# A NEW MINIMUM-CONSENSUS DISTRIBUTED PARTICLE FILTER FOR BLIND EQUALIZATION IN RECEIVER NETWORKS

*Claudio J. Bordin Jr.*

Universidade Federal do ABC
claudio.bordin@ufabc.edu.br

*Marcelo G. S. Bruno*

Instituto Tecnológico de Aeronáutica
bruno@ita.br

**ABSTRACT**

We describe in this paper a novel distributed particle filtering algorithm that performs blind equalization of frequency-selective channels in a setup with a single transmitter and multiple receivers. The algorithm employs parallel minimum consensus iterations to determine some a posteriori probability functions, providing equal approximations on all network nodes in a finite, deterministic, network-dependent number of steps. We verify via computer simulations that the new algorithm exhibits a bit error rate (BER) performance similar to that of the centralized particle-filter estimator with communication requirements milder than that of previous approaches, as the new method drops the need to evaluate quantities via average consensus.

***Index Terms*—** Distributed Algorithms, Particle Filters, Blind Equalization.

## 1. INTRODUCTION

Distributed estimation is currently a relevant research area due to the plethora of new devices equipped with sensing, computing, and communication capabilities. We consider in this paper a scenario in which a single transmitter broadcasts through frequency-selective channels a sequence of discrete-valued symbols which are received by a network of remote sensors. Instead of forwarding local observations to a remote fusion center, we develop a method in which the different nodes not only process their own observations independently, but also cooperate with each other to compute the joint optimal estimate of the transmitted symbols given all observations in the network.

Previous distributed filtering algorithms mostly perform linear estimation [1], [2], which is not ideally suited for blind equalization of digital channels as the minimum BER estimate of the transmitted data may be significantly different from the linear-mean-square error estimates provided by linear adaptive or Kalman filters. This restriction can be sidestepped by nonlinear techniques such as particle filters, which approximate the maximum a posteriori (MAP) estimate.

The development of distributed particle filters has been restrained, however, by the fact that, unless further approximations are made, all network nodes must agree on the same set of particles and weights. To abide by this restriction, most current methods [3],[4], broadcast messages across the network, an undesirable feature in many scenarios with communication or power constraints.

In this paper we introduce a new consensus-based, distributed particle filtering algorithm. The algorithm employs minimum consensus [5] to evaluate approximations to some node-dependent probabilities across the receiver network, dropping the need for broadcasts or to perform a random number of average consensus iterations [6].

The remaining text is organized as follows: in Sec. 2 we describe the problem setup, introducing in Sec. 3 a centralized particle filter approach to its solution. In Sec. 4, we introduce consensus-based techniques for distributed particle filtering and present the new minimum-consensus-based method, whose properties are discussed in Sec. 5 and performance assessed in Sec. 6. Our conclusions are summarized in Sec. 7.

## 2. PROBLEM SETUP

Denote by $\{b_n\}$ an independent, identically distributed (i.i.d.) binary bit sequence and by $\{x_n\}$, $x_n \in \{\pm 1\}$, the corresponding differentially encoded symbols. We assume that the observations $y_{r,0:n} \triangleq \{y_{r,0}, \ldots, y_{r,n}\}$ at the $r-$th node of a network of $R$ receivers are obtained as the output of the additive noise frequency-selective finite impulse response (FIR) channel

$$y_{r,n} = \mathbf{h}_r^T \mathbf{x}_n + v_{r,n} \ , \tag{1}$$

where $\mathbf{h}_r \in \mathbb{R}^{L\times 1}$ is a vector with the (time-invariant) channel impulse response terms, $\mathbf{x}_n \triangleq [x_n \ldots x_{n-L+1}]^T$, and $v_{r,n}$ represents an i.i.d. zero-mean Gaussian random process of known variance $\sigma_r^2$.

The *unknown, random* parameters $\mathbf{h}_r$, $1 \le r \le R$, are assumed to be independent for $r \neq s$, and distributed a priori as $\mathbf{h}_r \sim \mathcal{N}_L(\mathbf{h}_r|0; I/\varepsilon^2)$, where $\mathcal{N}_L$ denotes an $L-$variate Gaussian p.d.f., and $\varepsilon$ is the model's hyper-parameter.

Under these hypotheses, we aim at developing a recursive method for obtaining MAP estimates $\hat{b}_n = \arg\max_{b_n} p(b_n| y_{1:R,0:n})$, where $y_{1:R,0:n} \triangleq \{y_{1,0:n} \ldots y_{R,0:n}\}$.

