Distributed tracking of graph parameters in time-varying anonymous networks

D. Deplano, *Member, IEEE*, M. Franceschelli, *Member, IEEE*, and A. Giua, *Fellow, IEEE*

***Abstract*— In this paper we propose a distributed protocol for multi-agent systems to estimate and track changes to the diameter and radius of a time-varying network, as well as the eccentricity of each agent within it. The main strengths of the proposed protocol are its finite-time convergence and robustness to re-initialization, i.e., if there are changes in the network topology or in the agents' states during the protocol execution then it does not need to be re-initialized to converge to the correct estimation at the steady state. The expected error accuracy of the protocol can be traded-off by increasing the size of locally exchanged messages. We provide a theoretical characterization of the expected steady state error and some numerical simulations.**

I. Introduction

In networks of collaborating agents, the properties of their communication network topology are crucial. An effective model of the network is drawn by graph theory: each agent is modeled as a node and the interaction between two agents is modeled as an edge, thus constituting a graph. The properties of the graph modeling the communication network among the agents highly influences the behavior and the performance of almost any distributed algorithm executed by the network. Examples of distributed algorithms which estimate the spectrum and other graph properties such as controllability, observability and the Fiedler vector can be found in [2], [6], [7].

Among the many metrics that have been proposed to characterize the influence of the agents in the network [8], [17], one of particular significance is the *eccentricity* of a node, which is defined as the maximum distance to any other other node in the graph. This metric allows to easily define two important features of the graph, namely the *diameter* and *radius* of the network, which are formally equivalent to the maximum and minimum eccentricities among all nodes, respectively. Possible applications of such metrics are straightforward, such as selection of agents for maximizing the spread of influence in social networks [11], optimal coordination of cellular networks [13], maintaining a given efficiency in wireless networks [5], implementing a stopping criterion in distributed algorithms [12], and many others.

In the literature, much effort has been spent to go beyond centralized approaches [22] by focusing on the design of parallel [19] and distributed computation of these graph parameters, by means of flooding techniques [1], [18] and resorting to max-consensus protocols [9], [16]. The latter approach is particularly interesting since it allows to compute the desired parameters without the need of disclosing the identity of the agents within the networks, a framework known as *anonymous networks* [21]. However, these algorithms inherit the main drawback of max-consensus protocols: they can not be applied to time-varying networks because a change in the network or in the state of the agents during the algorithm execution requires to reinitialize the algorithm in the whole network, a feature not suitable to be implemented in large scale networks.

The main contribution of this paper is a novel distributed protocol for anonymous multi-agent networks which estimates and tracks the eccentricity of each node, diameter and radius of a graph modelling a time-varying network. The protocol exploits only anonymous local interactions among agents and is robust to re-initialization, thus it is suitable to be implemented in large time-varying networks. We characterize the steady-state estimation error and the algorithm convergence time.

The main novelty with respect to the current literature is the ability of the protocol to track changes in the graph parameters without the need to be re-initialized in the network.

This paper is organized as follows. In Section III we formulate the graph parameters tracking problem along with some working assumptions, then we present the proposed protocol and discuss it in plain words. In Section IV we characterize the steady state estimation error. Numerical simulations are presented in Section V and concluding remarks are given in Section VI.

II. Notation and preliminaries

We denote with $\mathbb{R}$ and $\mathbb{N}$ the sets of real and natural numbers respectively. Moreover, we denote with $\mathbb{R}_+$ and $\mathbb{N}_+$ their restriction to strictly positive numbers.

A. Multi-Agent systems

A multi-agent system (MAS) consists of a network of agents modeled as dynamical systems interacting among each other. The network is modeled by an undirected graph $\mathcal{G}(k) = (V, E(k))$ which represents the pattern of interactions among the agents at time $k \in \mathbb{N}$: $V \subset \mathbb{N}$ is the set of *nodes* modeling the agents and $E(k) \subseteq V \times V$ is the set of *edges* modeling interactions at time k between them. The total number of nodes in the network is constant and equal to $n = |V|$, where $|\cdot|$ denotes the cardinality of a set.

This work was partially supported by Region Sardinia (RAS) with project MOSIMA, RASSR05871, FSC 2014-2020, Annualità 2017, Area Tematica 3, Linea d'Azione 3.1. and by the Fondazione Banco di Sardegna with the grant "Formal Methods and Technologies for the Future of Energy Systems", cup F72F20000350007.

D. Deplano, M. Franceschelli and A. Giua are with DIEE, University of Cagliari, 09123 Cagliari, Italy. Emails: {diego.deplano,mauro.franceschelli,giua}@unica.it

A path between two nodes i and j in a graph is a sequence of consecutive edges $\pi_{ij} = (i,p),(p,q),\ldots,(r,s),(s,j)$ where each successive edge shares a node with its predecessor. An undirected graph $\mathcal{G}(k)$ is said to be *connected* if there exists a path π_{ij} between any pair of nodes $i,j \in V$. The *distance* between two nodes $i,j \in V$ at time k is denoted as $\text{dist}_{ij}(k)$ and it is defined as the length (number of edges) of the shortest path between nodes i and j.

The *eccentricity* of a node $i \in V$ at time k is denoted as $e_i(k)$ and it is defined as the maximal distance from i of any other node,

$$e_i(k) = \max_{j\in V} \text{dist}_{ij}(k).$$

The *diameter* of graph $\mathcal{G}(k)$ at time k is denoted as $d(k)$ and it is defined as the maximal eccentricity among the nodes,

$$d(k) = \max_{i\in V} e_i(k).$$

The *radius* of graph $\mathcal{G}(k)$ at time k is denoted as $r(k)$ and it is defined as the minimal eccentricity among the nodes,

$$r(k) = \min_{i\in V} e_i(k).$$

Agents i and j are said to be *neighbors* at time k if there exists an edge between i and j, i.e., $(i,j) \in E(k)$ or equivalently $(j,i) \in E(k)$. At any time k, the set of neighbors of the i-th agent is denoted as $\mathcal{N}_i(k) = \{j \in V : (i,j) \in E(k)\}$: it represents the agents in the graph sharing a point-to-point communication channel with agent i and interacting with it at time k. For sake of simplicity, we denote $\mathcal{N}_i^\circ(k) = \mathcal{N}_i(k) \cup \{i\}$. Similarly, the set of *h-hops neighbors* at time k is denoted as $\mathcal{N}_i^h(k)$ and it comprises the set of agents j which share a path π_{ij} between i and j and $\text{dist}_{ij} \equiv h$. Furthermore, $\mathcal{N}^h(k) = \{i \in V : e_i(k) = h\}$ denotes the set of all nodes with eccentricity equal to h.

B. Static and dynamic max-consensus protocols

Consider a MAS wherein each agent $i \in V$ with state $s_i(k) \in \mathbb{R}$ has access to a reference signals $v_i(k) \in \mathbb{R}$.

