Variational Soft Symbol Decoding for Sweep Spread Carrier Based Underwater Acoustic Communications

Arunkumar K.P. and Chandra R. Murthy

***Abstract*—We present a scheme for data detection in sweep spread carrier (S2C) based point-to-point communication over doubly-spread underwater acoustic (UWA) channels. The existing schemes for data detection – based on the gradient heterodyne receiver – are only effective when the path delay and Doppler spread are moderate. Based on the principle of variational Bayes' inference, we develop a new variational soft symbol decoding (VSSD) algorithm for a general linear channel model. We show that, in benign underwater channels with moderate delay and Doppler spreads, the VSSD algorithm is equivalent to the existing receivers for S2C communications. We apply the VSSD algorithm to the i.i.d. Gaussian multiple-input multiple-output (MIMO) channel and show, through numerical simulations, that it far outperforms the minimum mean squared error (MMSE) data detection in both uncoded and coded communications. Further, even with channel estimation errors, the VSSD retains its performance advantage over the MMSE receiver. In UWA channels where the existing S2C receivers completely fail, or must compromise on the data rate to maintain the bit error rate (BER) performance, the proposed VSSD algorithm recovers the data symbols at a signal-to-noise ratio (SNR) which is at least 10 dB (8 dB) lower than the MMSE decoder for uncoded (rate 2/3 LDPC coded) communications.**

***Index Terms*—Underwater acoustic communications, sweep spread carrier communication, variational Bayes.**

I. Introduction

The underwater acoustic (UWA) environment presents a formidable challenge for wireless communication due to large delay spreads and path-dependent Doppler shifts. Multipath propagation of sound results in a delay spread which is in the order of tens of milliseconds [2]. Time variations cause path-dependent Doppler shifts that are non-uniform over the bandwidth of the acoustic signal. Moreover, the sensor nodes used in an underwater sensor network comprising stationary or mobile underwater and surface vehicles are usually battery operated, and are therefore highly constrained on the amount of transmission power. High performance receiver processing, that recover data symbols at a low signal-to-noise ratio, is highly desirable in these applications.

Sweep spread carrier (S2C) communications, developed in [3] and used in a practical UWA modem [4], is well suited for battery operated modems due to its ideal peak-to-average power ratio (PAPR) for the transmitted waveform. The details of the S2C transmitter and receiver side processing, performance analysis, and experimental results are found in [5]–[9]. The S2C receiver in [3] relies on extracting the symbol arriving along the direct path. Consequently, the part of the transmitted symbol energy arriving along paths other than the direct path is not exploited. In [10], authors use a maximum ratio combiner (MRC) to leverage the multipath diversity and hence improve the performance of an S2C receiver. The receiver in [10] works well only when (a) the ratio of the maximum delay spread to minimum differential delay among path arrivals is below a certain value, and (b) the Doppler spread is small. If either condition is violated, the intersymbol interference (ISI) cancellation becomes imperfect and MRC becomes suboptimal and ineffective.

In this paper, we consider an S2C communication system similar to [3] and [10] but for a more general underwater channel model. We tie the existing S2C receivers – the gradient heterodyne (GradH) receiver and the path-matched gradient heterodyne (pGradH) receiver – that were heuristically motivated in [3] and [10], to the minimum mean squared error (MMSE) decoder for certain benign UWA channel models. In more severe channels, the performance of these detectors fall short of the MMSE decoder designed for such channels.

Several works have considered the MMSE equalizer for (hard) data symbol detection or joint channel estimation and data detection in orthogonal frequency division multiplex (OFDM) and code division multiple access (CDMA) based UWA communications [11]–[15]. However, in coded communications, it is more important to estimate the soft symbols rather than hard symbols. The variational Bayes' (VB) inference is a promising approach to obtain soft symbol estimates because it approximates the maximum a posteriori (MAP) decoder in an iterative manner. However, to the best of our knowledge, VB based soft symbol estimation has not been explored in the literature. We, therefore, develop a *new soft symbol decoder* based on the principle of variational Bayes' inference. The soft symbol estimate is found by solving a set of fixed points equations. We show that the fixed point iterations converge to a local minimum in the general case, and to a global minimum for orthogonal channel matrices whose important special cases are the AWGN and Rayleigh channels. Finally, we present the simulation results that evince the strong performance of the proposed soft symbol decoder in harsh channel conditions.

Arunkumar K. P. is with the Naval Physical Oceanographic Laboratory, Kochi 682 021, India (e-mail: arunkumar@npol.drdo.in). This work was carried out as part of the PhD program at Indian Institute of Science, Bangalore. A part of this work was published in [1]. Chandra R. Murthy is with the Dept. of ECE, Indian Institute of Science, Bangalore 560 012, India (e-mail: cmurthy@iisc.ac.in).

II. System Model

We consider an S2C system as in [3] and [10]. At the transmitter side, a succession of linear frequency modulated chirp pulses, each swept from a lower frequency limit f_L to an upper frequency limit f_H over a sweep duration T_{sw}, is employed as the carrier waveform:

$$c(t) = e^{j2\pi\left(f_L t_r(t) + m_c t_r(t)^2\right)}, 0 \le t \le T_c, \tag{1}$$

where

$$t_r(t) = t - \left\lfloor \frac{t}{T_{\text{sw}}} \right\rfloor T_{\text{sw}}, \tag{2}$$

is a periodic ramp function with period T_{sw}, $2m_c = \frac{f_H - f_L}{T_{\text{sw}}}$ is the chirp rate, $T_c = N_c T_{\text{sw}}$ is the overall carrier duration, and N_c is the number of chirp pulses comprising the carrier waveform.

The message signal, containing pilot and data symbols from a quadrature phase shift key (QPSK) constellation, is represented by:

$$s(t) = \sum_{k=0}^{N-1} s_k g(t - kT), \tag{3}$$

where $s_k, k = 0, \ldots, N-1$, are a sequence of QPSK symbols, T is the symbol duration, $N = \frac{T_c}{T}$ is the number of symbols mounted on the entire carrier waveform and $g(t)$ is a pulse shaping function, for example, a root-raised-cosine pulse with roll-off factor α. We denote the symbol bandwidth by B, and is given by $B \approxeq \frac{1+\alpha}{T}$. We also assume that T_{sw} is a multiple of T so that there are $M = \frac{T_{\text{sw}}}{T}$ symbols in every chirp pulse within a carrier waveform. Note that $N = MN_c$.

The modulated signal, to be transmitted, is given by:

$$x(t) = \text{Re}\left[s(t)c(t)\right], \tag{4}$$

which is prefixed with a preamble pulse and appended with a post-amble pulse to form a transmission frame. The preamble and post-amble are used for timing and synchronization, and for estimating the channel. The preamble (post-amble) is a chirp pulse modulated by a message signal made up of known pilot symbols. A guard interval of T_g is used after (before) the preamble (post-amble) pulse to allow the estimation of channel. Using $N_c > 1$ helps with amortizing the overhead due to guard interval over the total duration T_c.

The time-varying channel impulse response of the underwater acoustic channel is modeled as [16]:

$$h(t, \tau) = \sum_{p=0}^{N_p - 1} h_p(t) \delta\left(\tau - \tau_p(t)\right), \tag{5}$$

where $h_p(t)$ and $\tau_p(t)$ are the time-varying amplitude and delay, respectively, of the pth path, and N_p is the number of significant paths in the channel. As in [11], we assume that, the path amplitudes are constant within a data packet, that is, $h_p(t) = h_p$, and that the time variation of the path delays due to Doppler rate a_p can be approximated as

$$\tau_p(t) = \tau_p - a_p t. \tag{6}$$

The received signal after timing, synchronization, coarse Doppler scale estimation and compensation, is given by:

$$y(t) = \sum_{p=0}^{N_p - 1} y_p(t) + w(t), \tag{7}$$

where $w(t)$ is the additive white Gaussian noise (AWGN), $y_p(t) = h_p \mathcal{R}\{s(\tilde{t} - \tau_p(\tilde{t}))c(\tilde{t} - \tau_p(\tilde{t}))\}$ is the timing adjusted and Doppler compensated version of the S2C signal reaching via the pth path, $\tilde{t} = \frac{t}{1+\hat{a}} + \hat{\tau}$ is the shifted and rescaled time-axis, $\hat{\tau}$ is the starting time instance of the first (data) chirp pulse estimated from the preamble/post-amble as in [3], $\hat{a}$ is the coarse Doppler scale estimated using both the preamble and post-amble as in [16]. Using (6), we can write,

$$y_p(t) = h_p \sum_{k=0}^{N-1} \left(s_{k,\text{Re}} \cos\phi_p(t) - s_{k,\text{Im}} \sin\phi_p(t)\right) \times g\left(\overline{1+b_p}t - \tilde{\tau}_p - kT\right), \tag{8}$$

where $\tilde{\tau}_p = \tau_p - \hat{\tau}$ and $b_p = \frac{a_p - \hat{a}}{1+\hat{a}}$ are the residual delay and Doppler scale, respectively, of the pth path after compensation, $s_{k,\text{Re}}$ ($s_{k,\text{Im}}$) is the real (imaginary) part of the QPSK symbol s_k and,

$$\phi_p(t) \triangleq 2\pi\left(f_L t_r(\overline{1+b_p}t - \tilde{\tau}_p) + m_c t_r^2(\overline{1+b_p}t - \tilde{\tau}_p)\right). \tag{9}$$

