SYSTEM-THEORETIC FORMULATION AND ANALYSIS OF DYNAMIC CONSENSUS PROPAGATION

Valentin Schwarz and Gerald Matz

Institute of Communications and Radio-Frequency Engineering, Vienna University of Technology
Gußhausstraße 25/389, A-1040 Vienna, Austria; email: {vschwarz,gmatz}@nt.tuwien.ac.at

ABSTRACT

In our previous work we have proposed a dynamic version of the consensus propagation (CP) algorithm introduced by Moallemi and Van Roy. Here, we pursue a system theoretic approach to the analysis of CP. Specifically, we first develop a state-space model for CP and then use this model to prove stability of CP when applied to time-varying processes. We further show how the state-space model can be used to describe the transfer characteristics (spatio-temporal filtering) of CP in terms of attenuation and group delay, both for pure CP and an extended version that uses local linear predictors at the processing nodes. Numerical simulations illustrate our findings.

***Index Terms*—** Consensus propagation, state space model, wireless sensor networks, linear predictor, distributed inference

1. INTRODUCTION

Average consensus (AC) [1] is a popular method for distributed averaging based on consensus [2]. Gossip methods [3] can be viewed as asynchronous versions of AC. All these methods are well studied and tight bounds on the convergence speed have been derived. These analyses have been simplified by the simple structure of AC and gossip averaging in which messages, estimates, and states are identical. In this paper we consider a less well-known algorithm termed consensus propagation (CP) [4]. CP uses distributed Gaussian belief propagation to solve a convex optimization problem which yields the desired average. The convergence of CP has been proved in [4, 5]; the convergence proofs for Gaussian belief propagation in [6, 7] apply to CP as well. It is important to note that CP is algorithmically very different from AC in spite of the fact that both are distributed averaging algorithms (some authors incorrectly state that CP is identical to AC with adaptive weights). Thus, the results obtained for AC cannot be applied to CP. We note that a brief numerical comparison of AC and CP in the context of wireless sensor networks (WSNs) is provided in [8].

CP is applied on an undirected (connected) graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ denotes the vertex set and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ the edge set. The number of vertices and the number of edges are denoted by $I = |\mathcal{V}|$ and $E = |\mathcal{E}|$, respectively. In the formulation of CP, it is favorable to use the directed graph obtained from $\mathcal{G}$ by replacing each undirected edge with two directed edges in opposite directions (i.e., to an edge from node i to node j there exists an edge from node j to node i). The number of edges in this directed graph equals $E_d = 2E$. In many wireless communication scenarios the graph $\mathcal{G}$ is modeled as a random geometric graph [9]; in contrast, the world-wide web has the structure of a scale-free network [10]. CP can deal with arbitrary graph topologies but the latter affects the convergence behavior as shown in [11].

Funded by WWTF Grant ICT-044 and FWF Grant S10606.

In this paper, we provide a reformulation of dynamic CP in terms of a linear state space model. This state space representation is the basis for the core theoretical results of the paper which establishes the stability of dynamic CP. In addition, we use system theoretic concepts like the transfer function, amplitude response, and group delay to gain interesting insights regarding the behavior of CP. Using these tools it is finally shown that augmenting CP with local linear predictors at the nodes indeed has the potential to compensate the delays inherent to dynamic CP. We note that in the context of AC, prediction of the states has been proposed to improve the convergence speed (e.g. [12]), which is different from our case, where the predictors serve to compensate estimation time lags.

2. CP BASICS

In the dynamic version of CP, the ith node observes during the nth sampling interval the sample $z_n^{(i)}$ (e.g., a measurement of a physical quantity). Each node is assumed to know its set of neighbors $\mathcal{N}(i) = \{j : (i,j) \in \mathcal{E}\}$. CP estimates the arithmetic mean $\bar{z}_n = \frac{1}{I}\sum_{i=1}^{I} z_n^{(i)}$ in an iterative manner. To this end, two different messages are exchanged between neighboring nodes either synchronously[1] (at each iteration messages are exchanged on all edges) or asynchronously (message updates are performed arbitrarily, see [4]). In this work we consider the synchronous case. At iteration n, each node sends messages to each of its neighbor nodes, i.e., for node i and node $j \in \mathcal{N}(i)$ there is

