Optimal Worst-Case Dynamic Average Consensus

Bryan Van Scoy, Randy A. Freeman, and Kevin M. Lynch

***Abstract*— We formulate a method for designing dynamic average consensus estimators with optimal worst-case asymptotic convergence rate over a large set of undirected graphs. The estimators achieve average consensus for constant inputs and are robust to both initialization errors and changes in network topology. The structure of a general class of polynomial linear protocols is characterized and used to find global optimal parameters using polynomial matrix inequalities (PMIs). For the case of the PI estimator, these conditions are converted into convex linear matrix inequalities (LMIs) and solved efficiently.**

I. INTRODUCTION

We consider the dynamic average consensus problem where each agent in a network uses communication with its network neighbors along with a local estimator to calculate the average input of all the agents [1], [2], [3], [4], [5]. Average consensus is important for its central role in many applications in decentralized control of multi-agent systems.

Two challenges in applying decentralized average consensus to distributed control are (1) the slow convergence rate of most average consensus algorithms and (2) the requirement that the local estimators produce correct estimates even in the face of changing communication networks. In this paper we develop an estimator design procedure that addresses these two issues. We characterize the possible communication networks that can occur in terms of their minimum algebraic connectivity. Our design procedure is guaranteed to find the globally optimal estimator parameters which give the minimum worst-case asymptotic convergence time. In addition, the estimator achieves zero steady-state error for constant inputs and is robust to both initialization errors and changes in network topology (such as the addition/removal of agents). The design process contains two steps: 1) conditions are found on the state-transition matrix such that the system achieves average consensus for constant inputs and is robust to both the initial state and changes in network topology, 2) the worst-case convergence rate is then optimized over the set of connected undirected graphs whose Laplacian matrices have non-zero eigenvalues in a given range $[\lambda_{\text{min}}, \lambda_{\text{max}}]$ subject to the conditions found in step one.

The design process is developed for a general n-dimensional estimator of degree ℓ where n is the number of state variables on each agent and ℓ is the number of communication hops. Although the procedure is guaranteed to find the global optimum, the design is computationally challenging for $n > 2$. The case of $n = 2$ is shown to be convex and an efficient algorithm is given.

A. Related Work

Much work has been done on achieving fast static consensus. Xiao and Boyd [6] optimize the asymptotic convergence rate for a known graph. For undirected graphs, the global optimal asymptotic convergence rate is found using semidefinite programming (SDP). Improving on these results, Oreshkin et al. [7] use a local predictor at each node to enhance the convergence rate, and Erseghe et al. [8] use the alternating direction multipliers method (ADMM) and distributed optimization to select optimal parameters.

Static consensus, however, is inherently non-robust to changes in network topology [9]. Much less work has been done on optimizing dynamic average consensus estimators which are robust to changes in network topology. Elwin et al. [10] optimize the worst-case performance of dynamic average consensus estimators by applying numerical global optimization solvers, but no guarantees are given of finding the global optimum.

The rest of the paper is organized as follows. Section II sets up the average consensus problem, defines the polynomial linear protocol, and states the two problems to be solved. A general polynomial linear protocol of dimension n and degree l is studied in Sections III and IV; the first problem of developing conditions on the state-transition matrix is solved in Section III, and then Section IV formulates the problem of optimizing the worst-case asymptotic convergence rate to solve the second problem. The results are then used to design the optimal worst-case PI estimator in Section V, and conclusions are given in Section VI.

Notation: The vectors 1_n and 0_n represent the $n \times 1$ vectors with all entries 1 and 0, respectively. The symbol $\otimes$ represents the Kronecker product. The spectral radius is denoted $\rho(\cdot)$. The transpose of A is denoted by A^T. The Moore-Penrose pseudoinverse of a matrix A is denoted $A^\dagger$. A diagonal matrix with entries α_i on the diagonal is denoted $\text{diag}(\alpha)$. $A > 0$ and $A \geq 0$ mean that the matrix A is positive definite and positive semi-definite, respectively. A matrix M is said to be convergent if and only if its power sequence $\{M^k\}_{k=1}^\infty$ converges to a finite constant matrix as $k \to \infty$.

II. PROBLEM SETUP

A. Dynamic Average Consensus

Consider a group of N agents whose communication topology is modeled as a weighted undirected graph G.

This work is supported in part by the Office of Naval Research. The authors are with the Department of Electrical Engineering and Computer Science (Van Scoy and Freeman) and the Department of Mechanical Engineering (Lynch), Northwestern University, Evanston, IL 60208, USA. Emails: bryanvanscoy2012@u.northwestern.edu, freeman@eecs.northwestern.edu, kmlynch@northwestern.edu.

Define the adjacency matrix of G to be $A = [a_{ij}] \in \mathbb{R}^{N \times N}$ where $a_{ij} = a_{ji} > 0$ if agents i and j communicate information and zero otherwise (with $a_{ii} = 0$). The neighbors of agent i, denoted $\mathcal{N}_i$, is the set of agents with which agent i communicates. The degree of agent i, denoted $\deg(i)$, is the number of agents in $\mathcal{N}_i$. The Laplacian matrix is $L = \mathrm{diag}(A 1_N) - A$ which is symmetric, positive semidefinite, and satisfies $L 1_N = 0_N$. The algebraic connectivity of the graph is the second smallest eigenvalue of L, denoted $\lambda_2(L)$. The graph is connected if and only if $\lambda_2(L) > 0$.

