Message Passing Based Detection for Orthogonal Time Frequency Space Modulation

YUAN Zhengdao[1], LIU Fei[2], GUO Qinghua[3], WANG Zhongyong[2]

(1. The Open University of Henan, Zhengzhou 450000, China；
2. Zhengzhou University, Zhengzhou 450001, China；
3. University of Wollongong, Wollongong NSW 2522, Australia)

Abstract: The orthogonal time frequency space (OTFS) modulation has emerged as a promising modulation scheme for wireless communications in high-mobility scenarios. An efficient detector is of paramount importance to harvesting the time and frequency diversities promised by OTFS. Recently, some message passing based detectors have been developed by exploiting the features of the OTFS channel matrices. In this paper, we provide an overview of some recent message passing based OTFS detectors, compare their performance, and shed some light on potential research on the design of message passing based OTFS receivers.

Keywords: OTFS; detection; message passing; belief propagation; approximate message passing (AMP); unitary AMP (UAMP)

1 Introduction

Recently the orthogonal time frequency space (OTFS) modulation has attracted much attention due to its capability of achieving reliable communications in high mobility scenarios[1 – 6]. OTFS is a two-dimensional modulation scheme, and the information is modulated in the delay Doppler (DD) domain, which is in contrast to the time frequency (TF) domain modulation in the orthogonal frequency division multiplexing (OFDM). In OTFS, each symbol spreads over the time and frequency domains through the two dimensional (2D) inverse symplectic finite Fourier transform (SFFT), leading to both time and frequency diversities[1 – 2]. It has been shown that OTFS can significantly outperform OFDM in high mobility scenarios[7].

To harvest the diversities promised by OTFS, the design of a powerful detector is paramount. The optimal maximum a posteriori (MAP) detector is impractical due to its complexity growing exponentially with the length of the OTFS block. Recently, significant efforts have been devoted to the design of more efficient detectors. In Ref. [8], an effective channel matrix in the DD domain was derived, based on which a low-complexity two-stage detector was proposed. The first-order Neumann series was used in Ref. [9] to approximately solve the matrix inverse problem involved in the linear minimum mean squared error (MMSE) estimation based detection. A detection scheme was developed in Ref. [10], where the MMSE equalization was used in the first iteration, followed by parallel interference cancellation with a soft-output sphere decoder in subsequent iterations. A rectangular waveform was considered in Ref. [11], where the sparsity and quasi banded structure of channel matrices without fractional Doppler shifts were exploited to reduce the detection complexity. The linear equaliz-

ers were extended to the multiple input and multiple output (MIMO)-OTFS systems in Ref. [12]. A cross-domain method was proposed in Ref. [13], where a conventional linear MMSE estimator is adopted for equalization in the time domain and a low-complexity symbol-by-symbol detection is utilized in the DD domain. A low complexity iterative rake decision feedback detector was proposed in Ref. [14], which extracts and coherently combines the multiple copies of the symbols (due to multipath propagation) in the DD grid using maximal ratio combining (MRC).

Another line of OTFS detector design is based on factor graphs and message passing techniques[15, 23]. When the number of channel paths is small, the effective channel matrix in the DD domain is sparse, which allows efficient detection using the message passing algorithm (MPA)[2]. An expectation propagation (EP) algorithm was proposed in Ref. [16], where EP is used for message update with Gaussian approximation. A variational Bayes (VB) based detector was proposed in Ref. [17] to achieve better convergence. Studying the matched filtering processing, the authors in Ref. [18] proposed a message passing detector, which is combined with a probability clipping solution. The detectors in Refs. [2, 17, 19] take advantage of the sparsity of the channel matrix in the DD domain, and their complexity depends on the number of nonzero elements in each row of the channel matrix, which is denoted by S. Without considering fractional Doppler shifts, S is equal to the number of channel paths. In general, a wideband system can provide sufficient delay resolution. The Doppler resolution depends on the time duration of the OTFS block. To fulfill the low latency requirement in wireless communications, the time duration of an OTFS block should be relatively small, where it is necessary to consider fractional Doppler shifts[2, 20]. In this case, the value of S can be significantly larger than the number of channel paths. In the case of rich scattering environments, the complexity of these detectors can be a concern and the short loops in the corresponding system graph model may result in significant performance. To overcome the above issues, the design of OTFS detectors based approximate message passing (AMP)[21 – 22] was investigated in Ref. [25]. AMP works well for independent and identically distributed (sub-) Gaussian system transfer matrix, but it suffers from performance loss or even diverges for a general system transfer matrix[27 – 29]. Instead, the works in Refs. [25 – 26] resort to the unitary AMP (UAMP)[27 – 29], which is a variant of AMP and was formerly called UTAMP[27]. In UAMP, a unitary transformation of the original model is used, where the unitary matrix for transformation can be the conjugate transpose of the left singular matrix of the general system transfer matrix[27] obtained through singular value decomposition (SVD). It is shown in Ref. [25] that UAMP is well suitable for OTFS due to the structure of block circulant matrix with circulant block (BCCB) of the DD domain channel matrices, where the 2D discrete Fourier transform is used for the unitary transformation,

leading to very efficient implementation using the 2D fast Fourier transform (FFT) algorithm. In addition, as the noise variance is normally unknown, the noise variance estimation is also incorporated into the UAMP-based detector in Ref. [25].

In this paper, we provide an overview of the message passing based detectors, provide some comparison results, and discuss potential research on the design of message passing based OTFS receivers. The notations used in this paper are as follows. Boldface lower-case and upper-case letters denote vectors and matrices, respectively. We use $(\cdot)^H$ and $(\cdot)^T$ to denote the conjugate transpose and the transpose, respectively. The superscript * denotes the conjugate operation. We define $[\cdot]_M$ as the mod M operation. We use $N(x|\hat{x}, \nu_x)$ to denote the probability density function of a complex Gaussian variable with mean $\hat{x}$ and variance ν_x. The notation $\langle f(\mathbf{x}) \rangle_{q(\mathbf{x})}$ denotes the expectation of the function $f(\mathbf{x})$ with respect to the distribution $q(\mathbf{x})$. The relation $f(x) = c g(x)$ for some positive constant c is written as $f(x) \propto g(x)$. The notation $\otimes$ represents the Kronecker product, and $\mathbf{a} \cdot \mathbf{b}$ and $\mathbf{a} \cdot /\mathbf{b}$ represent the component-wise product and the division between vectors $\mathbf{a}$ and $\mathbf{b}$, respectively. We use $X = \text{reshape}_M(\mathbf{x})$ to denote that the vector $\mathbf{x}$ is reshaped as an $M \times N$ matrix X column by column, where the length of $\mathbf{x}$ is MN, and use $\mathbf{x} = \text{vec}(X)$ to represent vectorization of matrix X column by column. The notation $\text{diag}(\mathbf{a})$ represents a diagonal matrix with the elements of $\mathbf{a}$ as its diagonal. We use $|A|^2$ to denote the element-wise magnitude squared operation for matrix A. The notations $\mathbf{1}$ and $\mathbf{0}$ are used to denote an all-ones vector and an all-zeros vector with a proper length, respectively. The j-th entry of $\mathbf{q}$ is denoted by q_j. The superscript t of s^t denotes the iteration index of the variable s involved in an iterative algorithm.

2 System Model

The modulation and demodulation for OTFS are illustrated in Fig. 1, which are implemented with the 2D inverse SFFT (ISFFT) and SFFT at the transmitter and receiver, respectively[1, 24]. Before the OTFS modulation, a (coded) bit sequence is mapped to symbols $x[k,l], k = 0,...,N - 1, l = 0,...,M - 1$ in the DD domain, where $x[k,l] \in A = \{\alpha_1,...,\alpha_{|A|}\}$, $|A|$ is the cardinality of A, l and k denote the indices of the delay and Doppler shifts, respectively, and N and M are the number of grids
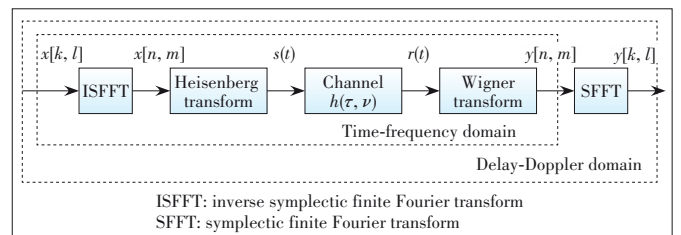
▲ Figure 1. Modulation and demodulation in an orthogonal time frequency space (OTFS) system[2]

of the DD plane. At the transmitter side, ISFFT is performed to convert the DD domain symbols to signals in the time-frequency (TF) domain.

