



Lower-Bound Study for Function Computation in Distributed Networks via Vertex-Eccentricity

H. K. Dai¹ · M. Toulouse²

Received: 17 April 2019 / Accepted: 23 May 2019 / Published online: 11 July 2019
© Springer Nature Singapore Pte Ltd 2019

Abstract

Distributed computing network-systems are modeled as graphs with vertices representing compute elements and adjacency-edges capturing their uni- or bi-directional communication. Distributed function computation has a wide spectrum of major applications in distributed systems. Distributed computation over a network-system proceeds in a sequence of time-steps in which vertices update and/or exchange their values based on the underlying algorithm constrained by the time-(in)variant network-topology. For finite convergence of distributed information dissemination and function computation in the model, we present a lower bound on the number of time-steps for vertices to receive (initial) vertex-values of all vertices regardless of underlying protocol or algorithmics in time-invariant networks via the notion of vertex-eccentricity in a graph-theoretic framework. We also address lower bounds on vertex-eccentricity and its maximum version in terms of common graph-parameters such as maximum degree, and order and size.

Keywords Distributed function computation · Linear iterative schemes · Information dissemination · Finite convergence · Vertex-eccentricity

Preliminaries

Distributed computation algorithms, decentralized data-fusion architectures, and multi-agent systems are modeled with a network of interconnected vertices that compute common value(s) based on initial values or observations at the vertices. Key computation and communication requirements for these network/system paradigms include that their vertices perform local/internal computations and regularly communicate with each other via an underlying protocol. Fundamental limitations and capabilities of these algorithms and systems are studied in the literature with viable applications

in computer science, communication, and control and optimization (see, for examples, [1, 4, 5, 11, 12]). We give brief and informal descriptions of some example studies below:

1. Quantized consensus [6]: Consider an order- n network with an initial network-state in which each vertex assumes an initial (integer) value $x_i[0]$ for $i = 1, 2, \dots, n$. The network achieves a quantized consensus when, at some later time, all the n vertices simultaneously arrive with almost equal values y_i for $i = 1, 2, \dots, n$ (that is, $|y_i - y_j| \leq 1$ for all $i, j \in \{1, 2, \dots, n\}$) while preserving the sum of all initial values (that is, $\sum_{i=1}^n x_i[0] = \sum_{i=1}^n y_i$).
2. Collaborative distributed hypothesis testing [7]: Consider a network-system of n vertices (sensors/agents) that collaboratively determine the probability measure of a random variable based on a number of available observations/measurements. For the binary setting in deciding two hypotheses, each vertex collects measurement(s) and makes a preliminary (local) decision $d_i \in \{0, 1\}$ in favor of the two hypotheses for $i = 1, 2, \dots, n$. The n vertices are allowed to communicate, and the network-system resolves with a final decision by, for example, the

This article is part of the topical collection “Future Data and Security Engineering” guest edited by Tran Khanh Dang.

✉ H. K. Dai
dai@cs.okstate.edu
M. Toulouse
michel.toulouse@vgu.edu.vn

¹ Computer Science Department, Oklahoma State University, Stillwater, OK 74078, USA

² Computer Science Department, Vietnamese-German University, Binh Duong New City, Vietnam

- majority rule (that is, computes the indicator function of the event $\sum_{i=1}^n d_i > \frac{n}{2}$) in distributed fashion.
- Solitude verification [4]: Consider an unlabeled network of n vertices (processes) in which each vertex is in one of a finite number of states: s_i for $i = 1, 2, \dots, n$. Solitude verification on the network checks if a unique vertex with a given state s exists in the network, that is, computes the Boolean function for the equality $|\{i \in \{1, 2, \dots, n\} \mid s_i = s\}| = 1$.
 - Fundamental iterative limits of distributed function computation [11, 18]: Consider a generic distributed information processing system to attain collective goals via iterative or non-iterative inter-vertex communication. Lower and upper bounds on numbers of iterations for achieving finite convergence of distributed information dissemination and function computation are studied via: structural-controllability and -observability theories [11] and information-theoretic techniques [18] in deterministic and probabilistic settings that capture initial-value/input distribution, network-topological and communication constraints, and/or estimation/output performance.

While there is a wide spectrum of algorithms in the literature that solve distributed computation problems such as the above, there are also studies that deal with algorithmic and complexity issues constrained by underlying time-(in)variant network-topology, resource-limitations associated with vertices, time/space and communication tradeoffs, convergence criteria and requirements, etc. We present below a model of distributed computing systems and address the motivation of our study.

Model of Distributed Computing Systems

Most graph-theoretic definitions in this article are given in [2]. We will abbreviate “directed graph” and “directed path” to digraph and dipath, respectively.

We consider the topological model and algorithmics detailed in [11] for distributed function computation, and provide its abstraction components as follows:

- Network-topology: A distributed computing system is modeled as a digraph G with $V(G)$ and $E(G)$ denoting its sets of vertices and directed edges, respectively. Uni-directional communication on $V(G)$ is captured by the adjacency relation represented by $E(G)$: for all distinct vertices, $u, v \in V(G)$, $(u, v) \in E(G)$ if and only if vertex u can send information to vertex v (and v can receive information from u). Note that bi-directional communication between u and v is viewed as the co-existence of the two directed edge (u, v) and (v, u) in $E(G)$.

Distributed computation over the network proceeds in a sequence of time-steps. At each time-step, all vertices

update and/or exchange their values based on the underlying algorithm constrained by the network-topology, which is assumed to be time-invariant.

- Resource capabilities in vertices: The digraph G of the network-topology is vertex-labeled such that messages are identified with senders and receivers. The vertices of $V(G)$ are assumed to have sufficient computational capabilities and local storage. Generally, we assume that: (1) all communications/transmissions between vertices are reliable and in correct sequence, and (2) each vertex may, in the current time-step, receive the prior-step transmission(s) from its in-neighbor(s), update, and send transmission(s) to its out-neighbor(s) in accordance to the underlying algorithm.

The domain of all initial/input and observed/output values of the vertices of G is assumed to be an algebraic field \mathbb{F} .

- Linear iterative scheme (for algorithmic lower- and upper-bound results): For a vertex $v \in V(G)$, denote by $x_v[k] \in \mathbb{F}$ the vertex-value of v at time-step $k = 0, 1, \dots$. A function with domain $\mathbb{F}^{|V(G)|}$ and codomain \mathbb{F} is computed in accordance to a linear iterative scheme. Given initial vertex-values $x_v[0] \in \mathbb{F}$ for all vertices $v \in V(G)$ as arguments to the function, at each time-step $k = 0, 1, \dots$, each vertex $v \in V(G)$ updates (and transmits) its vertex-value via a weighted linear combination of the prior-step vertex-values constrained by neighbor-structures: for all $v \in V(G)$ and $k = 0, 1, \dots$,

$$x_v[k+1] = \sum_{u \in V(G)} w_{vu} x_u[k],$$

where the prescribed weights $w_{vu} \in \mathbb{F}$ for all $v, u \in V(G)$ that are subject to the adjacency-constraints $w_{vu} = 0$ (the zero-element of \mathbb{F}) if u is not adjacent to v (that is, $(u, v) \notin E(G)$); equivalently,

$$\begin{aligned} & \text{transpose of } (x_v[k+1] \mid v \in V(G)) \\ &= W \cdot \text{transpose of } (x_v[k] \mid v \in V(G)) \end{aligned}$$

where the two vectors of vertex-values and W are indexed by a common discrete ordering of $V(G)$ with $W = [w_{vu}]_{(v,u) \in V(G) \times V(G)}$.

Motivation of Our Study

Based on the framework and its variants for distributed function computation, researches and studies are focused on mathematical interplays among:

- Time-(in)variance of network-topology.
- Granularity of time-step: discrete versus continuous.
- Choice of base field: special (real or complexes) versus arbitrary (finite or infinite).

- Characterization of calculable functions.
- Convergence criteria and rates (finite, asymptotic, and/or probabilistic).
- Adoption and algebraic properties of weight-matrix for linear interactive schemes: random weight-matrix, spectrum of eigenvalues, base field, etc.
- Resilience and robustness of computation algorithmics for network-topology in the presence/absence of malicious vertices.
- Lower and upper bounds on (linear) iteration required for the convergence of calculable functions.

Summarized results, research studies, and references are available in, for examples, [11–14, 16, 17].

Sundaram and Hadjicostis [11, 12] present their research findings in the finite convergence of distributed information dissemination and function computation in the model with linear iterative algorithmics stated above, among other contributions in distributed function computation and data-stream transmission in the presences of noise and malicious vertices. More specifically, (1) they employ structural theories in observability and invertibility of linear systems over arbitrary finite fields to obtain lower and upper bounds on the number of linear iterations for achieving network-consensus for finite convergence of arbitrary functions, and (2) the bounds are valid for all initial vertex-values of arbitrary finite fields as arguments to the functions in connected time-invariant topologies with almost all random weight-matrices.

