2014 American Control Conference (ACC)
June 4-6, 2014. Portland, Oregon, USA

Asymptotic Mean Ergodicity of Average Consensus Estimators

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

Abstract— **Dynamic average consensus estimators suitable for the decentralized computation of global averages of constant or slowly-varying local inputs include the proportional (P) and proportional-integral (PI) estimators. We analyze the convergence properties of these estimators when run on i.i.d. random graphs which are connected and balanced on average, but need not be connected or balanced at each time step. The statistics of the steady-state process are found using the Kronecker product covariance and an ergodic theorem is used to determine whether the steady-state process is mean ergodic. We show that for constant inputs the P estimator is asymptotically mean ergodic only for systems with non-zero forgetting factor which do not have zero steady-state error on average. The PI estimator has both the asymptotic mean ergodicity property and zero steady-state error in expectation for constant inputs independent of initial conditions, proving that the time-averaged output of each agent robustly converges to the correct average.**

I. INTRODUCTION

The average consensus problem is considered in which a group of agents calculates the global average of their inputs using only local interactions with neighboring agents. Applications of average consensus include formation control [1], distributed Kriged Kalman filtering [2], distributed merging of feature-based maps [3], and distributed estimation for motion coordination [4]. We study the problem of average consensus over a random graph topology using the polynomial linear protocol with focus on the proportional (P) and proportional-integral (PI) estimators [5].

Previous work has examined the convergence and robustness properties of the P and PI dynamic average consensus estimators for constant communication graphs [5]. In this paper, we consider the case where the graph is chosen randomly at each time step, modeling noisy communication channels where packets are dropped randomly. While the random graphs prevent asymptotic convergence of each agent's estimate to the average of the constant inputs, in this paper we pose a slightly weaker question: Under what conditions do the time-averaged values of the estimator outputs converge to the actual average of the inputs? With this property, estimator outputs can be low-pass filtered to obtain the correct average.

Convergence properties of average consensus algorithms over random graphs have been studied recently [6], [7], [8], [9], [10]. Much of the analysis has been performed on single-hop static consensus algorithms. Single-hop protocols are often used due to their simplicity in message passing, but higher degree protocols have also been shown to achieve average consensus [11]. We address multi-hop communication using the general polynomial linear protocol framework. This allows for the analysis of a general class of linear consensus protocols of arbitrary degree. Also, the current literature only addresses static consensus algorithms in which the inputs to the system are assigned as initial states. This simplifies the analysis by reducing the problem to the study of infinite products of stochastic matrices. However, this setup is not robust to initialization errors since a single fault at one time step can cause the system to converge to an incorrect value. The system is also not robust to agents joining or leaving the network and must be reinitialized whenever a change occurs. We use dynamic average consensus where the agent inputs are assigned not as initial states but as inputs into the networked dynamic system. Dynamic average consensus estimators are said to be *robust* if they converge to the average of all the inputs independent of the initial state.

To study the steady-state behavior of average consensus estimators over random graphs, we first state an ergodic theorem which gives necessary and sufficient conditions on the steady-state covariance for a discrete time random process to be asymptotically mean ergodic, meaning that the time-average converges to the ensemble average of the steady-state process. An expression for the steady-state covariance of the output of a polynomial linear protocol system is derived. The ergodic theorem is then applied to the polynomial linear protocol to identify conditions under which the estimator is ergodic. Results are applied to the P and PI estimators which are both polynomial linear protocols. The robustness and steady-state error properties of these estimators have already been established for constant graphs, so they are known for the expected system. In this paper, we show that if the randomly chosen graphs are i.i.d. and connected and balanced on average (but with no requirement on each graph being connected and balanced), then the PI estimator is ergodic, robust, and has expected steady-state error of zero, so the time-averaged output of each agent converges to the exact average independent of the initial state. Therefore the exact average can be obtained using a low-pass filter at the output of each agent.

The subsequent sections are organized as follows. Section II introduces the average consensus problem. Section III defines the covariance using the Kronecker product and establishes several useful results of the covariance, and Section IV gives sufficient conditions on the steady-state covariance for a random process to be asymptotically mean ergodic. The main theorem is presented in Section V which develops

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.

the steady-state covariance for a general polynomial linear protocol using the separated system. Results are applied to the P and PI estimators in Section VI, and Section VII summarizes the conclusions.

Notation: The vectors 1_n and 0_n represent the $n \times 1$ vectors with all entries 1 and 0, respectively. I_n represents the $n \times n$ identity matrix. The expectation of a random variable x is denoted $\mathrm{E}[x]$, and the expectation of a matrix A is denoted $\mathrm{E}[A]$ and is the element-wise expectation of each element of A. A diagonal matrix with entries α_i on the diagonal is denoted $\mathrm{diag}(\alpha_1, \ldots, \alpha_n)$. The symbol $\otimes$ denotes the Kronecker product, except in (6) where it represents the tensor product. The spectral radius is denoted $\rho(\cdot)$.

