Accelerated consensus in multi-agent networks via memory of local averages
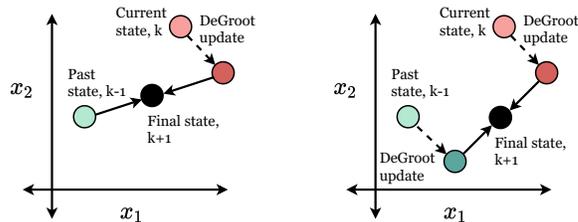
Aditya Bhaskar*, Shriya Rangarajan*, Vikram Shree*, Mark Campbell, and Francesca Parise

Abstract— Classical mathematical models of information sharing and updating in multi-agent networks use linear operators. In the paradigmatic DeGroot model, for example, agents update their states with linear combinations of their neighbors' current states. In prior work, an accelerated averaging model employing the use of memory has been suggested to accelerate convergence to a consensus state for undirected networks. There, the DeGroot update on the current states is followed by a linear combination with the previous states. We propose a modification where the DeGroot update is applied to the current and previous states and is then followed by a linear combination step. We show that this simple modification applied to undirected networks permits convergence even for periodic networks. Further, it allows for faster convergence than DeGroot and accelerated averaging models for suitable networks and model parameters.

I. INTRODUCTION

Linear models for information sharing play a crucial role in multi-agent systems and sensor networks. Among these, distributed averaging algorithms have been intensively studied as a way to reach global consensus with local computations. One of the first such models was proposed by DeGroot [1]. There, agents update their state by taking the weighted average of their neighbours' states at each time step. Many variations have since accounted for real-world phenomena (changing communication patterns or individuals' stubbornness), and optimized convergence rates.

As a starting point for this paper, we consider an accelerated averaging model first proposed by Muthukrishnan et al. in [2]. Therein, authors suggest a modification of the classic DeGroot scheme where agents update their states by first taking a DeGroot update of the current states and then performing a linear combination between this and their previous states, as shown in Fig. 1a. In [2] they show that this additional memory step allows for faster convergence in networks where the original DeGroot model also converges. In this paper, our modification of the Muthukrishnan et al. averaging scheme involves the agents performing DeGroot averaging on the current and previous states followed by a linear combination step, as shown in Fig. 1b. Hereafter,

(a) Accelerated averaging model   (b) MLA model

Fig. 1: Representation of discrete-time averaging algorithms in a two-dimensional state space (x_1, x_2). (a) In the accelerated averaging model, a DeGroot update is performed on current state at k. This is followed by a linear combination step to get the final state at $k+1$. (b) In MLA, we apply DeGroot update on the current and past states. This is followed by a linear combination of the intermediate states to obtain the final state at $k+1$.

we call this model the Memory of Local Averages (MLA) model. Besides deriving convergence guarantees for this modified scheme, our main objective is to show that this simple modification leads to two important consequences: First, we demonstrate conditions for undirected, connected and periodic networks under which MLA achieves consensus despite the DeGroot and the accelerated averaging models failing to do so. Second, we give sufficient conditions under which MLA achieves faster convergence than the DeGroot model and the accelerated averaging algorithm.

Our work is part of a rich literature on opinion dynamics, information sharing and consensus models. Besides classic DeGroot dynamics, several modifications have been proposed to include stubborn agents (e.g., [3], [4]) or dynamics on random or state-dependent networks (e.g. in [5] [6]) among others. We refer to [7] and [8] for recent surveys.

Our work is most related to the literature that focuses on capturing convergence rates while maintaining the original DeGroot structure. In general, faster convergence can be achieved either by tuning the network weights as suggested e.g. in [9] [10] or by modifying the update dynamics as suggested e.g. in [2]. Our work belongs to the latter class. We note that whereas in [2] and our work, faster convergence is obtained by introducing one step of memory, the use of a prediction step has also been investigated (see [11], [12]). Similar to these works, we retain linearity of the update equation which permits for analysis using techniques of linear algebra and Markov chains. We conclude by noting that significant work focused also on deriving finite-time consensus schemes [13], [14], [15]. Such schemes require non-linear updates, which may carry additional computational burden as compared to their linear counterparts.

*Equal contribution.

A. Bhaskar, V. Shree, and M. Campbell are with the Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY, USA. Email: {ab2823, vs476, mc288}@cornell.edu.

S. Rangarajan is with the Department of City and Regional Planning, Cornell University, Ithaca, NY, USA. Email: sr2248@cornell.edu.

F. Parise is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA. Email: fp264@cornell.edu.

II. THE MLA MODEL

A. Model formulation

Let $n \geq 2$ agents be connected in a network with corresponding weighted adjacency matrix, $A = [a_{ij}]_{n \times n}$. Here, a_{ij} denotes the weight that agent i assigns to agent j. The state of each agent, i, at time step, $k \in \mathbb{Z}_{\geq 0}$, is denoted by $x_i(k)$. Throughout the paper, we use the following assumption.

Assumption 1: The weighted adjacency matrix $A_{n \times n}$ is *non-negative* i.e. $a_{ij} \geq 0 \ \forall \ i, j \in \{1, \dots, n\}$ and *row-stochastic* i.e. $\sum_{j=1}^{n} a_{ij} = 1 \ \forall \ i \in \{1, \dots, n\}$.

In the DeGroot model, each agent updates its state with the weighted average of the current states of its local neighbors,

$$x_{i,\text{DG}}(k+1) = \sum_{j=1}^{n} a_{ij} x_j(k),$$
$$\Rightarrow \mathbf{x}_{\text{DG}}(k+1) = A\mathbf{x}(k), \qquad (1)$$

where $\mathbf{x}(k) = [x_1(k), \dots, x_n(k)]^T \in \mathbb{R}^n$.

