BROADCAST-BASED DYNAMIC CONSENSUS PROPAGATION IN WIRELESS SENSOR NETWORKS

Valentin Schwarz, Clemens Novak, and Gerald Matz

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gusshausstrasse 25/389, A-1040 Wien, Austria
Tel.: +43 1 58801 38919; Fax: +431 58801 38999; E-mail: valentin.schwarz@nt.tuwien.ac.at

Abstract — Consensus propagation (CP) is an attractive distributed asynchronous averaging algorithm. In this paper, we develop a modified CP protocol for wireless sensor networks (WSN) that uses message broadcast instead of bilateral message exchange, thereby significantly reducing transmit energy. Furthermore, we extend CP to dynamic environments. We show that the distributed nature of the algorithm causes an inherent tracking delay which we propose to compensate by incorporating an appropriately designed linear predictor. The performance of dynamic broadcast CP is assessed via numerical simulations for several WSN topologies.

1. INTRODUCTION

1.1. Background

Distributed inference in wireless sensor networks (WSN) has received much attention recently [1]. In this paper, we consider distributed averaging, which is the basis for more general tasks like distributed maximum likelihood estimation. Our focus will be on *consensus propagation (CP)* [2], which is a particularly promising iterative scheme for distributed averaging. To be more specific, consider a WSN with I sensors. The ith sensor obtains a measurement y_i and the goal is to compute the arithmetic mean $z = \frac{1}{I}\sum_{i=1}^{I} y_i$ in a distributed fashion. During the nth CP iteration, neighboring sensors exchange the messages

$$K_{i\to j}^{(n)} = \frac{1+\sum_{u\in\mathcal{N}(i)\setminus j} K_{u\to i}^{(n-1)}}{1+\frac{1}{\beta}\big(1+\sum_{u\in\mathcal{N}(i)\setminus j} K_{u\to i}^{(n-1)}\big)}, \tag{1a}$$

$$\mu_{i\to j}^{(n)} = \frac{y_i+\sum_{u\in\mathcal{N}(i)\setminus j} \mu_{u\to i}^{(n-1)} K_{u\to i}^{(n-1)}}{1+\sum_{u\in\mathcal{N}(i)\setminus j} K_{u\to i}^{(n-1)}}. \tag{1b}$$

Here, i is the source node, j is the destination node, $\mathcal{N}(i)$ denotes the set of neighbors of node i, and $\beta > 0$ is an algorithm parameter. Note that node i has to transmit $|\mathcal{N}(i)|$ messages during each iteration. The estimate of z at node i and iteration n is given by

$$\hat{z}_i^{(n)} = \frac{y_i+\sum_{u\in\mathcal{N}(i)} \mu_{u\to i}^{(n)} K_{u\to i}^{(n)}}{1+\sum_{u\in\mathcal{N}(i)} K_{u\to i}^{(n)}}. \tag{2}$$

In [2] it was shown that $\hat{z}_i^{(n)}$ asymptotically converges to the true mean z for any i as β and n go to infinity; furthermore, CP has the potential to convergence faster than gossip algorithms (i.e., pairwise averaging) [3]. Even though CP can be applied in an asynchronous manner, we here only consider synchronous operation where each node updates and transmits messages to each neighbor node in every iteration. Our aim in this paper is to investigate the following two topics that have not been previously addressed in the context of CP: i) conventional CP according to (1) and (2) is energy inefficient in the sense that the number of messages being exchanged scales with the number of links which can be very large, specifically for dense WSN; ii) CP up to now has been considered only for static scenarios (i.e., not for repeated measurements of dynamically changing states).

1.2. Contributions

In this paper, we develop extensions of CP that achieve significant energy savings and are applicable to dynamic (time-varying) scenarios. Our specific contributions are:

- In Section 2 we introduce a CP modification in which messages are broadcast to all neighbors simultaneously and we show that this broadcast CP strongly reduces the transmit energy without loss in performance.
- In Section 3 we extend CP to dynamic scenarios (i.e., time-varying parameters) and we study how the performance depends on various algorithm parameters. It turns out that CP acts like a spatio-temporal lowpass filter on the sensor measurements.
- In dynamic scenarios, we observed CP to feature a fundamental delay. To compensate for that delay, we augment CP with linear prediction at each sensor.
- Finally, we present extensive numerical simulations for different network topologies and different parameter dynamics (see Section 4).