Upon sampling at a rate F_s ($= 1/T_s$, where T_s is the sampling period), we may re-express the received signal in (7) in a vector form relevant to data detection, as:

$$\mathbf{y} = H\mathbf{s} + \mathbf{w}, \tag{10}$$

where,

$$\begin{aligned} H &= [C_0\mathbf{h}, -S_0\mathbf{h}, \ldots, C_{N-1}\mathbf{h}, -S_{N-1}\mathbf{h}] \in \mathbb{R}^{NL \times 2N}, \\ \mathbf{h} &= [h_0, h_1, \ldots, h_{N_p-1}]^T \in \mathbb{R}^{N_p \times 1}, \\ \mathbf{s} &= [s_{0,\text{Re}}, s_{0,\text{Im}}, \ldots, s_{N-1,\text{Re}}, s_{N-1,\text{Im}}]^T \in \mathbb{R}^{2N \times 1}, \\ \mathbf{w} &\sim \mathcal{N}\left(\mathbf{0}, \sigma^2 I_{2N}\right), \end{aligned}$$

$L = \lfloor F_s T \rfloor$ is the number of samples in the symbol duration, and $C_k \in \mathbb{R}^{NL \times N_p}$ and $S_k \in \mathbb{R}^{NL \times N_p}$ are matrices whose entries are given by:

$$\begin{aligned} C_k(l,p) &= \cos\phi_p(lT_s) g\left(\overline{1+b_p}lT_s - \tilde{\tau}_p - kT\right), \\ S_k(l,p) &= \sin\phi_p(lT_s) g\left(\overline{1+b_p}lT_s - \tilde{\tau}_p - kT\right), \end{aligned}$$

for $0 \le k \le N-1$, $0 \le l \le NL-1$ and $0 \le p \le N_p - 1$. Since $g(t) = 0, t \notin [0, T]$, entries of $C_k(:, p) \in \mathbb{R}^{NL \times 1}$ and $S_k(:, p) \in \mathbb{R}^{NL \times 1}$ are zeros except for $l \in \left\{\left\lceil \frac{\tilde{\tau}_p + kT}{1+b_p T_s} \right\rceil, \ldots, \left\lfloor \frac{\tilde{\tau}_p + \overline{k+1}T}{1+b_p T_s} \right\rfloor\right\}$.

We now address the problem of data detection for the S2C communication model. First, we examine the two existing S2C receivers in the literature – the GradH receiver, pioneered in [3], and the pGradH receiver proposed in [10].

III. Existing S2C Receivers: GradH and pGradH

We show that the GradH and pGradH based S2C receivers are minimum mean square error (MMSE) symbol detectors for the AWGN channel and a delay spread channel with well resolved path delays, respectively.

A. Optimality of GradH Receiver

Consider the received signal for the AWGN channel ($N_p = 1, \tilde{\tau}_0 = 0, b_0 = 0, h_0 = 1$) which is given by:

$$y(t) = \sum_{i=0}^{N-1} \left(s_{k,\text{Re}} \cos\phi_0(t) - s_{k,\text{Im}} \sin\phi_0(t)\right) g\left(t - kT\right) + w(t), \quad (11)$$

where $\phi_0(t) \triangleq 2\pi\left(f_L t_r(t) + m_c t_r^2(t)\right)$. Upon sampling, the received signal is as in (7) with the channel matrix taking the specific form:

$$H = Q = \begin{bmatrix} Q_0 & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & Q_1 & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & Q_{N-1} \end{bmatrix} \in \mathbb{R}^{NL \times 2N},$$

where,

$$Q_k = \text{diag}\left(\mathbf{g}\right) \begin{bmatrix} \cos\phi_0^{(k)}[0] & \sin\phi_0^{(k)}[0] \\ \cos\phi_0^{(k)}[1] & \sin\phi_0^{(k)}[1] \\ \vdots & \vdots \\ \cos\phi_0^{(k)}[L-1] & \sin\phi_0^{(k)}[L-1] \end{bmatrix} \in \mathbb{R}^{L \times 2},$$

$\mathbf{g} \in \mathbb{R}^L$ has entries that are samples of the pulse shaping function, $g_l = g(lT_s)$, and $\phi_0^{(k)}[l] = \phi_0\left((\tilde{k}-1)T + lT_s\right)$, $\tilde{k} = k - \lfloor \frac{k}{M} \rfloor M$, $l = 0, \dots, L-1$, $k = 0, 1, \dots, N-1$, and $\phi_0(t) = 2\pi\left(f_L t_r(t) + m_c t_r^2(t)\right)$. In this case, there is no inter-symbol interference (ISI), and the measurement corresponding to the kth symbol is given by:

$$\mathbf{y}_k = Q_k \mathbf{s}_k + \mathbf{w}_k, \quad (12)$$

where,

$$\begin{aligned} \mathbf{y}_k &= \left[\, y[(k-1)L], y[(k-1)L+1], \dots, y[kL-1] \,\right]^T, \\ \mathbf{s}_k &= \left[s_{k,\text{Re}}, s_{k,\text{Im}}\right]^T \in \{\left[\pm 1/\sqrt{2}, \pm 1/\sqrt{2}\right]^T\}, \\ \mathbf{w}_k &= \left[w_k[0], \dots, w_k[L-1]\right]^T \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 I_L\right), \end{aligned}$$

for $k = 0, \dots, N-1$.

For equiprobable symbols $\mathbf{s}_k$, the maximum a posteriori probability (MAP) solution to (12), which minimizes the probability of symbol error, is the maximum likelihood (ML) solution given by:

$$\hat{\mathbf{s}}_k^{(\text{ML})} = \arg\min_{\mathbf{s}_k \in \{\left[\pm 1/\sqrt{2}, \pm 1/\sqrt{2}\right]^T\}} \|\mathbf{y}_k - Q_k \mathbf{s}_k\|_2. \quad (13)$$

The MMSE solution to (12) is given by:

$$\hat{\mathbf{s}}_k^{(\text{MMSE})} = \mathcal{S}\left[\left(Q_k^T Q_k + \sigma^2 I_2\right)^{-1} Q_k^T \mathbf{y}_k\right], \quad (14)$$

where $\mathcal{S}[.]$ is the slicing operation that quantizes each entry of its argument vector to the nearest symbol in the QPSK constellation. For large enough symbol time, T (and hence L), and smoothly varying pulse shaping function, $g(t)$, such that $\sum_{l=0}^{L-1} g^2(lT_s)\cos^2\left(\phi_0^{(k)}[l]\right) \approxeq \sum_{l=0}^{L-1} g^2(lT_s)\sin^2\left(\phi_0^{(k)}[l]\right) = \beta$ (say) which is nearly the same for all $0 \le k \le N-1$, we have $Q_k^T Q_k \approx \beta I_2$. Under these conditions, the MMSE receiver in (14) simplifies to the GradH receiver in [3]:

$$\hat{\mathbf{s}}_k^{(\text{GradH})} = \mathcal{S}\left[\mathbf{z}_k\right], \quad (15)$$

where $\mathbf{z}_k = Q_k^T \mathbf{y}_k$. From (12), we see that $\mathbf{z}_k \approx \beta \mathbf{s}_k + \mathbf{v}_k$, where $\mathbf{v}_k = Q_k^T \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \beta\sigma^2 I_2)$, is affected only by the kth symbol. Notice that $\mathbf{z}_k$ is a sub-vector of $\mathbf{z} = Q^T \mathbf{y} \in \mathbb{R}^{2N \times 1}$, whose entries are precisely the sampled versions of the low-pass filtered in-phase and quadrature outputs of gradient heterodyne operation on the received signal [3].

