

# A Factor-Graph Clustering Approach for Detection of Underwater Acoustic Signals

Dror Kipnis and Roe Diamant, *Member, IEEE*,

**Abstract**—We address the challenge of detecting an arbitrary-shaped underwater acoustic signal. Instead of setting a detection threshold, which due to noise transients may result in a high false alarm rate, our method classifies each measured sample as either 'noise' or 'signal'. Utilizing a-priori knowledge of only the minimal duration of the signal, the decision is made using loopy belief propagation over a factor graph. Numerical simulations and sea experimental results show that our scheme achieves a favorable trade-off between the recall and false alarm rates, and noise robustness which far exceeds that of benchmark schemes.

**Index Terms**—Underwater acoustics, signal detection, factor graphs, clustering, loopy belief propagation, sea experiment.

## I. INTRODUCTION

Detection of underwater acoustic signals is the basis of almost all undersea operations involving remote sensing. This includes synchronization for underwater acoustic communication, active and passive remote sensing of objects through underwater acoustics, and acoustic ranging for localization. In all of these applications, the aim is to reliably detect the signal in the presence of ambient noise, with minimal false alarm rate (FAR). This is a challenging task since the ambient noise at sea contains non-Gaussian elements and noise transients. In this paper, we consider the case of passive detection when the signal's waveform modulation is unknown, and our goal is to attain reliable detection in the presence of noise transients.

Without pre-knowledge of the signal's waveform, detection can be preformed using e.g., cyclo-stationary analysis to detect periodic features in the received signal [1], or spectrogram correlation, which requires the time-frequency representation of the signal's rather than its exact waveform [2]. Yet, both alternatives assume some knowledge of the signal's characteristics, which, when mismatched, may yield a high false alarm rate. Instead, the common practice is to employ a constant false alarm rate (CFAR) system through e.g., an energy detector [3], whose detection threshold is set to achieve a desired false alarm rate for an assumed noise distribution. However, energy detection is sensitive to impulsing transients e.g., from snapping shrimps [4] such as in Fig. 1, and may induce false alarms when the noise distribution changes rapidly [4],[5].

Considering the challenge of detection in a rapidly changing signal-to-noise ratio (SNR), the minima controlled recursive averaging (MCRA) method [6] employs an adaptive threshold,

D. Kipnis (email: dkipnis@campus.haifa.ac.il) and R. Diamant (email: roeed@univ.haifa.ac.il) are with the Department of Marine Technology, University of Haifa. This work was sponsored in part by the European Union's Horizon 2020 research and innovation programme under grant agreement No 773753 (Symbiosis), and by the NATO Science for Peace and Security Programme under grant G5293.

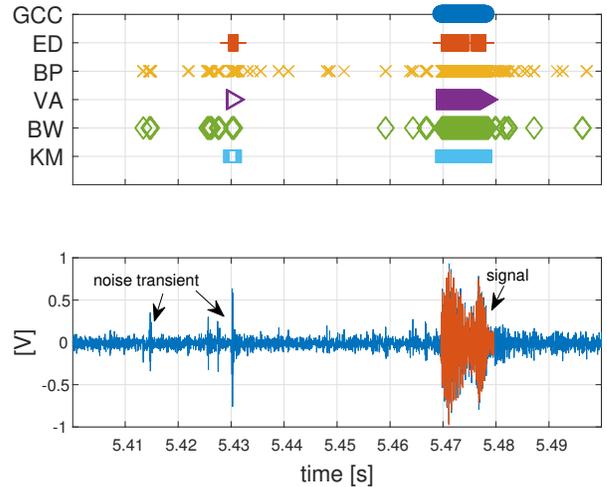


Figure 1: Top - Detection results of GCC and benchmark methods: energy detector (ED), belief propagation (BP), Viterbi algorithm (VA), Baum-Welch (BW) and k-means (KM). Bottom - received signal in the time domain. Results show that only GCC detected the signal with no false alarms.

determined according to an exponential averaging of the time-dependent estimated noise level. However, MCRA requires the adjustment of multiple parameters, which might affect its robustness to sea conditions and signal structure.

