CONVERGENCE OF A CLASS OF DECENTRALIZED BEAMFORMING ALGORITHMS

James A. Bucklew and William A. Sethares

Department of Electrical and Computer Engineering, University of Wisconsin-Madison

ABSTRACT

One of the key issues in decentralized beamforming is the need to phase-align the carriers of all the sensors in the network. Recent work in this area has shown the viability of certain methods that incorporate single-bit feedback from a beacon. This paper analyzes the behavior of the method (showing conditions for convergence in distribution and also giving a concrete way to calculate the final distribution of the convergent ball) and then generalizes the method in three ways. First, by incorporating both negative and positive feedback it is possible to double the convergence rate of the algorithm without adversely effecting the final variance. Second, a way of reducing the amount of energy required (by reducing the number of transmissions needed for convergence) is shown; its convergence and final variance can also be conveniently described. Finally, a wideband analog is proposed that operates in a decentralized manner to align the time delay (rather than the phase) between sensors.

Index Terms— Adaptive signal processing, beamforming, sensor networks

1. INTRODUCTION

A collection of sensors are scattered in unknown locations. The sensors wish to cooperatively transmit a common message signal as efficiently as possible using a beamforming method in order to be energy efficient. Significant gains occur when exploiting distributed beamforming ([6], [8]) because of the improved signal-to-noise ratio at the receiver: while the received signal magnitude increases with the number of transmitters J, the SNR increases with J^2. Since the total amount of power transmitted increases linearly in J, this represents an J-fold increase in energy efficiency. A key issue in the use of distributed beamforming systems [1] is that the phases of the carriers must be synchronized throughout the network.

A recently proposed scheme [9] accomplishes this phase synchronization using single-bit feedback from a base station. Each sensor broadcasts within each timeslot, perturbing the phase of its carrier slightly from the previous timestep. The base station replies with a signal that indicates whether the received signal is more (or less) coherent than the previous time. The sensors respond in the obvious way: if the signal is improved they keep the new phase while if the signal worsened they revert to their old phase. This scheme is shown, under certain conditions in [10], to asymptotically achieve perfect phase coherence in the noise-free case. However, even tiny disturbances (which may arise physically from thermal noise, from unmodeled dynamics, or from interference with other nearby communications systems) cause convergence to a ball about the correct answer (and not to the correct answer itself). Analysis in the present paper is able to concretely describe both the rate of convergence and the distribution of this convergent ball.

This paper begins by showing in Sect. 2 how the single-bit feedback mechanism for phase alignment can be written as a "small stepsize" μ-dependent algorithm with a discontinuous update term [3]. The discontinuity arises because the sensors either accept or reject the most recent phase change based on the single-bit feedback. Sect. 2 applies the analytical techniques of [4] and [5] to examine the convergence of the algorithm in terms of a related ordinary differential equation (ODE). An extension of these results allows derivation of the asymptotic variance. This concretely describes the final distribution of the algorithm about its equilibrium.

The method of [9], [10] either updates or freezes the estimated phases at each timestep. Sect. 3 observes that it may be possible to do better than to freeze the updates: if adding a small number makes things worse, then most likely subtracting a small number would improve things. This is an old idea [2], [7] in signal processing, and a "signed" algorithm for the decentralized phase alignment problem that uses both positive and negative feedback is described and analyzed: it is shown to converge twice as fast as the original, with the same final distribution.

One key requirement in a sensor network system is energy efficiency. Sect. 4 proposes the ρ-percent method in which only a subset of the sensors transmit at each timestep. Analysis shows that the savings in the number of transmissions (and hence in the energy) can be significant. Since the subset is chosen randomly at each epoch, there is no need to coordinate the sensors, and the method remains decentralized. The methods are extended (n the full paper) to the analysis of analogous algorithms that operate with wideband signals by aligning the received signals in time.

2. ALGORITHM STATEMENT AND CONVERGENCE

Let $\hat{\phi}_{j,n}$ be the phase of the carrier signal at sensor/transmitter j at timestep n and let α_j be the phase difference due to the (unknown) distance between the base station and sensor j. At each timeslot, each sensor randomly perturbs its phase by a small amount $\mu\gamma_{j,n}$. Further suppose that the received signal at the base station at iteration n is corrupted by a Gaussian noise g_n with mean zero and variance σ_t^2.

