Gossip along the way: Order-Optimal Consensus through Randomized Path Averaging

Florence Bénézit*, Alexandros G. Dimakis† , Patrick Thiran*, Martin Vetterli*†
*School of IC, EPFL, Lausanne CH-1015, Switzerland
†Department of Electrical Engineering and Computer Science (EECS)
University of California, Berkeley
Berkeley, CA 94720, USA

Abstract— **Gossip algorithms have recently received significant attention, mainly because they constitute simple and robust algorithms for distributed information processing over networks. However for many topologies that are realistic for wireless ad-hoc and sensor networks (like grids and random geometric graphs), the standard nearest-neighbor gossip converges very slowly.**

A recently proposed algorithm called geographic gossip improves gossip efficiency by a $\sqrt{n/\log n}$ factor for random geometric graphs, by exploiting geographic information of node locations. In this paper we prove that a variation of geographic gossip that averages along routed paths, improves efficiency by an additional $\sqrt{n/\log n}$ factor and is order optimal for grids and random geometric graphs. Our analysis provides some general techniques and can be used to provide bounds on the performance of randomized message passing algorithms operating over various graph topologies.

I. Introduction

Recently there has been significant interest in designing *completely distributed* algorithms for disseminating and processing information over graphs that model wireless sensor and ad-hoc networks. In particular, the problem of computing a global function of data that is distributed over a network, using only localized message-passing, is fundamental for numerous applications.

These problems and their connections to mixing rates of Markov chains have been extensively studied starting with the pioneering work of Tsitsiklis [18]. Earlier work studied mostly deterministic protocols, known as average consensus algorithms, in which each node communicates with each of its neighbors in every round. More recent work (e.g. [9], [2], [10]) has focused on so-called gossip algorithms, a class of randomized algorithms that solve the averaging problem by computing a sequence of randomly selected pairwise averages. Gossip and consensus algorithms have been the focus of renewed interest over the past several years [9], [3], [11], [12], [1], [6], motivated by applications in sensor networks and distributed control systems.

The simplest setup is the following: n nodes are placed on a graph where edges correspond to reliable communication links. Each node is initially given a scalar (which could correspond to some sensor measurement like temperature) and we are interested in solving the *distributed averaging* problem: namely, to find a distributed message-passing algorithm by which *all nodes* can compute the *average* of all n scalars. A scheme that computes the average can easily be modified to compute any linear function (projection) of the measurements as well as more general functions. Further, the scalars can be replaced with vectors and generalized to address problems like distributed filtering and optimization as well as distributed detection in sensor networks [17], [19], [15]. Random projections computed via gossip, can be used for compressive sensing of sensor measurements and field estimation as proposed in [14].

Gossip algorithms [9], [3] solve the averaging problem by first having each node randomly pick one of their one-hop neighbors and iteratively compute pairwise averages: Initially all the nodes start with an estimate of the average being their own measurement and update this estimate with a pairwise average of current estimates with a randomly selected neighbor, at each gossip round. An attractive property of gossip is that no coordination is required for the gossip algorithm to converge to the global average when the graph is connected – nodes can just randomly wake up, select one of their one-hop neighbors randomly, exchange estimates and update their estimate with the average. We will refer to this algorithm as *standard* or nearest-neighbor gossip.

A fundamental issue is the analysis of the performance of such algorithms, namely the communication (number of messages passed between one-hop neighboring nodes) required before a gossip algorithm converges to a sufficiently accurate estimate. For energy-constrained sensor network applications, communication corresponds to energy consumption and therefore should be minimal. Clearly the convergence time will depend on the graph connectivity, and we expect well-connected graphs to spread information faster and hence require fewer messages to converge.

This question was first analyzed for the complete graph [9], where it was shown that $\Theta(n \log \epsilon^{-1})$ gossip messages need to be exchanged to converge to the global average within ϵ accuracy. Boyd et al. [3] analyzed the convergence time

Work performed in part while A.G. Dimakis was visiting EPFL.

of standard gossip for any graph and showed that is closely linked to the mixing time of a Markov chain defined on the communication graph and further addressed the problem of optimizing the neighbor selection probabilities to accelerate convergence.

For certain types of well connected graphs (including complete graphs, expander graphs and small world graphs), standard gossip converges very quickly requiring the same number of messages ($\Theta(n \log \epsilon^{-1})$) as the fully connected graph. Note that any algorithm that averages n numbers will require $\Omega(n)$ messages.

Unfortunately, for random geometric graphs and grids, which are the relevant topologies for large wireless ad-hoc and sensor networks, standard gossip is extremely wasteful in terms of communication requirements. For instance, even optimized standard gossip algorithms on grids converge very slowly, requiring $\Theta(n^2 \log \epsilon^{-1})$ messages [3], [6]. Observe that this is of the same order as the energy required for every node to flood its estimate to all other nodes. On the contrary, the obvious algorithm of averaging numbers on a spanning tree and flooding back the average to all the nodes requires only $O(n)$ messages. Clearly, constructing and maintaining a spanning tree in dynamic and ad-hoc networks introduces significant overhead and complexity, but a quadratic number of messages is a high price to pay for fault tolerance.

Recently Dimakis et al [6] proposed *geographic gossip*, an alternative gossip scheme that reduces the number of required messages by a $\sqrt{n/\log n}$ factor over standard gossip on random geometric graphs, with slightly more complexity at the nodes. Assuming that the nodes have knowledge of their geographic location and under some assumptions in the network topology, greedy geographic routing can be used to build an *overlay network* and a gossip algorithm that iteratively averages *any pair* of nodes at the expense of routing the estimates through the network. In [6] the authors show that simple greedy routing towards a randomly selected position in the deployment field introduces enough randomness and that the benefit of fast mixing outweighs the extra cost of routing by a factor of $\sqrt{n/\log n}$.