The weights a_{ij} can be chosen to optimize the performance of the system if the communication graph is known [11]. When the graph is unknown, however, it is useful to use the weights to bound the eigenvalues of the Laplacian. For example, inverse out-degree (IOD) weighting assigns $a_{ij} = 1/[\deg(i) + \deg(j)]$. For undirected graphs, this decentralized weighting scheme restricts the eigenvalues of L to the interval $[0, 1]$. It also has the added advantage of producing symmetric (and therefore balanced) expected Laplacians under suitably symmetric packet-loss probability distributions (see [9] for details).

The average consensus problem is to design an estimator where the output of each agent converges to $(1/N) \sum_{i=1}^{N} u_{i,k}$ where $u_{i,k}$ is the input to agent i at time k.

B. Polynomial Linear Protocol

A general class of linear protocols which can be used for average consensus is the polynomial linear protocol [9].

Definition 1: A polynomial linear protocol of dimension n and degree l is the collection $\Sigma(L) = [A(L), B(L), C(L), D(L)]$ where

$$A(L) \equiv \sum_{i=0}^{l} L^i \otimes A_i \qquad B(L) \equiv \sum_{i=0}^{l} L^i \otimes B_i$$
$$C(L) \equiv \sum_{i=0}^{l} L^i \otimes C_i \qquad D(L) \equiv \sum_{i=0}^{l} L^i \otimes D_i$$

are polynomials in L which describe the linear system

$$x_{k+1} = A(L) x_k + B(L) u_k \tag{1}$$
$$y_k = C(L) x_k + D(L) u_k. \tag{2}$$

For a SISO system, the sizes of matrices and vectors are $L \in \mathbb{R}^{N \times N}$, $A_i \in \mathbb{R}^{n \times n}$, $B_i \in \mathbb{R}^n$, $C_i^T \in \mathbb{R}^n$, $D_i \in \mathbb{R}$, $x_k \in \mathbb{R}^{Nn}$, $y_k \in \mathbb{R}^N$.

Example 1: The standard PI estimator (see [12] for the continuous-time version) is a polynomial linear protocol of dimension two and degree one with parameters γ, k_p, and k_I where

$$\left[\begin{array}{c|c} A_0 & B_0 \\ \hline C_0 & D_0 \end{array}\right] = \left[\begin{array}{cc|c} 1-\gamma & 0 & \gamma \\ 0 & 1 & 0 \\ \hline 1 & 0 & 0 \end{array}\right]$$
$$\left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] = \left[\begin{array}{cc|c} -k_p & k_I & 0 \\ -1 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right].$$

For undirected graphs, the Laplacian is symmetric and can be diagonalized as $D = Q^T L Q$ where $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$, λ_i are the eigenvalues of L, and $Q \in \mathbb{R}^{N \times N}$ is orthogonal. Using the change of variable $\bar{x}_k = (Q \otimes I) x_k$, the system (1) and (2) may be separated as

$$\bar{x}_{i,k+1} = A(\lambda_i) \bar{x}_{i,k} + B(\lambda_i) u_{i,k} \tag{3}$$
$$y_{i,k} = C(\lambda_i) \bar{x}_{i,k} + D(\lambda_i) u_{i,k} \tag{4}$$

for $i = 1, \ldots, N$ (see [9, Theorem 5]).

C. Problem Statement

We wish to design an average consensus estimator with optimal worst-case asymptotic convergence rate over a large set of graphs while having other desired properties. Specifically, let $\mathcal{G}$ be the set of connected undirected graphs whose Laplacian matrices have eigenvalues in $\Lambda := \{0, [\lambda_{\text{min}}, \lambda_{\text{max}}]\}$. Using the separation principle in equations (3) and (4), we study the system

$$x_{k+1} = A(\lambda) x_k + B(\lambda) u_k \tag{5}$$
$$y_k = C(\lambda) x_k + D(\lambda) u_k \tag{6}$$

where $\lambda \in \Lambda$. The parameters λ_{min} and λ_{max} are bounds for the non-zero eigenvalues of the Laplacian which depend on the weighting scheme. For example, if IOD weighting is used then we can take $\lambda_{\text{max}} = 1$ if the graph is not known [9, Theorem 1]. The choice of λ_{min} depends on the minimum connectivity of the graphs which we expect to occur. A larger value of λ_{min} tightens the bounds on the range of networks which allows us to achieve better worst-case performance.

To obtain robustness to changes in network topology, we desire the output process of the estimator to be asymptotically mean ergodic in the presence of i.i.d. switching graphs, meaning that the time average of each agent's output converges to the statistical average as time approaches infinity. The statistical average is the output of the estimator when the expected Laplacian is used. If the estimator achieves average consensus for the expected Laplacian, then a local low-pass filter can be applied to the output of each agent to obtain the global average. Conditions for the output process of a polynomial linear protocol to be asymptotically mean ergodic are given in [13].

We now state the two problems to be solved.