In this paper, we describe a new passive detection scheme for underwater acoustic signals, referred to as *graph calibration clustering* (GCC). Our goal is to attain reliable detection in the presence of noise transients. Assuming knowledge of only the minimum duration of the desired signal, we represent the probabilistic dependencies between sequential observations through a graphical model. Instead of setting a threshold, detection decision is made based on the credibility assigned to each observation. In doing so, we do not assume any model for the noise or signal statistics. Our contribution is twofold:

- 1) A new probabilistic clustering method for passive detection of underwater acoustic signals.
- 2) A detection method that utilizes a lower bound on the signal's duration to probabilistically reject noise transients.

We test our method in numerical simulations and in sea experiments, and show that it achieves a favorable trade-off between the recall and false alarm rates.

The remainder of this paper is organized as follows: Section II specifies our system's model and assumptions. Our proposed

GCC method is described in Section III. Simulation and sea experimental results are presented in Sections IV and V, respectively. Additional technical issues are discussed in VI, and Conclusions are drawn in Section VII.

## II. SYSTEM MODEL AND ASSUMPTIONS

Our input is a sequence of observed time-domain samples  $e_i, i = 1 \dots N$ . The waveform modulation of the desired signal is assumed to be unknown, but we set a lower bound,  $\mathcal{L}$ , on its duration. No assumption is made on the ambient noise or the signal's distribution. We consider a binary hypothesis test for each sample  $i$ , such that,

$$e_i = \begin{cases} n[i], & \mathcal{H}_1 \\ n[i] + s[i](u[i - \tau] - u[i - \tau - \tau_s]), & \mathcal{H}_2 \end{cases} \quad (1)$$

where  $n[i]$  is a noise sample,  $s[i]$  is a signal sample,  $u[i]$  is the step function, and  $\tau$  is the time instance in which the signal is received,  $\tau_s > \mathcal{L}$  is the duration of the signal. For accurate detection, our goal is to correctly label each sample  $e_i$ .

## III. DESCRIPTION OF THE METHOD

### A. Key Idea

The key observation behind our scheme is that, in contrast to the observed noise samples, which are assumed to be independent, the signal's samples are actually statistically dependent. We utilize this dependency, as well as the prior knowledge of  $\mathcal{L}$ , by formalizing the problem as a bipartite factor graph, where samples correspond to variable nodes, and the factors favor sequences of at least  $\mathcal{L}$  variables. We then iterate along the factor graph, and make detection decisions per sample, based on the resulting beliefs. This setup resembles the architecture of the low density parity check (LDPC) coding scheme [7], where message nodes and check nodes are analogue to variable nodes and factor nodes, respectively. A legitimate codeword is a sequence of at least  $\mathcal{L}$  variables, rejecting shorter transients. The novelty in our approach is therefore the utilization of local statistical dependency of the signal's samples over time scales larger than  $\mathcal{L}$ , to reduce false alarms which otherwise exists due to noise transients.

### B. The Graphical Model

Our graphical model is illustrated in Fig. 2. Each sample  $e_i$  is connected to a variable node (also referred to as a hidden node)  $x_i$  that is in either state 1 (*noise*) or state 2 (*signal*). The variable nodes  $x_i$  are connected through factor nodes  $u_j$ , which form a Bethe cluster graph [8, ch.11]. The limit on the signal's duration  $\mathcal{L}$  is depicted by a length parameter  $L$ ,  $\mathcal{L} = 2L + 1$ . In our model, each factor node  $u_j$ ,  $L + 1 < j < N - L$  is attached to nodes  $x_{i-L}, x_{i-1}, x_i, x_{i+1}$  and  $x_{i+L}$ , (for example, in Fig. 2 we have  $N = 15$  and  $L = 5$ ). This graph architecture is chosen because, if a sample is classified as a 'signal' then, depending on the sample's position within the sequence, its neighbors would also be classified as such.

