Modal estimation in underwater acoustics by data-driven structured sparse decompositions

Clément Dorffer*, Thomas Paviet-Salomon*, Gilles Le Chenadec* and Angélique Drémeau*
*Lab-STICC & ENSTA-Bretagne, UMR CNRS 6285, Brest, France.
E-mail: clement.dorffer@ensta-bretagne.fr, thomas.paviet-salomon@ensta-bretagne.org,
gilles.le_chenadec@ensta-bretagne.fr, angelique.dremeau@ensta-bretagne.fr

***Abstract*—In underwater acoustics, shallow water environments act as modal dispersive waveguides when considering lowfrequency sources. In this context, propagating signals can be described as a sum of few modal components, each of them propagating according to its own wavenumber. Estimating these wavenumbers is of key interest to understand the propagating environment as well as the emitting source. To solve this problem, we proposed recently a Bayesian approach exploiting a sparsity-inforcing prior. When dealing with broadband sources, this model can be further improved by integrating the particular dependence linking the wavenumbers from one frequency to the other. In this contribution, we propose to resort to a new approach relying on a restricted Boltzmann machine, exploited as a generic structured sparsity-inforcing model. This model, derived from deep Bayesian networks, can indeed be efficiently learned on physically realistic simulated data using well-known and proven algorithms.**

***Index Terms*—Underwater modal estimation, structured sparse approximation, Restricted Boltzmann Machine, Bayesian algorithm.**

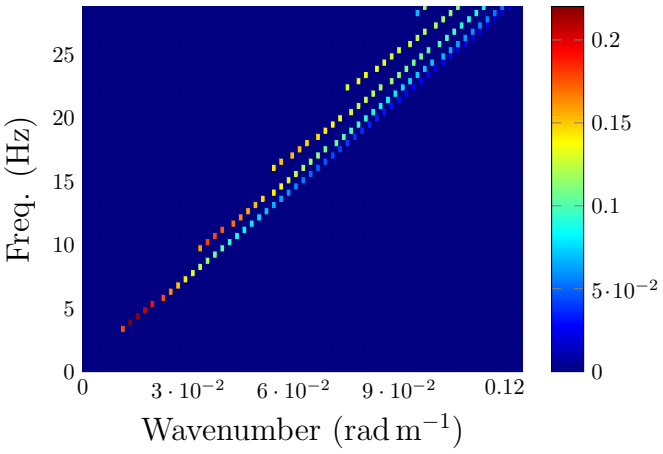

Fig. 1. Illustration of a (f-k) diagram obtained with a Pekeris waveguide.

I. Introduction

In underwater acoustics, shallow environments behave like dispersive waveguides when considering low-frequency sources. An acoustic field received on an antenna is then classically described by a small set of modes propagating longitudinally according to their horizonal wavenumbers. The knowledge of these modes is of great importance for the characterization of the observation environment and, consequently, for the source localization. Among the different methods used to discriminate these modal components, the *frequency-wavenumber* (f-k) representation (see Fig. 1) allows a direct observation of the dispersion (*i.e.,* the frequency dependence) of the wavenumbers. Inherently conceivable for a horizontal array of sensors aligned with the source, they are particularly used in geophysics [1]. Recent contributions have focused on the construction of (f-k) diagrams by exploiting less constrained acquisition schemes, allowing their use in underwater acoustics.

Since propagation is described by a small number of modes, the use of sparsity-inforcing models seems appropriate. In fact, some contributions (see *e.g.,* [2], [3]) have proposed the use of the "compressed sensing" paradigm to estimate modal dispersion. However, if these methods prove to be relevant, we argue that they can be further improved by precisely integrating the dispersion relation linking the wavenumbers from one frequency to another into the estimation process of the (f-k) diagram. To this end, we propose to build on some recent contributions in compressed sensing and machine learning.

On the one hand, research on compressed sensing [4]–[6] and more generally sparse decompositions have underlined the interest of taking into account the structures naturally living in signal representations [7]. On the other hand, restricted Boltzmann machines (RBM) [8] are at the heart of many recent contributions in machine learning that emphasize two precious qualities: *i)* they have been identified as generic models approximating any distribution over $\{0,1\}^n$ [9]; *ii)* efficient algorithms have been developed to train them making them powerful representational models when large data sets are available [10]–[12].

