ITERATIVE CHANNEL ESTIMATION AND DATA DETECTION ALGORITHM FOR OTFS MODULATION

Rabah Ouchikh[†] *Abdeldjalil Aïssa-El-Bey*[⋆] *Thierry Chonavel*[⋆] *Mustapha Djeddou*[†]

[†] Laboratoire Télécommunications, Ecole Militaire Polytechnique, Bordj El-Bahri, Algeria
[⋆]IMT Atlantique, Lab-STICC, UMR CNRS 6285, F-29238 Brest, France

ABSTRACT

In this paper, we design an iterative channel estimation and data detection algorithm in delay-Doppler domain for orthogonal time frequency space (OTFS) system by taking advantage of the sparse nature of the channel in this domain. The proposed algorithm iterates between message-passing-aided data detection and data-aided channel estimation. This sparse channel estimation is reformulated as a specific marginalization of maximum a posteriori (MAP) problem. To deal with the intractability of this problem, we provide a Bayesian approach based on the variational mean-field approximation via the variational Bayesian expectation maximization (VB-EM) algorithm. Finally, we compare the complexity and performance in term of Bit Error Rate (BER) and Normalized Mean Square Error (NMSE) of the proposed solution to a reference solution in the literature (SP-I).

***Index Terms*—** OTFS, channel estimation, Iterative algorithm, Bayesian approach.

1. INTRODUCTION

The OTFS modulation technique proposed in [1, 2] offers significant advantages over OFDM modulation. OTFS shows its robustness in high Doppler scenarios. The genius of OTFS is the transmission of data symbols in the delay-Doppler domain as opposed to OFDM where transmitted data are in the time-frequency domain [3, 4]. OTFS transforms a doubly-selective channel into a time-invariant one in the delay-Doppler domain, which makes it possible to reduce pilot overhead for estimating rapidly time-varying channel. In the delay-Doppler domain, the channel is sparse. This sparsity can be used to reduce the complexity of data detection and channel estimation algorithms [5, 6].

Several channel estimation and data detection schemes in the delay-Doppler domain have been proposed recently in the literature. Estimation schemes based on pilots pulses in delay-Doppler domain for OTFS systems have been proposed in [4, 7, 8, 9]. A channel estimation based on a thresholding and an appropriate arrangement for pilots, data and guard interval in delay-Doppler domain have been proposed in [6]. In each OTFS frame, pilots, guard and data symbols are arranged appropriately in the delay-Doppler domain in order to avoid interference between the data symbols and the pilots at the receiver side. This configuration decreases the spectral efficiency of the system because a lot of space is reserved for the guard interval in the delay-Doppler grid. Another disadvantage of this method is that it uses a very powerful pilot, which increases peak-to-average power ratio (PAPR). A low complexity Message Passing (MP) algorithm for data detection in OTFS systems have been proposed in [10]. A superimposed pilot (SP)-based channel estimation and data detection have been proposed in [11]. In [11], data and pilots are superimposed in the same locations in delay-Doppler domain. This configuration allows pilots energy to be scattered over all the delay-Doppler domain and high spectral efficiency compared to previous schemes where data and pilots are separated by guard interval. In this approach, channel estimation and data detection are improved alternatively in an iterative process.

In the present paper, we propose an iterative channel estimation and data detection algorithm for OTFS systems with superimposed pilots. The main contributions are summarized as follows:

(i) We propose an iterative channel estimation and data detection algorithm using a superimposed pilot, where channel estimation step is based on the variational mean-field approximation via the VB-EM algorithm, while data detection step is based on MP algorithm.

(ii) We study the performance of the proposed algorithm in terms of NMSE and BER as well as its complexity and we compare the results obtained with the method proposed in [11]. We show that our algorithm has almost the same performance as SP-I with slightly lower complexity and without need for any prior knowledge about the channel, unlike SP-I which assumes prior knowledge of channel taps.