The potential value  $\phi$  of a factor node corresponds to the probability of a legitimate sequence of states to occur. It is chosen such that a value  $\phi = 1$  is assigned to a legitimate

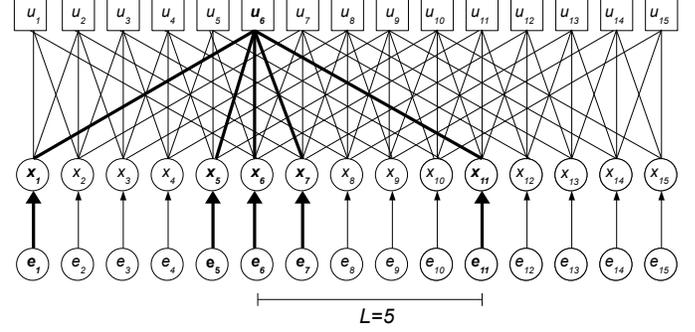


Figure 2: The graphical model for  $L = 5, N = 15$ : For example, factor node  $u_6$  connects variable nodes  $x_1, x_5, x_6, x_7$  and  $x_{11}$ .

Table I: Values for which  $\phi(x_{i-L}, x_{i-1}, x_i, x_{i+1}, x_{i+L}) = 1$ . The left column indicates the line entries in the full 32 rows table describing the 5-connected factors.

| line entry | $x_{i-L}$ | $x_{i-1}$ | $x_i$ | $x_{i+1}$ | $x_{i+L}$ | $\phi$ |
|------------|-----------|-----------|-------|-----------|-----------|--------|
| 1          | 1         | 1         | 1     | 1         | 1         | 1      |
| 2          | 1         | 1         | 1     | 1         | 2         | 1      |
| 4          | 1         | 1         | 1     | 2         | 2         | 1      |
| 8          | 1         | 1         | 2     | 2         | 2         | 1      |
| 16         | 1         | 2         | 2     | 2         | 2         | 1      |
| 17         | 2         | 1         | 1     | 1         | 1         | 1      |
| 25         | 2         | 2         | 1     | 1         | 1         | 1      |
| 29         | 2         | 2         | 2     | 1         | 1         | 1      |
| 31         | 2         | 2         | 2     | 2         | 1         | 1      |
| 32         | 2         | 2         | 2     | 2         | 2         | 1      |

sequence of states (i.e, all variables are at the same state, or a single transition between states), otherwise  $\phi = \epsilon \ll 1$ . We list legal variable node values' combinations for which  $\phi = 1$  in Table I. The rest of the table's entries are marked with  $\phi = \epsilon$ .

### C. Graph Calibration

To calibrate the graph in Fig. 2, we use the sum-product loopy belief propagation (LBP) scheme [8, ch.11]. The process starts by initializing the variable nodes  $x_i$  according to a posterior estimation  $p(x_i|e_i)$ . Messages  $\delta_{x_i \rightarrow u_j}, \delta_{u_j \rightarrow x_i}, i, j = 1 \dots N$  from  $x_i$  to  $u_j$  and vice versa, respectively, are initialized to 1.

Let  $\mathcal{K}(x_i, j)$  be the set of factor nodes neighboring to  $x_i$ , excluding  $u_j$ . For each variable node  $x_i$ , messages to each of its neighboring factor nodes  $u_j$  are calculated by:

$$\delta_{x_i \rightarrow u_j} = p_{i,s} \prod_{k \in \mathcal{K}(x_i, j)} \delta_{u_k \rightarrow x_i}, \quad (2)$$

where  $p_{i,s} = p(x_i = s|e_i)$  is the posterior for  $x_i$  to be  $s$  in  $\{1, 2\}$ . For each factor node  $u_j$ , message to its neighboring variable node  $x_i$  are calculated by:

$$\delta_{u_j \rightarrow x_i} = \psi_j(x_{\mathcal{K}(u_j, i)}) \prod_{k \in \mathcal{K}(u_j, i)} \delta_{x_k \rightarrow u_j}, \quad (3)$$

where  $x_{\mathcal{K}(u_j, i)}$  is the set of variables neighboring to  $u_j$  excluding  $x_i$ , and  $\psi_j(x_{\mathcal{K}(u_j, i)}) = \sum_{x_i} \phi_j(x_{\mathcal{K}(u_j, i)}, x_i)$  is the factor marginalization of  $x_i$  in  $\phi_j$  [8, ch.9].

The above procedure is repeated iteratively until convergence of the messages. Then, beliefs  $B(x_i, s)$  are calculated for each variable node  $x_i$  according to:

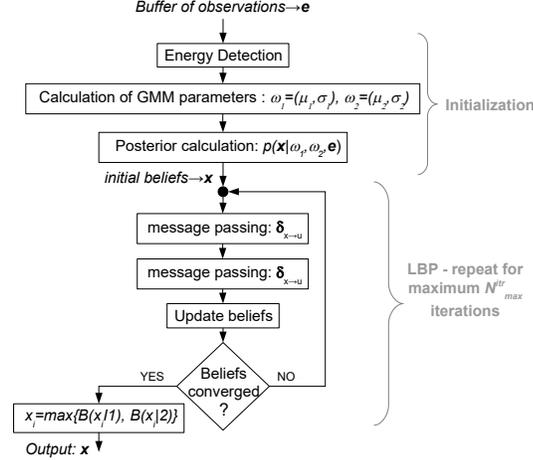


Figure 3: Diagram of the GCC algorithm.

$$B(x_i, s) = p_{i,s} \prod_k \delta_{u_k \rightarrow x_i}, \quad (4)$$

Finally,  $x_i = 1$  if  $B(x_i, 1) > B(x_i, 2)$ , otherwise  $x_i = 2$ .

#### D. Initialization

The initialization of the variable nodes is based on the posterior estimation,  $p(x_i|e_i)$ . To this end, we assume a Gaussian mixture model (GMM)

$$p(x_i|e_i) = \sum_{s=1}^2 k_s p(x_i|\omega_s, e_i), \quad (5)$$

where  $k_s$  is the prior of class  $s$ ,  $\omega_s = \{\mu_s, \sigma_s\}$ , and  $\mu_s$  and  $\sigma_s$  are the mean and variance of the Gaussian  $p(x_i|\omega_s, e_i)$ , respectively. Then,  $x_i$  are initialized by  $p(e_i|1) = p(x_i|\omega_1, e_i)$  and  $p(e_i|2) = p(x_i|\omega_2, e_i)$ .

To find  $\omega_s, s = 1, 2$ , we use a statistical test by pre-clustering each sample  $e_i$ . To achieve this, we use an energy detector tuned to a high desired false alarm rate (FAR), such that signals of low SNR are also detected. We calculate  $\omega_1$  statistically for all samples  $e_i$  not detected by the energy detection, and vice versa for  $\omega_2$ . We note that this preliminary energy detection is used only as a rough clustering, mainly to reduce computational costs.

#### E. Complexity analysis

Fig. 3 shows a flow chart of GCC. Our protocol involves message passing iterations from variable nodes to factor nodes and vice versa. Let  $\kappa = 5$  be the number of nodes attached to each variable/factor. The complexity of messages from variables to the factors,  $\delta_{x \rightarrow u}$  (2) is  $\kappa N$ , and the complexity of the  $\delta_{u \rightarrow x}$  messages (3), which involve the factor marginalization, is  $(\kappa + \kappa)N$ . Hence, the complexity of each message passing iteration is  $\kappa O(N)$ , and the complexity of running the whole algorithm for  $N_{itr}$  iterations is the linear term  $\kappa O(NN_{itr})$ .