These parallel researches in finally not so distant fields have recently led to the idea of exploiting RBMs learned from large databases to model the structures of sparse representations [13]–[15]. This paper follows on from these contributions. More particularly, we propose to deal with the (f-k) diagram estimation using a new Bayesian algorithm. Our approach considers a RBM as a model for structured sparse representations and exploits it through a mean-field approximation.

Thanks to DGA Naval Systems for funding.

II. Acoustic propagation in shallow water environments

In shallow water, acoustic propagation is described by modal theory. According to the latter, when considering a source at depth z_s and frequency f, the signal received by an hydrophone placed at depth z and distance r of the source can be formulated as

$$y(f,r,z,z_s) = \sum_{m=1}^{K(f)} A_m(f,z,z_s) e^{-irk_{rm}(f)} + n(f,r,z), \quad (1)$$

where $K(f)$ stands for the number of modal components propagating at frequency f, $A_m(f,z,z_s)$ is the modal amplitude associated to the m-th component, $k_{rm}(f)$ is the m-th wavenumber at frequency f and $n(f,r,z)$ stands for some additive measurement noise.

For a M sensor linear antenna, Eq. (1) can be written as

$$\mathbf{y}_f = \mathbf{F}\mathbf{z}_f + \mathbf{n}_f \quad (2)$$

where $\mathbf{F}$ is a $(M \times N)$-dictionary of Fourier discrete atoms, $\mathbf{z}_f$ contains a few non-zero coefficients indicating the wavenumbers propagating at frequency f and quantifying their amplitudes and $\mathbf{n}_f$ stands for the additive noise.

A common approach to estimate $\mathbf{z}_f$ from $\mathbf{y}_f$ is to use a simple (inverse) spatial Fourier transform. Although very popular, the method requires sensors that are finely spaced sufficiently to avoid aliasing. As an alternative, more recent contributions [2], [16], [17] have proposed to leverage the sparsity of $\mathbf{z}_f$ to design more robust methods.

When dealing with broadband sources, the generative model can be written as follows :

$$[\mathbf{y}_{f_1}, \ldots, \mathbf{y}_{f_P}] = \mathbf{F} \cdot [\mathbf{z}_{f_1}, \ldots, \mathbf{z}_{f_P}] + [\mathbf{n}_{f_1}, \ldots, \mathbf{n}_{f_P}], \quad (3)$$

where P is the number of (discretized) frequencies of the source. The estimation of the $\mathbf{z}_f$'s can then be performed frequency per frequency with the same tools as for the mono-frequency case. Stacking these estimates $\mathbf{z}_f$ one on top of the other leads to the (f-k) diagram (see Fig. 1) and makes appear some structures linking the wavenumbers propagating from one frequency to the other. These structures are well-known in practice under the physical name of dispersion relation. While some contributions in the literature work on an explicit combination of this relation with sparse approaches [2], [17], we propose here to resort to a Bayesian formulation of the (f-k) diagram estimation problem using a particular choice of prior holding the structure information.

To that end, a vectorized version of Eq. (3) is built by defining $\mathbf{y} = [\mathbf{y}_{f_1}^T, \ldots, \mathbf{y}_{f_P}^T]^T$, $\mathbf{z} = [\mathbf{z}_{f_1}^T, \ldots, \mathbf{z}_{f_P}^T]^T$, $\mathbf{n} = [\mathbf{n}_{f_1}^T, \ldots, \mathbf{n}_{f_P}^T]^T$ and $\mathbf{D}$ as a block-diagonal matrix with $\mathbf{F}$ P-times repeated on its diagonal. Adopting this formulation, Eq. 3 can be re-formulated as

$$\mathbf{y} = \mathbf{D}\mathbf{z} + \mathbf{n}, \quad (4)$$

where $\mathbf{y}$ (resp. $\mathbf{z}$) is of dimension MP (resp. NP) and $\mathbf{D}$ is a $(MP \times NP)$-dictionary.