In contrast, the MLA model is an averaging algorithm with memory which takes into account each agent's current and *previous* states. In particular, given the current states of the agents $\mathbf{x}(k)$, and the previous states $\mathbf{x}(k-1)$, we first apply Degroot model to get intermediate states,

$$\mathbf{z}(k) = A\mathbf{x}(k) \text{ and } \mathbf{z}(k-1) = A\mathbf{x}(k-1).$$

We then obtain the final state $\mathbf{x}(k+1)$ by taking a linear combination of $\mathbf{z}(k)$ and $\mathbf{z}(k-1)$ with weight γ

$$\mathbf{x}(k+1) = \gamma \mathbf{z}(k) + (1-\gamma)\mathbf{z}(k-1)$$
$$\Rightarrow \mathbf{x}(k+1) = \gamma A\mathbf{x}(k) + (1-\gamma)A\mathbf{x}(k-1), \qquad (2)$$

where the initial conditions are $\mathbf{x}(0) = \mathbf{x}(-1)$. In MLA, each agent only needs to store the local average at the previous time step, $\mathbf{z}(k-1)$, instead of the individual states of all its neighbors. Note that for $\gamma = 1$, Eq. (2) reduces to the DeGroot model.

B. Augmented system

To study convergence properties of the system given by Eq. 2 we define an augmented state, $\hat{\mathbf{x}}(k) \in \mathbb{R}^{2n}$ and an augmented iteration matrix, $\hat{A} \in \mathbb{R}^{2n \times 2n}$, as follows

$$\hat{\mathbf{x}}(k) \equiv \begin{bmatrix} \mathbf{x}(k) \\ \mathbf{x}(k-1) \end{bmatrix}, \qquad (3)$$

$$\Rightarrow \hat{\mathbf{x}}(k+1) = \begin{bmatrix} \gamma A & (1-\gamma)A \\ \mathbb{I}_{n \times n} & \mathbb{0}_{n \times n} \end{bmatrix} \hat{\mathbf{x}}(k) \equiv \hat{A}\hat{\mathbf{x}}(k). \qquad (4)$$

Spectral properties of $\hat{A}$ in relation to A are discussed in the following lemma.

Lemma 1: If $(\lambda, \mathbf{v})$ denotes an eigenpair of A, then the eigenpairs of $\hat{A}$ are $(\hat{\lambda}_{\pm}, \hat{\mathbf{v}}_{\pm})$ where,

$$\hat{\lambda}^2 - \gamma \lambda \hat{\lambda} + (\gamma - 1)\lambda = 0, \qquad (5)$$

$$\Rightarrow \hat{\lambda}_{\pm} = \frac{\gamma \lambda \pm \sqrt{\gamma^2 \lambda^2 - 4(\gamma - 1)\lambda}}{2}, \qquad (6)$$

$$\hat{\mathbf{v}}_{\pm} = \begin{bmatrix} \hat{\lambda}_{\pm} \mathbf{v} \\ \mathbf{v} \end{bmatrix}. \qquad (7)$$

Remark: Eqs. (6) and (7) are derived from an eigenanalysis of $\hat{A}$ as given in the companion arXiv paper [16].

III. CONVERGENCE GUARANTEES AND CONSENSUS VALUE

A. Convergence guarantees for connected undirected networks

Convergence properties of (4) (and thus (2)) can be connected to spectral properties of $\hat{A}$ as stated in Lemma 2.

Lemma 2 (Theorem 2.7 in [7]): A square matrix, $B \in \mathbb{R}^{n \times n}$, is semi-convergent and not convergent i.e. $\lim_{k \to \infty} B^k$ exists different from $\mathbb{0}_{n \times n}$ if and only if (i) 1 is an eigenvalue of B and is semi-simple (ii) all other eigenvalues of B have magnitude strictly less than 1.

Assumption 2: The weighted adjacency matrix A is *symmetric* i.e. $a_{ij} = a_{ji} \ \forall \ i, j \in \{1, 2, \dots, n\}$ and *irreducible* i.e. $\sum_{k=0}^{n-1} A^k > 0$.

Remark: Assumption 2 is equivalent to the network being undirected and connected. The spectrum of A is the set of all eigenvalues of A i.e. $\text{spec}(A) \equiv \{\lambda_1, \lambda_2, \dots, \lambda_n\}$ with the condition $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ and the corresponding eigenvectors to be $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. By Perron-Frobenius theorem for irreducible matrices, under Assumptions 1 and 2, (i) $\lambda_1 = 1$ is a simple eigenvalue of A, (ii) $|\lambda_i| \leq 1 \ \forall \ i \in \{2, 3, \dots, n\}$.

Theorem 3: Suppose Assumptions 1 and 2 hold. Then $\hat{A}$ is semi-convergent if and only if:

(i) $\gamma \in (0, 2)$,
(ii) $2\gamma\lambda_n - \lambda_n + 1 > 0$,

where λ_n is the smallest eigenvalue of A.

Proof: We proceed in three steps:

(1) $\hat{A}$ is semi-convergent $\Rightarrow \gamma \in (0, 2)$

We prove this by contradiction. Since $\lambda_1 = 1$ is an eigenvalue of A, from Lemma 1, $\{1, \gamma - 1\}$ are eigenvalues for $\hat{A}$ with corresponding eigenvectors $\{[\mathbf{v}_1; \mathbf{v}_1], [(\gamma - 1)\mathbf{v}_1; \mathbf{v}_1]\}$. Now if $\gamma \in \mathbb{R} \setminus (0, 2]$ then $\hat{A}$ has an eigenvalue $\hat{\lambda}_{1-} = \gamma - 1 \in \mathbb{R} \setminus (-1, 1]$, which is absurd from Lemma 2 since $\hat{A}$ is semi-convergent.