The max-consensus problem consists in the design of a local interaction rule enabling the agents' state to converge to the maximum among the reference signals. If the reference signals are assumed to be constant over time, i.e., $v(k) = v(0)$ for all $k \in \mathbb{N}_+$, the problem is solved by the *max-consensus protocol* protocol

$$s_i(k) = \max_{j\in\mathcal{N}_i^\circ}\{s_j(k-1)\}, \qquad s_i(0) = v_i(0), \tag{1}$$

which has been proved to converge in finite time and with zero error [10], [14], [15], [20]. On the other hand, if the reference signals are assumed to be time-varying, the problem can be approximately solved by the *dynamic max-consensus protocol* presented by Deplano et al. [3], [4],

$$s_i(k) = \max_{j\in\mathcal{N}_i^\circ(k-1)} \big\{s_j(k-1) - \alpha, v_i(k)\big\}. \tag{2}$$

which has been proved to converge in finite time and with bounded error.

III. Distributed tracking of graph parameters

A. Problem statement

We consider a network of n agents which synchronously gather state information from their neighbors and update their state at discrete instants of time. Time is divided in epochs, or iterations, indexed by the positive integer $k \in \mathbb{N}_+$.

At each iteration k the pattern of communication among the agents may change but the set of agents does not change. Thus, the network of agents can be effectively represented by a graph $\mathcal{G}(k) = (V, E(k))$ where V is the time-invariant set of nodes, representing the $n = |V|$ agents, and $E(k) \subseteq \{V \times V\}$ is the time-varying set of edges, representing the interactions among agents. We assume that there is a dwell time between two consecutive changes in the network topology, as stated in the next assumption.

Assumption 1 *There exists a minimum dwell time $\Upsilon \in \mathbb{N}_+$ between two consecutive changes of the graph $\mathcal{G}(k)$.*

In this paper we consider the problem of dynamically tracking the diameter $d(k)$, the radius $r(k)$ of the network, as well as the eccentricities $e_i(k)$ of the agents, which are time-varying parameters. We propose a local interaction rule that we call the EDR Protocol, given in the next page, to distributedly solve this problem within the framework of anonymous multi-agent networks, i.e., networks where information about the identity of the agents must be kept hidden.

B. Proposed estimation protocol

The proposed estimation methodology is detailed in the EDR Protocol, which envisages three operations/steps, described next. In the reminder of this section, we use the notation $a \to b$ to denote that a tracks the value b.

1) Initialization (lines $1-6$): The i-th agent selects $L \in \mathbb{N}$ random numbers $u_{i\ell}(0) \in [0,1]$ with $\ell = 1,\ldots,L$ with uniform distribution and initialize its state variables $x_{i\ell}, y_{i\ell}, z_{i\ell} \in \mathbb{R}$ according to

$$x_{i\ell}(0) = u_{i\ell}(0), \quad y_{i\ell}(0) = u_{i\ell}(0), \quad z_{i\ell}(0) = 0$$

2) Execution - Distances tracking (lines $8-16$): The strategy of this step makes use of an ingenious combination of static and dynamic max-consensus protocols in eq. (1)-(2). The following explanation has to be intended for each $\ell = 1,\ldots,L$.

Each variable $x_{i\ell} \in \mathbb{R}$ runs the max-consensus protocol in eq. (1) over the set $[u_{1\ell},\ldots,u_{n\ell}]$ and thus it tracks the maximum of the set with zero error,

$$x_{i\ell} \to \max_{i\in V} u_{i\ell} = u_{j_\ell^*\ell},$$

where

$$j_\ell^* = \operatorname*{argmax}_{j\in V} u_{j\ell}. \tag{3}$$

Assuming no quantization of the real numbers, the maximum number $u_{j_\ell^*\ell}$ is ensured to be unique with probability one by the continuity of the distribution. Thus, the update law in line 12 for $u_{i\ell}$ ensures that the signals will eventually be set to $-\infty$, unless $i = j^*$.

Each variable $y_{i\ell} \in \mathbb{R}$ runs the dynamic max-consensus protocol in eq. (2) over the set of numbers $[-\infty, \ldots, u_{j^*_\ell \ell}, \ldots, -\infty]$ and thus it tracks the maximum of the set up to an error, which is characterized in Lemma 2 in the next section,

$$y_{i\ell} \to u_{j^*\ell} - \alpha \cdot \mathrm{dist}_{ij^*}.$$

The agents i can now infer its distance $\mathrm{dist}_{ij^*_\ell}$ to the node j^*_ℓ attaining the maximum value by

$$\varepsilon_{i\ell} = \frac{|x_{i\ell} - y_{i\ell}|}{\alpha} \to \mathrm{dist}_{ij^*}.$$

Each variable $z_{i\ell} \in \mathbb{R}$ runs the dynamic max-consensus protocol in eq. (2) over the set of reference signals $[\varepsilon_{1\ell}, \ldots, \varepsilon_{n\ell}]$, whose maximum tracks the maximum distance from agent j^*_ℓ to all other agent in the network, in fact,

$$\max_{i\in V} \varepsilon_{i\ell} \to \max_{i\in V} \mathrm{dist}_{ij^*} = \mathrm{dist}_{i^*_\ell j^*_\ell} \tag{4}$$

with

$$i^*_\ell = \operatorname*{argmax}_{i\in V} \mathrm{dist}_{ij^*_\ell}. \tag{5}$$

Therefore, variable $z_{i\ell}$ tracks the value in eq. (4) up to an error, which is characterized in Lemma 2 in the next section,

$$z_{i\ell} \to \mathrm{dist}_{i^*_\ell j^*_\ell} - \alpha \cdot \mathrm{dist}_{ii^*_\ell}$$

In other words, the variable $z_{i\ell}$ tracks the maximum distance from any node in the network and the node attaining the maximum value, up to an error proportional to its distance to the node attaining such maximum distance.

3) Execution - Parameters inference (lines $17-20$): The i-th agent considers as an estimation of its eccentricity its maximum distance to all agents attaining a maximum,

$$\hat{e}_i = \max_{\ell=1,\ldots,L} \varepsilon_{i\ell}(k) \to \max_{\ell=1,\ldots,L} \mathrm{dist}_{ij^*_\ell},$$

Then, by assuming the parameter $\alpha \in \mathbb{R}_+$ to be small enough, it considers as an estimation of the network diameter the maximum distance of any agent in the network to all agents attaining a maximum,

$$\hat{d}_i = \max_{\ell=1,\ldots,L} \lceil z_{i\ell} \rceil \to \max_{\ell=1,\ldots,L} \mathrm{dist}_{i^*_\ell j^*_\ell}.$$

Finally, it considers as an estimation of the network radius the minimum distance of any agent in the network to all agents attaining a maximum,

$$\hat{r}_i = \min_{\ell=1,\ldots,L} \lceil z_{i\ell} \rceil \to \min_{\ell=1,\ldots,L} \mathrm{dist}_{i^*_\ell j^*_\ell}.$$

IV. Convergence analysis of the EDR Protocol

In this section we provide the characterization of the expected error on the estimation provided by the EDR Protocol. But first, we need two preliminary lemmas.