This paper: The main result of this paper is that geographic gossip with the additional modification of averaging all the nodes on the routed paths, requires a linear (in the number of nodes) number of messages for random geometric graphs and grids. Observe that averaging the whole route comes almost for free in multihop communication, since a packet can accumulate the sum and the number of nodes visted, compute the average when it reaches its final destination and follow the same route backwards to disseminate the average. More precisely, we show that the expected number of messages, i.e. the expected communication cost $\mathcal{E}(n, \epsilon)$, required for *geographic gossip with path averaging* to compute the average (within ϵ accuracy) scales like $\Theta(n \log \epsilon^{-1})$ for grids and random geometric graphs, with high probability over the graph realization. Therefore, our algorithm scales optimally in n, for any fixed accuracy ϵ. Note however that it is not necessarily optimal as a joint function of ϵ and n.

		Grid	RGG
Nearest neighbor		$\mathcal{E} = \Theta(n^2 \log \epsilon^{-1})$	$\mathcal{E} = \Theta(\frac{n^2 \log \epsilon^{-1}}{\log n})$
Route length		$R = \Theta(\sqrt{n})$	$R = \Theta(\sqrt{\frac{n}{\log n}})$
Geographic gossip		$T = \Theta(n \log \epsilon^{-1})$	$T = \Theta(n \log \epsilon^{-1})$
		$\mathcal{E} = \Theta(n^{1.5} \log \epsilon^{-1})$	$\mathcal{E} = \Theta\left(\frac{n^{1.5} \log \epsilon^{-1}}{\sqrt{\log n}}\right)$
Path averaging		$T = \Theta(\sqrt{n} \log \epsilon^{-1})$	$T = \Theta(\sqrt{n \log n} \log \epsilon^{-1})$
		$\mathcal{E} = \Theta(n \log \epsilon^{-1})$	$\mathcal{E} = \Theta(n \log \epsilon^{-1})$

TABLE I

PERFORMANCE OF DIFFERENT GOSSIP ALGORITHMS. T DENOTES AVERAGING TIME (IN GOSSIP ROUNDS) AND $\mathcal{E}$ DENOTES EXPECTED NUMBER OF MESSAGES REQUIRED TO ESTIMATE WITHIN ϵ ACCURACY.

The remainder of this paper is organized as follows: In Section II-A we provide a precise problem statement, describe existing gossip algorithms and their performance metrics. In Section II-C we describe the proposed algorithm and state our main results on its performance. Section III provides the theoretical analysis of the algorithm and proofs of our main results.

II. BACKGROUND AND SETUP

A. Model

1) Time model: We use the asynchronous time model [3], which is well-matched to the distributed nature of sensor networks. In particular, we assume that each sensor has an independent clock whose "ticks" are distributed as a rate λ Poisson process. However, our analysis is based on measuring time in terms of the number of ticks of an equivalent single virtual global clock ticking according to a rate $n\lambda$ Poisson process. An exact analysis of the time model can be found in [3]. Our analysis is based on measuring time in terms of the number of ticks of this (virtual) global clock and we will refer to the time between two consecutive clock ticks as one timeslot.

Throughout this paper we will be interested in minimizing the number of messages without worrying about delay. We can therefore adjust the length of the timeslots relative to the communication time so that only one packet exists in the network at each timeslot with high probability. Note that this assumption is made only for analytical convenience; in a practical implementation, several packets might co-exist in the network, but the associated congestion control issues are beyond the scope of this work.

2) Distributed averaging: At time-slot $k = 0, 1, 2 \ldots$, each node $i = 1, \ldots, n$ has an estimate $x_i(k)$ of the global average, and we use $x(k)$ to denote the n-vector of these estimates. The ultimate goal is to drive the estimate $x(k)$ to the vector of averages $\bar{x}_{\text{ave}}\vec{1}$, where $\bar{x}_{\text{ave}} := \frac{1}{n}\sum_{i=1}^{n} x_i(0)$, and $\vec{1}$ is an n-vector of ones.

For the algorithms of interest to us, the quantity $x(k)$ for $k > 0$ is a random vector, since the algorithms are randomized in their behavior. Accordingly, we measure the convergence of $x(k)$ to $\bar{x}_{\mathrm{ave}}$ in the following sense [9], [3]:

Definition 1: Given $\epsilon > 0$, the ϵ-averaging time is the earliest time at which the vector $x(k)$ is ϵ close to the normalized true average with probability greater than $1 - \epsilon$:

$$T_{ave}(n,\epsilon) = \sup_{x(0)} \inf_{k=0,1,2\ldots} \left\{ \mathbb{P}\left(\frac{\|x(k) - x_{ave}\vec{1}\|}{\|x(0)\|} \geq \epsilon \right) \leq \epsilon \right\}, \tag{1}$$

where $\|\cdot\|$ denotes the ℓ_2 norm. Note that this is essentially measuring a rate of convergence in probability.

We compare algorithms in terms of the amount of communication required. More specifically, let $R(k)$ represent the number of one-hop radio transmissions required for a given node to communicate with some other node in one time-slot. In a standard gossip protocol, the quantity $R(k) \equiv R$ is simply a constant, whereas for our protocol, $R(k)$ will be a random variable (with identical distribution for each node). The total communication cost, measured in one-hop transmissions, is given by the random variable

$$\mathcal{C}(n,\epsilon) = \sum_{k=1}^{T_{ave}(n,\epsilon)} R(k) \ . \tag{2}$$

We analyze the expected communication cost $\mathcal{E}(n,\epsilon) = E[R(k)]T_{ave}(n,\epsilon)$.

B. Graph topologies

We will analyze the performance of geographic gossip with path averaging on grids and random geometric graphs. To avoid edge effects our analysis will be performed on the torus (for both grid and random geometric graphs); recent simulation results show that the performance of the proposed algorithm is similar for graphs on the plane.

Random geometric graphs have been popular models for wireless network topologies [8], [13]. A random geometric graph $G(n,r)$ is formed by choosing n node locations uniformly and independently in the unit torus, with any pair of nodes s and t connected if their Euclidean distance is smaller than some transmission radius r It is well known [13], [8], [7] that in order to maintain connectivity and minimize interference, the transmission radius $r(n)$ should scale like $\Theta(\sqrt{\frac{\log n}{n}})$. For the purposes of analysis, we assume that communication within this transmission radius always succeeds.[1]

C. Proposed Algorithm

Geographic Gossip with path averaging: The proposed algorithm combines gossip with greedy geographic routing. One key assumption is that each node knows its location. In addition, each node can learn the geographic locations of its one-hop neighbors using a single transmission per node. Also the nodes will need to know the size of the space they are embedded in. Note that while our analysis is for the unit torus for grid and random geometric topologies, the algorithm can be applied on any set of nodes embedded on some compact region.