For a time-invariant topology with underlying digraph G and a vertex $u \in V(G)$, denote by $\text{deg}_{G,\text{in}}(u)$ the in-degree of u in G , and by $\Gamma_{G,\text{in}}(u)$ the in-neighbor of u in G ; hence $\Gamma_{G,\text{in}}^*(u)$ denotes the in-closure of u in G , that is,

$$\begin{aligned} \Gamma_{G,\text{in}}^*(u) &= \cup_{\eta \geq 0} \Gamma_{G,\text{in}}^\eta(u) \\ &= \{v \in V(G) \mid \text{there exists a dipath in } G \text{ from } v \text{ to } u\}. \end{aligned}$$

Consider all possible families of directed trees that are: (1) a vertex-decomposition of $\Gamma_{G,\text{in}}^*(u) - \{u\}$, and (2) rooted in (as subset of) $\Gamma_{G,\text{in}}(u)$. Denote by:

$$\begin{aligned} \alpha_{G,u} &= \min \{ \max \{ \text{order}(T_i) \mid 1 \leq i \leq n \} \mid \\ &\quad \{T_i\}_{i=1}^n \text{ is a family of directed trees that are:} \\ &\quad (1) \text{ a vertex-decomposition of } \Gamma_{G,\text{in}}^*(u) - \{u\}, \text{ and} \\ &\quad (2) \text{ rooted in (as subset of) } \Gamma_{G,\text{in}}(u) \}. \end{aligned}$$

Their upper-bound result for a vertex $u \in V(G)$ is stated as follows: for every linear iterative scheme with random weight-matrix over a finite base field \mathbb{F} of cardinality $|\mathbb{F}| \geq (\alpha_{G,u} - 1)(|\Gamma_{G,\text{in}}^*(u)| - \text{deg}_{G,\text{in}}(u) - \frac{1}{2}\alpha_{G,u})$, then, with probability at least $1 - \frac{1}{|\mathbb{F}|}(\alpha_{G,u} - 1)(|\Gamma_{G,\text{in}}^*(u)| - \text{deg}_{G,\text{in}}(u) - \frac{1}{2}\alpha_{G,u})$, the vertex u can calculate arbitrary functions of arbitrary initial vertex-values $x_v[0] \in \mathbb{F}$ for all $v \in \Gamma_{G,\text{in}}^*(u)$

via the linear iterative scheme within at most $\alpha_{G,u}$ time-steps.

Sundaram conjectures in [11] that $\alpha_{G,u}$ may also serve as a lower bound on the number of time-steps for a vertex $u \in V(G)$ to receive the initial vertex-values of all $v \in \Gamma_{G,\text{in}}^*(u)$ regardless of underlying protocol or algorithmics. Hence, linear iterative schemes are time-optimal in disseminating information over arbitrary time-invariant connected networks.

Toulouse and Minh [15] refute the conjecture via the notion of rank-step sequences for linear iterative schemes over connected network with an explicit counter-example in Fig. 1.

In accordance with the min–max formulation of $\alpha_{G,u}$, a direct argument justifying the counter-example is provided as follows:

1. We view the annotated graph G in Fig. 1 in which each edge represents the co-existence of its two directed versions. For the vertex $u \in V(G)$, the number of time-steps to receive the initial vertex-values of all $v \in \Gamma_{G,\text{in}}^*(u)$ ($= V(G)$), regardless of underlying protocol or algorithmics, is $\text{order}(S) + 2$ —which is realized by the dipath composed of the serial component S and vertices v_5, v_4 , and u .
2. We show below that $\alpha_{G,u} = \text{order}(S) + 3$, hence can not serve as a lower bound on the number of time-steps mentioned in item 1 above—as suggested in the conjecture [11].

To show that $\alpha_{G,u} \leq \text{order}(S) + 3$, we consider the following family of directed trees, $\{\hat{T}_i\}_{i=1}^3$, which are a vertex-decomposition of $\Gamma_{G,\text{in}}^*(u) - \{u\}$ and are rooted in (as subset of) $\Gamma_{G,\text{in}}(u)$: \hat{T}_1 is composed of the serial component S and vertices v_5, v_7 , and v_6 with $\text{order}(\hat{T}_1) = \text{order}(S) + 3$, and \hat{T}_2 and \hat{T}_3 saturate the remaining vertices in the parallel component P and vertices v_3, v_2, v_1 , and v_4 in an almost equipotent manner with $\text{order}(\hat{T}_2) = \lfloor \frac{\text{order}(P)+4}{2} \rfloor$ and $\text{order}(\hat{T}_3) = \lceil \frac{\text{order}(P)+4}{2} \rceil$. Since the two components P and S are equipotent, we have:

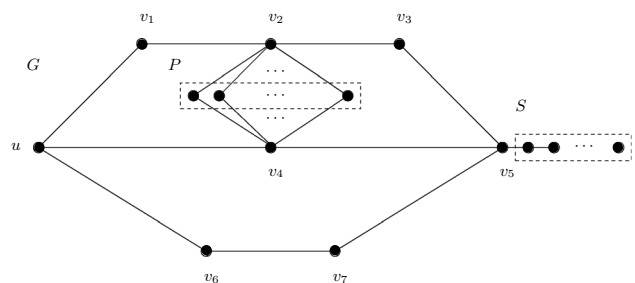


Fig. 1 A counter-example graph G , in which the embedded parallel component P and serial component S satisfying $\text{order}(P) = \text{order}(S)$, to the lower-bound conjecture in terms of $\alpha_{G,u}$ in [11]

$$\begin{aligned} \text{order}(\hat{T}_1) &= \text{order}(S) + 3 \\ &> \left\lceil \frac{\text{order}(P) + 4}{2} \right\rceil = \max\{\text{order}(\hat{T}_2), \text{order}(\hat{T}_3)\}. \end{aligned}$$

Therefore,

$$\begin{aligned} \alpha_{G,u} &\leq \max\{\text{order}(\hat{T}_i) \mid 1 \leq i \leq 3\} \\ &= \text{order}(\hat{T}_1) = \text{order}(S) + 3. \end{aligned}$$

To show the reverse inequality that $\alpha_{G,u} \geq \text{order}(S) + 3$, consider an arbitrary family of n (where $n \leq 3$) ordered trees that are a vertex-decomposition of $\Gamma_{G,\text{in}}^*(u) - \{u\}$ and are rooted in (as subset of) $\Gamma_{G,\text{in}}(u)$.

Case when $n = 1$: The single directed tree that saturates all the vertices of $\Gamma_{G,\text{in}}^*(u) - \{u\}$ is of order of:

$$\begin{aligned} \text{order}(P) + \text{order}(S) + 7 &> \text{order}(\hat{T}_1) (= \text{order}(S) + 3), \\ &\text{as desired.} \end{aligned}$$

Case when $n = 2$: Among the two directed trees that saturate all the vertices of $\Gamma_{G,\text{in}}^*(u) - \{u\}$, the larger-order one is of order bounded below by their average order of $\frac{\text{order}(P) + \text{order}(S) + 7}{2}$, therefore, by:

$$\begin{aligned} \left\lceil \frac{\text{order}(P) + \text{order}(S) + 7}{2} \right\rceil &= \left\lceil \frac{2 \text{order}(S) + 7}{2} \right\rceil \\ &= \text{order}(S) + 4 \\ &> \text{order}(\hat{T}_1) (= \text{order}(S) + 3), \end{aligned}$$

as desired.

Case when $n = 3$: Consider the three subcases of saturation of vertices v_6, v_7 , and v_5 (hence the serial component S).

Subcase when a directed tree consists of v_6 only: We proceed as in the case of $n = 2$. Among the two other directed trees that saturate the remaining vertices in the two components P and S , and vertices v_1, v_2, v_3, v_4, v_5 , and v_7 , the larger-order one is of order bounded below by:

$$\begin{aligned} \left\lceil \frac{\text{order}(P) + \text{order}(S) + 6}{2} \right\rceil &= \left\lceil \frac{2 \text{order}(S) + 6}{2} \right\rceil \\ &= \text{order}(S) + 3 = \text{order}(\hat{T}_1), \end{aligned}$$

as desired.

Subcase when a directed tree consists of v_6 and v_7 only: Similar to the previous subcase, the average order of the two other directed trees is $\frac{\text{order}(P) + \text{order}(S) + 5}{2}$, which yields a desired lower bound of:

$$\begin{aligned} \left\lceil \frac{\text{order}(P) + \text{order}(S) + 5}{2} \right\rceil &= \left\lceil \frac{2 \text{order}(S) + 5}{2} \right\rceil \\ &= \text{order}(S) + 3. \end{aligned}$$

Subcase when a directed tree saturates v_6, v_7 , and v_5 (hence, the serial component S): Such a directed tree is of order bounded below by:

$$\begin{aligned} \text{order}(S) + 3 &= \text{order}(\hat{T}_1), \\ &\text{as desired.} \end{aligned}$$

Combining the three cases of $n \in \{1, 2, 3\}$, we conclude that $\alpha_{G,u} \geq \text{order}(S) + 3$. Hence, the two inequalities on $\alpha_{G,u}$ yield that $\alpha_{G,u} = \text{order}(S) + 3$, which confirms the counter-example graph G in Fig. 1.

To complement the explicitly constructed counter-example to the lower-bound conjecture on the number of time-steps for distributed function computation and information dissemination with respect to a given vertex, we present in this article a lower bound on the number of time-steps for a vertex $u \in V(G)$ to receive the initial vertex-values of all $v \in \Gamma_{G,\text{in}}^*(u)$ regardless of underlying protocol or algorithmics in a time-invariant network via the notion of vertex-eccentricity.

Revised Lower Bound for Distributed Function Computation and Information Dissemination

Consider an arbitrary vertex $u \in V(G)$, and assume a non-trivial $\Gamma_{G,\text{in}}^*(u)$ ($|\Gamma_{G,\text{in}}^*(u)| > 1$) hereinafter. We develop a lower bound on the number of time-steps required for the vertex u to receive the (initial) vertex-values of all vertices of $\Gamma_{G,\text{in}}^*(u)$ (regardless of underlying protocol, including linear iterative schemes). See Fig. 2 for an example of $\Gamma_{G,\text{in}}^*(u)$.