Problem 1: Given the size of the polynomial linear protocol (n and l), determine conditions on $A(L)$ such that there exist $B(L)$, $C(L)$, and $D(L)$ where $\Sigma(L)$

- achieves exact average consensus for constant inputs,
- regardless of the choice of initial states (i.e., convergence is robust to initial conditions), and
- is asymptotically mean ergodic.

Problem 2: Given λ_{min}, λ_{max}, the size of the estimator (n and l), and the solution to Problem 1, determine $\Sigma(L)$ which optimizes the worst-case asymptotic convergence rate over all graphs in $\mathcal{G}$ and has the properties listed in Problem 1. That is, solve

$$\alpha = \min_{A_i} \max_{\lambda \in \Lambda} \rho(A(\lambda)) \tag{7}$$

subject to conditions from Problem 1

where $\rho(A(\lambda))$ does not include any unobservable eigenvalues of the pair $(A(\lambda), C(\lambda))$.

III. Structure of a Polynomial Linear Protocol

A characterization of a polynomial linear protocol is given in [9, Theorem 5] which gives necessary and sufficient conditions for an estimator to achieve consensus for constant inputs and to be robust to initial conditions. This characterization, however, gives no insight into how to design the system. In this section, we develop conditions on $A(L)$ which may be used for design and provide the solution to Problem 1. First, we give necessary and sufficient conditions on $A(L)$ for estimators which achieve average consensus for constant inputs and are robust to initial conditions. Then we show how to use the remaining degrees of freedom to obtain desired properties of the estimator.

Condition 1: Given $A(L)$ of degree l and dimension n, there exist $v, w, x_k, y_k \in \mathbb{R}^n$ for $k = 0, \dots, nl$ such that

$$0 = \begin{bmatrix} I - A_0 & & & & \\ \vdots & \ddots & & & \\ -A_l & & I - A_0 & & \\ & \ddots & & \ddots & \\ & & -A_l & \dots & I - A_0 \end{bmatrix} \begin{bmatrix} x_{nl} \\ \vdots \\ x_l \\ \vdots \\ x_0 - v \end{bmatrix} \tag{8}$$

$$0 = \begin{bmatrix} I - A_0^T & & & & \\ \vdots & \ddots & & & \\ -A_l^T & & I - A_0^T & & \\ & \ddots & & \ddots & \\ & & -A_l^T & \dots & I - A_0^T \end{bmatrix} \begin{bmatrix} y_{nl} \\ \vdots \\ y_l \\ \vdots \\ y_0 - w \end{bmatrix} \tag{9}$$

$$D_j = \begin{cases} 0, & j < 0 \\ 1 - w^T(I - A_0)v, & j = 0 \end{cases} \tag{10}$$

where

$$D_j = -\begin{bmatrix} y_0^T & \dots & y_l^T \end{bmatrix} \begin{bmatrix} -A_l & \dots & I - A_0 \\ & \ddots & \vdots \\ & & -A_l \end{bmatrix} \begin{bmatrix} x_{l-j} \\ \vdots \\ x_{-j} \end{bmatrix} \tag{11}$$

and $x_k = 0_n$ for $k > nl$.

We now state our main result which shows that $\Sigma(L)$ achieves average consensus for constant inputs and is robust to initial conditions if and only if Condition 1 is satisfied.

Theorem 1: Consider a polynomial linear protocol $\Sigma(L)$ with $A(\lambda)$ convergent for all $\lambda \in \text{eig}(L)$.

1) If $\Sigma(L)$ achieves average consensus for constant inputs and is robust to initial conditions, then $A(L)$ satisfies Condition 1.
2) If $A(L)$ satisfies Condition 1, then there exist $B(L)$, $C(L)$, and $D(L)$ such that $\Sigma(L)$ achieves average consensus for constant inputs and is robust to initial conditions.

Proof: 1) Let $\Sigma(L)$ be a robust protocol with $A(\lambda)$ convergent for all $\lambda \in \text{eig}(L)$. From [9, Theorem 5], we have that $B(\lambda) \in \text{Col}(I - A(\lambda))$, $C^T(\lambda) \in \text{Col}\big(I - A^T(\lambda)\big)$, $H(0) = 1$, and $H(\lambda) = 0$ for $\lambda > 0$ where

$$H(\lambda) = C(\lambda)[I - A(\lambda)]^\dagger B(\lambda) + D(\lambda).$$

Then there exist vectors $x(\lambda)$ and $y(\lambda)$ such that $B(\lambda) = [I - A(\lambda)]x(\lambda)$ and $C(\lambda) = y^T(\lambda)[I - A(\lambda)]$ for all λ, so

$$0 = \begin{bmatrix} I - A(\lambda) & B(\lambda) \end{bmatrix} \begin{bmatrix} x(\lambda) \\ -1 \end{bmatrix}$$

where $x(\lambda) = \sum_{k=0}^{\infty} \lambda^{-k} x_k$. The null space vector $\begin{bmatrix} x^T(\lambda) & -1 \end{bmatrix}^T$ can be written as a polynomial vector with degree bounded from above by [14, Lemma 5]

$$\sum_{i=1}^{n} \deg_i \left(\begin{bmatrix} I - A(\lambda) & B(\lambda) \end{bmatrix} \right) - \min_{i=1,\dots,n} \deg_i \left(\begin{bmatrix} I - A(\lambda) & B(\lambda) \end{bmatrix} \right) \le nl$$

where the notation $\deg_i(A(\lambda))$ denotes the degree of the ith column of the polynomial matrix $A(\lambda)$, so

$$0 = \begin{bmatrix} I - A(\lambda) & B(\lambda) \end{bmatrix} \begin{bmatrix} \lambda^{nl} x(\lambda) \\ -\lambda^{nl} \end{bmatrix}$$

where $\lambda^{nl} x(\lambda)$ is a polynomial vector of degree at most nl. Therefore, $x_k = 0$ for all $k > nl$. Similarly, $y_k = 0$ for $k > nl$.

