Lower bound performances for average consensus in open multi-agent systems

Charles Monnoyer de Galland and Julien M. Hendrickx

Abstract— We derive fundamental limitations on the performances of intrinsic averaging algorithms in open multi-agent systems, which are systems subject to random arrivals and departures of agents. Each agent holds a value, and their goal is to estimate the average of the values of the agents presently in the system. We provide a lower bound on the expected Mean Square Error for any estimation algorithm, assuming that the number of agents remains constant and that communications are random and pairwise. Our derivation is based on the expected error obtained with an optimal algorithm under conditions more favorable than those the actual problem allows, and relies on an analysis of the constraints on the information spreading mechanisms in the system, and relaxations of these.

I. INTRODUCTION

Multi-agent systems show great benefits for modeling and solving problems in various domains including sensor networks [1], [2], vehicle coordination [3], or social phenomena [4], [5]. Among the most cited properties of multi-agent systems are their flexibility, their scalability, or their robustness. Yet, most results around multi-agent systems stand for asymptotic properties under the convenient assumption that their composition remains unchanged. The increasing size of the systems challenges this assumption as it implies slower processes and higher probabilities of arrivals and departures, making those non-negligible. It is also challenged by the chaotic nature of some systems, where communications can be difficult, or happen at a time-scale highly comparable to that of the arrivals and departures: vehicles can for instance share a stretch of road before heading to different destinations in collaborative multi-vehicles systems.

All those reasons bring up the emergence of the study of *Open Multi-Agent Systems*. Results about closed systems do not easily extend to open ones: repeated arrivals and departures imply important differences in the design and analysis of such systems, and result in several challenges, see e.g. [6], [7]. First, with the frequent arrivals and departures of agents, the size of an open system changes with time, making the analysis of its state challenging. Moreover, incessant perturbations impact the state, but also in some cases the objective pursued by the agents of the system: algorithms then have to adapt to that variable objective, making their design challenging, and the usual convergence cannot be achieved anymore.

ICTEAM institute, UCLouvain (Belgium). J.H. is also with the CISE, Boston University (USA). This work was supported by "Communauté française de Belgique - Actions de Recherche Concertées". C.M. is a FRIA fellow (F.R.S.-FNRS), and J.H. holds a WBI.World excellence fellowship. Email adresses: `charles.monnoyer@uclouvain.be`, `julien.hendrickx@uclouvain.be`.

A. State of the art

There is little theoretical analysis of open multi-agent systems, as most multi-agent results rely on the assumption that systems are closed. However, some results considered arrivals and departures, such as simulation-based analyses of social phenomena performed in [8]. Also, it has been shown in [6] that systems subject to gossip interactions in open features can be analyzed through size-independent descriptors, whose evolution described by a dynamical system is shown to asymptotically converge to some steady states. Moreover, algorithm design has been explored for MAX-consensus problems with arrivals and departures in [7] through additional variables, and where performance was measured by the probability for the estimate to eventually converge if the system closes. Similarly, THOMAS architecture was designed to maintain connectivity into P2P networks [9]. Openness was also considered in applications such as VTL for autonomous cars to deal with cross-sections [10].

Nevertheless, up to now, efficiently studying performances of algorithms and analyzing open systems remains a challenge in general. As agents cannot instantaneously react to perturbations, obtaining an exact result is often out of range. Hence, the goal should be to remain near a time-varying objective, and usual convergence is no more relevant to be studied as it cannot be achieved. A step towards understanding open systems is the derivation of *fundamental performance limitations*: lower bounds on the performances that are possible to achieve. Through those limitations, one can obtain some quality criterion for algorithm design in open systems, but also get a better understanding of the possible bottlenecks that could arise.

B. Contribution

We establish fundamental performance limitations in open multi-agent systems for intrinsic averaging consensus problems (*i.e.* estimating the average of all intern values owned by agents *present in the system at that time*), where information is exchanged between agents through random pairwise communications. In this analysis, we focus on systems of fixed size: we assume that each departure of an agent is instantaneously followed by an arrival, so that *only replacements occur, and the system size remains constant*, see Section II for a formal definition. The approach used for deriving the fundamental limitations is detailed in Section III, and consists in evaluating the performances of an algorithm that is provably optimal under favorable settings where agents have access to more resources than decentralized problems

typically allow. These performances depend on some complicated distribution related to information propagation. Bounds on this distribution were obtained by considering relaxed conditions for the information exchange, namely the Ping model in Section IV and the Infection model in Section V, to properly define the performance limitations.

Since consensus problems are commonly a building block for various multi-agent complex applications (such as decentralized optimization [11] or several control applications [3]), we expect the techniques we use for deriving those limitations to be extendable to more advanced tasks on open multi-agent systems.