The inverse problem we are interested in is then formulated as the estimation of the (f-k) diagram $\mathbf{z}$ from the acoustic signal $\mathbf{y}$ measured over the M-sensor antenna at all of the P frequencies.

III. Problem formulation

According to the modal theory, in shallow water environments and at low frequencies, the vector $\mathbf{z}$ has few non-zero elements, corresponding to the propagating modal wavenumbers, that is $\mathbf{z}$ is assumed to be sparse. Modeling such a property in a Bayesian framework can be realized in different ways. A popular model considers $\mathbf{z}$ as the result of the Hadamard product of a centered circular Gaussian variable[1], say $\mathbf{x} \in \mathbb{C}^{NP}$ with variance σ_x^2 and a binary variable, say $\mathbf{s} \in \{0,1\}^{NP}$. This model explicitly expresses the *support* $\mathbf{s}$ of the sparse vector $\mathbf{z}$, which is a key component of sparse estimation[2]: we look for a good estimation of $\mathbf{s}$ knowing the observations $\mathbf{y}$.

When no structure is assumed in sparse representations, the Bernoulli law constitutes a natural choice for $\mathbf{s}$. Structures, for their part, can be modeled according to various distributions. In [7], [18]–[20], Boltzmann machines are envisaged as generic models encompassing many well-known models. These models depend on parameters which are difficult to train, due - among others - to the presence of a quadratic term. Instead, more recent works [13]–[15] propose to exploit restricted Boltzmann machines, for which efficient training algorithms exist.

Formally, for $\mathbf{s} \in \{0,1\}^{NP}$, RBM can be expressed as

$$p(\mathbf{s}) = \sum_{\mathbf{h}} p(\mathbf{s},\mathbf{h}) \propto \exp(\mathbf{a}^T\mathbf{h} + \mathbf{b}^T\mathbf{s} + \mathbf{s}^T\mathbf{W}\mathbf{h}), \quad (5)$$

where $\mathbf{h}$ is a $L-$dimensional binary hidden variable and $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{W}$ are the RBM parameters.

In [14], the authors propose to implement this model into a OMP-like framework, while in [15], a reweighted ℓ_1-like procedure is considered. Both procedures are deterministic, integrating RBM as an additional extra block and making hard decision on it. In [13], a Bayesian approach based on an approximate message passing procedure is proposed, leading to a more integrated vision of the restricted Boltzmann machine. However, RBM is not fully exploited in the sense that the support $\mathbf{s}$ is still assumed to follow a Bernoulli law, although informed by the RBM. In this paper, we propose a new Bayesian approach, leveraging both *i)* the probabilistic nature of the RBM, and *ii)* the explicit modeling of $\mathbf{s}$ as the commonly visible layer of a RBM (see Eq. (5)).

More particularly, considering model (4)-(5), we are interested in the following marginalized maximum a posteriori estimation problem :

$$(\hat{\mathbf{x}}, \hat{\mathbf{s}}) = \underset{\mathbf{x},\mathbf{s}}{\operatorname{argmax}} \log p(\mathbf{x},\mathbf{s}|\mathbf{y}) \quad (6)$$

[1]For a sake of simplicity, we will use the same notation for a random variable and its realizations.

[2]We note indeed that once the support is estimated, the corresponding coefficients emerge straightforwardly by least-squares estimation.

with $p(\mathbf{x},\mathbf{s}|\mathbf{y}) = \int p(\mathbf{x},\mathbf{s},\mathbf{h}|\mathbf{y})$. In the continuation of [7], we propose to resort to a mean-field approximation.