## 3. CENTRALIZED SOLUTION VIA PARTICLE FILTERS

Particle filters allow one to approximate the posterior probability mass function (p.m.f.) of the transmitted bits as

$$p(b_n|y_{1:R,0:n}) \approx \sum_{q=1}^{Q} w_n^{(q)} \mathcal{I}\left\{b_n = b_n^{(q)}\right\}, \quad (2)$$

where $\mathcal{I}\{\cdot\}$ denotes the indicator function, $Q$ the number of *particles* $b_n^{(q)}$ sampled from the importance function $\pi(\mathbf{x}_n^{(q)}|\mathbf{x}_{0:n-1}^{(q)}, y_{1:R,0:n})$, and $w_n^{(q)}$ are the importance weights. For simplicity, we employ the *prior* importance function, so that $\pi(\mathbf{x}_n^{(q)}|\mathbf{x}_{0:n-1}^{(q)}, y_{1:R,0:n}) \triangleq p(\mathbf{x}_n^{(q)}|\mathbf{x}_{n-1}^{(q)})$. The resulting *importance weights'* update expression is then given by

$$w_n^{(q)} \propto w_{n-1}^{(q)}\, p(y_{1:R,0:n}|\mathbf{x}_{0:n}^{(q)}). \quad (3)$$

From the a priori independence of the unknown parameters for each receiver's channel, it can be shown [7] that the quantity on the right-hand side (r.h.s.) of (3) can be factored as

$$p(y_{1:R,0:n}|\mathbf{x}_{0:n}^{(q)}) \propto \prod_{r=1}^{R} p(y_{r,0:n}|\mathbf{x}_{0:n}^{(q)}). \quad (4)$$

Finally, under the assumptions of Sec. 2, one can verify after algebraic manipulations [8] that

$$\begin{aligned} p(y_{r,0:n}|\mathbf{x}_{0:n}^{(q)}) &= \int_{\mathbb{R}^L} p(y_{r,0:n}, \mathbf{h}_r|\mathbf{x}_{0:n}^{(q)})\, d\mathbf{h}_r \\ &= \mathcal{N}\left(y_{r,n}|\ \hat{\mathbf{h}}_{r,n-1}^{(q)T}\mathbf{x}_n^{(q)}\ ; \gamma_{r,n}^{(q)}\right) \quad (5) \end{aligned}$$

where $\hat{\mathbf{h}}_{r,n-1}^{(q)T}$ and $\gamma_{r,n}^{(q)}$ can be computed via the Kalman filter recursions [8]

$$\gamma_{r,n}^{(q)} \triangleq \sigma_r^2 + \mathbf{x}_n^{(q)T}\mathbf{\Sigma}_{r,n-1}^{(q)}\mathbf{x}_n^{(q)}\ , \quad (6)$$

$$e_{r,n}^{(q)} \triangleq y_{r,n} - \hat{\mathbf{h}}_{r,n-1}^{(q)T}\mathbf{x}_n^{(q)}\ , \quad (7)$$

$$\hat{\mathbf{h}}_{r,n}^{(q)} = \hat{\mathbf{h}}_{r,n-1}^{(q)} + \mathbf{\Sigma}_{r,n-1}^{(q)}\mathbf{x}_n^{(q)}e_{r,n}^{(q)}/\gamma_{r,n}^{(q)}\ , \quad (8)$$

$$\mathbf{\Sigma}_{r,n}^{(q)} = \mathbf{\Sigma}_{r,n-1}^{(q)} - \mathbf{\Sigma}_{r,n-1}^{(q)}\mathbf{x}_n^{(q)}\mathbf{x}_n^{(q)T}\mathbf{\Sigma}_{r,n-1}^{(q)}/\gamma_{r,n}^{(q)}\ , \quad (9)$$

with $\hat{\mathbf{h}}_{r,-1}^{(q)} = \underline{\mathbf{0}}$ and $\mathbf{\Sigma}_{r,-1}^{(q)} = \mathbf{I}\varepsilon^{-2}$.

