Variational Bayesian Channel Tracking in High-speed Underwater Acoustic Communication

Peng Chen
Dept ECE
Curtin University
peng.ch@outlook.com

Yue Rong
Dept ECE
Curtin University
y.rong@curtin.edu.au

Sven Nordholm
Dept ECE
Curtin University
s.nordholm@curtin.edu.au

ABSTRACT

In this paper, two novel joint channel tracking and noise variance estimation algorithms for single-carrier underwater acoustic communication system are proposed. The new algorithms combine the advantages of the variational Bayesian (VB) method and the compressed sensing technique. The proposed algorithms overcome the over-fitting of conventional orthogonal matching pursuit (OMP) based methods, and thus, lead to an improved performance. The proposed algorithms are applied to detect data received during an experiment conducted in December 2012 in the Indian Ocean off the Rottnest Island, Western Australia. The results show that the proposed algorithms can reduce the uncoded system BER by 1% to 6%.

Keywords

Underwater communication, channel estimation, variational inference, compressed sensing

1. INTRODUCTION

The underwater acoustic (UA) channel is one of the most challenging channels for wireless communication. Due to the features of the medium, the UA channel is different from terrestrial radio channels in many aspects. Firstly, the propagation loss of acoustic waves in water is approximately proportional to square of the frequency [1]. Therefore, UA communication signals are usually transmitted at a low carrier frequency, and thus the bandwidth available for UA communication is extremely limited compared with that of terrestrial radio channels [2]. Secondly, the speed of UA wave near the sea surface is typically around 1520 m/s which is five orders of magnitude smaller than the speed of light [3]. Thirdly, the speed of UA wave is affected by many factors, such as temperature, salinity, and the pressure of water. These features of UA medium introduce rapid dispersion in both time and frequency domains to UA communication channels. The time-domain dispersion due to delay spread results in severe inter-symbol interference. The frequency-domain dispersion caused by the drift of the transmitter, receiver and/or the motion of water leads to rapidly time-varying communication channel.

It has been shown in [4] that many shallow-water UA channels have a sparse structure, which means that although the UA channel impulse response generally has extremely large delay spread, most of the channel energy is carried by only a few propagation paths. By exploiting the sparsity of the UA channel impulse response, channel estimators at receiver can have reduced number of taps, which reduces the noise involved in channel estimation. Consequently, the channel estimation can have an improved accuracy as well as a reduced computational complexity [5], [6].

One of the methods to exploit the sparse structure in channel estimation is the matching pursuit (MP) algorithm [7] or its orthogonal vision named the orthogonal matching pursuit (OMP) algorithm [8], both of which are considered as compressed sensing (CS) techniques. In particular, sparse channel estimation can be implemented by first selecting the most important paths of the sampled channel impulse response via a greedy L_p-norm regularized method and then estimating coefficients for all selected paths via the least-squares (LS) method. However, the LS method adopted in the coefficients estimation step is sensitive to noise [9].

On the other hand, a block-wise decision-directed channel tracking method has been developed in [10], where each received data frame is subdivided into data blocks, and each data block is decoded by using the channel state information (CSI) estimated from the detected symbols of the previous data block. The authors of [10] adopted the LS approach to perform channel estimation at each data block. In [11], the OMP approach has been applied to estimate the CSI at each data block by exploiting the sparsity of the UA channel, which yields a better performance than the LS approach. However, the paths selection step in the OMP algorithm may suffer from the error in detected symbols during channel tracking.

In this paper, we develop two novel channel estimation and tracking algorithms for single-carrier UA communication system. The proposed algorithms first select the most significant paths using the training sequence and then adopt the variational Bayesian (VB) method [9] to estimate the coefficient of the selected paths. The proposed algorithms overcome the over-fitting of conventional OMP based methods,

WUWNET '15, October 22-24 2015, Washington DC, USA

http://dx.doi.org/10.1145/2831296.2831314

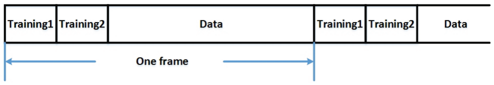

Figure 1: Frame structure of the system.

and thus, lead to an improved performance. The proposed algorithms are applied to detect data received during an experiment conducted in December 2012 in the Indian Ocean off the Rottnest Island, Western Australia. The results show that the proposed algorithms can reduce the uncoded system BER by 1% to 5%.