[1] However, we note that our proposed algorithm remains robust to communication and node failures.

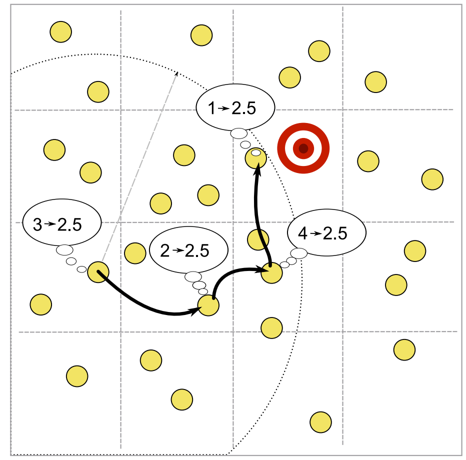


Fig. 1. Example of geographic gossip with path averaging: The node with inital value 3 selects a random position and places a target. Using $(\leftrightarrow, \updownarrow)$-box routing towards that target, all the nodes on the path replace their values with the average of the four nodes.

The algorithm operates as follows: at each timeslot one random node activates and selects a random position (target) on the unit torus (no node needs to be located there). It then creates a packet that contains its current estimate of the average, its position, the number of visited nodes so far (one), the target location, and passes the packet to the nearest neighbor that is closer to the target. As nodes receive the packet, greedily forwarding it towards the target, they add their value to the sum and increase the counter. When the packet reaches its destination node (the one whose nearest neighbors have larger distance to the target compared to it), the destination node computes the average of all the nodes on the path, and reroutes that information backwards on the same route. It is not hard to show [6] that for $G(n,r)$ when r scales like $\Theta(\sqrt{\frac{\log n}{n}})$, greedy forwarding succeeds with high probability over graphs– in other words there are no large 'holes' in the network. We will refer to this whole procedure of routing a message and averaging on a random path as one gossip round which lasts for one timeslot, after which $O(\sqrt{n/\log n})$ nodes will replace their estimates with their joint average (see Fig. 1 for a simple illustration).

Simplified routing: While greedy geographic routing will have good performance in practice, we simplify the analysis by using a slightly modified routing scheme we call $(\leftrightarrow, \updownarrow)$-box routing.

We will divide the unit torus in squares of size $c\frac{\log n}{n}$. It is well known [8], [13] that for a suitable constant c, each of these squares will contain one or more nodes with high probability (w.h.p.). In the appendix we prove a slightly stronger regularity condition, that in fact the number of nodes in each square will be $\Theta(\log n)$ nodes w.h.p. We will refer to random geometric graphs that satisfy this condition as *regular geometric graphs* and prove our results when under this regularity assumption.

By construction, the transmission radius $r(n)$ is selected so that a node can pass messages to all the nodes in the four squares adjacent to its own. In our analysis, when a node wakes up, it chooses uniformly at random its initial direction : horizontal or vertical. $(\leftrightarrow, \updownarrow)$-box routing operates by passing messages from square to square, first horizontally and then vertically (($\updownarrow, \leftrightarrow$)-box routing routs vertically first). Within each square the particular node is selected randomly (see Fig. 1). $(\leftrightarrow, \updownarrow)$-box routing resembles greedy geographic routing but guarantees certain regularity that simplifies the proof. A practical implementation of this routing scheme would require the nodes to know the location and the size of the boxes, which can be determined by the total number of nodes and their own location.

The use of this routing scheme should be seen as a proving technique (essentially binning the random geometric graph on a grid) rather than an algorithm to use in practice. Under this mapping, the performance analysis of gossip with $(\leftrightarrow, \updownarrow)$-box routing on a random geometric graphs can first rely on the analysis of the same protocol on a grid topology.

III. Analysis

A. Averaging scheme and eigenvalues.

Let $x(t)$ denote the vector of estimates of the global averages after the tth gossip round and $x(0)$ the vector of initial measurements. Any gossip algorithm can be described by an equation of the form

$$x(t+1) = W(t)x(t), \tag{3}$$

where $W(t)$ is the averaging matrix over the tth time round. We start with the result of Boyd et al [2], that states that if the matrices $W(t)$ are selected i.i.d. from a family of matrices that have the following three properties:

1) $\vec{1}^T W(t) = \vec{1}$ (the random matrices preserve the total sum of the numbers),
2) $W(t)\vec{1} = \vec{1}$ (the random matrices have the all-one vector as a fixed point),
3) $\rho(\mathcal{W} - \vec{1}\vec{1}^T/n) < 1$, where ρ denotes the spectral radius and where $\mathcal{W} := \mathbb{E}[W(1)] = \mathbb{E}[W(t)]$ is the expectation of the averaging matrix,

then $x(t)$ converges in expectation to $\bar{x}_{\text{ave}}$. Let μ_2 be the second largest eigenvalue in magnitude of $\mathbb{E}[W(t)^T W(t)]$ (the largest one being 1). If $\mu_2 < 1$, then $x(t)$ converges in second moment as well.

In our scheme, $\mathcal{W}$ is symmetric and doubly stochastic hence $\rho(\mathcal{W} - \vec{1}\vec{1}^T/n) = \lambda_2(\mathcal{W})$, where $\lambda_2(\mathcal{W})$ is the second largest eigenvalue in magnitude of $\mathcal{W}$. Moreover, at any time t, $W(t)$ is a symmetric projection matrix so $\mathbb{E}[W(t)^T W(t)] = \mathbb{E}[W(t)] = \mathcal{W}$ as well, hence $\mu_2 = \lambda_2(\mathcal{W})$. So if $\lambda_2(\mathcal{W}) < 1$, then $x(t)$ converges to $\bar{x}_{\text{ave}}$ in expectation and in second moment. Furthermore, the value of $\lambda_2(\mathcal{W})$ controls the speed of convergence; a straightforward extension of the proof of Boyd et al [3] from the case of pairwise averaging matrices to the case of symmetric projection matrices yields the following key bound on the averaging time in terms of the spectrum of the expected averaging matrix:

$$T_{ave}(\epsilon, \mathcal{W}) \leq \frac{3 \log \epsilon^{-1}}{\log(\frac{1}{\lambda_2(\mathcal{W})})} \leq \frac{3 \log \epsilon^{-1}}{1 - \lambda_2(\mathcal{W})}. \tag{4}$$

(There is also a lower bound of the same order.)

