REACHING CONSENSUS IN ASYNCHRONOUS WSNs: ALGEBRAIC APPROACH

Ondrej Slučiak and Markus Rupp
Institute of Telecommunications, Vienna University of Technology, Austria
{o-sluciak, m-rupp}@nt.tuwien.ac.at

ABSTRACT
Many models of wireless sensor networks (WSNs) assume a perfect synchronization along the graph of such network as a simplifying assumption. In our contribution we base our investigations of distributed algorithms solving consensus problems on more realistic, asynchronous networks in which nodes randomly transmit to their neighborhood. Following a linear algebraic approach we show conditions for convergence to a consensus and derive convergence properties in the mean and mean square sense.

Index Terms— asynchronous wireless sensor networks, consensus, convergence, solution space, linear algebra

I. INTRODUCTION
A typical simplification in modeling of a wireless sensor network (WSN) [1]–[3] and distributed algorithms design [4]–[6] is that at certain time instant $t_k$ all $N$ nodes of the WSN are transmitting synchronously so that a specific node $i$ receives from all its neighbors in the neighborhood $\mathcal{N}_i$ new information, i.e.,

$$x_i(t_k) = f(x_{\mathcal{N}_i}(t_{k-1})),$$

(1)

The notation here indicates that the state $x_i$ of node $i$ is changed by the state information collected in a vector $x_{\mathcal{N}_i}$ from its neighborhood $\mathcal{N}_i$. First of all, the sensor nodes are typically low energy devices that need to be woken up only if they are in need of operation. This makes it rather difficult to synchronize for communication between each other [7], [8]. On the other hand, even if such synchronization has been established, concurrent transmission of all nodes would lead to channel congestion and communication may not be possible. Note that typically nodes do not direct their transmission to a specific node, but rather transmit radially to all their neighborhood. Thus, in a small fraction of time a lot of channel or network capacity is required while at all other times no capacity would be needed. Alternatively, the instant $t_k$ could represent a time epoch in which all nodes have to transmit one after the other, thus avoiding interference and utilizing spectrum resources more efficiently. Then, after the last node’s operation is completed, the update would take place. This, however, requires careful synchronization and a controlling base station, just what is not available in this context.

Therefore, it is more feasible to operate on asynchronous transmissions, e.g., [9], that is, each node independently decides at a certain time instant, say $t_k$, that it transmits to its neighbors regardless of who receives this information. This shifts the model from the perspective of the receiving node to the transmitting node. Its neighbors may or may not receive the information and once they receive it correctly, they may cause an update themselves, possibly followed by another transmission of their own. Also, as the transmission may be unreliable, some error-control like automated repeat request (ARQ) protocol between nodes, which would handle duplicate and lost messages, might be required [10].

The paper is structured as follows. In Section II, we introduce asynchronous updates originating from the transmitter rather than the receiver. A relatively simple structured, so-called state transition matrix, of row-stochastic type is introduced as the key element. Section III analyses this central update matrix showing its main properties. In Section IV, we deliver probabilistic convergence conditions for the proposed asynchronous mode and in Section V we briefly discuss effects that arise when a single node fails. Section VI concludes the paper.

II. ASYNCHRONOUS MODEL
We will consider asynchronous linear updates every time an information is received in the neighborhood of a transmitting node $i$, i.e.,

$$x_i[k] = g(x_i[k-1]),$$

(2)

where we changed the notation for time to a simple event counter $k$ as the time itself is of no further importance. On the example of the average consensus algorithm [1] we will show the mechanics of such network. Take, for example, node $i$ that receives from its neighbor node $m$ a value $x_m$. The node $i$ would then take its internal state $x_i$ and compute

$$x_i[k] = \alpha x_i[k-1] + (1-\alpha) x_m[k-1],$$

(3)

where $\alpha \in (0,1)$ is so-called mixing parameter.

All other states remain unchanged by such operation. We can describe such basic state transition by a matrix operation $S_{im}$ on a current state vector $x[k]$ that contains the collected state info of the entire network:

$$S_{im} x[k-1] = x[k],$$

(4)

The state transition matrix $S_{im} \in \mathbb{R}^{N \times N}$ has a specific form: essentially a unit matrix with ones on its diagonal, except $S_{im}$ and $\{S_{im}\} = 1 - \alpha$ and $\{S_{im}\} \in \mathbb{R}$ respectively. We can thus define a matrix of transitions at time $k$, i.e., $S[k] \in \mathcal{S} = \{S_{im}\}$, where $\mathcal{S}$ is the set of all allowed transitions in the network.