$$X_{tf}[n,m] = \frac{1}{\sqrt{MN}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} x[k,l] e^{j2\pi(\frac{nk}{N} - \frac{ml}{M})}. \tag{1}$$

After that, the signals $X_{tf}[m,n]$ in the TF domain are converted to a continuous-time waveform $s(t)$ using the Heisenberg transform with a transmit waveform $g_{tx}(t)$[2], i.e.,

$$s(t) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} X_{tf}[n,m] g_{tx}(t - nT) e^{j2\pi m\Delta f(t - nT)}, \tag{2}$$

where Δf is subcarrier spacing and $T = 1/\Delta f$. Then the signal $s(t)$ is transmitted over a time-varying channel and the received signal in the time domain is given as:

$$r(t) = \iint h(\tau,\nu) s(t - \tau) e^{j2\pi\nu(t - \tau)} d\tau d\nu, \tag{3}$$

where $h(\tau,\nu)$ is the channel impulse response in the continuous DD domain, and it can be expressed as[1]:

$$h(\tau,\nu) = \sum_{i=0}^{P-1} h_i \delta(\tau - \tau_i) \delta(\nu - \nu_i), \tag{4}$$

with $\delta(\cdot)$ being the Dirac delta function, P being the number of channel paths, and h_i, τ_i and ν_i being the gain, delay and Doppler shift associated with the i-th path, respectively. The delay and Doppler-shift taps for the i-th path are given by

$$\tau_i = \frac{l_i}{M\Delta f}, \quad \nu_i = \frac{k_i + \kappa_i}{NT}, \tag{5}$$

where l_i and k_i are the delay and Doppler indices of the i-th path, and $\kappa_i \in [-1/2, 1/2]$ is a fractional Doppler associated with the i-th path. In the above equation, $M\Delta f$ is the system bandwidth and NT is the duration of an OTFS block.

At the receiver, a receive waveform $g_{rx}(t)$ is used to transform the received signal $r(t)$ to the TF domain, i.e.,

$$Y(t,f) = \int g_{rx}^*(t' - t) r(t') e^{-j2\pi f(t' - t)} dt', \tag{6}$$

which is then sampled at $t = nT$ and $f = m\Delta f$, yielding $Y[n,m]$. Then SFFT is applied to $Y[n,m]$ to generate the DD domain signal $y[k,l]$, i.e.,

$$y[k,l] = \frac{1}{\sqrt{MN}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} Y[n,m] e^{-j2\pi(\frac{nk}{N} - \frac{ml}{M})}. \tag{7}$$

Assuming that the transmitted waveform and the received waveform satisfy the bi-orthogonal property[1], in the DD domain we have the input-output relationship[2].

$$y[k,l] = \sum_{i=0}^{P-1} \sum_{c=-N_i}^{N_i} h_i x\left([k - k_i + c]_N, \right.$$
$$[l - l_i]_M) \frac{1}{N} \frac{1 - e^{-j2\pi(-c - \kappa_i)}}{1 - e^{-j2\pi\frac{-c - \kappa_i}{N}}} e^{-j2\pi\frac{l_i(k_i + \kappa_i)}{MN}} + \omega[k,l] \tag{8}$$

where $N_i < N$ is an integer, and $\omega[k,l]$ is the noise in the DD domain. We can see that for each path, the transmitted signal is circularly shifted, and scaled by a corresponding channel gain. We arrange $\{x[k,l]\}$ as a vector $\boldsymbol{x} \in C^{MN \times 1}$, where the j-th element x_j is $x[k,l]$ with $j = kM + l$. Similarly, a vector $\boldsymbol{y} \in C^{MN \times 1}$ can also be constructed based on $y[k,l]$. Then Eq. (8) can be rewritten in a vector form as:

$$\boldsymbol{y} = \boldsymbol{Hx} + \boldsymbol{\omega}, \tag{9}$$

where $\boldsymbol{H} \in C^{MN \times MN}$ is the effective channel matrix in the DD domain, and $\boldsymbol{\omega}$ denotes a white Gaussian noise with mean 0 and variance ϵ^{-1} (or precision ϵ). The channel matrix $\boldsymbol{H}$ in Eq. (9) can be represented as[25]:

$$\boldsymbol{H} = \sum_{i=0}^{P-1} \sum_{c=-N_i}^{N_i} \boldsymbol{I}_N(-[c - k_i]_N) \otimes [\boldsymbol{I}_M(l_i) h_i \times$$
$$\frac{1 - e^{-j2\pi(-c - \kappa_i)}}{N - Ne^{-j2\pi\frac{-c - \kappa_i}{N}}} e^{-j2\pi\frac{l_i(k_i + \kappa_i)}{MN}}], \tag{10}$$

where $\boldsymbol{I}_N(-[q - k_i]_N)$ denotes an $N \times N$ matrix obtained by circularly shifting the rows of the identity matrix by $-[q - k_i]_N$, and $\boldsymbol{I}_M(l_i)$ is obtained similarly. Without fractional Doppler, i.e., $\kappa_i = 0$, the channel matrix $\boldsymbol{H}$ is reduced to

$$\boldsymbol{H} = \sum_{i=0}^{P-1} \boldsymbol{I}_N(k_i) \otimes \left[\boldsymbol{I}_M(l_i) h_i e^{-j2\pi\frac{l_i k_i}{MN}}\right]. \tag{11}$$

3 Message Passing (MP) Based Detectors

Based on the model (9) in the DD domain, several detectors have been proposed using the message passing techniques.

3.1 MP Detector in Ref. [2]

In model (9), the $MN \times MN$ DD domain complex channel matrix $\boldsymbol{H}$ is sparse (especially in the case without fractional Doppler shifts), which makes belief propagation suitable for implementing the OTFS detectors. In Eq. (2), $\boldsymbol{y}$ and $\boldsymbol{\omega}$ are length-MN complex vectors with elements denoted by $y[d]$ and $\omega[d]$, $1 \leq d \leq MN$, the element of $\boldsymbol{H}$ is denoted by $H[d,c]$, $1 \leq d,c \leq MN$, $\boldsymbol{x}$ is a length-MN symbol vector with elements $x[c] \in A$, $1 \leq c \leq MN$, and $\mathcal{A}$ denotes the modulation alphabet.

Thanks to the sparsity of $\boldsymbol{H}$, the joint distribution of the random variables in model (9) can be represented with a sparsely-connected factor graph with MN variable nodes corresponding to $\boldsymbol{x}$ and MN observation nodes corresponding to $\boldsymbol{y}$. As shown in

Fig. 2, each observation node $y[d]$ is connected to a set of variable nodes $\{x[e_s], e_s \in \mathcal{I}(d)\}$, and similarly, each variable node $x[c]$ is connected to a set of observation nodes $y[e_s], e_s \in \mathcal{J}[c]$, where $\mathcal{I}(d)$ and $\mathcal{J}(c)$ respectively denote the sets of indexes of non-zero elements in the d-th row and c-th columns of H, $|\mathcal{I}(d)| = |\mathcal{J}(c)| = S$ and $1 \leqslant s \leqslant S$. The probability mass function (PMF) $\boldsymbol{p}_{c,e_s} = \{p_{c,e_s}(a_j)|a_j \in A\}$ represents the messages from variable nodes $x[c]$ to factor nodes $y[e_s]$.

Based on the factor graph in Fig. 2, a message passing algorithm was proposed in Ref. [2], and the detector is called MP detector in this paper. The following is a brief derivation of the message computations in the i-th iteration of the message computations.

