Sequential Bayesian Algorithms for Identification and Blind Equalization of Unit-Norm Channels

Claudio J. Bordin, *Member, IEEE*, and Marcelo G. S. Bruno, *Senior Member, IEEE*

***Abstract*—In many estimation problems of interest, the unknown parameters reside on spherical manifolds. As most common filtering algorithms assume that parameters have Gaussian prior distributions, their application to such problems leads to suboptimal performance. In this letter, we propose a model in which the unknown unit-norm parameter vectors have Fisher–Bingham (F-B) prior distributions. We show that if the observations relate to the parameters via Gaussian likelihoods, the F-B priors form a conjugate model that yields closed-form, recursive estimators that naturally take into account the restrictions on the unknowns. We apply this model to a communication setup with multiple gain-controlled FIR frequency-selective channels, deriving a novel maximum *a posteriori* (MAP) channel parameter estimator and a blind equalizer based on Rao–Blackwellized particle filters. As we verify via Monte Carlo numerical simulations, the F-B model leads to superior performance compared to previous algorithms that adopt mismatched Gaussian prior models.**

***Index Terms*—Bayes methods, blind equalizers, particle filters, system identification.**

I. Introduction

Wireless digital communication systems require some form of automatic gain control (AGC) to cope with variable channel gain [1]. In a (quasi)stationary environment in which the mean transmitted signal power and the mean noise power are constant, and the propagation channel is (almost) static, AGC asymptotically sets the equivalent baseband channel gain [2, p. 336]. Under the usual assumption of an FIR channel model [3], [4], [5], [6], [7], [8], this can be verified to constrain the channel parameter vector norm (without loss of generality, to unity).

Many previous blind channel equalization or identification algorithms rely on AGC to compensate for otherwise unidentifiable channel gains [3], [4], [5], [9, p.73], or as an auxiliary mechanism to avoid divergence [10]. However, in a *Bayesian* setup in which the channel parameters are unknown and are to be estimated (e.g., channel identification) or integrated out as nuisance parameters [11] (e.g., blind equalization), norm constraints preclude usual Gaussianity assumptions on which previous techniques [6]–[8], [12]–[23] are based. To remediate this, we introduce in this letter new Bayesian recursive methods for channel identification and blind equalization based on the Fisher-Bingham (F-B) distribution [24, p. 174].

The F-B distribution arises when a multivariate normal random vector is conditioned to have unit length [25]. Similar distributions defined on (hyper)spheres (e.g., Von Mises-Fisher, Bingham, Kent [24]) have found use in filtering problems. In [26], an approximate algorithm to track Von Mises-Fisher-distributed unit-norm rotation parameters is introduced. A mixture Bingham model is employed in [27] to fit the posterior distribution of unit-norm quaternions in a pose estimation problem.

This work innovates by introducing a conjugate Gaussian-F-B model. Namely, we show that if the *unknown* unit-norm channel parameter vector has F-B prior distribution and the likelihood of the observables is Gaussian, then the channel parameter vector posterior distribution is also a F-B distribution whose parameters can be *exactly* evaluated via recursive expressions. This allows us to derive a novel optimal MAP (*maximum a posteriori*) closed-form channel parameter estimator and a new blind equalizer based on Rao-Blackwellized particle filters [11] that analytically integrate out the unit-norm parameter vector.

The remainder of the text is organized as follows: in Section II we describe the examined estimation problems. In Section III, in turn, we prove the aforementioned conjugacy properties and derive the F-B MAP channel parameter estimator. Afterward, in Section IV, we derive the densities needed to operate a Rao-Blackwellized particle-filter-based blind equalizer. A computational validation of the methods is presented in Section V and our final comments are left to Section VI.

A) Notation: We employ the notation $y_n \triangleq \{y_{1,n}, \dots, y_{R,n}\}$, where R denotes the number of receivers and, likewise, $\mathbf{h} \triangleq \{\mathbf{h}_1, \dots, \mathbf{h}_R\}$. We (non-exclusively) use capitals to denote sequences, i.e., $\mathbf{X}_n \triangleq \{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$, $Y_{r,n} \triangleq \{y_{r,1}, \cdots, y_{r,n}\}$, and $Y_n \triangleq \{y_1, \cdots, y_n\}$. The symbol $\mathcal{N}(\mathbf{x}|\mu; \mathbf{\Sigma})$ denotes a (possibly multivariate) Gaussian density with mean μ and (co)variance matrix $\mathbf{\Sigma}$.

II. Problem Formulation

We consider a multiple-output FIR frequency-selective channel setup in which $y_{r,n}$, $r \in \{1, \dots, R\}$, the signal received by the $r-$th receiver at the time instant n, is modeled as

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

Manuscript received June 13, 2015; revised July 30, 2015; accepted July 30, 2015. Date of publication August 03, 2015; date of current version August 13, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Simon Maskell.