The measurement model at the output of GradH preprocessing, i.e., gradient heterodyne operation and low pass filtering is given by:

$$\mathbf{z} = G\mathbf{s} + \mathbf{v}, \quad (16)$$

where $G = Q^T H \in \mathbb{R}^{2N \times 2N}$ is the channel matrix at the output of GradH preprocessor and $\mathbf{v} = Q^T \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 Q^T Q)$. In the case of AWGN channel and large symbol duration, $G = \beta I_{2N}$ and $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \beta\sigma^2 I_{2N})$.

While the GradH receiver in (15) is an MMSE symbol detector for the AWGN channel, the receiver works reasonably well even for ISI channels with *moderate* delay spreads, as elaborated in [3]. The GradH receiver recovers the symbol arriving along the *direct path* when [10],

$$\frac{\mathcal{M}}{\mathcal{M}-1}\delta\tau_{\max} \le T_{sw} \le \mathcal{M}\delta\tau_{\min}, \quad (17)$$

where $\delta\tau_{\min} = \min_{0 \le i < j \le N_p - 1} |\tau_i - \tau_j|$ and $\delta\tau_{\max} = \max_{0 \le i,j \le N_p - 1} |\tau_i - \tau_j|$ are the smallest and largest separation between any two path arrival times τ_i and τ_j, and $\mathcal{M} \triangleq \frac{f_H - f_L}{B}$ is called the spreading factor.

B. Optimality of pGradH Receiver

The pGradH receiver in [10] additionally combines the symbol arriving along paths other than the direct path to leverage the multipath diversity. We briefly show that pGradH is the MMSE receiver when the path delays are well resolved and condition (17) holds. The MMSE receiver is given by:

$$\hat{\mathbf{s}}^{(\text{MMSE})} = \mathcal{S}\left[\left(H^T H + \sigma^2 I_{2N}\right)^{-1} H^T \mathbf{y}\right]. \quad (18)$$

When the path delays $\tau_p, p = 0, 1, \dots, N_p - 1$, are well resolved and condition (17) holds, then $C_i^T C_j \approx \kappa_C I_{N_p} \delta_{i,j}$, where $\kappa_C = C_i(:,p)^T C_i(:,p)$ is nearly the same for all $0 \le i \le N-1$ and $0 \le p \le N_p - 1$, and $\delta_{i,j}$ is the Kronecker-delta function. Similarly, $S_i^T S_j \approx \kappa_S I_{N_p} \delta_{i,j}$, where $\kappa_S = S_i(:,p)^T S_i(:,p)$, and $C_i^T S_j \approx 0$. Under these conditions, the

MMSE receiver in (18) simplifies to the pGradH receiver in [10]:

$$\hat{\mathbf{s}}_k^{(\text{pGradH})} = \mathcal{S}\left[\sum_{p=0}^{N_p-1} \frac{h_p}{|h_p|^2}\mathbf{z}_k^{(p)}\right], \tag{19}$$

where,

$$\mathbf{z}_k^{(p)} = Q_k^{(p)T}\mathbf{y}_k, \tag{20}$$

$$Q_k^{(p)} = \text{diag}\left(\mathbf{g}^{(p)}\right)\begin{bmatrix} \cos\phi_p^{(k)}[0] & \sin\phi_p^{(k)}[0] \\ \cos\phi_p^{(k)}[1] & \sin\phi_p^{(k)}[1] \\ \vdots & \vdots \\ \cos\phi_p^{(k)}[L-1] & \sin\phi_p^{(k)}[L-1] \end{bmatrix},$$

$\mathbf{g}^{(p)} \in \mathbb{R}^L$ has entries that are samples of the compressed/dilated and delayed pulse shaping function, $g_l^{(p)} = g\left(\overline{1+b_p}lT_s - \tilde{\tau}_p\right)$, $\phi_p^{(k)}[l] = \phi_p\left((\tilde{k}-1)T + lT_s\right)$, $\tilde{k} = k - \lfloor\frac{k}{M}\rfloor M$, $l = 0,\ldots,L-1$, $p = 0,\ldots,N_p-1$, and $k = 0,1,\ldots,N-1$. Stacking up $\mathbf{z}_k^{(p)}, k = 0,1,\ldots,N-1$, into a vector, we get:

$$\mathbf{z}^{(p)} = Q^{(p)T}\mathbf{y} \in \mathbb{R}^{2N\times 1}, \tag{21}$$

where,

$$Q^{(p)} = \begin{bmatrix} Q_0^{(p)} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & Q_1^{(p)} & \mathbf{0} & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & Q_{N-1}^{(p)} \end{bmatrix} \in \mathbb{R}^{NL\times 2N}.$$

The entries of $\mathbf{z}^{(p)}$ are precisely the sampled versions of the low-pass filtered in-phase and quadrature outputs of *path-matched* gradient heterodyne operation on the received signal [10].

The measurement model at the output of pGradH preprocessing, i.e., path-matched gradient heterodyne operation followed by low-pass filtering of the received signal, is given by:

$$\mathbf{z}^{(p)} = G^{(p)}\mathbf{s} + \mathbf{v}^{(p)}, p = 0,\ldots,N_p-1, \tag{22}$$

where $G^{(p)} = Q^{(p)T}H \in \mathbb{R}^{2N\times 2N}$ is the channel matrix at the output of MRC preprocessor and $\mathbf{v}^{(p)} = Q^{(p)T}\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 Q^{(p)T}Q^{(p)})$. For a moderately delay spread channel with well resolved path delays and large symbol duration, $G^{(p)} = \beta h_p I_{2N}$ and $\mathbf{v}^{(p)} \sim \mathcal{N}(\mathbf{0}, \beta\sigma^2 I_{2N})$. By stacking up $\mathbf{z}^{(p)}, p = 0,1,\ldots,N_p-1$, into a vector $\mathbf{z} \in \mathbb{R}^{2NN_p}$, we get the same form of measurement model as in (16) where the channel matrix G is obtained by vertically stacking $G^{(p)}$.

The channel matrix, G, at the output of the GradH and pGradH preprocessors are diagonal for a moderately delay spread channel with well resolved path delays, thanks to the condition in (17) that makes $Q^{(p)T}Q^{(q)} \approx \beta\delta_{p,q}I_{2N}, 0 \leq p,q \leq N_p-1$. In the next subsection, we bring out the need to consider alternate S2C receiver processing in large delay spread channels.

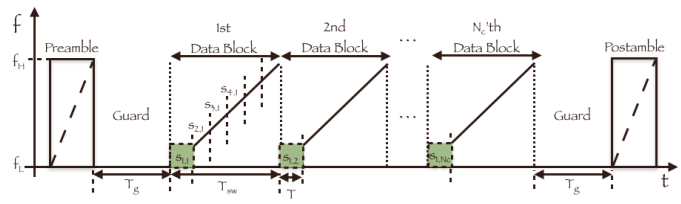

Fig. 1. An S2C frame consisting of preamble, N_c chirp pulses (data blocks), and post-amble. Choosing $T \geq \frac{1}{\sqrt{2m_c}}$ avoids ISI among adjacent symbols. But, inter-block interference (IBI) among the symbols mounted on the same frequency sweep slots can happen over a duration corresponding to channel delay spread.

C. Limitations Of GradH and pGradH Receivers

For both GradH and pGradH receivers, the condition in (17) is needed to ensure a negligible ISI after gradient heterodyne operation and low pass filtering. The condition (17) places a *lower limit* on the minimum differential path delay, $\delta\tau_{\min}$, of the multipath arrivals to avoid ISI ensuing from the mixing of adjacent symbols at the GradH and pGradH preprocessor outputs [10]. The condition (17) also places an *upper limit* on the channel delay spread, $\delta\tau_{\max}$, to avoid interference between the symbols on the corresponding frequency sweep slots of different chirp pulses. Together, these limits require the symbol rate, $R = 1/T$, of the existing S2C receivers to satisfy:

$$R \leq \left(\frac{f_H - f_L}{1+\alpha}\right)\frac{\min\{\delta\tau_{\min}, T_{\text{sw}} - \delta\tau_{\max}\}}{T_{\text{sw}}}. \tag{23}$$

The upper limit on the achievable rate, in (23), is maximized when $T_{\text{sw}} = \delta\tau_{\max} + \delta\tau_{\min}$, and the maximum rate achievable by the existing S2C receivers is given by:

$$R_{\max} = \left(\frac{f_H - f_L}{1+\alpha}\right)\frac{\delta\tau_{\min}}{\delta\tau_{\max} + \delta\tau_{\min}}. \tag{24}$$

Note that the rate limiting condition $R \leq R_{\max}$, to avoid ISI at the preprocessor output of the existing S2C receivers, is equivalent to the condition on the spreading factor: $\mathcal{M} \geq \frac{\delta\tau_{\max}}{\delta\tau_{\min}} + 1$. When the system is operated at a symbol rate $R = R_{\max}$, the spreading factor $\mathcal{M} = \frac{\delta\tau_{\max}}{\delta\tau_{\min}} + 1$.