II. PROBLEM STATEMENT

A. System description

We consider a set of N agents, labelled from 1 to N. Each agent $i \in \{1 \dots N\}$ holds an i.i.d. intrinsic value $x_i(t) \in \mathbb{R}$ randomly selected from a distribution which we assume without loss of generality to be zero mean and of variance $\mathbb{E}x_i^2 = \sigma^2$. This value remains constant except at a replacement, which we model as the complete erasure of the agent's memory and the attribution of a new value $x_i(t)$ drawn from the same distribution. Those replacements happen according to a Poisson clock defined by an individual replacement rate λ_r, so that on average $N\lambda_r$ replacements happen in the whole system per unit of time.

Agents interact *via* pairwise communications: every two agents interact with each other at random times determined by a Poisson clock with rate λ_c. There are thus on average $\frac{N(N-1)}{2}\lambda_c$ communications taking place in the system per unit of time. Agents are deterministic and have unlimited memory and computation power and there is no restriction on the size or nature of the messages they can exchange. They have a unique identifier that they can use and they have access to a common universal time. The way we model replacements as a memory erasure also means agents know the correspondence between the agents having left and those having replaced them since they share the same label. We assume moreover that they know the number of agents N, the parameters λ_r, λ_c, and the distribution of the x_i (although we will see that the results would be the same if they knew only the expected value of x_i, assumed here to be 0). Agents have thus access to significantly more information than in many works on multi-agent systems, but our lower bounds on the algorithm performance will of course also apply to these more usual situations, since our setting allows implementing any algorithm that could be implemented under more restrictive conditions.

B. Objective

In intrinsic averaging, agents try to estimate the average of the intern values of the agents present in the system at that time, denoted $\bar{x}(t)$. For that purpose, every agent i maintains its own estimate $y_i(t)$ of that average based on its knowledge about the system, and updates it in continuous time. Under ideal conditions, one would expect that estimate to become $y_i(t) = \bar{x}(t) = \frac{1}{N}\sum_{j=1}^N x_j(t)$.

This is not achievable in open systems because of the variable objective and the delays to transmit information in the system, leading to the absence of usual convergence. We thus need a quantitative measure of the *performance* of an algorithm solving this problem in open systems, and choose the classical Mean Square Error (MSE) of the estimation of all the agents at time t, presented in equation (1).

$$C(t) := \frac{1}{N}\sum_{j=1}^N (\bar{x}(t) - y_j(t))^2 \qquad (1)$$

In this analysis, we focus on the steady state-error: we assume the system has been running since $-\infty$, so that the effect of initial conditions has disappeared, and $\mathbb{E}[C(t)]$ can be considered independent of the time (one can verify that our processes remain well defined). We thus derive fundamental performance limitations for all algorithms that can be implemented in our setting as a (time-invariant) lower bound on (2).

$$\mathbb{E}[C(t)] = \mathbb{E}\left[(\bar{x}(t) - y_i(t))^2\right] \qquad (2)$$

C. Preliminary notions

Before stating our first result, we need to introduce certain notions related to the information available to an agent when computing its estimate, and to the age of its information.

We first formalize in the next definition all the information about events and values x_j to which an agent i could possibly access, and thus influence its estimate $y_i(t)$.

Definition 1: At its arrival at time t, an agent i whose value is set to x_i owns a *knowledge set* $\omega_i(t)$ such that $\omega_i(t^+) = \{i, x_i, t\}$. If agents i and j interact at time $\tilde{t}$, then $\omega_i(\tilde{t}^+) = \omega_j(\tilde{t}^+) = \omega_i(\tilde{t}^-) \cup \omega_j(\tilde{t}^-) \cup \{i\text{-}j; \tilde{t}, x_i(\tilde{t}), x_j(\tilde{t})\}$.

The last expression of the union defining an interaction denotes that i, j have interacted at time $\tilde{t}$ and confirms their values at that time.[1] In other words, the knowledge sets after interaction consist of the union of the knowledge sets, to which is added the information about the interaction and confirmation of the values. Standard results in distributed computation show that *any estimate $y_i(t)$ that an agent can compute in our settings can actually be computed based only on $\omega_i(t)$ and the time t*.

Observe that $\omega_i(t)$ may contain various values $x_j(t')$ for different $t' < t$. However, we will see later that only the most recent known values about the other agents are needed to build the estimate $y_i(t)$.

Definition 2: The *most recent value* known by the agent i about j at time t is denoted $\tilde{x}_j^{(i)}(t) = x_j(\tilde{t}_j^{(i)}(t))$, where $\tilde{t}_j^{(i)}(t) := \max\{s : x_j(s) \in \omega_i(t)\}$ is the time at which that most recent value was first obtained. We denote the *age of that information* by $T_j^{(i)}(t) := t - \tilde{t}_j^{(i)}(t)$.

By convention, if no value x_j lies in $\omega_i(t)$, we set x_j to 0 and the corresponding age to $+\infty$. Moreover, the estimate of an agent about itself is always correct, and the corresponding age of the information is 0.

