Exploiting Interference for Efficient Distributed Computation in Cluster-based Wireless Sensor Networks

Steffen Limmer*, Sławomir Stańczak*†, Mario Goldenbaum†, Renato L. G. Cavalcante *
* Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Einsteinufer 37, 10587 Berlin, Germany,
† Fachgebiet Informationstheorie und theoretische Informationstechnik,
Technische Universität Berlin, Einsteinufer 25, 10587 Berlin, Germany
Email: {steffen.limmer, slawomir.stanczak, renato.cavalcante}@hhi.fraunhofer.de; mario.goldenbaum@tu-berlin.de

Abstract—This invited paper presents some novel ideas on how to enhance the performance of consensus algorithms in distributed wireless sensor networks, when communication costs are considered. Of particular interest are consensus algorithms that exploit the broadcast property of the wireless channel to boost the performance in terms of convergence speeds. To this end, we propose a novel clustering based consensus algorithm that exploits interference for computation, while reducing the energy consumption in the network. The resulting optimization problem is a semidefinite program, which can be solved offline prior to system startup.

I. INTRODUCTION

In many wireless (sensor) applications, nodes cooperate for some common goal. One example is the localization of an acoustic source using a number of geographically distributed microphones that are equipped with wireless communication capabilities. In fire alarm networks, for instance, a number of wireless sensor nodes may be used to monitor maximum and average temperature values. In such applications, the goal is therefore not to share local measurements among network nodes but rather to compute one or multiple functions of these measurements (e.g., the maximum function or a weighted sum). The functions to be computed depend on the targeted application.

A key observation is that it is in general not necessary to exchange raw sensor measurements in order to compute a function thereof. It is further known that the broadcast property of the wireless channel can be beneficially exploited when the task is to compute or estimate function values at sensor nodes. In particular, the information-theoretic analysis in [1] shows how to encode messages for linear multiple-access channels to enhance the rate at which linear functions can be reliably computed at the receiver.

A different approach can be found in [2] where an *analog* computation scheme is proposed which 1) is able to efficiently compute *non-linear functions* over many multiple-access fading channels and 2) is robust against the lack of synchronization between different signals.[1] The idea was used in [3] as a building block of a cluster-based average

algorithm to improve the convergence speed of gossip-based algorithms for average consensus [4]. Reference [5] extended the approach of [3] to incorporate non-linear functions (f-consensus algorithms).

In this paper, we consider the approach of [3] but take into account different energy costs imposed on sensor nodes by different distant-dependent path losses. We neglect the energy consumption for feedback and channel estimation, which however is relatively low under certain conditions [6]. Given the setup, the objective is to cluster sensor nodes and then activate the clusters in such a way as to minimize the overall energy consumption for transmission. In each cluster, we neglect the impact of noise and assume that the average of sensor measurements within the cluster is computed using the CoMAC (Computation over Multiple-Access Channel) scheme of [2] (see also Figure 1). A consequence of this is that the energy consumption can be reduced by decreasing the size of the clusters; since this deteriorates the convergence speed and rate, there is an inherent trade-off between the energy consumption and the convergence behavior. This paper studies this trade-off by numerically solving a suitably formulated clustering and activation optimization problem. Different points on the trade-off curve are achieved by choosing different weights for a regularization term that takes into account the energy costs. The paper is organized as follows: We first introduce the system model and the cost model. This is followed by the problem statement. In Section IV, we reformulate the original problem of maximizing the convergence rate to take into account the energy consumption for transmission. This problem is solved using some standard optimization tools. Section VI presents some simulations, while the paper is completed by some conclusions and open problems.