C. J. Bordin is with the CECS, Universidade Federal do ABC, CEP 09210-580 Santo André, Brazil (e-mail: claudio.bordin@ufabc.edu.br).

M. G. S. Bruno is with the Divisão de Engenharia Eletrônica, Instituto Tecnológico de Aeronáutica, CEP 12228-900 São José dos Campos, Brazil (e-mail: bruno@ita.br).

This paper has supplemental downloadable multimedia material available at http://ieeexplore.ieee.org, provided by the authors. The Supplementary Material contains the C/MATLAB© source code for the algorithms described in the paper and the corresponding output data. This material is 1.3 MB in size.

Digital Object Identifier 10.1109/LSP.2015.2464154

where $\mathbf{x}_n \triangleq [x_n \cdots x_{n-L+1}]^T$ collects the last L transmitted symbols, L is the channel order (without loss of generality, considered equal for all receivers), $v_{r,n}$ is a zero-mean independent, identically distributed (i.i.d.) Gaussian process of known variance σ_r^2, and $\mathbf{h}_r \in \mathbb{R}^L$ collects the channel impulse response terms between the transmitter and the $r-$th receiver. The random quantities $\mathbf{h}_r$, $\{\mathbf{x}_n\}$, and $\{v_{r,n}\}$, $r \in \{1,\ldots,R\}$, are presumed to be mutually independent a priori.

Due to the aforementioned norm constraint, the channel parameters vector $\mathbf{h}_r$ is assumed to have an $L-$ variate F-B prior distribution

$$\begin{aligned} p(\mathbf{h}_r) = \mathcal{FB}(\mathbf{h}_r|\mathbf{a}_{r,0},\mathbf{B}_{r,0}) &\triangleq C(\mathbf{a}_{r,0},\mathbf{B}_{r,0})^{-1} \\ &\times \exp(\mathbf{h}_r^T\mathbf{B}_{r,0}\mathbf{h}_r + \mathbf{h}_r^T\mathbf{a}_{r,0})\mathcal{I}_{\|\mathbf{h}_r\|_2=1}, \end{aligned} \tag{2}$$

where $\mathbf{a}_{r,0} \in \mathbb{R}^L$ and the symmetrical matrix $\mathbf{B}_{r,0} \in \mathbb{R}^{L\times L}$ are given hyperparameters, $\mathcal{I}_{\{.\}}$ denotes the indicator function, and $C(\mathbf{a}_{r,0},\mathbf{B}_{r,0})$ is the normalization constant, i.e.,

$$\begin{aligned} C(\mathbf{a}_{r,0},\mathbf{B}_{r,0}) = \int_{\mathbf{h}_r\in\mathcal{S}^{L-1}} &\exp(\mathbf{h}_r^T\mathbf{B}_{r,0}\mathbf{h}_r + \mathbf{h}_r^T\mathbf{a}_{r,0}) \\ &\times d\mathcal{S}^{L-1}(\mathbf{h}_r), \end{aligned} \tag{3}$$

where $\mathcal{S}^{L-1} = \{\mathbf{h}_r \in \mathbb{R}^L : \|\mathbf{h}_r\|_2 = 1\}$ denotes the $L-1$ unit sphere and $d\mathcal{S}^{L-1}$ the volume element [28] on $\mathcal{S}^{L-1}$.

Although there is no known analytical expression for $C(\mathbf{a}_{r,0},\mathbf{B}_{r,0})$, numerical estimates can be computed in $\mathcal{O}(L^3)$ via the method described in [25, Section II-B], which employs so-called *saddlepoint* approximations. Observe that if $\mathbf{a}_{r,0} = \mathbf{0}$ and $\mathbf{B}_{r,0} = \mathbf{0}$, $\mathcal{FB}(\mathbf{h}_r|\mathbf{a}_{r,0},\mathbf{B}_{r,0})$ reduces to a uniform distribution on the unit sphere. Note also that, contrary to the Gaussian counterpart, $\mathbf{B}_{r,0}$ is not restricted to be positive definite.

III. Channel Identification

In this section, we derive a recursive method to evaluate the locally trained[1,2] MAP (*maximum a posteriori*) estimate

$$\hat{\mathbf{h}}_{r,n} \triangleq \arg\max_{\mathbf{h}_r} p(\mathbf{h}_r|\mathbf{X}_n, Y_{r,n}). \tag{4}$$

To this aim, we first demonstrate the following proposition.

Proposition: Under the assumptions (1) and (2), the posterior distribution

$$p(\mathbf{h}_r|\mathbf{X}_n, Y_{r,n}) = \mathcal{FB}(\mathbf{h}_r|\mathbf{a}_{r,n},\mathbf{B}_{r,n}), \tag{5}$$

where the parameters $\mathbf{a}_{r,n}$ and $\mathbf{B}_{r,n}$ can be recursively determined via (10)–(11).