IV. Deep structured soft Bayesian pursuit

Mean-field (MF) approximations aim at approximating a posterior distribution, here $p(\mathbf{x},\mathbf{s}|\mathbf{y})$, by a "simpler" distribution, say $q(\mathbf{x},\mathbf{s})$, having a easy-to-handle factorization. We consider in this paper the factorization

$$q(\mathbf{x},\mathbf{s},\mathbf{h}) = \prod_{i=1}^{NP} q(x_i,s_i) \prod_{j=1}^{L} q(h_j), \tag{7}$$

with x_i (resp. s_i) the i-th element in $\mathbf{x}$ (resp. $\mathbf{s}$) and h_j the j-th element in $\mathbf{h}$. The approximation $q(\mathbf{x},\mathbf{s},\mathbf{h})$ is chosen to be *as close as possible* to $p(\mathbf{x},\mathbf{s},\mathbf{h}|\mathbf{y})$ in the sense of the Kullback-Leibler (KL) divergence. In practice, this latter optimization problem can be efficiently solved by an iterative algorithm, called "variational Bayesian expectation maximization" (VBEM) algorithm, insuring a decreasing of the KL divergence at each iteration. We expose here below its particularization to model (4)-(5).

The VBEM procedure successively updates all factors in the MF approximation, namely here the $q(x_i,s_i)$'s and $q(h_j)$'s, according to the following iterative rules:

$$q^{(k+1)}(x_i,s_i) \propto \exp\left(\langle \log p(\mathbf{x},\mathbf{s},\mathbf{h},\mathbf{y})\rangle \; \prod_{j} q^{(k)}(h_j) \prod_{j>i} q^{(k)}(x_j,s_j) \prod_{j<i} q^{(k+1)}(x_j,s_j) \right)$$

$$q^{(k+1)}(h_i) \propto \exp\left(\langle \log p(\mathbf{x},\mathbf{s},\mathbf{h},\mathbf{y})\rangle \; \prod_{j>i} q^{(k)}(h_j) \prod_{j<i} q^{(k+1)}(h_j) \prod_{j} q^{(k+1)}(x_j,s_j) \right)$$

where $\langle f(\mathbf{u})\rangle_{q(u_j)} \triangleq \int_{u_j} q(u_j) f(\mathbf{u})\, \mathrm{d}u_j$ and k is the current iteration number. Developing these update rules according to model (4)-(5) gives $q^{(k+1)}(x_i,s_i) = q^{(k+1)}(x_i|s_i) q^{(k+1)}(s_i)$

$$\text{with } q^{(k+1)}(x_i|s_i) = \mathcal{CN}(m_x^{(k+1)}(s_i), \Sigma_x^{(k+1)}(s_i)) \tag{8}$$

$$m_x^{(k+1)}(s_i) = s_i \frac{\sigma_x^2}{\sigma_n^2 + \sigma_x^2 s_i \mathbf{d}_i^H \mathbf{d}_i} \mathbf{d}_i^H \langle \mathbf{r}_i \rangle^{(k+1)} \tag{9}$$

$$\Sigma_x^{(k+1)}(s_i) = \frac{\sigma_n^2 \sigma_x^2}{\sigma_n^2 + \sigma_x^2 s_i \mathbf{d}_i^H \mathbf{d}_i} \tag{10}$$

$$\langle \mathbf{r}_i \rangle^{(k+1)} = \mathbf{y} - \sum_{j\neq i} q^{(n)}(s_j{=}1)\; m_x^{(n)}(s_j{=}1) \mathbf{d}_j \tag{11}$$

$$\text{with } n{=}k+1 \text{ if } j{<}i \text{ and } n{=}k \text{ if } j{>}i$$

where $\mathcal{CN}(\mu,\Gamma)$ stands for the circular Gaussian distribution with mean μ and variance Γ and

$$q^{(k+1)}(s_i) \propto \exp\left(s_i \left(b_i + \sum_l w_{il} q^{(k)}(h_l = 1) \right) \right) \sqrt{\Sigma_x^{(k+1)}(s_i)} \exp\left(\frac{1}{2} \frac{|m_x^{(k+1)}(s_i)|^2}{\Sigma_x^{(k+1)}(s_i)} \right) \tag{12}$$

$$q^{(k+1)}(h_l) \propto \exp\left(h_l \left(a_l + \sum_i w_{il} q^{(k+1)}(s_i = 1) \right) \right) \tag{13}$$

where we have all along denoted by $\mathbf{d}_i$ the i-th column in $\mathbf{D}$ and w_{il} the (i,l)-th element in $\mathbf{W}$.