The first lemma concerns the specific steady state reached by a network running the dynamic max-consensus protocol in eq. (2) in the case of constant reference signals.

EDR Protocol : Distributed tracking of Eccentricities, Diameter, and Radius in time-varying networks

(Input): Tuning parameters $\alpha \in \mathbb{R}_+$ and $L \in \mathbb{N}$
(Output): $\hat{e}_i(k)$, $\hat{d}_i(k)$, $\hat{r}_i(k) \in \mathbb{R}$ for $i \in V$.

1 **(Initialization): for** $\ell = 1, \ldots, L$ **each node** i **does**
2 Select numbers with uniform distribution
3 $u_{i\ell}(0) \sim \mathcal{U}(0,1)$
4 Initialize state variables according to
5 $x_{i\ell}(0) = u_{i\ell}(0), \quad y_{i\ell}(0) = u_{i\ell}(0), \quad z_{i\ell}(0) = 0$
6 Send $x_{i\ell}(0)$, $y_{i\ell}(0)$, $z_{i\ell}(0)$ to its neighbors
7 **(Execution): for** $k = 1, 2, 3, \ldots$ **each node** i **does**
8 **for** $\ell = 1, \ldots, L$ **it does**
9 Gather $x_{j\ell}(k-1)$, $y_{j\ell}(k-1)$, $z_{j\ell}(k-1)$, from its neighbors $j \in \mathcal{N}_i(k-1)$
10 Update its state variables according to
11 $x_{i\ell}(k) = \max_{j\in\mathcal{N}^\circ_i(k-1)} \{x_{j\ell}(k-1)\}$
12 $u_{i\ell}(k) = \begin{cases} -\infty & \text{if } x_{i\ell}(k) > x_{i\ell}(k-1) \\ u_{i\ell}(k-1) & \text{otherwise} \end{cases}$
13 $y_{i\ell}(k) = \max_{j\in\mathcal{N}^\circ_i(k-1)} \{y_{j\ell}(k-1) - \alpha, u_{i\ell}(k)\}$
14 $\varepsilon_{i\ell}(k) = \dfrac{|x_{i\ell}(k) - y_{i\ell}(k)|}{\alpha}$
15 $z_{i\ell}(k) = \max_{j\in\mathcal{N}^\circ_i(k-1)} \{z_{j\ell}(k-1) - \alpha, \varepsilon_{i\ell}(k)\}$
16 Send $x_{i\ell}(k)$, $y_{i\ell}(k)$, $z_{i\ell}(k)$ to its neighbors $j \in \mathcal{N}_j(k)$
17 Estimate the graph parameters according to
18 $\hat{e}_i(k) = \max_{\ell=1,\ldots,L} \varepsilon_{i\ell}(k)$
19 $\hat{d}_i(k) = \max_{\ell=1,\ldots,L} \lceil z_{i\ell}(k) \rceil$
20 $\hat{r}_i(k) = \min_{\ell=1,\ldots,L} \lceil z_{i\ell}(k) \rceil$

Lemma 1 *Consider a network of n agents, each of which has access to a constant $v_i \in \mathbb{R}$ and updates its state $s_i(k)$ according to the dynamic max-consensus protocol in eq.* (2) *and let $k_0 \in \mathbb{N}$ be a generic instant of time.*

If graph $\mathcal{G}$ is time-invariant, then there exists a convercence time $T_c \in \mathbb{N}$ such that

$$T_c \le d + \max\left\{0, \left\lceil \max_{j\in V} \frac{s_j(k_0)}{\alpha} - \max_{j\in V} \frac{v_j}{\alpha} \right\rceil \right\} \tag{6}$$

so that each agent reaches an equilibrium state for $k \ge k_0 + T_c$ such that

$$s_i(k) \ge v_{j^*}(k) - \alpha \cdot \mathrm{dist}_{ij^*}, \tag{7}$$

where

$$j^* = \operatorname*{argmax}_{j\in V} v_j. \tag{8}$$

Moreover, if $v_{j^*} - v_i > \alpha \cdot \mathrm{dist}_{ij^*}$ holds for all $i \ne j^*$, then the inequality holds strictly.

Proof: Let $k_0 \in \mathbb{N}$ be a generic instant of time. By [3], [4, Theorem 1], there exists a time

$$T^* \leq \max\left\{0, \left\lceil \max_{j\in V} \frac{s_j(k_0)}{\alpha} - \max_{j\in V} \frac{v_j}{\alpha} \right\rceil \right\} \tag{9}$$

such that agent j^* as in eq. (8) satisfies

$$s_{j^*}(k) = v_{j^*}, \qquad k \geq k_0 + T_c. \tag{10}$$

and $s_{j^*}(k)$ is the maximum among all agents at time k,

$$s_{j^*}(k) \geq s_j(k), \qquad \forall j \in V. \tag{11}$$

At time $k^* = k_0 + T^*$ we define the set of one-hop neighbors of agent j^*, formally

$$\mathcal{V}_1 = \{i \in V : (i, j^*) \in E\}.$$

The dynamic max-consensus protocol in eq. (2) at k^*+1 for the agents belonging to this set reduces to

$$s_i(k^*+1) = \max\{v_{j^*} - \alpha, v_i\}, \qquad \forall i \in \mathcal{V}_1$$

because they have state as in eq. (11) and agent j^* as neighbor with state as in eq. (10). By induction, define the sets

$$\mathcal{V}_\ell = \left\{ i \in V : (i,j) \in E, j \in \bigcup_{s=0}^{\ell-1} \mathcal{V}_s \right\}, \quad \ell = 1, 2, \ldots.$$

It can be noticed that the parameter ℓ denotes the distance dist_{ij^*} of agent $i \in \mathcal{V}_\ell \setminus \mathcal{V}_{\ell-1}$ to agent j^*. Since the longest shortest path between two nodes in a connected graph is at most equal to its diameter d, then it holds that $\mathcal{V}_d = V$ and thus for $k \geq k^* + d$ it holds

$$s_i(k) = \max\{v_{j^*} - \alpha \cdot \text{dist}_{ij^*}, v_i\}, \qquad \forall i \in V.$$

This proves that for $k \geq k^* + d = k_0 + T^* + d$, i.e., after the convergence time $T_c = T^* + d$, which is therefore bounded as in eq. (6) due to eq. (9), the steady state reached by the network is lower bounded as in eq. (7) since the maximum of two values is greater or equal than one of them. But clearly if $v_i < v_{j^*} - \alpha \cdot \text{dist}_{ij^*}$ for all $i \neq j^*$ then it exactly holds $s_i(k) = v_{j^*}(k) - \alpha \cdot \text{dist}_{ij^*}(k)$. ■