Existing S2C receivers entail ISI when operating at a symbol rate greater than $R_{\max}$ in a channel with a minimum differential delay of $\delta\tau_{\min}$ and delay spread of $\delta\tau_{\max}$. Consider, for example, the S2C system in Table I operating in a UWA channel with a maximum delay spread of $\delta\tau_{\max} = 25$ ms. There are 20 QPSK symbols (i.e., 40 bits) in one chirp pulse (S2C block) of duration $T_{\text{sw}} = 10$ ms. Figure 1 shows a transmitted S2C frame, where the symbols $s_{i,j}$ and $s_{i,j+1}$ can potentially interfere with the detection of the symbol $s_{i,j+2}, j = 1,2,3$. Figure 2 and 3 show the images of the channel matrix H, in (10), and the effective channel matrix G, in (16), for such a channel. While the gradient heterodyne and low pass filtering operation has significantly reduced ISI, strong residual inter-block interference remains at the GradH/pGradH preprocessor output.

In the following section, we consider alternate receivers for S2C communications that can handle channel delay spreads

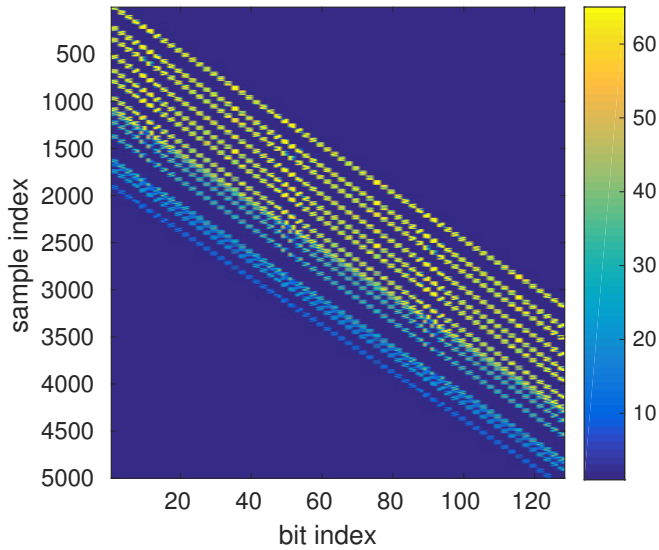

Fig. 2. Channel matrix image, $|H| \in \mathbb{R}^{NL \times 2N}$, before GradH processing.

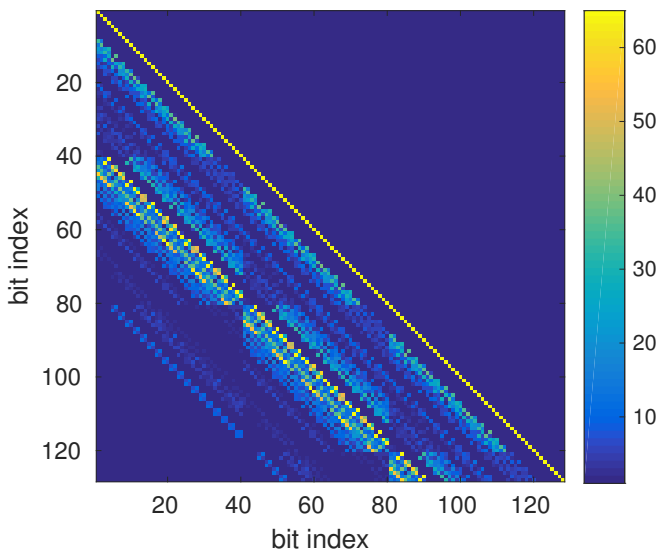

Fig. 3. Channel matrix image, $|G| \in \mathbb{R}^{2N \times 2N}$, after GradH processing.

greater than the chirp pulse duration and work well for symbol rates higher than the upper limit, $R_{\max}$, on the data rate in the existing receivers.

IV. Improved S2C Receivers

A. MMSE Based S2C Receiver

A solution to the symbol detection problem for S2C communications over fairly large delay and Doppler spread UWA channel model, as in (10), is to use the MMSE receiver in (18). Alternatively, we could use the MMSE receiver at the output of the GradH/pGradH preprocessor. The MMSE receiver for uncoded communication is given by:

$$\hat{\mathbf{s}}^{(\text{MMSE})} = \mathcal{S}\left[\left(G^T G + \sigma^2 I_{2N}\right)^{-1} G^T \mathbf{z}\right]. \tag{25}$$

In a coded communication, instead of slicing the MMSE equalizer output, $\left(G^T G + \sigma^2 I_{2N}\right)^{-1} G^T \mathbf{z}$, it is converted to log-likelihood ratio (LLR) vector and fed to the channel decoder that decides the final hard symbol vector.

Next, we develop a symbol detector based on the variational Bayes inference that approximates the optimum MAP decoder and offers a significantly improved performance over the MMSE receiver. The development of the VSSD is the main contribution of this work.

B. Variational Soft Symbol Decoder

The optimum (MAP) decoder is the symbol vector $\mathbf{s} \in \mathcal{P} = \{-\frac{1}{\sqrt{2}}, +\frac{1}{\sqrt{2}}\}^{2N}$ that maximizes the posterior $p(\mathbf{s}|G, \mathbf{z}) = p(\mathbf{z}|G, \mathbf{s})p(\mathbf{s})/p(\mathbf{z}|G)$. Direct maximization of the posterior requires a computationally intensive search over 2^{2N} lattice points in $\mathcal{P}$. Computing the posterior, that represents the soft symbol vector, is also hard since the marginalization over $\mathbf{s}$ in $p(\mathbf{z}|G) = \sum_{\mathbf{s}\in\mathcal{P}} p(\mathbf{s}, \mathbf{z}|G)$ is involved. We instead seek a good approximation to the posterior, $q_\phi(\mathbf{s}|G, \mathbf{z})$, called the variational decoder. Here, ϕ represents the model parameters, whose values are estimated based on the variational principle in order to maximize the evidence lower bound, as explained below.

To make the problem tractable, we assume that the approximate posterior is fully factorizable:

$$q_\phi(\mathbf{s}|G, \mathbf{z}) = \prod_{k=0}^{N-1} q_\phi(s_{k,\text{Re}}|G, \mathbf{z}) q_\phi(s_{k,\text{Im}}|G, \mathbf{z}). \tag{26}$$

Following Kingma et al. [17], the evidence lower bound (ELBO) on the log likelihood of observation is given by:

$$\mathcal{L}(\theta, \phi, \mathbf{z}) = \mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}} \log p_\theta(\mathbf{z}|G, \mathbf{s}) - \mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}} \left[\log \frac{q_\phi(\mathbf{s}|G, \mathbf{z})}{p_\theta(\mathbf{s})}\right], \tag{27}$$

where $\log p_\theta(\mathbf{z}|G, \mathbf{s})$ is the likelihood function and $p_\theta(\mathbf{s})$ is a prior on the symbol vector.