II. Problem Setup

Consider a group of n agents whose communication topology is modeled as a weighted directed graph G. The adjacency matrix of G is defined as $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ where $a_{ij} > 0$ if agent i can receive information from agent j and zero otherwise. The neighbors of agent i, denoted $\mathcal{N}_i$, is the set of agents from which agent i can receive information. The Laplacian matrix is $L = \mathrm{diag}(A1_n) - A$. Therefore, L is positive semidefinite and satisfies $Lv = 0_n$ where $v = 1_n/\sqrt{n}$. The graph is said to be balanced if and only if $v^T L = 0_n^T$. 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 [12]. When the graph is unknown, however, it is often useful to use the weights to bound the eigenvalues of the Laplacian. For example, inverse-out degree weighting assigns $a_{ij} = 1/[\deg(i)+\deg(j)]$ where $\deg(i)$ is the out-degree of agent i. This decentralized weighting scheme restricts the eigenvalues of L to the region $D_0 \cap D_1$ where $D_x \subset \mathbb{C}$ denotes the closed unit disc centered at x [11]. It also has the added advantage of producing symmetric (and therefore balanced) expected Laplacians under suitably symmetric packet-loss probability distributions (see [11] for details).

The consensus problem is to design an estimator whose output converges to the same signal for each agent. Average consensus requires the consensus signal to be $(1/n)\sum_{i=1}^{n} u_k^i$ where u_k^i is the input to agent i at time k. The P estimator is implemented on agent i using the equations

$$x_{k+1}^i = (1-\gamma)x_k^i - k_p \sum_{j \in \mathcal{N}_i} a_{ij}\left[y_k^i - y_k^j\right] \tag{1}$$

$$y_k^i = x_k^i + u_k^i \tag{2}$$

where x_k^i is the internal state and y_k^i is the output of agent i at time k, and γ and k_p are system parameters. The PI estimator equations are

$$\nu_{k+1}^i = (1-\gamma)\nu_k^i + \gamma u_k^i - k_p \sum_{j \in \mathcal{N}_i} a_{ij}\left[\nu_k^i - \nu_k^j\right] - k_I \sum_{j \in \mathcal{N}_i} a_{ij}\left[\eta_k^i - \eta_k^j\right] \tag{3}$$

$$\eta_{k+1}^i = \eta_k^i + k_I \sum_{j \in \mathcal{N}_i} a_{ij}\left[\nu_k^i - \nu_k^j\right] \tag{4}$$

where ν_k^i and η_k^i are the internal states of agent i at time k, ν_k^i is the output, and γ, k_p, and k_I are system parameters.

The convergence properties of the P and PI estimators for constant communication graphs have been studied [5]. In this paper we study the steady-state behavior of the estimators when the graph is not constant; that is, $L = L_k$ is time dependent. We seek to determine the steady-state behavior of the time-varying systems and give conditions under which the estimators achieve average consensus for constant inputs.

III. Kronecker Product Covariance

Given random matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, we would like a way of representing the covariances between $A_{i,j}$ and $B_{k,l}$ for all i, j, k, l. The covariance matrix between two random vectors $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ is generally defined as the $n \times m$ matrix

$$C(x,y) \equiv \mathrm{E}\left[(x - \mathrm{E}[x])\,(y - \mathrm{E}[y])^T\right]. \tag{5}$$

This definition, however, is limited to vectors (and scalars) and does not generalize to higher dimensions. By interpreting vectors as 1st-order tensors and matrices as 2nd-order tensors, the covariance matrix can be viewed as a 2nd-order tensor composed using the outer product of two 1st-order tensors. Basser and Pajevic generalize this concept to obtain the covariance of two 2nd-order tensors resulting in a 4th-order covariance tensor [13]. In general, the covariance between an n^{th}-order random tensor a and m^{th}-order random tensor b can be described by an mn^{th}-order tensor using the tensor product (or outer product),

$$C(a,b) \equiv \mathrm{E}[(a - \mathrm{E}[a]) \otimes (b - \mathrm{E}[b])], \tag{6}$$

where $\otimes$ denotes the tensor product. An intuitive definition for the covariance between two matrices is then given by Definition 1.

Definition 1: The covariance between $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$ is defined as

$$\mathrm{COV}[A,B] \equiv \mathrm{E}[(A - \mathrm{E}[A]) \otimes (B - \mathrm{E}[B])] \tag{7}$$

which has dimensions $mp \times nq$, and where $\otimes$ denotes the Kronecker product. The variance is the covariance between a matrix and itself and is denoted

$$\mathrm{VAR}[A] \equiv \mathrm{COV}[A,A]\,. \tag{8}$$

In the vector case, the covariances using the Kronecker product definition and the standard definition are related by $\mathrm{COV}[x,y] = \mathrm{vec}[C(y,x)]$ where $\mathrm{vec}[A]$ is the vectorization [14] of a matrix formed by stacking the columns of A. Equation (5) is valid only for random vectors while (7) is valid for both random vectors and matrices of any size, so the latter definition will be used throughout this paper. Theorems 1 and 2 provide useful results for the Kronecker product covariance which are analogous to those of the scalar case.