The second lemma concerns the specific steady state reached by a network running the EDR Protocol. But first, we recall that for each $\ell = 1, \ldots, L$, the agents run two instances of the dynamic max-consensus protocol, respectively at variables $y_{i\ell}$ and $z_{i\ell}$ for $i \in \mathcal{V}$, and we denote with $T_c(y, \ell)$ and $T_c(z, \ell)$ their convergence time, according to Lemma 1. More precisely, given a generic initial instant of time $k_0 \in \mathbb{N}$, they are bounded by the following

$$T_c(y,\ell) \leq d + \max\left\{0, \max_{j\in V} \frac{y_{j\ell}(k_0)}{\alpha} - \max_{j\in V} \frac{u_{j\ell}(k_0)}{\alpha}\right\} \tag{12}$$

$$T_c(z,\ell) \leq d + \max\left\{0, \max_{j\in V} \frac{z_{j\ell}(k_0)}{\alpha} - \max_{j\in V} \frac{\varepsilon_{j\ell}(k_0)}{\alpha}\right\} \tag{13}$$

Lemma 2 *Consider a network which implements the EDR Protocol under Assumption 1 and let $k_0 \in \mathbb{N}$ be a generic instant of time. If the tuning parameter $\alpha \in \mathbb{R}_+$ satisfies*

$$\alpha < 1/d, \tag{14}$$

and if the dwell time Υ is greater than the convergence time T_c of all dynamic max-consensus protocols, i.e.,

$$\Upsilon \geq \max_{\ell=1,\ldots,L} \{T_c(y,\ell), T_c(z,\ell)\} = T_c', \tag{15}$$

with $T_c(y,\ell)$ and $T_c(z,\ell)$ given in eq. (12)-(13)*, respectively, then for $k \in [k_0 + T_c, k_0 + \Upsilon]$ it holds*

$$\hat{e}_i(k) = \max_{\ell=1,\ldots,L} \text{dist}_{ij_\ell^*}(k_0), \tag{16}$$

$$\hat{d}_i(k) = \max_{\ell=1,\ldots,L} \text{dist}_{i_\ell^* j_\ell^*}(k_0), \tag{17}$$

$$\hat{r}_i(k) = \min_{\ell=1,\ldots,L} \text{dist}_{i_\ell^* j_\ell^*}(k_0). \tag{18}$$

Proof: For each $\ell \in [0, L]$, each node $i \in V$ runs at variable $x_{i\ell}$ the popular (static) max-consensus protocol in eq. (1), and thus is a standard result (cfr. [20]) that

$$x_{i\ell}(k) = u_{j_\ell^*\ell}(k), \quad k \in [k_0 + d, k_0 + \Upsilon],$$

with j_ℓ^* as in eq. (3). For each $\ell \in [0, L]$, each node $i \in V$ runs at variable $y_{i\ell}$ the dynamic max-consensus protocol in eq. (2) with reference signals $u_{i\ell}$. By the update rule of $u_{i\ell}$ is straightforward to conclude that for $k \geq k_0 + d$ all numbers are eventually set to $-\infty$, unless the unique maximum[1], i.e.,

$$u_{j_\ell^*\ell}(k) = u_{j_\ell^*\ell}(k_0), \quad u_{i\ell}(k) = -\infty, \qquad \forall i \neq j_\ell^*.$$

Since the distance $u_{j_\ell^*\ell}(k) - u_{i\ell}(k)$ is infinite, then Lemma 1 holds strictly and for $k \in [k_0 + T_c(y,\ell), k_0 + \Upsilon]$ it follows

$$y_{i\ell}(k) = u_{j_\ell^*\ell}(k) - \alpha \cdot \text{dist}_{ij_\ell^*}(k) = x_{i\ell}(k) - \alpha \cdot \text{dist}_{ij_\ell^*}(k)$$

Therefore, the estimation of the eccentricity converges to

$$\begin{aligned}\hat{e}_i(k) &= \max_{\ell=1,\ldots,L} \varepsilon_{i\ell}(k) = \max_{\ell=1,\ldots,L} \frac{|y_{i\ell}(k) - x_{i\ell}(k)|}{\alpha} \\ &= \max_{\ell=1,\ldots,L} \text{dist}_{ij_\ell^*}(k), \quad \forall k \in [k_0 + T_c(y,\ell), k_0 + \Upsilon],\end{aligned}$$

thus proving the veracity of eq. (16). For each $\ell \in [0, L]$, each node $i \in V$ runs at variable $z_{i\ell}$ the dynamic max-consensus protocol in eq. (2). The variable $z_{i\ell}$ tracks the value

$$\max_{i\in V} \varepsilon_{i\ell}(k) = \max_{i\in V} \text{dist}_{ij_\ell^*}(k) = \text{dist}_{i_\ell^* j_\ell^*}(k)$$

with i_ℓ^* as in eq. (5). Thus, by Lemma 1, for $k \in [k_0 + \max\{T_c(y,\ell), T_c(z,\ell)\}, k_0 + \Upsilon]$ it holds

$$z_{i\ell}(k) \geq \text{dist}_{i_\ell^* j_\ell^*}(k) - \alpha \cdot \text{dist}_{ij_\ell^*}(k).$$

By eq. (14), it holds $\lceil -\alpha \cdot \text{dist}_{ij_\ell^*}(k_0) \rceil = 0$ and $\lceil z_{i\ell}(k) \rceil = \text{dist}_{i_\ell^* j_\ell^*}(k)$. Thus, the diameter estimation converges to

$$\hat{d}_i(k) = \max_{\ell=1,\ldots,L} \lceil z_{i\ell}(k) \rceil = \max_{\ell=1,\ldots,L} \text{dist}_{i_\ell^* j_\ell^*}(k)$$

and the radius estimation converges to

$$\hat{r}_i(k) = \min_{\ell=1,\ldots,L} \lceil z_{i\ell}(k) \rceil = \min_{\ell=1,\ldots,L} \text{dist}_{i_\ell^* j_\ell^*}(k),$$

thus proving eq. (17)-(18). The above considerations are true if $\Upsilon \geq \max\{T_c(y,\ell), T_c(z,\ell)\}$, completing the proof. ■

[1] Assuming no quantization of the real numbers, the maximum is unique with probability one, by the continuity of the distribution.

Next, we discuss the expected error at the steady state of the estimations of the eccentricities, diameter and radius between any two consecutive changes of the network topology. In other words, we consider time windows $[k_0, k_0 + \Upsilon]$ such that $\Upsilon \geq T'_c$ where $k_0 \in \mathbb{N}$ is a generic initial time at which the network changes, T'_c is the convergence time of the protocol in eq. (15) and Υ is the dwell time before a new change may occur. The parameters of interest are constant over these time windows and this allows to provide a precise characterization of the expected errors in the next theorem.