Proof: The proof is inductive. First, observe that

$$p(\mathbf{h}_r|\mathbf{X}_n, Y_{r,n}) = \frac{p(\mathbf{h}_r, \mathbf{x}_n, y_{r,n}|\mathbf{X}_{n-1}, Y_{r,n-1})}{p(\mathbf{x}_n, y_{r,n}|\mathbf{X}_{n-1}, Y_{r,n-1})}. \tag{6}$$

Exploiting Markovian properties induced by (1), it follows that [19]

$$\begin{aligned} p(\mathbf{h}_r, \mathbf{x}_n, y_{r,n}|\mathbf{X}_{n-1}, Y_{r,n-1}) &= p(y_{r,n}|\mathbf{x}_n, \mathbf{h}_r) \\ &\times p(\mathbf{x}_n|\mathbf{x}_{n-1})p(\mathbf{h}_r|\mathbf{X}_{n-1}, Y_{r,n-1}), \end{aligned} \tag{7}$$

where $p(\mathbf{x}_n|\mathbf{x}_{n-1})$ is a discrete probability mass function (p.m.f.) assumed known to the receiver, and $p(y_{r,n}|\mathbf{x}_n,\mathbf{h}_r) = \mathcal{N}(y_{r,n}|\mathbf{h}_r^T\mathbf{x}_n;\sigma_r^2)$. We assume now the inductive hypothesis that

$$p(\mathbf{h}_r|\mathbf{X}_{n-1}, Y_{r,n-1}) = \mathcal{FB}(\mathbf{h}_r|\mathbf{a}_{r,n-1},\mathbf{B}_{r,n-1}). \tag{8}$$

Substituting (8) into (7), we obtain that

$$\begin{aligned} p(\mathbf{h}_r, \mathbf{x}_n, y_{r,n}|\mathbf{X}_{n-1}, Y_{r,n-1}) &= p(\mathbf{x}_n|\mathbf{x}_{n-1}) \\ \times \frac{\exp\left(-\frac{y_{r,n}^2}{2\sigma_r^2}\right)}{\sqrt{2\pi\sigma_r^2}} &\frac{\exp\left(\mathbf{h}_r^T B_{r,n}\mathbf{h}_r + \mathbf{h}_r^T\mathbf{a}_{r,n}\right)}{C(\mathbf{a}_{r,n-1},\mathbf{B}_{r,n-1})}\mathcal{I}_{\|\mathbf{h}_r\|_2=1}, \end{aligned} \tag{9}$$

where

$$B_{r,n} = B_{r,n-1} - \mathbf{x}_n\mathbf{x}_n^T/2\sigma_r^2, \tag{10}$$

$$\mathbf{a}_{r,n} = \mathbf{a}_{r,n-1} + \mathbf{x}_n y_{r,n}/\sigma_r^2. \tag{11}$$

Note that (9) depends on $\mathbf{h}_r$ through a F-B density (Eq. (2)) without the appropriate normalization term. Therefore, integrating $\mathbf{h}_r$ out of (9) results that

$$\begin{aligned} p(\mathbf{x}_n, y_{r,n}|\mathbf{X}_{n-1}, Y_{r,n-1}) &= p(\mathbf{x}_n|\mathbf{x}_{n-1}) \\ \times \frac{\exp\left(-\frac{y_{r,n}^2}{2\sigma_r^2}\right)}{\sqrt{2\pi\sigma_r^2}} &\frac{C(\mathbf{a}_{r,n},\mathbf{B}_{r,n})}{C(\mathbf{a}_{r,n-1},\mathbf{B}_{r,n-1})}. \end{aligned} \tag{12}$$

Dividing (9) by (12), it follows that

$$p(\mathbf{h}_r|\mathbf{X}_n, Y_{r,n}) = \frac{\exp\left(\mathbf{h}_r^T B_{r,n}\mathbf{h}_r + \mathbf{h}_r^T\mathbf{a}_{r,n}\right)}{C(\mathbf{a}_{r,n},\mathbf{B}_{r,n})}\mathcal{I}_{\|\mathbf{h}_r\|_2=1}, \tag{13}$$

which is equivalent to (5). This proves the result by the finite induction principle as, for $n = 1$, (8) reduces to (2). ■

The MAP solution sought in (4) can be recast as the constrained optimization problem

$$\begin{aligned} \hat{\mathbf{h}}_{r,n} = \arg\max_{\mathbf{h}_r} \quad & \exp(\mathbf{h}_r^T\mathbf{B}_{r,n}\mathbf{h}_r + \mathbf{h}_r^T\mathbf{a}_{r,n}), \\ \text{subject to} \quad & \|\mathbf{h}_r\|_2 = 1 \end{aligned} \tag{14}$$

as the normalization constant $C(\mathbf{a}_{r,n},\mathbf{B}_{r,n})$ does not depend on $\mathbf{h}_r$. The Lagrange function corresponding to (14) is given by

$$\Lambda(\mathbf{h}_r,\lambda) \triangleq \exp(\mathbf{h}_r^T\mathbf{B}_{r,n}\mathbf{h}_r + \mathbf{h}_r^T\mathbf{a}_{r,n}) + \lambda(\mathbf{h}_r^T\mathbf{h}_r - 1). \tag{15}$$

