spectrum estimation when inputs are random

Here we provide sufficient conditions that ensure correct spectrum estimation when the inputs employed to generate the data available to the agents are random. To this end, we build on the controllability condition in Lemma V.2. Note that, if A has an eigenvalue with geometric multiplicity larger than 1, then (A, b) is not controllable for any $b \in \mathbb{R}^n$, cf. [19]. In order to rule out this possibility, we make the following assumption.

Assumption 3 (Distinct system eigenvalues)**.** All the eigenvalues of the matrix A are distinct.

Assumption 3 implies that A is diagonalizable, which is generic. With this assumption, we are ready to state our main result.

Theorem V.3 (Almost sure correctness of distributed spectrum estimation for arbitrary inputs)**.** *Consider a discrete-time LTI system* (7) *and let Assumptions 1, 2 and 3 hold. Let* $\mathrm{spec}(A) := \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ *and denote*

$$\Lambda_i(t) := \begin{bmatrix} 1 & \lambda_i & \lambda_i^2 & \cdots & \lambda_i^t \end{bmatrix}^\top, \tag{18}$$

for $t \in \mathbb{N}$. Suppose data are collected for the system with an initial condition $x(0)$ which has a generic probability distribution over $\mathbb{R}^n$. In addition, assume that the initial conditions when running Algorithm 1 for finding g are also random with a generic probability distribution over $\prod_{i\in\mathcal{V}} \ker(H_{0,2n,T-2n+1}(v_i))$. If

$$\Lambda_i(T-2n) \notin \mathrm{im}(H_{0,2n,T-2n+1}(u)^\top),\ \forall i = 1, \cdots, n, \tag{19}$$

where $T \geq 2n$, then $\mathrm{spec}(A)$ can be accurately estimated with probability 1 by Algorithm 2 for sufficiently large K.

Proof. Given that A is diagonalizable, cf. Assumption 3, let $\Psi := \begin{bmatrix} \psi_1 & \psi_2 & \cdots & \psi_n \end{bmatrix} \in \mathbb{R}^{n\times n}$ be composed of right eigenvectors of A forming a basis of $\mathbb{R}^n$. Therefore, we can express $x(0) = \sum_{i=1}^n \alpha_i \psi_i$ and, for $s \in \mathbb{N}$, $Bu(s) = \sum_{i=1}^n \gamma_{s,i}\psi_i$. We then have

$$\begin{aligned} x(t) &= A^t x(0) + \sum_{s=0}^{t-1} A^{t-1-s}Bu(s) \\ &= \sum_{i=1}^n \lambda_i^t \alpha_i \psi_i + \sum_{s=0}^{t-1}\sum_{i=1}^n \lambda_i^{t-1-s}\gamma_{s,i}\psi_i \\ &= \sum_{i=1}^n \Big(\lambda_i^t\alpha_i + \sum_{s=0}^{t-1}\lambda_i^{t-1-s}\gamma_{s,i}\Big)\psi_i \\ &= \Psi \begin{bmatrix} \lambda_1^t\alpha_1 + \sum_{s=0}^{t-1}\lambda_1^{t-1-s}\gamma_{s,1} \\ \lambda_2^t\alpha_2 + \sum_{s=0}^{t-1}\lambda_2^{t-1-s}\gamma_{s,2} \\ \vdots \\ \lambda_n^t\alpha_n + \sum_{s=0}^{t-1}\lambda_n^{t-1-s}\gamma_{s,n} \end{bmatrix}. \end{aligned}$$

From (8), we have that $\bar{x}(0) = [x(0)\ x(1)\ \cdots\ x(T-2n)]g$. Using the equation above, we can write

$$\bar{x}(0) = \Psi \begin{bmatrix} \Lambda_1(T-2n)^\top(\alpha_1 I_{T-2n+1} + \Gamma_1(T)) \\ \Lambda_2(T-2n)^\top(\alpha_2 I_{T-2n+1} + \Gamma_2(T)) \\ \vdots \\ \Lambda_n(T-2n)^\top(\alpha_n I_{T-2n+1} + \Gamma_n(T)) \end{bmatrix} g, \tag{20}$$

where $\Gamma_i(T)$ is the $(T-2n+1)\times(T-2n+1)$ upper-triangular Toeplitz matrix given by

$$\Gamma_i(T) := \begin{bmatrix} 0 & \gamma_{0,i} & \gamma_{1,i} & \cdots & \gamma_{T-2n-1,i} \\ & 0 & \gamma_{0,i} & \cdots & \gamma_{T-2n-2,i} \\ & & \ddots & \ddots & \vdots \\ & & & 0 & \gamma_{0,i} \\ & & & & 0 \end{bmatrix}.$$

Let φ_i be a left eigenvector of A corresponding to the eigenvalue λ_i. According to Lemma A.1, the pair $(A, \bar{x}(0))$ is controllable if and only if $\varphi_i^\top \bar{x}(0) \neq 0$ for all $i = 1, \cdots, n$. The relation between left and right eigenvectors implies $\varphi_i^\top \psi_j \neq 0$ if and only if $i = j$. Therefore, from (20), it follows that

$$\begin{aligned} \varphi_i^\top \bar{x}(0) &= \varphi_i^\top \Psi \begin{bmatrix} \Lambda_1(T-2n)^\top(\alpha_1 I_{T-2n+1} + \Gamma_1(T)) \\ \Lambda_2(T-2n)^\top(\alpha_2 I_{T-2n+1} + \Gamma_2(T)) \\ \vdots \\ \Lambda_n(T-2n)^\top(\alpha_n I_{T-2n+1} + \Gamma_n(T)) \end{bmatrix} g \\ &= (\varphi_i^\top \psi_i)\Lambda_i(T-2n)^\top(\alpha_i I_{T-2n+1} + \Gamma_i(T))g, \end{aligned}$$