Theorem 1 *Consider the scenario of Lemma 2. Then, for all* $k \in [k_0 + T'_c, k_0 + \Upsilon]$ *the expected estimation errors of the EDR protocol are:*

$$\mathbb{E}[e_i(k) - \hat{e}_i(k)] = \sum_{\varepsilon=1}^{e_i} \left(1 - \frac{\sum_{h=\varepsilon}^{e_i} |\mathcal{N}_i^h(k)|}{|V|}\right)^L \tag{19}$$

$$\mathbb{E}[d(k) - \hat{d}_i(k)] = \sum_{\delta=r+1}^{d} \left(1 - \frac{\sum_{h=\delta}^{d} |\mathcal{N}^h(k)|}{|V|}\right)^L \tag{20}$$

$$\mathbb{E}[\hat{r}_i(k) - r(k)] = \sum_{\rho=r}^{d-1} \left(1 - \frac{\sum_{h=r}^{\rho} |\mathcal{N}^h(k)|}{|V|}\right)^L \tag{21}$$

Sketch of the Proof: We provide next a sketch of the proof for the expected error on the eccentricities. First of all, one needs to compute the probability that the estimated eccentricity greater or equal than a certain threshold $\varepsilon \in \{1, \ldots, e_i\}$, which is given by

$$\mathbb{P}[\hat{e}_i \geq \varepsilon] = 1 - \left(1 - \frac{|\mathcal{N}_i^{e_i}|}{|V|}\right)^L.$$

The expected error value can be calculated by multiplying each of the possible outcomes by the probability each outcome will occur and then summing all of those values,

$$\mathbb{E}[e_i - \hat{e}_i] = e_i - \mathbb{E}[\hat{e}_i] = e_i - \sum_{\varepsilon=1}^{e_i} \varepsilon \cdot \mathbb{P}[\hat{e}_i = \varepsilon].$$

The probability of the event $\hat{e}_i = \varepsilon$ can be computed by computing the probability of event $\hat{e}_i \geq \varepsilon$ and then subtracting the probability of event $\hat{e}_i \geq \varepsilon + 1$. This is possible since the the events $\hat{e}_i = \varepsilon$ for any $\varepsilon \in \{1, \ldots, e_i\}$ are independent, and thus one can write

$$\mathbb{E}[e_i - \hat{e}_i] = e_i - \sum_{\varepsilon=1}^{e_i} \varepsilon \cdot (\mathbb{P}[\hat{e}_i \geq \varepsilon] - \mathbb{P}[\hat{e}_i \geq \varepsilon + 1]),$$

from which eq. (19) follows. The other proofs follow the same strategy. Moreover, all proofs are discussed at length in the Appendix for the convenience of the reader. ■

The characterization of the expected errors provided in the previous theorem shows that the quality of the estimates heavily depends on the graph topology and on the choice of the parameter $L \in \mathbb{N}$. In fact, it can be noticed that not only for $L \to \infty$ the expected errors go to zero, but also that the decaying of each term is geometrical in L. We will further discuss this fact in the next section.

V. Simulation Results

In the first simulation we consider 10^3 random graphs having $n = 300$ nodes with diameter $d = 9$ and radius $r = 5$. For each graph, and for all choices of the parameter $L = 1, \ldots, n$ we compute the expected errors on the diameter and the radius according to eq. (20)-(21) and compute the average error made by executing the EDR Protocol for 10 times. The results of the simulation, which are given in Fig. 1, corroborates the theoretical results in the sense that the average actual error made by the the proposed protocol is exactly the expected error given by the characterization in Section IV.

As discussed in the previous section, the expected errors decay exponentially with L. In particular, while each term in the summation of eq. (20)-(21) decays exponentially, the convergence rate is determined by the largest addend in the summation, which by definition has always magnitude strictly less then one. It is also interesting to notice that it is possible to identify a threshold value, which in our case is $L^* \approx 50$, after which the decay has almost reached its final value.

Since the design of the parameter L trades-off memory burden and communication complexity with the estimation accuracy, a pragmatic choice of the parameter L could be around such pivot point $L^* \approx 50$ in the decaying curve, which ensures that both expected errors are strictly below 1.

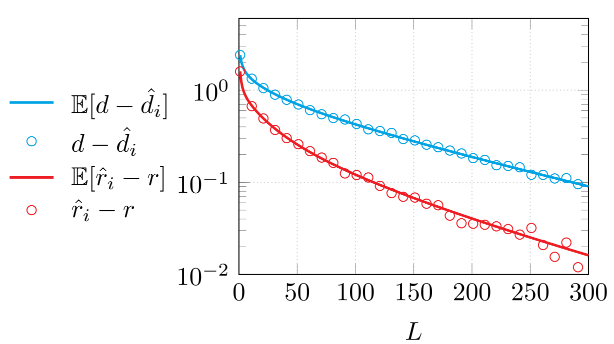

Fig. 1. Expected error and actual error for increasing values of L in a random network of $n = 300$ nodes.

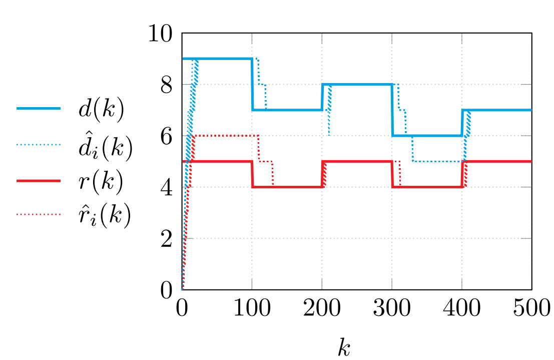

Fig. 2. Dynamic tracking of diameter and radius in a random network of $|V| = 300$ nodes.

As a second simulation we consider a random graph with the same number of $n = |V| = 300$ nodes but with time-varying topology, thus allowing fluctuations in the values of the radius and diameter. Changes in the network occur every $\Upsilon = 500$ iterations and the agents execute the EDR Protocol without the need of being re-initialized. Every time a change in the topology occurs, a transient behavior can be observed leading all agents to a new steady-state, thus validating the use of the proposed protocol in time-varying networks, theoretically characterized in terms of steady state in Lemma 2 and in terms of expected error in Theorem 1.

The estimations $\hat{d}_i(k)$, $\hat{r}_i(k)$ with the choice of $\alpha = 0.1$ and $L = 50$, the value suggested by the above analysis, are plotted in Fig. 2. It can be noticed that both the diameter and radius estimation show a transient behavior every time the network change its topology and then converge to a new steady state, failing the correct estimation only once. Moreover, it can be noticed that at the end of each time window of length $\Upsilon = 100$ the estimation error is not greater than 1, in accordance to the expected error analysis.