1) Messages passing from observation node $y[d]$ to variable node $x[e_s]$

The message is approximated to be Gaussian, and the mean μ_{d,e_s}^i and variance $(\sigma_{d,e_s}^i)^2$ are computed as

$$\mu_{d,e_s}^i = \sum_{e \in \mathcal{J}_d, e \neq e_s} \sum_{j=1}^{Q} p_{e,d}^{i-1}(a_j) a_j H[d,e], \tag{12}$$

$$(\sigma_{d,e_s}^i)^2 = \sum_{e \in \mathcal{J}_d, e \neq e_s} \left(\sum_{j=1}^{Q} p_{e,d}^{i-1}(a_j)|a_j|^2 |H[d,e]|^2 - \left| \sum_{j=1}^{Q} p_{e,d}^{i-1}(a_j) a_j H[d,e] \right|^2 \right) + \epsilon^{-1}. \tag{13}$$
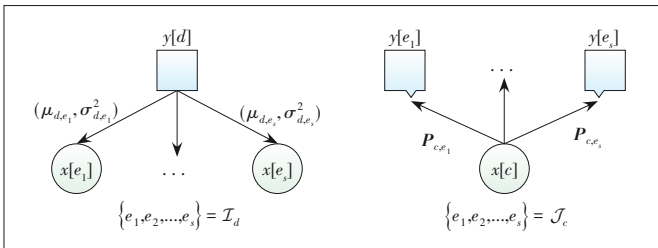
2) Messages passing from variable node $x[c]$ to observation node $y[e_s]$

The PMF $\boldsymbol{p}_{c,d}^i$ can be updated as

$$p_{c,e_s}^i(a_j) = \Delta \cdot \tilde{p}_{c,e_s}^i(a_j) + (1-\Delta) \cdot \tilde{p}_{c,e_s}^{i-1}(a_j), \tag{14}$$

where $\Delta \in [0,1]$ is the damping factor and

$$\tilde{p}_{c,e_s}^i(a_j) \propto \prod_{e \in \mathcal{J}(c), e \neq e_s} Pr(y[e]|x[c] = a_j, H) = \prod_{e \in \mathcal{J}(c), e \neq e_s} \frac{s^i(e,c,j)}{\sum_{k=1}^{Q} s^i(e,c,j)}, \tag{15}$$

▲ Figure 2. Graph representation used to derive the message passing (MP) detector in Ref. [2]

with

$$s^i(e,c,k) = \exp\left(\frac{-\left| y[e] - \mu_{e,c}^i - H[e,c]a_k \right|^2}{(\sigma_{e,c}^i)^2} \right). \tag{16}$$

After a certain number of iterations by repeating 1) and 2), the decision on the transmitted symbol can be obtained, i.e.,

$$\hat{x}[c] = \arg\min_{a_j \in A} p_c^i(a_j), \quad c = 1,\dots,MN, \tag{17}$$

where

$$p_c^i(a_j) = \prod_{e \in \mathcal{J}(c)} \frac{s^i(e,c,j)}{\sum_{k=1}^{Q} s^i(e,c,j)}. \tag{18}$$

The MP detector is summarized in Algorithm 1.

Algorithm 1. MPA detector in Ref. [2]
Input: y, H, Initialize: $p_{c,e_s}^0 = 1/|A|$, $c = 1,\dots,MN$, $e_s \in \mathcal{J}(c)$, $i = 1$
1: **Repeat**
2: $\forall d$: update μ_{d,e_s}^i and $(\sigma_{d,e_s}^i)^2$ with Eqs. (12) and (13)
3: $\forall c$: update $\boldsymbol{p}_{c,d}^i$ with Eq. (14)
4: $i = i + 1$
5: Until terminate
Output: The decision on transmitted symbols $\hat{x}[c]$ using Eq. (17)

The MP algorithm shown above is an approximation to loopy belief propagation since it approximates the interference to be Gaussian to achieve lower complexity. The complexity of the algorithm is $\mathcal{O}(MNS|A|)$ per iteration, which depends on the sparsity of the channel, i.e., the value of S. When S is small, the detector is very attractive because it has low complexity and the detector delivers a good performance as no short loops in the factor graph model. However, in the case of rich-scatting environments and fractional Doppler shifts, the value of S can be large, leading to a denser factor graph model, which can affect the performance of the MP detector and result in a significant increase in computational complexity.

3.2 VB Detector

The VB detector was proposed in Ref. [17] to guarantee the convergence of the iterative detector, which can be implemented with variational message passing. With model (9), the optimal MAP detection can be formulated as:

$$\hat{\boldsymbol{x}} = \arg\max_{\boldsymbol{x}} p(\boldsymbol{x}|\boldsymbol{y}). \tag{19}$$

However, the complexity of solving the above optimization problem increases exponentially with the size of $\boldsymbol{x}$. VB is ad-

opted to achieve low complexity approximate detection. In this method, a distribution $q(\boldsymbol{x})$ from a tractable distribution family $\mathcal{Q}$ is found as an approximation to the a posteriori distribution $p(\boldsymbol{x}|\boldsymbol{y})$. The trial distribution $q(\boldsymbol{x})$ can be obtained by minimizing the Kullback-Leibler divergence $\mathcal{D}(q\|p)$, i.e.,

$$q^*(\boldsymbol{x}) = \underset{q \in \mathcal{Q}}{\arg\max}\, \mathcal{D}(q\|p) =$$
$$\underset{q \in \mathcal{Q}}{\arg\max}\, \underbrace{\mathbb{E}_q[\,-\ln q(\boldsymbol{x}) + \ln p(\boldsymbol{x}|\mathbf{y})\,]}_{\mathcal{L}}, \tag{20}$$

where the expectation is taken over $\boldsymbol{x}$ according to the trial distribution $q(\boldsymbol{x})$.

To simplify the optimization problem, $q(\boldsymbol{x})$ is assumed to be fully factorized, i.e.,

$$q(\boldsymbol{x}) = \prod_{k,l} q_{k,l}(x_{k,l}), \tag{21}$$

where $k \in [0, N-1]$, $M \in [0, M-1]$ and $x_{k,l}$ denotes the $(kM + l)$-th entry of $\boldsymbol{x}$. With this assumption, $q(\boldsymbol{x})$ can be updated iteratively by maximizing $\mathcal{L}$. Since the noise sample $\omega_{k,l}$ and data symbol $x_{k,l}$, $\forall k,l$ are independent, and $\omega_{k,l} \sim \mathcal{CN}(\omega_{k,l}; 0, \epsilon^{-1})$, $p(\boldsymbol{x}|\boldsymbol{y})$ can be rewritten as:

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto \prod_{k,l} p(x_{k,l}) p(y_{k,l}|\boldsymbol{y}), \tag{22}$$

where $y_{k,l} = \boldsymbol{h}_{k,l}^T \boldsymbol{x} + \omega_{k,l}$, $\boldsymbol{h}_{k,l}$ denotes the equivalent channel vector whose $(kM + l)$-th entry is $h_{k,l}[k,l]$. Then the distribution $p(\boldsymbol{x}|\boldsymbol{y})$ can be further rewritten as:

$$p(\boldsymbol{x}|\boldsymbol{y}) \propto \prod_{k,l} \zeta_{k,l}(x_{k,l}) \prod_{k',l'} \psi_{k,l}(x_{k,l}, x_{k',l'}), \tag{23}$$

where

$$\zeta_{k,l}(x_{k,l}) = p(x_{k,l})\exp\left(-\frac{\rho_{k,l}|x_{k,l}|^2 + \eta_{k,l}x_{k,l}}{\epsilon^{-2}}\right), \tag{24}$$

$$\psi_{k,l}(x_{k,l}, x_{k',l'}) = \exp\left(-\frac{\varrho_{k,l,k',l'}x_{k,l}x_{k',l'}}{\epsilon^{-2}}\right), \tag{25}$$

with $\rho_{k,l} = \sum_{k',l'}|h_{k',l'}(k,l)|^2$, $\eta_{k,l} = 2\sum_{k',l'}\mathcal{R}\left[h_{k',l'}[k,l] \cdot y_{k',l'}\right]$, and $\varrho_{k,l,k',l'} = 2\mathcal{R}\left[h_{k,l}[k,l]h_{k,l}^*[k',l']\right]$. Substituting $p(\boldsymbol{x}|\boldsymbol{y})$ in Eq. (23) and $q(\boldsymbol{x})$ into $\mathcal{L}$ yields

$$\mathcal{L} = \mathbb{E}_q\left[\sum_{k,l}\ln\psi_{k,l}(x_{k,l}, x_{k',l'}) - \sum_{k,l}\ln\frac{q_{k,l}(x_{k,l})}{\zeta_{k,l}(x_{k,l})}\right] =$$
$$\mathbb{E}_q\left[-\frac{\sum_{k,l}\varrho_{k,l,k',l'}x_{k,l}x_{k',l'}}{\epsilon^{-2}} - \sum_{k,l}\ln\frac{q_{k,l}(x_{k,l})}{\zeta_{k,l}(x_{k,l})}\right]. \tag{26}$$

To find a stationary point of $\mathcal{L}$, the partial derivations of $\mathcal{L}$ with respect to all local functions $q_{k,l}(x_{k,l})$, $\forall k,l$ need to be zero. Take the latent variable $x_{k,l}$ as an example. Setting the partial derivation $\partial\mathcal{L}/\partial q_{k,l}(x_{k,l})$ to zero leads to:

$$\mathbb{E}_{q\backslash k,l}\left[-\frac{\sum_{k',l'}\varrho_{k,l,k',l'}x_{k,l}x_{k',l'}}{\epsilon^{-2}}\right] + \ln\zeta_{k,l}(x_{k,l}) - \ln q_{k,l}^{iter}(x_{k,l}) + C = 0, \tag{27}$$

where $q_{k,l} = \prod_{(k',l') \neq (k,l)} q_{k',l'}^{iter-1}(x_{k,l})$, $q_{k',l'}^{iter-1}(x_{k,l})$ is obtained in the $(iter - 1)$-th iteration and C denotes a constant.