If $\gamma = 2$, $\hat{A}$ has an eigenvalue $\hat{\lambda} = 1$ with algebraic multiplicity 2 (from Lemma 1). The associated eigenspace is $E_1 = \{\hat{\mathbf{v}} | \ \hat{A}\hat{\mathbf{v}} = \hat{\mathbf{v}}\}$. Note that for $\gamma = 2$,

$$\hat{A}\hat{\mathbf{v}} = \begin{bmatrix} 2A & -A \\ \mathbb{I}_{n \times n} & \mathbb{0}_{n \times n} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{v}}_1 \\ \hat{\mathbf{v}}_2 \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{v}}_1 \\ \hat{\mathbf{v}}_2 \end{bmatrix}, \qquad (8)$$

which implies $\hat{\mathbf{v}}_1 = \hat{\mathbf{v}}_2 = \mathbf{v}$. Thus, $2A\hat{\mathbf{v}}_1 - A\hat{\mathbf{v}}_2 = A\mathbf{v} = \hat{\mathbf{v}}_1 = \mathbf{v}$. Here $\mathbf{v}$ is an eigenvector of A. Since by assumption, the geometric multiplicity of $\lambda = 1$ in A is 1, the geometric multiplicity of $\hat{\lambda} = 1$ in $\hat{A}$ is also 1. This is absurd since $\hat{A}$ semi-convergent implies $\hat{\lambda} = 1$ has same algebraic and geometric multiplicity.

(2) $\hat{A}$ is semi-convergent $\Rightarrow 2\gamma\lambda_n - \lambda_n + 1 > 0$

By Eq. (5) $\text{spec}(\hat{A}) = \{\hat{\lambda}_{i\pm} | \ \lambda_i \in \text{spec}(A)\}$. Since $\hat{A}$ is semi-convergent $|\hat{\lambda}_{i\pm}| \leq 1 \ \forall \ i$. Moreover, since $\hat{\lambda} = 1$ has geometric multiplicity 1 and $\hat{A}$ is semi-convergent, it must be that $\hat{\lambda}_1$ has geometric multiplicity 1. Given that $\hat{\lambda}_{1+} = 1$, we conclude that $|\hat{\lambda}_{i\pm}| < 1 \ \forall \ i \in \{2, 3, \dots, n\}$.

Lemma 4 (Lemma 8.5 in [17]): The polynomial $z^2 + az + b = 0$, where $a, b \in \mathbb{C}$, has roots $|z_{1,2}| < 1$ if and only if roots s_1 and s_2 of $(1 + a + b)s^2 + 2(1 - b)s + b - a + 1 = 0$ satisfy $\mathfrak{Re}(s_1) < 0$ and $\mathfrak{Re}(s_2) < 0$.

Applying Lemma 4 to Eq. (5), $|\hat{\lambda}_{i\pm}| < 1$ if and only if the following Eq. (9) has roots with strictly negative real parts.

$$(1 - \gamma\lambda_i + (\gamma - 1)\lambda_i)s^2 + 2(1 - (\gamma - 1)\lambda_i)s + \ldots$$
$$\ldots + ((\gamma - 1)\lambda_i + \gamma\lambda_i + 1) = 0,$$
$$\Rightarrow (1 - \lambda_i)s^2 + 2(1 - \gamma\lambda_i + \lambda_i)s + (2\gamma\lambda_i - \lambda_i + 1) = 0, \quad (9)$$

For fixed $i \in \{2, 3, \ldots, n\}$, let the roots of Eq. (9) be $s_1, s_2 \in \mathbb{C}$. We have the sum of roots of Eq. (9),

$$s_1 + s_2 = \frac{-2(1 - \gamma\lambda_i + \lambda_i)}{1 - \lambda_i} \in \mathbb{R}. \quad (10)$$

Since the sum of roots is real, s_1 and s_2 are of the form, $s_1 = a_1 + bi$ and $s_2 = a_2 - bi$, where $a_1, a_2, b \in \mathbb{R}$. We then note that,

$$a_1 < 0, \ a_2 < 0 \Leftrightarrow c_1 : a_1 + a_2 < 0, \ c_2 : a_1 a_2 > 0. \quad (11)$$

We study conditions c_1 and c_2 independently. Note that for condition c_1,

$$a_1 + a_2 = s_1 + s_2 = \frac{-2(1 - (\gamma - 1)\lambda_i)}{1 - \lambda_i} < 0 \Leftrightarrow (\gamma - 1)\lambda_i < 1. \quad (12)$$

Since, as proved above, semi-convergence implies $\gamma \in (0, 2)$ and $|\lambda_i| \leq 1$, condition c_1 is automatically satisfied $\forall \ i \in \{2, 3, \ldots, n\}$. For condition c_2, when $b = 0$, using the product of roots of Eq. (9), we can write,

$$a_1 a_2 = s_1 s_2 = \frac{2\gamma\lambda_i - \lambda_i + 1}{1 - \lambda_i} > 0,$$
$$\Leftrightarrow 2\gamma\lambda_i - \lambda_i + 1 > 0, \ \forall \ i \in \{2, \ldots, n\}.$$

Specifically, for λ_n, we get,

$$2\gamma\lambda_n - \lambda_n + 1 > 0. \quad (13)$$

When $b \neq 0$, $a_1 = a_2 \Rightarrow a_1 a_2 = a_1^2 \geq 0$. Since, $a_1 + a_2 < 0$ from condition c_1, $a_1 \neq 0$. Hence, $a_1 a_2 > 0$.

(3) The conditions (i) $\gamma \in (0, 2)$ and (ii) $2\gamma\lambda_n - \lambda_n + 1 > 0$ $\Rightarrow \hat{A}$ is semi-convergent.

Recall that from Lemma 1, the eigenvalue λ of A gets mapped to eigenvalues $\hat{\lambda}_\pm$ in $\hat{A}$. Then $\lambda = 1$ is mapped to $\{1, \gamma - 1\}$, where $\gamma \in (0, 2)$.