For two vertices u and v of G , $\vec{d}_G(u, v)$ denotes the directed distance from u to v in G , that is,

$$\vec{d}_G(u, v) = \begin{cases} \text{length of a shortest dipath from } u \text{ to } v \text{ in } G & \text{if exists,} \\ \infty & \text{otherwise.} \end{cases}$$

For a vertex u of G , $e_{G,\text{in}}(u)$ denotes the in-eccentricity of u in G , which is the maximum directed distance from a vertex to u in G , that is,

$$e_{G,\text{in}}(u) = \max\left\{ \underbrace{\vec{d}_G(v, u)}_{\text{minimum length of a dipath from } v \text{ to } u \text{ in } G} \mid v \in V(G) \right\}.$$

Following the above-stated distributed computation framework as in [12] and for their conjecture, we develop a lower-bound result based on the notion of eccentricity (instead of ‘‘order’’ or ‘‘size’’ as in the conjecture):

1. For every (linear or non-linear) iteration scheme, in which a vertex’s value or information is transmitted to its out-neighbors via their incidence directed edges in unit time-step, requires at least $e_{G,\text{in}}(u)$ time-steps for vertex u to access values/information of all the vertices

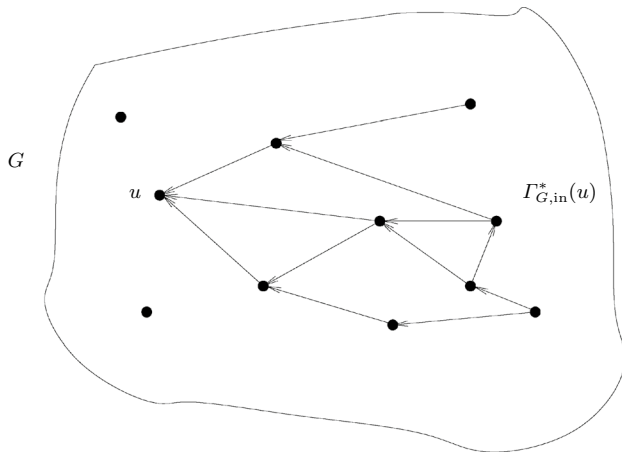


Fig. 2 For a vertex u in a digraph G : an example organization of the in-closure $\Gamma_{G,in}^*(u)$ of u in G

in $\Gamma_{G,in}^*(u)$. Thus, $e_{G,in}(u)$ serves as a lower bound on the number of time-steps required for function computation by vertex u via such iteration scheme.

- In accordance with the distributed framework for our function computation, we show below that:

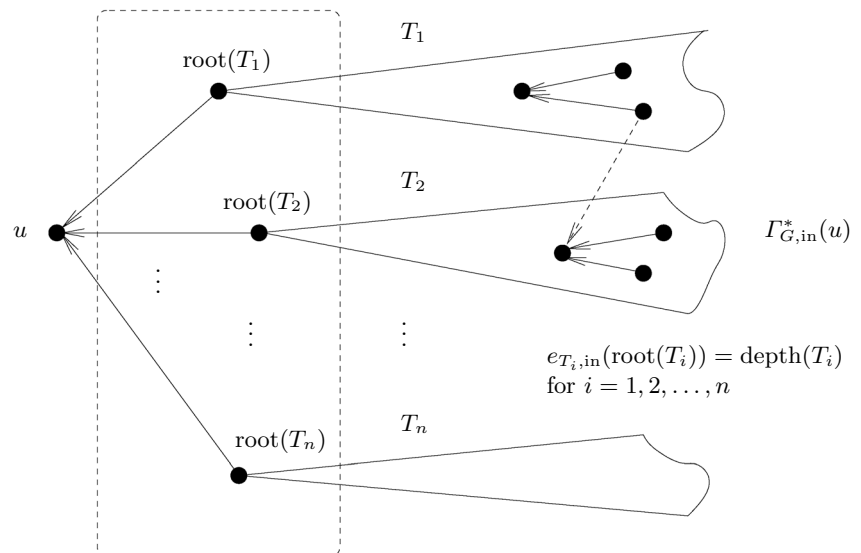
$$e_{G,in}(u) = 1 + \min\{\underbrace{\max\{e_{T_i,in}(\text{root}(T_i)) \mid 1 \leq i \leq n\}}_{=\text{depth}(T_i)}\}$$

$\{T_i\}_{i=1}^n$ is a family of directed trees that are:

- a vertex-decomposition of $\Gamma_{G,in}^*(u) - \{u\}$, and
- rooted in (as subset of) $\Gamma_{G,in}(u)$.

We illustrate an example organization of $\Gamma_{G,in}^*(u) - \{u\}$ in a family of vertex-disjoint directed trees in Fig. 3.

Fig. 3 For a vertex u in a digraph G : an example organization of $\Gamma_{G,in}^*(u) - \{u\}$ in a family $\{T_i\}_{i=1}^n$ of directed trees that are: (1) a vertex-decomposition of $\Gamma_{G,in}^*(u) - \{u\}$, and (2) rooted in (as subset of) $\Gamma_{G,in}(u)$



$\{\text{root}(T_i) \mid 1 \leq i \leq n\}$ is a subset (not necessarily proper) of $\Gamma_{G,in}^*(u)$

To show the above equality for $e_{G,in}(u)$, we prove the two embedded inequalities in the following sections.

Upper Bound for Vertex-Eccentricity

We first prove that:

$$e_{G,in}(u) \leq 1 + \min\{\underbrace{\max\{e_{T_i,in}(\text{root}(T_i)) \mid 1 \leq i \leq n\}}_{=\text{depth}(T_i)}\}$$

$\{T_i\}_{i=1}^n$ is a family of directed trees that are:

- a vertex-decomposition of $\Gamma_{G,in}^*(u) - \{u\}$, and
- rooted in (as subset of) $\Gamma_{G,in}(u)$;

equivalently,

$$e_{G,in}(u) \leq 1 + \underbrace{\max\{e_{T_i,in}(\text{root}(T_i)) \mid 1 \leq i \leq n\}}_{=\text{depth}(T_i)}$$

for arbitrary family of directed trees, $\{T_i\}_{i=1}^n$, which are a vertex-decomposition of $\Gamma_{G,in}^*(u) - \{u\}$ and are rooted in (as subset of) $\Gamma_{G,in}(u)$.

Consider an arbitrary family of directed trees, $\{T_i\}_{i=1}^n$, which are a vertex-decomposition of $\Gamma_{G,in}^*(u) - \{u\}$ and are rooted in (as subset of) $\Gamma_{G,in}(u)$. The in-eccentricity $e_{G,in}(u)$ of u in G is realized by a dipath P from a vertex $v \in \Gamma_{G,in}^*(u) - \{u\}$ to u in G . Since $\{T_i\}_{i=1}^n$ is a vertex-decomposition of $\Gamma_{G,in}^*(u) - \{u\}$, we have $v \in V(T_i)$ for some $i \in \{1, 2, \dots, n\}$. We depict the scenario in Fig. 4.

Now,

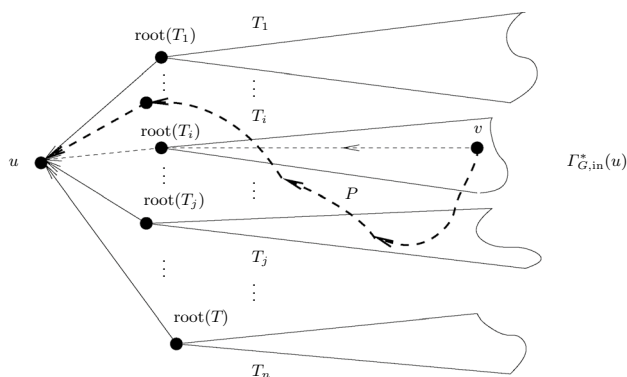


Fig. 4 For a vertex u in a digraph G : the in-eccentricity $e_{G,\text{in}}(u)$ of u in G is realized by a dipath P from a vertex $v \in \Gamma_{G,\text{in}}^*(u) - \{u\}$ to u in G

$$\begin{aligned}
 \underbrace{e_{G,\text{in}}(u)}_{\text{in } G} &= \underbrace{\text{length}(P)}_{\text{in } G} = \underbrace{\vec{d}_G(v, u)}_{\text{in } G} \\
 &\leq \text{length}(\text{unique dipath from } v \text{ to } \text{root}(T_i) \text{ in } T_i \text{ concatenated} \\
 &\quad \text{with directed edge } (\text{root}(T_i), u) \text{ since } T_i \text{ is a sub-digraph of} \\
 &\quad \text{the digraph vertex-spanned by } \Gamma_{G,\text{in}}^*(u) \\
 &\leq \text{depth}(T_i) + 1 \\
 &\leq 1 + \max\{\text{depth}(T_i) \mid 1 \leq i \leq n\}
 \end{aligned}$$

as desired.

Lower Bound for Vertex-Eccentricity

To show the reverse inequality:

$$\begin{aligned}
 e_{G,\text{in}}(u) &\geq 1 + \min\{\underbrace{\max\{e_{T_i,\text{in}}(\text{root}(T_i)) \mid 1 \leq i \leq n\}}_{=\text{depth}(T_i)}\} \\
 \{T_i\}_{i=1}^n &\text{ is a family of directed trees that are:} \\
 (1) &\text{ a vertex-decomposition of } \Gamma_{G,\text{in}}^*(u) - \{u\}, \text{ and} \\
 (2) &\text{ rooted in (as subset of) } \Gamma_{G,\text{in}}(u),
 \end{aligned}$$

it suffices to construct a family $\{T_i\}_{i=1}^n$ of directed trees that are a vertex-decomposition of $\Gamma_{G,\text{in}}^*(u) - \{u\}$, and are rooted in (as subset of) $\Gamma_{G,\text{in}}(u)$, such that:

$$e_{G,\text{in}}(u) \geq 1 + \max\{\text{depth}(T_i) \mid 1 \leq i \leq n\}.$$

We proceed with an inductive construction of a sequence (P_1, P_2, \dots) of dipaths with common terminal vertices u such that the sequence $(P_1 - \{u\}, P_2 - \{u\}, \dots)$ is organized as a family $\{T_1, T_2, \dots, T_i\}$, where $i \geq 1$, of directed trees such that:

1. The family $\{T_1, T_2, \dots, T_i\}$ consists of mutually vertex-disjoint directed trees with their roots in $\Gamma_{G,\text{in}}(u)$,