Theorem 1: Given matrices $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{p\times q}$ and vectors $x \in \mathbb{R}^n$, $y \in \mathbb{R}^q$ where A and B are uncorrelated with x and y, the covariance of the product is

$$\begin{aligned}\mathrm{COV}[Ax, By] &= \mathrm{COV}[A, B]\,(\mathrm{E}[x] \otimes \mathrm{E}[y]) \\ &\quad + (\mathrm{E}[A] \otimes \mathrm{E}[B])\,\mathrm{COV}[x, y] + \mathrm{COV}[A, B]\,\mathrm{COV}[x, y] \\ &= \mathrm{COV}[A, B]\,\mathrm{E}[x \otimes y] + (\mathrm{E}[A] \otimes \mathrm{E}[B])\,\mathrm{COV}[x, y] \\ &= \mathrm{COV}[A, B]\,(\mathrm{E}[x] \otimes \mathrm{E}[y]) + \mathrm{E}[A \otimes B]\,\mathrm{COV}[x, y]\,. \end{aligned} \quad (9)$$

For scalars A, B, x, y, this result reduces to

$$\begin{aligned}\mathrm{COV}[Ax, By] &= \mathrm{COV}[A, B]\,\mathrm{E}[x]\,\mathrm{E}[y] \\ &\quad + \mathrm{E}[A]\,\mathrm{E}[B]\,\mathrm{COV}[x, y] + \mathrm{COV}[A, B]\,\mathrm{COV}[x, y]\,. \end{aligned} \quad (10)$$

Theorem 2: Given sets of vectors $x_i \in \mathbb{R}^m$ and $y_j \in \mathbb{R}^n$, the covariance of the sum is

$$\mathrm{COV}\Big[\sum_i x_i, \sum_j y_j\Big] = \sum_i \sum_j \mathrm{COV}[x_i, y_j]. \quad (11)$$

The following definition is useful to simplify notation.

Definition 2: Let n be a non-negative integer. The n^{th} Kronecker power is defined as

$$X^{\otimes n} \equiv \underbrace{X \otimes \ldots \otimes X}_{n}. \quad (12)$$

IV. ASYMPTOTIC MEAN ERGODICITY

Consider a discrete-time random process $\{X_k\}_{k=k_0}^{\infty}$ where $X_k \in \mathbb{R}^n$ for $k \geq k_0$. We extend the ergodic theorem for a wide-sense stationary random process in [15] to an asymptotically wide-sense stationary random process. This establishes conditions under which the time average of the process converges to the ensemble average as k approaches infinity.

Definition 3 (Asymptotically Wide-Sense Stationary): The process X_k is asymptotically wide-sense stationary if and only if the mean and covariance of the steady-state process do not change with time; that is, the limits

$$m_X \equiv \lim_{n\to\infty} E[X_n] \quad (13)$$

and

$$C_X(k) \equiv \lim_{n\to\infty} \mathrm{COV}[X_{k+n}, X_n] \quad (14)$$

exist and are finite where m_X is the mean and $C_X(k)$ is the covariance of the steady-state process.

Definition 4 (Time Average): The time average mean of a random process X_k starting at k_0 is given by

$$\langle X_n \rangle_T = \frac{1}{T} \sum_{k=k_0}^{T+k_0-1} X_{k+n}. \quad (15)$$

We now state conditions under which the process is asymptotically ergodic in the mean. The proof is similar to that of [15] and is omitted for brevity.

Theorem 3 (Asymptotic Mean Ergodicity): Let $\{X_k\}_{k=k_0}^{\infty}$ be a single-sided asymptotically wide-sense stationary discrete-time random process with limiting mean m_X and limiting covariance $C_X(k)$. The process is asymptotically mean ergodic, that is,

$$\lim_{T\to\infty} \lim_{n\to\infty} \langle X_n \rangle_T = m_X, \quad (16)$$

in the mean square sense if and only if the quantity

$$\mathrm{AME}(X) \equiv \lim_{T\to\infty} \frac{1}{T} \sum_{k=-(T-1)}^{T-1} \left(1 - \frac{|k|}{T}\right) C_X(k) \quad (17)$$

is zero.

Corollary 1: An asymptotically wide-sense stationary random process X_k with steady-state covariance given by

$$C_X(k) = \lambda^{|k|} \quad (18)$$

is asymptotically mean ergodic if and only if $|\lambda| \leq 1$ and $\lambda \neq 1$.

Definition 5 (Convergent): A square matrix A is convergent when its power sequence $\{A^k\}_{k=1}^{\infty}$ converges to a finite constant matrix as $k \to \infty$. From the Jordan decomposition, A is convergent if and only if all Jordan blocks associated with eigenvalues at $\lambda = 1$ are of size one, and all other eigenvalues have magnitude less than one.

Corollary 2: An asymptotically wide-sense stationary random process X_k with steady-state covariance given by

$$C_X(k) = CA^{|k|}B \quad (19)$$

where $A \in \mathbb{R}^{n\times n}$ is convergent, $B \in \mathbb{R}^{n\times 1}$, $C \in \mathbb{R}^{1\times n}$, and any eigenvalue of A at one is either uncontrollable through B or unobservable through C, is asymptotically mean ergodic.