VI. Conclusion

In this work we have solved the problem of distributed tracking of important graph parameters in time-varying networks, namely the eccentricities of the nodes, the diameter and radius of the network. The proposed approach consists of a distributed estimation protocol which exploits static and dynamic max-consensus protocols as subroutines. The main advantages of the proposed method is the possibility to be implemented on time-varying networks without the need to reinitialize the algorithm after each change of the graph parameters, as it is required by other distributed algorithms at the current state of the art. Also, the proposed estimation protocol is developed within the framework of anonymous networks, i.e., each agent does not share its identity to its neighbors in the graph. As future work we aim to characterize and extend the current approach to open multi-agent networks, i.e., large scale networks where agents can also log in and out of the network.

References

[1] P. S. Almeida, C. Baquero, and A. Cunha, “Fast distributed computation of distances in networks,” in *IEEE 51st Conference on Decision and Control*, 2012, pp. 5215–5220.

[2] D. Deplano, M. Franceschelli, A. Giua, and L. Scardovi, “Distributed fiedler vector estimation with application to desynchronization of harmonic oscillator networks,” *IEEE Control Systems Letters*, vol. 5, no. 2, pp. 659–664, 2021.

[3] D. Deplano, M. Franceschelli, and A. Giua, “Dynamic max-consensus and size estimation of anonymous multi-agent networks,” *arXiv*, 2020.

[4] ——, “Dynamic min and max consensus and size estimation of anonymous multi-agent networks,” *IEEE Transaction on Automatic Control , under review*, 2021.

[5] L. M. Feeney, “Energy efficient communication in ad hoc wireless networks,” *Mobile Ad Hoc Networking*, pp. 301–327, 2004.

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

[7] M. Franceschelli, S. Martini, M. Egerstedt, A. Bicchi, and A. Giua, “Observability and controllability verification in multi-agent systems through decentralized laplacian spectrum estimation,” in *IEEE 49th Conference on Decision and Control*, 2010, pp. 5775–5780.

[8] L. C. Freeman, “Centrality in social networks conceptual clarification,” *Social networks*, vol. 1, no. 3, pp. 215–239, 1978.

[9] F. Garin, D. Varagnolo, and K. H. Johansson, “Distributed estimation of diameter, radius and eccentricities in anonymous networks,” *IFAC Proceedings Volumes*, vol. 45, no. 26, pp. 13–18, 2012.

[10] S. Giannini, A. Petitti, D. Di Paola, and A. Rizzo, “Asynchronous max-consensus protocol with time delays: Convergence results and applications,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 2, pp. 256–264, 2016.

[11] D. Kempe, J. Kleinberg, and É. Tardos, “Maximizing the spread of influence through a social network,” in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 2003, pp. 137–146.

[12] A. D. Kshemkalyani and M. Singhal, *Distributed computing: principles, algorithms, and systems*. Cambridge University Press, 2011.

[13] S.-J. Lee, E. M. Belding-Royer, and C. E. Perkins, “Scalability study of the ad hoc on-demand distance vector routing protocol,” *International journal of network management*, vol. 13, no. 2, pp. 97–114, 2003.

[14] G. Muniraju, C. Tepedelenlioglu, and A. Spanias, “Analysis and design of robust max consensus for wireless sensor networks,” *IEEE Transactions on Signal and Information Processing over Networks*, vol. 5, no. 4, pp. 779–791, 2019.

[15] B. M. Nejad, S. A. Attia, and J. Raisch, “Max-consensus in a max-plus algebraic setting: The case of fixed communication topologies,” in *2009 XXII International Symposium on Information, Communication and Automation Technologies*, 2009, pp. 1–7.

[16] G. Oliva, R. Setola, and C. N. Hadjicostis, “Distributed finite-time calculation of node eccentricities, graph radius and graph diameter,” *Systems & Control Letters*, vol. 92, pp. 20–27, 2016.

[17] T. Opsahl, F. Agneessens, and J. Skvoretz, “Node centrality in weighted networks: Generalizing degree and shortest paths,” *Social networks*, vol. 32, no. 3, pp. 245–251, 2010.

[18] D. Peleg, L. Roditty, and E. Tal, “Distributed algorithms for network diameter and girth,” in *International Colloquium on Automata, Languages, and Programming*, Springer, 2012, pp. 660–672.

[19] J. Shun, “An evaluation of parallel eccentricity estimation algorithms on undirected real-world graphs,” in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 1095–1104.

[20] A. Tahbaz-Salehi and A. Jadbabaie, “A one-parameter family of distributed consensus algorithms with boundary: From shortest paths to mean hitting times,” in *IEEE 45th Conference on Decision and Control*, 2006, pp. 4664–4669.

[21] M. Yamashita and T. Kameda, “Computing on an anonymous network,” in *7th ACM Symposium on Principles of distributed computing*, 1988, pp. 117–130.

[22] U. Zwick, “Exact and approximate distances in graphs—a survey,” in *European Symposium on Algorithms*, Springer, 2001, pp. 33–48.

APPENDIX

For the convenience of the reader, in this section we append the detailed proof of Theorem 1. It will not appear in the final version of this paper.

A. Expected estimation error of node eccentricities

We now prove that

$$\mathbb{E}[e_i(k) - \hat{e}_i(k)] = \sum_{\varepsilon=1}^{e_i} \left(1 - \frac{\sum_{h=\varepsilon}^{e_i} |\mathcal{N}_i^h(k)|}{|V|}\right)^L,$$

for $k \in [k_0 + T_c, k_0 + \Upsilon]$ and omitting in the following the time dependence (k). By Lemma 2, it holds

$$\hat{e}_i = \max_\ell \operatorname{dist}_{ij_\ell^*},$$

with j_ℓ^* given in eq. (3). Thus, $\hat{e}_i$ is at least $\varepsilon \in \{1, \ldots, e_i\}$ iff there is $\ell \in \{1, \ldots, L\}$ such that $\operatorname{dist}_{ij_\ell^*} \geq \varepsilon$, i.e.,

$$\mathbb{P}[\hat{e}_i \geq \varepsilon] = \mathbb{P}[\exists \ell : \operatorname{dist}_{ij_\ell^*} \geq \varepsilon].$$

The complementary event is that for all $\ell \in \{1, \ldots, L\}$ it holds that $\operatorname{dist}_{ij_\ell^*} < \varepsilon$, and therefore one can compute

$$\mathbb{P}[\hat{e}_i \geq \varepsilon] = 1 - \mathbb{P}[\forall \ell : \operatorname{dist}_{ij_\ell^*} < \varepsilon].$$