To help the convergence of the algorithm, an estimation of the noise variance σ_n^2 is implemented in the same way as the one proposed in SoBaP [7], leading to a dependence of $\Sigma_x^{(k+1)}(s_i)$ in the iteration number $(k+1)$ (see Eq. (10)). As stopping criterion, the KL divergence is computed at each iteration, between the target distribution $p(\mathbf{x},\mathbf{s},\mathbf{h}|\mathbf{y})$ and its current MF approximate $q(\mathbf{x},\mathbf{s},\mathbf{h})$: the algorithm stops when this divergence no longer decreases "sufficiently". Pseudo-code 1 gives a practical implementation of the above equations. The use of RBMs being a natural bridge towards deep networks, we will refer to the proposed procedure as the "Deep Structured Soft Bayesian Pursuit" (DSSoBaP).

Pseudo-code 1 DSSoBaP algorithm

Input: $\mathbf{y}, \mathbf{D}, \mathbf{W}, \mathbf{a}, \mathbf{b}, \sigma_x^2, \sigma_n^2$
Initialisation: $\{q^{(0)}(s_i), m_x^{(0)}(s_i)\}_{s_i\in\{0,1\}, i\in\{1,\dots,NP\}}$, $\{q^{(0)}(h_l)\}_{l\in\{1,\dots,L\}}$, $k=0$, KL $= 0$, $\text{KL}_{\text{old}} = \infty$
while $\text{KL}_{\text{old}} - \text{KL} > 10^{-1}$ **do**
  Optional: update σ_n^2 according to [7]
  for $i = 1 \dots NP$ **do**
    update $\langle \mathbf{r}_i \rangle^{(k+1)}$ using (11)
    Optional: if step 2, update $\Sigma_x^{(k+1)}(s_i)$ using (10)
    update $m_x^{(k+1)}(s_i)$ using (9) for $s_i \in \{0,1\}$
    update $q^{(k+1)}(s_i)$ using (12) for $s_i \in \{0,1\}$
  end for
  for $l = 1 \dots L$ **do**
    update $q^{(k+1)}(h_l)$ using (13) for $h_l \in \{0,1\}$
  end for
  $k \leftarrow k+1$
  $\text{KL}_{\text{old}} \leftarrow \text{KL}$
  compute current KL divergence
  if $\text{KL}_{\text{old}} - \text{KL} \leq 10^{-1}$ **then**, $k_{\max} = k$
  end if
end while
Output: $\{q^{(k_{\max})}(s_i=1), m_x^{(k_{\max})}(s_i=1)\}_{i\in\{1,\dots,NP\}}$

Note that DSSoBaP has an algorithmic complexity in the order of $\mathcal{O}(NP(MP+L))$ per iteration. This complexity is reasonable compared to other procedures in the literature, such as *e.g.*. RBM-OMP [14]. This drawback in particular prevented us from using RBM-OMP in our particular problem.

V. Experiment

In this section, we confront our approach to synthetic experiments and compare its performance to the SoBaP standard procedure.

A. Data simulation and training phase

Several models of the ocean have been proposed in the literature of modal theory [21]. One of the most popular is the Pekeris waveguide. In this model, the sea surface is assumed to be perfectly reflective, and the sea bottom is considered as a semi-infinite fluid. Its parameters therefore