For $\lambda > 0$, the matrix form of the equation $B(\lambda) = [I - A(\lambda)]x(\lambda)$ is given by

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ B_0 \\ \vdots \\ B_l \end{bmatrix} = \begin{bmatrix} I - A_0 & & & \\ \vdots & \ddots & & \\ -A_l & & I - A_0 & \\ & \ddots & & \ddots \\ & & -A_l & & I - A_0 \\ & & & \ddots & \vdots \\ & & & & -A_l \end{bmatrix} \begin{bmatrix} x_{nl} \\ \vdots \\ x_l \\ \vdots \\ x_0 \end{bmatrix}$$

which gives equation (8) and the structure for $B(\lambda)$,

$$\begin{bmatrix} B_0 \\ \vdots \\ B_l \end{bmatrix} = \begin{bmatrix} -A_l & \dots & I - A_0 \\ & \ddots & \vdots \\ & & -A_l \end{bmatrix} \begin{bmatrix} x_l \\ \vdots \\ x_0 \end{bmatrix} \tag{12}$$

with $B_0 = (I - A_0)v$. Similarly, the matrix form of $C^T(\lambda) = [I - A^T(\lambda)]y(\lambda)$ gives equation (9) and

$$\begin{bmatrix} C_0^T \\ \vdots \\ C_l^T \end{bmatrix} = \begin{bmatrix} -A_l^T & \dots & I - A_0^T \\ & \ddots & \vdots \\ & & -A_l^T \end{bmatrix} \begin{bmatrix} y_l \\ \vdots \\ y_0 \end{bmatrix} \tag{13}$$

with $C_0 = w^T(I - A_0)$. The condition $H(\lambda) = 1$ implies that $D(\lambda) = -y^T(\lambda)[I - A(\lambda)]x(\lambda)$. Writing this in matrix form gives equation (11). Since $D(\lambda)$ is a polynomial of degree l, we have $D_j = 0$ for $j < 0$. For $\lambda = 0$, the transfer function is

$$1 = H(0) = C_0(I - A_0)^\dagger B_0 + D_0 = w^T(I - A_0)v + D_0$$

which implies that equation (10) is satsified. Therefore, Condition 1 is satisfied and the estimator has the form given in equations (11), (12), and (13).

2) Let $\Sigma(L)$ be a polynomial linear protocol with $A(\lambda)$ convergent for all $\lambda \in \text{eig}(L)$ and where Condition 1 is satisfied. Define

$$x(\lambda) = \sum_{k=0}^{nl} \lambda^{-k} x_k, \qquad y(\lambda) = \sum_{k=0}^{nl} \lambda^{-k} y_k.$$

Then equations (8) and (12) combine to give the matrix form of the polynomial matrix equation $B(\lambda) = [I - A(\lambda)]x(\lambda)$ with $B_0 = (I - A_0)v$, so $B(\lambda) \in \text{Col}(I - A(\lambda))$ for all λ. Similarly, equations (9) and (13) combine to give the matrix form of the polynomial matrix equation $C(\lambda) = y^T(\lambda)[I - A(\lambda)]$ with $C_0 = w^T(I - A_0)$, so $C^T(\lambda) \in \text{Col}\big(I - A^T(\lambda)\big)$ for all λ. Equation (11) is the matrix form of the polynomial equation $D(\lambda) = -y^T(\lambda)[I - A(\lambda)]x(\lambda)$ which implies $H(\lambda) = 0$ for $\lambda > 0$. From equation (10), we have that

$$\begin{aligned} H(0) &= C_0(I - A_0)^\dagger B_0 + D_0 \\ &= w^T(I - A_0)v + (1 - w^T(I - A_0)v) = 1. \end{aligned}$$

Then by [9, Theorem 5], the protocol achieves average consensus for constant inputs and is robust to initialization errors. ■

The following corollary shows how to choose the vector w to achieve robustness to changes in network topology.

Corollary 1: A polynomial linear protocol of degree one with the form given in Theorem 1 is asymptotically mean ergodic if w is chosen such that $y_0 = 0$.

Proof: If $y_0 = 0$, then $C_1^T = -y_0^T A_1 = 0$ and $D_1 = y_1^T A_1 x_{-1} = 0$, so the protocol is asymptotically mean ergodic by [13, Theorem 4]. ■

From Theorem 1 and Corollary 1, Condition 1 is both necessary and sufficient for a polynomial linear protocol to have the properties listed in Problem 1. Therefore, we can robustly optimize the spectral radius of $A(\lambda)$ subject to Condition 1 and then choose $y_0 = 0$ to obtain all of the desired properties. We show how to choose v in Section V-A.