If node $i$ concurrently transmits to its neighbors $l,m,n$, we would have three concatenated updates, i.e., $S[i] = S_{il}S_{im}S_{in}$. The neighborhood of node $i$ defines how many different columns such an entry would potentially have.

Note that receiving new information does not necessarily mean there is an update required. A node could also collect receiving messages from neighborhood nodes and after a certain time period update [12]. However, this requires some logic to decide how long to wait and when to stop waiting. A simpler strategy is to perform an update every time new information is received. Whether this results in an
immediate transmission afterwards can remain open here. In our model we assume that a node \( i \) starts transmitting its next message (repeatedly) with probability \( p_i \) (for example \( p_1 = \cdots = p_N = p = 1/N \)). If indeed several nodes, say \( l, m, n \), transmit their results to node \( i \) we can describe this as a concatenation of \( S_lS_mS_n \), the order of which is relevant. Clearly, we can include synchronous updates in this formulation as well.

A proper working of such WSN thus depends on the sequence of \( S_{im} \) matrices that occur randomly. We like to know whether the result of such a random sequence with initial states \( x[0] \) leads to a consensus, that is \( x[\infty] = \gamma x[0] \top 1 \), with \( 1 \top = (1, \ldots, 1) \) or even more specifically the average consensus with \( \gamma = 1/N \). Without a formal proof we can already conclude intuitively one important property: there exists always a sequence of operations \( S_{im} \) that will not lead to such result. Thus, obviously, the WSN cannot be guaranteed to converge in worst case. This requires to analyze such networks in a stochastic context.

Note further that we do not restrict ourselves to fixed values of \( \alpha \). In fact every node pair \((i,m)\) can have its own \( \alpha_{im} \), which also can vary in time, denoted \( \alpha_{im}[k] \). Such variations on the updates clearly includes link failures [12], [13], for which \( \alpha \) simply turns to one for a certain time, and also include quantization effects [14], [15].

### III. STATE UPDATE ANALYSIS

Matrix \( S_{im} \) is thus the central element of our further analysis. Since the matrix is a so-called row-stochastic matrix, it is well known that the largest eigenvalue \( \lambda_1 \) is equal to 1 with the corresponding eigenvector \( \mathbf{1} \). As the matrix has a simple form it is not difficult to compute all eigenvalues and corresponding eigenvectors of it.

**Lemma 3.1:** Matrix \( S_{im} \) in Equation (4) has the following right eigenvalues and eigenvectors respectively:

\[
\lambda_1 = 1, \quad v_1 = 1 \top, \\
\lambda_2 = \alpha, \quad v_2 = \{0, \ldots, 0, 1, 0, \ldots, 0\}, \quad \text{at entries } j \neq \{i, m\}, \\
\lambda_{im} = \alpha, \quad v_{im} = \{0, \ldots, 0, 1, 0, \ldots, 0\}, \quad \text{at entry } i.
\]

**Proof:** Since matrix \( S_{im} \) is a block diagonal matrix, its characteristic polynomial is: \( \det(S_{im} - \lambda I) = (1 - \lambda)^N - (\alpha - \lambda) = 0 \), i.e., the eigenvalues are \( \lambda_1 = 1 \), with multiplicity \( N - 1 \), and \( \lambda_2 = \alpha \), with multiplicity 1. Since the rows of \( S_{im} \) sum to 1, one eigenvector is \( \mathbf{1} \). Other eigenvectors corresponding to these eigenvalues are then straightforward to find.

In other words, the first eigenvector has all ones entries. It is transmitted without any change, that is \( \lambda = 1 \). We find that the matrix of \( N \) states has \( N - 1 \) eigenvectors that are unit vectors, \( N - 2 \) of them (all but the \( i \)-th and the \( m \)-th entry) have eigenvalue one and thus are not changing on such entries. Finally, the unit vector with one entry at position \( m \) is also an eigenvector, but only \( \lambda_{im} = \alpha \) is transmitted of such input.

