ON THE CONVERGENCE OF AVERAGE CONSENSUS WITH GENERALIZED METROPOLIS-HASTING WEIGHTS

Valentin Schwarz, Gabor Hannak, and Gerald Matz

Institute of Telecommunications, Vienna University of Technology
Gusshausstrasse 25/389, 1040 Vienna, Austria; email: {vschwarz,ghannak,gmatz}@nt.tuwien.ac.at

ABSTRACT

Average consensus is a well-studied method for distributed averaging. The convergence properties of average consensus depend on the averaging weights. Examples for commonly used weight designs are Metropolis-Hastings (MH) weights and constant weights. In this paper, we provide a complete convergence analysis for a generalized MH weight design that encompasses conventional MH as special case. More specifically, we formulate sufficient and necessary conditions for convergence. A main conclusion is that AC with MH weights is guaranteed to converge unless the underlying network is a regular bipartite graph.

Index Terms— Average consensus, wireless sensor networks, distributed algorithms

1. INTRODUCTION

Wireless sensor networks (WSN) have attracted enormous interest over the past decade, with applications ranging from environmental monitoring to distributed localization and tracking (see, e.g., [1]). Distributed algorithms are particularly important tools in this context. Distributed averaging is one of the most fundamental methods; here, the goal is that all nodes compute in a distributed manner the arithmetic mean of a collection of numbers (typically the sensor measurements). distributed manner such that it is then available at all nodes. Overviews of existing iterative averaging algorithms problem are given in [2–4]. One of these algorithms is referred to as *average consensus* (AC) and was originally proposed in [5].

It is known that the convergence of AC is determined by the graph topology and by the underlying weight design. Latter has been topic of considerable research, e.g. [6] and [7] optimize the per-step convergence speed, where the second one takes the statistics of the measurements into account. Metropolis-Hastings (MH) weight design [8–10] is a particularly popular approach. There are two variants of MH that differ in their convergence properties [10–12]. According to [10, 13], convergence is guaranteed as long as the underlying communication graph is connected ”in the long term”. This condition is essential for AC in mobile WSN [14].

In this paper, we introduce a generalized MH (GMH) design for undirected graphs that comprises the two conventional variants as special cases. Based on GMH, we provide a unifying convergence analysis in terms of necessary and sufficient convergence conditions. We show that convergence is guaranteed unless the underlying communication graph is regular and bipartite. Numerical results illustrate our theoretical findings.

Funded by WWTF Grant ICT08-044, FWF Grant S10606, and by the EU FP7 project NEWCOM# (GA 318306).

2. GRAPH THEORY BACKGROUND

We consider a network of I agents that is modeled by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$; here, $\mathcal{V}$ is the node (vertex) set (i.e., the set of agents) and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of undirected edges (the communication links between agents). For simplicity we assume $\mathcal{V} = \{1, \ldots, I\}$ in what follows. Furthermore, we will use an enumeration of the edge set (the edge index is denoted by l). The $I \times I$ adjacency matrix $\mathbf{A}$ characterizes the graph topology. It is defined as $[\mathbf{A}]_{ij} = 1$ if $(i, j) \in \mathcal{E}$ and $[\mathbf{A}]_{ij} = 0$ otherwise. The neighbor set $\mathcal{N}_i$ of node i is the set of all nodes that form an edge with node i. The degree d_i of node i equals the number of its neighbors; formally, $d_i = \sum_{j=1}^{I} [\mathbf{A}]_{ij}$ or, in vector notation, $\mathbf{d} = \mathbf{A1}$, where $\mathbf{1}$ is the all-ones vector. We denote the largest degree by $d_{\max} = \max_i d_i$. A graph is called d-regular if all nodes have the same number d of neighbors, i.e., $\mathbf{d} = d\mathbf{1}$. The $I \times |\mathcal{E}|$ incidence matrix $\mathbf{B}$ describes the relation of nodes and edges; assuming that the lth edge consists of nodes i and j, $[\mathbf{B}]_{il} = 1$ and $[\mathbf{B}]_{jl} = -1$, i.e., in each column one entry equals plus one, one entry equals minus one, and all other entries are zero. Another matrix which is used to characterize graphs is the Laplacian, which is given by $\mathbf{L} = \mathbf{B}\mathbf{B}^{\mathrm{T}} = \mathrm{diag}\{\mathbf{d}\} - \mathbf{A}$; here, $\mathrm{diag}\{\mathbf{d}\}$ is the diagonal matrix whose diagonal is given by the node degree vector $\mathbf{d}$. A graph is called connected if there exists a path (i.e., a sequence of edges) between any two nodes, which is the case if and only if the Laplacian has no more than one eigenvalue that equals zero. A graph is termed bipartite if it can be separated into two groups of nodes such that no two nodes in each group form an edge. It can be shown that a graph is bipartite if and only if the smallest eigenvalue of the signless Laplacian [15] $|\mathbf{L}| = |\mathbf{B}||\mathbf{B}|^{\mathrm{T}} = \mathrm{diag}\{\mathbf{d}\} + \mathbf{A}$ is zero (here, the absolute value $|\cdot|$ is to be understood in an element-wise manner).