which is non-zero if and only if

$$\Lambda_i(T-2n)^\top(\alpha_i I_{T-2n+1} + \Gamma_i(T))g \neq 0, \tag{21}$$

for all $i = 1, \cdots, n$. From the condition (19), we have that $\Lambda_i(T-2n) \notin \ker(H_{0,n,T-n+1}(u))^\perp$. Taking the orthogonal space on both sides, we conclude $\ker(H_{0,n,T-n+1}(u)) \not\subset \mathrm{im}(\Lambda_i(T-2n))^\perp = \ker(\Lambda_i(T-2n)^\top)$. Hence, the dimension of $\ker(H_{0,n,T-n+1}(u)) \cap \ker(\Lambda_i(T-2n)^\top)$ is strictly smaller than the dimension of $\ker(H_{0,n,T-n+1}(u))$. Similarly to the proof of Proposition V.1, we conclude $\Pr(g \in \ker(H_{0,n,T-n+1}(u)) \cap \ker(\Lambda_i(T-2n)^\top)) = 0$, i.e., with probability 1, $(\Lambda_i(T-2n)^\top g) \neq 0$. In such case, (21) fails if and only if

$$\alpha_i = -\frac{\Lambda_i(T-2n)^\top\Gamma_i(T)g}{\Lambda_i(T-2n)^\top g}. \tag{22}$$

Recall that α_i's are the coefficients of $x(0)$ in the basis Ψ. Since $x(0)$ is random, the coefficients α_i match exactly the values in (22) with probability 0. In other words, when g is random in $\ker(H_{0,n,T-n+1}(u))$ and $x(0)$ is also random in $\mathbb{R}^n$, almost surely (21) holds, for all $i = 1, \cdots, n$. Hence, $(A, \bar{x}(0))$ is controllable with probability one and the statement follows by applying Lemma V.2. $\square$

The sufficient condition (19) for the correctness of Algorithm 2 in Theorem V.3 states that the vectors generated by the powers of the eigenvalues must not be contained in the row space of the Hankel matrix of the inputs. Informally speaking, this condition means that the inputs do not resonate with any of the system modes. Note that $\mathrm{rank}(H_{0,2n,T-2n+1}(u)^\top) \leq 2nm$ for any $T \in \mathbb{N}$. Therefore, when $T \geq 2(1+m)n$,

$\mathrm{im}(H_{0,2n,T-2n+1}(u)^\top)$ is a proper subspace of $\mathbb{R}^{T-2n+1}$, making condition (19) generically true.

For accurate spectrum estimation, we then need the input u to be sufficiently not "rich", in the sense that the row space spanned by $H_{0,2n,T-2n+1}(u)$ is small enough to avoid containing $\Lambda_i(T-2n)$. This is in contrast to the usual persistency of excitation conditions [35] required on the input for unique system identification. This can be attributed to the fact that spectrum estimation is easier to achieve than system identification, since it can be even done with zero input (but non-zero initial state), and the fact that our algorithm substracts the effects of the inputs when computing the unforced output trajectory.

We also point out an interesting observation here that Theorem V.3 does not impose additional constraints on the input matrix B. This means that we do not require (A,B) to be controllable for the spectrum estimation to work (in principle, B could even be the zero matrix, in which case the inputs would not influence the outputs).

D. *Sufficient conditions for correct spectrum estimation when inputs are generated by another LTI system*

The treatment of this section is motivated by scenarios where the system (7) might be interconnected with other dynamics (e.g., another component that is specifically designed to generate an input to (7) which satisfies certain performance optimality or operational persistently exciting specifications). Here we consider the particular scenario where the input to system (7) is generated by another LTI system of the form

$$z(t+1) = A_z z(t), \tag{23a}$$

$$u(t) = C_z z(t), \tag{23b}$$

where $A_z \in \mathbb{R}^{n_z\times n_z}$, $C_z \in \mathbb{R}^{m\times n_z}$, and initial condition $z(0)\in\mathbb{R}^{n_z}$. The next result analyzes under which cases the sufficient condition (19) on the inputs, guaranteing feasibility of the distributed spectrum estimation, is satisfied.

Proposition V.4 (Satisfaction of the sufficient condition for unique LAE solution as a function of spectra of system matrices)**.** *Consider the LTI system* (7) *with input signal u generated by* (23)*. Then,*

(i) *If $\mathrm{spec}(A)\cap\mathrm{spec}(A_z)=\emptyset$ and $T\geq 2n+n_z$, then the condition* (19) *always holds;*

(ii) *If $\mathrm{spec}(A)\cap\mathrm{spec}(A_z)\neq\emptyset$, let λ_k be a common eigenvalue and let $z(0)$ be a right eigenvector of A_z corresponding to λ_k. Further assume that $C_z z(0)\neq 0$. Then, the condition* (19) *fails for $i=k$ and all $T\geq 2n$.*

Proof. Regarding (i), note that $u(t)=C_zA_z^t z(0)$. Hence,

$$\begin{aligned}
&H_{0,2n,T-2n+1}(u)\\
&=\begin{bmatrix} C_z z(0) & C_zA_z z(0) & \cdots & C_zA_z^{T-2n}z(0)\\ \vdots & \vdots & \ddots & \vdots\\ C_zA_z^{2n-1}z(0) & C_zA_z^{2n}z(0) & \cdots & C_zA_z^{T-1}z(0)\end{bmatrix}\\
&=\begin{bmatrix} C_z\\ C_zA_z\\ \vdots\\ C_zA_z^{2n-1}\end{bmatrix}\begin{bmatrix} z(0) & A_z z(0) & \cdots & A_z^{T-2n}z(0)\end{bmatrix}.
\end{aligned}$$