In the following we investigate under which conditions such WSN can arrive at the desired result. We will not employ knowledge about graphs here, but will use well-known concepts of linear vector spaces instead.

**Definition 3.1:** We name the linear hull of all eigenvectors with corresponding eigenvalue \( \lambda = 1 \) the solution space of node \( i \):

\[
\mathbf{S}_i = \text{span} \{ \mathbf{1}, \{v_j\}_{j \neq \{i, m\}} \}.
\]

**Lemma 3.2:** A necessary condition for a WSN with asynchronous updates \( S_{im} \) to converge to a consensus is that the intersection of all solution spaces is spanned by exactly one eigenvector: \( \mathbf{1} \), that is:

\[
\bigcap_{i} \mathbf{S}_i = \text{span} \{ \mathbf{1} \}.
\]

**Proof:** Intuitively, if the zero vector would be the only solution, the WSN would result in all zeros asymptotically. If the intersection of all spaces would be spanned by more than \( \mathbf{1} \), other solutions would be possible and thus the solution would not be unique.

We note here that in [16] a more profound proof with a similar definition involving paracotracing Metropolis weight matrices has also been shown, offering very similar condition for reaching consensus as in Lemma 3.2.

**Example A1:** Let us assume a WSN with \( N = 4 \). Two pairs of nodes \((1, 2)\) and \((3, 4)\) are connected (bidirectional), while the pairs themselves are not connected. We then find the following matrices:

\[
S_{12} = \begin{pmatrix} \alpha & 1 - \alpha & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad S_{23} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 - \alpha & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\
S_{34} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad S_{41} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
\]

All four operations share the eigenvectors \( \mathbf{1} \) and \( \mathbf{v}^\top = (0, 0, 1, 1) \), the intersection of the solution spaces is thus span\{\mathbf{1}, \mathbf{v}\}, and the solution is not unique.

**Example A2:** Let us now consider Example A1 with a single additional connection from node 3 to node 1. Additionally, we then have

\[
S_{13} = \begin{pmatrix} \alpha & 0 & 1 - \alpha & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
\]

Now the intersection of the solution spaces reduces to the linear hull of the eigenvector \( \mathbf{1} \).

Note that these two examples clearly show the properties of such asynchronous algorithm. In our example A2, we have a weakly connected graph: nodes 3 and indirectly 4 (via 3) deliver information to node 1 and indirectly 2 (via 1) but from nodes 1 and 2 there is no information returned to nodes 3 and 4. Nevertheless a consensus is found.

Some further important properties of the state transition matrix \( S_{im} \) are summarized in the following lemma.

**Lemma 3.3:** The state transition matrix \( S_{im} \) defined in Eq. (4) has the following properties

\[
\left( \mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^\top \right) S_{im} = \left( \mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^\top \right) S_{im} \left( \mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^\top \right), \\
\left( \mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^\top \right) S_{im}^\top = \left( \mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^\top \right) S_{im}^\top \left( \mathbf{I} - \frac{1}{N} \mathbf{1} \mathbf{1}^\top \right).
\]

**Proof:** The proof follows straightforwardly from the fact that the matrix \( S_{im} \) is row-stochastic.

### IV. CONVERGENCE ANALYSIS

While there is some general convergence results for networks with fixed synchronous updates [1], the results for a time-variant (switching) network following the update (1) are more limited. Typically, results are known for convergence to the average consensus in the mean or mean square sense (e.g. [12], [17]). In [16] it was proposed to use the notion of the joint spectral radius of paracotraacting matrices to compute the convergence rate.