We next show that all the other eigenvalues of $\hat{A}$ are mapped inside a unit disk, which is enough to prove that $\hat{A}$ is semi-convergent by Lemma 2. Again, applying Lemma 4 to Eq. (5), we know that $\hat{\lambda}_{i\pm}$ lies within the unit-disk if and only if the roots of Eq. (9) lie in the open left half plane. As before, the roots of Eq. (9) are in the left half plane if and only if,

$$c_1 : -2(1 - (\gamma - 1)\lambda_i) < 0, \quad (14)$$
$$c_2 : (2\gamma - 1)\lambda_i + 1 > 0. \quad (15)$$

Recall that $\lambda_i \in [-1, 1)$ for $i \in \{2, \ldots, n\}$, hence for $\gamma \in (0, 2)$, criterion c_1 holds. For proving c_2, consider two cases: (a) $\gamma \in (0, 1)$ and (b) $\gamma \in [1, 2)$.

For case (a), $|2\gamma - 1| < 1$, and since $|\lambda_i| \leq 1$, the condition $(2\gamma - 1)\lambda_i + 1 > 0$, holds $\forall \ i \in \{2, \ldots, n\}$. For case (b), $2\gamma - 1 > 0$ and $f(\lambda_i) \equiv (2\gamma - 1)\lambda_i + 1$ is a monotonically increasing function. Since, $\lambda_i \geq \lambda_n \ \forall \ i \in \{2, \ldots, n\}$ and $f(\lambda_n) > 0$, we infer that $f(\lambda_i) = (2\gamma - 1)\lambda_i + 1 \geq f(\lambda_n) > 0 \ \forall \ i \in \{2, \ldots, n\}$.

Overall, the eigenvalues of $\hat{A}$ are $\{1, \gamma - 1, \hat{\lambda}_{i\pm}\}$, with $\hat{\lambda}_{1+} = 1$ simple and $(\gamma - 1)$, $\hat{\lambda}_{i\pm}$ inside the unit disk $\forall \ i \in \{2, \ldots, n\}$. From Lemma 2 $\hat{A}$ is semi-convergent. ∎

B. Special case of periodic networks

Note that Theorem 3 does not require the network to be aperiodic. If the network is periodic, then $\lambda_n = -1$ and convergence can be guaranteed by Theorem 3 for $\gamma \in (0, 1)$. This is in sharp contrast with the DeGroot model where, if the network is periodic, we can always find an initial condition that causes persistent oscillations.

For comparison, we next discuss the accelerated averaging model by Muthukrishnan et al. [2]. Similar to our model, it also takes into account agents' current and previous states, however the update equation in [2] is given by,

$$\mathbf{x}(k+1) = \beta A \mathbf{x}(k) + (1 - \beta)\mathbf{x}(k-1), \quad (16)$$

where $k \in \mathbb{Z}_{\geq 0}, \beta \in \mathbb{R}$ and initial states $\mathbf{x}(0) = \mathbf{x}(-1) = \mathbf{x}_0$.

Lemma 5: Under Assumptions 1 and 2, if the network associated with A is periodic, there exist initial conditions $\mathbf{x}_0$ s.t. the accelerated averaging model given by Eq. (16) does not converge.

Proof: The augmented iteration matrix for the model in Eq. (16) is $A_\beta \in \mathbb{R}^{2n \times 2n}$,

$$A_\beta = \left[\begin{array}{c|c} \beta A & (1 - \beta)\mathbb{I}_{n \times n} \\ \hline \mathbb{I}_{n \times n} & \mathbb{0}_{n \times n} \end{array} \right] \in \mathbb{R}^{2n \times 2n}. \quad (17)$$

Denote the eigenvalues of A_β by $\mu_{i\pm}$, for $i \in \{1, 2, \ldots, n\}$. As illustrated by Eq. (11.13) in [7], $\mu_{i\pm}$ can be obtained by performing eigenanalysis on A_β:

$$\mu_{i\pm} = \frac{\beta\lambda_i \pm \sqrt{\beta^2\lambda_i^2 - 4\beta + 4}}{2}. \quad (18)$$

The eigenvalues induced by $\lambda_n = -1$ are $\mu_{n\pm} = \{-1, \ 1 - \beta\}$. Thus, by Lemma 2, the model does not converge; selecting an $\mathbf{x}_0$ along the eigenvector corresponding to -1 leads to persistent oscillations. ∎

We next numerically compare convergence properties of the MLA model with the DeGroot model and the accelerated averaging model. To do so, we consider a periodic ring network with four nodes, shown in Fig. 2, and numerically simulate the system starting from 1000 randomly chosen initial conditions and plot the time series of the envelope of oscillations relative to the mean. It can be seen that for the chosen undirected, symmetric and periodic network, the DeGroot model as well as the accelerated averaging model oscillate whereas the MLA model reaches a consensus steady state.
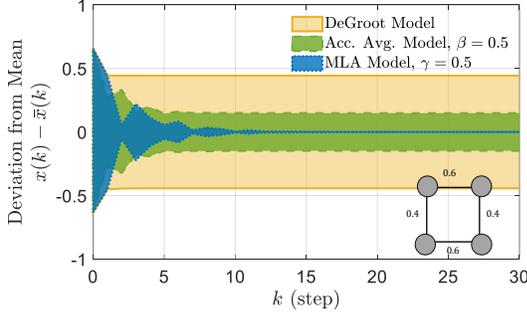
Fig. 2: Comparison of convergence in three averaging models for a periodic network. The plot shows the envelope enclosed between the maximum and minimum *deviation* from the mean state and is simulated using 1000 randomly chosen initial conditions. In contrast to others, the MLA model converges to consensus.

C. Value of convergence

Lemma 6: Under Assumptions 1 and 2, if the convergence criteria given in Theorem 3 are satisfied, the MLA model given by Eq. (2), with initial conditions $\hat{\mathbf{x}}(0) = [\mathbf{x}(0); \mathbf{x}(0)]$, converges to an average consensus value given by $\mathbb{1}_n^T \mathbf{x}(0)/n$.