In order to bring $q_\phi(\mathbf{s}|G, \mathbf{z})$ close to $p(\mathbf{s}|G, \mathbf{z})$, we maximize the ELBO, $\mathcal{L}(\theta, \phi, \mathbf{z})$. The ELBO consists of the likelihood term,

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}} \log p_\theta(\mathbf{z}|G, \mathbf{s}) = -N \log(2\pi\sigma^2) - \mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}} \left[\frac{\|\mathbf{z} - G\mathbf{s}\|^2}{2\sigma^2}\right], \tag{28}$$

and the regularizing term,

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}} \left[\log \frac{q_\phi(\mathbf{s}|G, \mathbf{z})}{p_\theta(\mathbf{s})}\right] = KL(q_\phi || p_\theta). \tag{29}$$

We assume a simple uniform prior $p_\theta(\mathbf{s}) = \frac{1}{2^{2N}}$. Therefore, when maximizing ELBO, the regularizing term acts to penalize the departure of the variational approximation q_ϕ from the uniform prior. On maximizing the ELBO, we get the following fixed point equations (see appendix for details):

$$\mathbf{q} = \varphi(\boldsymbol{\alpha}), \tag{30}$$

where

$$\alpha_j = \frac{\sqrt{2}}{\sigma^2}\left(\mathbf{z}^T G_{:,j} - \sum_{l=0}^{2N-1} G_{l,j}\left(\sum_i v_{l,i} - v_{l,j}\right)\right), \tag{31}$$

$$v_{l,j} = \frac{1}{\sqrt{2}} G_{l,j}\left(2q_j - 1\right), \tag{32}$$

$$\varphi(\alpha_j) = \frac{1}{1+e^{-\alpha_j}}, \tag{33}$$

for $j = 1, \ldots, 2N-1$.

Note that the fixed point iterations lead to soft symbol estimates in the form of the probability vector $\mathbf{q}$. We perform symbol detection by slicing the probability vector in uncoded communications. In coded communication, the soft symbols are converted to LLRs and fed to the channel decoder for deciding the hard symbol vector.

The fixed point updates do not involve any matrix inversion and their computational complexity, $\mathcal{O}(N^2)$, is an order of magnitude smaller than the computational complexity, $\mathcal{O}(N^3)$, of the MMSE receiver.

It is insightful to specialize the fixed point iterations for some simple channel models. Consider the case when the channel matrix is orthogonal, i.e.,

$$G_{:,i}^T G_{:,j} = \| G_{:,i} \|_2^2 \, \delta_{i,j}.$$

Note that the AWGN channel and Rayleigh fading channel are special cases of the orthogonal channels. In these cases, the fixed point iterations in (30) reduce to the following one point update:

$$\mathbf{q} = \frac{1}{1+e^{-\left(\frac{\sqrt{2}}{\sigma^2} G^T \mathbf{z}\right)}}. \tag{34}$$

Therefore, deciding the hard symbols from the probability vector $\mathbf{q}$ is tantamount to slicing the matched filtered observation: $\tilde{\mathbf{z}} = G^T \mathbf{z}$. Deciding $s_k = \pm\frac{1}{\sqrt{2}}$ based on $q_k \gtrless 0.5$ is equivalent to that based on $\tilde{z}_k \gtrless 0$. In other words, VSSD is an ML decoder for orthogonal channels.

Further, for any channel matrix, the ELBO is upper bounded by the marginal log likelihood, $\log p_\theta(\mathbf{z})$, and (as shown in the appendix) the fixed point updates never decrease the ELBO. Therefore, the fixed point iterations always converge to a local maximizer of the ELBO.

Finally, we propose to accelerate the fixed point updates to achieve a faster convergence. Specifically, we choose γ_n at the nth iterate so that the update,

$$\mathbf{q}_n = \mathbf{q}_{n-1} + \gamma_n \left[\varphi(\boldsymbol{\alpha}_{n-1}) - \mathbf{q}_{n-1}\right], \tag{35}$$

results in maximal ELBO increase. The optimum value of γ_n can be found through a 1-D search over a bounded interval in $\mathbb{R}$ (as provided in the appendix).

In the next section, we investigate the performance of the VSSD based symbol detection via Monte Carlo simulations.

TABLE I
S2C PARAMETERS USED IN THE SIMULATION.

Carrier frequency (f_c)	15 kHz
Bandwidth (B)	10 kHz
Chirp rate ($2m_c$)	1 MHz/s
Symbol duration (T)	0.5 ms
Sweep duration (T_{sw})	10 ms
Guard interval (T_g)	25 ms

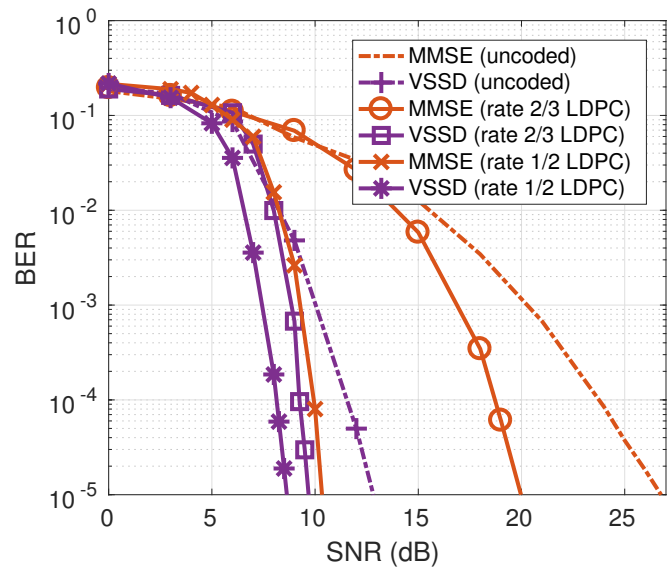

Fig. 4. BER performance of VSSD and MMSE receivers for i.i.d Gaussian channel matrix.

V. Numerical Simulations

First, we demonstrate the performance of VSSD over the benchmark i.i.d. Gaussian channel matrix, with entries $G_{i,j} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, and compare that with the MMSE receiver. We define the signal to noise ratio (SNR) as

$$\text{SNR} = \frac{E\{\|\mathbf{G}\mathbf{s}\|_2^2\}}{E\{\|\mathbf{w}\|_2^2\}}. \tag{36}$$

We compare the BER performance of the VSSD and MMSE receivers for $N = 288$ symbols, for uncoded and coded QPSK communications, assuming perfect channel knowledge. For coded communication, we use a rate $1/2$ and rate $2/3$ LDPC code from [18]. We terminate the VSSD iterations at the nth iteration if $\|\mathbf{q}_n - \mathbf{q}_{n-1}\|_2 < 10^{-3}$. Figure 4 shows the BER plots at different SNR values. In uncoded communication, the VSSD receiver achieves a BER of 10^{-3} at about 10 dB lower SNR than the MMSE receive. In the rate $2/3$ ($1/2$) coded communication, for a BER of 10^{-3}, VSSD outperforms MMSE receiver by an SNR margin of 8 dB (2 dB). For the same BER (10^{-3}), the VSSD receiver for a rate $2/3$ coded communication works at about 1 dB lower SNR than the MMSE receiver for a rate $1/2$ coded communication. Therefore, VSSD receiver offers 33% higher data rate than the MMSE receiver, while achieving the same BER.

The effect of imperfect channel knowledge, due to channel estimation error, on the BER performance is considered next.

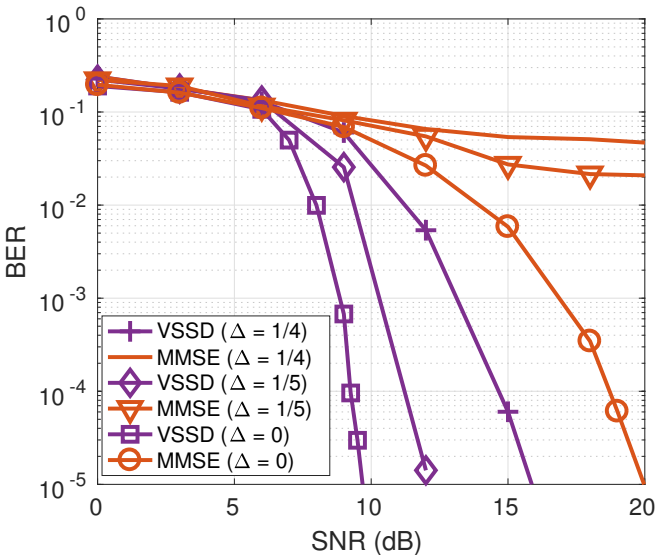

Fig. 5. BER performance of VSSD and MMSE receivers for i.i.d Gaussian channel matrix.