V. POLYNOMIAL LINEAR PROTOCOL

A general class of linear protocols which can be used for average consensus is the polynomial linear protocol [11]. Theorem 4 provides conditions under which the output of a polynomial linear protocol is asymptotically mean ergodic by examining the covariance of the output process. Ergodicity implies that time averages are equal to ensemble averages, so the low-pass filtered output of each agent converges to the ensemble average as time approaches infinity if the process is asymptotically mean ergodic. For protocols of degree one, Corollary 3 shows that the expected output is the output of the deterministic system using the expected Laplacian.

Definition 6: A polynomial linear protocol of degree l is the collection $\Sigma(X) = [A(X), B(X), C(X), D(X)]$ where

$$\begin{aligned} A(X) &\equiv \sum_{i=0}^{l} X^i \otimes A_i \qquad & B(X) &\equiv \sum_{i=0}^{l} X^i \otimes B_i \\ C(X) &\equiv \sum_{i=0}^{l} X^i \otimes C_i \qquad & D(X) &\equiv \sum_{i=0}^{l} X^i \otimes D_i \end{aligned} \quad (20)$$

are polynomials in X which describe the linear system

$$x_{k+1} = A(X)x_k + B(X)u_k \quad (21)$$
$$y_k = C(X)x_k + D(X)u_k \quad (22)$$

for $k \geq k_0$. The sizes of matrices and vectors are $X \in \mathbb{R}^{n\times n}$, $A_i \in \mathbb{R}^{p\times p}$, $B_i \in \mathbb{R}^{p\times q}$, $C_i \in \mathbb{R}^{m\times p}$, $D_i \in \mathbb{R}^{m\times q}$, $x_k \in \mathbb{R}^{np\times 1}$, $y_k \in \mathbb{R}^{nm\times 1}$.

Example 1: The P estimator is a polynomial linear protocol of degree one with parameters γ and k_p where $A_0 = 1 - \gamma$, $C_0 = D_0 = 1$, $A_1 = B_1 = -k_p$, and $B_0 = C_1 = D_1 = 0$.

Example 2: The PI estimator is a polynomial linear protocol of degree one with parameters γ, k_p, and k_I where

$$A_0 = \begin{bmatrix} 1-\gamma & 0 \\ 0 & 1 \end{bmatrix} \qquad B_0 = \begin{bmatrix} \gamma \\ 0 \end{bmatrix}$$
$$C_0 = \begin{bmatrix} 1 & 0 \end{bmatrix} \qquad D_0 = 0 \tag{23}$$

and

$$A_1 = \begin{bmatrix} -k_p & k_I \\ -k_I & 0 \end{bmatrix} \qquad B_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$C_1 = \begin{bmatrix} 0 & 0 \end{bmatrix} \qquad D_1 = 0. \tag{24}$$

Note that a polynomial linear protocol $\Sigma(X)$ of degree l may be written as

$$A(X) = \tilde{X}\tilde{A} \qquad B(X) = \tilde{X}\tilde{B}$$
$$C(X) = \bar{X}\tilde{C} \qquad D(X) = \bar{X}\tilde{D} \tag{25}$$

where

$$\tilde{X} = \begin{bmatrix} I_n \otimes I_p & X \otimes I_p & \dots & X^l \otimes I_p \end{bmatrix}, \tag{26}$$
$$\bar{X} = \begin{bmatrix} I_n \otimes I_m & X \otimes I_m & \dots & X^l \otimes I_m \end{bmatrix} \tag{27}$$

and

$$\tilde{A} = \begin{bmatrix} I_n \otimes A_0 \\ \vdots \\ I_n \otimes A_l \end{bmatrix} \qquad \tilde{B} = \begin{bmatrix} I_n \otimes B_0 \\ \vdots \\ I_n \otimes B_l \end{bmatrix}$$
$$\tilde{C} = \begin{bmatrix} I_n \otimes C_0 \\ \vdots \\ I_n \otimes C_l \end{bmatrix} \qquad \tilde{D} = \begin{bmatrix} I_n \otimes D_0 \\ \vdots \\ I_n \otimes D_l \end{bmatrix}. \tag{28}$$

Definition 6 is a permutation of that given by Freeman et al. in [11]. This reordering of the states allows for the separation of the system according to the eigenvalues of X. When X is a Laplacian matrix, it always has an eigenvalue at zero, so the subsystem corresponding to this eigenvalue can be analyzed separately.