If $\Lambda_i(T-2n)\in\mathrm{im}(H_{0,2n,T-2n+1}(u)^\top)$ for some $i=1,\cdots,n$, then there exists $w\in\mathbb{R}^{1\times 2nm}$ such that $wH_{0,2n,T-2n+1}(u)=\Lambda_i(T-2n)^\top$. If this is the case, let

$$\tilde{w} := w\begin{bmatrix} C_z\\ C_zA_z\\ \vdots\\ C_zA_z^{2n-1}\end{bmatrix}\in\mathbb{R}^{1\times n_z},$$

and note that

$$\tilde{w}\begin{bmatrix} z(0) & A_z z(0) & \cdots & A_z^{T-2n}z(0)\end{bmatrix} = \begin{bmatrix}1 & \lambda_i & \cdots & \lambda_i^{T-2n}\end{bmatrix}.$$

This means that $\tilde{w}A_z^t z(0)=\lambda_i^t$ for all $t=0,1,\cdots,T-2n$. It follows from the Cayley-Hamilton theorem that

$$p_{A_z}(A_z) = A_z^{n_z} + a_{n_z-1}A_z^{n_z-1}+\cdots+a_0 I = 0,$$

where p_{A_z} is the characteristic polynomial of A_z. Consider the term corresponding to $t=n_z$. We have

$$\begin{aligned}
\lambda_i^{n_z} = \tilde{w}A_z^{n_z}z(0) &= -\tilde{w}\sum_{t=0}^{n_z-1} a_t A_z^t z(0)\\
&= -\sum_{t=0}^{n_z-1} a_t\tilde{w}A_z^t z(0) = -\sum_{t=0}^{n_z-1} a_i\lambda_i^t.
\end{aligned}$$

This means that $p_{A_z}(\lambda_i)=0$, which implies λ_i is an eigenvalue of A_z. Therefore, if $\mathrm{spec}(A_z)\cap\mathrm{spec}(A)=\emptyset$, we deduce that condition (19) always holds with $T\geq 2n+n_z$.

Next we prove (ii). Let $\lambda_k\in\mathrm{spec}(A_z)\cap\mathrm{spec}(A)$ and $z(0)$ be as in the statement. We have $u(t)=C_zA_z^t z(0)=\lambda_k^t C_z z(0)$ for all $t\in\mathbb{N}$. Hence

$$\begin{aligned}
&H_{0,2n,T-2n+1}(u)\\
&=\begin{bmatrix} C_z z(0) & \lambda_k C_z z(0) & \cdots & \lambda_k^{T-2n}C_z z(0)\\ \vdots & \vdots & \ddots & \vdots\\ \lambda_k^{2n-1}C_z z(0) & \lambda_k^{2n}C_z z(0) & \cdots & \lambda_k^{T-1}C_z z(0)\end{bmatrix}\\
&=\left(\begin{bmatrix}1\\ \lambda_k\\ \vdots\\ \lambda_k^{2n-1}\end{bmatrix}\begin{bmatrix}1 & \lambda_k & \cdots & \lambda_k^{T-2n}\end{bmatrix}\right)\otimes(C_z z(0))\\
&=(\Lambda_k(2n-1)\Lambda_k(T-2n)^\top)\otimes(C_z z(0)).
\end{aligned}$$

Let $w=w_1\otimes w_2\in\mathbb{R}^{1\times 2nm}$, with $w_1=\begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix}\in\mathbb{R}^{1\times 2n}$ and $w_2\in\mathbb{R}^{1\times m}$ such that $w_2C_z z(0)=1$ (which is always possible since $C_z z(0)\neq 0$). From the mixed-product property, we conclude that

$$\begin{aligned}
&wH_{0,2n,T-2n+1}(u)\\
&=(w_1\otimes w_2)\left((\Lambda_k(2n-1)\Lambda_k(T-2n)^\top)\otimes(C_z z(0))\right)\\
&=\left(w_1\Lambda_k(2n-1)\Lambda_k(T-2n)^\top\right)\otimes(w_2C_z z(0))\\
&=\Lambda_k(T-2n)^\top,
\end{aligned}$$

which implies $\Lambda_k(T-2n)\in\mathrm{im}(H_{0,2n,T-2n+1}(u)^\top)$. □

The combination of this result with Theorem V.3 yields to the following conclusion.

Corollary V.5 (Almost sure correctness of distributed spectrum estimation for inputs generated by a LTI system)**.** *Consider the LTI system* (7) *with input signal u generated by* (23) *and let Assumptions 1, 2, and 3 hold. If* $\operatorname{spec}(A)\cap\operatorname{spec}(A_z)=\emptyset$ *and* $T\geq 2n+n_z$, *then* $\operatorname{spec}(A)$ *can be accurately estimated by Algorithm 2 with probability 1.*

Note that the condition on disjointness of the spectra of A and A_z in Corollary V.5 is generic. If the input-generating system (23) is unknown, we can also estimate the spectrum of A_z via a similar approach as Algorithm 2, where now the output is u and the input is 0. However, from our discussion in Section V-B, we know that a correct spectrum estimation may require A_z to have distinct eigenvalues and the pair $(A_z, z(0))$ to be controllable. Such assumptions are not required by Corollary V.5. It is therefore interesting to note that, although we might not be able to estimate the spectrum of A_z correctly, as long as it is not intersecting with that of A, we can almost always correctly estimate the spectrum of the original system by collecting data for a sufficiently long time.

VI. Illustration on mass/spring/damper system

In this section, we illustrate the proposed distributed spectrum estimation algorithm on a mass/spring/damper system. Ten unit masses are connected in series by pairs of spring and damper with unknown coefficients k_i, d_i. The left-end is connected to a stationary wall, cf. Figure 2. This mechanical system can be viewed as a simplification of finite element analysis (FEA) on a beam structure. We are interested in inspecting the stiffness of the mechanical system, which can be done via spectrum estimation. Note that it is unnecessary to determine the spring and damper coefficients; the springs and dampers arise from the FEA and hence have no physical meaning.

