A resistance-based approach to performance analysis of the consensus algorithm

E. Lovisari, F. Garin, and S. Zampieri

***Abstract*— We study the well–known linear consensus algorithm by means of a LQ-type performance cost. We want to understand how the communication topology influences this algorithm. In order to do this, we recall the analogy between Markov Chains and electrical resistive networks. By exploiting this analogy, we are able to rewrite the performance cost as the average effective resistance on a suitable network. We use this result to show that if the communication graph fulfills some local properties, then its behavior can be approximated with that of a suitable grid, over which the behavior of the cost is known.**

I. Introduction

In the last two decades a great attention has been devoted to multi-agent systems, where a large number of simple agents exchange information locally and distributedly coordinates the achievement of a global task. The local nature of the communication can be modeled as a graph $\mathcal{G}$ (the communication graph), with an edge among agents capable of exchanging information. A basic example of common task is to asymptotically reach the same value; a simple algorithm to this purpose is the linear consensus, described as follows. Denoting by a vector $\boldsymbol{x}(t) \in \mathbb{R}^N$ the states of the agents at time t, the evolution is: $\boldsymbol{x}(t+1) = P\,\boldsymbol{x}(t)$, where P is consistent with $\mathcal{G}$, i.e. has zero entries whenever (j, i) is not an edge. In this simple scenario the graph $\mathcal{G}$ and the matrix P are assumed to be time-invariant. Under the assumption that P is stochastic, aperiodic and irreducible, it is well–known that $x_i(t) \overset{t\to\infty}{\longrightarrow} \boldsymbol{\pi}^T \boldsymbol{x}(0)$ for each agent i, where $\boldsymbol{\pi}^T$ is such that $\boldsymbol{\pi}^T P = \boldsymbol{\pi}^T$, and $\boldsymbol{x}(0)$ is the vector of the initial values. We underline the fact that a stochastic matrix P is associated to a Markov Chain with transition probabilities given by the entries of P itself. In this paper we assume the matrix P to be symmetric, which yields $\boldsymbol{\pi}^T = \frac{1}{N}\mathbf{1}^T$ (by $\mathbf{1}$ we denote a vector with all entries equal to 1). The classical theory also gives the speed of convergence of the linear consensus algorithm, which is exponential, with speed given by the essential spectral radius (second largest singular value) of P. However, the convergence speed is not the only index for performance evaluation. When consensus is not an objective per se, but is used to solve an estimation or control problem, it is important to consider also other performance measures, more tightly related to the actual objective pursued. Different costs arise from different problems, and it can be shown by examples that considering a different performance measure can indeed lead to preferring different graph topologies. In this paper, we consider an LQ cost familiar to control theorists. We consider the case when the initial condition is a random variable with zero-mean and covariance matrix $\mathbb{E}\left[\boldsymbol{x}(0)^T \boldsymbol{x}(0)\right] = I$, and we study the expectation of the H^2-norm of the trajectory of the states:

$$J(P) = \frac{1}{N}\sum_{t\geq 0} \mathbb{E}\left[\|x(t) - x(\infty)\|^2\right] = \frac{1}{N}\,\mathrm{trace} \sum_{t\geq 0}\left(P^{2t} - \tfrac{1}{N}\mathbf{1}\mathbf{1}^T\right). \tag{1}$$

Proposition 1.1: Assume P is a row–stochastic symmetric matrix, irreducible and aperiodic. Then the cost defined in Eq. 1 can be rewritten as

$$J(P) = \frac{1}{N}\,\mathrm{trace} \sum_{t\geq 0}\left(P^{2t} - \tfrac{1}{N}\mathbf{1}\mathbf{1}^T\right). \tag{2}$$

□

The same cost arises also in quite different frameworks, e.g. from consensus algorithm in the presence of noise [1], or from a formation-control problem [2].

Clearly, for any given matrix the cost $J(P)$ can be computed, so that it is easy to compare different choices for P. However, it is interesting to unveil the fundamental role played by the graph topology, and moreover it is useful to establish -for all graphs within some family- the asymptotic behaviour when N grows, so as to give design guidelines. The behaviour of $J(P)$ is known for particular families of highly structured graphs and of matrices respecting the graph's symmetries. For example, for a d-dimensional grid of N vertices, under some symmetry assumption on the entries of P, it is known [1], [2] that $J(P)$ grows linearly in N if $d = 1$, logarithmically if $d = 2$, and is constant if $d \geq 3$. The

The research leading to these results has received funding from the European Communitys Seventh Framework Programme under agreement n. FP7-ICT-223866-FeedNetBack.