Since all events $\operatorname{dist}_{ij_\ell^*} < \varepsilon$ for $\ell \in \{1, \ldots, L\}$ are independent from each other, then their joint probability equals the product of their probabilities. Moreover, noting that for any given $\ell \in \{1, \ldots, L\}$ the probability of event $\operatorname{dist}_{ij_\ell^*} < \varepsilon$ is equal to the probability of event $\operatorname{dist}_{ij} < \varepsilon$ for all $j \in V$, one can write

$$\mathbb{P}[\hat{e}_i \geq \varepsilon] = 1 - \left(\mathbb{P}[\forall j : \operatorname{dist}_{ij} < \varepsilon]\right)^L.$$

By exploiting again the complementary event we write

$$\mathbb{P}[\hat{e}_i \geq \varepsilon] = 1 - \left(1 - \mathbb{P}[\exists j : \operatorname{dist}_{ij} \geq \varepsilon]\right)^L.$$

Finally, the probability that there exists $j \in V$ such that $\operatorname{dist}_{ij} \geq \varepsilon$ can be computed by taking the ratio between number of favourite outcomes, which is the sum of all nodes with distance to node i greater or equal than ε, and the total number of nodes, i.e.,

$$\mathbb{P}[\hat{e}_i \geq \varepsilon] = 1 - \left(1 - \frac{|\mathcal{N}_i^{e_i}|}{|V|}\right)^L. \tag{22}$$

By means of the previous result, we are going to compute the expected error, namely

$$\mathbb{E}[e_i - \hat{e}_i] = e_i - \mathbb{E}[\hat{e}_i].$$

In statistics and probability analysis, the expected value is calculated by multiplying each of the possible outcomes by the probability each outcome will occur and then summing all of those values, i.e.,

$$\mathbb{E}[e_i - \hat{e}_i] = e_i - \sum_{\varepsilon=1}^{e_i} \varepsilon \cdot \mathbb{P}[\hat{e}_i = \varepsilon].$$

It is easy to realize that the probability of the event $\hat{e}_i = \varepsilon$ can be computed by first computing the probability of event $\hat{e}_i \geq \varepsilon$ and then subtracting the probability of event $\hat{e}_i \geq \varepsilon + 1$. This is possible since the the events $\hat{e}_i = \varepsilon$ for any $\varepsilon \in \{1, \ldots, e_i\}$ are independent, and thus we can write

$$\mathbb{E}[e_i - \hat{e}_i] = e_i - \sum_{\varepsilon=1}^{e_i} \varepsilon \cdot \left(\mathbb{P}[\hat{e}_i \geq \varepsilon] - \mathbb{P}[\hat{e}_i \geq \varepsilon + 1]\right)$$

By some manipulation we compute

$$\begin{aligned}
\mathbb{E}[e_i - \hat{e}_i] &= e_i - \sum_{\varepsilon=1}^{e_i} \varepsilon \cdot \mathbb{P}[\hat{e}_i \geq \varepsilon] + \sum_{\varepsilon=1}^{e_i} \varepsilon \cdot \mathbb{P}[\hat{e}_i \geq \varepsilon + 1] \\
&= e_i - \sum_{\varepsilon=1}^{e_i} \varepsilon \cdot \mathbb{P}[\hat{e}_i \geq \varepsilon] + \sum_{\varepsilon=2}^{e_i} (\varepsilon - 1) \cdot \mathbb{P}[\hat{e}_i \geq \varepsilon] \\
&= e_i - \sum_{\varepsilon=1}^{e_i} \mathbb{P}[\hat{e}_i \geq \varepsilon].
\end{aligned}$$

By exploiting eq. (22) it follows

$$\begin{aligned}
\mathbb{E}[e_i - \hat{e}_i] &= e_i - \sum_{\varepsilon=1}^{e_i} \left(1 - \left(1 - \frac{\sum_{h=\varepsilon}^{e_i} |\mathcal{N}_i^h|}{|V|}\right)^L\right) \\
&= \sum_{\varepsilon=1}^{e_i} \left(1 - \frac{\sum_{h=\varepsilon}^{e_i} |\mathcal{N}_i^h|}{|V|}\right)^L,
\end{aligned}$$

completing the proof.

B. Expected estimation error of network diameter

We now prove that

$$\mathbb{E}[d(k) - \hat{d}_i(k)] = \sum_{\delta=r+1}^{d} \left(1 - \frac{\sum_{h=\delta}^{d} |\mathcal{N}^h(k)|}{|V|}\right)^L,$$

for $k \in [k_0 + T_c, k_0 + \Upsilon]$ and omitting in the following the time dependence (k). By Lemma 2, since $\alpha \leq d^{-1}$, it follows

$$\hat{d}_i = \max_\ell \operatorname{dist}_{i_\ell^* j_\ell^*}, \qquad \forall i \in V.$$

with j_ℓ^* and i_ℓ^* given in eq. (3)-(5), respectively. Thus, $\hat{d}_i$ is at least $\delta \in \{r, \ldots, d\}$ iff there is $\ell \in \{1, \ldots, L\}$ such that $\operatorname{dist}_{i_\ell^* j_\ell^*} \geq \delta$, i.e.,

$$\mathbb{P}[\hat{d}_i \geq \delta] = \mathbb{P}[\exists \ell : \operatorname{dist}_{i_\ell^* j_\ell^*} \geq \delta].$$

The complementary event is that for all $\ell \in \{1, \ldots, L\}$ it holds that $\operatorname{dist}_{i_\ell^* j_\ell^*} < \varepsilon$, and therefore one can compute

$$\mathbb{P}[\hat{d}_i \geq \delta] = 1 - \mathbb{P}[\forall \ell : \operatorname{dist}_{i_\ell^* j_\ell^*} < \delta]$$

Since all events $\operatorname{dist}_{i_\ell^* j_\ell^*} < \delta$ for $\ell \in \{1, \ldots, L\}$ are independent from each other, then their joint probability equals the product of their probabilities. Moreover, noting that for any given $\ell \in \{1, \ldots, L\}$ the probability of event $\operatorname{dist}_{i_\ell^* j_\ell^*} < \delta$ is equal to the probability of event $\operatorname{dist}_{ij} < \delta$ for all couples $p, q \in V$, one can write

$$\mathbb{P}[\hat{d}_i \geq \delta] = 1 - \left(\mathbb{P}[\forall (p,q) : \operatorname{dist}_{pq} < \delta]\right)^L.$$

By exploiting again the complementary event we write

$$\mathbb{P}[\hat{d}_i \geq \delta] = 1 - \left(1 - \mathbb{P}[\exists (p,q) : \operatorname{dist}_{pq} \geq \delta]\right)^L.$$

Finally, the probability that there exists a couple of nodes $p, q \in V$ such that $\text{dist}_{pq} \geq \delta$ can be computed by taking the ratio between number of favourite outcomes, which is the sum of all nodes with eccentricity greater or equal than δ, and the total number of nodes, i.e.,

$$\mathbb{P}[\hat{d}_i \geq \delta] = 1 - \left(1 - \frac{\sum_{h=\delta}^{d} |\mathcal{N}^h|}{|V|}\right)^L. \tag{23}$$