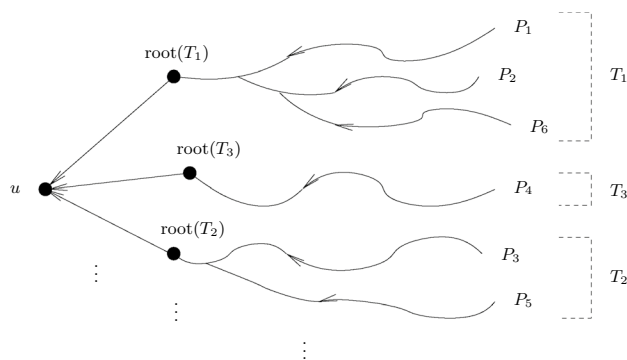


Fig. 5 For a vertex u in a digraph G : an inductive construction of a sequence (P_1, P_2, \dots) of dipaths with common terminal vertices u such that the sequence $(P_1 - \{u\}, P_2 - \{u\}, \dots)$ is organized as a family $\{T_1, T_2, \dots, T_i\}$, where $i \geq 1$, of directed trees that satisfies the stated conditions in items 1, 2, and 3

2. Each directed tree in the family provides a shortest dipath (in G) for each of its vertices to u , that is, for every vertex $v \in V(T_j)$ where $j \in \{1, 2, \dots, i\}$, the (unique) dipath from v to $\text{root}(T_j)$ in T_j yields $\vec{d}_G(v, u)$:

$$\text{length}(\text{unique dipath from } v \text{ to } \text{root}(T_j) \text{ concatenated with directed edge } (\text{root}(T_j), u)) = \vec{d}_G(v, u),$$

and

3. The in-eccentricity of u in G is bounded below as:

$$e_{G,\text{in}}(u) \geq 1 + \max\{\text{depth}(T_1), \text{depth}(T_2), \dots, \text{depth}(T_i)\}.$$

See an example configuration in Fig. 5.

Basis step: For P_1 , we may employ a dipath from a vertex, say v , in $\Gamma_{G,\text{in}}^*(u) - \{u\}$ to u in G that realizes $\vec{d}_G(v, u)$. Then, designate P_1 as such a path, and $T_1 = \{P_1 - \{u\}\}$.

For the family $\{T_1\}$, we can verify the above-stated three items 1, 2 (via “shortest dipath in G ” enjoys “optimal substructure property in G ” by typical cut-and-paste argument), and 3.

Induction step: Assume that we have constructed a sequence (P_1, P_2, \dots, P_j) of dipaths with common terminal vertices u such that the sequence $(P_1 - \{u\}, P_2 - \{u\}, \dots, P_j - \{u\})$ is organized as a family $\{T_1, T_2, \dots, T_i\}$, where $j \geq i \geq 1$, of directed trees that satisfies the above-stated items 1, 2, and 3.

If the family $\{T_1, T_2, \dots, T_i\}$ yields a vertex-decomposition of $\Gamma_{G,\text{in}}^*(u) - \{u\}$, then the inductive construction is complete. Thus, we may assume that there exists a vertex $v \in (\Gamma_{G,\text{in}}^*(u) - \{u\}) - \cup_{\eta=1}^i V(T_\eta)$. We construct a desired dipath P_{j+1} from v to u in G as follows.

First, consider a dipath P from v to u in G that realizes $\vec{d}_G(v, u)$ (that is, $\text{length}(P) = \vec{d}_G(v, u)$). Observe that,

$$\text{length}(P) = \vec{d}_G(v, u) \leq \underbrace{\max\{\vec{d}_G(v, u) \mid v \in V(G)\}}_{e_{G,\text{in}}(u)}.$$

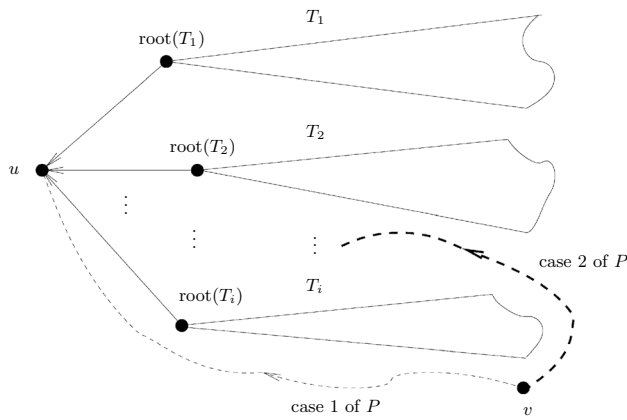


Fig. 6 For a vertex u in a digraph G : assume the inductive construction of a sequence (P_1, P_2, \dots, P_j) of dipaths that results in a family $\{T_1, T_2, \dots, T_i\}$, where $j \geq i \geq 1$, of mutually vertex-disjoint directed trees with their roots in $\Gamma_{G,\text{in}}^*(u) - \{u\}$ that satisfies the stated conditions in items 1, 2, and 3, then, for a vertex $v \in (\Gamma_{G,\text{in}}^*(u) - \{u\}) - \cup_{\eta=1}^i V(T_\eta)$, construct a desired dipath P_{j+1} from v to u in G by considering a dipath P from v to u in G with $\text{length}(P) = \vec{d}_G(v, u)$ in two cases

Consider the two cases of P based on its possible intersection with the constructed directed forest/family $\{T_1, T_2, \dots, T_i\}$ —which are shown in Fig. 6.

Case 1: $V(P) \cap \cup_{\eta=1}^i V(T_\eta) = \emptyset$. From the above observation that $\text{length}(P) \leq e_{G,\text{in}}(u)$, hence for P_{j+1} , we may employ P by designating $P_{j+1} = P$ and $T_{i+1} = \{P_{j+1} - \{u\}\}$ as in the basis step. We can verify the above-stated items 1, 2, and 3 for the augmented family $\{T_1, T_2, \dots, T_{i+1}\}$.

Case 2: $V(P) \cap \cup_{\eta=1}^i V(T_\eta) \neq \emptyset$. Denote the first entrance of the dipath P into $\cup_{\eta=1}^i V(T_\eta)$ by w , say $w \in V(P) \cap V(T_k)$ for some $k \in \{1, 2, \dots, i\}$.

With the denotations/labelings in Fig. 7, we have two possible dipaths from w to u : (1) the dipath:

$$Q = \underbrace{\text{(unique) dipath from } w \text{ to } \text{root}(T_k)}_{\text{contained in } T_k} \text{ concatenated with the directed edge } (\text{root}(T_k), u),$$

and (2) the dipath P_2 such that:

$$P = \underbrace{\text{dipath from } v \text{ to } w}_{\text{via vertices in } (\Gamma_{G,\text{in}}^*(u) - \{u\}) - \cup_{\eta=1}^i V(T_\eta)} \text{ concatenated with } \underbrace{\text{dipath } P_2}_{\text{from } w \text{ to } u \text{ in } G}.$$

What can we say about $\text{length}(Q)$ versus $\text{length}(P_2)$? They must be equal—via a proof by contradiction as follows:

1. Suppose that $\text{length}(Q) < \text{length}(P_2)$: The dipath from v , via w , to u formed by the concatenation of P_1 (v to w) and Q (w , via $\text{root}(T_k)$, to u) is a shorter dipath

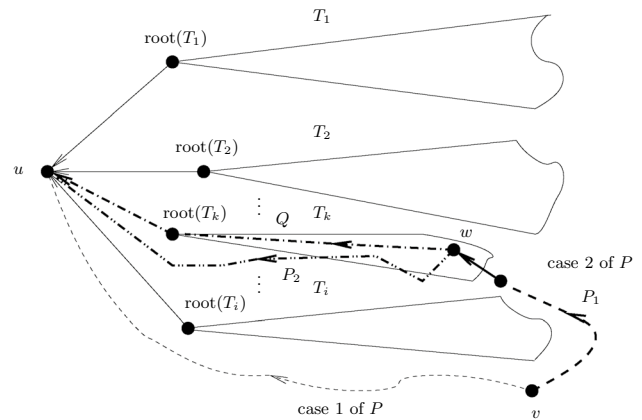


Fig. 7 For a vertex u in a digraph G : case 2 of P with $V(P) \cap \cup_{\eta=1}^i V(T_\eta) \neq \emptyset$ is considered

than P —which contradicts to the assumption that $\text{length}(P) = \vec{d}_G(v, u)$.

2. Suppose that $\text{length}(Q) > \text{length}(P_2)$: The existence of such dipath P_2 from v to u in G contradicts to the above item 2 that $\text{length}(Q) = \vec{d}_G(v, u)$.

Now, we let:

$$P_{j+1} = \underbrace{\text{dipath } P_1 \text{ from } v \text{ to } u \text{ in } G}_{\text{via vertices in } (\Gamma_{G,\text{in}}^*(u) - \{u\}) - \cup_{\eta=1}^i V(T_\eta)} \text{ concatenated with } \underbrace{\text{dipath } Q \text{ from } w \text{ to } u,}_{\text{via vertices in } T_k}$$

and include the dipath $P_{j+1} - \{u\}$ into the directed tree T_k .

We can check/verify the above-stated items 1, 2, and 3:

1. The statement is obvious,
2. The condition follows from that “shortest dipath in G ” enjoys “optimal substructure property in G ”, and
3. By noting that:

$$e_{G,\text{in}}(u) \geq \vec{d}_G(v, u) = \text{length}(P).$$

This completes the inductive construction, and we have shown the reverse inequality.

Lower Bounds for Vertex-Eccentricity via Graph-Parameters

The min-max formulation of $e_{G,\text{in}}(u)$, which was developed above for lower-bounding the number of time-steps for function computation by vertex u in $\Gamma_{G,\text{in}}^*(u)$, motivates us

to study lower bounds for (maximum) vertex-eccentricity in terms of common graph-parameters of the underlying graph G : (1) maximum in-degree, and (2) order and size.