Then, solving Eq. (27) for $q_{k,l}(x_{k,l})$ results in the local distribution, which can be expressed as:

$$q_{k,l}^{iter}(x_{k,l}) \propto \zeta_{k,l}(x_{k,l})\exp\left(\mathbb{E}_{q\backslash k,l}\left[-\frac{\sum_{k',l'}\varrho_{k,l,k',l'}x_{k,l}x_{k',l'}}{\epsilon^{-2}}\right]\right) \propto p(x_{k,l})\exp\left(-\frac{\rho_{k,l}|x_{k,l}|^2 - m_{k,l}d_{k,l}}{\epsilon^{-2}}\right), \tag{28}$$

where $m_{k,l} = \eta_{k,l} - \sum_{(k',l') \neq (k,l)}\varrho_{k,l,k',l'}\mathbb{E}_{q_{k',l'}^{iter-1}}x[k',l']$.

It is noted that the variance of $x_{k,l}$ is underestimated and only the noise variance is considered in Eq. (28). To fix the underestimation, a practical solution is to repeat the above procedure to approximate the a posteriori distribution for all the data symbols iteratively, resulting in the approximate marginal $q_{k,l}^*(x_{k,l})$, $\forall k,l$. Then, the decision on the symbols can be made by maximizing the approximate marginal distribution $q_{k,l}^*(x_{k,l})$, i.e.,

$$\hat{x}_{k,l} = \underset{x_{k,l} \in A}{\arg\max}\, q_{k,l}^*(x_{k,l}). \tag{29}$$

The complexity of the algorithm per iteration is $\mathcal{O}(MNS|\mathcal{A}|)$.

3.3 UAMP Detector

Leveraging the UAMP algorithm, the UAMP detector was developed in Ref.[25], where the BCCB structure of the DD domain channel matrix is exploited, leading to a highly efficient OTFS detector with 2D FFT. It can be seen from Eqs. (10) and (11) that the DD domain channel matrix $\boldsymbol{H}$ has a BCCB structure. A useful property of the BCCB matrix $\boldsymbol{H}$ is that it can be diagonalized using 2D Discrete Fourier Transform matrix, i.e.,

$$\boldsymbol{H} = \boldsymbol{F}^H \Lambda \boldsymbol{F}, \tag{30}$$

where $\boldsymbol{F} = \boldsymbol{F}_N \otimes \boldsymbol{F}_M$ with $\boldsymbol{F}_N$ and $\boldsymbol{F}_M$ being respectively the normalized N-point and M-point DFT matrices. In Eq. (30),

matrix $\boldsymbol{\Lambda}$ is a diagonal matrix, i.e., $\boldsymbol{\Lambda} = \mathrm{diag}(\boldsymbol{d})$, and $\boldsymbol{d}$ is a length-MN vector that can be computed using 2D FFT.

$$\boldsymbol{d} = \mathrm{vec}(\mathrm{FFT2}(\boldsymbol{C})), \tag{31}$$

where $\mathrm{FFT2}(\cdot)$ represents the 2D FFT operation, $\boldsymbol{C} = \mathrm{reshape}_M\big(\mathbf{H}(:,1)\big)$ is an $M \times N$ matrix, and $\mathbf{H}(:,1)$ with length-MN is the first column of matrix $\boldsymbol{H}$.

The above property is exploited in the design of the UAMP detector, leading to high computational efficiency while with outstanding performance compared with the existing detectors. Instead of using model (9) directly, the UAMP algorithm[27–29] works with the unitary transform of the model. The channel matrix $\boldsymbol{H}$ admits the diagonalization in Eq. (30), leading to the following unitary transform of the OTFS system model:

$$\boldsymbol{r} = \boldsymbol{\Lambda}\boldsymbol{F}\boldsymbol{x} + \boldsymbol{\omega}', \tag{32}$$

where $\boldsymbol{r} = \boldsymbol{F}\boldsymbol{y}$, $\boldsymbol{\omega}' = \boldsymbol{F}\boldsymbol{\omega}$, and the noise $\boldsymbol{\omega}'$ has the same distribution with $\boldsymbol{\omega}$ as $\boldsymbol{F}$ is an unitary matrix. The precision of the noise is still denoted by ϵ, which needs to be estimated. Define $\boldsymbol{\Phi} = \boldsymbol{\Lambda}\boldsymbol{F}$ and an auxiliary vector $\boldsymbol{z} = \boldsymbol{\Phi}\boldsymbol{x}$. Then we can factorize the joint distribution of the unknown variables $\boldsymbol{x}, \boldsymbol{z}, \epsilon$ given $\boldsymbol{r}$ as

$$
\begin{aligned}
p(\boldsymbol{x}, \boldsymbol{z}, \epsilon | \boldsymbol{r}) &= p(\epsilon)\, p(\boldsymbol{r}|\boldsymbol{z}, \epsilon)\, p(\boldsymbol{z}|\boldsymbol{x})\, p(\boldsymbol{x}) = \\
&p(\epsilon) \prod_j p(r_j|z_j, \epsilon)\, p(z_j|\boldsymbol{x}) \prod_i p(x_i) = \\
&f_\epsilon \prod_j f_{r_j}(z_j, \epsilon) f_{\delta_j}(z_j, \mathbf{x}) \prod_i f_{x_i}(x_i)
\end{aligned}, \tag{33}
$$

where indices $i, j \in [1{:}MN]$. To facilitate the factor graph representation of the factorization in Eq. (33), the relevant notations are listed in Table 1, which shows the correspondence between the factor nodes and their associated distributions. The factor graph representation for the factorization in Eq. (33) is depicted in Fig. 3.

Following the UAMP algorithm, a UAMP based iterative detector can be designed, which is summarized in Algorithm 2. According to the derivation of (U)AMP using loopy belief propagation, UAMP provides the message from variable node z_j to function node f_{r_j}, which is Gaussian and denoted by $m_{z_j \to f_{r_j}}(z_j) = \mathcal{N}(z_j | p_j, \nu_{p_j})$. Here, the mean p_j and the variance ν_{p_j} are given in Lines 1 and 2 of the Algorithm in a vector

▼Table 1. Factors, underlying distributions and functional forms associated with Eq. (31)

Factor	Distribution	Function Form
f_{r_j}	$p(r_j|z_j, \epsilon)$	$\mathcal{N}(z_j; r_j, \epsilon^{-1})$
f_{δ_j}	$p(z_j|\mathbf{x})$	$\delta(z_j - \boldsymbol{\Phi}_j \mathbf{x})$
f_{x_i}	$p(x_i)$	$(1/|\mathrm{A}|) \sum_{a=1}^{\mathrm{A}} \delta(x_i - \alpha_a)$
f_ϵ	$p(\epsilon)$	ϵ^{-1}
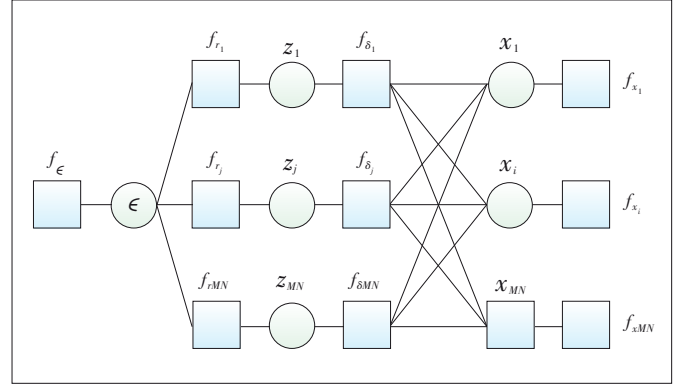
▲Figure 3. Factor graph representation of Eq. (31)

form. With the mean field rule[23] at the function node f_{r_j}, we can compute the message passed from function node f_{r_j} to variable node ϵ, i.e.,

$$
m_{f_{r_j} \to \epsilon}(\epsilon) \propto \exp\left\{\left\langle \log f_{r_j}(r_j|z_j, \epsilon)\right\rangle_{b(z_j)}\right\} \propto \epsilon \exp\left\{-\epsilon(|r_j - \hat{z}_j|^2 + v_{z_j})\right\}, \tag{34}
$$

where $b(z_j)$ is the belief of z_j. It turns out that $b(z_j)$ is also Gaussian with its variance and mean given by

$$
\nu_{z_j} = 1\big/\left(1/\nu_{p_j} + \hat{\epsilon}\right), \quad \hat{z} = \nu_{z_j}\left(p_j/\nu_{p_j} + \hat{\epsilon} r_j\right), \tag{35}
$$

respectively, where $\hat{\epsilon}$ is the estimate of ϵ in the last iteration. They can be expressed in a vector form shown in Lines 3 and 4 in Algorithm 2. The estimate of ϵ can be obtained based on the belief $b(\epsilon)$ at the variable node ϵ shown in Fig. 3, i.e.,

$$
b(\epsilon) \propto f_\epsilon(\epsilon) \prod_{j=1}^{MN} m_{f_{r_j} \to \epsilon}(\epsilon). \tag{36}
$$