In steady state, and due to the symmetry between the agents, the distribution of $T_j^{(i)}(t)$ is independent of i, j

---

[1] This confirmation is included for simplicity, but is actually redundant, as the values at that time could be obtained from $\omega_i(\tilde{t}^-)$ and $\omega_j(\tilde{t}^-)$.

and t. Hence, all agents share a common *cdf* and *pdf* for that variable, respectively denoted $F(s)$ and $f(s)$. Moreover, with a small abuse of language, we say that another *pdf* $f^*(s)$ *bounds* $f(s)$ when the corresponding *cdf* $F^*(s)$ satisfies $F^*(s) \geq F(s)$ $\forall s$, which is always satisfied for a random variable $T^* \leq_{st} T$, in the *usual stochastic order* [12].

III. BOUND IN TERMS OF THE INFORMATION SPREADING MECHANISM

We present in Theorem 1 a general lower bound on (2) depending on the way information spreads in the system. To properly define a bound, we will thus need to find relaxations of that information spreading to instantiate the expression.

Theorem 1: The performances of any algorithm in the setting defined in Section II are bounded from below as

$$\mathbb{E}[C(t)] \geq \frac{N-1}{N^2} \int_0^\infty f^*(t) \left(1 - e^{-2\lambda_r t}\right) \sigma^2 \mathrm{d}t \qquad (3)$$

where $f^*(t)$ bounds $f(t)$ as defined in Section II-C.

The bound derived in the above theorem is obtained by studying an optimal algorithm in the setting of Section II (Section III-A). It relies on the decomposition of the contributions of all the agents for the estimation of the average to obtain the MSE for some knowledge (Sections III-B to III-D), and then on the derivation of a steady state through the analysis of the way information spreads in the system (Sections III-E and III-F).

A. Optimal algorithm definition

We first build an algorithm that is optimal for solving intrinsic averaging based on a knowledge set.

Proposition 1: In the setting described in Section II, the following estimate is optimal in the sense of criterion (2).

$$y_i(t) = \mathbb{E}[\bar{x}(t)|\omega_i(t)] = \frac{1}{N} \sum_{j=1}^N \mathbb{E}[x_j(t)|\omega_i(t)] \qquad (4)$$

Proof: Conditional to the knowledge set $\omega_i(t)$, algorithm (4) is optimal since it minimizes the MSE:

$$\frac{d}{dy_i(t)} \mathbb{E}\left[(\bar{x}(t) - y_i(t))^2|\omega_i(t)\right] = 2y_i(t) - 2\mathbb{E}[\bar{x}(t)|\omega_i(t)] = 0.$$

Hence, since any other algorithm only depends on $\omega_i(t)$ and t, the result of any of them $y_i^*(t)$ is such that

$$\mathbb{E}\left[(\bar{x}(t) - y_i^*(t))^2|\omega_i(t)\right] \geq \mathbb{E}\left[(\bar{x}(t) - y_i(t))^2|\omega_i(t)\right].$$

Considering the expected value on all $\omega_i(t)$, it follows:

$$\mathbb{E}\left[\mathbb{E}\left[(\bar{x}(t) - y_i^*(t))^2|\omega_i(t)\right]\right] \geq \mathbb{E}\left[\mathbb{E}\left[(\bar{x}(t) - y_i(t))^2|\omega_i(t)\right]\right]. \qquad \blacksquare$$

Since the algorithm (4) is optimal in the sense of criterion (2), its performance

$$\mathbb{E}[C(t)] = \mathbb{E}\left[\mathbb{E}\left[(\bar{x}(t) - \mathbb{E}[\bar{x}(t)|\omega_i(t)])^2|\omega_i(t)\right]\right] \qquad (5)$$

provides a lower bound on the performance of all algorithms that can be deployed in our settings.

B. Individual contribution

We show in the following proposition that expression (5) conveniently reduces to the analysis of a single agent.

Proposition 2: The criterion (2) for the algorithm (4) conditional to the knowledge set $\omega_i(t)$ reduces to

$$\mathbb{E}\left[(\bar{x}(t) - y_i(t))^2|\omega_i(t)\right] = \frac{1}{N^2} \sum_{j=1}^N \mathbb{E}\left[(x_j(t) - \hat{x}_j(t))^2|\omega_i(t)\right] \quad (6)$$

where $\hat{x}_j(t) := \mathbb{E}[x_j(t)|\omega_i(t)]$ (we lighten the notation by dropping the dependence of $\hat{x}_j^{(i)}(t)$ on i).

Proof: With the optimal algorithm (4), one has

$$y_i(t) = \frac{1}{N} \sum_{j=1}^N \mathbb{E}[x_j(t)|\omega_i(t)]$$

It follows that the error given $\omega_i(t)$ is written

$$\mathbb{E}\left[(\bar{x}(t) - y_i(t))^2|\omega_i(t)\right] = \frac{1}{N^2} \mathbb{E}\left[\left(\sum_{j=1}^N (x_j(t) - \hat{x}_j(t))\right)^2 \Big| \omega_i(t)\right].$$

The absence of correlation between the agents values $x_i(t)$ finally allows to nullify the crossed-product terms of the squared sum, to obtain the final expression. $\blacksquare$

C. Single agent estimate

We can then explicitly write an expression for the estimate of a single agent $\hat{x}_j(t)$.