Nomenclature: Hadamard product and Kronecker product are represented by $\odot$ and $\otimes$, respectively. Operators $\mathrm{vec}(.)$ and $\mathrm{vec}^{-1}(.)$ denote the column vectorization of an $M \times N$ matrix into an MN column vector and the invectorization of an MN column vector to an $M \times N$ matrix, respectively.

Finally, $\boldsymbol{F}_n$, $\boldsymbol{F}_n^H$ and $\boldsymbol{I}_M$ denote the n-point DFT matrix, the n-point IDFT matrix, and the $M \times M$ identity matrix.

2. SYSTEM MODEL

In this section, we first recall the input/output equations of the OTFS system with a superimposed pilot. Then, we formulate the channel estimation and data detection problems.

2.1. OTFS input/output with superimposed pilot

Let NT, T/M, $g_{tx}(t)$ and $g_{rx}(t)$ denote the total duration of the transmitted signal frame, the sampling interval, the transmitted waveform of duration T, and the received filer impulse response. The transmitted signal $\boldsymbol{s}$ can be expressed in a matrix form $\boldsymbol{S}$ as follows:

$$\boldsymbol{S} = \boldsymbol{G_{tx}} \boldsymbol{F_M^H} (\boldsymbol{F_M} \boldsymbol{X} \boldsymbol{F_N^H}) = \boldsymbol{G_{tx}} \boldsymbol{X} \boldsymbol{F_N^H} \tag{1}$$

where $\boldsymbol{G}_{tx} = \mathrm{diag}(g_{tx}(0), g_{tx}(T/M), ..., g_{tx}((M-1)T/M)) \in \mathbb{C}^{M\times M}$ and $\boldsymbol{X} \in \mathbb{C}^{M\times N}$ is the two-dimensional superimposed symbols transmitted in the delay-Doppler domain, which can be written as $\boldsymbol{X} = \boldsymbol{X}_d + \boldsymbol{X}_p$. Where $\boldsymbol{X}_d$ and $\boldsymbol{X}_p$ are delay-Doppler data and pilot matrices formed by the arrangement of $x_d[k,l]$ and $x_p[k,l]$ for $k = 0 : N-1$ and $l = 0 : M-1$. The elements of $\boldsymbol{X}_p$ and $\boldsymbol{X}_d$ are assumed independent and identically distributed (i.i.d.) with zero-mean and we note $\mathbb{E}\{|x_p[k,l]|^2\} = \sigma_p^2$ and $\mathbb{E}\{|x_d[k,l]|^2\} = \sigma_d^2$ for $k = 0 : N-1$ and $l = 0 : M-1$. We also assume that $\mathbb{E}\{|x[k,l]|^2\} = \sigma_d^2 + \sigma_p^2 = \sigma_x^2$.

The channel is sparse in delay-Doppler domain and its response can be expressed as $h(\tau,\nu) = \sum_{i=1}^{P} h_i \delta(\tau-\tau_i)\delta(\nu-\nu_i)$, where P is the number taps, h_i, τ_i and ν_i represent the complex gain, delay, and Doppler shift associated with the ith path. The delay and the Doppler associated with the ith path are expressed as $\tau_i = \frac{l_i}{M\Delta f}$, $\nu_i = \frac{k_i}{NT}$.

In vectorized form, the received signal $\boldsymbol{Y}$ in delay-Doppler domain can be expressed as follows:

$$\boldsymbol{y} = \boldsymbol{H}(\boldsymbol{x}_d + \boldsymbol{x}_p) + \widetilde{\boldsymbol{w}} = \boldsymbol{H}\boldsymbol{x} + \widetilde{\boldsymbol{w}}, \tag{2}$$