Proof: The final value of the states, $\hat{\mathbf{x}}_{\text{final}}$, is,

$$\hat{\mathbf{x}}_{\text{final}} = \lim_{k\to\infty} \hat{\mathbf{x}}(k) = \lim_{k\to\infty} \hat{A}^k \hat{\mathbf{x}}(0). \quad (19)$$

Since A has a simple eigenvalue at $\lambda_1 = 1$ and $\hat{A}$ is semi-convergent, Eq. (5) implies that $\hat{\lambda}_1 = 1$ is a simple eigenvalue for $\hat{A}$ (see proof of Theorem 3). Using the Jordan decomposition, we can write,

$$\lim_{k\to\infty} \hat{A}^k = \lim_{k\to\infty} T \left[\begin{array}{c|c} 1 & \mathbb{0}_{1\times 2n-1} \\ \hline \mathbb{0}_{2n-1\times 1} & B_{2n-1\times 2n-1} \end{array} \right]^k T^{-1}, \quad (20)$$

where, $T = [\hat{\mathbf{v}}_{1\pm},\ldots,\hat{\mathbf{v}}_{n\pm}]$, $T^{-1} = [\hat{\mathbf{w}}_{1\pm}^T;\ldots;\hat{\mathbf{w}}_{n\pm}^T]$, and $\hat{\mathbf{v}}_{i\pm}, \hat{\mathbf{w}}_{i\pm}$ denote the right and left eigenvectors of $\hat{A}$ $\forall$ $i \in \{1,\ldots,n\}$, such that $\hat{\mathbf{w}}_{i\pm}^T \hat{\mathbf{v}}_{i\pm} = 1$, $\forall$ $i \in \{1,2,\ldots,n\}$. B consists of Jordan blocks corresponding to eigenvalues, $\hat{\lambda}_{i\pm}$, satisfying $|\hat{\lambda}_{i\pm}| < 1$, $\forall$ $i \in \{2,3\ldots,n\}$ and $|\lambda_{i-}| < 1$. Thus, we get,

$$\lim_{k\to\infty} \hat{A}^k = \lim_{k\to\infty} T \left[\begin{array}{c|c} 1 & \mathbb{0}_{1\times 2n-1} \\ \hline \mathbb{0}_{2n-1\times 1} & \mathbb{0}_{2n-1\times 2n-1} \end{array} \right] T^{-1}. \quad (21)$$

Combining Eqs. (19) and (21), we get:

$$\hat{\mathbf{x}}_{\text{final}} = \hat{\mathbf{v}}_{1+} \hat{\mathbf{w}}_{1+}^T \hat{\mathbf{x}}(0) = (\hat{\mathbf{w}}_{1+}^T \hat{\mathbf{x}}(0)) \hat{\mathbf{v}}_{1+}. \quad (22)$$

Since $\hat{A}$ is a row-stochastic matrix, $\hat{\mathbf{v}}_{1+} = \mathbb{1}_{2n}$, proving that the agents reach consensus. Define $\hat{\mathbf{w}}_{1+}^T = [\mathbf{w}_{1,1}^T, \mathbf{w}_{1,2}^T]$. Thus, we get the following eigenvalue problem:

$$\left[\mathbf{w}_{1,1}^T, \mathbf{w}_{1,2}^T\right] \left[\begin{array}{c|c} \gamma A & (1-\gamma)A \\ \hline \mathbb{I}_{n\times n} & \mathbb{0}_{n\times n} \end{array} \right] = 1 \left[\mathbf{w}_{1,1}^T, \mathbf{w}_{1,2}^T\right]. \quad (23)$$

Solving this system of equations yields:

$$\mathbf{w}_{1,1}^T A = \mathbf{w}_{1,1}^T. \quad (24)$$

Thus, $\mathbf{w}_{1,1}$ is a scaled dominant right eigenvector of A i.e. $\mathbf{w}_{1,1} = \zeta \mathbf{w}_1$ for some $\zeta \in \mathbb{R} \setminus \{0\}$. Combining this with

Eq. (23) yields $\mathbf{w}_{1,2} = (1-\gamma)\mathbf{w}_{1,1} = \zeta(1-\gamma)\mathbf{w}_1$. Additionally, by imposing the constraint $\hat{\mathbf{w}}_1^T \mathbb{1}_{2n\times 1} = 1$, we get:

$$\hat{\mathbf{w}}_1^T = \left[\left(\frac{1}{2-\gamma}\right) \mathbf{w}_1^T, \left(\frac{1-\gamma}{2-\gamma}\right) \mathbf{w}_1^T \right]. \quad (25)$$

Hence, the final consensus states attained by agents is:

$$\hat{\mathbf{x}}_{\text{final}} = \left[\left(\frac{1}{2-\gamma}\right) \mathbf{w}_1^T, \left(\frac{1-\gamma}{2-\gamma}\right) \mathbf{w}_1^T \right] \begin{bmatrix} \mathbf{x}(0) \\ \mathbf{x}(0) \end{bmatrix} \mathbb{1}_{2n} = \mathbf{w}_1^T \mathbf{x}(0) \mathbb{1}_{2n}. \quad (26)$$

Since A is symmetric, $\mathbf{w}_1 = (1/n)\mathbb{1}_n$ and MLA reaches a consensus to the average of the initial states of the agents. ∎

IV. AN ACCELERATED ROUTE TO CONSENSUS

In the previous section we showed that under certain conditions the MLA algorithm converges while the models by DeGroot and Muthukrishnan et al. do not. In this section we show networks and model parameters for which all three models converge but the MLA algorithm leads to faster convergence. To this end we consider a refinement of Assumption 2 that guarantees convergence for the DeGroot and the accelerated averaging models, as given below.

Assumption 3: The weighted adjacency matrix A is *symmetric* and *primitive*, that is $\exists$ $k \in \mathbb{N}$ such that A^k is positive.