Taking the gradient of (15) with respect to $\mathbf{h}_r$, we obtain that

$$\begin{aligned} \frac{\partial\Lambda(\mathbf{h}_r,\lambda)}{\partial\mathbf{h}_r} &= (2\mathbf{B}_{r,n}\mathbf{h}_r + \mathbf{a}_{r,n}) \\ &\times \exp(\mathbf{h}_r^T\mathbf{B}_{r,n}\mathbf{h}_r + \mathbf{h}_r^T\mathbf{a}_{r,n}) + 2\lambda\mathbf{h}_r. \end{aligned} \tag{16}$$

Dividing (16) by the exponential term and equating the result to zero, it follows that

$$\hat{\mathbf{h}}_{r,n} = -\frac{1}{2}\left(\mathbf{B}_{r,n} + \tilde{\lambda}\mathbf{I}\right)^{-1}\mathbf{a}_{r,n}. \tag{17}$$

where $\tilde{\lambda} \triangleq \lambda\exp(-\hat{\mathbf{h}}_{r,n}^T\mathbf{B}_{r,n}\hat{\mathbf{h}}_{r,n} - \hat{\mathbf{h}}_{r,n}^T\mathbf{a}_{r,n})$.

Equating the derivative of (15) with respect to λ to zero is equivalent to enforcing the constraint $\|\hat{\mathbf{h}}_{r,n}\|_2^2 = 1$. Noting that as $\mathbf{B}_{r,n}$ is by construction symmetrical, we can plug its

[1] The algorithm derived in Section III could in principle deal with arbitrary signals $\mathbf{X}_n$. Proper operation of AGC, however, may impose restrictions.

[2] Due to independence assumptions made in Section II, the use of remote observations $Y_{s,n}$, $s \neq r$, does not improve trained estimates of $\mathbf{h}_r$.

for $n > 0$, fixed r **do**
- Evaluate $\mathbf{a}_{r,n}$ and $\mathbf{B}_{r,n}$ via (10)-(11).
- Determine $\{\tilde{\lambda}^{<i>}\}$, the set of real $\tilde{\lambda}$ that solve (19).
- **for** all $\tilde{\lambda} \in \{\tilde{\lambda}^{<i>}\}$ **do**
  - Evaluate a candidate $\hat{\mathbf{h}}_{r,n}^{<i>}$ via (17).
  - Compute the unnormalized posterior $\tilde{p}^{<i>} \triangleq \exp((\hat{\mathbf{h}}_{r,n}^{<i>})^T \mathbf{B}_{r,n} \hat{\mathbf{h}}_{r,n}^{<i>} + (\hat{\mathbf{h}}_{r,n}^{<i>})^T \mathbf{a}_{r,n})$.
- **end for**
- Set $\hat{\mathbf{h}}_{r,n} = \hat{\mathbf{h}}_{r,n}^{<j>}$, where $j = \max_i \tilde{p}^{<i>}$.

end for

Fig. 1. Channel Identification Algorithm.

eigenvalue decomposition $\mathbf{B}_{r,n} \triangleq \mathbf{U}_{r,n}\mathbf{D}_{r,n}\mathbf{U}_{r,n}^T$ into (17) and rewrite the constraint[3] as

$$\mathbf{a}_{r,n}^T \mathbf{U}_{r,n}\left(\mathbf{D}_{r,n} + \tilde{\lambda}\mathbf{I}\right)^{-2}\mathbf{U}_{r,n}^T \mathbf{a}_{r,n} = 4, \tag{18}$$

Equation (18) is equivalent to

$$\sum_{k=1}^{L}\left(\frac{\left[\mathbf{a}_{r,n}^T \mathbf{U}_{r,n}\right]^{[k]}}{\left[\mathbf{D}_{r,n}\right]^{[k]} + \tilde{\lambda}}\right)^2 = 4 \tag{19}$$

which can be rewritten as a $2L$ degree polynomial equation and numerically solved for $\tilde{\lambda}$, where $[\cdot]^{[k]}$ denotes the $k-$th element of a vector or of a matrix diagonal.

The complete algorithm to determine $\hat{\mathbf{h}}_{r,n}$ is summarized in Fig. 1.

Remark: Note that (17) has the functional form of a regularized least-squares solution [29]. Moreover, if $\mathbf{B}_{r,n}$ is proportional to the identity, which happens asymptotically if x_n is i.i.d., it can be shown that $\hat{\mathbf{h}}_{r,n}$ is a normalized version of the least-squares [29] estimate $\hat{\mathbf{h}}_{r,n}^{LS} = -\mathbf{B}_{r,n}^{-1}\mathbf{a}_{r,n}/2$.