II. SYSTEM MODEL

A. Network Model

We consider a clustered wireless sensor network consisting of $N \in \mathbb{N}$ (sensor) nodes that are grouped in C clusters. Let $\mathcal{N} = \{1, \ldots, N\}$ and let $\mathcal{C}_i \subseteq \mathcal{N}$ be the index set of nodes belonging to cluster $i \in \mathcal{C} := \{1, \ldots, C\}$. Throughout the paper, we use $\mathrm{CH}_i \in \mathcal{C}_i$ to refer to the cluster head of the ith
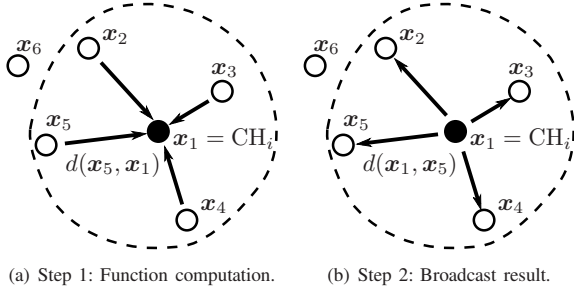
---

[1]Only a coarse frame synchronization is needed.

(a) Step 1: Function computation.   (b) Step 2: Broadcast result.

Fig. 1. Two-step approach to average consensus

cluster. Note that in general, $\mathcal{C}_i \cap \mathcal{C}_j \neq \emptyset$ for any $i, j \in \mathcal{C}$ and $\cup_{i=1}^{C} \mathcal{C}_i = \mathcal{N}$.

Remark 1. *Note that C denotes the number of possible clusters that are used in our optimization. C may be very large and, in extreme case, is equal to $2^N - 1$, the cardinality of the power set of $\mathcal{N}$ excluding the empty set. However, while operating, the network may activate only a subset of this set as a result of the optimization process.*

The sensor nodes are distributed over some geographical area and we use $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N] \in \mathbb{R}^{2 \times N}$ to denote the position matrix. Accordingly, $\boldsymbol{x}_n \in \mathbb{R}^2$ contains the coordinate of sensor node n with respect to some reference point. The vector of squared distances from node i to all nodes is $\boldsymbol{d}_i = \{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2\}_{j=1,\ldots,N}$ and we group these vectors in a matrix $\boldsymbol{D} = [\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N]$.

B. Time Model

To model the temporal behavior of cluster activation, the asynchronous time model in [4] is adopted. Furthermore we assume that the cluster heads wake up according to a rate $\mu_i \in \mathbb{R}_+$ Poisson process, where μ_i is chosen such that only a single cluster head is activated within a certain time window (with high probability).

C. Communication Model

Denote the reading of the kth sensor at time $t \in \mathbb{Z}_+$ by $y_k(t) \in \mathcal{Y} \subset \mathbb{R}$, $k \in 1, \ldots, N$, and the initial network state by $\boldsymbol{y}(0) \in \mathcal{Y}^N$. If cluster $\mathcal{C}_i$ is active at time $t \in \mathbb{Z}_+$, the received signal r_{CH_i} at time t by cluster head i (node 1 in Fig. 1), is given by

$$r_{\mathrm{CH}_i}(t) = \sum_{k \in \mathcal{C}_i \setminus \mathrm{CH}_i} h_{ik}(t) s_k(y_k(t)) + n_i(t), \tag{1}$$

where $h_{ik}(t)$, $n_i(t)$ and $s_k(\cdot)$ denote the corresponding channel coefficient from node k to cluster head, receiver-side noise and transmit signal of node k, respectively. This setting will be referred to as *noisy MAC*. Assume, nodes can estimate the channel to the cluster head by some wake up pilot symbols. Then, to compute the average within cluster $\mathcal{C}_i$, every node needs to invert it's own channel, which yields transmit signals

of the form

$$s_k(y_k(t)) = \frac{1}{h_{ik}(t)} y_k(t) \tag{2}$$

and simplifies the received signal to

$$r_{\mathrm{CH}_i}(t) = \sum_{k \in \mathcal{C}_i \setminus \mathrm{CH}_i} y_k(t) + n_i(t). \tag{3}$$

Hence, the channel inversion removes the impact of the channel. *In what follows, we assume an idealized noiseless setting; the noisy case remains an open problem for potential future research.*