E. Lovisari and S. Zampieri are with DEI (Department of Information Engineering), Universitá di Padova, Via Gradenigo 6/b, 35131 Padova, Italy, {lovisari, zampi}@dei.unipd.it

F. Garin is with INRIA Grenoble Rhône-Alpes, 655 av. de l'Europe, Montbonnot, 38334 Saint-Ismier cedex, France, federica.garin@inrialpes.fr

goal of this paper is to extend such results to a more general class of graphs and matrices.

II. Main results

The tool used in order to deal with asymmetries and irregularities of P is the well–known analogy [3] between Markov Chains and resistive electrical networks. The electrical network associated with a symmetric stochastic matrix P has the same graph $\mathcal{G}$ associated with P, and on any edge (i, j) it has a conductance of value $c_{ij} = P_{ij}$. We will denote an electric network as the couple $(\mathcal{G}, R)$, where $\mathcal{G} = (V, \mathcal{E})$ is an undirected graph, and where R is a function $\mathcal{E} \to \mathbb{R}^+$ which associates to each edge its resistance.

First, some notation. We call graphical distance between two node in $\mathcal{G}$, and we denote it with the symbol $d_{\mathcal{G}}(u, v)$, the length of the shortest path between u and v in $\mathcal{G}$. We assume all our graphs to be connected, so this distance is always finite.

We denote with $\mathcal{G}^{(h)}$ the h-fuzz of the graph $\mathcal{G}$: it has the same set of nodes, and an edge between u and v if the graphical distance between such nodes is less then h. For example, the graph associated to P^2 is the 2-fuzz of the graph $\mathcal{G}$ associated to P, so we will call it $\mathcal{G}^{(2)}$.

Proposition 2.1: Let P be a row–stochastic symmetric matrix, irreducible and aperiodic. Build the electric network associated to P^2, $(\mathcal{G}^{(2)}, R)$, and assume in this network the effective resistance between two nodes, u and v, is $\mathcal{R}_{uv}(\mathcal{G}, R)$. Then the cost Eq. 2 can be rewritten as

$$J(P) = \frac{1}{N^2} \sum_{u,\, v} \mathcal{R}_{uv}(\mathcal{G}, R) \stackrel{\triangle}{=} \overline{\mathcal{R}}(\mathcal{G}, R), \qquad (3)$$

which is the average effective resistance in the network. □

The interest of this result is that we can exploit the monotonicity properties of the effective resistances. It holds in fact the following result.

*Theorem 2.1 (**Rayleigh's monotonicity law**):* Given an electric network, if, in any edge, the resistance is increased (resp. decreased), then the effective resistance increases (resp. decreases). □

Moreover, it holds the following "scaling" property.

Lemma 2.1: Given, with a slight abuse of notation, a network $(\mathcal{G}, r_0)$ with all the resistances equal to r_0, the effective resistance between two generic nodes is exactly r_0 times the effective resistance among the same nodes, in the same graph but having set all resistances equal to 1, $(\mathcal{G}, 1)$. □

These properties imply the following result.

Proposition 2.2: Given an electric network $(\mathcal{G}, R)$, in which each resistance lies in an interval $[r_{min}, r_{max}]$, $0 < r_{min}, r_{max} < \infty$, it holds true

$$r_{min}\mathcal{R}_{uv}(\mathcal{G}, 1) \le \mathcal{R}_{uv}(\mathcal{G}, R) \le r_{max}\mathcal{R}_{uv}(\mathcal{G}, 1).$$

□

Because of this proposition, we will often use the short-hand notation $\mathcal{R}_{uv}(\mathcal{G})$ instead of $\mathcal{R}_{uv}(\mathcal{G}, 1)$.

To conclude, we can state the following result, whose proof can be found in [4], dealing with the h-fuzz of a graph.

Lemma 2.2: Assume to have the electrical network $\mathcal{G}$ (with all resistances equal to 1), and build its h-fuzz, h finite. Then there exists a parameter c, dependent only on h and on a geometric property of the graph, such that

$$c\mathcal{R}_{uv}(\mathcal{G}) \le \mathcal{R}_{uv}(\mathcal{G}^{(h)}) \le \mathcal{R}_{uv}(\mathcal{G}).$$

□

In this lemma, if h is bounded and the number of neighbors of any agent is bounded, then the parameter c is bounded away from zero: this turns out to be important in all those cases in which the number of agents N grows and we want to understand how our cost scales with N.