Computational Complexity. The algorithm of Fig. 1 requires an exhaustive search among the set of real solutions of (19) for the value $\tilde{\lambda}^{<i>}$ that, substituted into (17), maximizes the unnormalized posterior in (14). The complexity of the algorithm is dominated by the evaluation of (17), which demands $\mathcal{O}(L^3)$ but must be repeated $2L$ times (at most). The remaining operations, namely, the evaluation of the eigenvalue decomposition of B_n and the solution of the polynomial equation implicit in (19) also demand $\mathcal{O}(L^3)$, but need to be executed only once. The overall complexity of the algorithm is, therefore, $\mathcal{O}(L^4)$.

IV. Multichannel Blind Equalization

Under the additional assumption that the transmitted symbols $\{x_n\} \in \{\pm 1\}$ are binary and obtained by differentially encoding the i.i.d. bit stream $\{b_n\} \in \{\pm 1\}$, we aim at obtaining the joint MAP estimate $\hat{b}_n = \arg\max_{b_n} p(b_n|Y_n)$, that employs the observations available at all receiving *nodes* via a particle filtering method. Observing that $\mathbf{X}_n$ uniquely determines b_n, we employ a particle filter [11] to approximate the posterior probability mass function

$$p(\mathbf{X}_n|Y_n) \approx \sum_{p=1}^{P} w_n^{(p)} \mathcal{I}_{\{\mathbf{X}_n - \mathbf{X}_n^{(p)}\}}, \tag{20}$$

[3]Observe that the constraint $\|\hat{\mathbf{h}}_{r,n}\|^2 = 1$ can be enforced *without* explicitly evaluating λ, but only $\tilde{\lambda}$.

for $n > 0$ **do**
- **for** $p = 1 : P$ **do**
  - Draw $\mathbf{x}_n^{(p)} \sim p(\mathbf{x}_n|\mathbf{x}_{n-1}^{(p)})$.
  - **for** $r = 1 : R$ **do**
    - Evaluate $\mathbf{a}_{r,n}^{(p)}$ and $\mathbf{B}_{r,n}^{(p)}$ via (10)-(11).
    - Evaluate $p(y_{r,n}|\mathbf{X}_n^{(p)}, Y_{r,n-1})$ via (24)
  - **end for**
  - Evaluate $p(y_n|\mathbf{X}_n^{(p)}, Y_{n-1})$ via (22)
  - Update $w_n^{(p)}$ via (21)
- **end for**
- Normalize the weights, i.e., $\sum_{p=1}^{P} w_n^{(p)} = 1$.
- Estimate $p(b_n|Y_n) \approx \sum_{p=1}^{P} w_n^{(p)} \mathcal{I}_{\{b_n - b_n^{(p)}\}}$, where $b_n^{(p)} = x_n^{(p)} x_{n-1}^{(p)}$ and determine $\hat{b}_n$.
- Resample the particle set $\{\mathbf{X}_n^{(p)}, w_n^{(p)}\}_{p=1}^{P}$.

end for

Fig. 2. Multichannel Blind Equalization Algorithm.

where P denotes the number of particles $\mathbf{X}_n^{(p)}$, random samples of the state trajectory, and $w_n^{(p)}$ the corresponding normalized weights.

We adopt the so-called prior importance function [30], by which the particles are extended as $\mathbf{x}_n^{(p)} \sim p(\mathbf{x}_n|\mathbf{x}_{n-1}^{(p)})$. The corresponding weight update expression is the given by

$$w_n^{(p)} \propto w_{n-1}^{(p)} p(y_n|\mathbf{X}_n^{(p)}, Y_{n-1}), \tag{21}$$

where we remind the reader that the uppercase variables contain the complete time series from $n = 1$.

The previous assumptions of a priori independence of the channel parameters and noise samples at each receiver imply that [31]

$$p(y_n|\mathbf{X}_n^{(p)}, Y_{n-1}) = \prod_{r=1}^{R} p(y_{r,n}|\mathbf{X}_n^{(p)}, Y_{r,n-1}). \tag{22}$$

Likewise, conditional independence relations induced by (1) result [31] that

$$p(y_{r,n}|\mathbf{X}_n^{(p)}, Y_{r,n-1}) = \int_{\mathbf{h}_r \in \mathcal{S}^{L-1}} p(y_{r,n}|\mathbf{x}_n^{(p)}, \mathbf{h}_r) \times p(\mathbf{h}_r|\mathbf{X}_{n-1}^{(p)}, Y_{r,n-1}) d\mathcal{S}^{L-1}(\mathbf{h}_r). \tag{23}$$

Note that the integrand in (23) is very similar to (7) except for the term $p(\mathbf{x}_n|\mathbf{x}_{n-1})$, which is absent from (23). Therefore (12) implies that

$$p(y_{r,n}|\mathbf{X}_n^{(p)}, Y_{r,n-1}) = \frac{1}{\sqrt{2\pi\sigma_r^2}} \exp\left(-\frac{y_{r,n}^2}{2\sigma_r^2}\right) \frac{C(\mathbf{a}_{r,n}^{(p)}, \mathbf{B}_{r,n}^{(p)})}{C(\mathbf{a}_{r,n-1}^{(p)}, \mathbf{B}_{r,n-1}^{(p)})}, \tag{24}$$

where $\mathbf{a}_{r,n}^{(p)}$ and $\mathbf{B}_{r,n}^{(p)}$ are defined as in (10)–(11) replacing $\mathbf{x}_n$ with $\mathbf{x}_n^{(p)}$.