This work was supported by FWF Grant N10606 and WWTF Grant ICT08-44.

2. BROADCAST-BASED CONSENSUS PROPAGATION

2.1. Broadcast Messages

To reduce the energy transmitted, we propose a broadcast-based message exchange. In spite of the non-linearity of the CP message updates, this can be achieved by a decomposition of the original CP messages into a common message part followed by a post-processing step at the destination node. Specifically, it can be seen from (1) that each node sends almost identical messages (the only difference is the neighboring node excluded from the sum) to its various neighbors. We exploit this property by i) generating and transmitting a common broadcast message in which no term is excluded from the sum and ii) recovering the original CP messages at the destination nodes.

Mathematically, we rewrite the sums in the messages from source node i to destination node j in (1) by adding and subtracting $K_{j\to i}^{(n-1)}$ resp. $\mu_{j\to i}^{(n-1)} K_{j\to i}^{(n-1)}$:

$$K_{i\to j}^{(n)} = \frac{\left[1+\sum_{u\in\mathcal{N}(i)} K_{u\to i}^{(n-1)}\right] - K_{j\to i}^{(n-1)}}{1+\frac{1}{\beta}\left(\left[1+\sum_{u\in\mathcal{N}(i)} K_{u\to i}^{(n-1)}\right] - K_{j\to i}^{(n-1)}\right)},$$

$$\mu_{i\to j}^{(n)} = \frac{\left[y_i+\sum_{u\in\mathcal{N}(i)} \mu_{u\to i}^{(n-1)} K_{u\to i}^{(n-1)}\right] - \mu_{j\to i}^{(n-1)} K_{j\to i}^{(n-1)}}{\left[1+\sum_{u\in\mathcal{N}(i)} K_{u\to i}^{(n-1)}\right] - K_{j\to i}^{(n-1)}}.$$

By inspection of the above expression, it is seen that the terms in brackets are identical for all destination nodes $j \in \mathcal{N}(i)$ and thus constitute a common message part for all neighboring nodes. Furthermore, all terms outside the brackets, i.e., $K_{j\to i}^{(n-1)}$ and $\mu_{j\to i}^{(n-1)}$, are already known by the destination node j because those were the messages from node j to node i in the previous iteration. Consequently, it is sufficient for each node to broadcast the common messages to all neighbor nodes. During iteration n, node i thus only transmits the messages

$$\tilde{K}_i^{(n)} = 1 + \sum_{u\in\mathcal{N}(i)} K_{u\to i}^{(n-1)}, \qquad (3a)$$

$$\tilde{\mu}_i^{(n)} = y_i + \sum_{u\in\mathcal{N}(i)} \mu_{u\to i}^{(n-1)} K_{u\to i}^{(n-1)}. \qquad (3b)$$

This requires one broadcast transmission instead of $|\mathcal{N}(i)|$ transmissions from node i to all neighbor nodes $j \in \mathcal{N}(i)$. We note that the current estimate in (2) can now be simply obtained as $\hat{z}_i^{(n)} = \tilde{K}_i^{(n)}/\tilde{\mu}_i^{(n)}$. To recover the original CP messages, all destination nodes need to perform the following post-processing steps:

$$K_{i\to j}^{(n)} = \frac{\tilde{K}_i^{(n)} - K_{j\to i}^{(n-1)}}{1+\frac{1}{\beta}\left(\tilde{K}_i^{(n)} - K_{j\to i}^{(n-1)}\right)}, \qquad (4a)$$

$$\mu_{i\to j}^{(n)} = \frac{\tilde{\mu}_i^{(n)} - \mu_{j\to i}^{(n-1)} K_{j\to i}^{(n-1)}}{\tilde{K}_i^{(n)} - K_{j\to i}^{(n-1)}}, \qquad (4b)$$

A schematic illustration of the operational principle underlying conventional and broadcast CP is given in Fig. 1. We