$$K_n^{(i\to j)} = \frac{1 + \sum_{v\in\mathcal{N}(i)\setminus j} K_{n-1}^{(v\to i)}}{1 + \frac{1}{\beta}\left(1 + \sum_{v\in\mathcal{N}(i)\setminus j} K_{n-1}^{(v\to i)}\right)}, \tag{1}$$

$$\mu_n^{(i\to j)} = \frac{z_n^{(i)} + \sum_{v\in\mathcal{N}(i)\setminus j} \mu_{n-1}^{(v\to i)} K_{n-1}^{(v\to i)}}{1 + \sum_{v\in\mathcal{N}(i)\setminus j} K_{n-1}^{(v\to i)}}. \tag{2}$$

Here, β is a positive, real-valued parameter that determines the agility of CP in reacting to changing inputs. In the above formulation, the messages depend on the destination node j; in [11] we have developed a completely equivalent broadcast version that significantly reduces the overall transmit power in WSNs and requires each node to transmit only a single message pair per iteration. Each

[1] Note that this does not require accurate temporal synchronization; rather, only the iteration counter has to be coordinated. We furthermore tacitly assume that one CP update is performed per sampling interval.

node computes its local estimate of the average according to

$$\hat{\bar{z}}_n^{(i)} = \frac{z_n^{(i)} + \sum_{v\in\mathcal{N}(i)} \mu_{n-1}^{(v\to i)} K_{n-1}^{(v\to i)}}{1 + \sum_{v\in\mathcal{N}(i)} K_{n-1}^{(v\to i)}}. \qquad (3)$$

In the static case ($z_n^{(i)} = z^{(i)}$ for all n), it was proved in [4, 5] that all local estimates converge to the true mean, $\lim_{n\to\infty} \hat{\bar{z}}_n^{(i)} = \bar{z}$ as β goes to infinity.

3. MATRIX-VECTOR FORMULATION

Incidence matrix. For the derivation of the CP state space model we need modified definitions of the incidence matrix of a graph. The conventional undirected incidence matrix has size $I \times E$; its (i, j)th element indicates whether the i node is part of the jth edge. In contrast, we define two $I \times E_d$ incidence matrices $\boldsymbol{B}$ and $\boldsymbol{B}'$ that represent the incoming edges and the outgoing edges, respectively. We have $b_{ij} = 1$ if and only if edge j starts in node i; similarly, $b'_{ij} = 1$ if and only if edge j ends in node i. In our case, any two neighboring nodes are connected by two edges in opposite directions. Here, $\boldsymbol{B}$ and $\boldsymbol{B}'$ are related via a simple permutation, i.e., $\boldsymbol{B} = \boldsymbol{B}'\boldsymbol{P}$, where $\boldsymbol{P}$ denotes an $E_d \times E_d$ permutation matrix.

CP in matrix/vector notation. We first develop simple matrix-vector expressions for the CP messages defined in (1) and (2) by arranging all messages at iteration n into column vectors $\boldsymbol{k}_n$ and $\boldsymbol{\mu}_n$ of length E_d. Starting with (1), we obtain for the K-messages after some manipulation $\boldsymbol{k}_n = \mathrm{diag}^{-1}\{\mathbf{1} + \frac{1}{\beta}\boldsymbol{c}_n\}\, \boldsymbol{c}_n$. Here $\mathrm{diag}\{\mathbf{x}\}$ denotes the diagonal matrix having the elements of the vector $\mathbf{x}$ on its diagonal and the vector $\boldsymbol{c}_n$ is given by ($\mathbf{1}$ is the all-ones vector of appropriate dimension)

$$\boldsymbol{c}_n = \mathbf{1} + \boldsymbol{P}\big(\boldsymbol{B}^T\boldsymbol{B} - \boldsymbol{I}\big)\boldsymbol{k}_{n-1}\,.$$