The resulting blind equalization algorithm is summarized in Fig. 2. A *distributed* [32], [33], [34], [35], [36] version can be implemented as described in [31, Section III-A].

Computational Complexity. To run the particle filtering algorithm of Section IV, (21) must be evaluated P times. In each of

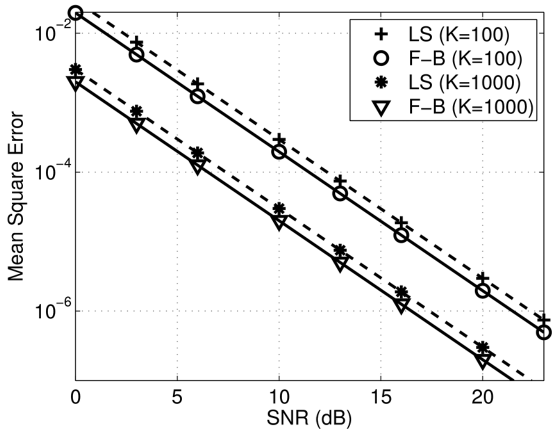

Fig. 3. Average mean square error (m.s.e.) in the identification of the parameter $\mathbf{h}_r$ as a function of the number of observed samples (K) and the signal-to-noise ratio (SNR) estimated in 5,000 Monte Carlo runs.

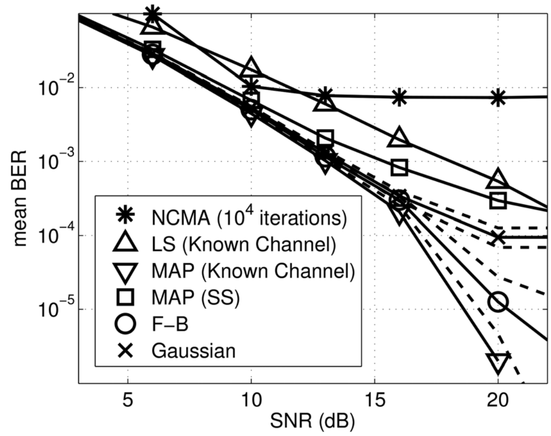

Fig. 4. Mean bit error rate (BER) estimated in 5,000 Monte Carlo runs as a function of the signal-to-noise ratio (SNR). The dashed lines surrounding the solid ones display the respective 95% confidence intervals.

those evaluations, (24) must be computed R times (one time for each receiving channel). The computational complexity of (24) is dominated by the evaluation of the normalization constant[4] $C(\mathbf{a}_{r,n}^{(p)}, \mathbf{B}_{r,n}^{(p)})$, which is $\mathcal{O}(L^3)$. The overall complexity of the algorithm is, therefore, $\mathcal{O}(PRL^3)$.

V. Experimental Results

A. Channel Identification

To evaluate the performance of the trained MAP identification algorithm (Section III), we ran 5,000 independent experiments, in each of which we generated K successive samples $y_{r,n}$ as defined by (1), using $L = 3$. We drew $\mathbf{h}_r$ by sampling independently in each realization from $\mathcal{N}(\mathbf{0}; \mathbf{I})$ and normalizing the result, so that $\mathbf{h}_r$ is uniformly distributed on the unit sphere [37]. We employed a matched prior density for $\mathbf{h}_r$ by setting the hyperparameters $\mathbf{a}_{r,0}$ and $\mathbf{B}_{r,0}$ to zero, which corresponds to a non-informative [38] uniform prior.

Fig. 3 shows the estimated mean square error (m.s.e.) $E\|\hat{\mathbf{h}}_{r,K} - \mathbf{h}_r\|^2$ resulting from the algorithm in Fig. 1 (F-B) using $K = 100$ and 1,000 samples. For comparison, we show in the same plot the performance of the least squares (trained) estimator (LS), which, for the considered setup, is equivalent to a MAP estimator that adopts a mismatched Gaussian prior for $\mathbf{h}_r$ [29]. As one may verify, the algorithm employing F-B prior led to a lower m.s.e. (about 66% of the LS estimator m.s.e.) for all the considered signal-to-noise ratio (SNR) levels; this performance advantage is also maintained when the observation window is enlarged ($K = 1,000$).