where $\boldsymbol{y} = \mathrm{vec}^{-1}(\boldsymbol{Y})$, $\boldsymbol{H}$ is a sparse matrix expressed as $\boldsymbol{H} = (\boldsymbol{F}_N \otimes \boldsymbol{I}_M)(\sum_{i=1}^{P} h_i \boldsymbol{\Pi}^{l_i} \boldsymbol{\Delta}^{k_i})(\boldsymbol{F}_N^H \otimes \boldsymbol{I}_M)$, with $\boldsymbol{\Delta} = \mathrm{diag}(\exp(j2\pi(0)/MN), ..., \exp(j2\pi(MN-1)/MN))$ and $\boldsymbol{\Pi}$ is the permutation matrix (forward cyclic shift). $\widetilde{\boldsymbol{w}} = (\boldsymbol{F}_N \otimes \boldsymbol{I}_M)\boldsymbol{w}$. We note $\boldsymbol{M}_{rx} = \boldsymbol{F}_N \otimes \boldsymbol{I}_M$ and $\boldsymbol{M}_{tx} = \boldsymbol{F}_N^H \otimes \boldsymbol{I}_M$.

2.2. Problem formulation

Channel estimation in the delay-Doppler domain amounts to estimate $\{h_i, \tau_i, \nu_i\}_{i=1}^{P}$, while data detection is the determination of $\boldsymbol{x}_d$. From (2), we have

$$\boldsymbol{y} = \boldsymbol{y}_p + \boldsymbol{y}_d + \widetilde{\boldsymbol{w}}, \tag{3}$$

where $\boldsymbol{y}_p = \boldsymbol{H}\boldsymbol{x}_p$ and $\boldsymbol{y}_d = \boldsymbol{H}\boldsymbol{x}_d$. The elements of $\boldsymbol{y}_p$ can be expressed as a circular convolution as follows:

$$y_p[k,l] = \sum_{k'=-k_\nu}^{k_\nu} \sum_{l'=0}^{l_\tau} b_{k',l'} h_{k',l'} \alpha_{k,l} x_p[[k-k']_N, [l-l']_M], \tag{4}$$

where k_ν and l_τ represent the maximum Doppler tap and the maximum delay tap, respectively. $b_{k',l'} \in \{0,1\}$ is the path indicator and $\alpha_{k,l}$ represents an additional phase shift caused by the imperfection of the rectangular waveform, where $\alpha_{k,l} = \exp(j2\pi(l-l)k'/MN)$ if $l' \leq l < M$ and $\exp(j2\pi((l-l)k'/MN)\exp(-j2\pi k/N)$ elsewhere [12].

Equation (4), can be written in the following vector form:

$$\boldsymbol{y_p} = (\boldsymbol{S}_p \odot \boldsymbol{\Psi})\boldsymbol{h} = \boldsymbol{A}\boldsymbol{h}, \tag{5}$$

where $\boldsymbol{\Psi} \in \mathbb{C}^{MN\times L}$ is an additional phase shift matrix given by $[\boldsymbol{\Psi}]_{(i,l'(2k_\nu+1)+k'+k_\nu)} = \alpha_{k,l}$, while $\boldsymbol{S}_p \in \mathbb{C}^{MN\times L}$ represents the pilots matrix, with $[\boldsymbol{S}_p]_{(i,l'(2k_\nu+1)+k'+k_\nu)} = x_p[[k-k']_N, [l-l']_M]$. $\boldsymbol{h} \in \mathbb{C}^L$ is a sparse channel vector containing only P non-zero elements and $L = (2k_\nu+1)(l_\tau+1)$. Thus, equation (2) can be written as follows:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{H}\boldsymbol{x}_d + \widetilde{\boldsymbol{w}}. \tag{6}$$

Finally, channel estimation in this context amounts to finding the channel support $\{l_i, k_i\}_{i=1}^{P}$ as well as the corresponding path parameters $\{h_i, \tau_i, \nu_i\}_{i=1}^{P}$. Whereas, data detection amounts to finding data vector $\boldsymbol{x}_d$ from (6).