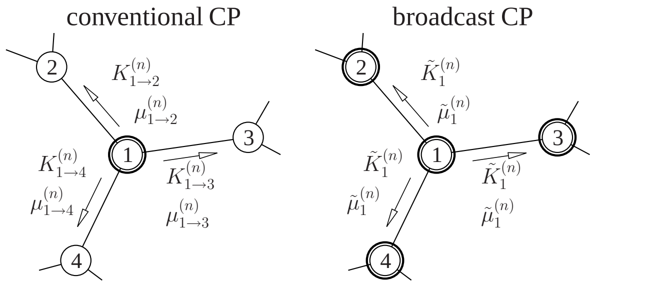

Fig. 1. *Exemplary comparison of conventional CP (left) and broadcast CP (right): with conventional CP, node #1 computes and transmits three different message pairs; with broadcast CP, node #1 computes and transmits only one message pair and nodes #2, #3, and #4 perform a post-processing. (Nodes performing computations are encircled with thick solid lines.)*

emphasize that broadcast CP yields exactly the same results as traditional CP and therefore all CP convergence studies apply to broadcast CP as well. However, if the messages are quantized or if there are transmission errors, the results do not coincide anymore. Due to lack of space, these issues are not addressed in this paper.

It can be verified that conventional CP and our proposed broadcast CP scheme have the same overall computational complexity that scales linearly with the number of edges (links) in the network. However, broadcast CP requires much fewer transmissions and hence less transmit energy (see next Section). In addition, broadcast CP has the advantage that energy consumption is completely homogenous: all nodes spend the same amount of energy per iteration. This is in striking contrast to conventional CP in which well-connected nodes (e.g. hubs in scale-free networks, see below) expend much more energy than nodes with few neighbors; this property of conventional CP is certainly undesirable in WSN since all nodes in general will be equipped with identical power supplies (batteries).

2.2. Analysis of Energy Savings

In this section we analyze the transmit energy savings achievable with broadcast CP. We consider different WSN topologies modeled by undirected graphs $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V} = \{1, \ldots, I\}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set (links between nodes). To quantify the energy savings, we assume that the amount of energy required to transmit one message pair (either (1) or (3)) is the same for conventional and broadcast CP. We further assume a flooding schedule in which all nodes update their messages simultaneously during each iteration. The energy saving factor η is then defined as the number of conventional CP message pairs per iteration divided by the number of broadcast CP message pairs per iteration.

With conventional CP, message pairs are transmitted in both direction across every edge (i.e., from node i to node j and from node j to node i, for $(i, j) \in \mathcal{E}$), accounting for a total of $2|\mathcal{E}|$ message pairs. In contrast, with broadcast CP

network model	$\mathrm{E}\{\eta\}$
random r-regular graphs	r
scale-free graphs	$2c$
random geometric graphs	pI

Table 1. *Mean energy saving factor of broadcast CP for different network models.*

each node transmits only one message pair per iteration to its neighbors, resulting in $|\mathcal{V}| = I$ message pairs in total. We emphasize that the number of messages for broadcast CP is independent of the network topology and depends only on the network size. Using these results, the energy saving factor is given by

$$\eta = \frac{2|\mathcal{E}|}{|\mathcal{V}|} = \frac{2|\mathcal{E}|}{I}\,. \tag{5}$$

It is interesting to observe that due to the relation $|\mathcal{E}| = \frac{1}{2}\sum_{i=1}^{I}|\mathcal{N}(i)|$ we have $\eta = \frac{1}{I}\sum_{i=1}^{I}|\mathcal{N}(i)|$, i.e., the energy saving factor equals the average number of neighbors. We note that for non-tree-structured connected networks $|\mathcal{E}| \geq I$ and hence $\eta \geq 2$, i.e., the energy saving factor is lower bounded by 2. We will next evaluate the expectation $\mathrm{E}\{\eta\}$ of (5) for various random network models. A summary of the results below is provided in Table 1.

Random r-regular networks [4]. With this type of network, each node is connected randomly with exactly $r \geq 2$ nodes, i.e., $|\mathcal{N}(i)| = r$ for all i. Hence, the number of edges is given by $|\mathcal{E}| = \frac{1}{2}rI$ for any network realization. This corresponds to an energy saving factor of $\mathrm{E}\{\eta\} = r$, independent of the network size.