Considering (3) and neglecting the noise term, we can write the time evolution of the state vector in a compact form. To this end, let $\boldsymbol{y}(t) = [y_1(t), \ldots, y_N(t)]$ be the state vector at time $t \in \mathbb{Z}_+$ and let $p_i \in [0, 1], i \in \mathcal{C}$, be the probability with which cluster i is activated. Then, $\boldsymbol{y}(t)$ evolves over time as

$$\boldsymbol{y}(t+1) = \boldsymbol{W}(t)\boldsymbol{y}(t), t \in \mathbb{Z}_+, \tag{4}$$

where for every $t \in \mathbb{Z}_+$, matrix $\boldsymbol{W}(t) := [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N] \in \mathbb{R}^{N \times N}$ is independently and randomly chosen according to the probability mass function/vector

$$\boldsymbol{p} = [p_1, \ldots, p_C] \in [0, 1]^C \tag{5}$$

from the set $\{\boldsymbol{W}^{(i)}\}_{i=1}^{C}$ of weight matrices defined by [3]

$$W_{j,k}^{(i)} := \begin{cases} 1, & \text{if } j \notin \mathcal{C}_i, \ k = j \\ \frac{1}{N_i}, & \text{if } j, k \in \mathcal{C}_i \\ 0, & \text{else.} \end{cases} \tag{6}$$

Remark 2. *To compute functions different from arithmetic or weighted mean, the consensus step can be adapted from weight matrices $\boldsymbol{W}^{(i)}$ to functionals $f : \mathcal{Y}^N \mapsto \mathbb{R}$. It has been shown in [7] that the resulting function space computable with the proposed scheme is essentially the space of nomographic functions, which contains also functions such as geometric mean as a special case.*

Remark 3. *For an analysis of the noisy setting, we refer the interested reader to [8].*

III. COST MODEL

We employ the energy cost model of [9], which, for a transmission of a k-bit message over a distance $d_{i,j}$ from node i to node j, is given by

$$E_{\mathrm{tot}} = E_{\mathrm{tx}} + E_{\mathrm{rx}} = kE_{\mathrm{elec}} + \epsilon_{\mathrm{amp}} k d_{i,j}^2 + kE_{\mathrm{elec}} \tag{7}$$

Notice that the circuitry for tx and rx is assumed to be identical for all nodes. *In what follows, we neglect the steady energy consumption (E_{elec}) and consider only dynamic energy consumption caused by amplifiers in the tx circuitry averaged over some time period.* We however point out that the results of this paper can be easily extended to capture the steady energy consumption, which may strongly influence the optimization results if the sensor nodes can be switched off after computing a function value. In such cases, it may be beneficial to prefer larger clusters when the steady energy consumption increases.

Moreover, for brevity, we neglect the receiver-side energy consumption for reception, which could be easily incorporated as well.

For an activated cluster $\mathcal{C}_i$ with respective weight matrix $\boldsymbol{W}^{(i)}$ and corresponding cluster head $\mathrm{CH}_i \in \mathcal{C}_i$, we have to differentiate between the energy cost associated with Step 1 $\boldsymbol{c}_{\mathrm{fc}}^{(i)}$ and the energy cost imposed by Step 2 $\boldsymbol{c}_{\mathrm{bc}}^{(i)}$ (see Figure 1). In the first case, the energy cost results from the transmission of sensor nodes to the ith cluster head and is therefore given by

$$\boldsymbol{c}_{\mathrm{fc}}^{(i)} = \boldsymbol{d}_{\mathrm{CH}_i} \odot \mathbf{1}_{\mathcal{C}_i} \tag{8}$$

where $\mathbf{1}_{\mathcal{C}_i}$ is the vector-valued indicator function of the set $\mathcal{C}_i$ defined as follows: if $i \in \mathcal{C}_i$, then the ith element of $\mathbf{1}_{\mathcal{C}_i}$ is 1; otherwise it is zero. Accordingly, the energy cost associated with Step 2 is determined by the maximum distance between the cluster head and cluster nodes, i.e. we have

$$\boldsymbol{c}_{\mathrm{bc}}^{(i)} = \|\boldsymbol{d}_{\mathrm{CH}_i} \odot \mathbf{1}_{\mathcal{C}_i}\|_\infty \cdot \boldsymbol{e}_{\mathrm{CH}_i} \tag{9}$$

where $\boldsymbol{e}_{\mathrm{CH}_i} \in \mathbb{R}^N$ is a unit vector with a 1 at the CH_ith position and zeros otherwise.