3. PROPOSED ITERATIVE CHANNEL ESTIMATION AND DATA DETECTION ALGORITHM

In this section, we detail the proposed algorithm. This algorithm iterate between data-assisted channel estimation and MP-assisted data detection.

3.1. Initial channel estimation

In this part of algorithm, we calculate an initial estimate of the channel. Equation (6) can be written as follows:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{v} = \sum_{i=1}^{L} b_i g_i \boldsymbol{A}_i + \boldsymbol{v}, \tag{7}$$

where $\boldsymbol{v} = \boldsymbol{H}\boldsymbol{x}_d + \widetilde{\boldsymbol{w}}$. It is shown in [11] that the mean of $\boldsymbol{v}$ is expressed as $\boldsymbol{\mu_v} = \mathbb{E}\{\boldsymbol{v}\} = \boldsymbol{0}_{MN\times 1}$ and its covariance matrix $\boldsymbol{C_v} = \mathbb{E}\{\boldsymbol{v}\boldsymbol{v}^H\} = \left(\left(\sum_{i=1}^{P} \sigma_{h_i}^2\right)\sigma_d^2 + \sigma_w^2\right)\boldsymbol{I}_{MN}$. $\boldsymbol{h} = \boldsymbol{b} \odot \boldsymbol{g}$, i.e. $h_k = b_k g_k$, where $\boldsymbol{b} = [b_1, b_2, ..., b_L]^T$ is the support vector and $\boldsymbol{g} = [g_1, g_2, ..., g_L]^T$ is the channel gains vector. $\boldsymbol{A}_i$ represents the ith column of $\boldsymbol{A}$. Therefore, $p(\boldsymbol{y}|\boldsymbol{g},\boldsymbol{b}) = \mathcal{CN}(\boldsymbol{A}_b\boldsymbol{g}_b, \sigma_v^2\boldsymbol{I}_L)$, where $\boldsymbol{A}_b \in \mathbb{C}^{MN\times P}$ and

$\boldsymbol{g}_b \in \mathbb{C}^P$ are made up from $\boldsymbol{A}$ and $\boldsymbol{g}$ considering indices i where $b_i \neq 0$.

To take into account the fact that most of the elements of $\boldsymbol{h}$ are zero except for P of them which are non-zero, we model its elements by a Bernoulli-Gaussian model. Therefore, $\boldsymbol{g}$ obeys the following model: $p(\boldsymbol{g}) = \prod_{i=1}^{L} p(g_i)$, where $p(g_i) = \mathcal{CN}(g_i; 0, \sigma_{g_i}^2)$ and $\boldsymbol{b}$ is modelled as $p(\boldsymbol{b}) = \prod_{i=1}^{L} p(b_i)$, where $p(b_i) = \mathrm{Ber}(p_i)$, where $p_i = p(b_i = 1) = 1 - p(b_k = 0)$.

We derive an estimator under a Maximum A Posteriori (MAP) criterion, which is the optimal Bayesian estimator using a Bayesian cost. Therefore, the estimation of $\boldsymbol{b}$ and $\boldsymbol{g}$ can take the form

$$(\hat{\boldsymbol{g}}, \hat{\boldsymbol{b}}) = \arg\max_{\boldsymbol{g},\boldsymbol{b}} \log(p(\boldsymbol{g}, \boldsymbol{b}|\boldsymbol{y})). \tag{8}$$

We start with the estimation of support vector $\boldsymbol{b}$, which can be made from marginalized MAP, leading to

$$\hat{b}_i = \arg\max_{b_i \in \{0,1\}} \log(p(b_i|\boldsymbol{y})). \tag{9}$$

The evaluation of $p(b_i|\boldsymbol{y})$ requires a costly marginalization of $p(\boldsymbol{b}|\boldsymbol{y})$. To avoid this problem, we use the VB-EM algorithm [13] that involves a tractable surrogate $q(b_i)$ of $p(b_i|\boldsymbol{y})$ via the mean-field variational approximation.