AC with symmetric weights can be characterized by a weighted graph in which each edge (i, j) is associated with a strictly positive weight w_{ij}. For weighted graphs, the nonzero elements of the adjacency matrix $\tilde{\mathbf{A}}$ are defined by $[\tilde{\mathbf{A}}]_{ij} = w_{ij}$. The weighted incidence matrix is given by $\tilde{\mathbf{B}} = \mathbf{B}\,\mathrm{diag}\{\sqrt{\mathbf{w}}\}$, where all weights w_{ij} are arranged into a length-$|\mathcal{E}|$ vector $\mathbf{w}$ and the square-root is to be understood element-wise. The weighted Laplacian reads $\tilde{\mathbf{L}} = \tilde{\mathbf{B}}\tilde{\mathbf{B}}^{\mathrm{T}}\,\mathrm{diag}\{\tilde{\mathbf{d}}\} - \tilde{\mathbf{A}}$ with the weighted degree vector $\tilde{\mathbf{d}} = \tilde{\mathbf{A}}\mathbf{1}$. The signless weighted Laplacian is given by $|\tilde{\mathbf{L}}| = \mathrm{diag}\{\tilde{\mathbf{d}}\} + \tilde{\mathbf{A}}$.

WSN are often modeled in terms of random geometric graphs [16]. Here, the I nodes are independently and uniformly distributed in a region $\mathcal{A}$ and communicate with nodes that lie within a certain (communication) range r. To avoid boundary effects, the region is often modeled as a torus.

3. GENERALIZED MH WEIGHTS

We assume a scenario in which agent i obtains an initial value s_i. The goal is to compute at each node the average of the measured values, i.e., $\bar{s} = \frac{1}{I}\sum_i s_i$. Let $x_i[k]$ denote the estimate of $\bar{s}$ at node i and time k. With the initialization $x_i[0] = s_i$, AC performs the following local updates:

$$x_i[k+1] = w_{ii}\, x_i[k] + \sum_{j \in \mathcal{N}_i} w_{ij}\, x_j[k].$$

Here, $w_{ij} = w_{ji}$ is a weight associated with the edge (i, j) and w_{ii} quantifies the persistence of the local node estimate. Defining the vector $\mathbf{x}[k] = (x_1[k] \ldots x_I[k])^T$, the AC update can be rewritten as

$$\mathbf{x}[k+1] = \mathbf{W}\mathbf{x}[k], \qquad (1)$$

where the symmetric weight matrix $\mathbf{W}$ is defined by $[\mathbf{W}]_{ij} = w_{ij}$ (we use the convention that $w_{ij} = 0$ unless $(i, j) \in \mathcal{E}$ or $i = j$). To guarantee convergence of AC, the following conditions need to be satisfied [10]: $\mathbf{W}\mathbf{1} = \mathbf{1}$ (equivalently, $\mathbf{1}^{\mathrm{T}}\mathbf{W} = \mathbf{1}^{\mathrm{T}}$ due to the symmetry of $\mathbf{W}$) and $\rho\big(\mathbf{W} - \frac{1}{I}\mathbf{1}\mathbf{1}^{\mathrm{T}}\big) < 1$, where $\rho(\cdot)$ is the spectral radius [17]. The first condition implies that $\mathbf{W}$ has to be doubly stochastic, which can be ensured implicitly by choosing the weights on the diagonal as $w_{ii} = 1 - \sum_{j \in \mathcal{N}_i} w_{ij}$. Using the weighted Laplacian, the AC weight matrix can then be written as $\mathbf{W} = \mathbf{I} - \tilde{\mathbf{L}}$.