The previous propositions allow us to state the following important result.

Theorem 2.2: Let P be a row–stochastic, symmetric, irreducible and aperiodic matrix. Assume that each entry of P satisfies $0 < p_{min} < p_{ij} < p_{max}$, and that the number of non-zero entries on each row is bounded[1]. Then there exist two constants, t_1 and t_2, independent from the number of agents N and from the graph topology, such that

$$t_1\overline{\mathcal{R}}(\mathcal{G}) \le J(P) \le t_2\overline{\mathcal{R}}(\mathcal{G}). \qquad (4)$$

□

The importance of this result is that the cost has been strongly related to the graph topology only, regardless of the actual values of the entries of the matrix P.

III. Geometric graphs: an application

The interpretation of $J(P)$ as an average effective resistance allows to give results for a quite general class of graphs, which we will call "geometric graphs" because they arise naturally when agents (e.g., sensors) are deployed in a bounded region of the Euclidean space, in a roughly uniform way but with possible irregularities, and communicate with their geographic neighbors. The intuition is that such graphs look like perturbed grids, and so their performance should resemble that of grids. This intuition is indeed correct, and our main result is a formal proof, which we obtain following a procedure inspired by [4]. We assume the components of P to be constrained between a minimum and a maximum fixed values, and $\mathcal{G}$ to satisfy the following assumptions: vertices of $\mathcal{G}$ are points lying in an hypercube Q of

[1] This corresponds to assuming that the number of neighbors of any node in the associated graph is bounded.

$\mathbb{R}^d$; each vertex has a number of neighbors limited by an upper bound independent of N; the following local properties hold true:

- the distance between two vertices is always larger than a parameter s;
- if two vertices communicate, then their relative distance is at most r;
- γ is the radius of the largest ball, centered in a point inside Q, containing no vertex;
- for all pairs of vertices within a suitable threshold distance, the ratio between their Euclidean and their graphical distance is larger than a parameter ρ.

Such geometric graphs can be approximated with grids, in the sense that we can build two grids $\mathcal{L}_1$ and $\mathcal{L}_2$ in $\mathbb{R}^d$, whose size only depend on the parameters s, r, γ and ρ, and on N, and four constants, q_1, k_1, q_2 and k_2, depending only on the four geometric parameters, such that

$$k_1 + q_1\overline{\mathcal{R}}(\mathcal{L}_1) \leq J(P) \leq k_2 + q_2\overline{\mathcal{R}}(\mathcal{L}_2),. \tag{5}$$

To prove this result we show that the following inequality holds true

$$\bar{k}_1 + \bar{q}_1\overline{\mathcal{R}}(\mathcal{L}_1) \leq \overline{\mathcal{R}}(\mathcal{G}) \leq \bar{k}_2 + \bar{q}_2\overline{\mathcal{R}}(\mathcal{L}_2). \tag{6}$$

for another set of constants. Then we use Prop. 2.1.

The technique to prove the last two inequalities is as follows: we tessellate Q with hypercubes in two ways, dependent on the parameters defined above. In the first tessellation the hypercubes are large enough to contain at least one node of $\mathcal{G}$, and then we identify all the nodes in the same hypercube. Then we consider the lattice whose nodes are the hypercubes and which has an edge between two nodes if the two corresponding hypercubes touch (not diagonally). Building a suitable h-fuzz of such lattice, we can embed it in $\mathcal{G}$ and use Prop. 2.2 in order to obtain the upper bound. In the second tessellation, instead, the hypercubes contain at most one node of $\mathcal{G}$, and the result follows identifying an hypercube with the closest node of $\mathcal{G}$ and following the same argument used for the upper bound.

More details will be given in a forthcoming paper.

REFERENCES

[1] R. Carli, F. Garin, and S. Zampieri, "Quadratic indices for the analysis of consensus algorithms," *ITA Workshop 2009*, 2009.

[2] B. Bamieh, M. Jovanovic, P. Mitra, and S. Patterson, "Coherence in large-scale networks: Dimension dependent limitations of local feedback," *IEEE Trans. Automat. Control*, 2010.

[3] P. Doyle and J. Snell, *Random Walks and Electric Networks*, ser. Carus Monographs. Mathematical Association of America, 1984.

[4] P. Barooah and J. Hespanha, "Estimation from relative measurements: Error bounds from electrical analogy," *ICISIP*, 2005.