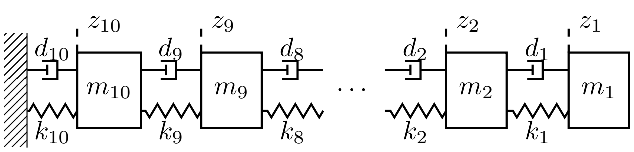


Fig. 2: Illustration of the mass/spring/damper system.

Denote the horizontal displacement of the i-th unit mass relative to its steady position as z_i. Suppose a force u_j is applied to the $(2j-1)$-th mass, $j=1,\cdots,5$. Let $x:=\begin{bmatrix} z_1 & \cdots & z_{10} & \dot z_1 & \cdots & \dot z_{10}\end{bmatrix}^\top$ denote the system state. The dynamics is given by $\dot x = A_c x + B_c u$, where the matrices take the form

$$A_c=\begin{bmatrix} 0_{10\times 10} & I_{10} \\ K_c & D_c \end{bmatrix},\quad B_c=\begin{bmatrix} 0_{10\times 5} \\ I_5\otimes\begin{bmatrix}1\\0\end{bmatrix}\end{bmatrix},$$

$$K_c=\begin{bmatrix} -k_1 & k_1 & 0 & \cdots & 0 \\ k_1 & -k_1-k_2 & k_2 & \ddots & \vdots \\ 0 & k_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & -k_8-k_9 & k_9 \\ 0 & \cdots & 0 & k_9 & -k_9-k_{10}\end{bmatrix},$$

$$D_c=\begin{bmatrix} -d_1 & d_1 & 0 & \cdots & 0 \\ d_1 & -d_1-d_2 & k_2 & \ddots & \vdots \\ 0 & d_2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & -d_8-d_9 & k_9 \\ 0 & \cdots & 0 & d_9 & -d_9-d_{10}\end{bmatrix}.$$

Consider a group of 10 agents that aims to collaboratively compute the system spectrum in a distributed way. Agent i is attached to mass m_i, such that it either measures the applied force $v_i = u_{\frac{i+1}{2}}$ if i is odd, or the displacement $y_i = z_i$ if i is even. The 10 agents form a line network; to be precise, Agents $i-1, i+1$ are the out-neighbors of Agent i, $i=2,3,\cdots,9$, while Agent 2 and itself are the out-neighbors of Agent 1, Agent 9 and itself are the out-neighbors of Agent 10. In addition, we assume that only Agent 1 has the computation capability of finding roots of polynomials. Thus, Agent 1 is responsible of determining the eigenvalues based on the exchange of information and finally report them.

Suppose the measurements are taken at a sampling rate of 10Hz; moreover, assume that the input u is a zero-order-hold signal with the same sampling rate. Then, the system can be converted into a discrete-time LTI system in the form (7), with system matrix and input matrix given by

$$A=e^{0.1A_c}, B=\int_0^{0.1} e^{A_c t}B_c \mathrm{d}t.$$

Assume that $k_i = 100, d_i = 0.3$ for all $i=1,2,\cdots,10$. We collect data for $T=1000$, starting with a random initial state $x(0)$ and inputs $u_i(t)=2+5\cos(0.02t+0.5i)+w_i(t)$, where $w_i(t)$ is white Gaussian noise of signal-to-noise ratio of 5. Figures 3 and 4 show the measurements of the agents. In this case, (19) holds and thus Theorem V.3 implies accurate estimation of the spectrum. This is indeed reflected by Figure 5, which shows a good match between the spectrum of the true system matrix and the estimated eigenvalues computed with the proposed distributed spectrum estimation strategy, cf. Algorithm 2, executed with $K=2000$. Figure 6 shows the evolution of the estimation errors of Agent 1 with respect to K, defined as $e(K) := \sum_{j=1}^n |\lambda_j^e(K)-\lambda_j|$, where $\lambda_j^e(K)$ is the estimation of the j-th eigenvalue for Agent 1 after K iterations and λ_j is the true j-th eigenvalue of A. One can observe that, as more data is exchanged between the agents, they quickly come to an agreement on the spectrum, which in turn converges to the true spectrum of A as the algorithm execution progresses. This observation is consistent with the convergence result in Theorem II.1.

We also consider the case when the input is noise-free; i.e., $u_i(t)=2+5\cos(0.02t+0.5j)$. In this case, $u(t)$ can also be viewed as the output of the system (23), with $n_z=15$, matrices

$$A_z = I_5\otimes\begin{bmatrix}\cos(0.02) & -\sin(0.02) & 0 \\ \sin(0.02) & \cos(0.02) & 0 \\ 0 & 0 & 1\end{bmatrix}, C_z = I_5\otimes\begin{bmatrix}1 & 0 & 1\end{bmatrix}$$

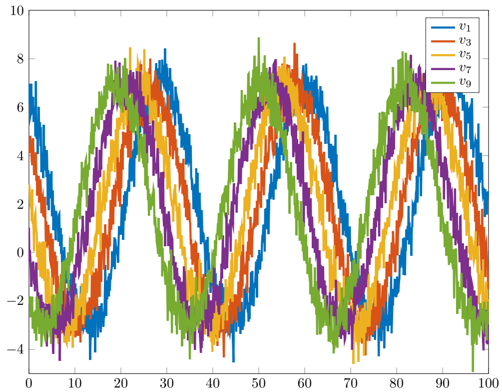


Fig. 3: Plots of $v_i, i = 1, 3, 5, 7, 9$.

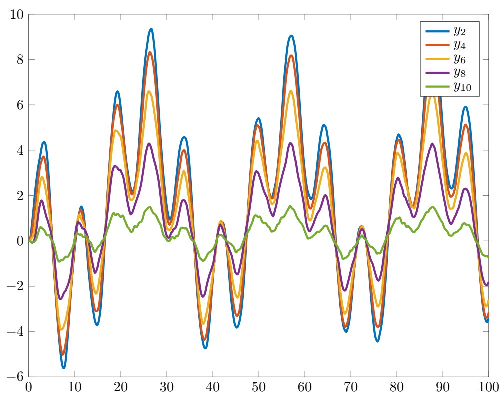


Fig. 4: Plots of $y_i, i = 2, 4, 6, 8, 10$.