And the estimate is given as

$$
\hat{\epsilon} = \int_0^\infty \epsilon\, b(\epsilon)\, d\epsilon = MN\Big/ \sum_{j=1}^{MN}\left(|r_j - \hat{z}_j|^2 + \nu_{z_j}\right), \tag{37}
$$

which can be rewritten in a vector form shown in Line 5 of the algorithm. With the mean field rule at the function node f_{r_j} again, the message passed from the function node f_{r_j} to the variable node z_j can be computed as:

$$
m_{f_{r_j} \to z_j}(z_j) \propto \exp\left\{\left\langle \log f_{r_j}(r_j|z_j, \hat{\epsilon})\right\rangle_{b(\epsilon)}\right\} \propto \mathcal{N}(h_j|r_j, \hat{\epsilon}^{-1}). \tag{38}
$$

Then the UAMP algorithm with known noise can be used as if the true noise precision is $\hat{\epsilon}$, leading to Lines 6–15 and

Lines 1 – 2 of the Algorithm 2. In Lines 10 – 13, the Gaussian message is combined with the discrete prior to obtain the MMSE estimates of the symbols in terms of their posterior means and variances. There is an extra operation in Line 14, which averages the variances of x_j. Thanks to the special form of the unitary matrix $\boldsymbol{F}$, 2D FFT is used in the implementations in Lines 2 and 9. It can be seen that the UAMP detector does not require any matrix-vector products, the algorithm requires only element-wise vector operations or scalar operations, except Lines 2 and 9, which are implemented with FFT. So the complexity of the UAMP detector is $\mathcal{O}(MN\log(MN)) + \mathcal{O}(MN|\mathcal{A}|)$ per OTFS block per iteration, which is independent of S.

Algorithm 2. UAMP detector for OTFS

Unitary transform: $\boldsymbol{r} = \boldsymbol{F}\boldsymbol{y} = \Lambda \boldsymbol{F}\boldsymbol{x} + \boldsymbol{\omega}$ with $\boldsymbol{F} = \boldsymbol{F}_N \otimes \boldsymbol{F}_M$.
Calculated $\boldsymbol{d}$ with Eq. (29), and define vector $\Lambda = \boldsymbol{d} \cdot \boldsymbol{d}^*$.
Initialize $\boldsymbol{s}^{-1} = 0$, $\hat{\boldsymbol{x}} = 0$, $\hat{\epsilon}^{(0)} = 1$, $\nu_x^{(0)} = 1$, and $t = 0$.
Input: $\boldsymbol{y}$, $\boldsymbol{H}$
Repeat
1: $\nu_p = \nu_x^t \Lambda$
2: $\boldsymbol{p} = \boldsymbol{d} \cdot \text{vec}\left(\text{FFT2}\left(\text{reshape}_M(\hat{\boldsymbol{x}}^t)\right)\right) - \nu_{\mathbf{p}} \cdot \boldsymbol{s}^{t-1}$
3: $\nu_z = 1./\left(1./\nu_{\mathbf{p}} + \hat{\epsilon}^t\right)$
4: $\boldsymbol{z} = \nu_z \cdot \left(\boldsymbol{p}./\nu_{\mathbf{p}} + \hat{\epsilon}^t \boldsymbol{r}\right)$
5: $\hat{\epsilon}^{t+1} = MN/\left(\|\boldsymbol{r} - \boldsymbol{z}\|_2^2 + \mathbf{1}^T \nu_{\mathbf{z}}\right)$
6: $\nu_s = 1./\left(\nu_{\mathbf{p}} + 1/\epsilon^{t+1}\mathbf{1}\right)$
7: $\boldsymbol{s}^t = \nu_s \cdot (\boldsymbol{r} - \hat{\boldsymbol{p}})$
8: $\nu_q = \Lambda^T \nu_s/(MN)$
9: $\boldsymbol{q} = \hat{\boldsymbol{x}}^{(t)} + \nu_q \text{vec}\left(\text{IFFT2}\left(\text{reshape}_M(\boldsymbol{d} \cdot \boldsymbol{s}^t)\right)\right)$
10: $\forall j: \xi_{j,a} = \exp\left(-\nu_q^{-1}|\alpha_a - q_j|^2\right)$
11: $\forall j: \beta_{j,a} = \xi_{j,a}/\sum_{a=1}^{|A|}\xi_{j,a}$
12: $\forall j: \hat{x}_j^{t+1} = \sum_{a=1}^{|A|}\alpha_a\beta_{j,a}$
13: $\forall j: \nu_{x_j}^{t+1} = \sum_{a=1}^{|A|}\beta_{j,a}|\alpha_a - \hat{x}_j^{t+1}|^2$
14: $\nu_x^{t+1} = \frac{1}{MN}\sum_{j=1}^{MN}\nu_{x_j}^{t+1}$
15: $t = t + 1$
Until terminated
Output: the estimate of $\boldsymbol{x}$ i.e., $\hat{\boldsymbol{x}}$

Compared with the UAMP detector, the MP and VB detectors have a complexity of $\mathcal{O}(MNS|\mathcal{A}|)$ per OTFS block per iteration, which can be considerably higher than that of the UAMP detector in the case of rich scattering environments and when fractional Doppler shifts have to be considered (leading to a large S). Moreover, the UAMP detector can deliver much better performance when the number of paths is relatively large. In particular, the UAMP detector with estimated noise precision can significantly outperform other detectors with perfect noise precision. We note that,

the OTFS detector can be implemented directly with the AMP algorithm. However, due to the deviation of the channel matrix from the i.i.d. Gaussian matrix, the AMP detector may perform poorly.

4 Turbo Processing in Coded Systems

It is well known that joint decoding and detection can bring significant system performance improvement, and it can be realized in a way that the detector and decoder exchange information iteratively, i.e., the turbo processing[30 – 31]. The OTFS detectors can be incorporated into a turbo receiver by endowing the OTFS detectors with the capabilities of taking the output log-likelihood ratios (LLRs) of the decoder as (soft) input and producing (soft) output in the form of extrinsic LLRs of the coded bits, i.e., the so-called soft input soft output (SISO) detector.

A typical turbo system is shown in Fig. 4, where Π and Π^{-1} represent interleaver and de-interleaver, respectively. The information bits are encoded and interleaved before symbol mapping, where each symbol $x_j \in A = \{\alpha_1,...,\alpha_{|A|}\}$ in the DD domain is mapped from a sub-sequence of the coded bit sequence, which is denoted by $\boldsymbol{c}_j = \left[c_j^1,...,c_j^{\log|A|}\right]$. Each α_a corresponds to a length-$\mathcal{A}\log|\mathcal{A}|$ binary sequence, which is denoted by $\{\alpha_a^1,...,\alpha_a^{\log|A|}\}$. Based on the LLRs provided by the SISO decoder and the output of the OTFS demodulator as shown in Fig. 4, the task of the SISO OTFS detector is to compute the extrinsic LLR for each coded bit, i.e.,

$$L^e(c_j^q) = \ln\frac{P(c_j^q = 0|\boldsymbol{r})}{P(c_j^q = 1|\boldsymbol{r})} - L^a(c_j^q), \tag{39}$$

where $L^a(c_j^q)$ is the output extrinsic LLR of the decoder in the last iteration. The extrinsic LLR $L^e(c_j^q)$ is passed to the decoder. The extrinsic LLR $L^e(c_j^q)$ can be expressed in terms of extrinsic mean and variance of the symbols[32], i.e.,

$$L^e(c_j^q) = \ln\frac{\sum_{\alpha_a \in A_q^0}\exp\left(-\frac{|\alpha_a - m_j^e|^2}{v_j^e}\right)\prod_{q' \neq q}P(c_j^{q'} = \alpha_a^{q'})}{\sum_{\alpha_a \in A_q^1}\exp\left(-\frac{|\alpha_a - m_j^e|^2}{v_j^e}\right)\prod_{q' \neq q}P(c_j^{q'} = \alpha_a^{q'})}, \tag{40}$$

where m_j^e and v_j^e are the extrinsic mean and variance of x_j, and $\mathcal{A}_q^0$ and $\mathcal{A}_q^0$ represent the subsets of all α_a corresponding to $c_j^q = 0$ and $c_j^q = 1$, respectively. The extrinsic variance and mean are defined in Ref. [32].