Similarly, we have $\boldsymbol{\mu}_n = \mathrm{diag}^{-1}\{\boldsymbol{c}_n\}\, \boldsymbol{d}_n$ where

$$\boldsymbol{d}_n = \boldsymbol{P}\boldsymbol{B}^T\boldsymbol{z}_n + \boldsymbol{P}\big(\boldsymbol{B}^T\boldsymbol{B} - \boldsymbol{I}\big)\mathrm{diag}\{\boldsymbol{k}_{n-1}\}\boldsymbol{\mu}_{n-1}\,.$$

and $\boldsymbol{z}_n = [z_n^{(1)}\; z_n^{(2)}\; \ldots\; z_n^{(I)}]^T$ is the length-I vector of observations. For the length-I vector of estimates $\hat{\bar{\boldsymbol{z}}}_n$, one can show that (cf. (3))

$$\hat{\bar{\boldsymbol{z}}}_n = \mathrm{diag}^{-1}\{\mathbf{1} + \boldsymbol{B}\boldsymbol{k}_{n-1}\}\big(\boldsymbol{z}_n + \boldsymbol{B}\mathrm{diag}\{\boldsymbol{k}_{n-1}\}\boldsymbol{\mu}_{n-1}\big).$$

We would like to model CP as a multiple-input multiple-output system that takes the measurement vector $\boldsymbol{z}_n$ as input and delivers the estimation vector $\hat{\bar{\boldsymbol{z}}}_n$ as output. Unfortunately, this is complicated by the fact that $\hat{\bar{\boldsymbol{z}}}_n$ and $\boldsymbol{\mu}_n$ all depend *non-linearly* on $\boldsymbol{k}_n$, which in addition depends on the iteration index. Hence, we can neither view $\boldsymbol{k}_n$ as a constant parameter, nor can we incorporate it into the state vector of a linear state space model.

4. CP SYSTEM THEORY

CP state space model. To establish a simple state space model for CP, we observe that the K-updates are independent of the observations $\boldsymbol{z}_n$. Furthermore, our dynamic setup necessitates a small value for β, which ensures that the K-updates converge much quicker than with large β in the static case (cf. [11]). Hence, it is reasonable to assume that the K-updates have converged to a fixed value $\boldsymbol{k} = \lim_{n\to\infty} \boldsymbol{k}_n$; essentially, this means that we start the iterations for $\hat{\bar{\boldsymbol{z}}}_n$ and $\boldsymbol{\mu}_n$ only after convergence of the K-messages so that $\boldsymbol{k}$ can be viewed as a fixed parameter. This allows us to view CP as a *linear time-invariant* spatio-temporal filter characterized by a state-space model in which the μ-messages are viewed as state parameters:

$$\boldsymbol{\mu}_n = \boldsymbol{\Phi}\boldsymbol{\mu}_{n-1} + \boldsymbol{\Gamma}\boldsymbol{z}_n \qquad (4a)$$

$$\hat{\bar{\boldsymbol{z}}}_n = \boldsymbol{C}\boldsymbol{\mu}_{n-1} + \boldsymbol{D}\boldsymbol{z}_n\,. \qquad (4b)$$

The matrices $\boldsymbol{\Phi}$ ($E_d \times E_d$), $\boldsymbol{\Gamma}$ ($E_d \times I$), $\boldsymbol{C}$ ($I \times E_d$), and $\boldsymbol{D}$ ($I \times I$) follow from the results in the previous section. To simplify notation, we define the $E_d \times E_d$ matrix

$$\boldsymbol{H} = \boldsymbol{P}\big(\boldsymbol{B}^T\boldsymbol{B} - \boldsymbol{I}\big),$$

which allows us two write

$$\begin{aligned}
\boldsymbol{\Phi} &= \mathrm{diag}^{-1}\{\mathbf{1} + \boldsymbol{H}\boldsymbol{k}\}\,\boldsymbol{H}\,\mathrm{diag}\{\boldsymbol{k}\}\,, \qquad (5)\\
\boldsymbol{\Gamma} &= \mathrm{diag}^{-1}\{\mathbf{1} + \boldsymbol{H}\boldsymbol{k}\}\,\boldsymbol{P}\boldsymbol{B}^T\,,\\
\boldsymbol{C} &= \mathrm{diag}^{-1}\{\mathbf{1} + \boldsymbol{B}\boldsymbol{k}\}\,\boldsymbol{B}\,\mathrm{diag}\{\boldsymbol{k}\}\,,\\
\boldsymbol{D} &= \mathrm{diag}^{-1}\{\mathbf{1} + \boldsymbol{B}\boldsymbol{k}\}\,.
\end{aligned}$$