2. SYSTEM MODEL

As shown in Fig. 1, each data frame contains two training blocks with identical sequences followed by the data load. Let us introduce T_s as the symbol duration, N_T as the length of each training sequence, and N_D as the length of each data block. Then the length of each frame is $N_f = 2N_T + N_D$ and its duration is $T_f = T_s N_f$. We assume that the channel state is quasi-stationary within a block of N symbols, and thus the data sequence in one frame can be divided into $P = N_D/N$ blocks. The channel impulse response of the tth data block, $t = 1, \cdots, P$, can be represented by

$$\boldsymbol{h}(t) = [h_0(t), h_1(t), \cdots, h_{L-1}(t)]^T$$

where L denotes the maximum delay spread of the channel impulse response, and $(\cdot)^T$ stands for the matrix transpose.

The tth data block $\boldsymbol{r}(t) = [r_0(t), r_1(t), \cdots, r_{N-1}(t)]^T$ at the receiver side can be written as

$$\boldsymbol{r}(t) = \boldsymbol{D}(t)\boldsymbol{h}(t) + \boldsymbol{n}(t) \tag{1}$$

where

$$\boldsymbol{D}(t) = \begin{pmatrix} d_0(t) & d_N(t-1) & \cdots & d_{N-L+2}(t-1) \\ d_1(t) & d_0(t) & \cdots & d_{N-L+3}(t-1) \\ \vdots & \vdots & \ddots & \vdots \\ d_{N-1}(t) & d_{N-2}(t) & \cdots & d_{N-L}(t) \end{pmatrix}$$

with $d_i(t)$ denoting the ith transmitted symbol of the tth data block, and $\boldsymbol{n}(t) = [n_0(t), n_1(t), \cdots, n_{N-1}(t)]^T$ is the zero mean complex white Gaussian additive noise with variance $1/\beta$ per dimension.

Fig. 2 shows the system block diagram. It can be seen that at the receiver, the frame head is firstly found by the synchronization element and then the receiver performs frequency offset estimation and compensation for all data blocks. The above processing is performed to the pass-band signals. After the frequency offset compensation, the receiver removes the carrier frequency and passes the signals through a matched filter followed by down sampling. Then channel estimation, tracking, and equalization algorithms are applied to the down-sampled signals to estimate the transmitted data. In this paper, we develop a block-wise channel tracking method to track the time-varying UA channel at the receiver which combines the advantages of the VB method and the CS technique.

3. COMPRESSED SENSING METHOD

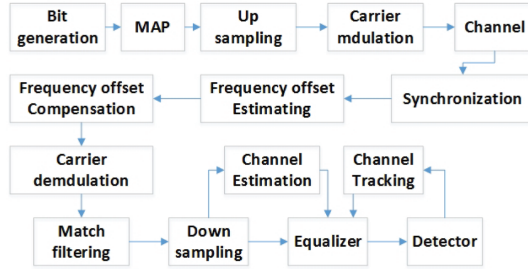

Figure 2: System block diagram.

Table 1: The OMP algorithm.