IV. ROBUST OPTIMIZATION OF SPECTRAL RADIUS

In this section we develop a solution to Problem 2. We start with the simplified problem of finding conditions such that $\rho(A(\lambda)) < \alpha$ for some fixed λ, and then add additional constraints to guarantee that $\rho(A(\lambda)) < \alpha$ for all $\lambda \in [\lambda_{\text{min}}, \lambda_{\text{max}}]$. Bisection is then performed on α to obtain the solution to Problem 2.

A. Fixed λ

First, we develop conditions such that $\rho(A(\lambda_0)) < \alpha$ for some fixed λ_0. Define the degree n polynomials in μ,

$$x(\mu) = \det(\mu I - A(\lambda_0)) = \sum_{i=0}^{n} x_i \mu^i$$

$$\tilde{x}(\mu) = \left(\frac{\mu}{\alpha}\right)^n x(\alpha^2/\mu) = \sum_{i=0}^{n} \tilde{x}_i \mu^i$$

and denote the vectors of coefficients as

$$\begin{aligned} \mathbf{x} &= \begin{bmatrix} x_0 & x_1 & \dots & x_n \end{bmatrix}^T \\ \tilde{\mathbf{x}} &= \begin{bmatrix} \tilde{x}_0 & \tilde{x}_1 & \dots & \tilde{x}_n \end{bmatrix}^T. \end{aligned}$$

We can now state the following lemma which follows immediately from [15, Lemma 1].

Lemma 1: For fixed λ_0, $\rho(A(\lambda_0)) < \alpha$ if and only if $H > 0$ where H satisfies the linear system of equations

$$\mathbf{x}\mathbf{x}^T - \tilde{\mathbf{x}}\tilde{\mathbf{x}}^T = \begin{bmatrix} -\alpha^2 H & 0_n \\ 0_n^T & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0_n^T \\ 0_n & H \end{bmatrix}.$$

B. Robust λ

Now we develop additional conditions which guarantee that $\rho(A(\lambda)) < \alpha$ for all $\lambda \in [\lambda_{\text{min}}, \lambda_{\text{max}}]$. Define the functions

$$\begin{aligned} f_1(\lambda) &= \det(\alpha I - A(\lambda)) \\ f_2(\lambda) &= \det(\alpha I + A(\lambda)) \\ f_3(\lambda) &= \det(\alpha^2 I - \Psi(A(\lambda))) \end{aligned}$$

where the matrix $\Psi(A(\lambda)) \in \mathbb{R}^{n(n+1)/2 \times n(n+1)/2}$ is defined as

$$(\Psi(A(\lambda)))_{j,k} = \det \begin{pmatrix} A(\lambda)_{y_1,z_1} & A(\lambda)_{y_1,z_2} \\ A(\lambda)_{y_2,z_1} & A(\lambda)_{y_2,z_2} \end{pmatrix}$$

where (y_1, y_2) and (z_1, z_2) are the j-th and k-th pairs in the sequence $(1,2), (1,3), \dots, (1,n), (2,3), (2,4), \dots, (2,n), (3,4), \dots, (n-1,n)$. Then f_i is a polynomial in λ of degree δ_i where $\delta_{1,2} = ln$ and $\delta_3 = ln(n-1)$, so $f_i(\lambda) = \sum_{j=0}^{2m_i} f_{i,j}\lambda^j$ where $m_i = \lceil \delta_i/2 \rceil$ for $i = 1,2,3$. Let $\Gamma_i = \Gamma_i^T \in \mathbb{R}^{m_i \times m_i}$ and $\Delta_i \in \mathbb{R}^{m_i(m_i-1)/2}$. Define the following matrices in $\mathbb{R}^{m_i+1 \times m_i+1}$:

$$(F_i)_{j,k} = \begin{cases} f_{i,j+k-2}, & j = k \\ f_{i,j+k-2}/2, & |j-k| = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$(R_i)_{j,k} = \begin{cases} 1, & j = k = 1 \\ 0, & \text{otherwise} \end{cases}$$

$$\begin{aligned} S_i(\Gamma_i) = \frac{\lambda_{\text{min}} + \lambda_{\text{max}}}{2} &\left(\begin{bmatrix} 0_{m_i} & \Gamma_i \\ 0 & 0_{m_i}^T \end{bmatrix} + \begin{bmatrix} 0_{m_i}^T & 0 \\ \Gamma_i & 0_{m_i} \end{bmatrix} \right) \\ &- \begin{bmatrix} 0 & 0_{m_i}^T \\ 0_{m_i} & \Gamma_i \end{bmatrix} - \lambda_{\text{min}}\lambda_{\text{max}} \begin{bmatrix} \Gamma_i & 0_{m_i} \\ 0_{m_i}^T & 0 \end{bmatrix} \end{aligned}$$

$$T_i(\Delta_i) \in \{T = T^T \in \mathbb{R}^{(m_i+1)\times(m_i+1)} : x^T T x = 0, \ \forall x\}.$$

Then we have the following robust performance condition which appears in [16, Theorem 1].