Stability of CP. Based on the state-space model formulated above, we next establish asymptotic stability [13] of dynamic CP in the sense that any transient input will eventually die out, which is equivalent to the condition that in case of zero inputs ($\boldsymbol{z}_n = \mathbf{0}$) the state vector tends to zero for any initial inner state $\boldsymbol{\mu}_0$ as time (iteration count) goes to infinity, i.e., $\lim_{n\to\infty} \boldsymbol{\mu}_n = \mathbf{0}$. We now state our main theoretical result:

Theorem. *For fixed non-negative K-values, the CP state-space formulation* (4) *is asymptotically stable for any graph $\mathcal{G}$.*

Proof. According to [13], a discrete-time system is asymptotically stable if and only if $|\lambda_i(\boldsymbol{\Phi})| < 1$, where $\lambda_i(\boldsymbol{\Phi})$ denotes the eigenvalues of $\boldsymbol{\Phi}$. The elements of $\boldsymbol{\Phi}$ can be shown to be given by (cf. (5))

$$\phi_{ij} = [\boldsymbol{\Phi}]_{ij} = \frac{h_{ij}k_j}{1 + \sum_l h_{il}k_l}\,.$$

The definition of the 0/1-valued incidence matrix $\boldsymbol{B}$ implies that the matrix $\boldsymbol{B}^T\boldsymbol{B}$ has non-negative elements and its diagonal elements all equal 1. Hence, $\boldsymbol{B}^T\boldsymbol{B} - \boldsymbol{I}$ has non-negative elements as well which in turn implies $h_{ij} \geq 0$ and further $\phi_{ij} \geq 0$. It follows that (cf. [14, Sections 2.3.1 and 2.3.2])

$$\|\boldsymbol{\Phi}\|_\infty \leq \max_i \sum_j |\phi_{ij}| = \max_i \sum_j \frac{h_{ij}k_j}{1 + \sum_l h_{il}k_l} < 1.$$

Since the spectral radius $\rho(\boldsymbol{\Phi}) = \max_i\{|\lambda_i(\boldsymbol{\Phi})|\}$ is upper bounded by any matrix norm, we have $|\lambda_i(\boldsymbol{\Phi})| \leq \rho(\boldsymbol{\Phi}) \leq \|\boldsymbol{\Phi}\|_\infty < 1$, thereby establishing asymptotic stability. □

Predictive CP. In [11] we proposed to augment CP with per-node linear predictors to compensate for the time lag observed for dynamic CP. The same predictor filter $\boldsymbol{a} = [a_0\, a_1\, \ldots\, a_{L-1}]^T$, characterized by a filter length L and a prediction horizon h, operates on the observed signals $z_n^{(i)}$, yielding the predicted signals

$$\tilde{\boldsymbol{z}}_{n+h} = [\boldsymbol{z}_n\; \boldsymbol{z}_{n-1}\; \ldots\; \boldsymbol{z}_{n-L+1}]\,\boldsymbol{a}\,, \qquad (6)$$

which now constitute the CP input. The predictor memory can be incorporated into our CP state space model by augmenting the state vector as

$$\tilde{\boldsymbol{\mu}}_n = \Big[\boldsymbol{\mu}_n^T\; \boldsymbol{z}_{n-1}^T\; \ldots\; \boldsymbol{z}_{n-L+1}^T\Big]^T\,.$$

The length of this extended state vector equals $E_d + I(L-1)$. The augmented state space model for predictive CP reads

$$\tilde{\boldsymbol{\mu}}_{n+1} = \tilde{\boldsymbol{\Phi}}\tilde{\boldsymbol{\mu}}_n + \tilde{\boldsymbol{\Gamma}}\boldsymbol{z}_n \tag{7a}$$

$$\hat{\tilde{\boldsymbol{z}}}_{n+h} = \tilde{\boldsymbol{C}}\tilde{\boldsymbol{\mu}}_n + \tilde{\boldsymbol{D}}\boldsymbol{z}_n\,, \tag{7b}$$