Remark 4. *Choosing an outer node as cluster head results in higher communication costs (as distances and therefore costs will be larger) even if the respective weight matrices are identical.*

IV. PROBLEM STATEMENT

Following the derivation in [3], [4], improving the convergence rate for a predefined clustered WSN can be achieved by optimizing the cluster activation probabilities $\boldsymbol{p} = [p_1, \ldots, p_C]$, corresponding to weight matrices $\boldsymbol{W}^{(i)}, i = 1 \cdots C$. The problem of optimizing the convergence rate can therefore be formulated as the problem of minimizing the second largest eigenvalue

$$[0,1]^C \mapsto [0,1] : \xi(\boldsymbol{p}) := \lambda_2(\boldsymbol{W}(\boldsymbol{p})) \tag{10}$$

of the stochastic weight matrix $\boldsymbol{W} := \boldsymbol{W}(\boldsymbol{p}) = \sum_i p_i \boldsymbol{W}^{(i)}$ (see [3], [4]). Taking into account the constraints, we arrive at the following optimization problem:

$$\min_{\boldsymbol{p} \in [0;1]^C} \xi(\boldsymbol{p}) \tag{11}$$
$$\text{s.t. } \boldsymbol{W} - \boldsymbol{J} \preceq \xi(\boldsymbol{p})\boldsymbol{I}_N$$
$$\boldsymbol{W} = \sum_{i=1}^C p_i \boldsymbol{W}^{(i)}$$
$$\sum_{i=1}^C p_i = 1$$

where $N\boldsymbol{J} = \mathbf{1}\mathbf{1}^T \in \mathbb{R}^{N \times N}$ is the all-one matrix. Note that the first constraint ensures that the second eigenvalue is smaller than or equal to the largest eigenvalue of $\boldsymbol{W}$, which is 1. An upper bound for the resulting mean squared error can be given with $\boldsymbol{\varepsilon}(t) := \boldsymbol{y}(t) - \overline{\boldsymbol{y}}(0)$ by [3], [4]

$$\mathbb{E}\left\{\boldsymbol{\varepsilon}(t)^T \boldsymbol{\varepsilon}(t)\right\} \leq \xi(\boldsymbol{p})\boldsymbol{\varepsilon}(t-1)^T \boldsymbol{\varepsilon}(t-1) \leq \xi(\boldsymbol{p})^t \boldsymbol{\varepsilon}(0)^T \boldsymbol{\varepsilon}(0)$$

V. JOINT ENERGY/CONVERGENCE OPTIMIZATION

Now we are in a position to extend the optimization problem in (11) to incorporate the additional energy cost. The problem of interest - called the joint convergence/expected lifetime optimization problem - takes the form

$$\min_{\boldsymbol{p} \in [0;1]^C} \xi(\boldsymbol{p}) + \alpha\|\boldsymbol{c}(\boldsymbol{p})\|_1 \tag{12}$$
$$\text{s.t. } \boldsymbol{W} - \boldsymbol{J} \preceq \xi(\boldsymbol{p})\boldsymbol{I}_N$$
$$\boldsymbol{W} = \sum_{i=1}^M p_i \boldsymbol{W}^{(i)}$$
$$\sum_{i=1}^C p_i = 1$$
$$\boldsymbol{c}(\boldsymbol{p}) = \sum_{i=1}^C p_i(\boldsymbol{c}_{\mathrm{fc}}^{(i)} + \boldsymbol{c}_{\mathrm{bc}}^{(i)}).$$

In addition, we assume that $\xi(\boldsymbol{p}) \leq 1 - \epsilon$ for some sufficiently small $\epsilon > 0$. This ensures that the resulting clusters are connected and therefore the algorithm converges. Furthermore, we point out that the associated constraint set $\{\boldsymbol{p} \in [0,1]^C : \xi(\boldsymbol{p}) \leq 1-\epsilon\}$ is convex for any choice of ϵ. This is because the largest eigenvalue of positive semidefinite symmetric matrices is a convex function of the matrix entries which depend linearly on $\boldsymbol{p}$. A practical problem is the choice of ϵ which cannot be too large since otherwise the problem is infeasible.