The essential spectral radius of a row-stochastic matrix A, denoted by $\rho_{\text{ess}}(A)$, is defined as follows [7]

$$\rho_{\text{ess}}(A) \equiv \begin{cases} 0, & \text{if } \text{spec}(A) = \{1,1,\ldots,1\}, \\ \max\{|\lambda| \mid \lambda \in \text{spec}(A) \setminus \{1\}\}, & \text{otherwise.} \end{cases}$$

The essential spectral radius determines the rate of convergence for linear discrete models. The lower the essential spectral radius the higher the rate of convergence [7].

A. Comparison with the DeGroot model

Lemma 7: Under Assumptions 1 and 3, and if $\lambda_2 + \lambda_n \neq 0$, $\exists$ $\gamma \in \mathbb{R}$ such that $\rho_{ess}(\hat{A}) < \rho_{ess}(A)$. This implies that the MLA model given by Eq. (4) converges to consensus faster than the DeGroot model given by Eq. (1).

Proof: Due to Perron-Frobenius theorem for primitive matrices, (i) $\lambda_1 = 1$ is a simple eigenvalue of A, (ii) $|\lambda_i| < 1$ $\forall$ $i \in \{2,3,\ldots,n\}$, with the convention $1 = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n > -1$. For $\gamma = 1 + \Delta$ such that $\Delta \approx 0$, the discriminant in Eq. (6) is approximately equal to λ^2, and is non-negative. Thus, all eigenvalues of $\hat{A}$ are real i.e. $\hat{\lambda}_{i\pm} \in \mathbb{R}$ $\forall$ $i \in \{1,2,..,n\}$. Additionally, for $\gamma \approx 1$, the convergence criteria given by Theorem 3 are satisfied i.e. A is row-stochastic, symmetric, irreducible, $\gamma \in (0,2)$, and $2\gamma\lambda_n - \lambda_n + 1 > 0$. Thus $\hat{A}$ is semi-convergent and reaches consensus. From Eq. (6), we get for $\lambda \neq 0$,

$$\hat{\lambda}_{\pm} = \frac{(1+\Delta)\lambda \pm |\lambda|\sqrt{1+2\Delta+\Delta^2-\frac{4\Delta}{\lambda}}}{2}$$

Since $\Delta^2 \ll 1$, we ignore this term and use the binomial approximation to get,

$$\hat{\lambda}_{\pm} \approx \frac{(1+\Delta)\lambda \pm |\lambda|\left(1+\Delta-\frac{2\Delta}{\lambda}\right)}{2}. \quad (27)$$

We then have,

$$\hat{\lambda}_{i\pm} \approx \{\lambda_i + \Delta(\lambda_i - 1), \ \Delta\}. \tag{28}$$

Since $\lambda_2 + \lambda_n \neq 0$ there exists a unique eigenvalue λ_{ess} such that $\rho_{ess}(A) = |\lambda_{ess}| \in (0,1)$. From Eq. (28) for $\Delta \approx 0$, we have $\rho_{ess}(\hat{A}) = |\lambda_{ess} + \Delta(\lambda_{ess} - 1)|$.

$$\lambda_{ess} \in (0,1) \Rightarrow \ \exists \ \Delta > 0 \ \text{s.t.} \ 0 < \lambda_{ess} + \Delta(\lambda_{ess} - 1) < \lambda_{ess},$$
$$\lambda_{ess} \in (-1,0) \Rightarrow \ \exists \ \Delta < 0 \ \text{s.t.} \ 0 > \lambda_{ess} + \Delta(\lambda_{ess} - 1) > \lambda_{ess}.$$

Overall, $\exists \ \Delta \in \mathbb{R}$ s.t. $\rho_{\text{ess}}(\hat{A}) < \rho_{\text{ess}}(A)$. $\tag{29}$

■

The technical assumption $\lambda_2 + \lambda_n \neq 0$ is generic and besides pathological cases, holds for almost all networks.

B. Comparison with the accelerated averaging model

Theorem 8: Under Assumptions 1, 3, if $\lambda_n < 0$ and $\lambda_2 \leq |\lambda_n|/3$, for $\gamma^* = \frac{2}{\rho_{ess}(A)}\left(\sqrt{1 + \rho_{ess}(A)} - 1\right)$, the essential spectral radius of $\hat{A}$ in the MLA model given by Eq. (4) is $\rho_{\text{ess}}(\hat{A}) = \sqrt{1 + \rho_{\text{ess}}(A)} - 1$. In this case, the system converges to consensus faster than both the DeGroot model and the accelerated averaging model for any value of the parameter β.

Proof: We will prove this in two parts: first, we show that γ^* satisfies the convergence criteria in Theorem 3. Second, we compute the convergence rate of the MLA model.

Part 1: Under the given assumption, the essential spectral radius of A, denoted by $\rho_{\text{ess}}(A)$, is given by,

$$\rho_{\text{ess}}(A) = -\lambda_n \in (0,1). \tag{30}$$

We need to verify that $\gamma^* \in (0,2)$.

$$\sqrt{1 + \rho_{ess}(A)} - 1 > 0 \Rightarrow \gamma^* = \frac{2}{\rho_{ess}(A)}\left(\sqrt{1 + \rho_{ess}(A)} - 1\right) > 0 \tag{31}$$

Additionally,

$$\gamma^* = \frac{2}{\rho_{ess}(A)}\left(\sqrt{1 + \rho_{ess}(A)} - 1\right) < 1,$$
$$\Leftrightarrow 1 + \rho_{ess}(A) < 1 + \frac{\rho_{ess}(A)^2}{4} + \rho_{ess}(A),$$
$$\Leftrightarrow \rho_{ess}(A)^2 > 0. \tag{32}$$

Hence $\gamma^* \in (0,1) \subset (0,2)$

To show that γ^* satisfies Theorem 3 (ii) we start noting that $\gamma^* \in (0,1)$, as proven before and $\rho_{ess}(A) > 0$, hence

$$0 < \gamma^* < 1 \Rightarrow \ 2\rho_{ess}(A) > 2\rho_{ess}(A)(1 - \gamma^*) > 0,$$
$$\Rightarrow \rho_{ess}(A) + 1 > 2\rho_{ess}(A)(1 - \gamma^*) - \rho_{ess}(A) + 1 > \dots$$
$$\dots > -\rho_{ess}(A) + 1 > 0. \tag{33}$$

Thus, $\lambda_n = -\rho_{ess}(A) \Rightarrow 2\gamma^*\lambda_n - \lambda_n + 1 > 0$, as desired for convergence. Note that the value of γ^* is obtained by setting the discriminant in Eq. (6) to zero i.e. $D(\lambda_n, \gamma^*) = 0$. From this expression we also get,

$$\lambda_n = \frac{4(\gamma^* - 1)}{\gamma^{*2}}, \ \text{since} \ \lambda_n < 0. \tag{34}$$

An explanation for setting $D(\lambda_n, \gamma^*) = 0$ is given in the companion arXiv paper [16].