$$v_j^e = (1/v_j^p - 1/v_j)^{-1}, m_j^e = v_j^e(m_j^p/v_j^p - m_j/v_j), \tag{41}$$

where m_j and v_j are the a priori mean and variance of x_j calcu-
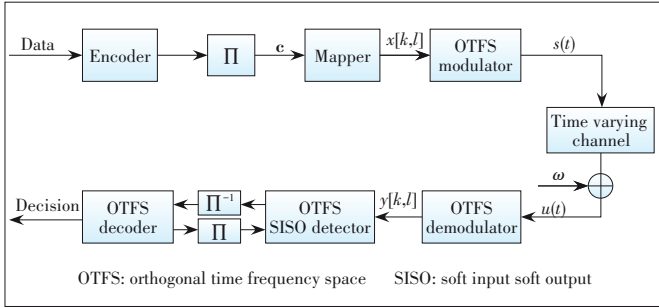
▲Figure 4. Iterative joint detection and decoding in a coded OTFS system[25]

lated based on the output LLRs of the SISO decoder[30] and m_j^p and v_j^p are a posteriori mean and variance of x_j.

Taking the UAMP detector as an example, we show the incorporation of the OTFS detector into a turbo receiver. According to the derivation of the UAMP algorithm, we can find that q and ν_q consist of the extrinsic means and variances of the symbols in x as they are the messages passed from the observation side and do not contain the immediate a priori information about x. Hence we have $m_j^e = q_j$ and $v_j^e = \nu_q$. Then Eq. (40) can be readily used to compute the extrinsic LLRs of the coded bits. With the LLRs provided by the SISO decoder, one can compute the probability $p(x_j = \alpha_a)$ for each x_j, which is no longer the "non-informative prior" in Algorithm 2. Therefore, $\xi_{j,a}$ in Line 7 of the algorithm is changed to

$$\xi_{j,a} = p(x_j = \alpha_a)\exp(-\nu_q^{-1}|\alpha_a - q_j|^2). \tag{42}$$

In addition, the iteration of the UAMP detector can be combined with the iteration between the SISO decoder and detector, which leads to a single loop iteration (i.e., inner iterations are not required).

The computational complexity of the detectors is summarized in Table 2. In the above discussion, we focus on the bi-orthogonal waveform. The detectors can be extended to OTFS systems with other waveforms, such as the simple rectangular waveform[25].

5 Simulation Results

In this section, we compare the performance of the message passing based detectors. The low complexity MRC detector in Ref. [14] is also included. We set $M = 256$ and $N =$

▼Table 2. Computational complexity of various detectors per iteration

Detectors	Complexity
MP detector	$\mathcal{O}(MNS|\mathcal{A}|)$
VB detector	$\mathcal{O}(MNS|\mathcal{A}|)$
UAMP detector	$\mathcal{O}(MN\log(MN)) + \mathcal{O}(MN|\mathcal{A}|)$

MP: message passing   UAMP: unitary approximate message passing
VB: variational Bayes

32, i.e., there are 32 time slots and 256 subcarriers in the TF domain. Both quadrature phase shift keying (QPSK) modulation and 16-quadrature amplitude modulation (QAM) are considered. The carrier frequency is 3 GHz, and the subcarrier spacing is 2 kHz. The speed of the mobile user is set to $v = 135$ km/h, leading to a maximum Doppler frequency shift index $k_{max} = 6$. We assume that the maximum delay index is $l_{max} = 14$. The Doppler index of the i-th path is uniformly drawn from the set $[-k_{max}, k_{max}]$ and the delay index is in the range of $[1, l_{max}]$ excluding the first path ($l_1 = 0$). We assume that the fractional Doppler κ_i is uniformly distributed within $[-1/2, 1/2]$, and the channel coefficients h_i are independently drawn from a complex Gaussian distribution with mean 0 and variance η^{l_i}, where the normalized power delay profile $\eta^i = \exp(-\alpha l_i)/\sum_i \exp(-\alpha l_i)$ with α being 0 or 0.1. The maximum number of iterations is set to 15 for all iterative detectors. We note that, all detectors except the MRC detector require the noise variance. The UAMP detector performs noise precision estimation, while the other detectors (except the MRC detector) including the AMP detector assume perfect noise precision. We evaluate the performance of the detectors in a variety of scenarios including the bi-orthogonal and rectangular waveforms with integer or fractional Doppler shifts, and QPSK or 16-QAM for modulations. In addition, both uncoded and coded systems are evaluated.

Fig. 5 shows the BER performance of various detectors in the case of the bi-orthogonal waveform with different numbers of paths, where we assume no fractional Doppler shifts, i. e., $S = P$. We also assume $\alpha = 0$, and QPSK is used. From this figure, we can see that, the MP detector performs well when $P = 6$, but with the increase of P, its performance becomes worse. The VB detector has a similar trend. The MRC detector performs similarly to the MP and VB detectors when $P=6$ and delivers better performance than the MP and VB detectors with larger P. The AMP and UAMP detectors perform well, where we can see that they enjoy the diversity gain and achieve better performance with the increase of P. In all cases, the UAMP based detector delivers the best performance and significantly outperforms other detectors.

With the rectangular waveform and factional Doppler shifts, we compare the bit error ratio (BER) performance of the AMP, UAMP and MRC detectors in Fig. 6, where the number of paths $P = 9$ and $\alpha = 0.1$ is used for the power delay profile. Both QPSK and 16-QAM are considered. Due to the deviation of the channel matrix from the i. i. d. (sub-) Gaussian matrix, AMP exhibits performance loss, leading to significantly worse performance compared with the UAMP detector. Thanks to the robustness of UAMP against a general matrix, UAMP performs well. We can see that the MRC detector performs better than the AMP detector. The UAMP detector performs the best and the gaps between other detectors with the UAMP detector be-
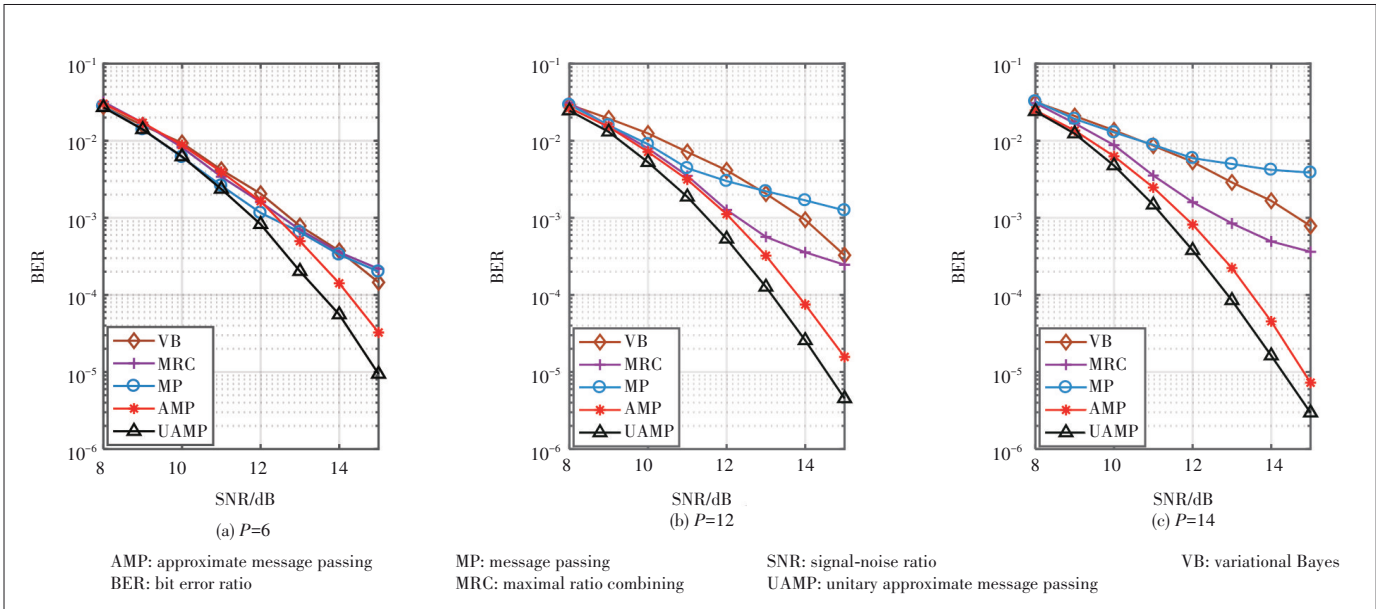
▲Figure 5. BER performance of detectors with bi-orthogonal waveform and integer Doppler shifts (results are based on Ref. [25])
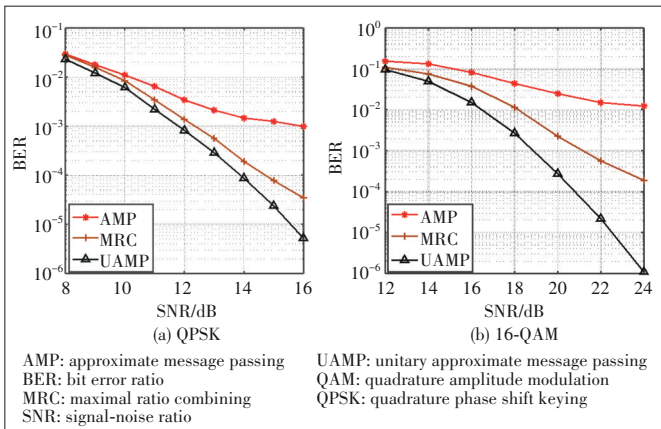
▲Figure 6. BER performance of detectors with the rectangular waveform and fractional Doppler shifts (results are based on Ref. [25])