Initialization
$\hat{\boldsymbol{h}} = \boldsymbol{0}, \boldsymbol{y}(0) = \boldsymbol{r}, \boldsymbol{u}(0) = \varnothing, \tilde{\boldsymbol{D}}(0) = \varnothing$
For $s = 1, \cdots, S$
Calculate the correlation vector $\boldsymbol{b}(s) = \boldsymbol{D}^H \boldsymbol{y}(s-1)$
Find the index $p = \arg\max_{j=1\cdots L, j\notin \boldsymbol{u}(s-1)}(b_j(s))$
Update the index set $\boldsymbol{u}(s) = \boldsymbol{u}(s-1) \cup p$
Update $\tilde{\boldsymbol{D}}(s) = \tilde{\boldsymbol{D}}(s-1) \cup \boldsymbol{D}_{:,p}$
Update $\tilde{\boldsymbol{h}} = [\tilde{\boldsymbol{D}}(s)^H \tilde{\boldsymbol{D}}(s)]^{-1} \tilde{\boldsymbol{D}}(s)^H \boldsymbol{r}$
Update the residual measurement $\boldsymbol{y}(s) = \boldsymbol{r} - \tilde{\boldsymbol{D}}(s)\tilde{\boldsymbol{h}}$
end for
$\hat{h}_{u_i} = \tilde{h}_i$ for $i = 1, 2, \cdots, S$

CS is a technique that can recover signal accurately from its measurements provided that the signal is sparse. Let us consider the measurement model (1). For the simplicity of notation, in the following, we neglect the block index t and rewrite (1) as

$$\boldsymbol{r} = \boldsymbol{D}\boldsymbol{h} + \boldsymbol{n}. \tag{2}$$

Generally, the LS and/or the minimal mean-squared error (MMSE) methods can be applied to recover $\boldsymbol{h}$. However, if $\boldsymbol{h}$ is S-sparse, which means that $\boldsymbol{h}$ has $S(S \leq L)$ non-zero entries, and $\boldsymbol{D}$ is designed to capture the dominant information of $\boldsymbol{h}$ into $\boldsymbol{r}$, $\boldsymbol{h}$ can be recovered by the CS technique.

Many algorithms such as OMP, basis pursuit (BP), and compressed sampling matching pursuit (CoSaMP) have been developed for sparse signal recovery. In this paper, the OMP algorithm is adopted to perform channel estimation. Details of this algorithm are shown in Table 1, where $b_j(s)$ is the jth element of $\boldsymbol{b}$ in the sth iteration and $\boldsymbol{D}_{:,p}$ denotes the pth column of $\boldsymbol{D}$.

It can be seen from Table 1 that the OMP algorithm includes two steps. The first step adopts a greedy L_p-norm regularized method to search the most significant paths of the channel impulse response in an energy decreasing order, resulting in the index vector $\boldsymbol{u}$. In the second step, the LS method is used to estimate coefficients for all selected paths, leading to $\tilde{\boldsymbol{h}}$. Finally, the algorithm maps $\tilde{\boldsymbol{h}}$ to the corresponding position of the estimated channel impulse response $\hat{\boldsymbol{h}}$ using the index vector $\boldsymbol{u}$. However, the LS method adopted in the coefficients estimation step is sensitive to noise [9]. To overcome this problem, we introduce the VB method in the next section.

4. VARIATIONAL BAYESIAN METHOD

Let us introduce $\boldsymbol{\theta}$ as the vector of parameters and latent variables and $\boldsymbol{z}$ as the observation vector. The estimation problem is to estimate $\boldsymbol{\theta}$ or part of $\boldsymbol{\theta}$ when $\boldsymbol{z}$ is provided. This can be achieved by maximizing the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{z})$ as

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}|\boldsymbol{z}) = \arg\max_{\boldsymbol{\theta}} \frac{p(\boldsymbol{z}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{z})} \tag{3}$$

where

$$p(\boldsymbol{z}) = \int p(\boldsymbol{z}, \boldsymbol{\theta})\, d\boldsymbol{\theta}. \tag{4}$$

However, in many cases, the integration (4) is analytically intractable, and the VB method can be used to bypass this integral. This method introduces a distribution $q(\boldsymbol{\theta})$ which provides an approximation to the true posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{z})$. From the Jensen's inequality, we have

$$\begin{aligned}\ln p(\boldsymbol{z}) &= \ln \int p(\boldsymbol{z}, \boldsymbol{\theta})\, d\boldsymbol{\theta} = \ln \int q(\boldsymbol{\theta}) \frac{p(\boldsymbol{z}, \boldsymbol{\theta})}{q(\boldsymbol{\theta})}\, d\boldsymbol{\theta} \\ &\geq \int q(\boldsymbol{\theta}) \ln \frac{p(\boldsymbol{z}, \boldsymbol{\theta})}{q(\boldsymbol{\theta})}\, d\boldsymbol{\theta} = \mathcal{L}(q). \end{aligned} \tag{5}$$

Note that $\mathcal{L}(q)$ in (5) is a rigorous lower bound of the true log marginal likelihood, which is tractable to compute provided that a suitable $q(\cdot)$ distribution is chosen.