## 4. CONSENSUS-BASED ALGORITHMS

The weight update rule can be derived by replacing (4) into (3), so that

$$w_n^{(q)} \propto w_{n-1}^{(q)} \prod_{r=1}^{R} p(y_{r,0:n}|\mathbf{x}_{0:n}^{(q)}), \quad (10)$$

which can be rewritten as

$$w_n^{(q)} \propto w_{n-1}^{(q)} \exp(\rho_n^{(q)}), \quad (11)$$

where $\rho_n^{(q)} \triangleq \sum_{r=1}^{R} \rho_{r,n}^{(q)}$ and $\rho_{r,n}^{(q)} \triangleq \ln p(y_{r,0:n}|\mathbf{x}_{0:n}^{(q)})$. The latter sum can be evaluated via $Q$ parallel *consensus averaging* iterations [9] as

$$\rho_n^{(q)} = \lim_{k\to\infty} \tilde{\rho}_{r,n}^{(k,q)}, \ \ \forall r \quad (12)$$

where $k$ is the consensus algorithm iteration index (independent of $n$), $\tilde{\rho}_{r,n}^{(0,q)} \triangleq R\rho_{r,n}^{(q)}$, and

$$\tilde{\rho}_{r,n}^{(k,q)} = \sum_{s\in\mathbf{N}(r)\cup r} a_{rs}\ \tilde{\rho}_{s,n}^{(k-1,q)}, \quad (13)$$

where $\mathbf{N}(r)$ denotes the *neighborhood* of the node $r$, the set of all nodes adjacent to $r$, and $a_{rs}$ are positive weights such that $\{\mathbf{A}\}_{rs} \triangleq a_{rs}$ is an $R \times R$ *primitive* [10] doubly stochastic matrix. Under these conditions, the iteration in (13) converges to the desired sum.

Consensus averaging allows (10) to be evaluated without the need of broadcasts or communications beyond the immediate neighborhood of any node. However, in general, for any finite $k$, $\tilde{\rho}_{r,n}^{(k,q)} \neq \tilde{\rho}_{s,n}^{(k,q)}$, $r \neq s$. As a result, direct application of consensus averaging to evaluate (11) will result in diverse importance weights for the same particle $q$ at different nodes, violating assumptions implicit in (2)-(3).

### 4.1. Minimum-Consensus-based Approach

To guarantee coherence in the particle sets across the network nodes, some form of approximation is required. Basically, one must create a mechanism for determining approximate importance weights that are equal on all nodes for a given particle $q$. Combined with synchronous sampling/resampling [3], this is a sufficient condition to assure coherence of the particle set.

In [6] and [7], equality of the importance weights was determined, directly or not, via *minimum/maximum consensus* [5] algorithms. Given node-dependent quantities $\mu_r$, these algorithms compute their minimum/maximum as $\left.\tilde{\mu}_r^{(k)}\right|_{k=D}$, where

$$\tilde{\mu}_r^{(k)} = \min_{s\in\mathbf{N}(r)\cup r}/\max\left\{\tilde{\mu}_s^{(k-1)}\right\}, \quad (14)$$

$D \leq R$ is dependent on the network's topology [5] and $\tilde{\mu}_r^{(0)} \triangleq \mu_r$.

In [7], the approximate weights $\tilde{w}_n^{(q)}$ are determined as

$$\tilde{w}_n^{(q)} = \tilde{w}_{r,n}^{(q)} \propto \mathcal{Q}\left(\tilde{w}_{n-1}^{(q)} \exp(\tilde{\rho}_n^{(k,q)})\right), \quad (15)$$

for $k = 0, 1, \ldots$ until $\tilde{w}_{r,n}^{(q)} = \tilde{w}_{s,n}^{(q)}$, $\forall r \neq s$, where $\mathcal{Q}(.)$ denotes a fixed quantizer, and $\tilde{\rho}_n^{(k,q)}$ is obtained via (13). The required equality of $\tilde{w}_{r,n}^{(q)}$ is verified by a method based on minimum consensus. In [6], alternatively, approximate weights are obtained directly as $\tilde{w}_n^{(q)} = \max_s \left\{ \tilde{w}_{n-1}^{(q)} \exp(\tilde{\rho}_{s,n}^{(k,q)}) \right\}$ via maximum consensus, for a fixed $k$.