with the matrices

$$\tilde{\boldsymbol{\Phi}} = \underbrace{\left[\begin{array}{c|ccc|c} \boldsymbol{\Phi} & & \boldsymbol{a}'\otimes\boldsymbol{\Gamma} & & \\ \hline \boldsymbol{0} & & \boldsymbol{0} & & \boldsymbol{0} \\ \hline & 1 & & & \\ \boldsymbol{0} & & \ddots & & \boldsymbol{0} \\ & & & 1 & \end{array}\right]}_{E_d \quad (L-2)I \quad I} \begin{array}{l} \} \; E_d \\ \} \; I \\ \\ \} \; (L-2)I \\ \\ \end{array}$$

$$\tilde{\boldsymbol{\Gamma}} = \left[\begin{array}{ccc} & a^{(0)}\boldsymbol{\Gamma} & \\ \hline 1 & & \\ & \ddots & \\ & & 1 \\ \hline & \boldsymbol{0} & \end{array}\right] \begin{array}{l} \} \; E_d \\ \\ \} \; (L-2)I \\ \\ \} \; I \end{array}$$

$$\tilde{\boldsymbol{C}} = \big[\, \underbrace{\boldsymbol{C}}_{E_d} \,\big|\, \underbrace{\boldsymbol{a}'\otimes\boldsymbol{\Gamma}}_{(L-1)I} \,\big]$$

$$\tilde{\boldsymbol{D}} = a^{(0)}\boldsymbol{D}$$

with $\boldsymbol{a}' = [a_1\, a_2\, \ldots\, a_{L-1}]^T$.

Stability of predictive CP follows by a simple argument (it could be proved directly based on (7) as well). Specifically, plain CP itself is stable according to the theorem in the previous subsection and the predictor (6) is stable since it has finite impulse response. Since the the concatenation of two stable systems is again stable, it follows that predictive CP is stable.

Frequency-domain analysis. The state space models (4) and (7) are useful since they allow us to apply system-theoretic analysis tools to CP. Specifically, we can obtain a frequency domain characterization of CP in terms of the $I \times I$ transfer function matrix

$$\boldsymbol{G}(\theta) = \boldsymbol{C}\big(e^{j2\pi\theta}\boldsymbol{I} - \boldsymbol{\Phi}\big)^{-1}\boldsymbol{\Gamma} + \boldsymbol{D}$$

where $\theta \in \big(-\frac{1}{2}, \frac{1}{2}\big]$ is the normalized frequency. Note that the ideal spatial averaging $\bar{\boldsymbol{z}}_n = \frac{1}{I}\boldsymbol{1}\boldsymbol{1}^T\boldsymbol{z}_n$ entails $\boldsymbol{G}_{\text{ideal}}(\theta) = \frac{1}{I}\boldsymbol{1}\boldsymbol{1}^T$, i.e., all elements of the transfer function equal $1/I$. However, on graphs that are not fully connected, $\boldsymbol{G}_{\text{ideal}}(\theta)$ cannot be realized which means that some spectral components of the measured data will get delayed and attenuated/amplified. To get a global characterization, we average the amplitude responses over all node pairs according to

$$A(\theta) = \frac{1}{I^2}\sum_{i=1}^{I}\sum_{j=1}^{I}\big|g_{ij}(\theta)\big|\,,$$

with $g_{ij}(\theta) = [\boldsymbol{G}(\theta)]_{ij}$. Ideally, $A(\theta)$ should be as close to $1/I$ as possible. In addition, we are interested in the phase response, or, more specifically in the (average) group delay,

$$\tau(\theta) = -\frac{1}{I^2}\sum_{i=1}^{I}\sum_{j=1}^{I}\frac{d}{d\theta}\arg\{g_{ij}(\theta)\}\,.$$

We note that in case of $\boldsymbol{G}_{\text{ideal}}(\theta)$ we have $\frac{d}{d\theta}\arg\{g_{ij}(\theta)\} = 0$, i.e., all frequencies have identical zero delay. While this is unrealistic for practical graphs, the goal is to have a group delay that varies as little as possible and is as small as possible.

For predictive CP, a transfer function matrix $\tilde{\boldsymbol{G}}(\theta)$, average amplitude response $\tilde{A}(\theta)$, and group delay $\tilde{\tau}(\theta)$ can be obtained by replacing the CP state space matrices with $\tilde{\boldsymbol{\Phi}}$, $\tilde{\boldsymbol{\Gamma}}$, $\tilde{\boldsymbol{C}}$, $\tilde{\boldsymbol{D}}$.