Scale-free graphs. We consider scale-free graphs which are constructed via the Barabási-Albert algorithm [5] as follows. The network construction is initialized using an arbitrary (but connected) small graph $\mathcal{G}_0 = (\mathcal{V}_0, \mathcal{E}_0)$. New nodes are then added in an iterative manner and connected to a prescribed number $c \geq 1$ of existing nodes; these c nodes are chosen with a probability that scales with their current number of neighbors, i.e., new nodes are more likely to attach to highly connected nodes. The procedure stops when a target size I of the network is reached. Even though the graphs are random, the number of edges is fixed and equals $|\mathcal{E}| = |\mathcal{E}_c| + c(I - |\mathcal{V}_c|) \approx c\,I$ for any network realization. The energy saving factor in this case is thus approximately equal to $\mathrm{E}\{\eta\} \approx 2c$, which is independent of the size of the network.

Random geometric graphs [6]. This type of networks is constructed based on the geometry of the node placement. In a first step, the nodes are randomly distributed in a d-dimensional region $\mathcal{A}$ according to a uniform distribution. In a second step, node pairs whose Euclidean distance is less than a given value R are connected by edges. Such graphs model well actual WSN deployments.

A randomly selected node is connected to a given node i if it lies within a hypersphere of radius R and volume $V_d = \frac{1}{\Gamma(d/2+1)}\pi^{d/2}R^d$ centered about node i. This event happens with probability[1] $p = \frac{V_d}{|\mathcal{A}|}$. This implies that the expected number of neighbors for any node i equals $\mathrm{E}\{\mathcal{N}(i)\} = p(I-1)$ and thus in turn $\mathrm{E}\{\eta\} = \frac{1}{I}\sum_{i=1}^{I}\mathrm{E}\{\mathcal{N}(i)\} = p(I-1)$. It is seen that the energy savings increase with increasing number of nodes (provided $|\mathcal{A}|$ and R remain unchanged). This indicates that in dense WSN broadcast CP has a huge advantage over conventional CP, which was indeed confirmed in our numerical simulations.

3. DYNAMIC CONSENSUS PROPAGATION

3.1. Basic Idea

We next discuss the application of broadcast CP to dynamic environments, i.e., averaging of time-varying signals. We assume that at sampling time mT (T is the sampling period) node i measures the signal

$$y_i[m] = s[m] + w_i[m]\,. \tag{6}$$

Here, $s[m]$ is the signal of interest and $w_i[m]$ represents additive noise. Approximate computation of the time-varying mean $z[m] = \frac{1}{I}\sum_{i=1}^{I}y_i[m]$ using CP is challenging since CP needs to effectively converge within one sampling period; for practical applications this is inefficient or even impossible. We therefore propose a different method in which we apply CP in an online manner (convergence to the true values will no longer be guaranteed in this case). Here, L CP iterations are performed during each sampling period using the last measurement; the messages from the Lth iteration are retained for the next sampling period. This amounts to replacing the μ-updates in (3b) with

$$\tilde{\mu}_i^{(n)} = y_i[m] + \sum_{u\in\mathcal{N}(i)} \mu_{u\to i}^{(n-1)} K_{u\to i}^{(n-1)}$$

with $m = \left\lfloor \frac{n}{L} \right\rfloor$. The K-update in (3a) and the post-processing in (4) do not depend on the measurements and hence remain unchanged. After the last CP iteration during the mth sampling period (i.e., for $n = mL+(L-1)$) the estimate of $z[m]$ at node i is obtained as

$$\hat{z}_i[m] = \frac{\tilde{\mu}_i^{(mL+L-1)}}{\tilde{K}_i^{(mL+L-1)}}\,. \tag{7}$$

The performance of dynamic CP strongly depends on the choice of the number L of iterations per sampling period, on the attenuation parameter β in the K-updates, and on the WSN topology. The parameter L amounts to trading off good estimation performance (requires large L) with low complexity (achieved for small L). If the signal dynamics are small (i.e., $s[m]$ varies only slowly) and/or the WSN has a small diameter (which entails faster convergence), tracking is not difficult and β should be chosen large to achieve small errors. In contrast, for rapidly varying signals and/or WSN with a large diameter it is important to sufficiently fast track the signal; this requires choosing a small β.

[1] We here assume that the hypersphere is always a subset of $\mathcal{A}$; this amounts to ignoring boundary effects.