The constant weight (CW) design for AC assigns identical weights $w_{ij}^{\mathrm{CW}} = \alpha > 0$ to all edges [2]. Bounds on α that ensure convergence are discussed in the next section. With CW, we have

$$\mathbf{W}^{\mathrm{CW}} = \mathbf{I} - \alpha\mathbf{L}\,.$$

An AC design with improved performance is based on the MH algorithm (e.g., [9]). Using a non-negative regularization parameter $\epsilon \geq 0$, we introduce the following generalized version of MH weights:

$$w_{ij}^{\mathrm{MH}}(\epsilon) = \begin{cases} \frac{1}{\max\{d_i, d_j\} + \epsilon}\,, & \text{for } (i,j) \in \mathcal{E}, \\ 0\,, & \text{for } (i,j) \notin \mathcal{E} \text{ and } i \neq j, \\ 1 - \sum_{j \neq i} w_{ij}^{\mathrm{MH}}(\epsilon)\,, & \text{for } i = j\,. \end{cases} \qquad (2)$$

Recall that d_i denotes the degree of node i. The conventional MH weight designs are obtained with $\epsilon = 0$ [9] and $\epsilon = 1$ [12]. The latter was proposed to guarantee convergence for any graph topology (see next section). Decreasing ϵ also decreases the "self-loop" weights $w_{ii}^{\mathrm{MH}}(\epsilon)$ and implies that the nodes tend less to preserve their own current state.

GMH weights have the advantage that they are very easy to compute and require only local information about the graph topology. Better averaging performance can be achieved with weight designs that require global graph parameters or are computationally more demanding (e.g., [6]).

4. CONVERGENCE ANALYSIS

In this section we investigate the convergence of AC with GMH weights. The convergence of AC with constant weights is addressed, e.g., in [2]. Specifically, it is shown that $0 < \alpha < 1/d_{\max}$ is necessary and sufficient for convergence of AC/CW in arbitrary graphs. Furthermore, it is known that with $\alpha = 1/d_{\max}$, AC/CW does not converge for regular bipartite graphs. Since for d-regular graphs $w_{ij}^{\mathrm{MH}}(0) = w_{ij}^{\mathrm{CW}} = 1/d$, $(i, j) \in \mathcal{E}$, AC/GMH with $\epsilon = 0$ does not converge either in regular bipartite graphs. This can be illustrated using the structure of the weight matrix. For a d-regular bipartite graph, we have $\mathbf{L} = d\mathbf{I} - \mathbf{A}$ and hence $\mathbf{W}^{\mathrm{CW}} = \mathbf{A}/d$ for $\alpha = 1/d$. If the graph in addition is bipartite, the nodes can always be re-ordered such that the adjacency matrix has a block structure with diagonal blocks identical to zero; this further implies

$$\mathbf{W}^{\mathrm{CW}} = \mathbf{W}^{\mathrm{MH}}(0) = \frac{1}{d}\mathbf{A} = \frac{1}{d}\begin{pmatrix} \mathbf{0} & \mathbf{A}_0 \\ \mathbf{A}_0^{\mathrm{T}} & \mathbf{0} \end{pmatrix}.$$

In view of (1), the nodes in each group average the current estimates of the nodes in the other group. Asymptotically, each node estimate oscillates between the two means of the measurements in each group (unless the graph is $2p$-partite with $p > 1$).

We next address the open question whether there are other classes of graphs for which AC/GMH does not converge, This will lead to necessary and sufficient convergence conditions. In the literature, e.g. [6], it is only stated that bipartite graphs violate the stability constraints.

4.1. Sufficient Conditions for Convergence

We first show that $\epsilon > 0$ is sufficient for convergence of AC/GMH, regardless of the graph topology.