Proposition 3: There holds,

$$\hat{x}_j(t) = \mathbb{E}[x_j(t)|\omega_i(t)] = e^{-\lambda_r T_j^{(i)}} x_j(t - T_j^{(i)}). \qquad (7)$$

Hence, the estimate $\hat{x}_j(t)$ only depends on the most recent information about the agent j (note that the time-dependence of $T_j^{(i)}(t)$ is removed to lighten the notations).

Proof: The most recent information we know about j is given by Definition 2, and we have no information about whether it was replaced since then. Hence, denoting R the event that j has been replaced, and $\bar{R}$ that it has not,

$$\hat{x}_j(t) = \mathbb{E}[x_j(t)|\bar{R}] \cdot P(\bar{R}) + \mathbb{E}[x_j(t)|R] \cdot (1 - P(\bar{R})).$$

By definition, $\mathbb{E}[x_j(t)|R] = 0$. Then, from Poisson properties:

$$\hat{x}_j(t) = 0 + \tilde{x}_j^{(i)}(t) \; e^{-\lambda_r \left(t - \tilde{t}_j^{(i)}(t)\right)}$$

and the conclusion follows from the definition of the age of the most recent information. $\blacksquare$

The result above also stands when an agent estimates its own value or that of an agent for which it has no information, leading respectively to $\hat{x}_i^{(i)}(t) = x_i(t)$ and $\hat{x}_j^{(i)}(t) = 0$.

D. Individual error

For concision matters, we denote

$$C_j^{(i)}(t) := (x_j(t) - \hat{x}_j(t))^2 \qquad (8)$$

We can then write the MSE of the estimation of a single agent by injecting (7) into (8).

Proposition 4: The MSE when estimating a single agent for algorithm (4) conditional to the knowledge set $\omega_i(t)$ is

$$\mathbb{E}\left[(x_j(t) - \hat{x}_j(t))^2|\omega_i(t)\right] = \left(1 - e^{-2\lambda_r T_j^{(i)}}\right)\sigma^2 \qquad (9)$$

and is thus entirely characterized by the age of the most recent information about it.

Proof: Denoting R the event of at least one replacement of the agent j during $T_j^{(i)}$ and by $\bar{R}$ the event of no replacement, we can develop (8) with a case-by-case analysis:

$$\mathbb{E}\left[C_j^{(i)}(t)|\omega_i(t)\right] = \mathbb{E}\left[C_j^{(i)}(t)|R\right] \cdot P(R) + \mathbb{E}\left[C_j^{(i)}(t)|\bar{R}\right] \cdot P(\bar{R}).$$

We develop each term to obtain the final result:

$$P(R) = 1 - e^{-\lambda_r T_j^{(i)}} \qquad \mathbb{E}\left[C_j^{(i)}(t)|R\right] = \left(1 + e^{-2\lambda_r T_j^{(i)}}\right)\sigma^2$$

$$P(\bar{R}) = e^{-\lambda_r T_j^{(i)}} \qquad \mathbb{E}\left[C_j^{(i)}(t)|\bar{R}\right] = \left(1 - e^{-\lambda_r T_j^{(i)}}\right)^2 \sigma^2$$

This last result allows to write the following:

$$\mathbb{E}[C(t)|\omega_i(t)] = \frac{1}{N^2} \sum_{j=1}^{N} \left(1 - e^{-2\lambda_r T_j^{(i)}}\right)\sigma^2. \quad (10)$$

E. Global expected value

We have developed an expression for the error in terms of the age of the most recent information about the other agents in (10). We can now obtain an expression for criterion (2) by computing the expected value of that result.

Proposition 5: The MSE defined in (2) is given by

$$\mathbb{E}[\bar{x}(t) - y_i(t)]^2 = \frac{N-1}{N^2} \int_0^\infty f(t)\left(1 - e^{-2\lambda_r t}\right)\sigma^2 \, dt, \quad (11)$$

where we remind that $\mathbb{E}x_i = 0$ and $\mathbb{E}x_i^2 = \sigma^2$. It is thus entirely characterized by the distribution of the age of an information, and knowing $f(t)$ leads to a proper bound.