come larger in the case of higher order modulation 16-QAM, compared with QPSK.

We then evaluate the performance of the detectors in a coded OTFS system, where the turbo receiver in Fig. 4 is employed. The number of paths $P=14$, and a rectangular waveform is used. In Fig. 7(a), we show the performance of the uncoded system with the AMP and UAMP detectors. In Fig. 7(b), we use a rate-1/2 convolutional code with a generator $[5,7]_8$ followed by a random interleaver and QPSK modulation. The length of the codeword is MN. The BCJR algorithm is used for the SISO decoder. We can find that the performance gaps between the AMP detector and the UAMP detector become larger in the coded system. The turbo receiver can achieve much better performance (about $3.5-4$ dB at the BER of 10^{-4})
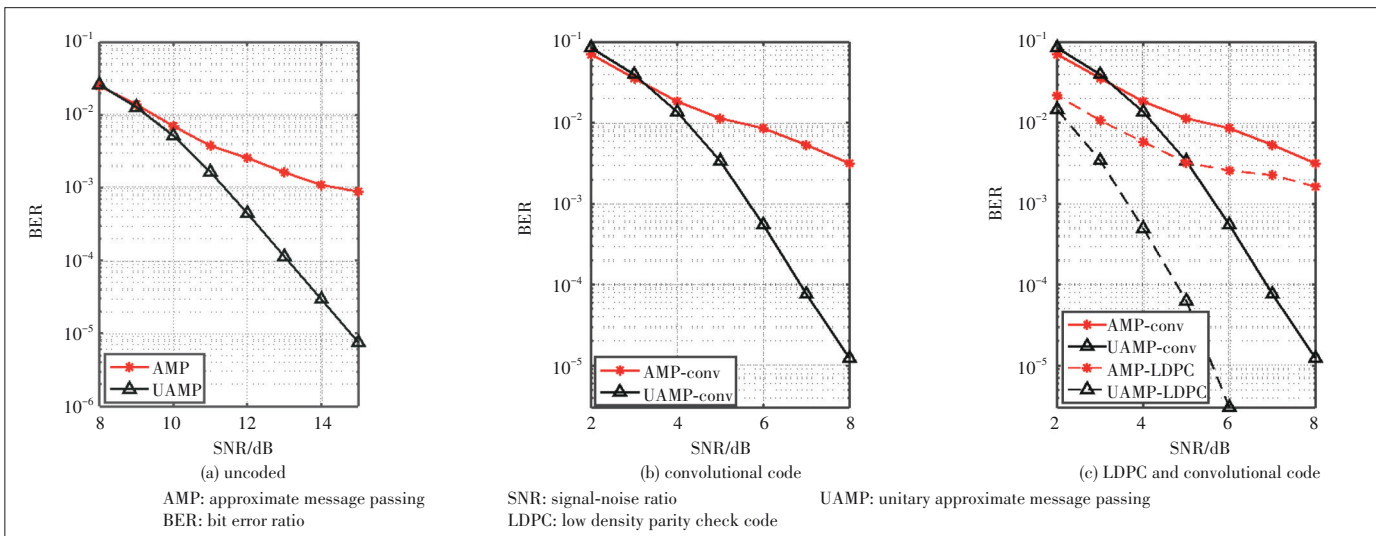
▲Figure 7. BER performance comparison of coded and uncoded system with rectangular waveform (part of the results is based on Ref. [25])

thanks to the joint processing of decoding and detection. In Fig. 7(c), we investigate the performance of the system with a more powerful LDPC. The 8 192 information bits are coded at rate R=1/2 by an irregular LDPC code with an average column weight of 3, then the coded bits are randomly interleaved and mapped. As expected, the system performance is improved considerably when the LDPC is used. From Fig. 7(c), we can see that the use of the LDPC code can improve the performance of the UAMP based detector significantly and the performance gap between AMP and UAMP increases when the LDPC is used.

6 Conclusions and Potential Future Work

In this paper, we review and compare the recently proposed message passing based OTFS detectors, which exploit the structures of the OTFS channel matrices, such as sparsity and BCCB. According to the results, the MP and VB detectors are more suitable in the scenarios that the number of paths is relatively small and the modulation order is low, where they deliver good performance while with relatively low complexity. The UAMP detector seems very promising especially in the case of rich-scattering environments and/or when fractional Doppler shifts have to be considered, where the UAMP detector is attractive in both computational complexity and performance. The results also show that the OTFS system with a turbo receiver can provide significant performance gain.

The message passing techniques seem promising in the design of OTFS receivers. In this paper, we assume the OTFS channel matrix is known, which however has to be estimated for practical applications. Message passing based OTFS channel estimation has been investigated in the literature, such as the work in Ref. [33]. With the message passing techniques, channel estimation and detection can be integrated for joint channel estimation and detection, which is expected to lead to superior system performance and/or significant reduction of the training overhead. This is because the data symbols can be used to serve as a virtual training sequence and the guard band between the training symbols and data symbols is not necessary.

It has been shown that joint decoding and detection based on a turbo receiver can significantly improve the system performance. The system performance can be potentially further improved by optimizing the error control codes. This requires fast and accurate performance prediction of the iterative receiver, so that the error control codes, e.g., LDPC, can be optimized.

The message passing techniques could be used to implement sophisticated receivers in more complex systems, such as multi-user OTFS systems, grant-free multiple access with OTFS, multiple-output-multiple-input (MIMO)-OTFS, integrated sensing and communication with OTFS.

References

[1] HADANI R, RAKIB S, TSATSANIS M, et al. Orthogonal time frequency space modulation [C]//2017 IEEE Wireless Communications and Networking Conference (WCNC). San Francisco, USA: IEEE, 2017: 1 – 6. DOI: 10.1109/WCNC.2017.7925924

[2] RAVITEJA P, PHAN K T, HONG Y, et al. Interference cancellation and iterative detection for orthogonal time frequency space modulation [J]. IEEE transactions on wireless communications, 2018, 17(10): 6501 – 6515. DOI: 10.1109/TWC.2018.2860011

[3] SURABHI G D, AUGUSTINE R M, CHOCKALINGAM A. On the diversity of uncoded OTFS modulation in doubly-dispersive channels [J]. IEEE transactions on wireless communications, 2019, 18(6): 3049 – 3063. DOI: 10.1109/TWC.2019.2909205

[4] HADANI R, MONK A. OTFS: A new generation of modulation addressing the challenges of 5G [EB/OL]. [2021-10-01]. https: //arxiv.org/abs/1802.02623

[5] LI S Y, YUAN J H, YUAN W J, et al. Performance analysis of coded OTFS systems over high-mobility channels [J]. IEEE transactions on wireless communications, 2021, 20(9): 6033 – 6048. DOI: 10.1109/TWC.2021.3071493

[6] WEI Z Q, YUAN W J, LI S Y, et al. Orthogonal time-frequency space modulation: a promising next-generation waveform [J]. IEEE wireless communications, 2021, 28(4): 136 – 144. DOI: 10.1109/MWC.001.2000408