By using this approximation, (9) can easily be solved by simple thresholding operation, i.e, $\hat{b}_i = 1$ if $q(b_i = 1) > \rho$ and $\hat{b}_i = 0$ otherwise, with $\rho = 0.5$. The estimation of the sparse vector of channel coefficients $\boldsymbol{g}$ is given as its MAP estimate as

$$\hat{\boldsymbol{g}} = \arg\max_{\boldsymbol{g}} \log(p(\boldsymbol{g}|\hat{\boldsymbol{b}}, \boldsymbol{y})). \tag{10}$$

Letting $\hat{\boldsymbol{g}}_{\hat{\boldsymbol{b}}}$ denote the entries of $\boldsymbol{g}$ limited to the estimated channel support and $\boldsymbol{A}_{\hat{\boldsymbol{b}}}$ the corresponding columns of $\boldsymbol{A}$, the solution of this problem is given by

$$\begin{array}{rll} & \hat{\boldsymbol{g}}_{\hat{\boldsymbol{b}}} = & (\boldsymbol{A}_{\hat{\boldsymbol{b}}}^T \boldsymbol{A}_{\hat{\boldsymbol{b}}} + \boldsymbol{\Delta})^{-1} \boldsymbol{A}_{\hat{\boldsymbol{b}}}^T \boldsymbol{y} \\ \text{and} & \hat{g}_i = 0 & \text{if } \hat{b}_i = 0, \end{array} \tag{11}$$

where $\boldsymbol{\Delta} = \mathrm{diag}(\sigma_v^2/\sigma_{g_1}^2, \sigma_v^2/\sigma_{g_2}^2, ..., \sigma_v^2/\sigma_{g_L}^2)$.

3.2. Message-passing (MP) data detection

Once the first channel estimation is done, the pilot signal is subtracted from the received signal $\boldsymbol{y}$ to get $\boldsymbol{y}_d$ which will be used for detection and which is given as follows:

$$\boldsymbol{y}_d = \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{h}} = \hat{\boldsymbol{H}}\boldsymbol{x}_d + \boldsymbol{w}_e \tag{12}$$

where $\boldsymbol{w}_e = \boldsymbol{v} + \boldsymbol{A}(\boldsymbol{h} - \hat{\boldsymbol{h}})$ consist of noise and channel estimation error. The objective here is to estimate the vector of data symbols $\boldsymbol{x}_d$ from $\boldsymbol{y}_d$ and $\hat{\boldsymbol{H}}$. For this purpose, we use a low-complexity message passing (MP) detection algorithm proposed in [10].

3.3. Data-aided Channel estimation

After estimating the data vector $\hat{\boldsymbol{x}}_d$, equation (6) becomes as follows:

$$\boldsymbol{y}_e = \boldsymbol{A}\boldsymbol{h} + \boldsymbol{v}_e, \tag{13}$$

where $\boldsymbol{v}_e = \hat{\boldsymbol{H}}\hat{\boldsymbol{x}}_d + \widetilde{\boldsymbol{w}}$ consists of noise and estimate of channel matrix and data symbols vector. The MP data detection steps yield independent variables for the entries of $\hat{\boldsymbol{x}}_d$. Then, for channel update we propose the Bayesian approach described in subsection (3.1) with the following noise approximation in (7): $\boldsymbol{v}_e \sim \mathcal{N}(\hat{\boldsymbol{H}}\mathbb{E}\{\hat{\boldsymbol{x}}_d\}, \hat{\boldsymbol{H}}\hat{\boldsymbol{R}}_x\hat{\boldsymbol{H}}^H + \sigma_{\widetilde{\boldsymbol{w}}}^2 \boldsymbol{I})$, where $\hat{\boldsymbol{R}}_x = \mathbb{E}\{\hat{\boldsymbol{x}}_{\boldsymbol{d}}\hat{\boldsymbol{x}}_{\boldsymbol{d}}^H\}$ is the covariance matrix of $\hat{\boldsymbol{x}}_d$.