Proof: By definition, the global MSE is

$$\mathbb{E}\left[(\bar{x}(t) - y_i(t))^2\right] = \mathbb{E}\left[\mathbb{E}\left[(\bar{x}(t) - y_i(t))^2|\omega_i(t)\right]\right].$$

Using the result (10), it becomes

$$\mathbb{E}\left[(\bar{x}(t) - y_i(t))^2\right] = \frac{1}{N^2} \sum_{j=1}^{N} \mathbb{E}\left[\left(1 - e^{-2\lambda_r T_j^{(i)}}\right)\right]\sigma^2.$$

Finally, since all $T_j^{(i)}$ follow $f(t)$ from Section II-C,

$$\mathbb{E}\left[\left(1 - e^{-2\lambda_r T_j^{(i)}}\right)\right] = \int_0^\infty f(t)\left(1 - e^{-2\lambda_r t}\right)dt,$$

and the conclusion is then direct, where we remind that $T_i^{(i)} = 0$ by definition (*cfr* Section II-C). ∎

F. Relaxation of the communication process

The previous result is actually that of Theorem 1 where the *pdf* is exactly $f(s)$. Since this *pdf* reveals to be hard to compute, we show in the next proposition that a *pdf* $f^*(s)$ bounding $f(s)$ as defined in Section II-C leads to a proper lower bound on (11), which concludes the development of Theorem 1. In the next two Sections, we will then present two possible relaxations of the information spreading mechanism that will both lead to an appropriate $f^*(s)$ to instantiate the bound.

Proposition 6: Given some value $E = \int_0^\infty F'(t) \cdot err(t)dt$ where $F(t)$ is a *cdf*, and where $err(t)$ is a positive non-decreasing function, then for any other *cdf* $F^*(t) \geq F(t) \, \forall t$,

$$E^* = \int_0^\infty (F^*)'(t) \cdot err(t)dt \leq E \quad (12)$$

Proof: By defining $F^* = F + \Delta$ with $\Delta \geq 0$, one has

$$E^* = E - \int_0^\infty \Delta \cdot err'(t)dt$$

thanks to *cdf* properties. Using then the positivity of Δ and $err'(t)$, the conclusion is direct. ∎

IV. STRONG ASSUMPTION: PING MODEL

A. Assumption description

The first relaxation we consider relies on a strong simplification of the communication mechanism. We assume that each time it communicates, an agent *acquires all the information about all the agents presently in the system,* and never forgets it, even at replacements. We refer to this assumption as the *Ping model,* in reference to the *ping* software used

to test the reachability of machines in a network. The most recent value known by the agent i about j at time t is then given by $\tilde{x}_j^{(i)}(t) = x_j(\tilde{t}_j^{(i)}(t))$ where $\tilde{t}_j^{(i)}(t)$ is the time of the very last communication in which agent i was involved, and the age of that information $T_{j,Ping}^{(i)}(t) = t - \tilde{t}_j^{(i)}(t)$ is the time spent since then.

Proposition 7: In steady state, $\forall i \neq j$, the age of the most recent information about j held by i, noted $T_{j,Ping}^{(i)}(t)$ follows

$$f^{Ping}(s) = (N-1)\lambda_c e^{-(N-1)\lambda_c s}, \quad (13)$$

which bounds $f(s)$ as defined in Section II-C.

Proof: The random variable $T_{j,Ping}^{(i)}(t)$ defines the time spent since the last communication involving the agent i, and thus follows by definition of the communications a Poisson process of rate $(N-1)\lambda_c$. Its *cdf* is then given by

$$F^{Ping}(s) = P\left[T_{j,Ping}^{(i)}(t) \leq s\right] = 1 - e^{-(N-1)\lambda_c s}.$$

The *pdf* is then obtained with $f^{Ping}(s) = \frac{dF^{Ping}(s)}{ds}$.

$T_{j,Ping}^{(i)}(t)$ is actually the minimal value that can take $T_j^{(i)}(t)$. Hence $T_{j,Ping}^{(i)}(t) \leq_{st} T_j^{(i)}(t)$, and $f^{Ping}(s)$ bounds $f(s)$. ∎

B. Results

Theorem 2: The performances of any algorithm in the setting defined in Section II are bounded from below as

$$\mathbb{E}[C(t)] \geq \frac{N-1}{N^2}\left(\frac{1}{1 + \frac{N-1}{2}\frac{\lambda_c}{\lambda_r}}\right)\sigma^2 \quad (14)$$

Proof: Applying Theorem 1 with the *pdf* $f^{Ping}(t)$, the bound is given by performing the integration

$$\mathbb{E}[C(t)] \geq \frac{N-1}{N^2} \int_0^\infty (N-1)\lambda_c e^{-(N-1)\lambda_c s}\left(1 - e^{-2\lambda_r s}\right)\sigma^2 ds.$$

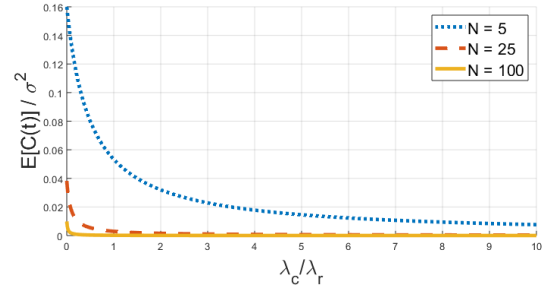A few algebraic operations lead to the conclusion. ∎

Fig. 1: Bound derived with the Ping model (14) in terms of the rate ratio between pairwise communications and replacements. Each curve stands for a different number of agents N, with a unit variance $\sigma^2 = 1$.

