Data-Driven Bandpass Filter Design for Estimating Symbol Rate of Sporadic Signal at Low SNR

Can Pei *Student Member, IEEE*, Suzhi Bi, *Senior Member, IEEE*, and Zhi Quan, *Senior Member, IEEE*

***Abstract*—Symbol rate is one of the most important parameters in signal demodulation process. In real-time signal processing, traditional symbol rate estimation algorithms for the Multiple Phase Shift Keying (M-PSK) and the Multiple Quadrature Amplitude Modulation (M-QAM) are based on the Fourier transform of signal's complex envelope. At the low signal-to-noise ratio (SNR), the accuracy of symbol rate estimation can be improved by increasing the number of symbols as much as possible. However, this improvement is infeasible in many applications such as the energy-limited Internet of Things devices and sporadic noncooperative transmissions. In this paper, we propose a data-driven bandpass filter (BPF) design scheme for accurate estimation of symbol rate under low SNR with only a small number of symbols available. The proposed scheme considerably improves the estimation performance by optimizing the BPF design using the equivalent dynamic linearization model with time-varying pseudo-partial derivatives. Specifically, the proposed scheme iteratively optimizes the upper and lower cut-off frequencies of the BPF based on the measured complex envelope spectrum until achieving the optimal BPF. Therefore, the peaks of the complex envelope spectrum are extracted as the estimate of the symbol rate by applying the optimal BPF. Experimental results indicate the promise of the proposed scheme as an efficient symbol rate estimator for sporadic signal at low SNR and with a small number of symbols.**

***Index Terms*—Symbol rate estimation, data-driven signal processing, complex envelope spectrum, digital filter design, and wireless communication.**

I. Introduction

IN non-cooperative communication environments such as military reconnaissance and surveillance, radio monitoring, and intelligent signal identification, symbol rate parameter estimation is a necessary condition for the signal decoding of blind receiver [1]–[8]. In electronic reconnaissance and interference identification, parameter estimation is used to intercept enemy information. The symbol rate estimation is the first step in the signal analysis process. At the same time, most blind demodulation parameters (for example symbol timing recovery) rely on the correct estimation of the symbol rate. In a multi-rate system with burst transmissions [9], transmitter can vary the symbol rate based on various criteria, such as channel conditions, dynamic bandwidth constraints, content types, etc. The receiver then estimates the changing symbol rate based on the captured data packet when no pilot data sequence is available [10]. Symbol rate estimation is also an important task when performing passive signal analysis on smart devices such as low-power IoT devices [11]. In addition, the symbol rate in commercial systems such as cable digital video broadcasting, satellite digital video broadcasting and the second-generation satellite digital video broadcasting is not fixed in order to send the satellite signal to the cable network without any processing at the cable front [4], [12], [13]. Therefore, the effective and robust symbol rate estimation is essential for adaptive communication systems before processing other signal parameters [10].

As the first step in the many signal analysis applications, blind symbol rate estimation has been studies for decades. Using the estimated bandwidth as the estimate of the symbol rate has low computational complexity but results in coarse estimation [14], [15]. The wavelet-based methods [16] extract the symbol rate from the spectrum of wavelet transform coefficients. However, the wavelet transform generally requires high sampling rates and high signal-to-noise ratios (SNR), and the wavelet scale might experience blind spots which affect the blind estimation performance. [17] proposed using the cyclic spectrum for symbol rate estimation, in which the cyclostationary characteristics of a digitally modulated single-carrier signal are processed using autocorrelation function to find the maximum cyclic correlation value. Like the wavelet-based methods, the cyclic correlation based methods need a high sampling rate and high SNR to ensure the estimation accuracy, which implies high computational complexity and long processing delay. In [18]–[20], inverse Fourier transform was used to estimate the symbol rate of M-ary modulation signals, but these estimators need a large number of symbols otherwise the estimation accuracy deteriorates significantly. Square complex envelope spectrum is efficient for symbol rate estimation [21], [22] with low computational complexity, but nonetheless its accuracy becomes worse at lower SNR and/or with a small number of symbols.