B. Blind Equalization

We assessed the steady-state performance of the proposed blind equalization algorithm (Section IV) via Monte Carlo simulations consisting of 5,000 independent runs, in each of which the mean bit error rate (BER) was estimated as a function of the SNR, assumed for simplicity equal on all receivers. In each realization, an i.i.d. sequence of 250 differentially encoded binary symbols was transmitted, being the first 150 bits discarded to allow for convergence. The simulated system has $R = 4$ receivers. The parameters $\mathbf{h}_r$, $r \in \{1,,4\}$, were sampled independently for each r as described in Section V-A.

As in Section V-A, the hyperparameters $\mathbf{a}_{r,0}$ and $\mathbf{B}_{r,0}$ were set to zero. The particle filter employs $P = 300$ particles and performs systematic resampling [39] at each time step.

Fig. 4 displays the mean performance of the blind algorithm of Section IV (F-B) ($\circ$) and that of the equivalent blind method that employs mismatched Gaussian priors [31, Section III-A] ($\times$). The 95% confidence intervals for both means are displayed as dashed lines. For comparison, Fig. 4 also depicts the performance of zero-delay equalizers based on the MAP criterion (grid filter [40]) using exact channel parameters (∇) and parameters estimated via the Subspace technique [4] ($\square$). Fig. 4 also shows the mean performance of the zero-delay 15-tap linear least-squares equalizer [29] (LS - Δ), and that of the NCMA[5] [29] ($*$) after 10^4 iterations, using 7 taps and $\mu = 1$.

As one may note, the performance of the new algorithm (F-B) is similar to that of the optimal MAP equalizer and that of the Gaussian particle-filter-based method for low and medium SNR levels. For high SNR levels, on the other hand, the performance of the new method exceeds that of [31] by a statistically significant margin.

VI. Conclusion

In this work we described new algorithms for channel identification and blind equalization using a Fisher-Bingham prior model for the unknown parameters. We introduced a conjugate model that led to closed-form expressions for the parameters of some required posterior densities, dropping with the need for further approximations as those performed by previous works that employed other sphere-constrained distributions (Von Mises-Fisher [26], Bingham [27]).

As we assessed via Monte Carlo simulations, the new channel identification and blind equalization algorithms outperformed conventional algorithms that adopt mismatched Gaussian priors, at the cost of increased computational complexity. A limitation of our simulations is that they generate synthetic signals with statistics exactly matched to the proposed Fisher-Bingham prior model. As future work, we plan to test the proposed algorithms with real data taking into account AGC power oscillation, time-varying channels and synchronization issues.

[4]Observe that $C(\mathbf{a}_{r,n-1}^{(p)}, \mathbf{B}_{r,n-1}^{(p)})$ in (24) is evaluated at time $n-1$ and, therefore, does not need to be re-evaluated at time n.

[5]CMA-type algorithms displayed BER's of about 30% for the considered scenario after 150 iterations. Note that the performance of NCMA may surpass that of the optimal zero-delay LS equalizer due to convergence to a solution with different delay.

References

[1] P. Baudin, *Wireless Transceiver Architecture: Bridging RF And Digital Communications*. New York, NY, USA: Wiley, 2014.

[2] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw-Hill, 2000.

[3] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *25th Asilomar Conf. Signals, Syst., Comput., Conf. Rec.*, 1991, pp. 856–860.

[4] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[5] A. Chevreuil, E. Serpedin, P. Loubaton, and G. B. Giannakis, "Blind channel identification and equalization using periodic modulation precoders: Performance analysis," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1570–1586, Jun. 2000.

[6] J. S. Liu and R. Chen, "Blind deconvolution via sequential imputations," *J. Amer. Statist. Assoc.*, vol. 90, no. 430, pp. 567–576, 1995.

[7] R. Chen, J. S. Liu, and X. Wang, "Convergence analyses and comparisons of Markov chain Monte Carlo algorithms in digital communications," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 255–270, Feb. 2002.

[8] J. Míguez and P. M. Djurić, "Blind equalization of frequency-selective channels by sequential importance sampling," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 2738–2748, Oct. 2004.

[9] Z. Ding and Y. Li, *Blind Equalization and Identification*, ser. Signal Processing and Communications. Boca Raton, FL, USA: CRC Press, 2001.

[10] S. Verdu, B. D. O. Anderson, and R. A. Kennedy, "Blind equalization without gain identification," *IEEE Trans. Inf. Theory*, vol. 39, no. 1, pp. 292–297, Jan. 1993.

[11] A. Doucet, A. Smith, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*, ser. Information Science and Statistics. New York, NY, USA: Springer, 2010.

[12] T. C. Clapp and S. J. Godsill, "Fixed-lag blind equalization and sequence estimation in digital communications systems using sequential importance sampling," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 1999, vol. 5, pp. 2495–2498.

[13] T. Clapp, "Statistical methods for the processing of communications data," Ph.D. dissertation, Univ. Cambridge, Cambridge, U.K., 2000.

[14] J. Míguez and P. M. Djurić, "Blind equalization by sequential importance sampling," in *Proc. IEEE Int. Symp. Circuits Syst., (ISCAS'02)*, 2002, vol. 1, pp. I–845–I–848.