A. Separated System

Consider the average consensus problem in Section II. Given a Laplacian matrix $L \in \mathbb{R}^{n\times n}$, consider the polynomial linear protocol $\Sigma(L)$ of degree l. Let $Q \in \mathbb{R}^{n\times n}$ be an orthogonal matrix such that $Q = \begin{bmatrix} v & S \end{bmatrix}$ where $v = 1_n/\sqrt{n}$ and $S \in \mathbb{R}^{n\times(n-1)}$. Define the reduced Laplacian to be $\hat{L} \equiv S^T L S$. Since S and v are orthogonal, we have $v^T S = 0$. To simplify notation, let $\tilde{v} = v \otimes I$ and $\tilde{S} = S \otimes I$ so that $\tilde{Q} = Q \otimes I = \begin{bmatrix} \tilde{v} & \tilde{S} \end{bmatrix}$. Performing the change of variable $\tilde{x}_k = \tilde{Q}^T x_k$, the separated system $\tilde{\Sigma}(L)$ is

$$\tilde{A}(L) = \begin{bmatrix} A_0 & \tilde{v}^T A(L)\tilde{S} \\ 0 & A(\hat{L}) \end{bmatrix} \qquad \tilde{B}(L) = \begin{bmatrix} \tilde{v}^T B(L) \\ \tilde{S}^T B(L) \end{bmatrix} \tag{29}$$
$$\tilde{C}(L) = \begin{bmatrix} v \otimes C_0 & C(L)\tilde{S} \end{bmatrix} \qquad \tilde{D}(L) = D(L) \tag{30}$$

which is equivalent to the original system.

B. Asymptotic Mean Ergodicity

We now state our main theorem which gives conditions under which a time-varying polynomial linear protocol is asymptotically mean ergodic.

Theorem 4: Consider the time-varying polynomial linear protocol $\Sigma(L_k)$ of degree l based on the time-varying Laplacian L_k where $\mathrm{E}[L_k]$ is balanced and connected, and L_k are i.i.d. and independent of the initial state for all k. The output process due to a constant input is asymptotically mean ergodic if the following hold:

1) A_0 is convergent,
2) any eigenvalues of A_0 with magnitude one are unobservable through C_0,
3) $\rho\big(\mathrm{E}\big[A(\hat{L}_k)\big]\big) < 1$, and
4) $C_i = D_i = 0$ for $0 < i \leq l$.

Remark 1: The assumptions on the Laplacian in Theorem 4 do not require the Laplacian to be balanced or connected at any individual time step.

Remark 2: Requirements (1) and (2) in Theorem 4 are also necessary for Σ to achieve robust average consensus [11]. Requirement (3) eliminates the possibility of $\mathrm{E}[A(\hat{L}_k)]$ having simple eigenvalues at one, and requirement (4) is needed to evaluate the expression for the covariance of the output.

Proof: [Theorem 4] Using the separated system in (29) and (30),

$$\begin{bmatrix} z_{k+1} \\ w_{k+1} \end{bmatrix} = \begin{bmatrix} A_0 & \tilde{v}^T A(L_k)\tilde{S} \\ 0 & A(\hat{L}_k) \end{bmatrix} \begin{bmatrix} z_k \\ w_k \end{bmatrix} + \begin{bmatrix} \tilde{v}^T B(L_k) \\ \tilde{S}^T B(L_k) \end{bmatrix} u \tag{31}$$
$$y_k = \begin{bmatrix} v \otimes C_0 & C(L_k)\tilde{S} \end{bmatrix} \begin{bmatrix} z_k \\ w_k \end{bmatrix} + D(L_k)u \tag{32}$$

for $k \geq 0$ and initial conditions z_0 and w_0. The system can also be written as in (25) where $\tilde{A}$, $\tilde{B}$, $\tilde{C}$,$\tilde{D}$, $\tilde{L}_k$ and $\bar{L}_k$ are defined in (26) - (28) (where X is replaced by L_k). Define the steady-state covariance between two random vectors x_k and y_k as

$$\sigma^2_{x,y} \equiv \lim_{k\to\infty} \mathrm{COV}[x_k, y_k]. \tag{33}$$

We want to determine the steady-state covariance of the output in (32),

$$C_y(k) \equiv \lim_{n\to\infty} \mathrm{COV}[y_{n+k}, y_n], \tag{34}$$

to determine if the system is asymptotically mean ergodic using Theorem 3. Note that z_k and w_k are uncorrelated with L_j for $j \geq k$. Using the Kronecker product relations from Section III, the following covariances are zero,

$$\mathrm{COV}\Big[A_0 z_k, \tilde{v}^T A(L_k)\tilde{S} w_k\Big] = (A_0 \otimes \tilde{v}^T \mathrm{E}\Big[\tilde{L}_k\Big]\tilde{A}\tilde{S})\mathrm{COV}[z_k, w_k] = 0 \tag{35}$$

and

$$\mathrm{COV}\big[A_0 z_k, \tilde{v}^T B(L_k)u\big] = (A_0 \otimes \tilde{v}^T)\mathrm{COV}\Big[z_k, \tilde{L}_k\Big](I \otimes \tilde{B}u) = 0 \tag{36}$$

since $\mathrm{E}[L_k]$ is balanced and z_k is uncorrelated with L_k, respectively. Define the variance

$$\begin{aligned}\hat{\sigma}^2(k) &\equiv \mathrm{VAR}\Big[A(L_k)\tilde{S}w_k + B(L_k)u\Big] \\ &= \mathrm{VAR}\Big[\tilde{L}_k(\tilde{A}\tilde{S}w_k + \tilde{B}u)\Big] \\ &= \mathrm{E}\Big[\tilde{L}_k^{\otimes 2}\Big]\,(\tilde{A}\tilde{S})^{\otimes 2}\mathrm{VAR}[w_k] \\ &\quad + \mathrm{VAR}\Big[\tilde{L}_k\Big]\,(\tilde{A}\tilde{S}\ \mathrm{E}[w_k] + \tilde{B}u)^{\otimes 2}. \qquad (37)\end{aligned}$$