## IV. NUMERICAL ANALYSIS

Our numerical analysis attempts to evaluate the performance of our GCC detector in terms of the detection and false alarm rates. Since GCC clusters the data, we measure detection performance by the Recall,

$$\text{Recall} = \frac{n_D}{n_{2tot}}, \quad (6)$$

where  $n_D$  is the number of correctly detected samples with  $s = 2$ , and  $n_{2tot}$  is the ground-truth for state 2 (signal). FAR is calculated on a sequence of noise, according to  $\frac{n_F}{N_{noise}}$ , where  $n_F$  is the number of falsely detected samples, and  $N_{noise}$  is the total length of the noise sequence.

Since we have neither a-priori knowledge of the likelihood of the hypotheses or their corresponding cost, we cannot compare our method to Bayesian or minmax decision-making [3, ch.3.2]. Instead, together with the simple energy detection, we consider benchmark methods that do not depend on the noise distribution. This includes the Viterbi algorithm (VA) [9][10], the Baum-Welch (BW) [10], the belief propagation (BP) method [11], and K-means clustering (KM) [12].

The specifications of the benchmark schemes are as follows: For the energy detector (ED), we set the detection threshold for a target FAR of  $10^{-8}$ , and the averaging window to be  $2L + 1$  as in GCC; for the VA, we perform detection by clustering the received samples after convergence. We use the prior estimation from the initial ED (see Section III-D) to set the transition matrix  $T$  and emission matrix  $E$  according to:

$$T = \begin{bmatrix} 1 - 1/N & 1/N \\ 1/N & 1 - 1/N \end{bmatrix} \quad (7)$$

$$E = \begin{bmatrix} p(x_i = 1|e_i = 1) & p(x_i = 2|e_i = 1) \\ p(x_i = 1|e_i = 2) & p(x_i = 2|e_i = 2) \end{bmatrix} \quad (8)$$

The BW and the KM schemes are initialized by the initial ED; In contrast, BP is a simple version of GCC, where no factor graph is used and the dependencies form a causal Markov chain [11, ch. 4.2]. For BP, we use  $E$  from (8) as the conditional probability matrix  $M_{yx}$ . We note that we were unable to find a benchmark that utilizes the information on the lower bound  $\mathcal{L}$ . This is in addition to the trivial case of requiring  $\mathcal{L}$  consecutive detection decisions to declare detection, which is by no means either efficient or optimal.

#### A. Simulation Setup

Detection performance is tested over 100 buffers, each containing 4,800 samples of both noise and a synthesized sine wave of length  $\tau_s$  samples and of frequency 12kHz, sampled at 48kHz. To examine the impact of the sea ambient noise, we carry out simulations for both the Gaussian noise and the real sea noise we recorded in a sea experiment in shallow water of 10 m depth<sup>1</sup>. The length parameter of the GCC is  $L = 50$ .

Recall as a function of SNR is measured for  $\tau_s = 480$  samples. Since we measure performance per sample decision, the statistics obtained is for 48,000 trials. Dependence of

<sup>1</sup>While there are models that treat ambient sea noise as  $\alpha$  distribution [4], we used real recorded noise to demonstrate the effectiveness of our algorithm.

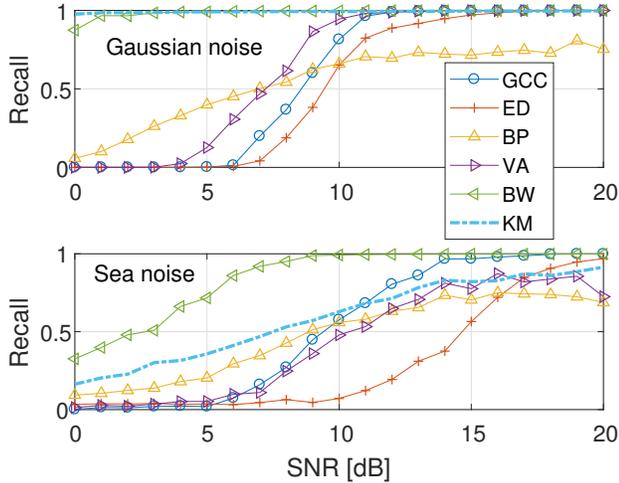


Figure 4: Recall for synthesized signals with Gaussian noise (top) and recorded sea noise (bottom).  $L = 50$ . Results show that the GCC achieves reasonable results compared to the benchmarks, and is less affected by sea noise.