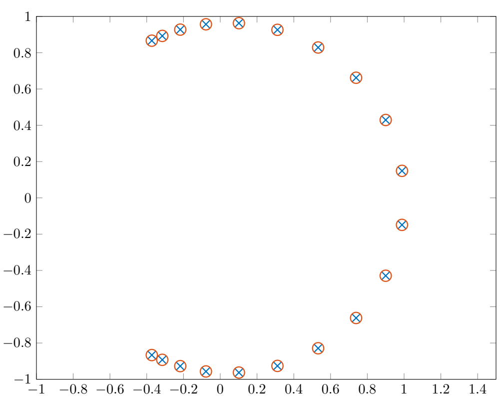


Fig. 5: True eigenvalues (blue cross marks) and estimated eigenvalues (red circles) of A.

and initial state

$$z_i(0) = \begin{cases} 5\cos(0.5k) & \text{if } i = 3k-2, k \in \mathbb{N}, \\ 0 & \text{if } i = 3k-1, k \in \mathbb{N}, \\ 2 & \text{if } i = 3k, k \in \mathbb{N}. \end{cases}$$

Because $\text{spec}(A) \cap \text{spec}(A_z) = \emptyset$ and $T \geq 2n + n_z$, Corollary V.5 implies that Algorithm 2 can accurately estimate $\text{spec}(A)$.

VII. Conclusions

We have introduced a data-driven algorithm for distributed spectrum estimation of linear-time invariant systems. The

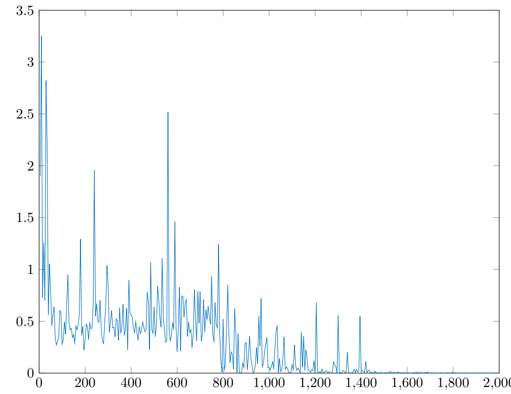


Fig. 6: Estimation error of Agent 1 vs. number of iterations.

agents have no knowledge of the system or control matrices and no control over its inputs. They can however collectively observe the inputs and outputs along a trajectory of the system. Our algorithm relies on this data to compute the characteristic polynomial of the system matrix. Our technical approach to establish convergence relies on the formulation of the estimation problem as suitable systems of linear equations amenable to distributed algorithmic solutions. Future work will explore extensions to scenarios where the dimension of the linear system is unknown and measurement errors are present in collected data, and employ our results in multi-agent coordination problems involving the evaluation of network resilience metrics in a distributed fashion.

Appendix

Proof of Theorem II.1. Multiply $P_{\ker(M_i)^\perp}$ on the left on both sides of (4) and note that $P_{\ker(M_i)^\perp} P_{\ker(M_i)} = 0$, we have $P_{\ker(M_i)^\perp}(z_i(k+1) - z_i(k)) = 0$. In other words $z_i(k+1) - z_i(k) \in \ker(M_i)$ and, recursively, we have $z_i(k) - z_i(0) \in \ker(M_i)$. It follows from the initial condition (3) that $M_i z_i(k) = l_i$, for all $k \in \mathbb{N}, i \in \mathcal{V}$. In other words, the update law (4) always gives feasible solutions to each LAE of (1) at each time k. Let z^* be a solution to (2). It holds that $z_{\text{mn}} = P_{\ker(M)^\perp} z^*$. For each $i \in \mathcal{V}$, denote $e_i(k) := z_i(k) - z^*$. We have $M_i e_i(k) = M_i z_i(k) - M_i z^* = 0$ so $e_i(k) \in \ker(M_i)$. Therefore, $e_i(k) = P_{\ker(M_i)} e_i(k)$. It then follows from (4) that

$$\begin{aligned} e_i(k+1) &= P_{\ker(M_i)} e_i(k) + \frac{1}{d_i} P_{\ker(M_i)} \sum_{j \in N_i^{\text{out}}} (e_j(k) - e_i(k)) \\ &= \frac{1}{d_i} P_{\ker(M_i)} \sum_{j \in N_i^{\text{out}}} e_j(k). \end{aligned} \tag{24}$$

Denote by $d_{\text{null}} := \text{nullity}(M)$, and define $\bar{e}_i(k) := Q e_i(k)$, or in the stacked from $\bar{e}(k) = (I_p \otimes Q) e(k)$, where $Q \in \mathbb{R}^{(n-d_{\text{null}}) \times n}$ is a matrix such that $\ker(Q) = \ker(M) = \bigcap_{i \in \mathcal{V}} \ker(M_i)$, and $QQ^\top = I_{n-d_{\text{null}}}$. For each $i \in \mathcal{V}$, define $\bar{P}_i := Q P_{\ker(M_i)} Q^\top$. Then it follows from [24, Lemma 1] such that $\bar{P}_i$ is also an orthogonal projection matrix,

$$Q P_{\ker(M_i)} = \bar{P}_i Q, \tag{25}$$

and

$$\bigcap_{i \in \mathcal{V}} \ker \bar{P}_i = \{0\}. \tag{26}$$

Also, since both $P_{\ker(M_i)}, \bar{P}_i$ are orthogonal projection matrices, they are symmetric and we conclude

$$P_{\ker(M_i)} Q^\top = Q^\top \bar{P}_i \tag{27}$$

by taking the transpose of (25) on both sides. Multiply Q on the left on both sides of (24) and use the property (25), we conclude that