Taking the limit as $k \to \infty$ of (37),

$$\begin{aligned}\hat{\sigma}^2 &\equiv \lim_{k\to\infty} \hat{\sigma}^2(k) \qquad (38)\\ &= \mathrm{E}\Big[\tilde{L}_k^{\otimes 2}\Big]\,(\tilde{A}\tilde{S})^{\otimes 2}\sigma_w^2 + \mathrm{VAR}\Big[\tilde{L}_k\Big]\,(\tilde{A}\tilde{S}\bar{w} + \tilde{B}u)^{\otimes 2}\end{aligned}$$

where $\bar{w} = \lim_{k\to\infty} w_k$. The variances and covariances of the state can be found recursively using the Kronecker product relations from Section III. Using (35) and (36), the variance of z_{k+1} is

$$\mathrm{VAR}[z_{k+1}] = A_0^{\otimes 2}\mathrm{VAR}[z_k] + (\tilde{v}^T)^{\otimes 2}\hat{\sigma}_x^2(k). \qquad (39)$$

Similarly, the other variances and covariances of the state are

$$\begin{aligned}\mathrm{VAR}[w_{k+1}] &= (\tilde{S}^T)^{\otimes 2}\hat{\sigma}_x^2(k) \\ &= \mathrm{E}\Big[A(\hat{L}_k)^{\otimes 2}\Big]\,\mathrm{VAR}[w_k] + (\tilde{S}^T)^{\otimes 2} \\ &\quad \mathrm{VAR}\Big[\tilde{L}_k\Big]\,(\tilde{A}\tilde{S}\ \mathrm{E}[w_k] + \tilde{B}u)^{\otimes 2}, \qquad (40)\\ \mathrm{COV}[z_{k+1}, w_{k+1}] &= \Big(A_0 \otimes \mathrm{E}\Big[A(\hat{L}_k)\Big]\Big)\,\mathrm{COV}[z_k, w_k] \\ &\quad + (\tilde{v}^T \otimes \tilde{S}^T)\hat{\sigma}_x^2(k), \qquad (41)\\ \mathrm{COV}[w_{k+1}, z_{k+1}] &= \Big(\mathrm{E}\Big[A(\hat{L}_k)\Big] \otimes A_0\Big)\,\mathrm{COV}[w_k, z_k] \\ &\quad + (\tilde{S}^T \otimes \tilde{v}^T)\hat{\sigma}_x^2(k). \qquad (42)\end{aligned}$$

The steady-state variance of z_k may be infinite due to the possible eigenvalue of A_0 at one. The other systems all have eigenvalues with magnitude less than one, so the steady-state variances and covariances are given by

$$\begin{aligned}\sigma_w^2 &= \Big[I - \mathrm{E}\big[A(\hat{L}_k)^{\otimes 2}\big]\Big]^{-1}(\tilde{S}^T)^{\otimes 2}\mathrm{VAR}\Big[\tilde{L}_k\Big] \\ &\quad (\tilde{A}\tilde{S}\bar{w} + \tilde{B}\bar{u})^{\otimes 2}, \qquad (43)\\ \sigma_{z,w}^2 &= \Big[I - A_0 \otimes \mathrm{E}\Big[A(\hat{L}_k)\Big]\Big]^{-1}(\tilde{v}^T \otimes \tilde{S}^T)\hat{\sigma}_x^2, \qquad (44)\\ \sigma_{w,z}^2 &= \Big[I - \mathrm{E}\Big[A(\hat{L}_k)\Big] \otimes A_0\Big]^{-1}(\tilde{S}^T \otimes \tilde{v}^T)\hat{\sigma}_x^2. \qquad (45)\end{aligned}$$

The covariance between the state at iteration $k+i$ and k is

$$\begin{aligned}\mathrm{COV}[z_{k+i}, z_k] &= \big(A_0^i \otimes I\big)\,\mathrm{VAR}[z_k], \qquad (46)\\ \mathrm{COV}[w_{k+i}, w_k] &= \left(\mathrm{E}\Big[A(\hat{L}_k)\Big]^i \otimes I\right)\mathrm{VAR}[w_k], \qquad (47)\\ \mathrm{COV}[z_{k+i}, w_k] &= \big(A_0^i \otimes I\big)\,\mathrm{COV}[z_k, w_k], \qquad (48)\\ \mathrm{COV}[w_{k+i}, z_k] &= \left(\mathrm{E}\Big[A(\hat{L}_k)\Big]^i \otimes I\right)\mathrm{COV}[w_k, z_k]. \qquad (49)\end{aligned}$$