In practice, estimation accuracy improvement by increasing the number of symbols might not be feasible in many applications, especially for sporadic and low latency transmission. For example, the IoT devices under harsh communication conditions, i.e., underwater, stringent latency, military reconnaissance, and power constrained, usually transmit very short messages [23], [24]. To reduce communication overheads and power consumption, sporadic transmission with ultra-low delay and short symbol length is mostly used by the next generation mobile wireless networks [25], [26]. [26] introduces a modulation on conjugate-reciprocal zeros technique that can transmit sporadic short-packets at high spectral efficiency and

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1800800, in part by the National Natural Science Foundation of China (Project 61871271), the Guangdong Province Pearl River Scholar Funding Scheme 2018, and the Key Project of Department of Education of Guangdong Province (No. 2020ZDZX3050). (*Corresponding Author: Zhi Quan*)

C. Pei, S. Bi, and Z. Quan are with the College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China, 518060(e-mail: 2160130447@email.szu.edu.cn; bsz@szu.edu.cn; zquan@szu.edu.cn). S. Bi is also with the Peng Cheng Laboratory, Shenzhen, China, 518066.

low-latency via unknown wireless multipath channels. Furthermore, frequency-hopping military communication systems usually change modulation parameters between short bursts to avoid interception and recognition [27]. Consequently, the accuracy of symbol rate estimation for sporadic signal is limited because the number of symbols thus collected is small.

Recently, scholars have proposed some symbol rate estimate methods based on machine learning [28], [29]. [29] proposed an automated symbol rate estimator for Binary Phase Shift Keying (BPSK), Quadrature phase shift keying (QPSK), 16 phase shift keying (16-PSK), 16 Quadrature Amplitude Modulation (16-QAM), and 64 Quadrature Amplitude Modulation (64-QAM) signals by using deep neural network (DNN) model. However, the symbol rate estimator for each modulation scheme needs 5×10^5 sample data. Meanwhile, the DNN-based method for symbol rate estimation needs to be trained in batches of 10^4. This time-consuming estimation method cannot better meet the requirements of real-time signal processing. At the same time, a large number of data samples are not within the scope of our discussion. In this paper, we present an accurate estimator for the symbol rate at low SNR with a small number of symbols. The strategy is to apply a bandpass filter (BPF) on the received signal to diminish the noise in frequency domain as much as possible before processing. The main challenge is how to design a BPF with optimal upper and lower cut-off frequencies that can best identify the spectral lines of the symbol rate. We propose a data-driven scheme to derive the optimal BPF design parameters for symbol rate estimation without an explicit mathematical model [30]–[32]. This scheme iteratively optimizes the cut-off frequencies of the BPF based on a series of local equivalent dynamic linearization models with time-varying parameters. Specifically, we initialize the upper and lower cut-off frequencies based on the coarse symbol rate estimation that is obtained using the 3dB bandwidth estimation method. The upper and lower cut-off frequencies are iteratively updated by using the proposed data-driven scheme based on pseudo-partial derivative (PPD) parameter vector [32] to further reduce the average-to-peak power ratio of the complex envelope spectrum. This procedure is repeated a few times over the same set of measurement data until an optimal BPF is obtained. Finally, the estimated optimal BPF is used to extract the complex envelope spectrum peaks as the symbol rate spectral lines.

The major contributions of the data-driven symbol rate estimate scheme proposed in this paper are as follows. Firstly, aiming at the problem that the symbol rate estimation of the short burst signal is not accurate at low SNR and a small number of symbols, we proposed a data-driven BPF design scheme in the paper for symbol rate estimation of the short burst signal. The data-driven BPF is flexible in design and has a small bandwidth. It can reduce the influence of interference signals and noise on the symbol rate spectral lines, and increase the SNR, thereby effectively improving the performance of the symbol rate estimation of the short-time burst signal at low SNR and with a small number of symbols. Secondly, our proposed data-driven based algorithm does not need to establish an accurate model of the system and only require the input/output data of the system, i.e., the input data are the BPF upper and lower cut-off frequencies $\boldsymbol{u}(k)$, and the output average-to-peak ratio $r(k)$ of the complex envelope spectrum as a result of the filter specifications $\boldsymbol{u}(k)$ obtained from measurement data. Compared with the DNN-based symbol rate estimation algorithm [29], our proposed algorithm does not require large training data and calculation time. It only needs online iterative optimization, which can well meet the requirements of real-time signal processing.

The rest of this paper is organized as follows. Section II introduces the signal model of symbol rate estimation based on the square complex envelope spectrum. Section III illustrates the data-driven BPF design scheme for symbol rate estimation at low SNR with a small number of symbols. In Section IV, the proposed estimator is experimentally validated in a system testbed. Section V concludes the paper.