3.2. Dynamic CP with Linear Prediction

Our simulation results (cf. Section 4) indicated that the spatio-temporal memory of dynamic CP causes $\hat{z}_i[m]$ to lag behind the actual mean $z[m]$, i.e., there is a fundamental delay τ between $z[m]$ and $\hat{z}_i[m]$. In time-critical applications this delay is undesirable, which motivates us to augment CP with linear prediction (LP), i.e., we replace $y_i[m]$ in the CP iterations by

$$\hat{y}_i[m+h] = \sum_{k=0}^{K-1} a_k\, y_i[m-k]. \quad (8)$$

Here, h and K denote the predictor horizon and memory, respectively, and $\mathbf{a} = (a_0 \dots a_{K-1})^T$ are appropriately designed predictor coefficients. The idea is to choose $h^* = \tau$ to compensate for the CP delay.

We assume that signal and noise in (6) are stationary and that the noise is spatio-temporally i.i.d. and zero-mean:

$$\mathrm{E}\{w_i[m] w_j[m']\} = \sigma_w^2\, \delta_{m-m'}\, \delta_{i-j}\,.$$

Our first approach aims at scenarios where the noise variance is unknown and the predictor can only be designed to predict the noiseless signal $s[m]$ based on the vector $\mathbf{s}[m] = (s[m] \;\dots s[m-K+1])^T$. The MMSE-optimal predictor coefficients are obtained by solving the Wiener-Hopf equation

$$\mathbf{R}_\mathbf{s}\, \mathbf{a} = \mathbf{r}_\mathbf{s}$$

with

$$\mathbf{R}_\mathbf{s} = \mathrm{E}\left\{\mathbf{s}[m]\, \mathbf{s}^T[m]\right\} \quad \text{and} \quad \mathbf{r}_\mathbf{s} = \mathrm{E}\left\{\mathbf{s}[m] s[m+h]\right\}\,.$$

Since $z[m] = s[m] + \bar{w}[m]$ with $\bar{w}[m] = \frac{1}{I}\sum_{i=1}^{I} w_i[m]$, the resulting predictor coefficients can also be used to predict $z[m]$. Using (8) it follows that

$$\hat{z}[m+h] = \sum_{k=0}^{K-1} a_k\, z[m-k] = \sum_{k=0}^{K-1} a_k\, \frac{1}{I}\sum_{i=1}^{I} y_i[m-k]$$
$$= \frac{1}{I}\sum_{i=1}^{I} \hat{y}_i[m+h],$$

which shows that predicting the CP output is equivalent to applying CP to the predicted measurements.

In case the noise variance σ_w^2 is known, additional noise suppression can be built into the predictor. The corresponding coefficients for such a noisy predictor are obtained by solving the new Wiener-Hopf equation $(\mathbf{R}_s + \sigma_w^2 \mathbf{I}_K)\mathbf{a} = \mathbf{r}_\mathbf{s}$ ($\mathbf{I}_K$ denotes the $K \times K$ identity matrix). Here we further assumed that the noise $w_i[m]$ and the signal $s[m]$ are statistically independent.

4. NUMERICAL RESULTS

We next show simulation results for dynamic broadcast CP for different WSN topologies obtained with three random graph models: i) random 3-regular graphs [4], ii) random geometric graphs [6], and iii) Barabási-Albert scale-free graphs [5]. Unless stated otherwise, the results shown pertain to the case $L = 1$, i.e., one CP iteration per sampling period.

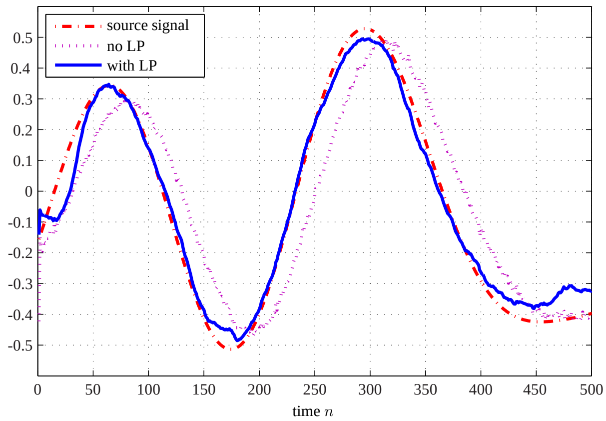

Fig. 2. *Illustration of dynamic broadcast CP at one node for a random geometric graph ($I = 100$, $\beta = 2$, SNR $= 3$ dB.*