Consequently, the rate at which the *spectral gap* $1 - \lambda_2(\mathcal{W})$ approaches zero, as n increases controls the averaging time. For example, in the case of a complete graph and uniform pairwise gossiping, one can show that $\lambda_2(\mathcal{W}) = 1 - 1/n$ so, as previously mentioned, this scheme's averaging time is $O(n \log \epsilon^{-1})$. In pairwise gossiping, the convergence time and the number of messages have the same order because there are a constant number R of transmissions per time-slot. Our scheme reaches the same efficiency even though one round uses many messages for the path routing (on average $\sqrt{n}$ messages in the grid, and $\sqrt{n/\log n}$ messages in the regular geometric graph). The main results of this paper are the following bounds on the averaging time and expected communication cost:

Theorem 1 (Averaging time on grids): On a $\sqrt{n} \times \sqrt{n}$ torus grid, the averaging time $T_{ave}(\epsilon, \mathcal{W})$ of geographic gossip with path averaging (using $(\leftrightarrow, \updownarrow)$-box routing) is $O(\sqrt{n} \log(\epsilon^{-1}))$.

Theorem 2 (Averaging time on random geometric graphs.): Consider a random geometric graph $G(n, r)$ on the unit torus with $r(n) = c\sqrt{\frac{\log n}{n}}$. With high probability over graphs, the averaging time $T_{ave}(\epsilon, \mathcal{W})$ of geographic gossip with path averaging (using $(\leftrightarrow, \updownarrow)$-box routing is $O(\sqrt{n \log n} \log(\epsilon^{-1}))$.

Corollary 1 (Expected communication cost): In both cases (grids and regular geometric graphs), the expected communication cost $\mathcal{E}(n, \epsilon)$ for geographic gossip with path averaging is $O(n \log(\epsilon^{-1}))$, where n is the number of nodes.

B. Sketch of the proofs

The main two steps in our proofs are the estimation of the expected averaging matrix $\mathcal{W}$ and the upper bounding of its second largest eigenvalue $\lambda_2(\mathcal{W})$. We first present the techniques in the proof for the grid (Theorem 1) and then refine the argument to establish Theorem 2.

The first step is to obtain good estimates on $\mathcal{W}_{ij}$ for any fixed pair of nodes (i, j). Contrary to standard nearest-neighbor gossip, $\mathcal{W}$ in our models is a dense matrix ($\Theta(n^2)$ nonzero elements), because any pair of nodes can be averaged in some routes. Moreover, close nodes will communicate a lot more often than far away nodes because there are many more possible routes averaging them together (see Fig. 2). To take this phenomenon into account, the first part of the proof consists in counting the routes averaging some fixed pair of nodes (i, j). Placing our nodes on a torus greatly simplifies this counting because the result only depends on the distance and not on their location. We will have to count separately the routes of different length, because the contribution of a route to $\mathcal{W}$ is the inverse of its length. Indeed, if a route contains nodes i and j and is of length ℓ, then $W_{ij} = 1/\ell$. Counting the routes is easy on the grid but very difficult on a random geometric graph with the distributed algorithm we propose. This is why we study models and start by analyzing the grid case, which will greatly help the analysis of the second model based on regular geometric graphs. For random geometric graphs we only obtain lower bounds on the entries of $\mathcal{W}$, which still suffice to guarantee a minimum level of information exchange between the nodes.

To give a sense of the orders of magnitude of the expectation matrices, let us take again the simple example of a complete graph with uniform pairwise gossiping. Nodes i and j communicate with probability $2/n^2$ because there are $n^2/2$ unordered pairs of nodes (with repetitions), and $W_{ij} = 0.5$ when $i \neq j$, which implies that $\mathcal{W}_{ij} = 1/n^2$ for all such pair of nodes (i, j). One key part of our proof is showing that for most[2] pairs of nodes (i, j), $\mathcal{W}_{ij} = \Omega(1/n^{1.5})$ for path averaging on the grid, and $\mathcal{W}_{ij} = \Omega(1/(n^{1.5}\sqrt{\log n}))$ for regular geometric graphs. Since these lower bounds are larger than $1/n^2$, we expect a faster convergence in gossip with path averaging compared to standard gossip. Notice however that each gossip round costs $O(\sqrt{n})$ message transmissions for the grid and $O(\sqrt{\frac{n}{\log n}})$ for the random geometric graph.

The second ingredient is using these lower bounds on the entries of $\mathcal{W}$ to bound $\lambda_2(\mathcal{W})$. It is pretty surprising that one can obtain asymptotically tight bounds on the spectral gap of $\mathcal{W}$ when even determining the exact entries of the matrix is very difficult. The main problem we need to overcome is that far away nodes are not averaged together often. For example, in the grid, consider the extreme case of a distance $\sqrt{n}$ between two nodes: a node i and a node j in a corner of the grid when it is represented centered on i. There are only two routes that will average them: the route that goes from i to j, and the reverse one. These routes are selected with probability $1/n^2$ and $W_{ij} = 1/\sqrt{n}$, implying that $\mathcal{W}_{ij} = 2/n^{2.5}$. *The key idea is to notice that, although they do not directly exchange information often, they both do communicate very often with the node k that is half way between i and j* ($\delta = 0.5$). This node k acts like a diffusion relay between i and j. Every node can actually be seen as a relay for some far away pair of nodes, in such a way that we can appropriately balance the relaying load over all the nodes to obtain our bound. We use the the Poincaré inequality of Diaconis and Stroock [5], [16] to use the geometry of these relays and directly bound the spectral gap of $\mathcal{W}$.