5. NUMERICAL RESULTS

Simulation scenario. In this section we numerically illustrate the concepts introduced in the previous section. To this end, we focus on random geometric graphs [9], which are often used to model wireless sensor networks. More specifically, we placed $I = 100$ nodes in the unit square and set the maximum communication range to $r = \frac{\gamma}{\sqrt{I}}$, with some $\gamma > 0$. Unless stated otherwise, $\gamma = 1.6$. For predictive CP, we designed a simple lowpass predictor for bandwith $\theta_c = 1/(2I) = 5 \cdot 10^{-3}$, using a filter length of $L = 50$ and a prediction horizon chosen to equal the DC group delay $\tau(0)$.

Illustrative example. The left panel in Fig. 1 shows the mean (and the standard deviation) of the amplitude response of plain CP and CP with predictor for one realization of a random geometric graph. Since CP performs averaging, only the lowpass band $\theta = [-2/I, 2/I]$ is shown. Evidently, for both CP versions $T(\theta) \approx 1/I = 0.01$ for $|\theta| \leq \theta_c$ such that appropriate averaging can be expected for signals that vary reasonably slow. The right panel in Fig. 1 shows the corresponding group delay (mean and standard deviation). It is seen that for $|\theta| \leq I/2 = 0.005$, plain CP suffers from frequency-dependent delays in the order of 20-30 iterations (this agrees with numerical results in [15]). By augmenting CP with prediction, these group delays are significantly reduced and flattened out so that CP does not lag behind and delivers dispersion-free averages.

Impact of CP agility and graph connectivity. We next assess the dependence of the CP transfer characteristics on the CP parameter β and the graph connectivity γ. For each (β, γ)-pair we averaged over 20 graph realizations. The left panel of Fig. 2 shows the 90% bandwidth, i.e., the bandwidth within which $A(0) - A(\theta) \leq 0.1$, normalized by $1/\beta$. The mean group delay and the standard deviation of the group, evaluated within the frequency band $\theta \in [-2\cdot10^{-3}, 2\cdot10^{-3}]$, are respectively shown in the middle and right panel of Fig. 2.

It is seen that the bandwidth of CP is almost exactly inversely proportional to β, with the proportionality factor decreasing with increasing γ. which implies that dynamic CP works better if β and γ are small. This is in contrast to the static case, where larger β and better graph connectivity improve averaging accuracy.

The mean and standard deviation of the group delay increase with β and γ (at least up to $\beta \approx 50$ in case of the mean). While the mean can be compensated using predictive CP (see the middle panel), this is not true for the standard variation, which essentially measures the dispersion. Hence, small β is desirable also form this perspective.

6. CONCLUSION

In this paper, we derived a state space model for the dynamic version of consensus propagation (CP) and for its predictive extension. This state space model enabled us to prove that CP is stable for any graph topology. Furthermore, the state space model was used to characterize CP in the frequency domain in terms of a mean amplitude response and a mean group delay. Using these concepts, we provided numerical evidence that contrary to the static case initially considered in [4], the dynamic case calls for low β and low network

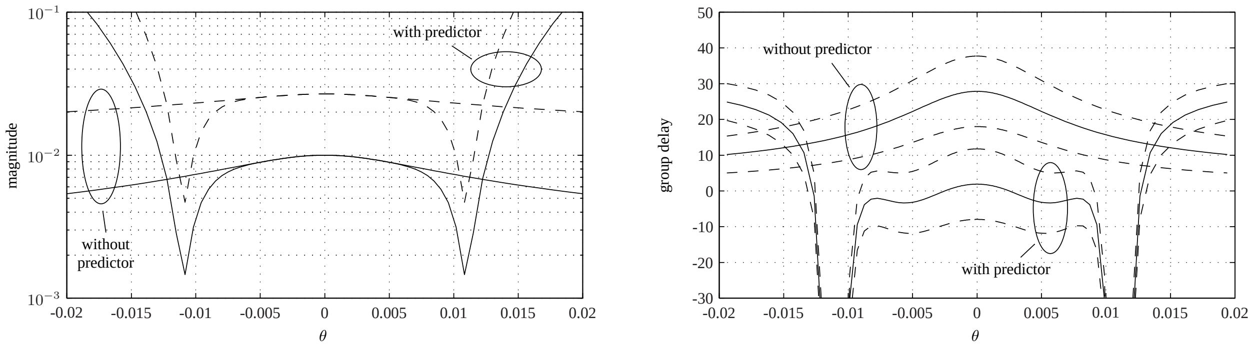

Fig. 1. *The average transfer function $A(\theta)$ (left) and the average group delay $\tau(\theta)$ (right) versus θ. The dashed lines indicate the standard deviation. The prediction bandwidth here was chosen as $\theta_c = 5 \cdot 10^{-3}$.*