It is easy to find that the Kullback-Leibler (KL) divergence between the approximation distribution $q(\boldsymbol{\theta})$ and the true posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{z})$ is given by

$$KL(q||p) = \ln p(\boldsymbol{z}) - \mathcal{L}(q) = -\int q(\boldsymbol{\theta}) \ln \frac{p(\boldsymbol{\theta}|\boldsymbol{z})}{q(\boldsymbol{\theta})}\, d\boldsymbol{\theta}. \tag{6}$$

It can be seen from (6) that maximizing the lower bound $\mathcal{L}(q)$ is equivalent to minimize the KL divergence. A suitable form of $q(\boldsymbol{\theta})$ should be sufficiently simple so that the lower bound $\mathcal{L}(q)$ can be readily evaluated. Moreover, $q(\boldsymbol{\theta})$ should be sufficiently flexible so that the lower bound $\mathcal{L}(q)$ can be reasonably tight.

However, the best $q(\boldsymbol{\theta})$ may be difficult to compute in many cases of interest. In this situation, the VB method considers a particular $q(\cdot)$ by assuming that it can be factorized over the component variables in $\boldsymbol{\theta}$, which is called the mean field approximation in the theoretical physics area. Then $q(\boldsymbol{\theta})$ can be represented as

$$q(\boldsymbol{\theta}) = \prod_i q_i(\theta_i)$$

where the parameters in $\boldsymbol{\theta}$ is collected into separate groups θ_i each with their own approximate posterior distribution $q(\theta_i)$. Assuming that the posterior distributions of all groups are independent, the computation of $q(\theta_i)$ is to maximize the lower bound $\mathcal{L}(q)$ over $q(\theta_i)$, which results in

$$\ln q(\theta_i) = \langle \ln p(\boldsymbol{z}, \boldsymbol{\theta}) \rangle_{k \neq i} \tag{7}$$

where $\langle \cdot \rangle_{k\neq i}$ denotes the expectation with respect to the distributions $q_k(\theta_k)$ for all $k \neq i$. The VB method iteratively maximizes the lower bound $\mathcal{L}(q)$ with respect to the distributions $q_k(\theta_k)$ for all k, which is essentially the coordinate ascent in the function space of variational distributions.

5. VB AND OMP BASED CHANNEL ESTIMATION/TRACKING ALGORITHMS

In this section, we propose two novel channel estimation and tracking algorithms, namely the VB-based OMP and OMP-based VB algorithms. Both proposed algorithms apply the OMP method to obtain the channel paths index vector $\boldsymbol{u}$ and then use the VB technique to update the vector $\tilde{\boldsymbol{h}}$ before mapping $\tilde{\boldsymbol{h}}$ to $\hat{\boldsymbol{h}}$. The difference between two algorithms is that, in the OMP-based VB algorithm, the paths searching step is performed during the training sequence only, while this step is applied during both the training and data stream in the VB-based OMP method. The additional update step increases the computational complexity.