The algorithm described above can be written

$$\begin{aligned}\hat{\phi}_{j,n+1} &= \hat{\phi}_{j,n} + \mu\gamma_{j,n} \\ &\cdot 1_{\{g_{n+1}+\sum_{r=1}^{J}\cos(\hat{\phi}_{r,n}+\mu\gamma_{r,n}-\alpha_r)>g_n+\sum_{r=1}^{J}\cos(\hat{\phi}_{r,n}-\alpha_r)\}}\end{aligned} \quad (1)$$

for $j = 1, 2, \ldots, J$ where $1_{\{A\}}$ is an indicator function taking on value one if A is true and is zero otherwise. The sum of cosines terms represent the received carrier wave and take on maxima when the $\hat{\phi}$ are phase-aligned. Thus the indicator function is unity if the perturbed phases $\hat{\phi}_{r,n}+\mu\gamma_{r,n}$ are better aligned than the unperturbed phases, and is zero otherwise. The goal of the algorithm is to drive the $\hat{\phi}$ terms to a value at which the sum is maximum, which occurs when all of the terms are maximized, i.e., when $\hat{\phi}_r$ is equal to α_r.

For the purpose of analysis, it is more convenient to rewrite the algorithm in "error system" form by letting $\phi_{j,n} = \hat{\phi}_{j,n} - \alpha_j$. We also suppose that the i.i.d. perturbation random variables $\{\gamma_{j,n}\}$ are chosen to have a symmetric distribution (about zero) with finite variance σ_γ^2. Then for small μ (keeping just the first terms in the Taylor series), $\cos(\phi_{r,n} + \mu\gamma_{r,n})$ can be approximated by $\cos(\phi_{r,n}) - \mu\gamma_{r,n}\sin(\phi_{r,n})$. The algorithm is then

$$\phi_{j,n+1} = \phi_{j,n} + \mu\gamma_{j,n}1_{\{\tilde{g}+\sum_{r=1}^{J}\gamma_{r,n}\sin(\phi_{r,n})<0\}} \quad (2)$$

where $\tilde{g}$ is normal with mean zero and variance $2\sigma_t^2/\mu^2$.

In order to investigate the behavior of the algorithm, observe that convergence of ϕ_r to zero is equivalent to convergence of the phase estimates $\hat{\phi}_r$ to their unknown values α_r. The analysis requires developing some new technical machinery. The basis of the analytical approach is to find an ordinary differential equation (ODE) that accurately mimics the behavior of the algorithm for small values of μ. Studying the ODE then gives information regarding the behavior of the algorithm. For example, if the ODE is stable, the algorithm is convergent (at least in distribution). If the ODE is unstable, the algorithm is divergent. The approach grows out of results in [3] and [4], which are themselves based on the techniques of [5]. The approach is conceptually similar to stochastic approximation but its assumptions (and hence conclusions) are somewhat different. First, the stepsize μ in (2) is fixed, unlike in stochastic approximations where the stepsize is required to converge to zero. Thus the algorithms do not necessarily converge to a fixed vector; rather, they converge in distribution. Moreover, the analysis is capable of delivering concrete values for the convergent distribution; as far as we know, this is not possible with other methods. Second, no continuity assumptions need to be made on the update terms; this is crucial because of the discontinuity caused by the indicator function in (2), and is also more general than other methods that require differentiability of the update term.

Using the Central Limit Theorem and the techniques of [4] it is easy to show that the limiting differential equation is

$$\frac{d\phi_j(t)}{dt} = -\frac{\sigma_\gamma^2\sin(\phi_j(t))}{\sqrt{2\pi}\sqrt{\sigma_\gamma^2||\sin(\phi_{/(j)}(t))||^2 + 2\sigma_t^2/\mu^2}}. \quad (3)$$

A straightforward linearization argument shows that this ODE is stable about zero. Simulations in Sect. 5 show that the ODE accurately tracks the trajectories of the algorithm.

Once the algorithm has converged, it is important to be able to characterize the final distribution. Using similar techniques, we can show that the asymptotic variance is given by

$$VAR = \mu\frac{\sqrt{\pi\tilde{\sigma}_t^2}}{2}. \quad (4)$$

Somewhat surprisingly, the asymptotic variance is independent of the size of the phase perturbations σ_γ^2. Sect. 5 shows that this calculated variance matches closely to the empirical variance derived from simulations.

3. THE SIGNED ALGORITHM

In the distributed phase alignment algorithm (1), the phase of each sensor is updated by the perturbed value (if the feedback from the beacon says that the overall alignment improved) or else it remains fixed. Accordingly, in many iterations, no changes are made. Since each of the individual phase updates are scalar, it seems reasonable that when the feedback indicates no improvement, an update in the opposite direction might be useful. Effectively, this replaces the indicator function in the update with a signum function. Following the logic of (1)-(2) leads to the error system which is valid for small μ

$$\phi_{j,n+1} = \phi_{j,n} - \mu\gamma_{j,n}\text{sgn}(\tilde{g} + \sum_{r=1}^{J}\gamma_{r,n}\sin(\phi_{r,n})). \quad (5)$$

Carrying out a similar analysis as in Sect. 2, the ODE is almost exactly as before but with a factor of two in the right hand side of (3). Thus the corresponding ODE for the signed algorithm converges twice as fast as when using the indicator function.