Several observations on the evolution of the error in terms of the ratio of the rates λ_c/λ_r and N arise from Fig. 1. When communications are rather rare ($\lambda_c/\lambda_r \to 0$), the error tends to $\frac{N-1}{N^2}\sigma^2$, *i.e.* the error that an agent would obtain only considering itself in the estimation. As the communications become more frequent, the second factor of (14) makes the bound decrease, to ultimately bring it to 0 as $\lambda_c/\lambda_r \to +\infty$, which is the error to which we converge in a close system.

Observe that $\bar{\lambda}_c = (N-1)\lambda_c$ is the communication rate for one agent. Hence, the bound is actually separated into

two factors: $\frac{N-1}{N^2}\sigma^2$ and the second one that only depends on $\bar{\lambda}_c/\lambda_r$ which characterizes the expected number of interactions involving one agent before it is replaced.

V. WEAKER ASSUMPTION: INFECTION MODEL

A. Assumption description

The random variable $T_j^{(i)}(t)$ denotes the age of the most recent information about j available to i at time t, *i.e.* the time since this information was emitted by j. Since the distribution of $T_j^{(i)}(t)$ is independent of the time t, its *cdf* $F(s)$ corresponds to the probability that an information $x_j(\tau)$ (with $0 \le \tau \le s$) is available to i at time s. This probability is lower than what it would be if the agents were never replaced (as information is erased with replacements). Hence we use this latter replacement-free situation as a relaxation to obtain a bound on the *cdf*. One can verify that when agents are never replaced, the information propagation follows a simple infection process: j is infected at time 0, and agents get infected as soon as they communicate with an infected agent. The probability of i having that information about j at time s is then equal to the probability of i being infected at time s, denoted by $P[T_{j,Inf}^{(i)}(t) \le s]$ where $T_{j,Inf}^{(i)}$ is the time of its first infection. Based on this, one can prove that the *cdf* of $T_{j,Inf}^{(i)}$ bounds $F(s)$ as defined in Section II-C.

Proposition 8: In steady state, the pdf of $T_{j,Inf}^{(i)}(t)$, noted $f^{Inf}(s)$, bounds $f(s)$ as defined in Section II-C.

Proof: Considering one realization of communications, there exists a sequence of events where $T_j^{(i)}(t)$ is such that it did not suffer replacements. Since that sequence always also exists with the Infection model, $T_{j,Inf}^{(i)}(t) \le_{st} T_j^{(i)}(t)$. ∎

B. Result

Proposition 9: In steady state, $\forall i \ne j$, the age of the most recent information agent i has about agent j with the Infection model, denoted $T_{j,Inf}^{(i)}(t)$, follows the *cdf*

$$F^{Inf}(s) = \sum_{k=1}^{N} \frac{k-1}{N-1} P_k(s) \quad (15)$$

where $P_k(s)$ is defined by an ODE system:

$$\dot{P}_k(s) = (k-1)(N-k+1)\lambda_c P_{k-1}(s) - k(N-k)\lambda_c P_k(s) \quad (16)$$

such that $P_1(0) = 1$ and $P_{k \ne 1}(0) = 0$.

Proof: Let $I(s)$ be the set containing the labels of the infected agents at time s, $N_{I(s)}$ its size, and denote $P_k(s) = P[N_{I(s)} = k]$. Then, one has

$$F^{Inf}(s) = P[i \in I(s)] = \sum_{k=1}^{N} P_k(s) \cdot P[i \in I(s)|N_{I(s)} = k].$$

From the symmetry between the agents, and since they are all interchangeable, one has

$$P[i \in I(s)|N_{I(s)} = k] = \frac{k-1}{N-1}.$$

Finally, the ODE system defining $P_k(s)$ is a standard result from continuous time Markov chains theory. ∎

Theorem 3: The performances of any algorithm in the setting defined in Section II are bounded from below as

$$\mathbb{E}[C(t)] \ge \frac{N-1}{N^2}\left(1 - w^T A(2\lambda_r I - A)^{-1}e_1\right)\sigma^2 \quad (17)$$

where A is the bidiagonal matrix defining the ODE system (16) from Proposition 9 i.e., $A_{ii} = -i(N-i)\lambda_c$ and

$A_{i+1,i} = i(N-i)\lambda_c$, and $w^T = \begin{bmatrix} 0 & \frac{1}{N-1} & \frac{2}{N-1} & \dots & 1 \end{bmatrix}$, with $\lambda_r \ne 0$.