We emphasize that the set of all possible weight matrices (i.e. clusterings) $\{\boldsymbol{W}^{(i)}\}_{i=1,\ldots,C}$ comprises $C = (N-1)N$ elements (each node can be cluster head of a cluster consisting of $\{2, \ldots, N\}$ nodes), which might result in optimization problems that are very expensive from a computational perspective. However, the set of weight matrices can be reduced if we restrict the number of nodes per cluster. In addition, we can exclude candidate clusters with an outer node as cluster head because these choices correspond to identical consensus steps only at a higher cost, i.e $\boldsymbol{W}^{(i)} = \boldsymbol{W}^{(j)}$, with $\|\boldsymbol{c}^{(i)}\|_1 \geq \|\boldsymbol{c}^{(j)}\|_1$ for some $\mathcal{C}_i = \mathcal{C}_j$, $\mathrm{CH}_i \neq \mathrm{CH}_j$.

VI. SIMULATION RESULTS

In this Section we present simulation results for the proposed algorithm (12) using different predefined cluster sizes $|\mathcal{C}_i| \subseteq \{2, \ldots, 30\}$. The simulated WSN consists of $N = 30$ nodes that are placed uniformly at random in a 50×50 square and to solve the optimiziation problem (12) we use CVX [10]. To ensure convergence of the resulting consensus scheme, the optimization variable $\xi(\boldsymbol{p})$ is bounded by $\xi(\boldsymbol{p}) \leq 1 - \epsilon$. However, this bound must not be set too small, otherwise the optimization problem will not be feasible. We found $\epsilon = 10^{-2}$ to yield good results.

To evaluate the resulting consensus scheme consisting of weight matrices $\boldsymbol{W}^{(i)}$, respective costs $\boldsymbol{c}^{(i)}$ and activation probabilities $\boldsymbol{p}$, we simulate a physical application, where the nodes are deployed to monitor temperature values. The initial network state $\boldsymbol{y}(0)$ is drawn uniformly from $[0\ {}^\circ\mathrm{C}, 30\ {}^\circ\mathrm{C}]^N$. We monitor the error defined by $\frac{\|\boldsymbol{\varepsilon}(t)\|_2^2}{\|\boldsymbol{y}(0)\|_2^2}$ and the consumed

energy for a certain realization of the sequence of activated clusters. The results are averaged over 10^3 runs. For the simulations, we also assume that the nodes have expert knowledge of the current estimation error to terminate communication, and therefore also dynamic energy consumption, once the error falls below a threshold of 10^{-1}.
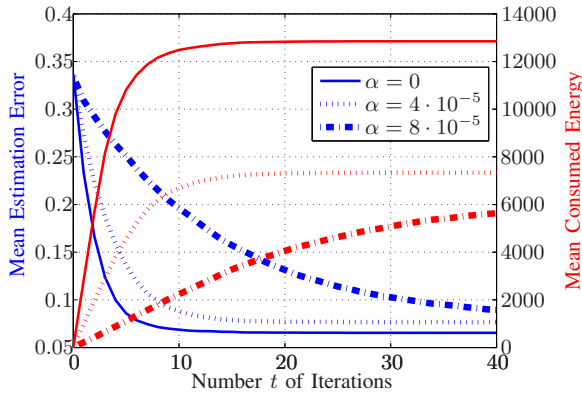