[15] E. Punskaya, "Sequential Monte Carlo methods for digital communications," Ph.D. dissertation, Cambridge Univ. Eng. Dept., Cambridge, U.K., 2003.

[16] D. Yee, J. P. Reilly, and T. Kirubarajan, "Blind particle filtering for detection in a time-varying frequency selective channel with non-Gaussian noise," in *Proc. IEEE 6th Workshop Signal Process. Adv. Wireless Commun.*, 2005, pp. 925–929.

[17] T. Ghirmai, M. F. Bugallo, J. Míguez, and P. M. Djurić, "A sequential Monte Carlo method for adaptive blind timing estimation and data detection," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2855–2865, Aug. 2005.

[18] S. Barembruch, A. Garivier, and E. Moulines, "On optimal sampling for particle filtering in digital communication," in *Proc. IEEE 9th Workshop Signal Process. Adv. Wireless Commun., (SPAWC'08)*, 2008, pp. 634–638.

[19] C. J. Bordin and M. G. S. Bruno, "Particle filters for joint blind equalization and decoding in frequency-selective channels," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2395–2405, Jun. 2008.

[20] P. Closas, C. Fernandez-Prades, and J. A. Fernandez-Rubio, "A Bayesian approach to multipath mitigation in GNSS receivers," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 4, pp. 695–706, Aug. 2009.

[21] S. Barembruch, "A comparison of approximate viterbi techniques and particle filtering for data estimation in digital communications," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2010, pp. 3826–3829.

[22] A. A. Nasir, S. Durrani, and R. A. Kennedy, "Particle filter for joint blind carrier frequency offset estimation and data detection," in *Proc. 4th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, 2010, pp. 1–5.

[23] F. Lehmann, "A Gaussian mixture approach to blind equalization of block-oriented wireless communications," *EURASIP J. Adv. Signal Process*, 2010, 2010, doi:10.1155/2010/340417.

[24] K. V. Mardia and P. E. Jupp, *Directional Statistics*. New York, NY, USA: Wiley, 2000.

[25] A. Kume, "Saddlepoint approximations for the Bingham and Fisher-Bingham normalising constants," *Biometrika*, vol. 92, no. 2, pp. 465–476, Jun. 2005.

[26] J. Traa and P. Smaragdis, "Multiple speaker tracking with the factorial von mises-fisher filter," in *Proc. IEEE Int. Workshop Mach. Learn. for Signal Process. (MLSP)*, Sep. 2014, pp. 1–6.

[27] J. Glover, G. Bradski, and R. B. Rusu, *Robotics:Science and Systems VII*. Cambridge, MA, USA: MIT Press, 2011, ch. Monte Carlo Pose Estimation with Quaternion Kernels and the Bingham Distribution, pp. 97–104.

[28] X. Pennec, "Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements," in *Proc. Int. Workshop Nonlinear Signal and Image Process. (NSIP99)*, Jun. 1999, pp. 194–198.

[29] A. H. Sayed, *Adaptive Filters*. New York, NY, USA: Wiley, 2011.

[30] A. Doucet, S. Godsill, and C. Andrieu, "On Sequential Monte Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, no. 3, pp. 197–208, Jul. 2000.

[31] C. J. Bordin and M. G. S. Bruno, "Consensus-based distributed particle filtering algorithms for cooperative blind equalization in receiver networks," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, May 2011, pp. 3968–3971.

[32] M. Coates, "Distributed particle filters for sensor networks," in *Proc. 3rd Int. Symp. Inf. Process. Sens. Netw.*, 2004, pp. 99–107.

[33] F. S. Cattivelli, C. G. Lopes, and A. H. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1865–1877, May 2008.

[34] A. Mohammadi and A. Asif, "Consensus-based distributed unscented particle filter," in *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Jun. 2011, pp. 237–240.

[35] S. Farahmand, S. I. Roumeliotis, and G. B. Giannakis, "Set-membership constrained particle filter: Distributed adaptation for sensor networks," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4122–4138, Sep. 2011.

[36] O. Hlinka, F. Hlawatsch, and P. M. Djurić, "Distributed particle filtering in agent networks: A survey, classification, and comparison," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 61–81, Jan. 2013.

[37] M. E. Muller, "A note on a method for generating points uniformly on N-dimensional spheres," *Commun. ACM*, vol. 2, no. 4, pp. 19–20, Apr. 1959.

[38] H. S. Migon, D. Gamerman, and F. Louzada, *Statistical Inference: An Integrated Approach*, ser. Chapman & Hall/CRC Texts in Statistical Science, 2nd ed. Boca Raton, FL, USA: CRC Press, 2015.

[39] R. Douc and O. Cappé, "Comparison of resampling schemes for particle filtering," in *Proc. 4th Intl. Symp. Image Signal Process. Anal. (ISPA)*, 2005, pp. 64–69.

[40] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.