The proposed algorithm for channel estimation and data detection is summarized in Algorithm 1.

Algorithm 1 Iterative channel estimation and data detection algorithm.

Input: measurements $\boldsymbol{y} \in \mathbb{C}^{MN}$, pilot matrix $\boldsymbol{A} \in \mathbb{C}^{MN \times L}$, initial channel estimation $\hat{\boldsymbol{h}}^{(0)}$,

repeat

Compute $\hat{\boldsymbol{H}}^{(i)} = \boldsymbol{M}_{rx} \left(\sum_{n=1}^{P} \hat{h}_n^{(i)} \boldsymbol{\Pi}^{l_n} \boldsymbol{\Delta}^{k_n} \right) \boldsymbol{M}_{tx}$,

Compute $\boldsymbol{y}_d^{(i)} = \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{h}}^{(i)} = \hat{\boldsymbol{H}}^{(i)}\boldsymbol{x}_d + \boldsymbol{v}_e$,

Compute $\hat{\boldsymbol{x}}_d$ by feeding the MP algorithm with $\hat{\boldsymbol{H}}^{(i)}$ and $\boldsymbol{y}_d^{(i)}$,

Compute $\boldsymbol{y}_e^{(i)} = \boldsymbol{y} - \hat{\boldsymbol{H}}^{(i)}\hat{\boldsymbol{x}}_d^{(i)}$,

Compute $\hat{\boldsymbol{h}}^{(i+1)}$ by feeding $\boldsymbol{y}_e^{(i)}$ and $\boldsymbol{A}$ to (13),

until stopping condition

Output: $\hat{\boldsymbol{h}}$, $\hat{\boldsymbol{x}}_d$.

---

4. SIMULATION RESULTS AND COMPLEXITY ANALYSIS

In this section, we first set the simulation parameters. Then, we study the performance of the proposed algorithm in terms of NMSE and BER by comparing it with the state of the art method [11]. Finally, we study the complexity of the proposed solution.

4.1. Simulation parameters

We consider the following simulation parameters: the carrier frequency is 4 GHz and spacing between sub-carriers is 15 kHz. The OTFS frame used is of size $M = N = 16$. A rectangular pulse and BPSK modulation are considered. Parameters of the 5-tap channel used here are given in Table 1 of [11]. Pilot and data power used here are $\sigma_p^2 = 0.3\sigma_x^2$ and $\sigma_d^2 = 0.7\sigma_x^2$, respectively. It has been shown in [11] that $\sigma_p^2 = 0.3\sigma_x^2$ is an optimal value that minimises the BER. The proposed algorithm terminates when $|\hat{\boldsymbol{h}}^{(n)} - \hat{\boldsymbol{h}}^{(n+1)}| < \epsilon$ ($\epsilon = 10^{-3}$) or when the number of iterations reaches 10.

4.2. Normalized Mean square error (NMSE) versus SNR

The NMSE expression we used is given as NMSE (dB) $= 10\log_{10}(1-(|\boldsymbol{h}^H\hat{\boldsymbol{h}}|/\|\boldsymbol{h}\|_2\|\hat{\boldsymbol{h}}\|_2)^2)$.

Fig. 1 compares NMSE for the proposed iterative channel estimation and data detection algorithm with the SP-NI and SP-I designs proposed in [11]. It is very clear that SP-I has better performance than SP-NI, which is the non-iterative version used for its initialization. We observe that the NMSE of our proposed algorithm is close to that of the SP-I design. It should be noted that unlike SP-I the channel estimation performed in our work does not assume prior knowledge of channel delays and Doppler frequencies (l_i, k_i). In the SP-I algorithm (l_i, k_i) are assumed constant for several OTFS frames, and thus to estimate them, a super-frame architecture is used. In the first frame, the arrangement of pilot, data and guard interval proposed in [6] is used. The estimation of (l_i, k_i) in this first frame is done using a thresholding. Thus, in [11], once (l_i, k_i) are estimated, the estimation of h_i is performed on the other frames of the super-frame architecture.