4.1. Dynamic CP without Prediction

We first consider the application of dynamic CP without linear prediction to the signal model (6). The signal $s[m]$ was a stationary lowpass process of bandwidth Θ and the signal-to-noise ratio (SNR) was 3 dB. The network size was $I = 100$. A realization of $s[m]$ and a corresponding CP estimate are shown in Fig. 2 (dashed and dotted line, respectively) for the case of a random geometric graph. In order for CP to be able to track the signal variations, the attenuation parameter in (4a) had to be chosen as $\beta = 2$.

To systematically assess CP performance, we use the mean-squared error (MSE) $\bar{\epsilon}(\beta, \tau)$ which is obtained by averaging

$$\epsilon(\beta,\tau) = \frac{1}{MI}\sum_{m=1}^{M}\sum_{i=1}^{I}\frac{(s[m] - \hat{z}_i[m+\tau])^2}{\mathrm{E}\{s^2[m]\}}\,,$$

over 1000 network realizations (the signal length was $M = 700$). Note that the non-causal time shift of $\hat{z}_i[m+\tau]$ by τ cannot be realized in practice.

Fig. 3 shows the optimal β^* and τ^* minimizing $\bar{\epsilon}(\beta,\tau)$ as functions of Θ for various WSN topologies. The minimal MSE $\bar{\epsilon}(\beta^*, \tau^*)$ versus Θ is displayed in Fig. 4. It is seen that β^* and τ^* essentially scale inversely proportionally with the signal bandwidth Θ. Furthermore, increasing signal dynamics result in poorer estimation performance (i.e., increasing MSE $\bar{\epsilon}(\beta^*, \tau^*)$). The most favorable topology is 3-regular, closely followed by scale-free. The noticeably poorer performance with random geometric graphs can be improved significantly by increasing the number of CP iterations to $L = 20$ (this also allows for larger β^* and smaller delay τ^*).

Fig. 5 illustrates how β^* and τ^* jointly depend on network size I and signal dynamics Θ for random geometric graphs. Since β^* and τ^* both depend inversely proportionally on Θ, they together approximately form a straight line with slope 1. More surprisingly, the pair (β^*, τ^*) turns out to be approximately independent of the network size I. This can be explained by the fact that (β^*, τ^*) is determined by the local network structure which does not change with the network size.

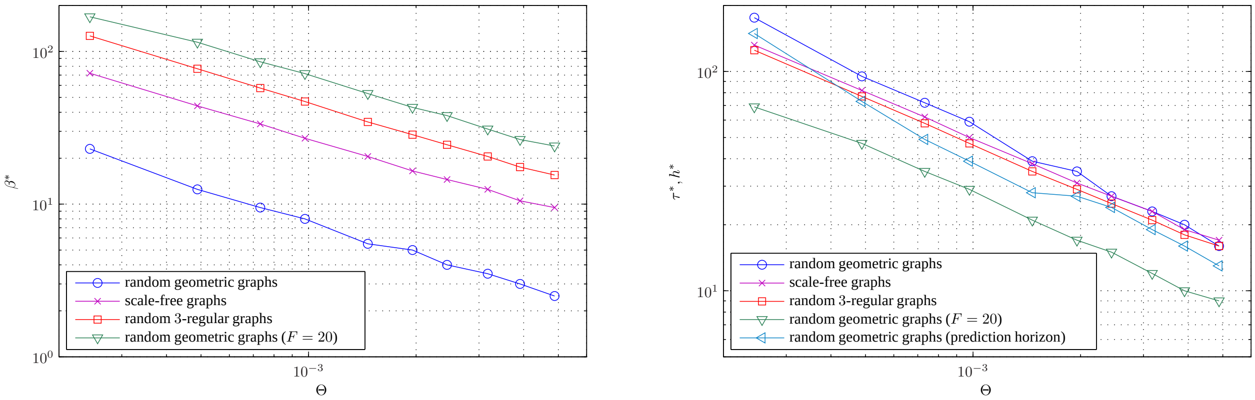

Fig. 3. *Optimal β^* (left) and τ^* (right) versus signal bandwidth Θ for different types of WSN. The right-hand side plot also shows the optimal prediction horizon h^*.*