For the steady-state process, the covariances are

$$\begin{aligned}\lim_{k\to\infty}\mathrm{COV}[z_{k+i}, z_k] &= \big(A_0^i \otimes I\big)\,\sigma_z^2, \qquad (50)\\ \lim_{k\to\infty}\mathrm{COV}[w_{k+i}, w_k] &= \left(\mathrm{E}\Big[A(\hat{L}_k)\Big]^i \otimes I\right)\sigma_w^2, \qquad (51)\\ \lim_{k\to\infty}\mathrm{COV}[z_{k+i}, w_k] &= \big(A_0^i \otimes I\big)\,\sigma_{z,w}^2, \qquad (52)\\ \lim_{k\to\infty}\mathrm{COV}[w_{k+i}, z_k] &= \left(\mathrm{E}\Big[A(\hat{L}_k)\Big]^i \otimes I\right)\sigma_{w,z}^2. \qquad (53)\end{aligned}$$

Using the assumption that $C_i = D_i = 0$ for $0 < i \leq l$, the steady-state covariance of the output at time step $k+i$ and k is

$$\begin{aligned}C_y(i) &= \lim_{k\to\infty}\mathrm{COV}[y_{k+i}, y_k] \qquad (54)\\ &= \lim_{k\to\infty}\,[(v_e \otimes C_0) \otimes (v_e \otimes C_0)]\,\mathrm{COV}[z_{k+i}, z_k] \\ &\quad + [(v_e \otimes C_0) \otimes (S \otimes C_0)]\,\mathrm{COV}[z_{k+i}, w_k] \\ &\quad + [(S \otimes C_0) \otimes (v_e \otimes C_0)]\,\mathrm{COV}[w_{k+i}, z_k] \\ &\quad + [(S \otimes C_0) \otimes (S \otimes C_0)]\,\mathrm{COV}[w_{k+i}, w_k]. \qquad (55)\end{aligned}$$

Using the expressions in (50) to (53),

$$\begin{aligned}C_y(i) &= \big[(v_e \otimes C_0 A_0^i) \otimes (v_e \otimes C_0)\big]\,\sigma_z^2 \\ &\quad + \big[(v_e \otimes C_0 A_0^i) \otimes (S \otimes C_0)\big]\,\sigma_{z,w}^2 \\ &\quad + [(S \otimes C_0) \otimes (v_e \otimes C_0)]\,\Big(\mathrm{E}\Big[A(\hat{L}_k)\Big] \otimes I\Big)^i\,\sigma_{w,z}^2 \\ &\quad + [(S \otimes C_0) \otimes (S \otimes C_0)]\,\Big(\mathrm{E}\Big[A(\hat{L}_k)\Big] \otimes I\Big)^i\,\sigma_w^2. \qquad (56)\end{aligned}$$

Both A_0 and $\mathrm{E}[\hat{L}_k]$ are convergent. Any eigenvalues at one of A_0 are unobservable through C_0, and $\mathrm{E}[\hat{L}_k]$ has no eigenvalues at one. Therefore we can apply Corollary 2. The steady-state variance σ_z^2 may be infinite, however, which would cause the system to not be ergodic. σ_z^2 may only have infinite values in positions corresponding to an eigenvalue at one of A_0, but this does not affect the output since the eigenvalue at one is unobservable through C_0. Therefore the output y_k is ergodic by Corollary 2. ■

For protocols of degree one, the system is linear in L_k which results in the following.

Corollary 3: Under the same assumptions as Theorem 4, the time-averaged output of each agent converges to the same output as $\Sigma(\mathrm{E}[L_k])$ if the protocol is degree one.

VI. DYNAMIC AVERAGE CONSENSUS

Both the P and PI dynamic average consensus estimators are polynomial linear protocols of degree one, so asymptotic mean ergodicity implies that the time-averaged output of each agent converges to the output of the estimator using the expected Laplacian by Corollary 3. The expected steady-state error and robustness to initial conditions is known for the deterministic system [11]. The ergodicity property of each estimator is discussed below, and the results are shown in Table I.

TABLE I
SUMMARY OF PROPERTIES FOR THE P AND PI ESTIMATORS WITH $\mathrm{E}[L_k]$ BALANCED AND CONNECTED, AND L_k I.I.D. FOR ALL k.

Estimator	**Ergodic**	**Robust**	$\lim_{k\to\infty} \mathrm{E}[e_k]$
P, $\gamma = 0$	No	No	Zero[1]
P, $\gamma \neq 0$	Yes	Yes	Non-zero
PI	Yes	Yes	Zero

[1] If the expectation of the initial state is zero.

A. P Estimator

Consider the P estimator in example 1. The estimator has different properties depending on the forgetting factor γ since A_0 has an eigenvalue at one if $\gamma = 0$. We analyze the two cases separately.