Vertex-Eccentricity and Maximum In-Degree

We derive a lower bound on $e_{G,\text{in}}(u)$ from the knowledge of the maximum in-degree of G (vertex-spanned by $\Gamma_{G,\text{in}}^*(u)$), which yields a (possibly weaker) lower bound on the number of time-steps for vertex u to access values/information of all the vertices in $\Gamma_{G,\text{in}}^*(u)$.

Denote by $\Delta_{G,\text{in}}(u) (\geq 1)$ the maximum in-degree of the subdigraph of G vertex-spanned by $\Gamma_{G,\text{in}}^*(u)$.

Theorem 1 For a digraph G and a vertex $u \in V(G)$,

$$e_{G,\text{in}}(u) \geq \begin{cases} |\Gamma_{G,\text{in}}^*(u)| - 1 & \text{if } \Delta_{G,\text{in}}(u) = 1, \\ \log_{\Delta_{G,\text{in}}(u)}((\Delta_{G,\text{in}}(u) - 1)|\Gamma_{G,\text{in}}^*(u)| + 1) - 1 & \text{otherwise} \\ & (\Delta_{G,\text{in}}(u) \geq 2). \end{cases}$$

Proof Consider an arbitrary digraph G with a vertex $u \in V(G)$. Organize the in-closure of u in G as the sequence of successive in-neighbors as illustrated in Fig. 8: at most $\Delta_{G,\text{in}}(u)^i$ vertices of $\Gamma_{G,\text{in}}^*(u)$ are at a directed distance of i to u for $i = 0, 1, \dots, e_{G,\text{in}}(u)$, and we have the following inequality:

$$|\Gamma_{G,\text{in}}^*(u)| \leq 1 + \Delta_{G,\text{in}}(u) + \Delta_{G,\text{in}}(u)^2 + \dots + \Delta_{G,\text{in}}(u)^{e_{G,\text{in}}(u)}.$$

Consider the two cases for the $\Delta_{G,\text{in}}(u)$ -value.

Case when $\Delta_{G,\text{in}}(u) = 1$: The above inequality on $|\Gamma_{G,\text{in}}^*(u)|$ is:

$$|\Gamma_{G,\text{in}}^*(u)| \leq 1 + e_{G,\text{in}}(u),$$

which gives that $e_{G,\text{in}}(u) \geq |\Gamma_{G,\text{in}}^*(u)| - 1$.

Case when $\Delta_{G,\text{in}}(u) \geq 2$: The above inequality on $|\Gamma_{G,\text{in}}^*(u)|$ is:

$$|\Gamma_{G,\text{in}}^*(u)| \leq \frac{\Delta_{G,\text{in}}(u)^{e_{G,\text{in}}(u)+1} - 1}{\Delta_{G,\text{in}}(u) - 1},$$

which yields a lower bound on $e_{G,\text{in}}(u)$:

$$e_{G,\text{in}}(u) \geq \log_{\Delta_{G,\text{in}}(u)}((\Delta_{G,\text{in}}(u) - 1)|\Gamma_{G,\text{in}}^*(u)| + 1) - 1.$$

□

We can obtain desired lower bounds in analogous fashion with similar graph-parameters such a regularity in-degree, and maximum and regularity degrees.

Maximum Vertex-Eccentricity and Graph-Order and -Size

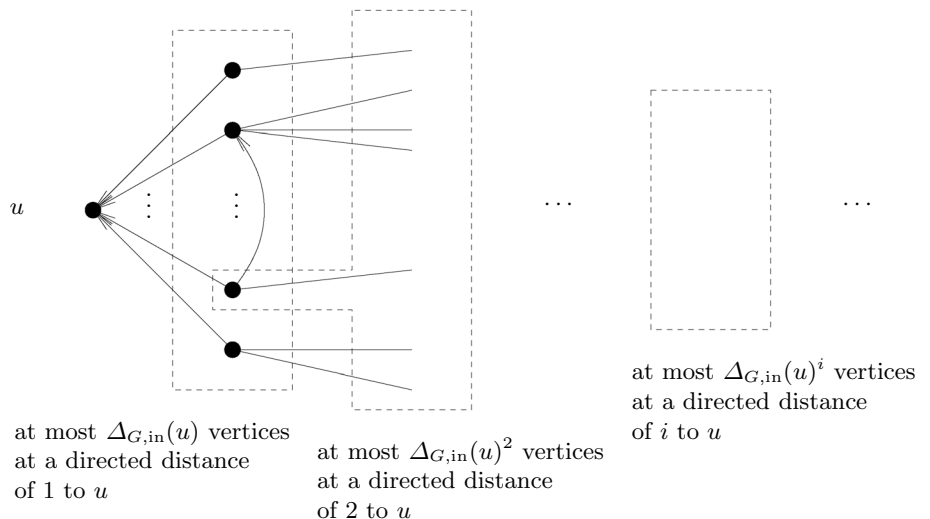
The diameter of a digraph G , denoted by $\text{dia}(G)$, is the maximum in-eccentricity of all the vertices of G ; that is,

$$\begin{aligned} \text{dia}(G) &= \max\{e_{G,\text{in}}(u) \mid u \in V(G)\} \\ &= \max\{\vec{d}_G(u, v) \mid u, v \in V(G)\}. \end{aligned}$$

A digraph G is strongly connected if for every pair of vertices $u, v \in V(G)$, there exists a dipath from u to v (and vice-versa) in G .

For a strongly connected digraph, we study a lower bound on its diameter in terms of its order and size, and show the optimality of the diameter-bound with a family of explicitly constructed strongly connected digraphs.

Fig. 8 For a vertex u in a digraph G : organize vertices of the in-closure $\Gamma_{G,\text{in}}^*(u)$ of u in G according to their directed distances to u



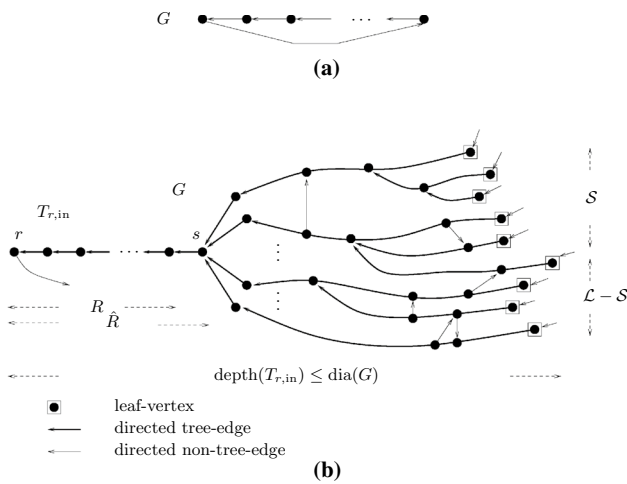


Fig. 9 For a strongly connected digraph G : **a** when $\text{dia}(G) = |V(G)| - 1$ (full diameter): $|E(G)| \geq |V(G)|$; **b** when $\text{dia}(G) \leq |V(G)| - 2$: $\text{dia}(G) \geq \frac{2(|V(G)|-1)}{|E(G)|-|V(G)|+1}$

Theorem 2 For a strongly connected digraph G ,

$$|E(G)| \geq \begin{cases} |V(G)| & \text{if } \text{dia}(G) = |V(G)| - 1, \\ |V(G)| - 1 + \frac{2(|V(G)|-1)}{\text{dia}(G)} & \text{otherwise} \\ \text{equivalently, } \text{dia}(G) \geq \frac{2(|V(G)|-1)}{|E(G)|-|V(G)|+1} & (\text{dia}(G) \leq |V(G)| - 2). \end{cases}$$

Proof For the case of full diameter: $\text{dia}(G) = |V(G)| - 1$, the extremity of $\text{dia}(G)$ gives that $|E(G)| \geq |V(G)| - 1$, and the strong connectedness of G increases the lower bound by 1: $|E(G)| \geq |V(G)|$. A size-optimal strongly connected digraph of full diameter is shown in Fig. 9a.

We study the general case, henceforth, $\text{dia}(G) \leq |V(G)| - 2$. Our approach in deriving a desired lower bound on $|E(G)|$, hence on $\text{dia}(G)$, relies on: (1) embedding in G a vertex-spanning rooted directed tree with depth of $\text{dia}(G)$, and (2) then relating the order $|V(G)|$, size $|E(G)|$, and diameter $\text{dia}(G)$ of G via a classification/enumeration of all leaf-to-root dipaths into families of dipaths with shared versus non-shared suffixes (towards the root).

To have a refined derivation, the rooted directed tree is seeded with a maximum-length stemming dipath to ensure the consideration of near-leaf proximity of the first/lowest common descendants of shared suffixes of the leaf-to-root dipaths.

Denote by R a longest dipath of vertices with in-degree of 1 in G , that is, (1) for every vertex $u \in V(R)$, $\text{deg}_{G,\text{in}}(u) = 1$, and (2) $\text{length}(R)$ is the maximum among those of dipaths satisfying item (1). Note that $0 \leq \text{length}(R) \leq \text{dia}(G) - 1$. Denote by r the terminal vertex of R , and by \hat{R} the dipath that concatenates the unique vertex s (and its directed edge) adjacent to the initial vertex of R together with the dipath R .

Employing \hat{R} as a seed-structure, we grow a vertex-spanning rooted (at the terminal vertex r of R) directed in-tree $T_{r,\text{in}}$ of G [9] in which for every vertex $u \in T_{r,\text{in}}$ ($= V(G)$), there exists a unique dipath from u to r in $T_{r,\text{in}}$.