Let us now consider a chain of update events assuming that the transmission from node \( i \) to node \( m \) occurs with probability \( p_{im} \). For simplicity let us assume that the various probabilities \( p_{1,2}, \ldots, p_{N,N-1} \) remain constant over time. After, say, \( K \) updates, we have the following state vector:

\[
x[K] = \prod_{k=1}^{K} S[k] x[0]
\]

where the matrices \( S[k] \in \{ S_{im} \} \) are selected randomly with probability \( p_{im} \). As we have observed before the order in which the matrices
occur plays an important role. As time average and ensemble average lead to different results we can conclude that $x[k]$ is a non-ergodic random process. From here we find convergence in the mean

$$E[x[k]] = \prod_{k=1}^{K} E[S[k]] x[0] = \left( \sum_{i,m} p_{i,m} S_{im} \right)^{K} x[0].$$

We recognize that matrix $A$ maintains its largest eigenvalue at one with the eigenvector $1$. Let us decompose $A = QAQ^{-1}$ with $A$ containing the eigenvalues ordered from largest to smallest. With $\lim_{k \to \infty} A^k = L$ a matrix with a single one element on its top left corner we obtain

$$\lim_{k \to \infty} E[x[k]] = QLQ^{-1} x[0] \triangleq x.$$

Due to the independence of the events, the expectation appears in the various product terms. Thus, given the various $\alpha_{im}$ and corresponding probabilities $p_{i,m}$ the sum term can be computed and the eigenvalues analyzed. For sure one eigenvalue remains one associated with its eigenvector $1$. The remaining eigenvalues and corresponding eigenvectors may change though. For the above Example A2, we obtain with $p_{i,m} = 1/5$ and $\alpha = 0.5$ the eigenvalues $\{0.74, 0.8, 0.96, 1\}$. Note, however, that in this example we find

$$QLQ^{-1} = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

thus obtaining the average of the third and fourth node, not including the information of the first two nodes. Whether the result is the desired average of all nodes’ state information, strongly depends on the structure of the WSN as it directly influences $Q$. As soon as the WSN has symmetric bidirectional connections, more precisely if $p_{i,m} (1-\alpha_{im}) = p_{m,i} (1-\alpha_{im}), Q$ becomes unitary and the ave-rage consensus is found in the mean (independent of the choice of $\alpha$) [12]. This allows for summarizing the result.

**Theorem 4.1:** The asynchronous consensus updates as described in Update (4) converges to any consensus in the mean if the conditions of Lemma 3.2 are satisfied. If furthermore $p_{i,m} (1-\alpha_{im}) = p_{m,i} (1-\alpha_{im})$ the average consensus $1/N x[0] \cdot 1$ is obtained asymptotically in the mean.

Note that the last condition causes a certain symmetry so that the state transition matrix in the mean becomes a stochastic matrix, in particular that a left eigenvector $1$ exists. Once such symmetry is guaranteed, many more explicit statements on convergence and convergence time are possible [12], [18].

Practically speaking, the convergence in the mean has not much impact. As it is an ensemble average it just means the averaging over an ensemble of runs (or even graphs) results in the desired mean. Due to the, in general, non-ergodic behavior each realization may be far off from the desired consensus.

This is the reason to analyze better in terms of convergence in the mean square sense. In [17] such analysis has been applied to the standard average consensus algorithm obtained by ensemble average-raging over a set of graphs representing switching networks. As the obtained MSE value is not zero, it shows that for such networks the individual result may be off from the desired average consensus. In [17] the MSE with respect to a fixed consensus, say $1$, was computed, thus

$$\text{MSE}_f[k] = E \left[ \|x[k] - 1\|^2 \right].$$

Note, however, that the value $\gamma$ is not fixed for all runs. As the runs are not ergodic, asymptotically the outcome will be different every time and in consequence a different $\gamma$ appears for each run. Thus, the previous metric $\text{MSE}_f[k]$ is not describing the behavior rightfully in terms of any consensus, but explains the deviation with respect to a fixed (expected) consensus.

Now let us also compute the MSE $E[k]$ under the condition that we are satisfied with any consensus, no matter what value it is. In this case the consensus can be the average of all values at the time instant $k$

$$\gamma_k = \frac{1}{N} \sum_{i=1}^{N} x[k] \triangleq \bar{x}_k.$$

We find for the MSE:

$$\text{MSE}[k] = E \left[ \|x[k] - \bar{x}_k \|^2 \right] = E \left[ \|x[k] - \bar{x}_k - 1\|^2 \right] = E \left[ x[k] - \bar{x}_k \right]^\top \left( I - \frac{1}{N} \bar{1} \bar{1}^\top \right) E \left[ x[k] - \bar{x}_k \right].$$

The last term is obtained using the properties of Lemma 3.3. We thus have just proven the following.