We recall that $\mathbf{W} = \mathbf{I} - \tilde{\mathbf{L}}$ and hence $\mathbf{W}$ is a doubly stochastic matrix, which has an eigenvalue with magnitude one and associated eigenvector $\mathbf{1}/\sqrt{I}$. Observe that $\tilde{\mathbf{L}}$ is positive semidefinite and $\lambda_k(\mathbf{W}) = 1 - \lambda_k(\tilde{\mathbf{L}})$ (where $\lambda_k(\mathbf{A})$ denotes the sorted eigenvalues of $\mathbf{A}$). The convergence condition $\rho\big(\mathbf{W} - \frac{1}{I}\mathbf{1}\mathbf{1}^{\mathrm{T}}\big) < 1$ is thus equivalent to the requirement that (i) the maximum eigenvalue of $\tilde{\mathbf{L}}$ is less than two, $\rho(\tilde{\mathbf{L}}) = \max_k \lambda_k(\tilde{\mathbf{L}}) < 2$, (ii) the zero eigenvalue of $\tilde{\mathbf{L}}$ has multiplicity one. The latter condition is satisfied if and only if the underlying graph is connected. This can easily be confirmed via the quadratic form

$$\mathbf{v}^{\mathrm{T}}\tilde{\mathbf{L}}\mathbf{v} = \sum_{(i,j) \in \mathcal{E}} w_{ij}(v_i - v_j)^2.$$

For positive weights, this quadratic form equals zero if and only if $v_i = v_j$ for any two connected nodes. For a connected graph, the unique normalized vector meeting this requirement is $\mathbf{v} = \mathbf{1}\sqrt{I}$. It remains to study the constraint $\rho(\tilde{\mathbf{L}}) < 2$. Since $\tilde{\mathbf{L}} = \mathrm{diag}\{\tilde{\mathbf{d}}\} - \tilde{\mathbf{A}}$ and $\tilde{\mathbf{A}}\mathbf{1} = \tilde{\mathbf{d}}$, the Gershgorin circle theorem [18] implies that the eigenvalues $\lambda(\tilde{\mathbf{L}})$ of $\tilde{\mathbf{L}}$ satisfy

$$\big|\lambda(\tilde{\mathbf{L}}) - \tilde{d}_i\big| \leq \tilde{d}_i\,.$$

Here, the ith element of $\tilde{\mathbf{d}}$, $\tilde{d}_i = \sum_{j \in \mathcal{N}_i} w_{ij}$, is the weighted degree of node i. In particular, we have the bound $\lambda(\tilde{\mathbf{L}}) \leq 2\tilde{d}_i$. Therefore, $\tilde{d}_i < 1$ for all $i \in \mathcal{V}$ is a sufficient condition for AC convergence.

With GMH weights, we have $\max\{d_i, d_j\} \geq d_i$ and hence $w_{ij}^{\mathrm{MH}}(\epsilon) \leq \frac{1}{d_i + \epsilon}$, which in turn implies

$$\tilde{d}_i \leq \sum_{j \in \mathcal{N}_i} \frac{1}{d_i + \epsilon} = \frac{d_i}{d_i + \epsilon}\,.$$

Thus, $\epsilon > 0$ is a sufficient condition for convergence because it ensures $\tilde{d}_i < 1$. It is seen that $\epsilon = 1$ as in conventional MH [12] is actually a very conservative choice. For the case of constant weights, we have $\tilde{d}_i = \alpha d_i$ and hence recover the sufficient convergence condition $\alpha < 1/d_{\max}$.

4.2. Necessary Conditions for Convergence

We next provide necessary conditions for AC convergence and we identify the corresponding critical graph topologies.

Theorem 1: *The conditions $\epsilon > 0$ (AC/GMH) and $\alpha < 1/d_{\max}$ (AC/CW) are necessary for convergence. AC/GMH with $\epsilon = 0$ and*

AC/CW with $\alpha = 1/d_{\max}$ converge unless the graph is regular and bipartite.