Furthermore, we limit the depth of $T_{r,\text{in}}$ with $\text{depth}(T_{r,\text{in}}) \leq \text{dia}(G)$. For instance, we apply Dijkstra’s single-source shortest-dipaths algorithm (see, for example, [3]) with the seed-structure of \hat{R} , which results in a vertex-spanning directed in-tree $T_{r,\text{in}}$ of G such that: (1) $\text{root}(T_{r,\text{in}}) = r$, and (2) for every vertex $u \in V(T_{r,\text{in}})$ ($= V(G)$), the unique dipath P from u to the root r of $T_{r,\text{in}}$ satisfies that: (2.1) P is a shortest dipath from u to r in G with $\text{length}(P) \leq \text{dia}(G)$, and (2.2) if $u \in V(\hat{R})$ then P is the suffix of \hat{R} (with initial vertex u) else P contains \hat{R} as its suffix. We depict in Fig. 9b the topological structure of a vertex-spanning rooted directed in-tree $T_{r,\text{in}}$ of a strongly connected digraph G .

Note that, due to the maximality of $\text{length}(R)$ imposed above, every leaf-to-root dipath P must contain a vertex $u \in V(P)$ that satisfies the following sharing-condition: (1) $\text{deg}_{G,\text{in}}(u) \geq 2$, and (2) the (first) appearance of such u is within a directed distance of $\text{length}(\hat{R})$ from the initial leaf-

vertex of P , that is, $\vec{d}_{T_{r,\text{in}}}(\text{initial}(P), u) \leq \text{length}(\hat{R})$.

Denote by \mathcal{L} the family of all leaf-to-root dipaths of $T_{r,\text{in}}$ (hence the number of leaf-vertices of $T_{r,\text{in}}$ is $|\mathcal{L}|$). We partition \mathcal{L} into two disjoint subfamilies \mathcal{S} and $\mathcal{L} - \mathcal{S}$ based on the sharing-condition constrained to $T_{r,\text{in}}$:

$$\mathcal{S} = \{P \in \mathcal{L} \mid \text{there exists } u \in V(P) \text{ such that } \vec{d}_{T_{r,\text{in}}}(\text{initial}(P), u) \leq \text{length}(\hat{R}) \text{ and } \text{deg}_{T_{r,\text{in}},\text{in}}(u) \geq 2\};$$

that is, the degree-constraint is satisfied in the context of $T_{r,\text{in}}$ —with at least two tree-edges of $E(T_{r,\text{in}})$ incident/convergent to the first/lowest common descendant u shared with other dipath(s) of \mathcal{L} . An annotated configuration of leaf-to-root dipaths of \mathcal{S} versus $\mathcal{L} - \mathcal{S}$ is illustrated in Fig. 9b.

After embedding $T_{r,\text{in}}$ in G , we enumerate $V(G)$ and $E(G)$ with respect to the partition $\{\mathcal{S}, \mathcal{L} - \mathcal{S}\}$ and establish an upper and lower bounds on $|V(G)|$ and $|E(G)|$, respectively, as follows.

1. Upper-bounding $|V(G)|$:

- 1.1. For every leaf-to-root dipath $P \in \mathcal{S}$, there exists a vertex $u \in V(P)$ such that $\vec{d}_{T_{r,\text{in}}}(\text{initial}(P), u) \leq \text{length}(\hat{R})$ and u is a first/lowest common descendant shared with at least one other leaf-to-root dipath $P' \in \mathcal{S}$. For the dipath-

pair P and P' , we include: (1) all the vertices of the prefix of P preceding to s ($= \text{initial}(\hat{R})$), that is, $V(P - \hat{R})$, in the enumeration of $V(G)$ with:

$$\begin{aligned} |V(P - \hat{R})| &= (\text{length}(P) + 1) - (\text{length}(R) + 2) \\ &= \text{length}(P) - \text{length}(R) - 1 \\ &\leq \text{depth}(T_{r,\text{in}}) - \text{length}(R) - 1 \\ &= \text{dia}(G) - \text{length}(R) - 1, \end{aligned}$$

and (2) all the vertices of the prefix of P' preceding to u , that is, $\vec{d}_{T_{r,\text{in}}}(\text{initial}(P), u)$ such vertices, in the enumeration of $V(G)$ with:

$$\vec{d}_{T_{r,\text{in}}}(\text{initial}(P), u) \leq \text{length}(\hat{R}) = \text{length}(R) + 1.$$

Therefore, an upper bound on the number of vertices of the prefixes $P - \hat{R}$ for all $P \in \mathcal{S}$ that are included in the enumeration of $V(G)$ is:

$$\begin{aligned} \frac{|\mathcal{S}|}{2}(\text{dia}(G) - \text{length}(R) - 1) \\ + \frac{|\mathcal{S}|}{2}(\text{length}(R) + 1) = \frac{|\mathcal{S}|}{2}\text{dia}(G). \end{aligned}$$

- 1.2. For every leaf-to-root dipath $P \in \mathcal{L} - \mathcal{S}$, the vertex s ($= \text{initial}(\hat{R})$) may be the first/lowest common descendant that is shared with other leaf-to-root dipaths of \mathcal{S} . Therefore, an upper bound on the number of vertices of the prefixes $P - \hat{R}$ for all $P \in \mathcal{L} - \mathcal{S}$ that are included in the enumeration of $V(G)$ is:

$$|\mathcal{L} - \mathcal{S}|(\text{dia}(G) - \text{length}(R) - 1).$$

- 1.3. All the leaf-to-root dipaths of \mathcal{L} share the common suffix \hat{R} , which contributes $\text{length}(\hat{R}) + 1 = \text{length}(R) + 2$ vertices in the enumeration of $V(G)$.

Hence, the three above-derived upper bounds yield that:

$$\begin{aligned} |V(G)| \leq \frac{|\mathcal{S}|}{2}\text{dia}(G) + |\mathcal{L} - \mathcal{S}|(\text{dia}(G) - \text{length}(R) - 1) \\ + \text{length}(R) + 2. \end{aligned}$$

2. Lower-bounding $|E(G)|$:

- 2.1. The vertex-spanning rooted directed in-tree $T_{r,\text{in}}$ of G contributes $|E(T_{r,\text{in}})| = |V(T_{r,\text{in}})| - 1 = |V(G)| - 1$ tree-edges in the enumeration of $E(G)$.

- 2.2. Every leaf-to-root dipath P of \mathcal{L} satisfies the above-stated sharing-condition in the context of G : there exists a vertex u of P , within a directed distance of $\text{length}(\hat{R})$ from $\text{initial}(P)$, has $\text{deg}_{G,\text{in}}(u) \geq 2$. The membership of a leaf-to-root dipath P of \mathcal{S} demands the degree-requirement in the context of $T_{r,\text{in}}$: $\text{deg}_{T_{r,\text{in}},\text{in}}(u) \geq 2$. Hence, every leaf-to-root dipath $P \in \mathcal{L} - \mathcal{S}$ must contain a vertex u (within a directed distance of $\text{length}(\hat{R})$ from $\text{initial}(P)$) to which at least one directed non-tree-edge of $E(G) - E(T_{r,\text{in}})$ is incident/convergent. Therefore, a lower bound on the number of directed non-tree-edges of $E(G) - E(T_{r,\text{in}})$ incident to P (within a directed distance of $\text{length}(\hat{R})$ from $\text{initial}(P)$) for all $P \in \mathcal{L} - \mathcal{S}$ that are included in the enumeration of $E(G)$ is $|\mathcal{L} - \mathcal{S}|$.

- 2.3. The strong connectedness of G implies that for every leaf-vertex of $T_{r,\text{in}}$ has its in-degree of at least 1. Therefore, a lower bound on the number of directed non-tree-edges incident to all leaf-vertices of $T_{r,\text{in}}$ that are included in the enumeration of $E(G)$ is $|\mathcal{L}|$.

- 2.4. The strong connectedness of G also entails that for the root-vertex r of $T_{r,\text{in}}$, there exists a directed non-tree-edge incident from r to a vertex $u \in V(G)$ ($= V(T_{r,\text{in}})$). Consider the two cases of the containment of \mathcal{S} in \mathcal{L} that decides if the directed non-tree-edge (r, u) may contribute additionally to the enumeration of $E(G)$.

Case when $\mathcal{S} \subsetneq \mathcal{L}$: Since $\mathcal{L} - \mathcal{S} \neq \emptyset$, the vertex u may appear in a leaf-to-root dipath $P \in \mathcal{L} - \mathcal{S}$ (within a directed distance of $\text{length}(\hat{R})$ from $\text{initial}(P)$), and the directed non-tree-edge (r, u) may have been included in the enumeration of $E(G)$ above in item 2.2. In addition, if u is a leaf-vertex in $T_{r,\text{in}}$, the directed non-tree-edge (r, u) may have been accounted for in item 2.3.

Case when $\mathcal{S} = \mathcal{L}$: The vertex u must appear in a leaf-to-root dipath $P \in \mathcal{S}$. We claim that there exists a directed non-tree-edge that contributes additionally to the enumeration of $E(G)$. Consider the two subcases of $u \in V(P)$: (1) If u is not the initial leaf-vertex of P , then the directed non-tree-edge (r, u) has not been included in the enumeration of $E(G)$ above in item 2.3 nor item 2.2 (as $\mathcal{L} - \mathcal{S} = \emptyset$), or (2) If $u = \text{initial}(P)$, then the maximality of $\text{length}(R)$ imposed above implies that $\text{deg}_{G,\text{in}}(u) \neq 1$ (> 1 , in fact, due to the directed non-tree-edge (r, u) incident to u). Thus, there exists an additional directed non-tree-edge incident to u that has not been included in the enumeration of $E(G)$ above in item 2.3.

Hence, the four above-derived lower bounds yield that:

$$|E(G)| \geq \begin{cases} |V(G)| - 1 + |\mathcal{L} - \mathcal{S}| + |\mathcal{L}| & \text{if } \mathcal{S} \subsetneq \mathcal{L}, \\ |V(G)| - 1 + |\mathcal{L} - \mathcal{S}| + |\mathcal{L}| + 1 & \text{otherwise } (\mathcal{S} = \mathcal{L}). \end{cases}$$

We now combine the above upper bound on $|V(G)|$ and lower bound on $|E(G)|$ to achieve the desired lower bound on $\text{dia}(G)$ in terms of $|V(G)|$ and $|E(G)|$. In parallel to the consideration of the containment of \mathcal{S} in \mathcal{L} , we consider the following two cases.