Using the channel paths index vector $\boldsymbol{u}$ obtained by the OMP method as shown in Table 1, the received signal vector $\boldsymbol{r}$ in (2) can be represented as

$$\boldsymbol{r} = \tilde{\boldsymbol{D}}\tilde{\boldsymbol{h}} + \boldsymbol{n}$$

where the measurement matrix $\tilde{\boldsymbol{D}}$ is generated from columns of $\boldsymbol{D}$ in (2) using $\boldsymbol{u}$. The likelihood distribution of $\tilde{\boldsymbol{h}}$ and β thus can be written as

$$p(\boldsymbol{r}|\tilde{\boldsymbol{h}}, \beta) = \frac{\beta^N}{\pi^N} \exp(-\beta(\boldsymbol{r} - \tilde{\boldsymbol{D}}\tilde{\boldsymbol{h}})^H(\boldsymbol{r} - \tilde{\boldsymbol{D}}\tilde{\boldsymbol{h}})) \tag{8}$$

where $(\cdot)^H$ stands for the matrix Hermitian transpose. To utilize the VB method for joint channel tracking and channel noise variance estimation, we first choose a Gaussian prior distribution with a distinct inverse variance α_l, $l = 1, \cdots, S$, for each selected path as

$$p(\tilde{\boldsymbol{h}}|\boldsymbol{\alpha}) = \frac{\prod_{l=1}^{S} \alpha_l}{\pi} \exp(-\tilde{\boldsymbol{h}}^H \operatorname{diag}(\boldsymbol{\alpha})\tilde{\boldsymbol{h}}) \tag{9}$$

where $\operatorname{diag}(\boldsymbol{\alpha})$ stands for a diagonal matrix taking $\boldsymbol{\alpha}$ as the diagonal elements. Obviously, (9) is an over-parameterized model with almost as many observations as parameters to be estimated. So we treat the precision parameter vector $\boldsymbol{\alpha}$ as random variables and impose a Gamma prior distribution to them because Gamma distribution is conjugate to the Gaussian distribution. We have

$$\begin{aligned} p(\boldsymbol{\alpha}) &= \prod_{l=1}^{S} \Gamma(\alpha_l|a, b) \\ &= \prod_{l=1}^{S} \frac{1}{b^a \Gamma(a)} \alpha_l^{a-1} \mathrm{e}^{-\alpha_l/b} \end{aligned} \tag{10}$$

where $\Gamma(\cdot)$ stands for the Gamma distribution. The prior distribution for the inverse of the channel noise variance β is also imposed as a Gamma distribution as

$$\begin{aligned} p(\beta) &= \Gamma(\beta|c, d) \\ &= \frac{1}{d^c \Gamma(c)} \beta^{c-1} \mathrm{e}^{-\beta/d}. \end{aligned} \tag{11}$$

The joint tracking of the channel impulse response and the channel noise variance requires the computation of the posterior distribution

$$p(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta|\boldsymbol{r}) = \frac{p(\boldsymbol{r}|\tilde{\boldsymbol{h}}, \beta)p(\tilde{\boldsymbol{h}}|\boldsymbol{\alpha})p(\boldsymbol{\alpha})p(\beta)}{p(\boldsymbol{r})}. \tag{12}$$

However, the marginal likelihood $p(\boldsymbol{r})$ is analytically intractable here. Thus, following the VB method in Section 4, we assume an approximate posterior distribution $q(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta) \approx p(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta|\boldsymbol{r})$ and apply the mean field approximation to enforce independency between channel $\tilde{\boldsymbol{h}}$ and the variance parameters $\boldsymbol{\alpha}$ and β as

$$q(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta) = q(\tilde{\boldsymbol{h}})q(\boldsymbol{\alpha})q(\beta). \tag{13}$$

From (6), the KL divergence between the posterior distribution $p(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta|\boldsymbol{r})$ and its approximate posterior distribution $q(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta)$ can be written as

$$KL(q||p) = -\int q(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta) \ln \frac{p(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta|\boldsymbol{r})}{q(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta)} \, d\tilde{\boldsymbol{h}} d\boldsymbol{\alpha} d\beta. \tag{14}$$