1) Case 1: $\gamma \neq 0$: In this case, the eigenvalue of A_0 is $1-\gamma$ which must have magnitude less than one, so we require $0 < \gamma < 2$. In order for the eigenvalues of $\mathrm{E}[A(\hat{L}_k)]$ to have magnitude less than one, k_p must be chosen such that $|(1-\gamma) - k_p\lambda_i| < 1$ for each eigenvalue λ_i of $\mathrm{E}[\hat{L}_k]$. If both constraints are satisfied, then the four conditions in Theorem 4 are satisfied so the protocol is asymptotically mean ergodic. The expected output, however, does not converge to the correct average with zero steady-state error. Therefore the time average converges to the ensemble average, but neither is the correct average.

2) Case 2: $\gamma = 0$: In this case, the pair (A_0, C_0) has an observable eigenvalue at one, so the estimator is not ergodic. The system can be made ergodic, however, with extra restrictions. If the Laplacian is balanced at each time step, then $\sigma^2_{z,w} = 0$. In addition, $\sigma^2_z = 0$ if $\mathrm{VAR}[z_0] = 0$. Using (56), the output is ergodic if L_k is balanced for all k and $\mathrm{VAR}[z_0] = 0$. In this case the estimator does converge to the correct average, so ergodicity gives that the time-averaged output of each agent converges to the correct average, but this requires the extra restrictions on the Laplacian and initial state.

B. PI Estimator

Consider the PI estimator in example 2. The eigenvalue at one of A_0 is unobservable through C_0, and the other eigenvalue of A_0 is $1-\gamma$ so we require $0 < \gamma < 2$. The constants k_p and k_I must be chosen such that the eigenvalues of $A(\mathrm{E}[\hat{L}_k])$ have magnitude less than one. If these conditions are satisfied, then the output y_k is asymptotically mean ergodic by Theorem 4, so the time-averaged output of each agent converges to the ensemble average. Since the expected steady-state error is zero independent of initial conditions, the time-averaged output of each agent robustly converges to the correct average.

VII. CONCLUSIONS

We studied the convergence properties of the P and PI estimators when packets are dropped at random. Dropped packets were modeled by i.i.d. random Laplacians assumed to be balanced and connected on average, although the Laplacian need not be balanced or connected at any time step. This model is limited to situations in which the packet drop probabilities are symmetric between two agents, so that the expected Laplacian is balanced.

To study the convergence properties of average consensus estimators over random graphs, the covariance was defined using the Kronecker product in order to obtain an expression for the steady-state covariance of the system output. An ergodic theorem then gave conditions under which the output of a polynomial linear protocol is asymptotically mean ergodic. Results were applied to the P and PI estimators, and it was shown that the P estimator is ergodic either if $\gamma \neq 0$, or the Laplacian is balanced at each time step and the initial state is deterministic. The PI estimator, however, is always ergodic (provided that γ, k_p, k_I satisfy the conditions in VI-B), has zero expected steady-state error, and is robust to initial conditions. For the PI estimator, the time-averaged output of each agent converges with zero steady-state error in the presence of dropped packets independent of initial conditions.

REFERENCES

[1] P. Yang, R. Freeman, and K. Lynch, "Multi-agent coordination by decentralized estimation and control," *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2480–2496, 2008.

[2] J. Cortes, "Distributed kriged kalman filter for spatial estimation," *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2816–2827, 2009.

[3] R. Aragues, J. Cortes, and C. Sagues, "Distributed consensus on robot networks for dynamically merging feature-based maps," *IEEE Transactions on Robotics*, vol. 28, no. 4, pp. 840–854, 2012.

[4] C. K. Peterson and D. A. Paley, "Distributed estimation for motion coordination in an unknown spatially varying flowfield," *Journal of Guidance, Control, and Dynamics*, vol. 36, no. 3, pp. 894–898, 2013.

[5] 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.

[6] K. Cai and H. Ishii, "Average consensus on arbitrary strongly connected digraphs with dynamic topologies," in *Proceedings of the 2012 American Control Conference*, 2012, pp. 14–19.

[7] N. Vaidya, C. Hadjicostis, and A. Dominguez-Garcia, "Robust average consensus over packet dropping links: Analysis via coefficients of ergodicity," in *Proceedings of the 51st IEEE Conference on Decision and Control*, 2012, pp. 2761–2766.

[8] J.-Y. Chen and J. Hu, "On the convergence of distributed random grouping for average consensus on sensor networks with time-varying graphs," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 4233–4238.

[9] Y. Chen, R. Tron, A. Terzis, and R. Vidal, "Corrective consensus: Converging to the exact average," in *Proceedings of the 49th IEEE Conference on Decision and Control*, 2010, pp. 1221–1228.

[10] T. Li and J.-F. Zhang, "Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2043–2057, 2010.

[11] 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.

[12] 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.

[13] P. J. Basser and S. Pajevic, "Spectral decomposition of a 4th-order covariance tensor: Applications to diffusion tensor MRI," *Signal Processing*, vol. 87, no. 2, pp. 220 – 236, 2007.

[14] H. Neudecker, "Some theorems on matrix differentiation with special reference to kronecker matrix products," *Journal of the American Statistical Association*, vol. 64, no. 327, pp. 953–963, 1969.

[15] A. Leon-Garcia, *Probability, Statistics, and Random Processes for Electrical Engineering*. Pearson/Prentice Hall, 2008.