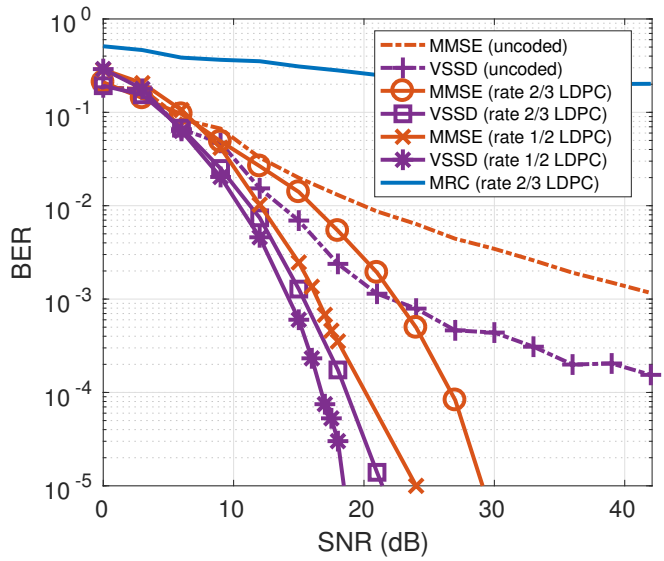

Fig. 6. BER performance of VSSD and MMSE receivers over a UWA channel simulated according to the model in Berger et al. [11].

To do so, we perturb the entries of the i.i.d. Gaussian channel matrix with i.i.d. Gaussian noise, i.e. $G_{i,j} = G_{i,j} + \epsilon_{i,j}$, where $\epsilon_{i,j} \sim \mathcal{N}(0, \Delta), 1 \leq i, j, \leq 2N$. Figure 5 shows the effect on the BER performance of VSSD and MMSE decoders for $\Delta = 1/4, 1/5$ and coded communications using a rate $2/3$ LDPC code. VSSD receiver retains its performance advantage over MMSE even with channel estimation errors.

We next consider the performance of VSSD based receiver for the S2C communication system in Table I. Note that the symbol rate is twice the limit $R_{\max} = \sqrt{2m_c} = 1$ kHz on the existing S2C receivers. A total of $N = 288$ QPSK symbols are mounted on a train of $N_c = 15$ chirp pulses. The channel is generated, as in [11], [12], with $N_p = 16$ discrete paths whose inter-arrival times are exponentially distributed with a mean of 1 ms. The path amplitudes are Rayleigh distributed with the average power decreasing exponentially with delay, where the difference between the beginning and the end of the guard time is 20 dB. Figure 6 shows the BER performance of the VSSD and MMSE based data detection algorithms assuming perfect channel knowledge. Again, from these plots, we notice a strong performance of the VSSD based symbol detection in an S2C receiver. The VSSD receiver attains a BER $= 10^{-3}$ at about 18 dB lower SNR than MMSE in uncoded communication. In coded communication, the SNR margin of VSSD over MMSE receiver is 8 dB (3 dB) for rate 2/3 (1/2) LDPC code. Finally, figure 7 shows the number of VSSD iterations (averaged over at least 1000 trials) for different SNR. On an average, the number of iterations stay below 10 and the maximum number of iterations never crossed 15.

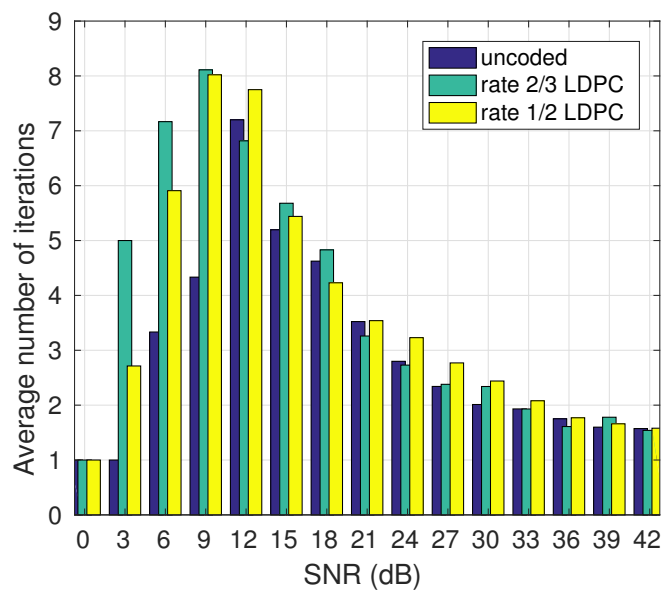

Fig. 7. Number of VSSD iterations averaged over at least 1000 trials for each SNR.

VI. Conclusions

In this work, we considered data symbol detection in an S2C receiver for doubly spread UWA channels. We linked the two existing S2C receivers to the MMSE decoder for certain benign UWA channels. For more severe channels, where the existing receivers either completely fail or must compromise on the data rate, we developed a new soft symbol decoder based on the variational Bayes' inference. Our proposed VSSD decoder estimates a probability vector (soft symbols) whose KL-distance to the true posterior of the symbol vector is minimized by iterating through a fixed point equation. Simulation results showed that VSSD significantly outperforms the MMSE decoder and maintains robust performance under channel estimation errors.

APPENDIX

Evidence Lower Bound (ELBO): We derive the ELBO for the soft symbol estimation. The likelihood term in (27) is given by,

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}} \log p_\theta(\mathbf{z}|G,\mathbf{s}) = -N\log(2\pi\sigma^2) - \mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}\left[\frac{\|\mathbf{z}-G\mathbf{s}\|^2}{2\sigma^2}\right]. \quad (37)$$

On expanding the last term in (37), we get:

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}\left[\|\mathbf{z}-G\mathbf{s}\|^2\right] = \|\mathbf{z}\|^2 - 2\mathbf{z}^T G.\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}[\mathbf{s}] + \mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}\left[\|G\mathbf{s}\|^2\right]. \quad (38)$$

We define:

$$q_{k,\text{Re}} \triangleq q_\phi\left(s_{k,\text{Re}} = \frac{1}{\sqrt{2}}\middle| G,\mathbf{z}\right) \in [0,1], \quad (39)$$

$$q_{k,\text{Im}} \triangleq q_\phi\left(s_{k,\text{Im}} = \frac{1}{\sqrt{2}}\middle| G,\mathbf{z}\right) \in [0,1]. \quad (40)$$

Note that the approximate posterior is completely specified by the soft symbol vector $\mathbf{q} \in \mathbb{R}^{2N}$ formed by stacking up $\mathbf{q}_k = [q_{k^*,\text{Re}}, q_{k^*,\text{Im}}]^T \in \mathbb{R}^2, k = 0,1,\ldots,N-1$. For our problem, we let the parameter $\phi \triangleq \mathbf{q}$.

Expectations in the expression in (38) can be readily evaluated as follows:

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}[s_{k,\text{Re}}] = \frac{1}{\sqrt{2}}(2q_{k,\text{Re}}-1), \quad (41)$$

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}[s_{k,\text{Im}}] = \frac{1}{\sqrt{2}}(2q_{k,\text{Im}}-1), \quad (42)$$

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{y})}}\left[\|G\mathbf{s}\|^2\right] = \text{trace}\left[G^T G \mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}\mathbf{s}\mathbf{s}^T\right] = \sum_{l=0}^{2N-1}\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}[G\mathbf{s}]_l^2 \quad (43)$$

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}[G\mathbf{s}]_l^2 = \sum_{k=0}^{N-1}\left(\eta_{l,k} + \nu_{l,k}\sum_{m\neq k}\nu_{l,m}\right), \quad (44)$$

where,

$$\eta_{l,k} = \frac{1}{2}G_{l,k,\text{Re}}^2 + \frac{1}{2}G_{l,k,\text{Im}}^2 + G_{l,k,\text{Re}}G_{l,k,\text{Im}}(2q_{k,\text{Re}}-1)(2q_{k,\text{Im}}-1), \quad (45)$$

$$\nu_{l,m} = \frac{1}{\sqrt{2}}G_{l,m,\text{Re}}(2q_{m,\text{Re}}-1) + \frac{1}{\sqrt{2}}G_{l,m,\text{Im}}(2q_{m,\text{Im}}-1). \quad (46)$$

The ELBO regularizing term in (27) is:

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}\left[\log\frac{q_\phi(\mathbf{s}|G,\mathbf{z})}{p_\theta(\mathbf{s})}\right] = KL(q_\phi||p_\theta). \quad (47)$$