Part 2: For the convergence rate, we evaluate the essential spectral radius of $\hat{A}$, denoted by $\rho_{\text{ess}}(\hat{A})$, corresponding to $\gamma = \gamma^*$. Note that for $\gamma^* \in (0,1)$,

$$D(\lambda_n, \gamma^*) = 0 \Rightarrow \begin{cases} D(\lambda_i, \gamma^*) \leq 0 \ \forall \ \lambda_i \in [\lambda_n, 0], \\ D(\lambda_i, \gamma^*) > 0 \ \forall \ \lambda_i \in (0,1]. \end{cases} \tag{35}$$

We examine the ordering in modulus of eigenvalues of $\hat{A}$. For $\gamma = \gamma^*$, and for some a and b, consider two scenarios: (i) $1 \geq \lambda_a > \lambda_b > 0$ (ii) $\lambda_n \leq \lambda_a < \lambda_b \leq 0$. In (i), the expression for $\hat{\lambda}_+$, given in Eq. (6) is monotonically increasing which implies $\hat{\lambda}_{a+} > \hat{\lambda}_{b+} > 0$. In (ii), the modulus of complex roots of Eq. (5), is $|(\gamma^* - 1)\lambda|$ which implies $|\hat{\lambda}_{a\pm}| > |\hat{\lambda}_{b\pm}| > 0$. Thus, for the essential spectral radius of $\hat{A}$, we then have,

$$\rho_{\text{ess}}(\hat{A}) = \max\{|\hat{\lambda}_{1-}|, |\hat{\lambda}_{n\pm}|, |\hat{\lambda}_{2+}|\}, \tag{36}$$

where $|\hat{\lambda}_{1-}| = 1 - \gamma^*$, $|\hat{\lambda}_{n\pm}| = \sqrt{(\gamma^* - 1)\lambda_n} = -\gamma^*\lambda_n/2$ and $|\hat{\lambda}_{2+}| = \left(\gamma^*\lambda_2 + \sqrt{\gamma^{*2}\lambda_2^2 - 4(\gamma^* - 1)\lambda_2}\right)/2$. We show that $|\hat{\lambda}_{n\pm}| > |\hat{\lambda}_{1-}|$. In fact, for $\gamma^* \in (0,1)$ and using the value of λ_n from Eq. (34), we arrive at the following statements,

$$\sqrt{(\gamma^* - 1)\lambda_n} > |\gamma^* - 1|$$
$$\Leftrightarrow \sqrt{(\gamma^* - 1)\left(\frac{4(\gamma^* - 1)}{\gamma^{*2}}\right)} > |\gamma^* - 1| \Leftrightarrow -2 < \gamma^* < 2.$$

which clearly holds $\forall \ \gamma^* \in (0,1)$. Thus, $\rho_{\text{ess}}(\hat{A}) = \max\{|\hat{\lambda}_{n\pm}|, |\hat{\lambda}_{2+}|\}$. We next show that $\rho_{\text{ess}}(\hat{A}) = |\hat{\lambda}_{n\pm}|$ if and only if

$$|\hat{\lambda}_{n\pm}| \geq |\hat{\lambda}_{2+}|,$$
$$\Leftrightarrow -\frac{\gamma^*\lambda_n}{2} \geq \frac{\gamma^*\lambda_2 + \sqrt{\gamma^{*2}\lambda_2^2 - 4(\gamma^* - 1)\lambda_2}}{2},$$
$$\Leftrightarrow -\lambda_n - \lambda_2 \geq \sqrt{\lambda_2^2 - \lambda_n\lambda_2},$$
$$\Leftrightarrow c_1 : \lambda_n + \lambda_2 \leq 0 \ \text{and} \ c_2 : |\lambda_n + \lambda_2| \geq \sqrt{\lambda_2^2 - \lambda_n\lambda_2}. \tag{37}$$

Condition c_1 holds for $\lambda_n < 0$ and $|\lambda_n| \geq |\lambda_2|$. Condition c_2 in Eq. (37) can be further simplified as shown below:

$$c_2 \Leftrightarrow (\lambda_n + \lambda_2)^2 \geq \lambda_2^2 - \lambda_n\lambda_2,$$
$$\Leftrightarrow \lambda_n(\lambda_n + 3\lambda_2) \geq 0,$$
$$\Leftrightarrow \lambda_n \leq -3\lambda_2, \ \text{since} \ \lambda_n < 0,$$
$$\Leftrightarrow \lambda_2 \leq \frac{|\lambda_n|}{3}. \tag{38}$$

Overall, the expression for the spectral radius of $\hat{A}$ is then given by,

$$\rho_{\text{ess}}(\hat{A}) = |\hat{\lambda}_{n\pm}| = -\gamma^*\lambda_n/2. \tag{39}$$