Recall on  $\tau_s$  is evaluated for SNR=15dB and  $L = 50$ . FAR is measured over 359 buffers, each containing 4,000 noise samples, obtaining statistics for  $\sim 1.4$  million trials. Sensitivity of the algorithm to  $L$  is explored by receiver operation characteristic (ROC) curve, using  $L$  as a configuration parameter.

### B. Simulation Results

Fig. 4 shows the recall measure for the various tested methods vs. the SNR for both Gaussian noise and sea noise. The results show that GCC achieves reasonable detection performance compared to the benchmarks. GCC performance is also the most robust for the two noise source cases, since its performance are least affected by the noise distribution.

Fig. 5 shows the cumulative distribution function (CDF) of the FAR. Results show that, except for KM, whose performance is generally poor, the FAR of all methods highly increases in the presence of sea noise. This is mostly due to the noise transients as seen in Fig. 1. Still, we observe that GCC performance is much less affected by the sea noise, and that its FAR is much better than that of the benchmarks.

Fig. 6 shows the dependence of Recall on  $\tau_s$ , evaluated for both Gaussian and sea noise, for SNR=15dB. It appears that very short signals with  $\tau_s < 2L$  are effectively rejected, while robust detection is achieved for  $\tau_s > 8L$  (working point for Fig. 4 is  $\tau_s/L = 9.6$ ). These results suggests that the GCC should be applied to signals which are at least an order of magnitude longer than the noise transients we wish to reject.

The ROC curve of Fig. 7 explores trade-off between recall and FAR for sea noise<sup>2</sup>, obtained for various values of  $L$  while  $\tau_s = 480$ . It shows that performance increase when  $L$  becomes larger, and optimum occurs for  $8 < L < 30$ . Degradation of recall for larger values of  $L$  is due to the larger factor graph, which rejects detections more effectively.

<sup>2</sup>The ROC curve is only for sea noise since the FAR for Gaussian noise is too low for our simulation

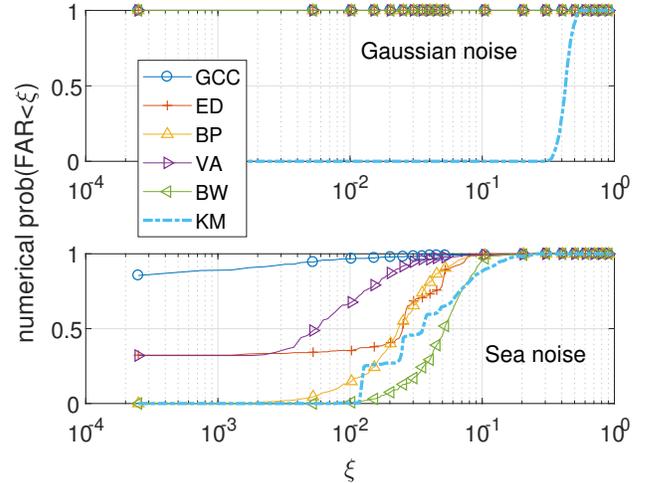


Figure 5: FAR CDF for Gaussian noise (top), and recorded sea noise (bottom).  $L = 50$ .