Proof: Equation (16) defines a linear system $\dot{P}(s) = AP(s)$ with $P(s) = \begin{bmatrix} P_1(s) & P_2(s) & \dots & P_N(s) \end{bmatrix}^T$, whose solution is given by

$$P(s) = e^{As}P(0) \quad \text{with} \quad P(0) = e_1.$$

It follows from (15) that the *cdf* and *pdf* are given by

$$F^{Inf}(s) = w^T e^{As} e_1 \quad \text{and} \quad f^{Inf}(s) = w^T A e^{As} e_1.$$

Injecting that result in the integral from Theorem 1, which we will see is well defined, one has

$$\int_0^\infty f^{Inf}(s)\left(1 - e^{-2\lambda_r s}\right)ds = 1 - w^T A \int_0^\infty \left(e^{As}e^{-2\lambda_r s}\right)ds\, e_1.$$

Since A and $-2\lambda_r I$ commute, the integral reduces to

$$\int_0^\infty \left(e^{As}e^{-2\lambda_r Is}\right)ds = \int_0^\infty e^{(A-2\lambda_r I)s}ds.$$

Since $\lambda_r \ne 0$, then $(A - 2\lambda_r I)$ is invertible and has strictly negative eigenvalues, leading to

$$\int_0^\infty e^{(A-2\lambda_r I)s}ds = (A - 2\lambda_r I)^{-1}\left[e^{(A-2\lambda_r I)s}\right]_0^\infty = (2\lambda_r I - A)^{-1}.$$

Hence, the integral from Theorem 1 is given by

$$\int_0^\infty f^{Inf}(s)\left(1 - e^{-2\lambda_r s}\right)ds = 1 - w^T A(2\lambda_r I - A)^{-1}e_1,$$

and the conclusion is direct from Theorem 1. ∎

Corollary 1: The expression below is equivalent to (17).

$$\mathbb{E}[C(t)] \ge \frac{N-1}{N^2}\left(1 - \frac{1}{1 + \frac{N-1}{2}\frac{\lambda_c}{\lambda_r}} h\left(N, \frac{\lambda_c}{\lambda_r}\right)\right)\sigma^2 \quad (18)$$

with

$$h(N,L) = \sum_{k=2}^{N}\left(\frac{k-1}{N-1}\prod_{j=1}^{k-1}\frac{j(N-j)L}{2 + (j+1)(N-j-1)L}\right) \quad (19)$$

Corollary 1 presents an algebraic expression for the bound (17) from Theorem 1, and is proven in Appendix A of [13]. Detailed empirical explorations suggest that $h(N,L) \le \frac{L}{2} + \frac{L}{2}\sum_{k=1}^{N-2}\frac{2kL}{2+2kL}$. Using results on bounds in numerical integration, we can then find a compact bound on that expression, and derive the following empirical expression for bounding $\mathbb{E}[C(t)]$ from below.

$$\frac{N-1}{N^2}\left(\frac{1}{1 + \frac{N-1}{2}\frac{\lambda_c}{\lambda_r}}\right)\left(1 + \frac{1}{2}\log\left(\frac{1 + (N-2)\frac{\lambda_c}{\lambda_r}}{1 + \frac{\lambda_c}{\lambda_r}}\right)\right)\sigma^2 \quad (20)$$

C. Discussion

Fig. 2 compares the bounds derived with the Ping and Infection models, and the expression (20) based on the Infection bound, in logarithmic scale. The first observation is the improvement of both bounds from the Infection model, which relies on a weaker relaxation, in comparison with the Ping model. That improvement is even more apparent when the rate ratio λ_c/λ_r increases, since the impact of the corresponding factor gets more important. In particular, the expression (20) is actually exactly that of the bound from the Ping model multiplied by a logarithmic factor that scales to $\log N$ when λ_c/λ_r is large. This could relate to the fact that information propagates in one hop in the Ping model, whereas it needs time with the Infection model.

Observe also in Fig. 2 that the bound follows two distinct regimes: when λ_c/λ_r is very small, it has a very low impact
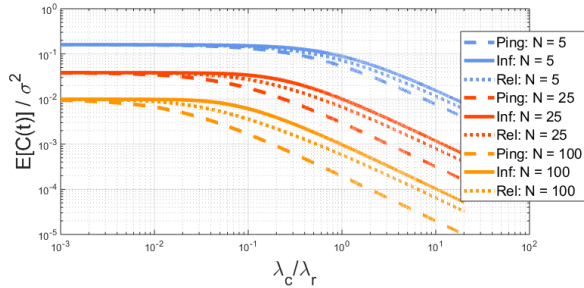
Fig. 2: Bounds derived with the Ping model (dashed), the Infection model (plain), and the expression (20) (dotted) in terms of the rate ratio λ_c/λ_r, for different numbers of agents N, and with a unit variance $\sigma^2 = 1$.
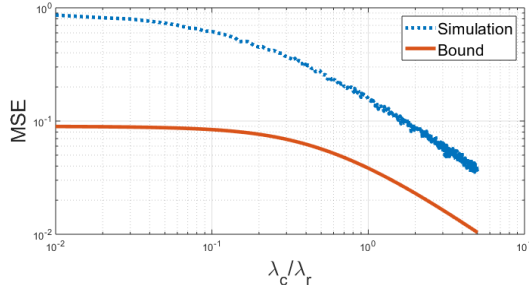

Fig. 3: Performances comparison for solving an intrinsic consensus problem with 10 agents between the bound from the Infection model (plain) and a simulation using a simple Gossip algorithm (dotted), both in logarithmic scale. The simulations are performed with 10000 iterations of the simulation over 200 events, which is empirically shown to be enough for the system to be considered in steady state in that setting.

on the bound, which stays around $\frac{N-1}{N^2}\sigma^2$; whereas after some threshold, the bound decays according to $1/(\lambda_c/\lambda_r)$.