[7] FARHANG A, REZAZADEHREYHANI A, DOYLE L E, et al. Low complexity modem structure for OFDM-based orthogonal time frequency space modulation [J]. IEEE wireless communications letters, 2018, 7(3): 344 – 347. DOI: 10.1109/LWC.2017.2776942

[8] LI L, WEI H, HUANG Y, et al. A simple two-stage equalizer with simplified orthogonal time frequency space modulation over rapidly time-varying channels [EB/OL]. [2021-10-10]. https://arxiv.org/abs/1709.02505

[9] LONG F, NIU K, DONG C, et al. Low complexity iterative LMMSE-PIC equalizer for OTFS [C]//2019 IEEE International Conference on Communications (ICC). Shanghai, China: IEEE, 2019: 1 – 6. DOI: 10.1109/ICC.2019.8761635

[10] ZEMEN T, HOFER M, LOESCHENBRAND D. Low-complexity equalization for orthogonal time and frequency signaling (OTFS) [EB/OL]. [2021-10-10]. https://arxiv.org/abs/1710.09916v1

[11] SURABHI G D, CHOCKALINGAM A. Low-complexity linear equalization for OTFS modulation [J]. IEEE communications letters, 2020, 24(2): 330 – 334. DOI: 10.1109/LCOMM.2019.2956709

[12] SINGH P, MISHRA H B, BUDHIRAJA R. Low-complexity linear MIMO-OTFS receivers [C]//2021 IEEE International Conference on Communications Workshops (ICC Workshops). Montreal, Canada: IEEE, 2021: 1 – 6. DOI: 10.1109/ICCWorkshops50388.2021.9473839

[13] LI S Y, YUAN W J, WEI Z Q, et al. Cross domain iterative detection for orthogonal time frequency space modulation [J]. IEEE transactions on wireless communications, 2021, (99): 1. DOI: 10.1109/TWC.2021.3110125

[14] THAJ T, VITERBO E. Low complexity iterative rake decision feedback equalizer for zero-padded OTFS systems [J]. IEEE transactions on vehicular technology, 2020, 69(12): 15606 – 15622. DOI: 10.1109/TVT.2020.3044276

[15] KSCHISCHANG F R, FREY B J, LOELIGER H A. Factor graphs and the sum-product algorithm [J]. IEEE transactions on information theory, 2001, 47(2): 498 – 519. DOI: 10.1109/18.910572

[16] LI H, DONG Y Y, GONG C H, et al. Low complexity receiver via expectation propagation for OTFS modulation [J]. IEEE communications letters, 2021, 25(10): 3180 – 3184. DOI: 10.1109/LCOMM.2021.3101827

[17] YUAN W J, WEI Z Q, YUAN J H, et al. A simple variational Bayes detector for orthogonal time frequency space (OTFS) modulation [J]. IEEE transactions on vehicular technology, 2020, 69(7): 7976 – 7980. DOI: 10.1109/TVT.2020.2991443

[18] ZHANG H J, ZHANG T T. A low-complexity message passing detector for OTFS modulation with probability clipping [J]. IEEE wireless communications letters, 2021, 10(6): 1271 – 1275. DOI: 10.1109/LWC.2021.3063904

[19] TIWARI S, DAS S S, RANGAMGARI V. Low complexity LMMSE Receiver for OTFS [J]. IEEE communications letters, 2019, 23(12): 2205 – 2209. DOI: 10.1109/LCOMM.2019.2945564

[20] RAVITEJA P, VITERBO E, HONG Y. OTFS performance on static multipath channels [J]. IEEE wireless communications letters, 2019, 8(3): 745 – 748. DOI: 10.1109/LWC.2018.2890643

[21] DONOHO D L, MALEKI A, MONTANARI A. Message passing algorithms for compressed sensing: motivation and construction [C]//2010 IEEE Information Theory Workshop on Information Theory (ITW 2010, Cairo). Cairo, Egypt: IEEE, 2010: 1 – 5. DOI: 10.1109/ITWKSPS.2010.5503193

[22] DONOHO D L, MALEKI A, MONTANARI A. Message passing algorithms for compressed sensing: analysis and validation [C]//2010 IEEE Information Theory Workshop on Information Theory (ITW 2010, Cairo). Cairo, Egypt: IEEE, 2010: 1 – 5. DOI: 10.1109/ITWKSPS.2010.5503228

[23] WINN J, BISHOP C M. Variational message passing [J]. Journal of machine learning research, 2005, 6(4): 661 – 694.

[24] MONK A, HADANI R, TSATSANIS M, et al. OTFS - Orthogonal Time Frequency Space [EB-OL]. [2021-10-10]. https://arxiv.org/abs/1608.02993v1.

[25] YUAN Z D, LIU F, YUAN W J, et al. Iterative detection for orthogonal time frequency space modulation with unitary approximate message passing [J]. IEEE transactions on wireless communications, 2021. DOI: 10.1109/TWC.2021.3097173

[26] LIU F, YUAN Z D, GUO Q H, WANG Z Y, et al. Multi-block UAMP based detection for OTFS with rectangular waveform [J]. IEEE wireless communications letters, 2021. DOI: 10.1109/LWC.2021.3126871

[27] GUO Q H, XI J T. Approximate message passing with unitary transformation [EB/OL]. [2021-10-10]. https://arxiv.org/abs/1504.04799

[28] YUAN Z D, GUO Q H, LUO M. Approximate message passing with unitary transformation for robust bilinear recovery [J]. IEEE transactions on signal processing, 2021, 69: 617 – 630. DOI: 10.1109/TSP.2020.3044847

[29] LUO M, GUO Q H, JIN M, et al. Unitary approximate message passing for sparse Bayesian learning [J]. IEEE transactions on signal processing, 2021, 69: 6023 – 6039. DOI: 10.1109/TSP.2021.3114985

[30] TUCHLER M, SINGER A C, KOETTER R. Minimum mean squared error equalization using a priori information [J]. IEEE transactions on signal processing, 2002, 50(3): 673 – 683. DOI: 10.1109/78.984761

[31] GUO Q H, PING L. LMMSE turbo equalization based on factor graphs [J]. IEEE journal on selected areas in communications, 2008, 26(2): 311 – 319. DOI: 10.1109/JSAC.2008.080208

[32] GUO Q H, HUANG D D. A concise representation for the soft-in soft-out LMMSE detector [J]. IEEE communications letters, 2011, 15(5): 566 – 568. DOI: 10.1109/LCOMM.2011.032811.102073

[33] LIU F, YUAN Z D, GUO Q H, et al. Message passing based structured sparse signal recovery for estimation of OTFS channels with fractional Doppler shifts [J]. IEEE transactions on wireless communications, 2021. DOI: 10.1109/TWC.2021.3087501

Biographies

YUAN Zhengdao received the B.E. degree in communication and information system from Henan University of Science and Technology, China in 2006, the M.E. degree in communication engineering from Soochow University, China in 2009, and the Ph.D. degree in information and communication engineering from the National Digital Switching System Engineering and Technological Research Center, China in 2018. He is currently an associate professor with the Open University of Henan. He was a visiting scholar with the University of Wollongong, Australia in 2019. His research interests are mainly in massive MIMO, sparse channel estimation, message passing algorithm, and iterative receiver.

LIU Fei received the B.E. and M.E. degrees in information and communication engineering from Zhengzhou University, China in 2015 and 2017, respectively. He is currently working toward the Ph.D. degree with the School of Information and Engineering, Zhengzhou University, China. His research interests are message passing algorithm, sparse signal recovery, and OTFS.

GUO Qinghua (qguo@uow.edu.au) received the B.E. degree in electronic engineering and the M.E. degree in signal and information processing from Xidian University, China in 2001 and 2004, respectively, and the Ph.D. degree in electronic engineering from the City University of Hong Kong, China in 2008. He is currently an associate professor with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Australia and an adjunct associate professor with the School of Engineering, The University of Western Australia. His research interests include signal processing, machine learning and telecommunications. He was a recipient of the Australian Research Council's inaugural Discovery Early Career Researcher Award in 2012.

WANG Zhongyong received the B.S. and M.S. degrees in automatic control from Harbin Shipbuilding Engineering Institute, China in 1986 and 1988, respectively, and the Ph.D. degree in automatic control theory and application from Xi'an Jiaotong University, China in 1998. From 1988 to 2002, he has been with the Department of Electronics, Zhengzhou University, China. Now he is a professor with the Department of Communication Engineering, Zhengzhou University. His research interests include numerous aspects within embedded systems, signal processing, and communication theory.