Case when $\mathcal{S} \subsetneq \mathcal{L}$: The upper and lower bounds on $|V(G)|$ and $|E(G)|$, respectively, are:

$$|V(G)| \leq \frac{|\mathcal{S}|}{2} \text{dia}(G) + |\mathcal{L} - \mathcal{S}|(\text{dia}(G) - \text{length}(R) - 1) + \text{length}(R) + 2,$$

and

$$|E(G)| \geq |V(G)| - 1 + |\mathcal{L} - \mathcal{S}| + |\mathcal{L}| = |V(G)| - 1 + |\mathcal{L} - \mathcal{S}| + |\mathcal{S}| + |\mathcal{L} - \mathcal{S}|.$$

Eliminating $|\mathcal{S}|$ (and $\text{length}(R)$) from the two inequalities above, we obtain that:

$$\frac{2|V(G)| - 2|\mathcal{L} - \mathcal{S}|(\text{dia}(G) - \text{length}(R) - 1) - 2\text{length}(R) - 4}{\text{dia}(G)} \leq |\mathcal{S}| \leq |E(G)| - |V(G)| + 1 - 2|\mathcal{L} - \mathcal{S}|,$$

which yields that:

$$(|V(G)| - 1)\text{dia}(G) + 2(|V(G)| - 1) + 2(\text{length}(R) - 1)(|\mathcal{L} - \mathcal{S}| - 1) \leq |E(G)|\text{dia}(G).$$

From the assumption of $\mathcal{S} \subsetneq \mathcal{L}$,

$$|E(G)|\text{dia}(G) \geq (|V(G)| - 1)\text{dia}(G) + 2(|V(G)| - 1),$$

that is,

$$|E(G)| \geq |V(G)| - 1 + \frac{2(|V(G)| - 1)}{\text{dia}(G)},$$

as desired.

Case when $\mathcal{S} = \mathcal{L}$: The upper and lower bounds on $|V(G)|$ and $|E(G)|$, respectively, are:

$$|V(G)| \leq \frac{|\mathcal{S}|}{2} \text{dia}(G) + |\mathcal{L} - \mathcal{S}|(\text{dia}(G) - \text{length}(R) - 1) + \text{length}(R) + 2 = \frac{|\mathcal{L}|}{2} \text{dia}(G) + \text{length}(R) + 2,$$

and

$$|E(G)| \geq |V(G)| - 1 + |\mathcal{L} - \mathcal{S}| + |\mathcal{L}| + 1 = |V(G)| - 1 + |\mathcal{L}| + 1.$$

The two inequalities above give that:

$$\frac{2|V(G)| - 2\text{length}(R) - 4}{\text{dia}(G)} \leq |\mathcal{L}| \leq |E(G)| - |V(G)|,$$

therefore,

$$|E(G)| \geq |V(G)| + \frac{2|V(G)| - 2\text{length}(R) - 4}{\text{dia}(G)} = |V(G)| + \frac{2(|V(G)| - 1)}{\text{dia}(G)} - \frac{2(\text{length}(R) + 1)}{\text{dia}(G)} = |V(G)| - 1 + \frac{2(|V(G)| - 1)}{\text{dia}(G)} + \left(1 - \frac{2(\text{length}(R) + 1)}{\text{dia}(G)}\right).$$

Consider the algebraic sign of $1 - \frac{2(\text{length}(R)+1)}{\text{dia}(G)}$.

If $1 - \frac{2(\text{length}(R)+1)}{\text{dia}(G)} \geq 0$ (the seed-structure of $T_{r,\text{in}}$ is a relatively short maximum-length stemming dipath R), then:

$$|E(G)| \geq |V(G)| - 1 + \frac{2(|V(G)| - 1)}{\text{dia}(G)} + \left(1 - \frac{2(\text{length}(R) + 1)}{\text{dia}(G)}\right) \geq |V(G)| - 1 + \frac{2(|V(G)| - 1)}{\text{dia}(G)},$$

as desired.

Otherwise, assume that $1 - \frac{2(\text{length}(R)+1)}{\text{dia}(G)} < 0$. We revise our approach above in upper-bounding $|V(G)|$ (items 1.2 and 1.3) and derive a lower bound on $|\mathcal{L}|$ as follows.

For every leaf-to-root dipath $P \in \mathcal{L}$ ($= \mathcal{S}$), the vertex-set of the prefix P preceding to s ($= \text{initial}(\hat{R})$) is constrained by:

$$|V(P - \hat{R})| \leq \text{dia}(G) - \text{length}(R) - 1.$$

Therefore, an upper bound on the number of vertices of the prefixes $P - \hat{R}$ for all $P \in \mathcal{L}$ that are included in this enumeration of $V(G)$ is:

$$|\mathcal{L}|(\text{dia}(G) - \text{length}(R) - 1).$$

All the leaf-to-root dipaths of \mathcal{L} share the common suffix \hat{R} , which contributes $\text{length}(R) + 2$ vertices in this enumeration of $V(G)$.

Hence, the two above-stated upper bounds yield that:

$$|V(G)| \leq |\mathcal{L}|(\text{dia}(G) - \text{length}(R) - 1) + \text{length}(R) + 2,$$

which provides a lower bound on $|\mathcal{L}|$ with further simplification:

$$\begin{aligned} |\mathcal{L}| &\geq \frac{|V(G)| - \text{length}(R) - 2}{\text{dia}(G) - \text{length}(R) - 1} \\ &= \frac{|V(G)| - 1 - \text{dia}(G) + (\text{dia}(G) - \text{length}(R) - 1)}{\text{dia}(G) - \text{length}(R) - 1} \\ &= \frac{|V(G)| - 1 - \text{dia}(G)}{\text{dia}(G) - \text{length}(R) - 1} + 1. \end{aligned}$$

From the assumption that $1 - \frac{2(\text{length}(R)+1)}{\text{dia}(G)} < 0$, which implies that:

$$(0 \leq) \text{dia}(G) - (\text{length}(R) + 1) < \frac{\text{dia}(G)}{2},$$

the above lower bound on $|\mathcal{L}|$ gives that:

$$\begin{aligned} |\mathcal{L}| &\geq \frac{|V(G)| - 1 - \text{dia}(G)}{\text{dia}(G) - (\text{length}(R) + 1)} + 1 \\ &> \frac{|V(G)| - 1 - \text{dia}(G)}{\frac{\text{dia}(G)}{2}} + 1 \\ &= \frac{2(|V(G)| - 1)}{\text{dia}(G)} - 1. \end{aligned}$$

Hence, the lower bound on $|E(G)|$ derived above (for this case of $\mathcal{S} = \mathcal{L}$) is revised as follows:

$$\begin{aligned} |E(G)| &\geq |V(G)| - 1 + |\mathcal{L} - \mathcal{S}| + |\mathcal{L}| + 1 \\ &= |V(G)| - 1 + |\mathcal{L}| + 1 \\ &> |V(G)| - 1 + \left(\frac{2(|V(G)| - 1)}{\text{dia}(G)} - 1 \right) + 1 \\ &= |V(G)| - 1 + \frac{2(|V(G)| - 1)}{\text{dia}(G)}, \end{aligned}$$

as desired.

Thus, for $\text{dia}(G) \leq |V(G)| - 2$ and in either case ($\mathcal{S} \subsetneq \mathcal{L}$ or $\mathcal{S} = \mathcal{L}$), we arrive at the same relationship for the order, size, and diameter of G :

$$|E(G)| \geq |V(G)| - 1 + \frac{2(|V(G)| - 1)}{\text{dia}(G)};$$

equivalently,

$$\text{dia}(G) \geq \frac{2(|V(G)| - 1)}{|E(G)| - |V(G)| + 1}.$$

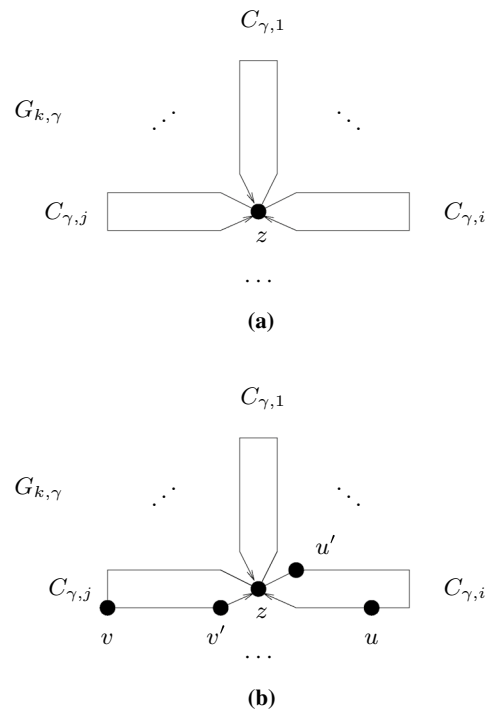


Fig. 10 For each positive integers k and γ : **a** the strongly connected digraph $G_{k,\gamma}$ is an amalgamation of k (mutually edge-disjoint) copies of a directed cycle of $\gamma + 1$ vertices: $C_{\gamma,1}, C_{\gamma,2}, \dots, C_{\gamma,k}$ that share a common vertex; **b** when $\gamma \geq 2$ and $i, j \in \{1, 2, \dots, k\}$ with $i \neq j$, (u', v') is a diametrical pair of vertices in $C_{\gamma,i}$ and $C_{\gamma,j}$, respectively, with $d_{G_{k,\gamma}}(u', v') = 2\gamma$

We construct a family of strongly connected digraphs that achieve the optimality of the above-derived diameter-bound. For each positive integer γ , denote by C_γ a directed cycle of $\gamma + 1$ vertices, and for each positive integer k , define a strongly connected digraph $G_{k,\gamma}$ to be an amalgamation of k (mutually edge-disjoint) copies of C_γ : $C_{\gamma,1}, C_{\gamma,2}, \dots, C_{\gamma,k}$ that share a common vertex z . Figure 10a show the topological structure of $G_{k,\gamma}$.