We introduce here a new approach that, contrary to [6] and [7], does not aim at approximating the product on the r.h.s. of (4). Instead, it relies on the alternate approximation

$$\tilde{p}(y_{1:R,0:n}|\mathbf{x}_{0:n}^{(q)}) \propto \min_r \left( p(y_{r,0:n}|\mathbf{x}_{0:n}^{(q)}) \right), \tag{16}$$

which can be exactly evaluated via minimum consensus only. The approximate importance weights are then determined as

$$\tilde{w}_n^{(q)} \propto \tilde{w}_{n-1}^{(q)} \, \tilde{p}(y_{1:R,0:n}|\mathbf{x}_{0:n}^{(q)}). \tag{17}$$

## 5. ASYMPTOTIC PROPERTIES OF THE PROPOSED APPROXIMATION

Applying the matrix inversion lemma [11] to (9), it follows that

$$(\Sigma_{r,n}^{(q)})^{-1} = (\Sigma_{r,n-1}^{(q)})^{-1} + \frac{1}{\sigma^2} \mathbf{x}_n^{(q)} \, (\mathbf{x}_n^{(q)})^T, \tag{18}$$

where $\mathbf{x}_n^{(q)}$ is a vector with binary entries equal to +1 or -1. From (18), we note that $\Sigma_{r,n}^{(q)} \to \mathbf{0}$ as $n \to \infty$. From Eq.(6) then, for $n$ sufficiently large, we can make the approximation $\gamma_{r,n}^{(q)} \approx \sigma_r^2$. Assuming additionally that $\sigma_r^2 = \sigma^2$, $\forall r$, we obtain that, asymptotically,

$$p(y_{1:R,0:n}|\mathbf{x}_{0:n}^{(q)}) = \mathcal{N}_R \left( \mathbf{y}_n | \hat{\mathbf{y}}_n^{(q)} \, ; \mathbf{I}\sigma^2 \right), \tag{19}$$

where $\mathbf{y}_n \triangleq [y_{1,n} \cdots y_{R,n}]^T$ and $\hat{\mathbf{y}}_n^{(q)} \triangleq \left[ \hat{\mathbf{h}}_{1,n-1}^{(q)T} \mathbf{x}_n^{(q)} \cdots \hat{\mathbf{h}}_{R,n-1}^{(q)T} \mathbf{x}_n^{(q)} \right]^T$.

It follows from (19) that (16) is equivalent to approximating $\|\mathbf{y_n} - \hat{\mathbf{y}}_\mathbf{n}^{(\mathbf{q})}\|_\mathbf{2}^\mathbf{2}$ by its *maximum* squared component, i.e., by $\|\mathbf{y_n} - \hat{\mathbf{y}}_\mathbf{n}^{(\mathbf{q})}\|_\infty^\mathbf{2}$. It can be verified [11] that $\|\mathbf{y_n} - \hat{\mathbf{y}}_\mathbf{n}^{(\mathbf{q})}\|_\infty^\mathbf{2} \leq \|\mathbf{y_n} - \hat{\mathbf{y}}_\mathbf{n}^{(\mathbf{q})}\|_\mathbf{2}^\mathbf{2}$, so that the resulting *unnormalized* weights are overestimated.

The Kullback-Leibler [12] divergence between $p(y_{1:R,0:n}|\mathbf{x}_{0:n}^{(q)})$ and $\tilde{p}(y_{1:R,0:n}|\mathbf{x}_{0:n}^{(q)})$ is given by

$$\begin{aligned} D_{KL}(p||\tilde{p}) &= E_{p(\cdot)} \left[ \ln \left( p(\cdot) \frac{\kappa(\sigma^2, R)}{\tilde{p}(\cdot)} \right) \right] \\ &= \frac{1}{2\sigma^2} E_{p(\cdot)} \left[ \|\mathbf{y_n} - \hat{\mathbf{y}}_\mathbf{n}^{(\mathbf{q})}\|_\infty^\mathbf{2} - \|\mathbf{y_n} - \hat{\mathbf{y}}_\mathbf{n}^{(\mathbf{q})}\|_\mathbf{2}^\mathbf{2} \right] \\ &\quad - (R-1) \ln(\sigma\sqrt{2\pi}) + \ln \kappa(\sigma^2, R), \end{aligned} \tag{20}$$

where

$$\kappa(\sigma^2, R) \triangleq \int_{\mathbb{R}^R} \min_r \mathcal{N}_1 \left( \mathbf{y}_{r,n} | \hat{\mathbf{y}}_{r,n}^{(q)} \, ; \sigma^2 \right) \, d\mathbf{y}_{r,n}. \tag{21}$$

is the normalization term for $\tilde{p}(.)$.