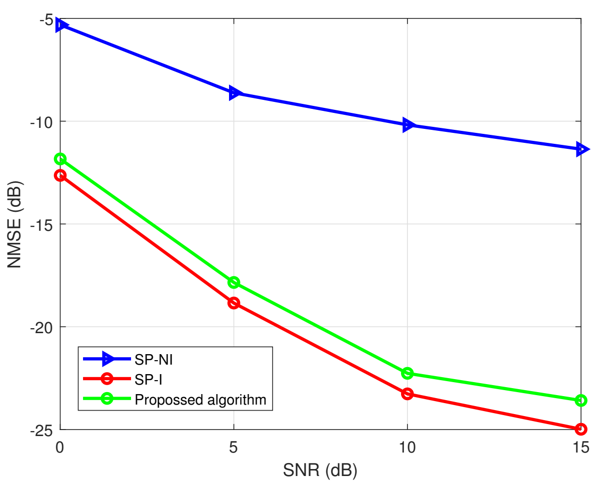

Fig. 1. NMSE comparison for the proposed algorithm and existing SP-I and SP-NI designs.

4.3. Bit Error Rate (BER) versus SNR

In Fig. 2, we perform a comparison of BER versus SNR for our proposed algorithm, SP-NI, SP-I designs and OTFS with known channel state information (CSI). It can be seen from Fig. 2 that the BER of our proposed algorithm is close to that of SP-I. This small marginal difference is due to the small difference in the NMSE because we use the same MP detector. We note that both proposed solution and SP-I algorithm are close to the oracle.

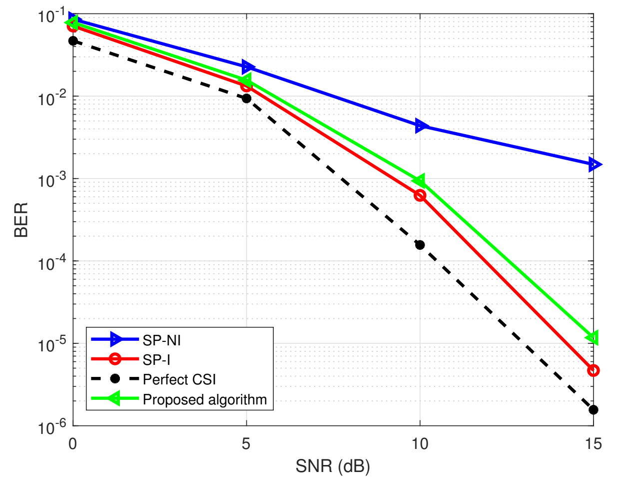

Fig. 2. BER comparison for the proposed algorithm and existing SP-I, SP-NI designs and oracle.

4.4. Complexity analysis

For the channel estimation step, the computational coast per iteration is $\mu_e = (4P^2+6L)MN+(P^3+L)$. For the MP detection step, the overall cost over n_{iter} is $\mu_d = n_{iter}NMSP$. The overall complexity of the proposed algorithm is $C = \mathcal{O}(\mu_e) + \mathcal{O}(\mu_d) + \mathcal{O}(\mu_i)$, where $\mu_i = \mu_e$ is the complexity of the first estimate. In practice, we have $P \ll MN$, the overall complexity of the proposed algorithm is $\mu = (N_{iter}+1)\mathcal{O}(MN) + N_{iter}\mathcal{O}(n_{iter}NMSP)$. The complexity of our algorithm is slightly lower than that of SP-I which is given by $\mu_{SPI} = (N_{SPI}+1)\mathcal{O}(MN)+(N_{SPI}+1)\mathcal{O}(n_{iter}NMSP)$. We note that the number of iterations N_{iter} and N_{SPI} after convergence of the two algorithms are of the same order of magnitude. Finally, we observe a difference in overall computational cost of $< 1\%$ between the two methods and we manage to estimate the channel without significant additional computational cost compared to [11] where the delay and Doppler taps of the channel are known.