$$\bar{e}_i(k+1) = \frac{1}{d_i} \bar{P}_i \sum_{j \in N_i^{\text{out}}} \bar{e}_j(k), \tag{28}$$

which can be compactly written as

$$\bar{e}(k+1) = \bar{P}((\mathsf{D}^{-1}\mathsf{A}) \otimes I_n)\bar{e}(k), \tag{29}$$

where recall the definitions of D, A in Theorem II.1, and the matrix $\bar{P} := \operatorname{diag}(\bar{P}_1, \cdots, \bar{P}_p)$. Because of (26) and the fact that $\mathsf{D}^{-1}\mathsf{A}$ is row stochastic, we appeal to [24, Theorem 3] and conclude that there exists $\lambda_1 \in (0,1)$ such that

$$\left\| \left(\bar{P}((\mathsf{D}^{-1}\mathsf{A}) \otimes I) \right)^k \right\| \leq \lambda_1^{k-(p-1)^2}.$$

As a result, the iterative equation (29) implies that

$$|\bar{e}(k)| \leq c_1 \lambda_1^k, \tag{30}$$

where $c_1 := \lambda_1^{-(p-1)^2} |\bar{e}(0)|$.

Lastly, define $r_i(k) := e_i(k) - Q^\top \bar{e}_i(k)$, or in the stacked form $r(k) := e(k) - (I_p \otimes Q^\top)\bar{e}(k)$. We have $Qr_i(k) = Qe_i(k) - \bar{e}_i(k) = 0$. In other words, $r_i(k) \in \ker(Q) = \bigcap_{i \in \mathcal{V}} \ker(M_i)$ so $P_{\ker(M_i)} r_j(k) = r_j(k)$ for all $i, j \in \mathcal{V}$. As a result, it follows from (24), (27) and (28) that

$$\begin{aligned}
r_i(k+1) &= e_i(k+1) - Q^\top \bar{e}_i(k+1) \\
&= \frac{1}{d_i} P_{\ker(M_i)} \sum_{j \in N_i^{\text{out}}} e_j(k) - \frac{1}{d_i} Q^\top \bar{P}_i \sum_{j \in N_i^{\text{out}}} \bar{e}_j(k) \\
&= \frac{1}{d_i} P_{\ker(M_i)} \sum_{j \in N_i^{\text{out}}} (e_j(k) - Q^\top \bar{e}_j(k)) \\
&= \frac{1}{d_i} P_{\ker(M_i)} \sum_{j \in N_i^{\text{out}}} r_j(k) \\
&= \frac{1}{d_i} \sum_{j \in N_i^{\text{out}}} r_j(k),
\end{aligned}$$

which can be compactly written as

$$r(k+1) = ((\mathsf{D}^{-1}\mathsf{A}) \otimes I_n) r(k).$$

Since $\mathcal{G}$ is strongly connected and aperiodic, $\mathsf{D}^{-1}\mathsf{A}$ is primitive. Also recall that $w \in \mathbb{R}^n$ is the left Perron-Frobenius eigenvector of $\mathsf{D}^{-1}\mathsf{A}$. Thus, it follows from standard results on discrete-time consensus [36, Theorem 11.2] that there exists $\lambda_2 \in (0,1), c_2 \geq 1$ such that

$$|r(k) - 1 \otimes (R(0)w)| \leq c_2 \lambda_2^k,$$

where

$$\begin{aligned}
R(0) :&= \begin{bmatrix} r_1(0) & \cdots & r_p(0) \end{bmatrix} \\
&= \begin{bmatrix} e_1(0) & \cdots & e_p(0) \end{bmatrix} - Q^\top \begin{bmatrix} \bar{e}_1(0) & \cdots & \bar{e}_p(0) \end{bmatrix} \\
&= (I_n - Q^\top Q) \begin{bmatrix} e_1(0) & \cdots & e_p(0) \end{bmatrix} \\
&= (I_n - Q^\top Q)(Z(0) - 1^\top \otimes z^*).
\end{aligned}$$

Moreover, from the definition of Q we have $P_{\ker(M)} = P_{\ker(Q)} = I_n - P_{\ker(Q)^\perp} = I_n - P_{\operatorname{im}(Q^\top)} = I_n - Q^\top Q$ and $(1^\top \otimes z^*)w = z^* w^\top 1 = z^*$, we further conclude that

$$\left| r(k) - 1 \otimes \left(P_{\ker(M)}(Z(0)w - z^*) \right) \right| \leq c_2 \lambda_2^k. \tag{31}$$

Now recall the definition of z_{c} in (6). It follows from (30) and (31) that

$$\begin{aligned}
&|z(k) - z_{\text{c}}| \\
&\quad = |z(k) - 1 \otimes (P_{\ker(M)} Z(0)w + z_{\text{mn}})| \\
&\quad = |(z(k) - 1 \otimes z^*) + 1 \otimes (z^* - P_{\ker(M)} Z(0)w - z_{\text{mn}})| \\
&\quad = |e(k) + 1 \otimes (z^* - P_{\ker(M)} Z(0)w - P_{\ker(M)^\perp} z^*)| \\
&\quad = \left| (I_p \otimes Q^\top)\bar{e}(k) + r(k) - 1 \otimes \left(P_{\ker(M)}(Z(0)w - z^* \right) \right| \\
&\quad \leq |\bar{e}(k)| + \left| r(k) - 1 \otimes (P_{\ker(M)}(Z(0)w - z^*) \right| \\
&\quad \leq c_1 \lambda_1^k + c_2 \lambda_2^k,
\end{aligned}$$

where we have used the relation between $e(k), \bar{e}(k)$ and $r(k)$ and the fact that $z^* = P_{\ker(M)} z^* + P_{\ker(M)^\perp} z^*$ for the fourth equality, and the orthonormal property of Q such that $\|I_p \otimes Q^\top\| \leq 1$ for the second last line. Hence (5) is shown with $c = c_1 + c_2, \lambda = \max\{\lambda_1, \lambda_2\}$. ☐

The following result is invoked in the proof of Theorem V.3.