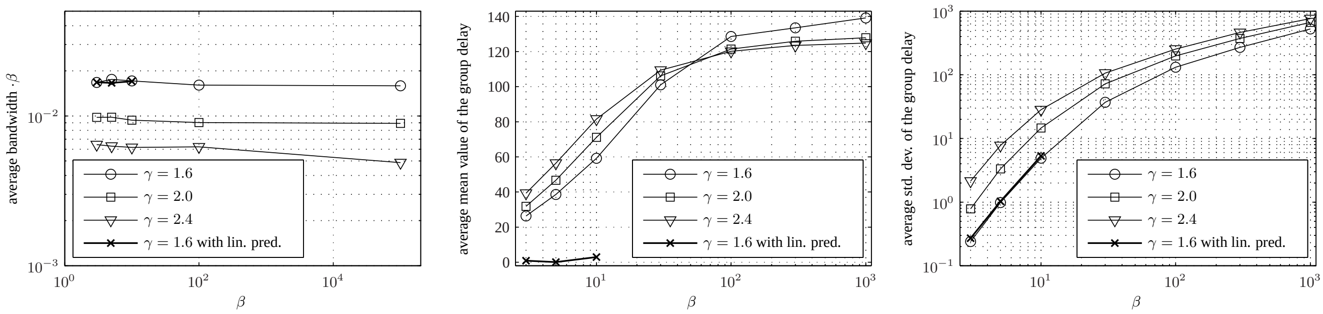

Fig. 2. *Average β-normalized 90%-bandwidth (left), mean group delay (middle), and standard deviation of group delay (right), all versus β for different connectivity values γ.*

connectivity. Our results also confirmed that predictive CP has the potential to compensate for the delays inherent to CP without compromising averaging performance.

REFERENCES

[1] 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.

[2] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*, Englewood Cliffs, NJ: Prentice-Hall, 1989.

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

[5] C. C. Moallemi, *A Message-Passing Paradigm For Optimization*, Ph.D. thesis, Stanford University, Sept. 2007.

[6] Y. Weiss and W. T. Freeman, "Correctness of belief propagation in Gaussian graphical models of arbitrary topology," *Neural Computation*, vol. 13, no. 10, pp. 2173–2200, Oct. 2001.

[7] J. Johnson, D. Malioutov, and A. Willsky, "Walk-sum interpretation and analysis of Gaussian belief propagation," in *Advances in Neural Information Processing Systems 18*, Y. Weiss, B. Schölkopf, and J. Platt, Eds., pp. 579–586. MIT Press, Cambridge, MA, 2006.

[8] V. Schwarz and G. Matz, "Distributed averaging in wireless sensor networks under an aloha-like communication protocol," in *Proc. 44th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2010.

[9] Mathew Penrose, *Random Geometric Graphs*, Oxford University Press, 2003.

[10] Réka Albert, Albert-László Barabási, and Eric Bonabeau, "Scale-free networks," *Scientific American*, vol. 288, pp. 60–69, May 2003.

[11] V. Schwarz, C. Novak, and G. Matz, "Broadcast-based dynamic consensus propagation in wireless sensor networks," in *Proc. 43th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2009.

[12] T. C. Aysal, B. N. Oreshkin, and M. J. Coates, "Accelerated distributed average consensus via localized node state prediction," *IEEE Trans. Signal Processing*, vol. 57, pp. 1563–1576, April 2009.

[13] T. Kailath, *Linear Systems*, Prentice Hall, Englewood Cliffs (NJ), 1980.

[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 3rd edition, 1996.

[15] V. Schwarz and G. Matz, "Distributed reconstruction of time-varying spatial fields based on consensus propagation," in *Proc. IEEE ICASSP-2010*, Dallas, TX, March 2010, pp. 2926–2929.