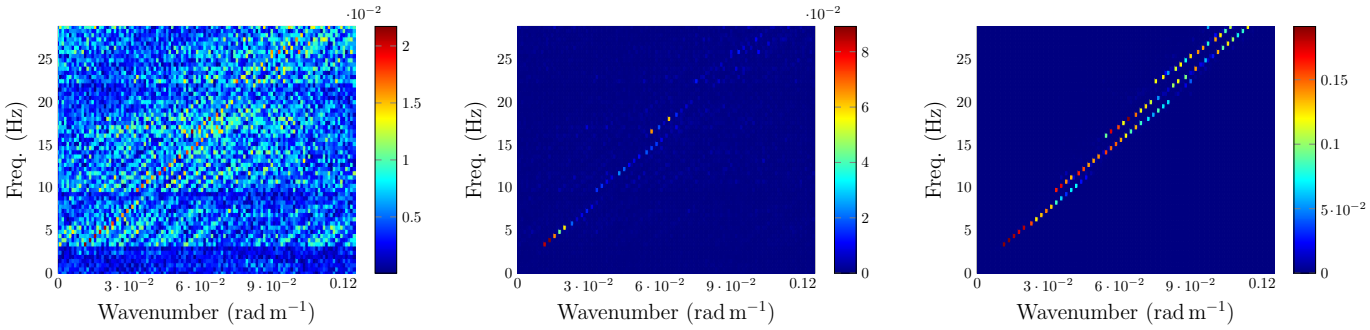

Fig. 2. (f-k) diagrams obtained for $M/N = 0.1$ and SNR= 3 dB by (left) inverse Fourier transform; (middle) SoBaP; (right) DSSoBaP.

include water layer thickness D (constant over the source-receiver distance), water density ρ_w and celerity c_w, ground density ρ_g and celerity c_g. Considering this simple model, we propose to simulate the propagation of an acoustic signal in various Pekeris environments and to learn the structures in the corresponding (f-k) diagrams. The considered setup is as follows.

Source and antenna are assumed to lie on the ground at same depth $z_s = z = D$ varying from $100\,\mathrm{m}$ to $200\,\mathrm{m}$. We simulate the water celerity c_w from $1450\,\mathrm{m\,s^{-1}}$ to $1550\,\mathrm{m\,s^{-1}}$, and the water density ρ_w from $950\,\mathrm{kg\,m^{-3}}$ to $1050\,\mathrm{kg\,m^{-3}}$. The studied frequencies range from $0\,\mathrm{Hz}$ to $30\,\mathrm{Hz}$ with a step-size of $0.5\,\mathrm{Hz}$, *i.e.,* $P = 60$ frequencies are considered, and we set the ground properties c_g and ρ_g from the sediments geoacoustic parameters synthesis in [22]. The simulated antenna is composed of $M = 120$ sensors regularly spaced from $50\,\mathrm{m}$.

With such parameters, we generated a set of 3575 signals – of size 120×60, *i.e.,* 120 sensors, 60 frequencies–and their theoretical sparse decomposition (f-k) diagrams – of size 120×60, *i.e.,* $N = 120$ potential wavenumbers for each 60 frequencies – whose sparsity ratios fluctuate between 3.10^{-3} and 3.10^{-2}. We then randomly divided these simulations into a training set (3000 samples) and a testing set (575 samples) and binarized the (f-k) diagrams from the training set so as to obtain their theoretical supports.

These supports were used to train a RBM with $L = 30$ hidden units. Training parameters were fixed to 10^{-2} for the learning rate, 50 for the minibatch size, 2000 for the number of epochs and the Persistent Contrastive Divergence algorithm [10] was employed with an l_1 regularization so as to inforce the sparsity of the learned features. The obtained RBM is then integrated as prior into DSSoBaP.

B. Tests and results analysis

To assess the relevance of RBM into the considered inverse problem, we compare the performance of DSSoBaP with that of SoBaP, introduced in [7]. This algorithm is based on the same MF approximation as DSSoBaP but instead of the RBM prior (see Eq. (5)), SoBaP exploits a *i.i.d.* Bernoulli model on the SR support. It has to be noticed that we do not compare the approach with those proposed in [7], [14], mainly because of their computational cost.

Two figures of merit are considered: *i)* the estimation of the modal amplitudes (*i.e.,* the non-zero coefficients in the $\mathbf{z}_f$'s) is evaluated according to the normalized mean square error (NMSE), *ii)* the detection of the wavenumbers (*i.e.,* the support of the sparse $\mathbf{z}_f$'s) is evaluated according to the True Positives Rate (TPR) which is the proportion of the theoretical support that has been correctly recovered and the False Discovery Rate (FDR) which is the proportion of the estimated support that should not have been activated (*i.e.,* the proportion of false detection within the estimated support)Each metric is valued according to the signal-to-noise ratio (SNR) on the one hand, and the number of sensors on the other hand. For the tests related to the SNR, we fix the number of observations to 1440 by randomly selecting 24 sensors and let the SNR range from $0\,\mathrm{dB}$ to $5\,\mathrm{dB}$ by adding Gaussian noise $\mathbf{n}_f$ with modulated variance σ_n^2. For the tests related to the number of sensors – which is classical in compressed sensing – we choose to fix the SNR to $3\,\mathrm{dB}$ and let the proportion of hydrophones range from 0.1 to 1 by randomly selecting from 12 to 120 sensors of the antenna.