In Fig. 1, we display the behavior of $D_{KL}(p||\tilde{p})$ as a function of $\sigma^2$ for some values of $R$. To that aim, the expectation in (20) was evaluated via Monte Carlo simulations in which we generated 10.000 independent samples from $p(.)$ according to (19). To determine $\kappa(\sigma^2, R)$, in turn, we obtained numerical estimates of (21) via Monte Carlo integration; the results obtained for $1 \leq R \leq 8$ fitted the expression

$$\kappa(\sigma^2, R) = \begin{cases} \sigma^{R-1} 2^{(3R-1)/2} \pi^{-1/2} (R/2)!, & R \text{ even} \\ \sigma^{R-1} 2^{(R-3)/2} \dfrac{(R+1)!}{(\frac{R+1}{2})!}, & R \text{ odd} \end{cases} \tag{22}$$

As one may note, $D_{KL}(p||\tilde{p})$ grows with $R$, so that the approximation in (16) progressively degrades for larger networks.

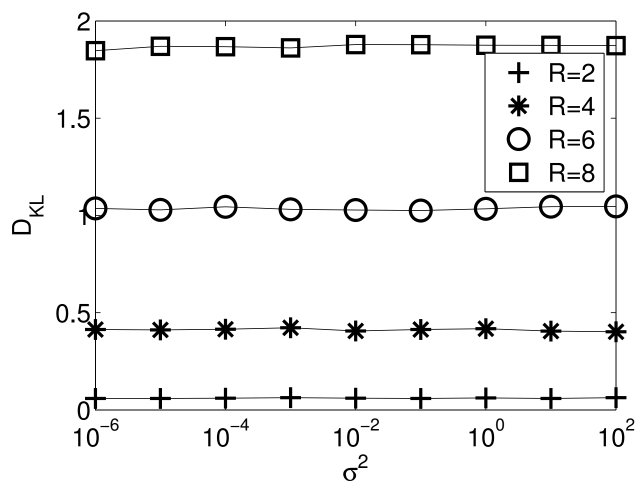


**Fig. 1**. $D_{KL}(p||\tilde{p})$ as a function of $\sigma^2$ and $R$.

## 6. SIMULATION RESULTS

The steady state performance of the proposed algorithm was assessed via simulations consisting of 300 independent Monte Carlo runs. In each realization, we computed the mean bit error rate (BER) as a function of $E_B/N_0$, transmitting a random sequence of 300 i.i.d bits, with the first 150 bits discarded to allow for convergence.

The simulated system has $R = 4$ receiving nodes and the filters employed $Q = 300$ particles. All algorithms perform synchronized systematic resampling [3] whenever the *effective sample size* estimated as $\left( \sum_{q=1}^{Q} \tilde{w}_n^{(q)} \right)^{-1}$ falls below $Q/2$.

The transmission channels $\mathbf{h}_r$ have $L = 3$ coefficients, and were obtained by sampling independently in each realization and for each receiver from a Gaussian p.d.f. $\mathcal{N}(\mathbf{0}; \mathbf{I})$ and normalized so that $\|\mathbf{h}_r\|^2 = 1$. The noise variances were determined as $\sigma^2 = \|\mathbf{h}_r\|^2 N_0/E_B$. The model hyper-parameter was set to $\varepsilon = 1$. The average consensus weight matrix employed was

$$\mathbf{A} = \frac{1}{5}\begin{bmatrix} 3 & 2 & 0 & 0 \\ 2 & 1 & 2 & 0 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 2 & 3 \end{bmatrix}.$$

In Fig. 2, we show the performance of the proposed algorithm ($\nabla$). For comparison, we ran with the same setup the optimal joint particle-filter-based algorithm that exactly computes (4) (dash-dotted line). Fig. 2 also displays the performance of isolated receivers ($+$) that do not cooperate and that of a version of the algorithm proposed in [6] ($\circ$) that employs 20 average-consensus steps to estimated the required density product and does not discard any particle. In addition, we display the results for an alternative algorithm ($\triangle$) in which the minimization in (16) is replaced with maximization. As one may verify, the proposed algorithm performed