5. CONCLUSION

In this paper, we have developed an iterative algorithm for channel estimation and data detection. We addressed the channel estimation step from a Bayesian point of view, using mean-field approximation as well as VB-EM algorithm. For the detection step we used the low complexity MP algorithm proposed in OTFS literature. In addition to the gain in terms of spectral efficiency compared to schemes where pilots and data are separated by guard intervals, our proposed algorithm shows BER and MSE performance close to those of the SP-I algorithm proposed in the literature and this without any prior assumption upon channel taps knowledge. Our algorithm is also less complex compared to the SP-I algorithm.

6. REFERENCES

[1] Anton Monk, Ronny Hadani, Michail Tsatsanis, and Shlomo Rakib, "OTFS-orthogonal time frequency space," *arXiv preprint arXiv:1608.02993*, 2016.

[2] Ronny Hadani, Shlomo Rakib, Michail Tsatsanis, Anton Monk, Andrea J Goldsmith, Andreas F Molisch, and R Calderbank, "Orthogonal time frequency space modulation," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2017, pp. 1–6.

[3] GD Surabhi, Rose Mary Augustine, and A Chockalingam, "On the diversity of uncoded OTFS modulation in doubly-dispersive channels," *IEEE Transactions on Wireless Communications*, vol. 18, no. 6, pp. 3049–3063, 2019.

[4] M Kollengode Ramachandran and A Chockalingam, "MIMO-OTFS in high-doppler fading channels: Signal detection and channel estimation," in *IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 206–212.

[5] Patchava Raviteja, Khoa T Phan, Yi Hong, and Emanuele Viterbo, "Interference cancellation and iterative detection for orthogonal time frequency space modulation," *IEEE Transactions on Wireless Communications*, vol. 17, no. 10, pp. 6501–6515, 2018.

[6] Patchava Raviteja, Khoa T Phan, and Yi Hong, "Embedded pilot-aided channel estimation for OTFS in delay–doppler channels," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4906–4917, 2019.

[7] Mingchen Zhang, Fanggang Wang, Xiaojun Yuan, and Lei Chen, "2D structured turbo compressed sensing for channel estimation in OTFS systems," in *IEEE International Conference on Communication Systems (ICCS)*, 2018, pp. 45–49.

[8] Yushan Liu, Shun Zhang, Feifei Gao, Jianpeng Ma, and Xianbin Wang, "Uplink-aided high mobility downlink channel estimation over massive MIMO-OTFS system," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 9, pp. 1994–2009, 2020.

[9] Wenqian Shen, Linglong Dai, Jianping An, Pingzhi Fan, and Robert W Heath, "Channel estimation for orthogonal time frequency space (OTFS) massive MIMO," *IEEE Transactions on Signal Processing*, vol. 67, no. 16, pp. 4204–4217, 2019.

[10] Patchava Raviteja, Khoa T Phan, Qianyu Jin, Yi Hong, and Emanuele Viterbo, "Low-complexity iterative detection for orthogonal time frequency space modulation," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2018, pp. 1–6.

[11] Himanshu B Mishra, Prem Singh, Abhishek K Prasad, and Rohit Budhiraja, "OTFS channel estimation and data detection designs with superimposed pilots," *IEEE Transactions on Wireless Communications*, 2021.

[12] Siqiang Wang, Jing Guo, Xinyi Wang, Weijie Yuan, and Zesong Fei, "Pilot design and optimization for OTFS modulation," *IEEE Wireless Communications Letters*, 2021.

[13] Angélique Drémeau, Cédric Herzet, and Laurent Daudet, "Boltzmann machine and mean-field approximation for structured sparse decompositions," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3425–3438, 2012.