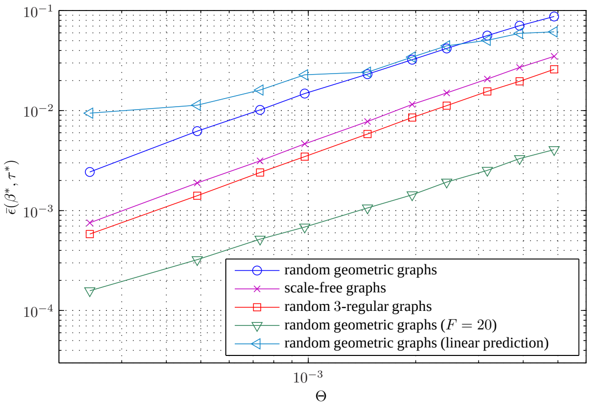

Fig. 4. *Optimal MSE versus signal bandwidth Θ for different types of WSN.*

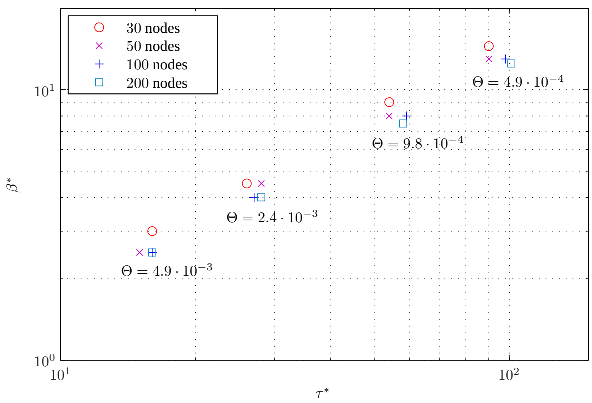

Fig. 5. *Optimal CP parameters (β^*, τ^*) for various network sizes I and signal bandwidths Θ.*

4.2. Dynamic CP with Prediction

We next consider the same setup as before and augment CP with optimal (noisy) prediction. An example realization is shown in Fig. 2 (solid line). We here provide results for random geometric graphs with $I = 100$, $M = 1000$ and a predictor length of $K = 50$, using the optimal β^* found previously. The optimal prediction horizon h^*, which is approximately equal to τ^*, is depicted on the right-hand side of Fig. 3. The corresponding MSE, shown in Fig. 4, is smaller than the MSE achieved by delay-compensated CP without prediction for $\Theta > 3 \cdot 10^{-3}$. This is due to the additional noise suppression achieved by the predictor. For small bandwidths Θ, the MSE of CP with prediction degrades since smaller Θ requires larger prediction horizons and hence entails poorer prediction performance.

5. CONCLUSIONS

The adaption of consensus propagation (CP) to a broadcast based message passing algorithm increases its attractiveness for wireless sensor networks. In particular we showed that for random regular, scale-free and random geometric graph topologies the transmit energy can be reduced significantly. Moreover we proposed a modification of CP which is applicable to dynamic/time-varying scenarios. Our simulation results confirmed that dynamic broadcast CP offers very attractive tracking abilities, specifcially when augmented with a linear predictor to compensate for the tracking delay.

6. REFERENCES

[1] A. Swami, Q. Zhao, Y.-W. Hong, and L. Tong, *Wireless Sensor Networks: Signal Processing and Communications Perspectives*. John Wiley and Sons Ltd, 2007.

[2] C. C. Moallemi and B. Van Roy, “Consensus propagation,” *IEEE Trans. Inf. Theory*, vol. 52, pp. 4753–4766, Nov. 2006.

[3] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, “Randomized gossip algorithms,” *IEEE Trans. Inf. Theory*, vol. 52, pp. 2508–2530, June 2006.

[4] B. Bollobás, *Random Graphs*. Cambridge Univ Press, 2001.

[5] R. Albert and A.-L. Barabási, “Emergence of scaling in random networks,” *Science*, vol. 286, pp. 509–512, Oct. 1999.

[6] M. Penrose, *Random Geometric Graphs*. Oxford University Press, 2003.