It can be shown that the minimum of (14) is achieved when $q(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta) = p(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta|\boldsymbol{r})$, which yields the flowing results considering (7), as derived in [9]

$$q(\tilde{\boldsymbol{h}}) = \exp(\langle \ln p(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta|\boldsymbol{r})\rangle_{q(\boldsymbol{\alpha})q(\beta)}) = \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{15}$$

$$q(\boldsymbol{\alpha}) = \exp(\langle \ln p(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta|\boldsymbol{r})\rangle_{q(\tilde{\boldsymbol{h}})q(\beta)}) = \prod_{l=1}^{S} \Gamma(\tilde{a}, \tilde{b}_l) \tag{16}$$

$$q(\beta) = \exp(\langle \ln p(\tilde{\boldsymbol{h}}, \boldsymbol{\alpha}, \beta|\boldsymbol{r})\rangle_{q(\tilde{\boldsymbol{h}})q(\boldsymbol{\alpha})}) = \Gamma(\tilde{c}, \tilde{d}) \tag{17}$$

where $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ stands for a multivariate Gaussian distribution with the mean vector $\boldsymbol{\mu}$ and the covariance matrix $\boldsymbol{\Sigma}$ and

$$\begin{aligned}
\boldsymbol{\Sigma} &= (\bar{\beta}\tilde{\boldsymbol{D}}^H\tilde{\boldsymbol{D}} + \mathrm{diag}(\bar{\boldsymbol{\alpha}}))^{-1} \\
\boldsymbol{\mu} &= \bar{\beta}\boldsymbol{\Sigma}\tilde{\boldsymbol{D}}^H\boldsymbol{r} \\
\tilde{a} &= a + 1 \\
\tilde{b}_l &= b + \overline{|h_l|^2}, \quad l = 1, \cdots, S \\
\tilde{c} &= c + N \\
\tilde{d} &= d + \boldsymbol{r}^H\boldsymbol{r} - \boldsymbol{r}^H\tilde{\boldsymbol{D}}\boldsymbol{\mu} - \boldsymbol{\mu}^H\tilde{\boldsymbol{D}}^H\boldsymbol{r} + \mathrm{tr}(\tilde{\boldsymbol{D}}^H\tilde{\boldsymbol{D}}\boldsymbol{\Sigma}) + \boldsymbol{\mu}^H\tilde{\boldsymbol{D}}^H\tilde{\boldsymbol{D}}\boldsymbol{\mu}.
\end{aligned}$$

Here $\mathrm{tr}(\cdot)$ stands for the matrix trace, $\bar{\beta}$ and $\bar{\boldsymbol{\alpha}}$ denote the mean values of β and $\boldsymbol{\alpha}$, respectively, $|h_l|$ represents the amplitude of h_l, and $\overline{|h_l|^2}$ is the mean value of $|h_l|^2$. By setting the parameters a, b, c, and d to very small values to make prior distributions non-informative, the approximate posterior distributions in (15)-(17) are then iteratively updated until convergence, since they depend on the statistics of each other. The parameter updating procedure is summarized in Table 2, where $(\cdot)$ denotes complex conjugate and $[\boldsymbol{\Sigma}(i)]_{l,l}$ is the (l, l)-th element of $\boldsymbol{\Sigma}(i)$. After the convergence of the VB procedure, the estimated $\tilde{\boldsymbol{h}}$ is given by $\boldsymbol{\mu}$.

6. EXPERIMENT ARRANGEMENT

An UA communication experiment was conducted in December 2012, in the Indian Ocean off the coast of the Rottnest Island, Western Australia (Fig. 3). The average water depth was about 50 meter. As shown in Fig. 4, a single hydrophone at the receiver was attached through a cable at one meter above the seabed. A transducer attached to a drifting vessel through cable for data transmission was located approximately 20 meters below the sea surface. Signals were transmitted when the vessel and transducer were at the positions as denoted by the red dots with labels of T52, T54, T55, T56, T57, T58, T59 , T60, and T61 in Fig. 3, which correspond to 125m, 250m, 500m, 1km, 2km, 4km, 6km,

Table 2: Procedure of the VB algorithm.