Lemma 2: $\rho(A(\lambda)) < \alpha$ for all $\lambda \in [\lambda_{\text{min}}, \lambda_{\text{max}}]$ if and only if the following conditions hold:

- $\rho(A(\lambda_0)) < \alpha$ for some $\lambda_0 \in [\lambda_{\text{min}}, \lambda_{\text{max}}]$;
- for $i = 1,2,3$, there exist $\beta_i \in \mathbb{R}$, $\Gamma_i = \Gamma_i^T \in \mathbb{R}^{m_i \times m_i}$, and $\Delta_i \in \mathbb{R}^{m_i(m_i-1)/2}$ which satisfy the PMIs

$$\begin{cases} 0 < \beta_i \\ 0 \le \Gamma_i \\ 0 \le F_i - \beta_i R_i - S_i(\Gamma_i) + T_i(\Delta_i). \end{cases}$$

Note that Lemma 1 can be used to ensure that the first condition is satisfied. In general, the conditions in Lemma 1 and Lemma 2 are polynomial matrix inequalities (PMIs), i.e., matrix inequalities whose coefficients are polynomials in the variables. Global solutions to this type of problem can be found by solving convex LMI relaxations whose solution is guaranteed to converge to the global solution of the original PMI as the size of the relaxation increases. Finite convergence can be detected in many cases and conditions exist to guarantee that the global optimum has been achieved [17]. Due to the number of variables and order of relaxation required, this technique was found to be computationally challenging for problems larger than the PI estimator ($n = 2$ and $l = 1$). For the case of the PI estimator, however, these conditions can be made convex as described in Section V-B.

V. Example: PI Estimator

In this section, the results from Sections III and IV are used to design the optimal worst-case robust PI estimator.

A. Solution to Problem 1

Condition 1 requires that A_0 have a simple (since A_0 must be convergent) eigenvalue at one. To satisfy Condition 1, we can use the parameterization

$$A(\lambda) = \begin{bmatrix} 1-\gamma & 0 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} -k_p & k_I \\ -1 & 0 \end{bmatrix}$$

without loss of generality which gives

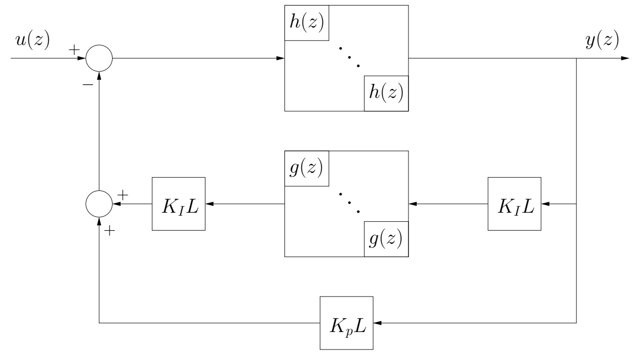
$$\begin{bmatrix} x_1 \\ x_0 - v \end{bmatrix} = \begin{bmatrix} 0 \\ -\gamma/k_I \\ -1 \\ 0 \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ y_0 - w \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1/\gamma \\ 0 \end{bmatrix}.$$

Substituting these into equations (11), (12), and (13) give

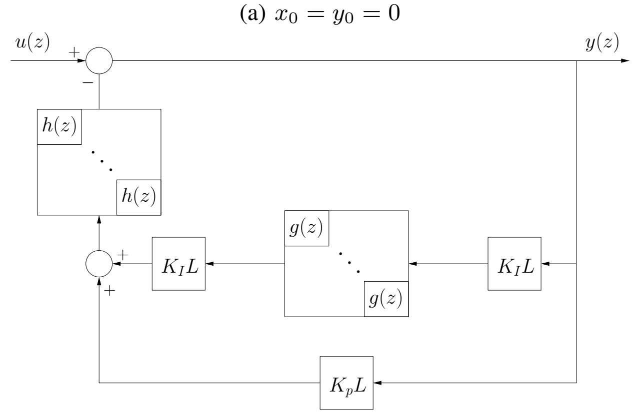
$$\begin{aligned} B(\lambda) &= \begin{bmatrix} \gamma v_1 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} k_p(v_1 - 1) - k_I v_2 \\ v_1 - 1 \end{bmatrix} \\ C^T(\lambda) &= \begin{bmatrix} \gamma w_1 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} k_p(w_1 - 1/\gamma) + w_2 \\ -k_I(w_1 - 1/\gamma) \end{bmatrix} \\ D(\lambda) &= (1 - w_1\gamma v_1) + \lambda\Big[k_I v_2(w_1 - 1/\gamma) \\ &\qquad - [k_p(w_1 - 1/\gamma) + w_2](v_1 - 1)\Big] \end{aligned}$$

for any $v, w \in \mathbb{R}^2$ in order for the estimator to be robust to the initial conditions. Using Corollary 1, we choose $w = \begin{bmatrix} 1/\gamma & 0 \end{bmatrix}^T$ to obtain robustness to changes in network topology.