II. Unbiased Symbol Rate Estimation

A. Signal Model

Consider a single-carrier digitally modulated signal, such as Multiple Phase Shift Keying (M-PSK) or the Multiple Quadrature Amplitude Modulation (M-QAM), transmitted over a stationary Gaussian white noise channel. The analytical form of the received modulated signal is given by [21], [22]

$$x(t) = e^{-j2\pi f_c t} \sum_{i=-\infty}^{+\infty} s_i h(t - iT_s) + n(t), \tag{1}$$

where f_c is the carrier frequency, $\left\{s_i = c_i e^{j\phi_i}\right\}$ is a sequence of identically distributed independent symbols with zero mean and unit variance, c_i and ϕ_i are the amplitude and phase of transmitted symbol in the i-th symbol period respectively. The pulse shape filter $h(t)$ is the root raised cosine (RRC) function, T_s is the symbol duration, and $n(t)$ is the additive white Gaussian noise (AWGN) with two sided power spectral density (PSD) having a variance of σ_v^2. When $f_c = 0$, $x(t)$ is a baseband signal. Thus, the baseband information of the signal $x(t)$ is

$$A(t) = \sum_{i=-\infty}^{+\infty} s_i h(t - iT_s), \tag{2}$$

where $|A(t)|$ is the module of the complex envelope of $x(t)$ with a period of T_s. The frequency response of the shaping filter function $h(t)$ is given by [15]

$$H(w) = \begin{cases} T_s, \quad 0 \le |w| \le \frac{(1-\alpha)\pi}{T_s} \\ \frac{T_s}{2}[1 + \sin\frac{T_s}{2}(\frac{\pi}{T_s} - w)], \quad \frac{(1-\alpha)\pi}{T_s} \le |w| \le \frac{(1+\alpha)\pi}{T_s} \\ 0, \quad |w| \ge \frac{(1+\alpha)\pi}{T_s} \end{cases} \tag{3}$$

where $0 \le \alpha \le 1$ is the roll-off factor. The symbol rate estimator based on the square complex envelope spectrum is given by [21], [22]

$$\hat{R}_s = \frac{1}{2\pi} \underset{w\in(0,+\infty)}{\arg\max} \left[\int_{t=-\infty}^{\infty} |x(t)|^2 e^{-jwt} dt \right]. \tag{4}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3114678, IEEE Transactions on Wireless Communications

B. Unbiasedness of the Symbol Rate Estimator

The mean value of the estimated symbol rate $\hat{R}_s$ is

$$E\left[\hat{R}_s\right] = E\left[\frac{1}{2\pi}\underset{w\in(0,+\infty)}{\arg\max}\int_{t=-\infty}^{\infty}|x(t)|^2 e^{-jwt}dt\right] = \frac{1}{2\pi}\underset{w\in(0,+\infty)}{\arg\max}\int_{t=-\infty}^{\infty}E\left[|x(t)|^2\right]e^{-jwt}dt. \quad (5)$$

From (1) and (2)

$$\begin{aligned} E\left[|x(t)|^2\right] &= E\left[x(t)x^*(t)\right] \\ &= E\left[(A(t)e^{-j2\pi f_c t}+n(t))(A^*(t)e^{j2\pi f_c t}+n(t))\right] \\ &= E[A(t)A^*(t)]+E[n(t)A^*(t)e^{j2\pi f_c t}] \\ &\quad +E[A(t)e^{-j2\pi f_c t}n^*(t)]+E[n(t)n^*(t)] \end{aligned} \quad (6)$$

where $(.)^*$ denotes the complex conjugate operator. As $n(t)$ is Gaussian with zero mean and variance σ^2, (6) can be simplified as

$$E\left[|x(t)|^2\right] = E\left[|A(t)|\,|A^*(t)|\right]+\sigma^2. \quad (7)$$

Substituting (2) into (7) gives

$$E\left[|x(t)|^2\right] = E\left[\sum_{i=-\infty}^{+\infty}s_i h(t-iT_s)\times\sum_{i=-\infty}^{+\infty}s_i^* h^*(t-iT_s)\right]+\sigma^2. \quad (8)$$

Because $\{s_i\}$ is an independent identically distributed zero-mean unit-variance random sequence, $E\left[s_m s_i^*\right]=\delta(m-i)$. For the response function of the shaping filter $h(t)=h^*(t)$, (8) can be rewritten as

$$E\left[|x(t)|^2\right] = \sum_{i=-\infty}^{+\infty}\left[h(t-iT_s)\right]^2+\sigma^2. \quad (9)$$