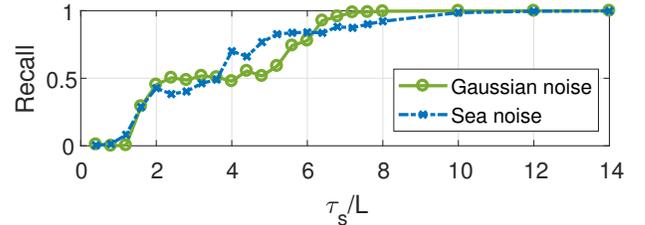


Figure 6: Recall for various values of  $\tau_s/L$ . SNR=15dB.

## V. SEA EXPERIMENT

We now discuss the results we obtained in a sea experiment. The experiment was carried out in Aug 2017, in the Achziv nature reserve in Northern Israel, at a water depth of 10 m. This reserve hosts a population of lobsters and snapping shrimps that emit strong acoustic transients. The experimental setup is illustrated in Fig. 8. We transmitted a sequence of 10 ms chirp-modulated signals at frequencies between 3 kHz and 17 kHz. To transmit the signals, we used the Evologics S2CR 7/17 software-defined acoustic modem, deployed from a boat to a depth of 5 m. The receiver was based on the  $\mu$ Radar-mkII acoustic recorder, deployed at a 5 m depth from a buoy located about 500 m away from the boat. Processing was made

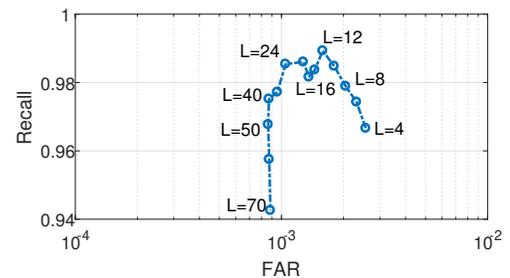


Figure 7: ROC curve for sea noise, SNR=15dB.

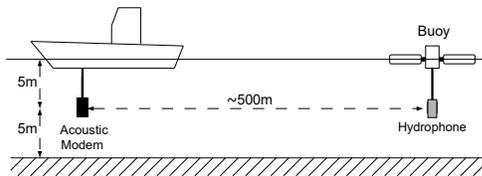


Figure 8: Illustration of the experimental setup.

Table II: Experimental results: Top - FAR. Bottom - recall. Results show that the GCC's FAR is much lower than that of the benchmark, and that its recall is similar.

|        | SNR  | GCC  | ED   | BP   | VA   | BW   | KM   |
|--------|------|------|------|------|------|------|------|
| FAR(%) | N/A  | 0.24 | 2.57 | 2.88 | 1.05 | 6.7  | 5.08 |
| Recall | 10dB | 0.64 | 0.69 | 0.57 | 0.66 | 0.78 | 0.65 |
|        | 15dB | 0.84 | 0.83 | 0.72 | 0.80 | 0.89 | 0.64 |
|        | 20dB | 0.94 | 0.87 | 0.80 | 0.89 | 0.93 | 0.64 |

off-line, using  $L=50$ . An example of detection results on a single buffer is shown in Fig. 1. A matched filter analysis of the received signal showed a delay spread of approx. 1.12ms with 5 significant non-direct arrival paths. Yet, since all detection schemes do not assume knowledge of the signal, all replicas are detected, and multipath have no affect on performance.

The experiment results are summarized in Table II. FAR is defined as the number of false negative detections per 4000 samples, and is averaged over 359 buffers. To evaluate the recall for different SNR values, during the analysis we added recorded ambient sea noise to the obtained buffers. The results show that the GCC's FAR is significantly lower than that of the benchmark methods, while for reasonable SNR of above 10dB, the GCC's recall is at the same range as the leading benchmarks. These results demonstrate that GCC achieves a favorable trade-off between detection and false-alarm performance. It can be seen that our GCC method is the only one to detect the signal without false alarms.