The block diagrams for two simple choices of v and w are shown in Figure 1. Figure 1a is the discrete-time counterpart to the standard PI estimator in [5], and Figure 1b is equivalent to the SOI-DC algorithm in [4] which has the additional benefit of rejecting any signal common to all agents since the input is passed through the Laplacian before filtering. To reject common signals, we choose $v = 0$. Note that the standard PI estimator from Example 1 uses $v = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$



(a) $x_0 = y_0 = 0$



(b) $v = y_0 = 0$

Fig. 1: Block diagram of the PI estimator for three choices of v, w, x_0, and y_0, where $K_p = \frac{k_p}{\gamma}$, $K_I = \frac{k_I}{\gamma}$, $h(z) = \frac{\gamma}{z-(1-\gamma)}$, and $g(z) = \frac{\gamma/k_I}{z-1}$.

and $w = \begin{bmatrix} 1/\gamma & 0 \end{bmatrix}^T$. The form of the estimator is then

$$\left[\begin{array}{c|c} A(\lambda) & B(\lambda) \\ \hline C(\lambda) & D(\lambda) \end{array}\right] = \left[\begin{array}{cc|c} 1-\gamma & 0 & 0 \\ 0 & 1 & 0 \\ \hline 1 & 0 & 1 \end{array}\right] + \lambda \left[\begin{array}{cc|c} -k_p & k_I & -k_p \\ -1 & 0 & -1 \\ \hline 0 & 0 & 0 \end{array}\right]. \quad (14)$$

In the following section, the parameters γ, k_p, and k_I are chosen to solve Problem 2.

B. Solution to Problem 2

For control systems of dimension two, the stability region is convex in the coefficients of the characteristic polynomial [18]. Therefore, using the parameterization of $A(\lambda)$ from Subsection V-A, the condition $H > 0$ in Lemma 1 can be made convex as shown in the following lemma. The proof, omitted for brevity, relies on the fact that the matrix H from Lemma 1 factors when $n = 2$.

Lemma 3: For the PI estimator, $\rho(A(\lambda)) < \alpha$ if and only if

$$\begin{aligned} 0 &< \alpha^2 - \kappa \\ 0 &< \begin{bmatrix} \alpha^2 + \kappa & -(2 - \gamma - k_p\lambda) \\ -(2 - \gamma - k_p\lambda) & (\alpha^2 + \kappa)/\alpha^2 \end{bmatrix} \end{aligned}$$

where $\kappa = 1 - \gamma - k_p\lambda + k_I\lambda^2$.

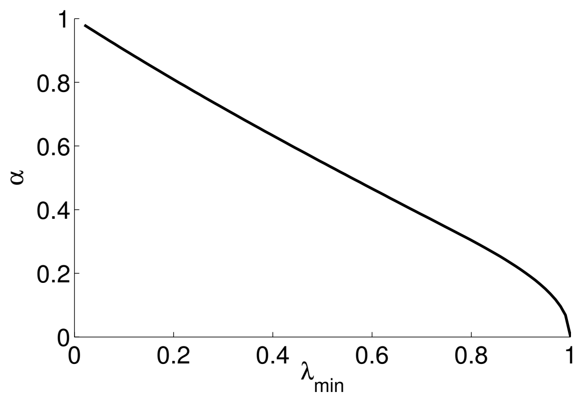


Fig. 2: Worst-case spectral radius of the PI estimator as a function of λ_{min} with $\lambda_{\text{max}} = 1$.

Combining this result with Lemma 2, we have the following.

Theorem 2: For the PI estimator, $\rho(A(\lambda)) < \alpha$ for all $\lambda \in [\lambda_{\text{min}}, \lambda_{\text{max}}]$ if and only if

$$\begin{aligned}
&-\alpha < 1-\gamma < \alpha \\
&0 < \alpha^2 - \kappa \\
&0 < \begin{bmatrix} \alpha^2+\kappa & -(2-\gamma-k_p\lambda_0) \\ -(2-\gamma-k_p\lambda_0) & (\alpha^2+\kappa)/\alpha^2 \end{bmatrix} \\
&0 \leq F_0 + \alpha F_1 - \beta_1 R - \Gamma_1 S \\
&0 \leq F_0 - \alpha F_1 - \beta_2 R - \Gamma_2 S \\
&0 \leq \begin{bmatrix} \alpha^2-(1-\gamma) & k_p/2 \\ k_p/2 & -k_I \end{bmatrix} - \beta_3 R - \Gamma_3 S \\
&0 \leq \Gamma_i, \quad i = 1,2,3 \\
&0 < \beta_i, \quad i = 1,2,3
\end{aligned}$$

where

$$\begin{aligned}
&\kappa = 1-\gamma-k_p\lambda_0+k_I\lambda_0^2, \\
&F_0 = \begin{bmatrix} 1-\gamma+\alpha^2 & -k_p/2 \\ -k_p/2 & k_I \end{bmatrix}, \quad F_1 = \begin{bmatrix} 2-\gamma & -k_p/2 \\ -k_p/2 & 0 \end{bmatrix}, \\
&R = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad S = \begin{bmatrix} -\lambda_{\text{min}}\lambda_{\text{max}} & (\lambda_{\text{min}}+\lambda_{\text{max}})/2 \\ (\lambda_{\text{min}}+\lambda_{\text{max}})/2 & -1 \end{bmatrix},
\end{aligned}$$

and $\lambda_0 \in [\lambda_{\text{min}}, \lambda_{\text{max}}]$ and $\Gamma_i, \beta_i \in \mathbb{R}$ for $i = 1, 2, 3$.