Proof: Lemma 1 below establishes

$$\rho(\tilde{\mathbf{L}}) \leq \rho(|\tilde{\mathbf{L}}|),$$

with equality if and only if the graph is bipartite. Furthermore, for AC/GMH with $\epsilon \geq 0$ and for AC/CW with $\alpha \leq 1/d_{\max}$, Lemma 2 below states that

$$\rho(|\tilde{\mathbf{L}}|) \leq 2,$$

with equality if and only if the graph is regular and $\epsilon = 0$ (AC/GMH) or $\alpha = 1/d_{\max}$ (AC/CW). Hence, for regular bipartite graphs and $\epsilon = 0$ (AC/GMH) or $\alpha = 1/d_{\max}$ (AC/CW), we have $\rho(\tilde{\mathbf{L}}) = 2$, which implies $\rho\big(\mathbf{W} - \frac{1}{I}\mathbf{1}\mathbf{1}^{\mathrm{T}}\big) = 1$ und hence divergence of AC. If the graph is not regular or not bipartite, AC converges even with $\epsilon = 0$ (AC/GMH) and $\alpha = 1/d_{\max}$ (AC/CW) since here $\rho(\tilde{\mathbf{L}}) < 2$. □

The next Lemma is a specialization of [19, Lemma 2.1] to weighted Laplacians.

Lemma 1: *For any weighted connected graph with strictly positive weights, the weighted Laplacian $\tilde{\mathbf{L}}$ satisfies*

$$\rho(\tilde{\mathbf{L}}) \leq \rho(|\tilde{\mathbf{L}}|).$$

Equality holds if and only if the graph is bipartite.

Proof: According to [20, Theorem 8.1.18], $\rho(\tilde{\mathbf{L}}) \leq \rho(|\tilde{\mathbf{L}}|)$. Furthermore, [20, Theorem 8.4.5] states that equality holds if $|\tilde{\mathbf{L}}|$ is irreducible (in our case this is the case since the underlying graph is assumed to be connected) and there exists a diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}\{e^{j\varphi_1}, \ldots, e^{j\varphi_I}\}$ such that

$$\tilde{\mathbf{L}} = e^{j\theta}\mathbf{\Lambda}|\tilde{\mathbf{L}}|\mathbf{\Lambda}^{-1}. \quad (3)$$

Here, $e^{j\theta}$ is the phase of the largest eigenvalue, which in our case equals plus one since $\tilde{\mathbf{L}}$ is positive semi-definite. Denoting the elements of $\tilde{\mathbf{L}}$ by $\tilde{l}_{ij}$, (3) can be rewritten as

$$\tilde{l}_{ij} = e^{j(\varphi_i - \varphi_j)}|\tilde{l}_{ij}|. \quad (4)$$

Due to $\mathbf{W} = \mathbf{I} - \tilde{\mathbf{L}}$, we furthermore have

$$\tilde{l}_{ij} = \begin{cases} -w_{ij}, & (i,j) \in \mathcal{E}, \\ \sum_{k \in \mathcal{N}_i} w_{ik}, & i = j, \\ 0, & \text{else.} \end{cases}$$

Condition (4) is trivially satisfied for $i = j$ and for $\tilde{l}_{ij} = 0$. It remains to study the case $(i,j) \in \mathcal{E}$; here, (4) is equivalent to $w_{ij} = -e^{j(\varphi_i - \varphi_j)}|w_{ij}|$, which in turn necessitates $\varphi_i - \varphi_j = \pi \bmod 2\pi$, $(i,j) \in \mathcal{E}$. Suppose we pick an arbitrary node i_0 with associated phase φ_{i_0}. Then, $\varphi_j = \varphi_{i_0} + \pi \bmod 2\pi$ for all neighboring nodes $j \in \mathcal{N}_{i_0}$. Furthermore, the neighbors $i \in \mathcal{N}_j$ for $j \in \mathcal{N}_{i_0}$, i.e., all nodes that are two hops away from i_0, must have $\varphi_i = \varphi_{i_0} \bmod 2\pi$. Continuing this argument iteratively, it follows that $\varphi_j = \varphi_{i_0} \bmod 2\pi$ if there is an even number of edges between i_0 and j and $\varphi_j = \varphi_{i_0} + \pi \bmod 2\pi$ if there is an odd number of edges between i_0 and j. This implies that there are two groups of nodes, i.e., one with $\varphi_j = \varphi_{i_0} + \pi \bmod 2\pi$ and one with $\varphi_j = \varphi_{i_0} \bmod 2\pi$, where none of the nodes in a group are neighbors. Thus, the underlying graph has to be bipartite. □

The following Lemma establishes an upper bound on the spectral radius of the signless Laplacian.