The final variance can also be calculated as before and is found to be indentical to (4). Hence, we find that the signed algorithm converges twice as fast as (1) yet has the same residual error variance.

4. THE "ρ % SOLUTION" ALGORITHM

One of the key requirements in a sensor network system is energy efficiency. Suppose that at each epoch, sensors independently transmit to the beacon with probability $p = \rho/100$ using its current phase value. This transmission is then immediately followed by another transmission of a "perturbed" phase angle. The beacon then feeds back a single bit which specifies which of the two transmissions had greater power. Each transmitting sensor then updates its current phase value based on the feedback. Thus, in each transmission epoch, only pJ sensors transmit on average; but they must transmit twice.

This strategy can also be written as a small stepsize μ-dependent algorithm. Let $\{B_{1,n}, B_{2,n}, \ldots, B_{J,n}\}$ be independent zero-one Bernoulli random variables with $P(B_{j,n} = 1) = p$. The event $\{B_{j,n} = 1\}$ indicates that at time n, sensor j will transmit. It is possible to mimic the analysis given in (1)-(3) to obtain the limiting ODE

$$\frac{d\phi_j(t)}{dt} = -pE[\frac{\sigma_\gamma^2\sin(\phi_j(t))}{\sqrt{2\pi(\sigma_\gamma^2\sum_{r=1(r\neq j)}^{J}B_r[\sin(\phi_r(t))]^2 + 2\sigma_t^2/\mu^2)}}]. \quad (6)$$

The presence of the Bernoulli random variables in the denominator makes a simple closed form solution impossible (though an infinite power series could be developed). Sect. 5 shows that the total number of transmissions (and hence the total energy consumed in the phase alignment process) can be decreased when following this strategy.

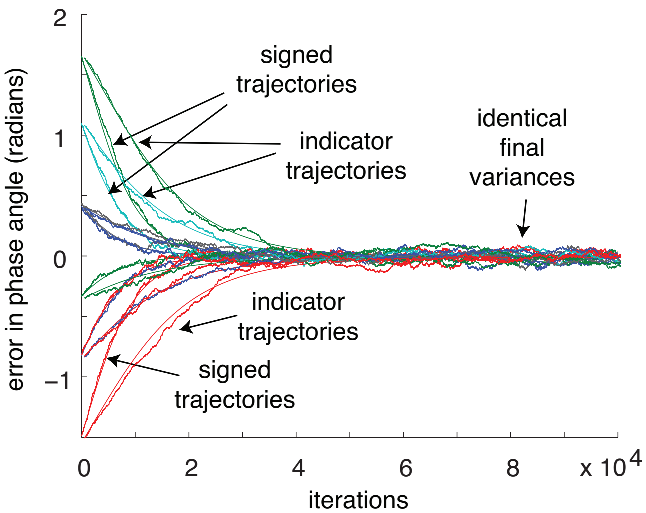
In addition, it is possible to combine the idea of the signed update from Sect. 3 with the ρ-percent algorithm. Following the same procedure shows that this algorithm has the same ODE, but multiplied by a factor of two, indicating a doubling of the convergence rate.

The variance analysis of the ρ-percent algorithm is straightforward and we obtain an identical expression. Importantly, this expression is independent of p (and hence ρ).

5. SIMULATIONS

This section illustrates the relationship between the trajectories of the algorithm(s) and the behavior of the ODE(s) and shows that the calculated variances accurately reflect the behavior of the algorithm. Recall that convergence of the error system to a region about zero is equivalent to convergence of the actual trajectories of the phase estimates to a region about their (unknown) values. The final error variance can be calculated directly from the simulation. Fig. 5

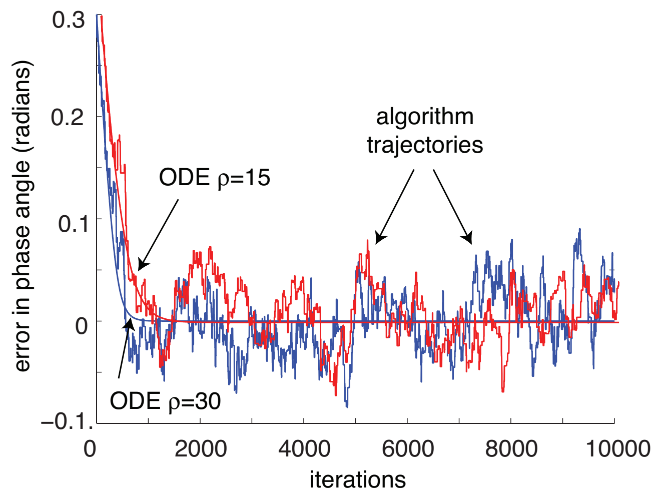
shows trajectories of the error systems for the indicator algorithm and its ODE and for the signed version and its ODE. $J = 10$ sensors were used though only six are shown in the figure to reduce clutter. The two algorithms were initialized at the same values and allowed to iterate. Observe that in all cases, the signed algorithm converges faster, at about twice the rate of the indicator version, as suggested by the corresponding ODEs. Parameters for the simulation are $\mu = .0005$, $\tilde{\sigma}_t^2 = 10$, and $\sigma_\gamma^2 = 2$. The final variance, calculated to be .0014, agrees with the empirical value (measured from the simulations) to four decimal places.