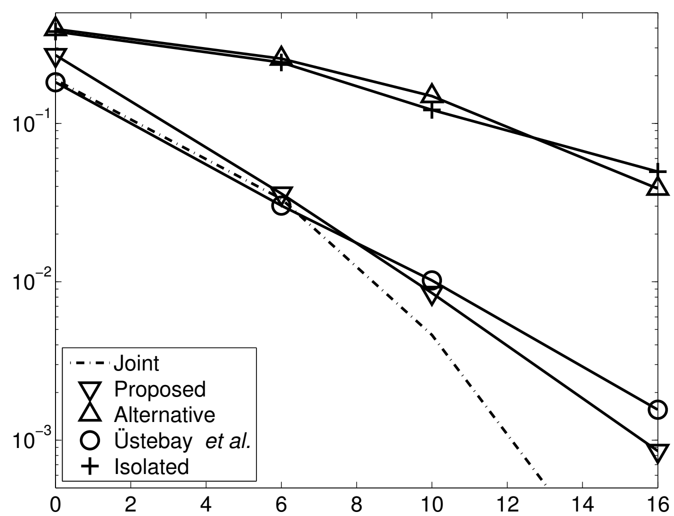


**Fig. 2**. Mean bit error rate (BER) estimated in 300 independent runs.

worse than the centralized particle filter estimator, but outperformed the isolated receiver by a great margin. One may also note that the proposed algorithm performed similarly to the algorithm of [6] at a much lower communication complexity, since the proposed algorithm completely eliminates the need for average consensus iterations. Finally, one may notice that although the alternative approximation ($\triangle$) might seem plausible, it led to a performance similar to that of the isolated receiver.

## 7. CONCLUSIONS

We introduced in this paper a new distributed particle filtering algorithm based on minimum consensus. The algorithm determines approximations to some required posterior probability functions that converge to the same value on all network nodes within a finite number of consensus iterations. Compared to previous approaches, the proposed method attains similar performances with reduced, deterministic communication and computational requirements. The method introduced in this paper can be applied to any filtering problem with conditionally independent linear Gaussian observations and discrete-valued variables. In simulated Monte Carlo experiments of a distributed blind equalization problem with multiple remote receivers, the new algorithm exhibited a BER performance similar to that of the optimal joint particle-filter estimator.

## 8. REFERENCES


[1] A. Ribeiro, G. B. Giannakis, and S. I. Roumeliotis, "SOI-KF: Distributed Kalman Filtering With Low-Cost Communications Using the Sign of Innovations," *IEEE Trans. on Sig. Proc.*, vol. 54, no. 12, pp. 4782–4795, 2006.

[2] C. G. Lopes and A. H. Sayed, "Incremental Adaptive Strategies Over Distributed Networks," *IEEE Trans. on Sig. Proc.*, vol. 55, no. 8, pp. 4064–4077, Aug. 2007.

[3] M. J. Coates, "Distributed particle filtering for sensor networks," in *3rd Intl. Symp. on Inf. Proc. in Sensor Networks*, pp. 99–107, Berkeley - CA, April 2004.

[4] H. Sangjin, M. Hong and P. M. Djurić, "An efficient fixed-point implementation of residual resampling scheme for high-speed particle filters," *IEEE Sig. Proc. Lett.*, vol. 11, no. 5, pp. 482–485, May 2004.

[5] V. Yadav and M. V. Salapaka, "Distributed protocol for determining when averaging consensus is reached," in *45th Annual Allerton Conf.*, pp. 715–720, Allerton House - UIUC, Sep. 2007.

[6] D. Üstebay, M. Coates, and M. Rabbat, "Distributed auxiliary particle filters using selective gossip," in *2011 Intl. Conf. on Acoustics, Speech, and Sig. Proc.*, pp. 3296–3299, Prague - Czech Rep., May 2011.

[7] C. J. Bordin Jr. and M. G. S. Bruno, "Consensus-based Distributed Particle Filtering Algorithms for Cooperative Blind Equalization in Receiver Networks," in *2011 Intl. Conf. on Acoustics, Speech, and Sig. Proc.*, pp. 3968–3971, Prague - Czech Rep., May 2011.

[8] C. J. Bordin Jr. and L. A. Baccalá, "Deterministic Particle Filters for Joint Blind Equalization and Decoding on Frequency Selective Channels," in *2005 Intl. Conf. on Acoustics, Speech, and Sig. Proc.*, pp. 497–500, Philadelphia - PA, Mar. 2005.

[9] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Syst. Control Lett.*, vol. 53, no. 1, pp. 65-78, Sep. 2004.

[10] J. R. Norris, *Markov Chains*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 1997.

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1990.

[12] S. Kullback, *Information Theory and Statistics*, Wiley, 1959.