## VI. DISCUSSION

Here we discuss the technical details of the LBP implementation that is a major part of GCC. The implementation we use is based on the simultaneous propagation of state beliefs. The LBP stops when the changes in transferred beliefs are smaller than a pre-defined threshold (0.001), or if convergence is not achieved after  $\gamma=N_{itr}^{max}$  iterations (200 in our implementation). Monitoring the number of iterations required for convergence reveals that, as illustrated in Fig. 9, convergence is achieved in most of the cases. For the other few cases, detailed analysis of the results showed that the length of the signal relative to  $L$  affects the convergence of the algorithm. We also observed that often beliefs corresponding to the majority of a buffer converge while beliefs of relatively few nodes, especially those corresponding to the signal edges, converge slower. Yet, this phenomena has no effect over the detection performance. Future research would focus on identifying redundant links and reduce the complexity load, e.g using informed message scheduling and residual belief propagation [8, ch. 11].

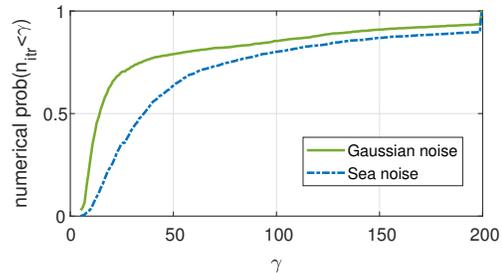


Figure 9: CDF of the number of iterations required for convergence,  $n_{itr}$ . Data are 4800 samples long buffers of Gaussian/sea noise with synthesized transmission signals of various lengths. Statistics were obtained over 2200 runs per noise type. Stopping criteria at  $\gamma=200$  iterations without convergence allows conversion of 90% of the buffers in these simulations.

## VII. CONCLUSION

In this work, we proposed a probabilistic graphical model for detecting a non-stationary arbitrary-shaped underwater acoustic signal of assumed minimal duration. We proposed a novel detector, GCC, which implements loopy belief propagation over a factor graph to assign credibility to each sample. This clustering is then followed by a maximum likelihood detection. We compared the false alarm and detection rates of the GCC to a number of benchmark detection methods in both simulations and in a real sea environment. Our results show that when sea noise was considered, GCC obtained a far better trade-off between the false-alarm rate and the recall rate than the benchmarks. This suggests that the GCC is preferable especially for high noise with transients as appears in a realistic sea environment.

## REFERENCES

- [1] J. Antoni, "Cyclostationarity by examples," *Mechanical Systems and Signal Processing*, vol. 23, pp. 987–1036, 2009.
- [2] D. K. Mellinger, "A comparison of methods for detecting right whale calls," *Canadian Acoustics*, vol. 32, no. 2, pp. 55–65, 2004.
- [3] D. Kazakos and P. Papantoni-Kazakos, *Detection and Estimation*. Computer Science Press, 1990.
- [4] J. R. P. Mandar A. Chitre and S.-H. Ong, "Optimal and near-optimal signal detection in snapping shrimp dominated ambient noise," *IEEE Journal of Oceanic Engineering*, vol. 31, pp. 497–503, 2006.
- [5] R. Diamant, D. Kipnis, and M. Zorzi, "A clustering approach for the detection of acoustic/seismic signals of unknown structure," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, pp. 1017–1029, 2018.
- [6] I. Cohen and B. Berdugo, "Noise estimation by minima controlled recursive averaging for robust speech enhancement," *IEEE Signal Processing Letters*, vol. 9, pp. 12–15, 2002.
- [7] R. G. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, pp. 21–28, 1962.
- [8] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. Cambridge, Massachusetts: MIT Press, 2009.
- [9] A. J. Viterbi, *CDMA: Principles of spread spectrum communication*. Addison-Wesley, 1995.
- [10] L. R. Rabiner and B. H. Juang, "An introduction to hidden markov models," *IEEE ASSP Magazine*, pp. 4–16, 1986.
- [11] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann Publishers, inc., 1988.
- [12] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. Berkeley, Calif.: University of California Press, 1967, pp. 281–297.