[2] Provided that the distance between i and j is not larger than a fixed constant fraction δ of the maximum possible distance

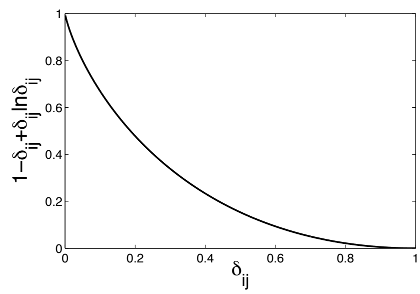


Fig. 2. Behavior of $\mathcal{W}_{i,j}$ as a function of the distance in norm 1 between i and j: $f(\delta_{ij}) = 1 - \delta_{ij} + \delta_{ij} \ln \delta_{ij}$.

C. Grid: proof of Theorem 1

1) Evaluating $\mathcal{W}$: We work on a torus of size $\sqrt{n} \times \sqrt{n}$. To each route r, we assign a generalized gossip $n \times n$ matrix $W^{(r)}$ that averages the current estimates of the nodes on the route. At each iteration t, we choose one of the two shorstest routes starting from a random node I and ending at a random destination node J (see Figure 3(a)). The pair (I, J) is uniformly chosen, as well as the first direction: we flip a coin and depending on the result of the coin toss, we choose to route first horizontally and next vertically, which we will denote by $(\leftrightarrow, \updownarrow)$, or conversely first vertically and next horizontally $(\updownarrow, \leftrightarrow)$. Consequently, at iteration t, $W(t) = W^{r(t)}$, where $r(t)$ was randomly chosen. We call R the route random variable. As we choose the shortest route, the maximum number of nodes a route can contain is $\sqrt{n}$ if $\sqrt{n}$ is odd, $\sqrt{n}+1$ if $\sqrt{n}$ is even, which can be written as $2\lfloor\sqrt{n}/2\rfloor + 1$ in short. We need to define the shortest distance on a torus. To this end, we introduce a torus short absolute value $|.|_\mathcal{T}$ and a torus short norm 1 $\|.\|_1$. For any algebraic value x on a one dimensional torus (circle) and any vector i on a two dimensional torus,

$$\begin{aligned} |x|_\mathcal{T} &= \min(|x|, |x-\sqrt{n}|, |x+\sqrt{n}|) \\ \|i\|_1 &= |i_x|_\mathcal{T} + |i_y|_\mathcal{T}. \end{aligned}$$

We call $\ell_{ij} = \|j - i\|_1$ the L_1 distance between nodes i and j. The shortest routes between I and J have $\alpha = \ell_{IJ} + 1 = |J_x - I_x|_\mathcal{T} + |J_y - I_y|_\mathcal{T} + 1$ nodes to be averaged, thus the non-zero coefficients of their corresponding matrices W are all equal to $1/\alpha$.

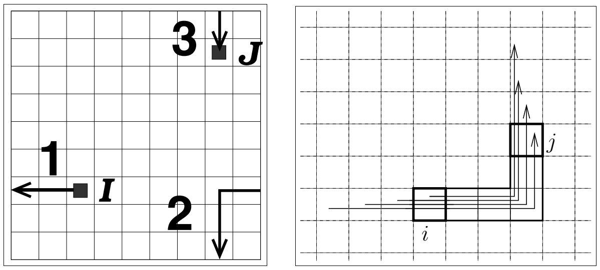
Lemma 1 (Expected $\mathcal{W}$ on the grid): The entries of the expectation of the gossip matrices averaging along the way verify

$$\mathcal{W}_{i,j} \geq \frac{2}{n^2}\left(\sqrt{n} - \ell_{ij} - \ell_{ij}\ln\frac{\sqrt{n}}{\ell_{ij}}\right). \tag{5}$$

Moreover, with $\delta_{ij} = \|j - i\|_1/\sqrt{n}$ denoting the distance between nodes i and j normalized by $\sqrt{n}$,

$$\mathcal{W}_{i,j} \geq \frac{2(1 - \delta_{ij} + \delta_{ij}\ln\delta_{ij})}{n\sqrt{n}}.$$

The intuition behind this result is that far away nodes are less likely to be averaged out than neighboring ones (see Figure 2).



(a) Shortest route (b) Number of routes

Fig. 3. (a) Shortest $(\leftrightarrow, \updownarrow)$-path from I to J on a torus. (b) $\ell_{ij} = 5$, $\ell = 9$ nodes. There are $\ell - \ell_{ij} = 9 - 5 = 4$ possible routes with exactly ℓ nodes going through node i then through node j. We admit only routes going horizontally first then vertically .

Proof: Observing that $\mathbb{E}[W^{(R)}|(\leftrightarrow, \updownarrow)] = \mathbb{E}[W^{(R)}|(\updownarrow, \leftrightarrow)]$ because the route from a node I to a node J horizontally first has the same nodes as the route from J to I vertically first, we get

$$\begin{aligned}\mathcal{W} &= \mathbb{E}[W^{(R)}] \\ &= \frac{1}{2}\mathbb{E}[W^{(R)}|(\leftrightarrow, \updownarrow)] + \frac{1}{2}\mathbb{E}[W^{(R)}|(\updownarrow, \leftrightarrow)] \\ &= \mathbb{E}[W^{(R)}|(\leftrightarrow, \updownarrow)].\end{aligned}$$

So, for a given pair of nodes (i, j), we can compute the (i, j)th entry of the matrix expectation $\mathcal{W}$ by systematically routing first horizontally. Only the $(\leftrightarrow, \updownarrow)$-routes which contain both these two nodes i and j will have a non-zero contribution in $\mathcal{W}_{ij}$. Pick such a route r, and call $s(r)$ its starting node, $d(r)$ its destination node, and $\ell(r) = \ell_{s(r)d(r)} + 1$ its number of nodes. The (i, j)th entry of the corresponding averaging matrix is $W^{(r)}_{i,j} = 1/\ell(r)$. We call $\mathcal{R}^{\ell}_{ij}$ the set of $(\leftrightarrow, \updownarrow)$-routes with ℓ nodes passing by node i and by node j, and denote $x^+ = \max(x, 0)$. It is not hard to see that $(\ell - \ell_{ij})^+$ is the number of routes of length ℓ passing by i first and j next (see 3(b)), so $|\mathcal{R}^{\ell}_{ij}| = 2(\ell - \ell_{ij})^+$. We thus have for any $i \neq j$:

$$\begin{aligned}\mathcal{W}_{i,j} &= \textstyle\sum_r W^{(r)}_{i,j}\mathbb{P}[R = r] = \frac{1}{n^2}\sum_r W^{(r)}_{i,j} \\ = \frac{1}{n^2} & \textstyle\sum_{\ell=\ell_{ij}+1}^{2\lfloor\frac{\sqrt{n}}{2}\rfloor+1} \frac{|\mathcal{R}^{\ell}_{ij}|}{\ell} = \frac{2}{n^2}\sum_{k=\ell_{ij}+1}^{2\lfloor\frac{\sqrt{n}}{2}\rfloor+1} \frac{k-\ell_{ij}}{k},\end{aligned}$$

from which we can deduce that for $i \neq j$

$$\begin{aligned}\mathcal{W}_{i,j} &\leq \frac{2}{n^2}\int_{\ell_{ij}+1}^{\sqrt{n}+2} \frac{x - \ell_{ij}}{x}dx \\ &= \frac{2}{n^2}\left(\sqrt{n} - \ell_{ij} + 1 - \ell_{ij}\ln\frac{\sqrt{n}+2}{\ell_{ij}+1}\right) \\ \mathcal{W}_{i,j} &\geq \frac{2}{n^2}\int_{\ell_{ij}}^{\sqrt{n}} \frac{x - \ell_{ij}}{x}dx \\ &= \frac{2}{n^2}\left(\sqrt{n} - \ell_{ij} - \ell_{ij}\ln\frac{\sqrt{n}}{\ell_{ij}}\right).\end{aligned}$$

$\mathcal{W}_{i,j}$ decreases from $\frac{2}{n\sqrt{n}}$ to $o(\frac{1}{n^2})$ as a function of ℓ_{ij}. To get a normalized expression with respect to $\sqrt{n}$, we use the coefficient δ_{ij} defined in the statement of Lemma 1.

$$\begin{aligned}&\frac{2}{n\sqrt{n}}\left(1 - \delta_{ij} + \delta_{ij}\ln\delta_{ij}\right) \leq \mathcal{W}_{i,j} \leq \\ &\qquad \frac{2}{n\sqrt{n}}\left(1 - \delta_{ij} + \delta_{ij}\ln\delta_{ij} + \frac{1}{\sqrt{n}} - \delta_{ij}\ln\frac{\sqrt{n}+2}{\sqrt{n}+\frac{1}{\delta_{ij}}}\right).\end{aligned}$$

This establishes the claim. In particular, if $\delta_{ij} = 1/2$, then $\mathcal{W}_{i,j} \sim \frac{1-\ln 2}{n\sqrt{n}}$.

■

2) Bounding $\lambda_2(\mathcal{W})$: We need now to upperbound the second largest eigenvalue in magnitude of $\mathcal{W}$, or equivalently, the relaxation time $1/(1 - \lambda_2(\mathcal{W}))$.

Lemma 2 (Relaxation time):

$$\frac{1}{1 - \lambda_2(\mathcal{W})} = O(\sqrt{n}). \tag{6}$$

Proof: Since $\mathcal{W}$ is symmetric, all its eigenvalues are real. The sum of all the entries along the lines of $\mathcal{W}$ without counting the diagonal element is $O(1/\sqrt{n})$, whereas the diagonal elements are $\Theta(1)$, so by Gershgorin bound [4], all the eigenvalues of $\mathcal{W}$ are positive. In particular, the second largest eigenvalue in magnitude is the second largest positive eigenvalue.

We now use Poincaré inequality [5] (see also [4], p. 212-213) to bound the spectral gap of $\mathcal{W}$: Let P denote the transition matrix of an irreducible reversible Markov chain with stationary distribution π. In order to apply this theorem, for each ordered pair of distinct nodes (i, j), choose one and only one path $\gamma_{ij} = (i, i_1, \ldots, i_m, j)$ between i and j, and define

$$|\gamma_{ij}| = \frac{1}{\pi(i)p_{ii_1}} + \frac{1}{\pi(i_1)p_{i_1i_2}} + \ldots + \frac{1}{\pi(i_m)p_{i_mj}}. \tag{7}$$

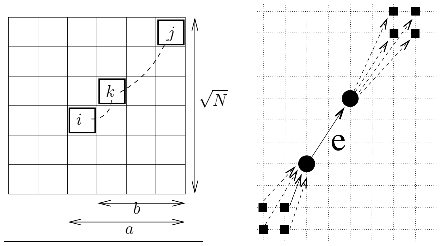
The Poincaré coefficient is

$$\kappa = \max_{\text{edge } e}\sum_{\gamma_{ij}\ni e}|\gamma_{ij}|\pi(i)\pi(j), \tag{8}$$

and the theorem states that $\lambda_2(P) \leq 1 - \frac{1}{\kappa}$.

$\mathcal{W}$ is a doubly stochastic matrix, and therefore can be regarded as a symmetric transition matrix of a fully connected Markov chain. It is thus irreducible and reversible with uniform stationary distribution $\pi(i) = 1/n$ for all i. For each ordered pair of distinct nodes (i, j), we choose a 2-hop path γ_{ij} from i to j stopping by node k chosen to be located approximatively half way between nodes i and j. To be more precise, we define direction functions σ_x and σ_y, where $\sigma_x(i, j) = 1$ (respectively, $\sigma_y(i, j) = 1$) if the horizontal (resp., vertical) part of the route from i to j goes to the right (resp., up) and $\sigma_x(i, j) = -1$ (resp., $\sigma_y(i, j) = -1$) if it goes left (resp., down). The coordinates of node k in the torus are

$$\begin{aligned}k_x &= \left(i_x + \sigma_x(i, j)\lfloor\frac{|j_x - i_x|_{\mathcal{T}}}{2}\rfloor\right) \pmod{\sqrt{n}} \\ k_y &= \left(i_y + \sigma_y(i, j)\lfloor\frac{|j_y - i_y|_{\mathcal{T}}}{2}\rfloor\right) \pmod{\sqrt{n}}.\end{aligned}$$



(a) Longest hop (b) 8 possible paths

Fig. 4. (a) a is the maximum horizontal hop size between i and j; $a = \lfloor\frac{\sqrt{n}}{2}\rfloor+1$. b is the maximal horizontal hop size between k and j; $b = \lfloor\frac{a}{2}\rfloor+1$. The largest hop length in norm 1 is then $\ell_{kj} = 2(b-1) = 2\lfloor\frac{a}{2}\rfloor$. We conclude that $\ell_{kj} \leq \frac{\sqrt{n}}{2} + 1$. Here $\sqrt{n} = 6$, $a = 4$, $b = 3$ and $\max \ell = 4$. (b) For a given edge e there are 8 paths including e.