Lemma A.1 (Controllability and left eigenvectors)**.** *Under Assumption 3, the pair* (A, b) *is controllable if and only if* $\varphi_\lambda^\top b \neq 0$ *for any* $\lambda \in \operatorname{spec}(A)$*, where* φ_λ *is a non-zero left eigenvector of* A *with respect to the eigenvalue* λ*.*

Proof. We rely on the Popov-Belevitch-Hautus (PBH) test [34, Corollary 4.7] for controllability, that states that the pair (A, b) is controllable if and only if

$$\operatorname{rank}(\begin{bmatrix} A - \lambda I_n & b \end{bmatrix}) = n$$

for all $\lambda \in \mathbb{C}$. To prove the implication from left to right, suppose $\varphi_\lambda^\top b = 0$ for some $\lambda \in \operatorname{spec}(A)$. Then, we have $\varphi_\lambda^\top \begin{bmatrix} A - \lambda I & b \end{bmatrix} = 0$, so $\begin{bmatrix} A - \lambda I & b \end{bmatrix}$ is not full rank. Thus by the PBH test, (A, b) is not controllable.

To show the implication from right to left, assume that $\varphi_\lambda^\top b \neq 0$ for all $\lambda \in \operatorname{spec}(A)$. For any $\lambda^* \in \mathbb{C}$, pick $w \in \ker \begin{bmatrix} A - \lambda^* I & b \end{bmatrix}^\top$, that is,

$$w^\top (A - \lambda^* I) = 0, \tag{32a}$$
$$w^\top b = 0. \tag{32b}$$

If $\lambda^* \notin \operatorname{spec}(A)$, (32a) only holds with $w = 0$ since $A - \lambda^* I$ is full rank. Otherwise, if $\lambda^* \in \operatorname{spec}(A)$, (32a) implies that w must be a left eigenvector of A with respect to the eigenvalue λ^*; in other words, $w = c\varphi_{\lambda^*}$ for some $c \in \mathbb{R}$. Since $\varphi_\lambda^\top b \neq 0$ for all $\lambda \in \operatorname{spec}(A)$, (32b) implies that $c = 0$; that is, $w = 0$. Hence $\ker \begin{bmatrix} A - \lambda^* I & b \end{bmatrix}^\top = \{0\}$ for all $\lambda^* \in \mathbb{C}$, which by the PBH test implies that (A, b) is controllable. ☐

References

[1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[2] J. H. Chow, *Power System Coherency and Model Reduction*. New York, NY: Springer New York, 2013.

[3] W. Favoreel, B. De Moor, and P. Van Overschee, "Subspace state space system identification for industrial processes," *Journal of Process Control*, vol. 10, no. 2, pp. 149–155, 2000.

[4] B. Peeters and G. De Roeck, "Stochastic System Identification for Operational Modal Analysis: A Review," *Journal of Dynamic Systems, Measurement, and Control*, vol. 123, no. 4, pp. 659–667, 02 2001.

[5] E. Reynders, "System identification methods for (operational) modal analysis: Review and comparison," *Archives of Computational Methods in Engineering*, vol. 19, pp. 51–124, 2012.

[6] L. Ljung, *System Identification: Theory for the User*, ser. Prentice Hall Information and System Sciences Series. Prentice Hall, 1999.

[7] B. L. Ho and R. E. Kalman, "Editorial: Effective construction of linear state-variable models from input/output functions," *at - Automatisierungstechnik*, vol. 14, no. 1-12, pp. 545–548, 1966.

[8] J.-N. Juang and R. S. Pappa, "An eigensystem realization algorithm for modal parameter identification and model reduction," *Journal of Guidance, Control, and Dynamics*, vol. 8, no. 5, pp. 620–627, 1985.

[9] A.-J. Van Der Veen, E. Deprettere, and A. Swindlehurst, "Subspace-based signal analysis using singular value decomposition," *Proceedings of the IEEE*, vol. 81, no. 9, pp. 1277–1308, 1993.

[10] M. Viberg, "Subspace-based methods for the identification of linear time-invariant systems," *Automatica*, vol. 31, no. 12, pp. 1835–1851, 1995.

[11] R. Brincker, L. Zhang, and P. Andersen, "Modal identification of output-only systems using frequency domain decomposition," *Smart Materials and Structures*, vol. 10, no. 3, pp. 441–445, jun 2001.

[12] T. Sahai, A. Speranzon, and A. Banaszuk, "Hearing the clusters of a graph: A distributed algorithm," *Automatica*, vol. 48, no. 1, pp. 15–24, 2012.

[13] M. Franceschelli, A. Gasparri, A. Giua, and C. Seatzu, "Decentralized estimation of Laplacian eigenvalues in multi-agent systems," *Automatica*, vol. 49, no. 4, pp. 1031–1036, 2013.

[14] D. Kempe and F. McSherry, "A decentralized algorithm for spectral analysis," *Journal of Computer and System Sciences*, vol. 74, no. 1, pp. 70–83, 2008.

[15] Y. Saad, *Numerical Methods for Large Eigenvalue Problems*. Philadelphia, PA: SIAM, 2011.

[16] P. Yang, R. A. Freeman, G. J. Gordon, K. M. Lynch, S. S. Srinivasa, and R. Sukthankar, "Decentralized estimation and control of graph connectivity for mobile sensor networks," *Automatica*, vol. 46, no. 2, pp. 390–396, 2010.

[17] C. Li and Z. Qu, "Distributed estimation of algebraic connectivity of directed networks," *Systems & Control Letters*, vol. 62, no. 6, pp. 517–524, 2013.

[18] F. Deroo, M. Meinel, M. Ulbrich, and S. Hirche, "Distributed stability tests for large-scale systems with limited model information," *IEEE Transactions on Control of Network Systems*, vol. 2, no. 3, pp. 298–309, 2015.