Let

$$G(t) = \sum_{i=-\infty}^{+\infty}\left[h(t-iT_s)\right]^2. \quad (10)$$

Its frequency response function can be given as [21]

$$\begin{aligned} G(w) &= \frac{1}{T_s}comb_{1/T_s}\left[H(w)\otimes H(w)\right] \\ &= \frac{1}{T_s}\sum_{i=-\infty}^{+\infty}\left[H(w)\otimes H(w)\right]\delta\left(w-\frac{2\pi i}{T_s}\right), \end{aligned} \quad (11)$$

where $comb_{1/T_s}(.)$ is the comb of the unit impulse function with intervals of $2\pi/T_s$ and $\otimes$ denotes the convolution operator. When $w=2\pi i/T_s$, $G(w)$ obtains the maximum value at $w\in(0,+\infty)$. $G(w)$ is a periodic pulse sequence and can be used to determine the symbol rate for $i=1$ [22]. We set $w=2\pi/T_s$ in (11) and then substitute (11) into (5), and obtain

$$E\left[\hat{R}_s\right] = 1/T_s. \quad (12)$$

Thus, $\hat{R}_s$ is unbiased.

C. Complex Envelope Spectrum

In real-time signal processing, the estimate of the symbol rate can be obtained by scanning the square complex envelope spectrum for spectral peaks. Fig. 1 shows the square complex envelope spectrum of a 16-PSK signal with 1000 symbols, where the symbol energy-to-noise power ratio $Es/No=8$ dB, the roll-off factor of the pulse shaping filter is 0.35, and the number of samples per symbol is 4. Denote the sampling frequency as f_s, the two spectral lines of the symbol rate as N_1 and N_2, and the size of the fast Fourier transform (FFT) as N. The symbol rate estimator can be calculated as

$$\hat{R}_s = \frac{|N_2-N_1|}{2N}f_s. \quad (13)$$

However, with a small number of symbols and at low SNR, the symbol rate spectral lines are difficult to identify in the complex envelope spectrum. As shown in Fig. 2, for the spectrum envelope of a 16-PSK with only 200 symbols, where Es/No, the roll-off factor, and the number of samples per symbol are the same as in Fig. 1, the spectral peaks are overwhelmed by noise and thus are difficult to recognize.

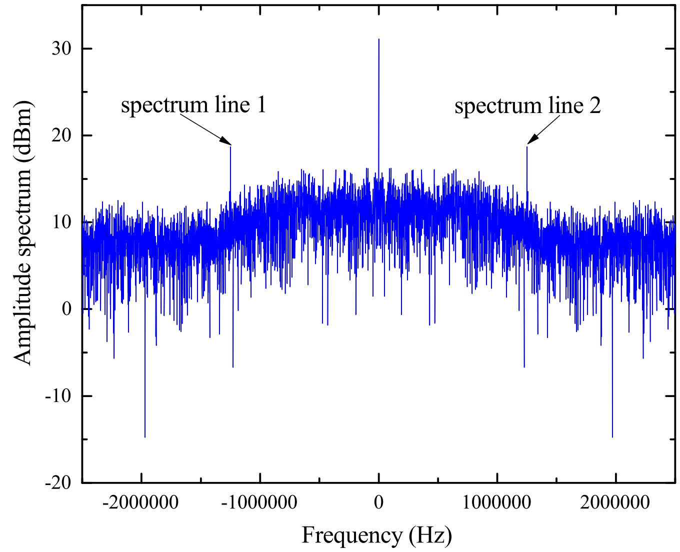

Fig. 1: Complex envelope spectrum of a 16-PSK signal with 1000 symbols.

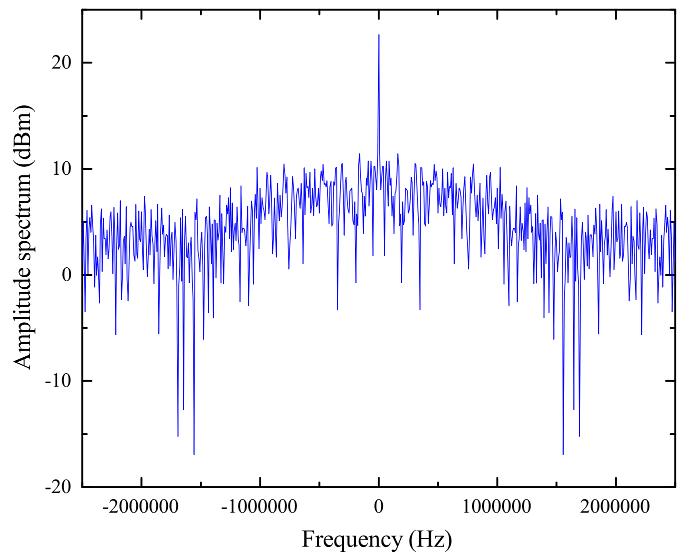

Fig. 2: Complex envelope spectrum of a 16-PSK signal with 200 symbols.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3114678, IEEE Transactions on Wireless Communications