For each path we have:

$$\begin{aligned} |\gamma_{ij}| &= \frac{1}{\pi(i)\mathcal{W}_{i,k}} + \frac{1}{\pi(k)\mathcal{W}_{k,j}} \\ &= n\left(\frac{1}{\mathcal{W}_{i,k}} + \frac{1}{\mathcal{W}_{k,j}}\right) \\ &\leq \frac{2n^2\sqrt{n}}{1-\ln 2-\epsilon}, \end{aligned} \tag{9}$$

for any small $\epsilon \geq 0$ and with n large enough for $2(1+\ln 2)/\sqrt{n}$ to be smaller than ϵ. Inequality (9) comes from our choice of intermediate node k at mid-distance between nodes i and j, which implies that each edge of γ_{ij} is not longer than $\ell \leq \frac{\sqrt{n}}{2} + 1$ in L_1 distance (see Figure 4(a)), and from the decrease of $\mathcal{W}_{i,j}$ with ℓ_{ij}. Inequality (11) with $\ell_{ij} = \frac{\sqrt{n}}{2} + 1$ gives a useful lower bound to the entries of $\mathcal{W}$, which leads to our upper bound (9) of $|\gamma_{ij}|$:

$$\mathcal{W}_{i,j} \geq \frac{1}{n\sqrt{n}}\left(1 - \ln 2 - \frac{2}{\sqrt{n}}(1+\ln 2)\right).$$

We can now compute the Poincaré coefficient:

$$\kappa = \max_e \sum_{\gamma_{ij}\ni e} |\gamma_{ij}|\pi_i\pi_j = \frac{1}{n^2}\max_e \sum_{\gamma_{ij}\ni e} |\gamma_{ij}|. \tag{10}$$

In our construction, we wisely balanced the relaying load over all the edges of the graph: an edge e belongs to at most 8 paths! Note that even though there are only $\Theta(n)$ edges on the original grid (and $\Theta(n^2)$ paths), *the graph that corresponds to the markov chain defined by $\mathcal{W}$ has an edge between any two nodes that can be averaged together and therefore has $\Theta(n^2)$ edges*. This is the key reason that allows a constant number of paths per edge and leads to fast averaging.

In particular, e can be followed by 4 different edges in a path and can be preceded by 4 different edges as well (see Figure 4(b)). Combining (9) and (10), we get that for n large enough: $\kappa \leq \frac{16}{1-\ln 2-\epsilon}\sqrt{n}$. As a result, for n large enough, $\lambda_2 \leq 1 - \frac{1-\ln 2-\epsilon}{16\sqrt{n}}$, which yields Lemma 2. ∎

D. Regular geometric graph: proof of Theorem 2

1) Evaluating $\mathcal{W}$: In this section, we work on a regular geometric graph lying on a torus. In other words, we have k boxes forming a torus grid like in previous section and $k = \lceil\sqrt{n/(\alpha\log n)}\rceil^2 \simeq n/(\alpha\log n)$, for some $\alpha > 2$. We first need to show that the number of nodes in each box is $\Theta(\log n)$ with high probability:

Lemma 3 (Regularity of random geometric graphs): Consider a random geometric graph with n nodes and partition the unit square in boxes of size $c\frac{\log n}{n}$. Then, all the boxes contain $\Theta(\log n)$ nodes, with high probability as $n \to \infty$.

The proof of this lemma follows the standard [8], [13] approach of using Chernoff lower and upper bounds on the number of nodes on a particular box a union bound to make a claim for all k boxes.

We use the $(\leftrightarrow, \updownarrow)$-box routing scheme presented in Section II-C and we can now establish a lower bound on the entries of the expected averaging matrix:

Lemma 4 (Expected $\mathcal{W}$ on the regular geometric graph): The expectation of the gossip matrices averaging along the way on regular geometric graphs verify for any nodes i and j belonging to different boxes

$$\mathcal{W}_{i,j} \geq \frac{4a}{b^2}\frac{2}{n^2}\left(\sqrt{k} - \ell_{ij} - \ell_{ij}\ln\frac{\sqrt{k}}{\ell_{ij}}\right), \tag{11}$$

where ℓ_{ij} is the L_1 distance between boxes $b(i)$ and $b(j)$.

We omit this proof due to space constraints. Note that there are only a few modifications on the argument for the grid that are sufficient to establish this lemma. The idea is to notice that for any route $r = (r_1, r_2, \cdots, r_\ell)$, we can attribute a box route $\widetilde{r}$ consisting of the boxes the nodes of r belong to. We are therefore dealing with a case very similar to a grid of size $k = n/(\alpha\log n)$ boxes but the routes are no longer uniform since each box contains a different number of nodes. However using the regularity lemma, the number of nodes on each box cannot vary too much and this suffices to establish the lower bound.

2) Bounding $\lambda_2(\mathcal{W})$:

Lemma 5 (Relaxation time RGG):

$$\frac{1}{1-\lambda_2(\mathcal{W})} = O(\sqrt{n\log n}). \tag{12}$$

Proof: We want use the weighted path upper bound again: for each ordered pair of distinct nodes (i,j), we have to define a path. There are two additional difficulties compared to the grid case.