[19] M. Yang and C. Y. Tang, "Distributed estimation of graph spectrum," in *2015 American Control Conference (ACC)*, 2015, pp. 2703–2708.

[20] A. Gusrialdi and Z. Qu, "Distributed estimation of all the eigenvalues and eigenvectors of matrices associated with strongly connected digraphs," *IEEE Control Systems Letters*, vol. 1, no. 2, pp. 328–333, 2017.

[21] N. Emirov, C. Cheng, Q. Sun, and Z. Qu, "Distributed algorithms to determine eigenvectors of matrices on spatially distributed networks," *Signal Processing*, vol. 196, p. 108530, 2022.

[22] D. P. Bertsekas, *Parallel and Distributed Computations: Numerical Methods*. Belmont, MA: Athena Scientific, 1997.

[23] A. Nedic, A. Ozdaglar, and P. A. Parrilo, "Constrained consensus and optimization in multi-agent networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 922–938, 2010.

[24] S. Mou, J. Liu, and A. S. Morse, "A distributed algorithm for solving a linear algebraic equation," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 2863–2878, 2015.

[25] G. Shi, B. D. O. Anderson, and U. Helmke, "Network flows that solve linear equations," *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 2659–2674, 2017.

[26] X. Wang, J. Zhou, S. Mou, and M. J. Corless, "A distributed algorithm for least squares solutions," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4217–4222, 2019.

[27] A. Nedic and A. Ozdaglar, "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, vol. 54, no. 1, pp. 48–61, 2009.

[28] S. S. Kia, J. Cortés, and S. Martinez, "Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication," *Automatica*, vol. 55, pp. 254–264, 2015.

[29] W. Shi, Q. Ling, G. Wu, and W. Yin, "EXTRA: an exact first-order algorithm for decentralized consensus optimization," *SIAM Journal on Optimization*, vol. 25, no. 2, pp. 944–966, 2015.

[30] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[31] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009, available at http://www.coordinationbook.info.

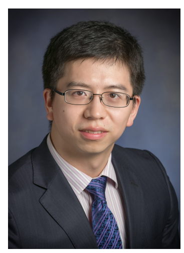
[32] H. J. van Waarde, C. De Persis, M. K. Camlibel, and P. Tesi, "Willems' fundamental lemma for state-space systems and its extension to multiple datasets," *IEEE Control Systems Letters*, vol. 4, no. 3, pp. 602–607, 2020.

[33] O. Aberth, "Iteration methods for finding all zeros of a polynomial simultaneously," *Mathematics of Computation*, vol. 27, pp. 339–344, 1973.

[34] P. J. Antsaklis and A. N. Michel, *Linear Systems*. Boston, MA: Birkhäuser Boston, 2006.

[35] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. M. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.

[36] F. Bullo, *Lectures on Network Systems*, 1st ed. Kindle Direct Publishing, 2019, with contributions by J. Cortes, F. Dorfler, and S. Martinez.



Shenyu Liu (S'16-M'20) received his B. Eng. degree in Mechanical Engineering and B.S. degree in Mathematics from the University of Singapore, Singapore, in 2014. He received his M.S. degree in Mechanical Engineering from the University of Illinois at Urbana-Champaign, in 2015, where he also received his Ph.D. degree in Electrical Engineering in 2020, under the supervision of Prof. Daniel Liberzon and Prof. Mohamed-Ali Belabbas. He then spent two years in the Department of Mechanical and Aerospace Engineering at University of California, San Diego, as a postdoctoral researcher under the supervision of Prof. Jorge Cortés and Prof. Sonia Martínez. He is now an Assistant Professor in the School of Automation at Beijing Institute of Technology, Beijing, China. His current research interest includes stability theory of switched/hybrid systems, Lyapunov methods for nonlinear systems, matrix perturbation theory, uncertain systems and data-driven controller design.

Jorge Cortés (M'02, SM'06, F'14) received the Licenciatura degree in mathematics from Universidad de Zaragoza, Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from Universidad Carlos III de Madrid, Madrid, Spain, in 2001. He held postdoctoral positions with the University of Twente, Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, Urbana, IL, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics, University of California, Santa Cruz, CA, USA, from 2004 to 2007. He is currently a Professor in the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA. He is the author of Geometric, Control and Numerical Aspects of Nonholonomic Systems (Springer-Verlag, 2002) and co-author (together with F. Bullo and S. Martínez) of Distributed Control of Robotic Networks (Princeton University Press, 2009). He is a Fellow of IEEE, SIAM, and IFAC. His current research interests include distributed control and optimization, network science, nonsmooth analysis, reasoning and decision making under uncertainty, network neuroscience, and multi-agent coordination in robotic, power, and transportation networks.

Sonia Martínez (M'02-SM'07-F'18) is a Professor of Mechanical and Aerospace Engineering at the University of California, San Diego, CA, USA. She received the Ph.D. degree in Engineering Mathematics from the Universidad Carlos III de Madrid, Spain, in May 2002. She was a Visiting Assistant Professor of Applied Mathematics at the Technical University of Catalonia, Spain (2002-2003) and a Postdoctoral Fulbright Fellowship at the Coordinated Science Laboratory of the University of Illinois, Urbana-Champaign (2003-2004) and the Center for Control, Dynamical systems and Computation of the University of California, Santa Barbara (2004-2005). Her research interests include the control of network systems, multi-agent systems, nonlinear control theory, and robotics. She received the Best Student Paper award at the 2002 IEEE Conference on Decision and Control for her work on the control of underactuated mechanical systems and was the recipient of a NSF CAREER Award in 2007. For the paper "Motion coordination with Distributed Information," co-authored with Jorge Cortés and Francesco Bullo, she received the 2008 Control Systems Magazine Outstanding Paper Award. She is the Editor in Chief of the recently launched Open Journal of Control Systems.