Initialization
set a,b,c,d with sufficient small value
set $\bar{\boldsymbol{\alpha}}(0)$ and $\bar{\beta}(0)$
For iteration i
update $\boldsymbol{\Sigma}(i)$ using $\bar{\boldsymbol{\alpha}}(i-1)$ and $\bar{\beta}(i-1)$
update $\boldsymbol{\mu}(i)$ using $\boldsymbol{\Sigma}(i)$ and $\beta(i-1)$
calculate $\overline{|h_l|^2}(i) = [\boldsymbol{\Sigma}(i)]_{l,l} + \mu_l(i)\mu_l^*(i)$, $l = 1, \cdots, S$
update $\tilde{a} = a + 1$
update $\tilde{b}_l(i) = b + \overline{|h_l|^2}(i)$ for $l = 1, \cdots, S$
then update $\bar{\boldsymbol{\alpha}}(i)$
update $\tilde{c} = c + N$
update $\tilde{d}(i)$ using $\boldsymbol{\Sigma}(i)$ and $\boldsymbol{\mu}(i)$
then update $\bar{\beta}(i)$

Figure 3: General location of the experiment environment.

8km, and 10km from the receiver, respectively. It should be noted that both transmitter and receiver were drifting while signals were transmitted, which leads to significant Doppler spreading. GPS data showed that the average drifting speed of the vessel was 0.96 m/s and the peak drift speed was 1.7 m/s when the communication distance is 1 km.

The transmitted signal occupied the frequency band between 10 kHz and 14 kHz and the system sampling rate was 96 kHz. Each training sequence involved 511 symbols while the data sequence included 2046 symbols. BPSK modulated pseudo random sequence was used for the training sequence. For data sequences, 8PSK and QPSK signals were transmitted at ranges of 125m, 250m, 500m, 1km, 2km, and 4km. QPSK and BPSK signals were transmitted at the ranges of 6km and 8km.

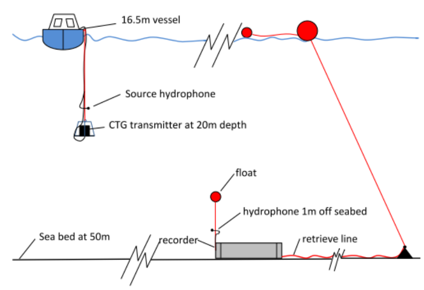

Figure 4: Transmitter and receiver diagram.

Table 3: Average uncoded BER using the LS equalization.

Distance	1km	2km	4km	6km	8km
Trad. OMP	17.3%	29.4%	27.6%	12.7%	29.8%
VB-based OMP	13.1%	28.3%	27.9%	12.3%	29.7%
OMP-based VB	10.9%	25.9%	25.3%	9.4%	26.2%

Table 4: Average uncoded BER using the MMSE equalization.

Distance	1km	2km	4km	6km	8km
Trad. OMP	7.0%	14.7%	10.9%	2.8%	10.1%
VB-based OMP	6.5%	13.9%	10.7%	2.7%	10%
OMP-based VB	6.0%	12.2%	9.7%	2.6%	9.3%

7. EXPERIMENT RESULTS

In this section, we study the BER performance of the experimental system. The channel estimation algorithm during the channel training period is developed assuming that the UA channel has 75 taps. While during the data transmission stage, the channel tracking method uses the information of the first 25 taps, as the energy of the remaining 50 channel taps is sufficiently small to be neglected. Table 3 shows the uncoded system BER when the LS method is applied for channel equalization, while Table 4 demonstrates the system BER results using the MMSE-based channel equalization.

It can be seen from Tables 3 and 4 that both proposed channel tracking algorithms almost outperform the traditional OMP algorithm. When the LS-based channel equalization is used, the OMP-based VB channel tracking algorithm reduces the system BER by over 6% for the 1km range, and around 3% for the 2km–8km ranges. With the MMSE-based channel equalization, the OMP-based VB algorithm reduces the BER by about 1%. The VB-based OMP method slightly increases the system BER by 0.3% at the 4km range with the LS equalization method, which is the only case where the new algorithm results in a slightly worse performance during our experiments. This is because the BER performance may also be affected by many other factors, such as frequency offset and additive noise.