By means of the previous result, we are going to compute the expected error, namely

$$\mathbb{E}[d - \hat{d}_i] = d - \mathbb{E}[\hat{d}_i].$$

We use the same proof strategy as for the eccentricities, which follows

$$\begin{aligned}
\mathbb{E}[d - \hat{d}_i] &= d - \sum_{\delta=r}^{d} \delta \cdot \mathbb{P}[\hat{d}_i = \delta] \\
&= d - \sum_{\delta=r}^{d} \delta \cdot \left(\mathbb{P}[\hat{d}_i \geq \delta] - \mathbb{P}[\hat{d}_i \geq \delta + 1]\right) \\
&= d - \sum_{\delta=r}^{d} \delta \cdot \mathbb{P}[\hat{d}_i \geq \delta] + \sum_{\delta=r+1}^{d} (\delta - 1) \cdot \mathbb{P}[\hat{d}_i \geq \delta] \\
&= d - r - \sum_{\delta=r+1}^{d} \mathbb{P}[\hat{d}_i \geq \delta].
\end{aligned}$$

By exploiting eq. (23) it follows

$$\begin{aligned}
\mathbb{E}[d - \hat{d}_i] &= d - r - \sum_{\delta=r+1}^{d} \left(1 - \left(1 - \frac{\sum_{h=\delta}^{d} |\mathcal{N}^h|}{|V|}\right)^L\right) \\
&= \sum_{\delta=r+1}^{d} \left(1 - \frac{\sum_{h=\delta}^{d} |\mathcal{N}^h|}{|V|}\right)^L,
\end{aligned}$$

completing the proof.

C. Expected estimation error on the network radius

We now prove that

$$\mathbb{E}[\hat{r}_i(k) - r(k)] = \sum_{\rho=r}^{d-1} \left(1 - \frac{\sum_{h=r}^{\rho} |\mathcal{N}^h(k)|}{|V|}\right)^L,$$

for $k \in [k_0 + T_c, k_0 + \Upsilon]$ and omitting in the following the time dependence (k). By Lemma 2, since $\alpha \leq d^{-1}$ it follows

$$\hat{r}_i = \lceil \hat{r}_i \rceil = \min_{\ell} \text{dist}_{i_\ell^* j_\ell^*}.$$

with j_ℓ^* and i_ℓ^* given in eq. (3)-(5), respectively. Thus, $\hat{r}_i$ is at most $\rho \in \{r, \ldots, d\}$ iff there is $\ell \in \{1, \ldots, L\}$ such that $\text{dist}_{i_\ell^* j_\ell^*} \geq \rho$, i.e.,

$$\mathbb{P}[\hat{r}_i \leq \rho] = \mathbb{P}[\exists \ell : \text{dist}_{i_\ell^* j_\ell^*} \leq \rho].$$

The complementary event is that for all $\ell \in \{1, \ldots, L\}$ it holds that $\text{dist}_{i_\ell^* j_\ell^*} > \varepsilon$, and therefore one can compute

$$\mathbb{P}[\hat{r}_i \leq \rho] = 1 - \mathbb{P}[\forall \ell : \text{dist}_{i_\ell^* j_\ell^*} > \rho]$$

Since all events $\text{dist}_{i_\ell^* j_\ell^*} < \rho$ for $\ell \in \{1, \ldots, L\}$ are independent from each other, then their joint probability equals the product of their probabilities. Moreover, noting that for any given $\ell \in \{1, \ldots, L\}$ the probability of event $\text{dist}_{i_\ell^* j_\ell^*} > \rho$ is equal to the probability of event $\text{dist}_{ij} > \rho$ for all couples $p, q \in V$, one can write

$$\mathbb{P}[\hat{r}_i \geq \rho] = 1 - \left(\mathbb{P}[\forall (p, q) : \text{dist}_{pq} > \rho]\right)^L.$$

By exploiting again the complementary event we write

$$\mathbb{P}[\hat{r}_i \leq \rho] = 1 - \left(1 - \mathbb{P}[\exists (p, q) : \text{dist}_{pq} \leq \rho]\right)^L.$$

Finally, the probability that there exists a couple of nodes $p, q \in V$ such that $\text{dist}_{ij} \leq \rho$ can be computed by taking the ratio between number of favourite outcomes, which is the sum of all nodes with eccentricity lesser or equal than ρ, and the total number of nodes, i.e.,

$$\mathbb{P}[\hat{r}_i \leq \rho] = 1 - \left(1 - \frac{\sum_{h=r}^{\rho} |\mathcal{N}^h|}{|V|}\right)^L. \tag{24}$$

By means of the previous result, we are going to compute the expected error, namely

$$\mathbb{E}[\hat{r}_i - r] = \mathbb{E}[\hat{r}_i] - r.$$

We use the same proof strategy as for the eccentricities, thus it holds

$$\begin{aligned}
\mathbb{E}[\hat{r}_i - r] &= -r + \sum_{\rho=r}^{d} \rho \cdot \mathbb{P}[\hat{r}_i = \rho] \\
&= -r + \sum_{\rho=r}^{d} \rho \cdot \left(\mathbb{P}[\hat{r}_i \leq \rho] - \mathbb{P}[\hat{r}_i \leq \rho - 1]\right) \\
&= -r + \sum_{\rho=r+1}^{d+1} (\rho - 1) \cdot \mathbb{P}[\hat{r}_i \leq \rho - 1] \\
&\quad - \sum_{\rho=r}^{d} \rho \cdot \mathbb{P}[\hat{r}_i \leq \rho - 1] \\
&= -r + d \cdot \mathbb{P}[\hat{r}_i \leq d] + \sum_{\rho=r+1}^{d} \rho \cdot \mathbb{P}[\hat{r}_i \leq \rho - 1] \\
&\quad - \sum_{\rho=r+1}^{d} \mathbb{P}[\hat{r}_i \leq \rho - 1] - \sum_{\rho=r+1}^{d} \rho \cdot \mathbb{P}[\hat{r}_i \leq \rho - 1] \\
&\quad - r \cdot \mathbb{P}[\hat{r}_i \leq r - 1] \\
&= d - r - \sum_{\rho=r}^{d-1} \mathbb{P}[\hat{r}_i \leq \rho].
\end{aligned}$$

By exploiting eq. (24) it follows

$$\begin{aligned}
\mathbb{E}[\hat{r}_i - r] &= d - r - \sum_{\rho=r}^{d-1} \left(1 - \left(1 - \frac{\sum_{h=r}^{\rho} |\mathcal{N}^h|}{|V|}\right)^L\right) \\
&= \sum_{\rho=r}^{d-1} \left(1 - \frac{\sum_{h=r}^{\rho} |\mathcal{N}^h|}{|V|}\right)^L,
\end{aligned}$$

completing the proof.