First we have to make sure that we can refine the trick of creating two-hop paths from box level to node level without overloading some edges. More precisely, an edge should not be used more than a constant number of times (it was 8 for the grid). It is actually possible to design such paths because the number of nodes in the boxes does not vary more than by a constant multiplicative factor b/a from one box to the other. Let's assume for simplicity that every box has $\log n$ nodes. There are $(\log n)^2$ paths to find between nodes of box

1 and nodes of box 3, but happily enough there are $(\log n)^2$ edges between box 1 and box 2 and also between box 2 and box 3 and these edges are uniformly shared out among the nodes. Therefore, the box-path (box 1, box 2, box 3) can correspond to $(\log n)^2$ node-paths all using different edges. This path allocation technique can easily be extended to cases where the boxes do not have the same number of nodes by using some edges at most $\lceil b/a \rceil$ times each time it is used at a box level.

Secondly, nodes that are in the same box do not average together enough ($\mathcal{W}_{ij} \leqslant 2/(an^2)$). However they communicate often with nodes in their neighboring boxes. Just as middle nodes are diffusion relays for far away nodes, here, neighboring box's nodes are relays for nodes that live in the same box. So formally, if node i and node j are in the same box, we design the path from i to j to be a two hop path stopping at a node situated in the box above their box. By sharing fairly the available relays, the short vertical edges might be used in $\lceil b/a \rceil$ extra paths.
We can thus construct paths for any pair of nodes that will use at most $9\lceil b/a \rceil$ times each edge. The rest of the proof is identical to the grid proof.

For each path we have:

$$\begin{aligned} |\gamma_{ij}| &= \frac{1}{\pi(i)\mathcal{W}_{i,k}} + \frac{1}{\pi(k)\mathcal{W}_{k,j}} \\ &= n\left(\frac{1}{\mathcal{W}_{i,k}} + \frac{1}{\mathcal{W}_{k,j}}\right) \\ &\leq cn^2\sqrt{n \log n}, \end{aligned} \tag{13}$$

for some constant c. Inequality 13 was obtained with the same reasoning as in the grid. We therefore conclude, using the Poincaré coefficient argument that $\kappa \leq 9\lceil \frac{b}{a} \rceil c\sqrt{n \log n}$. As a result, for n large enough, and some constant c'. $\lambda_2 \leq 1 - \frac{1}{c'\sqrt{n \log n}}$, which yields the lemma. ■

IV. Conclusions

We introduced a novel gossip algorithm for distributed averaging. The proposed algorithm operates in a distributed and asynchronous manner on locally connected graphs and requires an order-optimal number of communicated messages. The execution of geographic gossip with path averaging relies on knowledge of geographic locations; this location information is independently useful and likely to exist in many application scenarios. The key idea that makes path averaging so efficient is the opportunistic combination of routing and averaging.

We believe that the idea of greedily routing towards a randomly selected target (and possibly processing information on the routed paths) is a very useful primitive for designing message-passing algorithms on networks that have planar geometry. Other than computing linear functions, such path-processing algorithms can be designed for information dissemination or more general message passing computations such as marginal computations or MAP estimates for probabilistic graphical models. Processing and forwarding the messages on random paths can avoid the diffusive nature of random walks and accelerate the convergence of message-passing. We plan to investigate such protocols in future work.

References

[1] M. Alanyali, V. Saligrama, and O. Savas. A random-walk model for distributed computation in energy-limited networks. In *Proceedings of the 1st Workshop on Information Theory and its Applications*, San Diego, CA, 2006.

[2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Analysis and optimization of randomized gossip algorithms. In *Proceedings of the 43rd Conference on Decision and Control (CDC 2004)*, 2004.

[3] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. In *IEEE Transactions on Information Theory, Special issue of IEEE Transactions on Information Theory and IEEE/ACM Transactions on Networking*, 2006.

[4] P. Brémaud. *Markov Chains. Gibbs Fields, Monte Carlo Simulation, and Queues*. Springer, 1999.

[5] P. Diaconis and D. Stroock. Geometric bounds for eigenvalues of markov chains. In *Annals of Applied Probability*, volume 1, 1991.

[6] A. G. Dimakis, A. D. Sarwate, and M. J. Wainwright. Geographic gossip: efficient aggregation for sensor networks. In *ACM/IEEE Symposium on Information Processing in Sensor Networks*, 2006.

[7] A. E. Gamal, J. Mammen, B. Prabhakar, and D. Shah. Throughput-delay trade-off in wireless networks. In *Proceedings of the 24th Conference of the IEEE Communications Society (INFOCOM 2004)*, 2004.

[8] P. Gupta and P. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, March 2000.

[9] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proc. IEEE Conference of Foundations of Computer Science, (FOCS)*, 2003.

[10] W. Li and H. Dai. Location-aided fast distributed averaging. In *IEEE Statistical Signal Processing Workshop*, 2007.

[11] C. Moallemi and B. V. Roy. Consensus propagation. Technical report, Stanford University, June 2005.

[12] D. Mosk-Aoyama and D. Shah. Information dissemination via gossip: Applications to averaging and coding. http://arxiv.org/cs.NI/0504029, April 2005.

[13] M. Penrose. *Random Geometric Graphs*. Oxford studies in probability. Oxford University Press, Oxford, 2003.

[14] M. Rabbat, J. Haupt, A.Singh, and R. Nowak. Decentralized compression and predistribution via randomized gossiping. In *ACM/IEEE Conference on Information Processing in Sensor Networks (IPSN'06)*, April 2006.

[15] V. Saligrama, M. Alanyali, and O. Savas. Distributed detection in sensor networks with packet losses and finite capacity links. In *IEEE Transactions on Signal Processing, to appear*, 2007.

[16] A. Sinclair. Improved bounds for mixing rates of markov chains and multicommodity flow. In *Combinatorics, Probability and Computing*, volume 1, 1992.

[17] D. Spanos, R. Olfati-Saber, and R. Murray. Distributed Kalman filtering in sensor networks with quantifiable performance. In *2005 Fourth International Symposium on Information Processing in Sensor Networks*, 2005.

[18] J. Tsitsiklis. *Problems in decentralized decision-making and computation*. PhD thesis, Department of EECS, MIT, 1984.

[19] L. Xiao, S. Boyd, and S. Lall. A scheme for asynchronous distributed sensor fusion based on average consensus. In *2005 Fourth International Symposium on Information Processing in Sensor Networks*, 2005.