Lemma 2: *For a connected graph with GMH weights $\epsilon \geq 0$ or CW with $\alpha \leq 1/d_{\max}$, the spectral radius of the the signless Laplacian is bounded as*

$$\rho(|\tilde{\mathbf{L}}|) \leq 2.$$

Equality is obtained if and only if the graph is regular and $\epsilon = 0$ (GMH) or $\alpha = 1/d_{\max}$ (CW).

Proof: According to [20, Theorem 8.1.22 and Theorem 8.4.4], we have $\rho(|\tilde{\mathbf{L}}|) \leq \max_i \sum_{j=1}^{I} |\tilde{l}_{ij}|$ with equality if and only if all row sums of $|\tilde{\mathbf{L}}| = \mathrm{diag}\{\tilde{\mathbf{d}}\} + \tilde{\mathbf{A}}$ are equal. These row sums are given by $\sum_{j=1}^{I} |\tilde{l}_{ij}| = 2\tilde{d}_i = 2\sum_{j=1}^{I} w_{ij}$.

For AC/CW with $\alpha \leq 1/d_{\max}$, $\tilde{d}_i = \alpha d_i \leq 1$ and hence $\rho(|\tilde{\mathbf{L}}|) \leq 2$. Equality holds if and only if (i) all degrees d_i are identical, i.e., $d_i = d$, and thus the graph is d-regular, and (ii) $\alpha = 1/d$.

For AC/GMH with $\epsilon \geq 0$, $\max\{d_i, d_j\} \geq d_i$ and hence $w_{ij}^{\mathrm{MH}}(\epsilon) \leq \frac{1}{d_i + \epsilon}$, which in turn implies $\tilde{d}_i \leq \frac{d_i}{d_i + \epsilon} \leq 1$. Again, $\rho(|\tilde{\mathbf{L}}|) \leq 2$ with equality if and only if (i) $\epsilon = 0$ and (ii) the graph is regular (i.e., all degrees d_i are identical. □

4.3. Discussion

According to the results above, $\epsilon > 0$ and $\alpha < 1/d_{\max}$ are necessary and sufficient conditions for convergence of AC/GMH and AC/CW, respectively, on arbitrary connected graphs. However, on graphs that are not regular or not bipartite, convergence is obtained even with $\epsilon = 0$ and $\alpha = 1/d_{\max}$. Furthermore, as will be seen in the simulation results, GMH with $\epsilon = 1$ is not necessarily optimal for regular bipartite graphs. More specifically, for the case of 4-regular graphs the optimum ϵ is roughly inversely proportional to the number of nodes.

In contrast to applications like grid computing, it is rather unlikely in the context of WSN that the communication graph is regular and bipartite. If the WSN is modeled via random geometric graphs, the probability for a regular topology can be shown to decrease exponentially with increasing network size. Thus, the probability for a regular topology is non-vanishing only for very small random geometric graphs. If in addition we require bipartiteness, geometric constraints imply that the worst case are 4-regular graphs on the torus that are bipartite. In a finite two-dimensional region, 4-regular bipartite graphs are not feasible. In higher dimensions, regular bipartite graphs with larger degrees are possible but very unlikely to occur with random geometric graph models.

In summary, in small WSN the regularization parameter ϵ of AC/GMH should be chosen larger than in huge WSN. For the latter, AC/GMH with $\epsilon = 0$ will usually be superior.