Substituting Eq. (30) and γ^* from Theorem 8 in Eq. (39), we get:

$$\rho_{\text{ess}}(\hat{A}) = \sqrt{1 + \rho_{\text{ess}}(A)} - 1. \tag{40}$$

The optimized essential spectral radius of the accelerated averaging model given by Eq. (16) is given by [7],

$$\min_{\beta \in (0,2)} \rho_{\text{ess}}(A_\beta) = \rho_{\text{ess}}(A_{\beta^*}) = \frac{\rho_{\text{ess}}(A)}{1 + \sqrt{1 - \rho_{\text{ess}}^2(A)}} < \rho_{\text{ess}}(A),$$
(41)

for $\rho_{\text{ess}}(A) \in (0,1)$. Using further algebraic manipulations it can then be shown that, for $\rho_{\text{ess}}(A) \in (0,1)$, MLA converges faster than the accelerated averaging model and DeGroot since,

$$\rho_{\text{ess}}(\hat{A}) = \sqrt{1 + \rho_{\text{ess}}(A)} - 1 < \rho_{\text{ess}}(A_{\beta^*}) < \rho_{\text{ess}}(A). \quad (42)$$
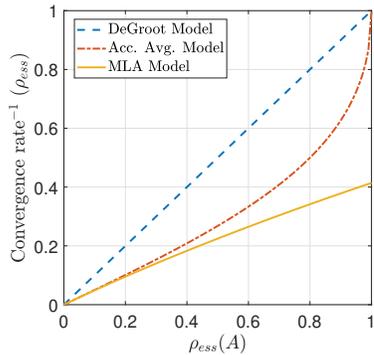
∎

Fig. 3: Comparison of essential spectral radii and the optimal rates of convergence of the three models of network dynamics; DeGroot model, accelerated averaging model, and the MLA model under the criteria given in Theorem 8. MLA model yields faster convergence.

Eq. (42) is plotted in Fig. 3. It can be noted that for a network that satisfies the criteria given in Theorem 8 with an essential spectral radius close to unity, the MLA model reaches consensus significantly faster than the classic DeGroot model and the accelerated averaging model. Good candidates for such matrices are perturbations of a periodic network. One such perturbation is shown in Fig. 4 which is obtained by adding self-loops of low weight in the 4-node ring network. The corresponding time series for the three models with optimized parameters are also shown.

V. CONCLUSION

We propose a simple modification to the accelerated averaging scheme introduced in [2]. In the proposed model (MLA), we apply the DeGroot update to the the current and past states of neighboring nodes followed by a linear combination step. The MLA model is applicable to networks that are symmetric, primitive and row-stochastic. We find the optimal model parameter, γ^*, for which MLA converges faster than both the DeGroot and accelerated averaging algorithms under certain network constraints. Another important contribution is that unlike the other two algorithms, MLA converges even for periodic networks. Future work could involve extending the results presented to a larger class of matrices such as asymmetric matrices.
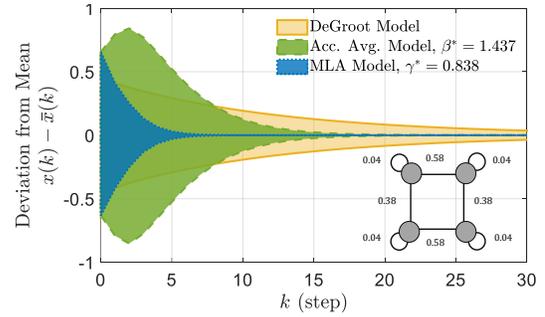

Fig. 4: Comparison of rates of convergence in three averaging models for a ring network with self-loops. The plot shows the envelope enclosed between the maximum and minimum *deviation* from the mean state and is simulated using 1000 randomly chosen initial conditions. MLA converges faster.

REFERENCES

[1] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[2] S. Muthukrishnan, B. Ghosh, and M. H. Schultz, "First-and second-order diffusive methods for rapid, coarse, distributed load balancing," *Theory of Computing Systems*, vol. 31, no. 4, pp. 331–354, 1998.

[3] D. Acemoğlu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Mathematics of Operations Research*, vol. 38, no. 1, pp. 1–27, 2013.

[4] N. E. Friedkin and E. C. Johnsen, "Social influence and opinions," *Journal of Mathematical Sociology*, vol. 15, no. 3-4, pp. 193–206, 1990.

[5] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents," *Advances in Complex Systems*, vol. 3, no. 01n04, pp. 87–98, 2000.

[6] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence models, analysis, and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, no. 3, 2002.

[7] F. Bullo, *Lectures on Network Systems*, 1st ed. Kindle Direct Publishing, 2020, with contributions by J. Cortes, F. Dorfler, and S. Martinez. [Online]. Available: http://motion.me.ucsb.edu/book-lns

[8] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[9] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[10] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *Journal of Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33–46, 2007.

[11] T. C. Aysal, B. N. Oreshkin, and M. J. Coates, "Accelerated distributed average consensus via localized node state prediction," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1563–1576, 2008.

[12] H. Wang, X. Liao, and T. Huang, "Accelerated consensus to accurate average in multi-agent networks via state prediction," *Nonlinear Dynamics*, vol. 73, no. 1-2, pp. 551–563, 2013.

[13] L. Wang and F. Xiao, "Finite-time consensus problems for networks of dynamic agents," *IEEE Transactions on Automatic Control*, vol. 55, no. 4, pp. 950–955, 2010.

[14] S. Li, H. Du, and X. Lin, "Finite-time consensus algorithm for multi-agent systems with double-integrator dynamics," *Automatica*, vol. 47, no. 8, pp. 1706–1712, 2011.

[15] Y. Cao and W. Ren, "Finite-time consensus for multi-agent networks with unknown inherent nonlinear dynamics," *Automatica*, vol. 50, no. 10, pp. 2648–2656, 2014.

[16] A. Bhaskar, S. Rangarajan, V. Shree, M. Campbell, and F. Parise, "Accelerated consensus in multi-agent networks via memory of local averages," *arXiv preprint arXiv:2109.12441*, 2021.

[17] W. Ren and Y. Cao, *Distributed coordination of multi-agent networks: emergent problems, models, and issues.* Springer Science & Business Media, 2010.