We assume a simple uniform prior $p_\theta(\mathbf{s}) = \frac{1}{2^{2N}}$. Therefore, when maximizing ELBO, the regularizing term acts to penalize the departure of the variational approximation q_ϕ from the simple uniform prior model. We have:

$$\mathbb{E}_{q_{\phi(\mathbf{s}|G,\mathbf{z})}}\left[\log\frac{q_\phi(\mathbf{s}|G,\mathbf{z})}{p_\theta(\mathbf{s})}\right] = \log 2^{2N} - \sum_{k=0}^{N-1}\left[\mathcal{H}(q_{k,\text{Re}}) + \mathcal{H}(q_{k,\text{Im}})\right], \quad (48)$$

where

$$\mathcal{H}(q) = -q\log q - (1-q)\log(1-q). \quad (49)$$

On combining the likelihood and regularization terms, we find the overall ELBO to be:

$$\begin{aligned}\mathcal{L}(\theta,\mathbf{q},\mathbf{z}) = &-N\log(2\pi\sigma^2) - \frac{\|\mathbf{z}\|^2}{2\sigma^2} + \frac{1}{\sqrt{2}\sigma^2}\mathbf{z}^T G(2\mathbf{q}-1) \\ &- \frac{1}{2\sigma^2}\sum_{l=0}^{2N-1}\sum_{k=0}^{N-1}\left(\eta_{l,k} + \nu_{l,k}\sum_{m\neq k}\nu_{l,m}\right) - \log 2^{2N} \\ &+ \sum_{k=0}^{N-1} -q_{k,\text{Re}}\log q_{k,\text{Re}} - (1-q_{k,\text{Re}})\log(1-q_{k,\text{Re}}) \\ &+ \sum_{k=0}^{N-1} -q_{k,\text{Im}}\log q_{k,\text{Im}} - (1-q_{k,\text{Im}})\log(1-q_{k,\text{Im}}). \quad (50)\end{aligned}$$

Known Noise Variance: In this case, we take θ to be the empty set. The derivative of the overall cost function with respect to $q_{k^*,\text{Re}}$ is given by:

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial q_{k^*,\text{Re}}} = &\frac{\sqrt{2}}{\sigma^2}\mathbf{z}^T G_{:,k^*,\text{Re}} \\ &- \frac{1}{2\sigma^2}\sum_{l=0}^{2N-1}\left(\frac{\partial\eta_{l,k^*}}{\partial q_{k^*,\text{Re}}} + 2\frac{\partial\nu_{l,k^*}}{\partial q_{k^*,\text{Re}}}\sum_{m\neq k^*}\nu_{l,m}\right) \\ &- \log q_{k,\text{Re}} + \log(1-q_{k,\text{Re}}). \quad (51)\end{aligned}$$

We have $\frac{\partial\eta_{l,k}}{\partial q_{k,\text{Re}}} = 2G_{l,k,\text{Re}}G_{l,k,\text{Im}}(2q_{k,\text{Im}}-1)$ and $\frac{\partial\nu_{l,k}}{\partial q_{k,\text{Re}}} = \sqrt{2}G_{l,k,\text{Re}}$. Setting $\frac{\partial\mathcal{L}}{\partial q_{k^*,\text{Re}}} = 0$ and solving for $q_{k^*,\text{Re}}$, we get

$$q_{k^*,\text{Re}} = \varphi(\alpha_{k^*,\text{Re}}),$$

where $\varphi(x) = \frac{1}{1+e^{-x}}$ and

$$\begin{aligned}\alpha_{k^*,\text{Re}} = &\frac{\sqrt{2}}{\sigma^2}\mathbf{z}^T G_{:,k^*,\text{Re}} \\ &- \frac{1}{\sigma^2}\sum_{l=0}^{2N-1}G_{l,k^*,\text{Re}}G_{l,k^*,\text{Im}}(2q_{k^*,\text{Im}}-1) \\ &- \frac{\sqrt{2}}{\sigma^2}\sum_{l=0}^{2N-1}G_{l,k^*,\text{Re}}\sum_{m\neq k^*}\nu_{l,m}. \quad (52)\end{aligned}$$

Similarly, setting $\frac{\partial\mathcal{L}}{\partial q_{k^*,\text{Im}}} = 0$, we get $q_{k^*,\text{Im}} = \varphi(\alpha_{k^*,\text{Im}})$

where

$$\begin{aligned}\alpha_{k^*,\text{Im}} &= \frac{\sqrt{2}}{\sigma^2}\mathbf{z}^T G_{:,k^*,\text{Im}} \\ &- \frac{1}{\sigma^2}\sum_{l=0}^{2N-1} G_{l,k^*,\text{Im}} G_{l,k^*,\text{Re}}(2q_{k^*,\text{Re}}-1) \\ &- \frac{\sqrt{2}}{\sigma^2}\sum_{l=0}^{2N-1} G_{l,k^*,\text{Im}} \sum_{m\neq k^*} \nu_{l,m}. \end{aligned} \tag{53}$$

Stacking up $\mathbf{q}_k = [q_{k^*,\text{Re}}, q_{k^*,\text{Im}}]^T \in \mathbb{R}^2$ into a vector, we get the following fixed point equations:

$$\mathbf{q} = \varphi(\boldsymbol{\alpha}), \tag{54}$$

where the vector $\boldsymbol{\alpha} \in \mathbb{R}^{2N}$ is formed by stacking $\boldsymbol{\alpha}_k = [\alpha_{k^*,\text{Re}}, \alpha_{k^*,\text{Im}}]^T \in \mathbb{R}^2, k = 0, 1, \ldots, N-1$.

Unknown Noise Variance: In this case, we take $\theta = \{\sigma^2\}$. Differentiating the ELBO in (50) with respect to σ^2, we get:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \sigma^2} &= -\frac{N}{\sigma^2} + \frac{\|\mathbf{z}\|^2}{2\sigma^4} - \frac{1}{\sqrt{2}\sigma^4}\mathbf{z}^T G(2\mathbf{q}-1) \\ &+ \frac{1}{2\sigma^4}\sum_{l=0}^{2N-1}\sum_{k=0}^{N-1}\left(\eta_{l,k} + \nu_{l,k}\sum_{m\neq k}\nu_{l,m}\right). \end{aligned} \tag{55}$$

Setting $\frac{\partial \mathcal{L}}{\partial \sigma^2} = 0$ and solving for σ^2, we find:

$$\begin{aligned}\hat{\sigma}^2 &= \frac{\|\mathbf{z}\|^2}{2N} - \frac{1}{\sqrt{2}N}\mathbf{z}^T G(2\mathbf{q}-1) \\ &+ \frac{1}{2N}\sum_{l=0}^{2N-1}\sum_{k=0}^{N-1}\left(\eta_{l,k} + \nu_{l,k}\sum_{m\neq k}\nu_{l,m}\right). \end{aligned} \tag{56}$$

Convergence: We show that every update of the fixed point iteration is along the ELBO gradient (ascent direction) and therefore cannot decrease the ELBO. To see this, consider the inner product of $\varphi(\boldsymbol{\alpha}) - \mathbf{q}$ and $\nabla\mathcal{L}$:

$$(\varphi(\boldsymbol{\alpha}) - \mathbf{q})^T \nabla\mathcal{L} = \sum_{j=0}^{2N-1}(\varphi(\alpha_j) - q_j)\nabla\mathcal{L}_j. \tag{57}$$

We recognize, from equations (51) and (53), that $\nabla\mathcal{L}_j = \alpha_j - \log q_j + \log(1-q_j)$. Each term in (57) is nonnegative since $\sigma(\alpha_j) - q_j \gtrless 0 \Leftrightarrow \alpha_j - \log q_j + \log(1-q_j) \gtrless 0$. Therefore, the inner product is nonnegative and hence the update $\mathbf{q} \to \varphi(\boldsymbol{\alpha}(\mathbf{q}))$ cannot decrease ELBO. Since the ELBO is also upper bounded by the marginal log likelihood, $\log p_\theta(\mathbf{z})$, the fixed point iterations always converge.