Comparison of the indicator and the signed algorithms.

Fig. 5 shows the trajectories of the ρ-percent error system 4 and the corresponding ODE (6). Again, the algorithm follows the ODE as it converges exponentially towards its stable point. In these simulations the stepsize $\mu = 0.01$ and the standard deviation of the thermal noise was 0.001. The predicted final variance for the $\rho = 30$ case was $8.862E^{-4}$ while the actual variance, computed over all sensors, was $8.587E^{-4}$. The predicted final variance for the $\rho = 15$ case was $8.862E^{-4}$ (the same as for $\rho = 30$) while the actual variance was $9.32E^{-4}$.

It is also necessary to verify that the convergence of the ρ-percent algorithm is rapid enough that the total number of transmissions needed is less than for the corresponding algorithm where all sensors transmit at every time step (which is essentially the $\rho = 100$ case). With $J = 10$ sensors, a thermal noise with standard deviation 0.001, and a stepsize of $\mu = 0.01$. We conduct an experiment to show how many iterations are needed for convergence as a function of the ρ value. The experiment is conducted by setting the phase error for sensor $\#1$ at 1.0 radian, and checking how many iterations are needed before the sensor converges 95% of the way to zero. As might be expected, we find that the number of iterations decreases as ρ increases, but so do the number of transmissions. In this experiment about 720 transmissions are needed for the $\rho = 100$ case. Since the algorithm requires two transmissions in each epoch, any ρ that requires fewer than half this number (i.e., 360) transmissions will be more efficient. In this case, the crossover point occurs at about $\rho = 30$. The purpose here is not to try and elucidate the best parameters to use, only to demonstrate that significant gains in energy usage, as reflected in the number of transmissions required, are possible when using the ρ-percent algorithm.



Typical trajectories of $\rho\%$ algorithm (with $\rho = 15$ and $\rho = 30$) and their corresponding ODEs.

6. REFERENCES

[1] G. Barriac, R. Mudumbai, U. Madhow, "Distributed beamforming for information transfer in sensor networks," IPSN'04, April 26–27, 2004, Berkeley, California, USA.

[2] D. L. Duttweiler, "Adaptive filter performance with nonlinearities," *IEEE Trans. on Acoustics, Speech, and Signal Processing,* Vol. 30, pp. 578–586, Aug 1982.

[3] J. A. Bucklew, T. Kurtz, and W. A. Sethares, "Weak convergence and local stability of fixed step size recursive algorithms,"*IEEE Trans. on Info. Theory*, Vol. 39, No. 3, pp. 966–978, May, 1993.

[4] J.A. Bucklew, W.A. Sethares, "The covering problem and μ-dependent adaptive algorithms," *IEEE Trans. Sig. Proc.*, Vol. 42, No. 10, pp. 2616–2627, October, 1994.

[5] S. Ethier and T. Kurtz, *Markov Processes - Characterization and Convergence,* Wiley-Interscience, New York, 1986.

[6] M. Gastpar and M. Vetterli, "On the capacity of large gaussian relay networks," *IEEE Trans. on Inform. Theory*, vol. 51, no. 3, pp. 765–779, 2005.

[7] A. Gersho, "Adaptive filtering with binary reinforcement," *IEEE Trans. on Info. Theory*, Vol. IT-30, No. 2, pp. 191–198, March 1984.

[8] B. Hassibi and A. Dana, "On the power efficiency of sensory and ad-hoc wireless networks," *IEEE Trans. on Inform. Theory*, July 2006.

[9] R. Mudumbai, B. Wild, U. Madhow and K. Ramchandran, "Distributed beamforming using 1 bit feedback: from concept to realization," *Proc. of 44'th Allerton Conference on Communication Control and Computing,* Sept. 2006.

[10] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Distributed Transmit Beamforming using Feedback Control," submitted to *IEEE Trans. on Info. Theory* (see http://www.ece.ucsb.edu/∼raghu/research/pubs.html)

[11] H. Ochiai, P. Mitran, H. V. Poor, and V. Tarokh, "Collaborative beamforming for distributed wireless ad hoc sensor networks, *IEEE Trans. Signal Proc.*, Vol. 53, No. 11, Nov. 2005.