5. NUMERICAL RESULTS

In general, the convergence rate of AC/GMH depends on ϵ. To study this dependence, Fig. 1 shows the mean-square error (MSE) in dB versus ϵ after 30 AC iterations for different graphs and independent, identically distributed (i.i.d.) initial values s_i. In a 4-regular bipartite graph with 16 nodes, the optimum regularization turns out to be $\epsilon \approx 1$. In the case of the regular bipartite graph with 100 nodes, however, the minimum MSE is obtained with $\epsilon \approx 0.25$. These results agree with the analytical findings regarding optimal CW in the

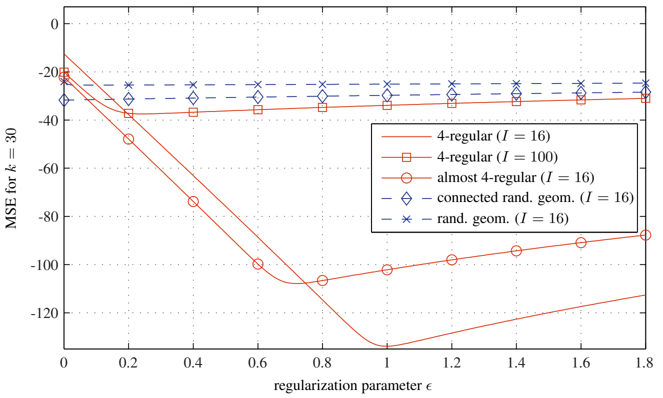

Fig. 1. *MSE (in dB) after* 30 *iterations versus GMH parameter* ϵ *for different graph topologies; initial measurements were i.i.d.*

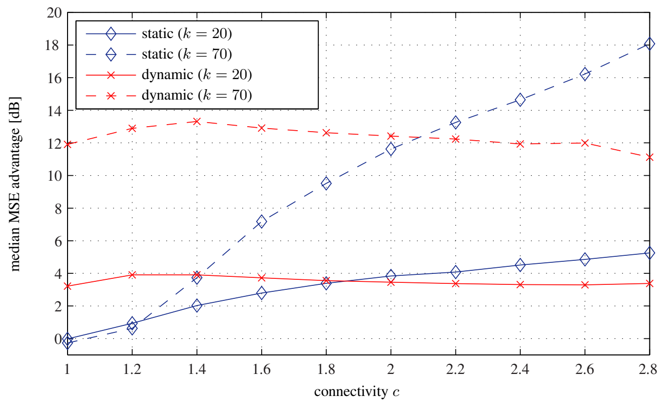

Fig. 2. *Median MSE advantage of AC/GMH with* $\epsilon = 0$ *relative to* $\epsilon = 1$ *versus connectivity* c *after* 20 *and* 70 *iterations in static and dynamic RGGs with* $I = 100$ *nodes.*

asymptotic regime [6]. For an "almost regular graph," obtained by deleting one edge from the 16-node 4-regular bipartite graph, the optimum MSE is achieved with $\epsilon \approx 0.7$. In the case of connected random geometric graphs, (almost) regular topologies are highly unlikely; here, the optimum is at $\epsilon = 0$, i.e., a non-zero regularization term only slows down convergence. For possibly unconnected random geometric graphs (resulting, e.g., from a smaller communication radius), the MSE depends weakly on the GMH parameter ϵ, with a poorly pronounce optimum larger than 0; this can be attributed to the fact that here smaller unconnected components (mostly consisting of 2 nodes) arise in the graph. We also note that in all our simulations choosing $\epsilon > 1$ never lead to performance improvements.

We next compare the performance of the two extreme cases of GMH, i.e., $\epsilon = 0$ and $\epsilon = 1$. We consider random geometric graphs with $I = 100$ nodes in a square region $\mathcal{A}$. Denoting the communication range by r, we define the connectivity c as $c = r\sqrt{I/|\mathcal{A}|}$; note that c is proportional to the average number of neighbors in the graph. The initial values s_i were again i.i.d. Fig. 2 shows the median MSE advantage in AC/GMH obtained with $\epsilon = 0$ relative to $\epsilon = 1$ as a function of network connectivity c. Results are shown for $k = 20$ and $k = 70$ iterations and for static and dynamic networks. In static networks, the topology and hence the GMH weights do not change over time. In the dynamic case, 20 nodes randomly changed their position during each time step, resulting in time-varying GMH weights (further details on this node mobility model and its impact on the averaging performance can be found in [21]). It is seen that in this scenario, $\epsilon = 0$ is superior to $\epsilon = 1$ in all operating conditions except for poorly connected static networks, where $\epsilon = 1$ is slightly superior after 70 iterations. In the static case, the MSE advantage increases with increasing connectivity c whereas in the dynamic case the MSE advantage is almost independent of c. For strongly connected graphs ($c = 2.8$), $\epsilon = 0$ performs better than $\epsilon = 1$ after 70 iterations by about 18 dB (static networks) and 11 dB (dynamic networks). This can be attributed to the fact that for well-connected graphs it is highly unlikely to have (almost) regular or bipartite topologies. We performed the same experiments also with (spatially) correlated initial values (results not shown). In this case we observed that $\epsilon = 0$ is superior even at low connectivity ($c \approx 1$).