Global Maxima: The entries of the Hessian matrix of $\mathcal{L}$ with respect to $\mathbf{q}$, i.e., $\nabla^2_{\mathbf{q}}\mathcal{L} \in \mathbb{R}^{2N\times 2N}$, are given by:

$$\begin{aligned}\frac{\partial^2\mathcal{L}}{\partial q_j^2} &= -\frac{1}{q_j(1-q_j)} < 0, \\ \frac{\partial^2\mathcal{L}}{\partial q_i \partial q_j} &= \frac{\partial^2\mathcal{L}}{\partial q_j \partial q_i} = -\frac{2}{\sigma^2}\sum_l G_{l,i}G_{l,j}, i \neq j, \end{aligned}$$

where $i, j \in \{0, 1, \ldots, 2N-1\}$. For orthogonal channel matrices, the matrix G satisfies $\sum_l G_{l,i}G_{l,j} = 0$, which makes the Hessian negative definite and therefore the stationary point $\mathbf{q}_\star$ a global maximizer of the ELBO.

A larger class of channel matrices, that guarantees global convergence, can be found by requiring $-\nabla^2_{\mathbf{q}}\mathcal{L}$ to be diagonally dominant, i.e.,

$$\eta_j \triangleq \frac{2}{\sigma^2}\left|\sum_{i\neq j}\sum_l G_{l,i}G_{l,j}\right| < \frac{1}{q_j(1-q_j)}, \forall j, \tag{58}$$

which implies:

$$q_j^2 - q_j + 1/\eta_j > 0, \forall j. \tag{59}$$

Now, the condition in (59) holds for every $0 \leq q_j \leq 1$ if and only if $0 \leq \eta_j < 4$. Since diagonal dominance implies positive definiteness (p.d.), $-\nabla^2_{\mathbf{q}}\mathcal{L}$ is p.d. for the class of channel matrices $\mathcal{G} = \{G \in \mathbb{R}^{2N\times 2N} : |\sum_{i\neq j}\sum_l G_{l,i}G_{l,j}| < 2\sigma^2, \forall j\}$ and therefore global convergence is guaranteed whenever $G \in \mathcal{G}$.

Local Maxima: If $G \in \mathcal{G}$, the limit point $\mathbf{q}_\star$ is a global maximizer. Or else, if $G \notin \mathcal{G}$ and $q_{j,\star} \notin (\kappa_j^{(1)}, \kappa_j^{(2)}) \subset [0,1], \forall j$, where $\kappa_j^{(1,2)}$ are the roots of the equation $q_j^2 - q_j + 1/\eta_j = 0$ ($\eta_j > 4$) given by:

$$\kappa_j^{(1,2)} = \frac{1 \pm \sqrt{1 - 4/\eta_j}}{2}, \forall j, \tag{60}$$

then the limit point $\mathbf{q}_\star$ is a local maxima.

Either Local Maxima or Saddle Point: If $G \notin \mathcal{G}$ and $q_{j,\star} \in (\kappa_j^{(1)}, \kappa_j^{(2)})$, then the limit point $\mathbf{q}_\star$ is either a local maxima or a saddle point. Consider, for example, the random channel matrix whose entries are i.i.d. $\mathcal{N}(0,1)$. The root mean square (RMS) length of the interval $(\kappa_j^{(1)}, \kappa_j^{(2)})$ is given by:

$$l_j = \sqrt{\mathbb{E}\left[(\kappa_j^{(1)} - \kappa_j^{(2)})^2\right]} = \sqrt{1 - 4/\mathbb{E}(\eta_j)}. \tag{61}$$

From the definition of η_j in (58), triangle inequality, and i.i.d property of the entries of G, we have:

$$\mathbb{E}[\eta_j] \leq \frac{2}{\sigma^2}\sum_{i\neq j}\sum_l \mathbb{E}\left[|G_{l,i}|\right]\mathbb{E}\left[|G_{l,j}|\right] = \frac{2}{\sigma^2}\sqrt{\frac{2}{\pi}}2N(2N-1), \tag{62}$$

and therefore,

$$l_j \leq \sqrt{1 - \sqrt{\frac{\pi}{2}}\frac{\sigma^2}{N(2N-1)}}. \tag{63}$$

Since $l_j \to 1$ and $\mathbb{P}\{\eta_j > 4\} \to 1$, as $N \to \infty$, for i.i.d. Gaussian channel matrices, the fixed point is (almost surely) in $(\kappa_j^{(1)}, \kappa_j^{(2)})$. Since, in this case, fixed point $\mathbf{q}_\star$ could be a saddle point, we seek to perturb $\mathbf{q}_\star$ so as to move out of the saddle region in an attempt to further increase the ELBO.

Acceleration: The optimum value of γ_n in (35), that best increases ELBO, lies within $[\gamma_{\min}, \gamma_{\max}] \in \mathbb{R}$, where

$$\begin{aligned}\gamma_{\min} &= \max_{0 \le j \le 2N-1} \frac{-q_j}{\varphi(\alpha_j) - q_j}, \\ \gamma_{\max} &= \min_{0 \le j \le 2N-1} \frac{1 - q_j}{\varphi(\alpha_j) - q_j}.\end{aligned}$$

REFERENCES

[1] K. P. Arunkumar and C. R. Murthy, "Variational soft symbol decoding for sweep spread carrier based underwater acoustic communications," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC) (IEEE SPAWC 2019)*, Cannes, France, Jul. 2019.

[2] 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.

[3] K. G. Kebkal and R. Bannasch, "Sweep-spread carrier for underwater communication over acoustic channels with strong multipath propagation," *J. Acoust. Soc. Am.*, vol. 112, no. 5, pp. 2043–2052, Nov. 2002.

[4] "Underwater acoustic modems," in *Product Information Guide*. Evologics GmbH.

[5] K. G. Kebkal and R. Bannasch, "Method and devices for transmitting and receiving information," *U.S. Patent*, vol. 6,985,749 B2, Jan. 2006.

[6] K. Kebkal, A. K. Kebkal, and G. A. Ermolin, "Mathematic and experimental evaluation of phase errors when receiving hydro-acoustic PSK-signals with sweep-spread carrier in reverberant underwater environments," in *Proc. MTS/IEEE OCEANS Conf.*, June 2013.

[7] K. G. Kebkal, O. G. Kebkal, and R. Bannasch, "Synchronisation of underwater communication receivers by means of swept pulses," in *Proc. of the 4th Int. Conf. on Underwater Acou. Measurements: Technologies and Results*, June 2011.

[8] K. G. Kebkal, V. K. Kebkal, O. G. Kebkal, and R. Petroccia, "Clock synchronization in underwater acoustic networks during payload data exchange," in *Proc. 2nd Int. Conf. Exhibit. Underwater Acoust.*, June 2014.

[9] K. Kebkal, A. Kebkal, and V. Kebkal, "Synchronization tools of acoustic communication devices in control of underwater sensors, distributed antennas, autonomous underwater vehicles," in *Gyroscopy and Navigation*, vol. 5, no. 4. Pleiades Publishing, 2014, pp. 257–265.

[10] L. Marchetti and R. Reggiannini, "An efficient receiver structure for sweep-spread-carrier underwater acoustic links," *IEEE J. of Ocean. Eng.*, vol. 41, no. 2, pp. 440–449, April 2016.

[11] C. R. Berger, S. Zhou, J. C. Preisig, and P. Willett, "Sparse channel estimation for multicarrier underwater acoustic communication: From subspace methods to compressed sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 2941–2965, May 2011.

[12] J. Z. Huang, S. Zhou, J. Huang, C. R. Berger, and P. Willett, "Progressive inter-carrier interference equalization for OFDM transmission over time-varying underwater acoustic channels," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 8, pp. 1524–1536, Dec. 2011.

[13] S. Zhou and Z. Wang, "OFDM for underwater acoustic communications." Wiley, 2014.

[14] M. Stojanovic and L. Freitag, "Multichannel detection for wideband underwater acoustic CDMA communications," *IEEE J. of Ocean. Eng.*, vol. 31, no. 3, pp. 685–695, Jul. 2006.

[15] Z. Wang, S. Zhou, G. B. Giannakis, C. R. Berger, and J. Huang, "Frequency-domain oversampling for zero-padded OFDM in underwater acoustic communications," *IEEE J. of Ocean. Eng.*, vol. 37, no. 1, pp. 14–24, Jan. 2012.

[16] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts," *IEEE J. Ocean. Eng.*, vol. 33, no. 2, pp. 1638–1649, Apr. 2008.

[17] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," in *Proceedings of the 2nd International Conference on Learning Representations*, April 2014.

[18] M. Helmling, S. Scholl, F. Gensheimer, T. Dietz, K. Kraft, S. Ruzika, and N. Wehn, "Database of Channel Codes and ML Simulation Results," www.uni-kl.de/channel-codes, 2019.