**Theorem 4.2:** The WSN with asynchronous updates according to Eq. (3), satisfying the necessary condition of Lemma 3.2 converges to a consensus in the mean square sense iff all eigenvalues of

$$E[B] = \left( \sum_{i,m} p_{i,m} S_{im} \right) \left( I - \frac{1}{N} \bar{1} \bar{1}^\top \right) S_{im} \triangleq B$$

are smaller than one.

For our Example A2 we find the largest eigenvalue to be 0.95 while for Example A1 the largest eigenvalue is 1 and thus no consensus is reached.

We note here that according to [19] the rate of convergence of the so-called mean-square deviation (cf. $\text{MSE}[k]$) is determined by the eigenvalue $\lambda_{\text{max}}(E[\bar{W}^\top \left( I - \frac{1}{N} \bar{1} \bar{1}^\top \right) \bar{W}])$, where $\bar{W}$ is a random weight matrix.

For the case of a fixed value $\gamma$, i.e., the true average,

$$\gamma = \frac{1}{N} \sum_{i=1}^{N} x_i \triangleq x_0$$

the MSE takes the form

$$\text{MSE}_f[k] = E \left[ \|x[k] - x_0 \|^2 \right] = E \left[ \|x[k] - x_0 + (\bar{x}_k - x_0) \|^2 \right] = E \left[ \|x[k] - x_0 \|^2 \right] + E \left[ N \|\bar{x}_k - x_0 \|^2 \right] = \text{MSE}[k] + NE \left[ \|\bar{x}_k - x_0 \|^2 \right].$$

It follows straightforwardly that for the steady-state we obtain

$$\lim_{k \to \infty} E \left[ \|x[k] - x_0 \|^2 \right] = \lim_{k \to \infty} \left( \text{MSE}[k] + NE \left[ \|\bar{x}_k - x_0 \|^2 \right] \right) = \lim_{k \to \infty} NE \left[ \|\bar{x}_k - x_0 \|^2 \right] = N \left( \lim_{k \to \infty} E \left[ \bar{x}_k^2 \right] - \bar{x}_0^2 \right)$$

Thus, the $\text{MSE}_f[k]$ grows linearly with the size of the network and, in general, is determined by the $\lim_{k \to \infty} E \left[ \bar{x}_k^2 \right]$. Obviously, in case when $\lim_{k \to \infty} E \left[ \bar{x}_k^2 \right] = \bar{x}_0^2$, the MSE of $\text{MSE}_f[k]$ asymptotically goes to zero.

**V. WHEN A NODE FAILS**

In case of a node failure, several scenarios are of interest. 1) We assume the node $m$ is dead, thus not transmitting any more. In this case the desired solution, that is the average of $N$ values, has to shrink to the new average of $N-1$ values. The solution space then shrinks by one dimension. Thus we expect the WSN to adapt to a new solution. However, if the condition of Lemma 3.2 becomes violated, a consensus cannot be reached anymore. This allows to define a robustness condition by counting how many nodes can die out at minimum before the condition is violated.
2) The node \( m \) is not updating internally but sending a fixed value, say \( x_f \), all the time it is activated to transmit. This would change the solution space \( S_m \), but not necessarily the intersection of all solution spaces, as the eigenvector \( I \) remains in \( S_m \). A consensus is indeed reached but instead of the average consensus it will be \( x_f \). This shows how simple it is to make a WSN to misbehave. Simply by planting a “bad” node, it would dominate the entire WSN and solely define its result.

3) The node \( m \) is misbehaving by sending wrong values, for example, it passes the received values without any change. This would also change the solution space significantly. As long as the condition of Lemma 3.2 is not violated this case may not lead to severe problems unless the node turns into the previous case.

VI. CONCLUSION

In this contribution we argued that most models of WSN are not very realistic given the fact that they should operate in a low power mode and thus cannot consume too much energy by coordination of data transmission. To overcome this problem we propose a very simple asynchronous update and analyze its behavior in the mean and mean square sense. Many even stronger statements like almost sure convergence can be derived under some symmetry conditions following the concepts of [18]. Some considerations on node failures conclude the paper.

VII. REFERENCES