All of the inequalities in Theorem 2 are linear in the parameters γ, k_p, and k_I, so these are convex LMIs. Therefore, bisection can be performed on α to find the minimum α such that the conditions in Theorem 2 have a solution. The package CVX was used to specify and solve the LMIs [19]. Figure 2 shows the optimal worst-case spectral radius as a function of λ_{min} with $\lambda_{\text{max}} = 1$. When $\lambda_{\text{min}} = 0$, the worst-case asymptotic convergence rate is optimized over all undirected graphs, including disconnected graphs which cannot achieve consensus, so $\alpha = 1$. For the other extreme case of $\lambda_{\text{min}} = 1$, the optimization is over only fully-connected undirected graphs which are able to achieve consensus in finite time, so $\alpha = 0$.

VI. Conclusions

We developed a design process for finding optimal worst-case robust dynamic average consensus estimators. By characterizing the structure of polynomial linear protocols which achieve average consensus for constant inputs and are robust to intialization errors, we developed necessary and sufficient conditions on $A(L)$. The problem of optimizing the worst-case asymptotic convergence rate for a general polynomial linear protocol was then formulated as PMIs which have globally optimal methods, although the problem is difficult for $n > 2$. For the special case of the PI estimator, the conditions can be made convex and the global optimal parameters calculated efficiently.

Closed-form solutions for the PI estimator and the case $n = 4$ are known and will be published in future work.

References

[1] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Dynamic consensus on mobile networks," *IFAC World Congress*, pp. 1–6, 2005.

[2] F. Chen, Y. Cao, and W. Ren, "Distributed average tracking of multiple time-varying reference signals with bounded derivatives," *IEEE Trans. Autom. Control*, vol. 57, no. 12, pp. 3169–3174, Dec. 2012.

[3] M. Zhu and S. Martnez, "Discrete-time dynamic average consensus," *Automatica*, vol. 46, no. 2, pp. 322 – 329, 2010.

[4] S. Kia, J. Cortes, and S. Martinez, "Singularly perturbed algorithms for dynamic average consensus," in *Proceedings of the 2013 European Control Conference*, July 2013, pp. 1758–1763.

[5] H. Bai, R. Freeman, and K. Lynch, "Robust dynamic average consensus of time-varying inputs," in *Proceedings of the 49th IEEE Conference on Decision and Control*, 2010, pp. 3104–3109.

[6] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, vol. 5, Dec 2003, pp. 4997–5002.

[7] B. N. Oreshkin, M. J. Coates, and M. G. Rabbat, "Optimization and analysis of distributed averaging with short node memory," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2850–2865, May 2010.

[8] T. Erseghe, D. Zennaro, E. Dall'Anese, and L. Vangelista, "Fast consensus by the alternating direction multipliers method," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5523–5537, Nov. 2011.

[9] R. Freeman, T. Nelson, and K. Lynch, "A complete characterization of a class of robust linear average consensus protocols," in *Proceedings of the 2010 American Control Conference*, 2010, pp. 3198–3203.

[10] M. Elwin, R. Freeman, and K. Lynch, "A systematic design process for internal model average consensus estimators," in *Proceedings of the 52nd IEEE Conference on Decision and Control*, Dec 2013, pp. 5878–5883.

[11] P. Yang, R. Freeman, and K. Lynch, "Optimal information propagation in sensor networks," in *Proceedings of the 2006 IEEE International Conference on Robotics and Automation*, 2006, pp. 3122–3127.

[12] R. Freeman, P. Yang, and K. Lynch, "Stability and convergence properties of dynamic average consensus estimators," in *Proceedings of the 45th IEEE Conference on Decision and Control*, 2006, pp. 338–343.

[13] B. Van Scoy, R. Freeman, and K. Lynch, "Asymptotic mean ergodicity of average consensus estimators," in *Proceedings of the 2014 American Control Conference*, June 2014, pp. 4696–4701.

[14] D. Henrion and M. Sebek, "Numerical methods for polynomial rank evaluation," in *Systems Structure and Control, IFAC,*, 1998, pp. 385–390.

[15] D. Henrion, D. Peaucelle, D. Arzelier, and M. Sebek, "Ellipsoidal approximation of the stability domain of a polynomial," *IEEE Trans. Autom. Control*, vol. 48, no. 12, pp. 2255–2259, Dec 2003.

[16] G. Chesi, "Exact robust stability analysis of uncertain systems with a scalar parameter via LMIs," *Automatica*, vol. 49, no. 4, pp. 1083 – 1086, 2013.

[17] D. Henrion and J.-B. Lasserre, "Convergent relaxations of polynomial matrix inequalities and static output feedback," *IEEE Trans. Autom. Control*, vol. 51, no. 2, pp. 192–202, Feb 2006.

[18] J. Ackermann, "Parameter space design of robust control systems," *IEEE Trans. Autom. Control*, vol. 25, no. 6, pp. 1058–1072, Dec 1980.

[19] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.