SoBaP and DSSoBaP are run using the same stopping criterium as described in pseudo-code 1. For SoBaP, the Bernoulli parameter is set for each test sample to its exact value that is the theoretical proportion of nonzero coefficients in $\mathbf{z}$, while for DSSoBaP the prior is provided by the same trained RBM for all test samples. Finally, the performance metrics are averaged over the entire test dataset.

Fig. 3 illustrates the obtained results. One can see that DSSoBaP performance curves follow the same tendencies as SoBaP while being better in most cases, especially in the worst configurations, *i.e.,* with low SNR and low M/N ratio. In particular, the gap between DSSoBaP and SoBaP TPR curves shows that the RBM-based prior has been more effective than the Bernoulli prior for recovering the theoretical support, while the similar FDR curves show that it also conserves the same sparsity properties of the Bernoulli prior by producing only slightly more false detections. Interestingly, we can observe that the improvement of the detection performance leads to an improvement of the NMSE metric, that is the amplitudes of the modes are consequently better recovered with DSSobaP

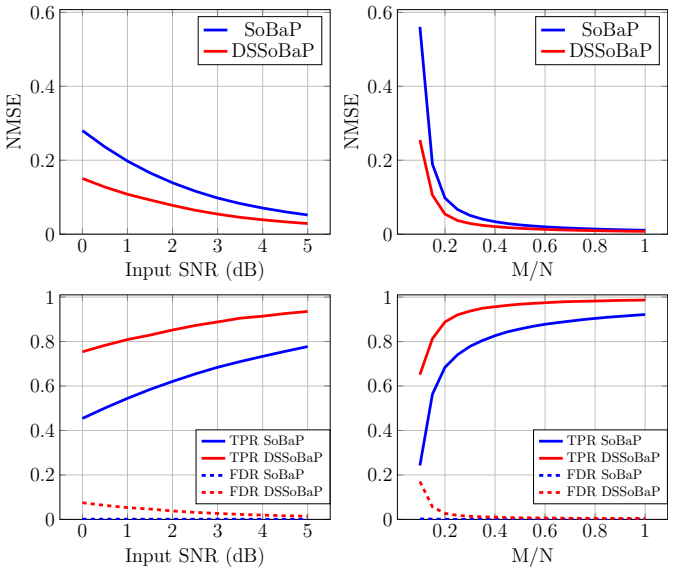

Fig. 3. NMSE (top) and TPR/FDR (bottom) w.r.t. the SNR (left) and the sensor ratio M/N (right).

than with SoBaP.

An example of recovered (f-k) diagrams by inverse Fourier transform, SoBaP and DSSoBaP is shown on Fig. 2. The considered test parameters are SNR=3 dB and sensor ratio $M/N = 0.1$. One can see that DSSoBaP is able to detect much elements from the ground truth support than SoBaP. In particular, structures that are characteristic of modal propagation – the line shapes induced by wavenumber propagation along frequencies – are better recognized using the RBM-based prior, thus showing that RBM are, as expected, well suited for learning spatial dependencies in the support density.

VI. Conclusion

In this work, we introduced a new Bayesian sparse decomposition algorithm able to exploit an RBM – learned in a previous step – as a structure model on the support of a sparse representation. The proposed algorithm relies on an MF approximation and allows a natural integration of the RBM. Its performance on a current underwater acoustic problem is very promising compared to its "standard" counterpart which exploits non-structured Bernoulli variables. Moreover, as the training phase is performed on a collection of physically-realistic simulations, it naturally incites the prior to respect some theoretical structures, which might be helpful when dealing with real datasets.