We can also observe from Tables 3 and 4 that for both the LS and MMSE equalization cases, the OMP-based VB method yields a larger BER improvement than the VB-based OMP method. Moreover, both proposed algorithms have a larger BER improvement when the LS-based equalization is used compared with the case of the MMSE equalizer. For instance, when the LS equalization method is used, the VB-based OMP method reduces the system BER by around 4% at the 1km range, while the OMP-based VB method reduces the BER by 6%. However, when the MMSE equalizer is adopted, the BER benefit of the VB-based OMP method is only 0.5%, while the advantage of the OMP-based VB method is only 1% for the 1km range. The reason is that, in general, detected symbols with higher BER are more likely to cause paths positions detecting error during paths searching stage of the OMP algorithm. Thus, the proposed algorithms can bring more benefit in the cases where the system already has a low BER than the cases with high BER.

8. CONCLUSIONS

Two novel joint channel tracking and noise variance estimation algorithms, which combine the advantages of the VB method and the CS technique, are developed for single-carrier underwater acoustic communication system in this paper. It is shown by our experiment that the VB-based OMP method overcomes the over-fitting of the conventional OMP method, and the OMP-based VB method can also avoid the detecting error of channel paths positions, resulting in a better BER performance. The results show that the proposed algorithms can reduce the uncoded system BER by 1% to 6%.

Acknowledgment

This research was supported under the Australian Research Council's Discovery Projects funding scheme (project number DP140102131).

9. REFERENCES

[1] D. B. Kilfoyle and A. B. Baggeroer, The state of the art in underwater acoustic telemetry, *IEEE J. Ocean. Eng.*, vol. 25, no. 1, pp. 4-27, Jan. 2000.

[2] T. H. Eggen, A. B. Baggeroer, and J. C. Preisig, Communication over Doppler spread channels – Part I: Channel and receiver presentation, *IEEE J. Ocean. Eng.*, vol. 25, no. 1, pp. 62-71, Jan. 2000.

[3] S. Zhou and Z. Wang, *OFDM for Underwater Acoustic Communications.* John Wiley & Sons, June 2014.

[4] M. Stojanovic and J. Preisig, Underwater acoustic communication channels: Propagation models and statistical characterization, *IEEE Commun. Mag*, vol. 47, no. 1, pp. 84-89, Jan. 2009.

[5] M. Stojanovic, L. Freitag, and M. Johnson, Channel-estimation-based adaptive equalization of underwater acoustic signals, in *Proc. MTS/IEEE OCEANS 1999*, , vol. 2, pp. 590-595.

[6] W. Li and J. C. Preisig, Estimation of rapidly time-varying sparse channels, *IEEE J. Ocean. Eng.*, vol. 32, no. 4, pp. 927-939, Oct. 2007.

[7] S. F. Cotter and B. D. Rao, Sparse channel estimation via matching pursuit with application to equalization, *IEEE Trans. Commun.*, vol. 50, no. 3, pp. 374-377, Mar. 2002.

[8] G. Z. Karabulut and A. Yongacoglu, Sparse channel estimation using orthogonal matching pursuit algorithm, in *Proc. IEEE Veh. Technol. Conf.*, (Los Angeles, CA, 2004), ,vol. 6, pp.3880 -3884.

[9] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, The variational approximation for Bayesian inference, *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 131-146, Nov. 2008.

[10] J. Zhang and Y. R. Zheng, Bandwidth-efficient MIMO underwater acoustic communications with frequency-domain equalization, in *Proc. MTS/IEEE OCEANS 2010*, (Sydney, Australia, 2010), , pp. 1-5.

[11] S. Guo, Z. He, W. Jiang, Y. Ou, K. Niu, Y. Rong, M. Caley, and A. J. Duncan, Channel estimation based on compressed sensing in high-speed underwater acoustic communication, in *Proc. 9th Int. Conf. Inform. Commun. Signal Process.*, (Tainan, Taiwan, 2013), , pp. 1-5.