Corollary 3 *The family of strongly connected digraphs $G_{k,\gamma}$ for all positive integers k and γ is optimal for the relationship of the graph-parameters: order, size, and diameter established in Theorem 2.*

Proof For the case of full parameter (graph-diameter = graph-order - 1), we consider the family of $G_{k,\gamma}$ for $k = 1$ and all positive integers γ . The strongly connected digraph $G_{1,\gamma}$ is (a copy of) C_γ , which exhibits the following graph-parameters:

$$|V(G_{1,\gamma})| = \gamma + 1 \text{ and } |E(G_{1,\gamma})| = \gamma + 1, \text{ and } \text{dia}(G_{1,\gamma}) = \gamma = |V(G_{1,\gamma})| - 1,$$

□ which satisfy that $|E(G_{1,\gamma})| = |V(G_{1,\gamma})|$.

For the general case of graph-diameter \leq graph-order $- 2$, we consider the family of $C_{k,\gamma}$ for all positive integers $k \geq 2$ and all positive integers γ . The strongly connected digraph $G_{k,\gamma}$ is an amalgamation of k (mutually edge-independent) copies of C_γ sharing a common vertex z , which exhibits the following graph-parameters:

$$|V(G_{1,\gamma})| = k\gamma + 1 \text{ and } |E(G_{1,\gamma})| = k(\gamma + 1).$$

To compute the diameter of $G_{k,\gamma}$, $\text{dia}(G)$, it suffices to notice that, for all $i, j \in \{1, 2, \dots, k\}$: if $i = j$, then for all vertices $u, v \in V(C_{\gamma,i}) (= V(C_{\gamma,j}))$, we have: $\vec{d}_{G_{k,\gamma}}(u, v) \leq \gamma$; otherwise ($i \neq j$), for all vertices $u \in V(C_{\gamma,i}) - \{z\}$ and $v \in V(C_{\gamma,j}) - \{z\}$, and with u' and v' denoting the (unique) vertices of $V(C_{\gamma,i}) - \{z\}$ and $V(C_{\gamma,j}) - \{z\}$, respectively, which are adjacent from and to, respectively, the vertex z , we have: $\vec{d}_{G_{k,\gamma}}(u, v) \leq \vec{d}_{G_{k,\gamma}}(u', v') = \gamma + \gamma = 2\gamma$. Figure 10b illustrates a diametrical pair of vertices in $G_{k,\gamma}$. The three graph-parameters satisfy that:

$$\text{dia}(G) = 2\gamma = \frac{2(|V(G)| - 1)}{|E(G)| - |V(G)| + 1}.$$

□

Conclusion

Distributed function computation has a wide spectrum of major applications in distributed systems. There is a natural need to understand and approximate, if possible, lower and upper bounds on the number of time-steps for some or all vertices to receive (initial) vertex-values of all vertices of the network-graph, regardless of the underlying protocol or algorithmics.

The number of time-steps for a vertex u of a network-graph G to collect values/information from all the vertices in $\Gamma_{G,\text{in}}^*(u)$ is lower-bounded by the in-eccentricity of u , $e_{G,\text{in}}(u)$, in G . In accordance with the above-stated min-max distributed framework, we have proved that $e_{G,\text{in}}(u)$ is the minimum of the maximum tree-depth of a directed forest among all possible directed forests that are vertex-decomposition of $\Gamma_{G,\text{in}}^*(u) - \{u\}$ and rooted in (as subset of) $\Gamma_{G,\text{in}}(u)$.

However, it is not necessary to compute $e_{G,\text{in}}(u)$, directly or indirectly, through an underlying optimal directed forest whose maximum tree-depth yields $e_{G,\text{in}}(u)$. Instead, with a given/fixed topology of a directed forest (described in above fashion) underlying a distributed iteration scheme, we can obtain a stronger lower bound with a distributed computation of the maximum depth D_{max} among all the directed trees in the forest, and note that:

number of time-steps for vertex u to access values or information from all vertices in the given directed forest $\geq 1 + D_{\text{max}} \geq e_{G,\text{in}}(u)$.

We also address lower bounds on vertex-eccentricity and its maximum version, graph-diameter, in terms of common graph-parameters of the underlying graph in a graph-theoretic framework:

1. For a digraph G and a vertex $u \in V(G)$, we have shown a lower bound on $e_{G,\text{in}}(u)$ in terms of the graph-order and maximum in-degree of G , and
2. For a strongly connected digraph G , we have proved a lower bound on $\text{dia}(G)$ in terms of the graph-order and -size of G , and have demonstrated the optimality of the diameter-bound for a family of (explicitly constructed) strongly connected digraphs.

In addition to the probabilistic upper-bound result on the number of time-steps for (general) distributed function computation via linear iterative schemes with random weight-matrix, Sundaram and Hadjicostis [11, 12] employ observability theory of linear systems to study the linear-functional case for distributed computation (of linear functions), and achieve an upper bound via the minimal polynomial of the underlying weight-matrix.

Toulouse and Minh [15] study the linear functional case with prescribed time-invariant network-topology over random weight-matrices, and obtain various empirical upper-bound results.

In accordance with an information-theoretic framework, Xu and Raginsky [18, 19] study the fundamental time-step limits of distributed function computation in a constrained probabilistic setting. The lower- and upper-bound results are based on tradeoffs between: (1) the minimal amount of information necessarily extracted about the function value by any accuracy- and confidence-constrained algorithm, and (2) the maximal amount of information about the function value obtained by any algorithm within specified time-step and communication bounds. The lower-bound analysis indicates the dependence of computation time-steps on the diameter of the underlying network-graph, while the upper-bound one relies on cutset-capacity arguments.

In addition, there have been several other recent theoretical developments in distributed computation and optimization. Olshevsky and Tsitsiklis [10] prove lower bounds on the worst-case convergence time for various classes of linear, time-invariant (in network-topology), distributed consensus methods. Kuhn, Moscibroda, and Wattenhofer [8] study lower and upper bounds on local/distributed computability and approximability (amount of local information, approximation ratio, communication round) for a large class of optimization problems: minimum vertex cover, minimum (connected) dominating set, maximum matching, maximal independent set, and maximal matching.

Funding This study was not supported by any funding.

Compliance with ethical standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Ayaso O, Shah D, Dahleh MA. Information theoretic bounds for distributed computation over networks of point-to-point channels. *IEEE Trans Inf Theory*. 2010;56(12):6020–39.
2. Bondy JA, Murty USR. Graph theory, volume 244 of graduate texts in mathematics. London: Springer; 2008.
3. Cormen TH, Leiserson CE, Rivest RL, Stein C. Introduction to algorithms. 3rd ed. Cambridge: MIT Press; 2009.
4. Fich FE, Ruppert E. Hundreds of impossibility results for distributed computing. *Distrib Comput*. 2003;16(2–3):121–63.
5. Hendrickx JM, Olshevsky A, Tsitsiklis JN. Distributed anonymous discrete function computation. *IEEE Trans Autom Control*. 2011;56(10):2276–89.
6. Kashyap A, Basar T, Srikant R. Quantized consensus. *Automatica*. 2007;43(7):1192–203.
7. Katz G, Piantanida P, Debbah M. Collaborative distributed hypothesis testing. *Computing Research Repository*, [arXiv:abs/1604.01292](https://arxiv.org/abs/1604.01292), 2016.
8. Kuhn F, Moscibroda T, Wattenhofer R. Local computation: lower and upper bounds. *J ACM*. 2016;63(2):17:1–44.
9. Mehlhorn K, Sanders P. Algorithms and data structures: the basic toolbox. New York: Springer; 2008.
10. Olshevsky A, Tsitsiklis JN. Convergence speed in distributed consensus and averaging. *SIAM J Control Optim*. 2009;48(1):33–55.
11. Sundaram, S. Linear iterative strategies for information dissemination and processing in distributed systems. PhD thesis, University of Illinois at Urbana-Champaign;2009.
12. Sundaram S, Hadjicostis CN. Distributed function calculation and consensus using linear iterative strategies. *IEEE J Sel Areas Commun*. 2008;26(4):650–60.
13. Sundaram S, Hadjicostis CN. Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Trans Autom Control*. 2011;56(7):1495–508.
14. Toulouse M, Minh B Q. Applicability and resilience of a linear encoding scheme for computing consensus. In: Muñoz V M, Wills G, Walters R J, Firouzi F, Chang V, editors. Proceedings of the Third International Conference on Internet of Things, Big Data and Security, IoTBDS 2018, Funchal, Madeira, Portugal, March 19-21, 2018. p. 173–184. SciTePress.
15. Toulouse M, Minh BQ, Minh QT. Invariant properties and bounds on a finite time consensus algorithm. *Trans Large Scale Data Knowl Cent Syst*. 2019;41:32–58.
16. Wang L, Xiao F. Finite-time consensus problems for networks of dynamic agents. *IEEE Trans Autom Control*. 2010;55(4):950–5.
17. Xiao L, Boyd SP, Kim S-J. Distributed average consensus with least-mean-square deviation. *J Parallel Distrib Comput*. 2007;67(1):33–46.
18. Xu A. Information-theoretic limitations of distributed information processing. PhD thesis, University of Illinois at Urbana-Champaign;2016.
19. Xu A, Raginsky M. Information-theoretic lower bounds for distributed function computation. *IEEE Trans Inf Theory*. 2017;63(4):2314–37.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.