References

[1] L. Amundsen and B. Ursin, "Frequency-wavenumber inversion of acoustic data," *Geophysics*, vol. 56, no. 7, pp. 1027–1039, 1991.

[2] A. Drémeau, F. Le Courtois, and J. Bonnel, "Reconstruction of dispersion curves in the frequency-wavenumber domain using compressed sensing on a random array," *IEEE Journal of Oceanic Engineering*, vol. 42, no. 4, pp. 914–922, 2017.

[3] J. B. Harley and J. M. Moura, "Dispersion curve recovery with orthogonal matching pursuit," *The Journal of the Acoustical Society of America*, vol. 137, no. 1, pp. EL1–EL7, 2015.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[6] Y. C. Eldar and G. Kutyniok, *Compressed sensing: theory and applications*. Cambridge university press, 2012.

[7] A. Drémeau, C. Herzet, and L. Daudet, "Boltzmann machine and mean-field approximation for structured sparse decompositions," *IEEE Transactions On Signal Processing*, vol. 60, no. 7, p. 3425–3438, 2012.

[8] P. Smolensky, "Information processing in dynamical systems: Foundations of harmony theory," Colorado Univ at Boulder Dept of Computer Science, Tech. Rep., 1986.

[9] N. Le Roux and Y. Bengio, "Representational power of restricted boltzmann machines and deep belief networks," *Neural computation*, vol. 20, no. 6, pp. 1631–1649, 2008.

[10] T. Tieleman, "Training restricted boltzmann machines using approximations to the likelihood gradient," in *Proc. 25th international conference on Machine learning*, 2008, pp. 1064–1071.

[11] G. E. Hinton, "A practical guide to training restricted boltzmann machines," in *Neural networks: Tricks of the trade*. Springer, 2012, pp. 599–619.

[12] A. Fischer and C. Igel, "Training restricted boltzmann machines: An introduction," *Pattern Recognition*, vol. 47, no. 1, pp. 25–39, 2014.

[13] E. W. Tramel, A. Manoel, F. Caltagirone, M. Gabrié, and F. Krzakala, "Inferring sparsity: Compressed sensing using generalized restricted boltzmann machines," in *2016 IEEE Information Theory Workshop (ITW)*. IEEE, 2016, pp. 265–269.

[14] L. F. Polania and K. E. Barner, "Exploiting restricted boltzmann machines and deep belief networks in compressed sensing," *IEEE Transactions on Signal Processing*, vol. 65, no. 17, pp. 4538–4550, 2017.

[15] Z. Liao, J. Zhang, D. Hu, C. Li, L. Zhu, and Y. Li, "Sparse signal reconstruction with statistical prior information: A data-driven method," *IEEE Access*, vol. 7, pp. 157 037–157 045, 2019.

[16] F. Le Courtois and J. Bonnel, "Compressed sensing for wideband wavenumber tracking in dispersive shallow water," *The Journal of the Acoustical Society of America*, vol. 138, no. 2, pp. 575–583, 2015.

[17] T. Paviet-Salomon, C. Dorffer, J. Bonnel, B. Nicolas, T. Chonavel, and A. Drémeau, "Dispersive grid-free orthogonal matching pursuit for modal estimation in ocean acoutics," *Proc. Int'l Conference on acoustics, speech and signal processing*, 2019.

[18] T. Peleg, Y. Eldar, and M. Elad, "Exploiting statistical dependencies in sparse representations for signal recovery," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2286–2303, 2012.

[19] V. Cevher, M. Duarte, C. Hegde, and R. Baraniuk, "Sparse signal recovery using markov random fields," *Proc. Neural Information Processing Systems (NeurIPS)*, 2008.

[20] J. Garrigues and B. A. Olshausen, "Learning horizontal connectionsin a sparse coding model of natural images," *Proc. Neural Information Processing Systems (NeurIPS)*, 2008.

[21] F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational ocean acoustics*. Springer Science & Business Media, 2011.

[22] D. Jackson, "High-frequency ocean environmental acoustic models handbook," *Tech. report, Applied Physics Laboratory, Univ. of Washington*, 1994.