Finally, we have compared our bound with the results of the simplest version of average gossip [14]: agents take their own value x_i as initial estimate $y_i(t)$, and whenever two agents i, j communicate, their estimates become

$$y_i(t^+) = y_j(t^+) = \frac{y_i(t^-) + y_j(t^-)}{2}. \tag{21}$$

This algorithm computes the average in closed system. Observe that it is consistent with our setting presented in Section II, and that we did not make any adaptation to the open character. Fig. 3 compares the performances of the algorithm (21) with the bound from the Infection model (17). It is interesting to notice that even though there is a gap between the curves, the behavior of the error in terms of λ_c/λ_r is well captured by the bound, especially when that ratio gets large enough. Interestingly, the algorithm relies on only one variable, and does not make use of identifiers nor of any other information provided to the agents as described in Section II, while our bound is valid for algorithms potentially using this information. Hence one could wonder about the precise impact of that information and the efficiency gains it allows.

VI. CONCLUSIONS

We considered in this paper the possibility of arrivals and departures of agents in the study of multi-agent systems, and highlighted several challenges arising from this property. In particular, it prevents algorithms to converge for many problems, making their analysis challenging.

We focused on the analysis of the performances of intrinsic averaging algorithms for open systems of constant size, for which we derived *fundamental performance limitations* as lower bounds on the Mean Square Error of estimation in steady state. This was done by studying the performances of an algorithm optimal in a context more favorable to the agents than what is usually assumed in multi-agent systems, and relied then on the relaxation of the information spreading mechanism in the system. Two of such relaxations were studied, leading to a conservative but readable bound, and to a tighter but more complex bound.

Possible extensions include the study of more complicated problems (e.g. decentralized optimization [11]), more structured constraints on the inter-agent communications, and the consideration of variable-size systems. More fundamentally, our bounds purely rely on limitations on the information propagation, and are valid even if the agents have some strong knowledge on the system and use identifiers or heterogeneous algorithms. Stronger lower bounds can be obtained under more restrictive assumptions, e.g. by improving the Infection model from Section V with healing. Yet, it would be interesting to investigate what factors significantly impact the bound: in particular, an anonymous algorithm was already shown in Section V-C to exhibit performances not too different from our bound, questioning the impact of anonymity of the agents.

REFERENCES

[1] J. Predd, S. Kulkarni, and H. V. Poor, "Distributed learning in wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 23, pp. 56–69, 07 2006.

[2] R. Olfati-Saber and N. F. Sandell, "Distributed tracking in sensor networks with limited sensing range," pp. 3157 – 3162, 07 2008.

[3] Y. Cao, W. Yu, W. Ren, and G. Chen, "An overview of recent progress in the study of distributed multi-agent coordination," *IEEE Transactions on Industrial Informatics*, vol. 9, 07 2012.

[4] V. Blondel, J. M. Hendrickx, and J. Tsitsiklis, "Continuous-time average-preserving opinion dynamics with opinion-dependent communications," *SIAM Journal on Control and Optimization*, vol. 48, 07 2009.

[5] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence models, analysis and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, 07 2002.

[6] J. M. Hendrickx and S. Martin, "Open multi-agent systems: Gossiping with deterministic arrivals and departures," in *Proceedings of the 56th IEEE CDC*, pp. 1094–1101, 09 2016.

[7] M. Abdelrahim, J. M. Hendrickx, and W. M. Heemels, "Max-consensus in open multi-agent systems with gossip interactions," in *Proceedings of the 56th IEEE CDC*, pp. 4753–4758, 12 2017.

[8] P. Sen and B. K. Chakrabarti, *Sociophysics: an introduction*. Oxford University Press, 2013.

[9] A. Giret, V. Julin, M. Rebollo, E. Argente, C. Carrascosa, and V. Botti, "An open architecture for service-oriented virtual organizations," pp. 118–132, 09 2010.

[10] O. K. Tonguz, "Red light, green lightno light: Tomorrow's communicative cars could take turns at intersections," *IEEE Spectrum*, vol. 55, pp. 24–29, Oct 2018.

[11] J. C. Duchi, A. Agarwal, and M. J. Wainwright, "Dual averaging for distributed optimization: Convergence analysis and network scaling," *IEEE Transactions on Automatic Control*, vol. 57, 05 2010.

[12] A. Müller and D. Stoyan, *Comparison methods for stochastic models and risks*. Chichester ; New York, NY : Wiley, 2002.

[13] *arXiv e-prints*, p. arXiv:1909.02475 [cs.MA], Sep 2019.

[14] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE/ACM Trans. Netw.*, vol. 14, pp. 2508–2530, June 2006.