6. CONCLUSIONS

We introduced a generalized Metropolis-Hastings weight design for average consensus that uses an arbitrary non-negative regularization parameter ϵ. Existing MH weight designs are re-obtained as special cases with $\epsilon = 0$ and $\epsilon = 1$. We further provided a comprehensive convergence analysis and explicitly stated necessary and sufficient convergence conditions. Our theoretical and numerical results indicate that conventional MH with $\epsilon = 1$ in most case is overly conservative. In particular, in realistic sensor network scenarios GMH with $\epsilon = 0$ outperforms $\epsilon = 1$ significantly under virtually any operating conditions.

REFERENCES

[1] Yingshu Li and My T. Thai, Eds., *Wireless sensor networks and applications*, Signals and Communication Technology. Springer Berlin, 2008.

[2] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[3] A. G. Dimakis, S. Kar, J. M. F. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847–1864, November 2010.

[4] C. C. Moallemi and B. Van Roy, "Consensus propagation," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4753–4766, Nov. 2006.

[5] J. N. Tsitsiklis, *Problems in Decentralized Decision making and Computation*, Ph.D. thesis, Massachusetts Institute of Technology, Dec. 1984.

[6] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, Sept. 2004.

[7] V. Schwarz and G. Matz, "Mean-square optimal weight design for average consensus," in *Proc. IEEE SPAWC-2012*, Cesme, TR, June 2012, pp. 374–378.

[8] P. Bremaud, *Markov Chains: Gibbs Fields, Monte Carlo Simulation and Queues*, Springer New York, 1999.

[9] S. Boyd, P. Diaconis, and L. Xiao, "Fastest mixing Markov chain on a graph," *SIAM Review*, vol. 46, no. 4, pp. 667–689, Dec. 2004.

[10] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Int. Conf. on Information Processing in Sensor Networks*, Los Angeles, CA, 2005, pp. 63–70.

[11] P. Denantes, "Performance of Averaging Algorithms in Time-Varying Networks," Tech. Rep., EPFL, 2007.

[12] L. Xiao, S. Boyd, and S. Lall, "Distributed average consensus with time-varying metropolis weights," *Automatica*, 2006.

[13] A. Olshevsky and J.N. Tsitsiklis, "Degree fluctuations and the convergence time of consensus algorithms," in *Proc. IEEE CDC-ECC-2011*, Dec 2011, pp. 6602–6607.

[14] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *Automatic Control, IEEE Transactions on*, vol. 50, no. 2, pp. 169–182, February 2005.

[15] Cvetković D., "New theorems for signless laplacian eigenvalues," *Bulletin*, vol. 137, no. 33, pp. 131–146, 2008.

[16] H. Kenniche and V. Ravelomananana, "Random geometric graphs as model of wireless sensor networks," in *Computer and Automation Eng. (ICCAE), 2010 2nd Int. Conf. on*, Singapore, vol. 4, pp. 103–107.

[17] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 2nd edition, 1989.

[18] Richard S Varga, *Geršgorin and His Circles*, vol. 36 of *Springer Series in Computational Mathematics*, Springer, 2004.

[19] Xiao-Dong Zhang and Rong Luo, "The spectral radius of triangle-free graphs," *Australasian Journal of Combinatorics*, vol. 26, pp. 33–40, Sept. 2002.

[20] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge (UK), 1999.

[21] V. Schwarz and G. Matz, "On the performance of average consensus in mobile wireless sensor networks," in *Proc. IEEE SPAWC-2013*, Darmstadt, GER, June 2013, pp. 175–179.