Fig. 2. Proposed consensus algorithm for a WSN with $N = 30$ nodes and small clusters $|\mathcal{C}_i| \subseteq \{2, \ldots, 10\}$.
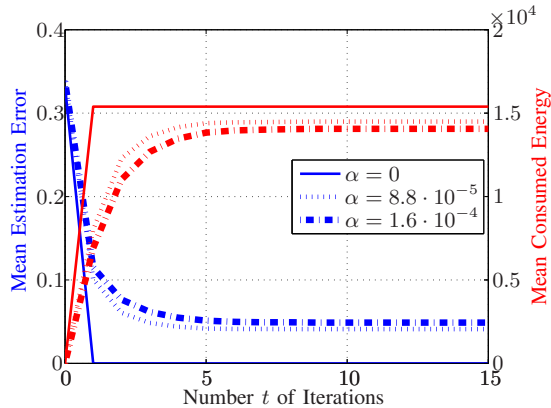

Fig. 3. Proposed consensus algorithm for a WSN with $N = 30$ nodes and large clusters $|\mathcal{C}_i| \subseteq \{20, \ldots, 30\}$.

As can be seen from Fig. 2 and 3, the obtained simulation results confirm the intended behaviour. In fact, if the regularization parameter is chosen to be $\alpha = 0$, the resulting consensus scheme is only tailored towards fast convergence as in (11), neglecting however the energy consumption. Conversely, moving along the path of regularization parameter results in a lower energy consumption at the cost of a slower convergence rate. As an example, the results for a choice $\alpha \in \{4; 8\} \cdot 10^{-5}$ in the small cluster setting and $\alpha \in \{8.8; 16\} \cdot 10^{-5}$ in the large cluster setting are also depicted in Fig. 2 and 3, respectively.

VII. CONCLUSION

In this paper, we presented a novel algorithm for the joint optimization of convergence and energy consumption for consensus algorithms in wireless sensor networks. The proposed algorithm takes into account distance dependent transmit energies and clusters the network nodes according to user-defined cluster sizes that may depend on application and site specific characteristics. By incorporating a regularization term, we investigated the trade-off between convergence speed and energy consumption. This is achieved by a series of numerical simulations of a temperature monitoring application under the assumption of noiseless communication links. The examples show that the naive choice of a single cluster containing all nodes can be outperformed in terms of network energy consumption if a certain excess error can be tolerated.

Of particular interest for our subsequent work will be scenarios involving noisy communications links and time-varying measurement objectives. This can be achieved by using adaptive subgradient methods that harness interference for computation [8]. Also, by using more knowledge about the target application and leveraging ideas from *compressed sensing*, the number of required measurements in the network can be reduced. In turn, this will result in additional savings in energy consumption and introduce new degrees of freedom to the tradeoff between convergence speed and estimation accuracy on one hand, and energy consumption and network robustness on the other hand.

VIII. ACKNOWLEDGMENTS

REFERENCES

[1] B. Nazer and M. Gastpar, "Computation over Multiple-Access Channels," *IEEE Trans. on Inf. Theory*, vol. 53, no. 10, Oct. 2007.

[2] M. Goldenbaum and S. Stanczak, "Robust analog function computation via wireless sensor multiple-access channels," *IEEE Trans. Commun.*, 2013, (to appear, preprint available at http://arxiv.org/abs/1210.2967).

[3] M. Zheng, M. Goldenbaum, S. Stanczak, and H. Yu, "Fast average consensus in clustered wireless sensor networks by superposition gossiping," in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, 2012, pp. 1982–1987.

[4] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transaction on Information Theory*, vol. 52, no. 6, pp. 2508–2530, 2006.

[5] M. Goldenbaum, H. Boche, and S. Stanczak, "Nomographic Gossiping for f-Consensus," in *Proc. 10th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Paderborn, Germany, May 2012.

[6] M. Goldenbaum and S. Stanczak, "Computing Functions via SIMO Multiple-Access Channels: How Much Channel Knowledge Is Needed?," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, USA, March 14–19 2010.

[7] M. Goldenbaum, H. Boche, and S. Stanczak, "Harnessing interference for analog function computation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 61, no. 20, pp. 4893–4906, 2013.

[8] R. L. G. Cavalcante and S. Stanczak, "A distributed subgradient method for dynamic convex optimization problems under noisy information exchange," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 243–256, 2013.

[9] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on*, 2000.

[10] Michael Grant and Stephen Boyd, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," sep 2012.