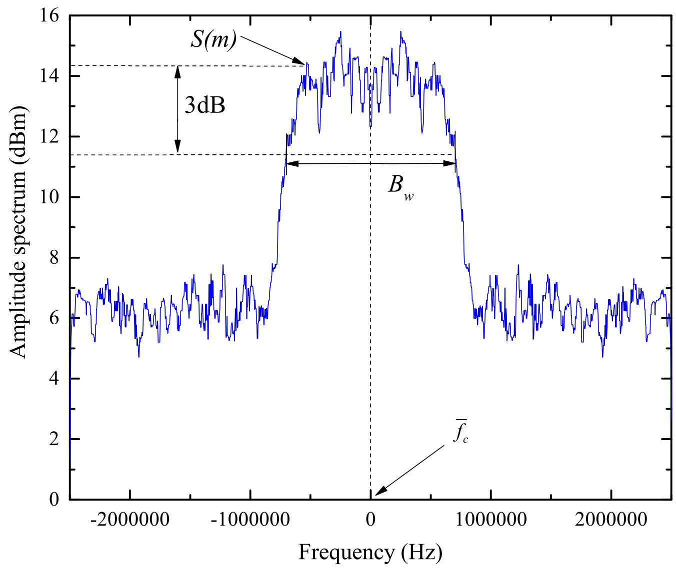

Fig. 3: The smoothed amplitude spectrum $S(f)$.

III. Iterative Optimization of Symbol Rate Estimation

A. Coarse Bandwidth Estimation

As discussed above, the symbol rate can be roughly estimated from the received signal bandwidth B_w according to the Nyquist criterion [14], [15]. The estimated symbol rate is coarsely determined by the 3 dB bandwidth method [14], [15], i.e.,

Step 1: Perform Fourier transform on the received signal $x(t)$ to get its amplitude spectrum $X(f)$.

Step 2: Apply a median filter to smooth the amplitude spectrum $X(f)$ to obtain the smoothed amplitude spectrum $S(f)$ as shown in Fig. 3.

Step 3: Use center of gravity method [15] to estimate the carrier frequency $\overline{f_c}$, the formula is as follows:

$$\overline{f_c} = \frac{\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} X(f).fdf}{\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} X(f)df}. \tag{14}$$

Step 4: Calculate the average value of the smoothed amplitude spectrum $S(f)$ in the range of $[\overline{f_c} - f_s/12, \overline{f_c} + f_s/12]$, denoted as $S(m)$ [2].

Step 5: Find the amplitude spectrum which is 3 dB smaller than $S(m)$ and calculate its corresponding frequency. Therefore, the estimated bandwidth B_w of $x(t)$ is calculated according to the following formula:

$$\begin{aligned} B_w =& \mathrm{Max}[\arg\{f : 10\log[S(m)] - 10\log[S(f)] = 3\}] - \\ & \mathrm{Min}[\arg\{f : 10\log[S(m)] - 10\log[S(f)] = 3\}]. \end{aligned} \tag{15}$$

The estimate of B_w is not accurate due to interference noise in short-time burst signals and the smoothing filter. Refer to [2], [14], using this estimated bandwidth in symbol rate estimation will bring an estimation error close to 20%, but can serve as a rough estimate of the symbol rate for further refinement.

B. Iteratively Refining Estimation

In this subsection, we introduce the iterative BPF optimization, which can start from the rough 3 dB bandwidth estimation above. The two key parameters of a BPF are the lower and upper cut-off frequencies, denoted as f_1 and f_2 respectively. Thus, its bandwidth is given by $(f_2 - f_1)$. The normalized cut-off frequencies of the corresponding BPF are given by:

$$u_1 = \frac{f_1}{f_s/2}, u_2 = \frac{f_2}{f_s/2}, \tag{16}$$

where f_s represents the sampling rate of the signal. The cut-off frequencies can be rewritten in the vector form $\boldsymbol{u} = [u_1, u_2]^T$, where $u_1 < u_2$ and $0 < u_i < 1, i = 1, 2$. Therefore, the normalized upper and lower cut-off frequencies vectors in the k-th iteration are $\boldsymbol{u}(k) = [u_1(k), u_2(k)]^T$. The initial cut-off frequencies vector is given by $\boldsymbol{u}(1) = 2/f_s \cdot [0.75B_w, 1.25B_w]^T$. Since the initial value $\boldsymbol{u}(1)$ can not be arbitrarily set, the position of the passband cut-off frequency and the size of the BPF bandwidth will affect the accuracy of symbol rate estimation. Therefore, to obtain an appropriate filter upper and lower cutoff frequencies and find the maximum symbol rate spectrum line, we develop a data-driven scheme to achieve fine estimation performance by iteratively optimizing the BPF upper and lower cut-off frequencies.

The digital filter design process usually involves the impulse response function calculation based on the given specifications. There are various frequency tunning techniques for digital filters in the literature [33], [34], but they are out of scope of this paper. In this paper, we will use a data-driven method based on model-free adaptive control [35], [36] to dynamically optimize the cut-off frequencies of the BPF based on the complex envelope spectrum. By iteratively adjusting the upper and lower cut-off frequencies, the passband center frequency of the BPF $r_c(k) = [u_1(k) + u_2(k)]/2$ gradually approaches to the spectral peak of the complex envelope. This peak value of the complex envelope spectrum filtered through the BPF is used as the estimate of the symbol rate. In order to quantitatively evaluate the BPF, we use the average-to-peak power ratio r as the performance metric, which is defined as the ratio of the mean of the signal square complex envelope spectrum within the passband to the peak. The average-to-peak power ratio in the k-th iteration is denoted by $r(k)$.

To summarize, the flow diagram of the proposed symbol rate estimation scheme is displayed in Fig. 4. Firstly, the rough estimate of the symbol rate is determined by the complex signal bandwidth B_w obtained by the 3 dB bandwidth method, based on which the cut-off frequencies $\boldsymbol{u}(1)$ and the average-to-peak ratio $r(1)$ of a BPF are initialized in the next step. Then, the data-driven method iteratively updates the cut-off frequencies $\boldsymbol{u}(k)$ based on the same measurement dataset. The average-to-peak power ratio $r(k)$ will reach to a certain minimum value after a few iterations, where the optimal BPF is achieved. Therefore, the symbol rate is estimated by (13) from the spectral peaks filtered by the optimal BPF.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3114678, IEEE Transactions on Wireless Communications

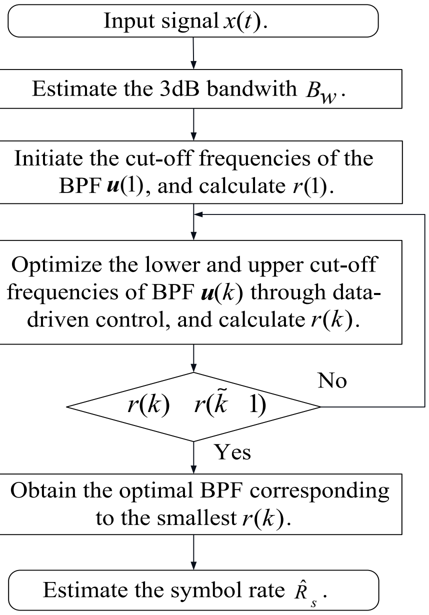

Fig. 4: Signal processing diagram of the symbol rate estimation.

C. Dynamic Linearization Data Model

In this section, we develop the data-driven BPF design scheme based on real-time measurement data. We start by introducing the equivalent dynamic linearization data model of BPF design. The average-to-peak power ratio of signal complex envelope spectrum in k-th iteration is expressed as:

$$r(k) = f(r(k-1), ..., r(k-n_p), \boldsymbol{u}(k), ..., \boldsymbol{u}(k-n_s)), \quad (17)$$

where $f(\cdot)$ is an unknown nonlinear function, n_p and n_s are the unknown orders, $\boldsymbol{u}(k)$ is cut-off frequency vector of BPF at the k-th iteration. It can be seen from (17) that the average-to-peak power ratio $r(k)$ of the complex envelope spectrum at the k-th iteration is related to the cut-off frequencies $\boldsymbol{u}$ from iterations $k-n_s$ to k and the average-to-peak power ratio r from iterations $k-n_p$ to $k-1$. The nonlinear multiple input single output (MISO) system (17) can be represented as a dynamic linearization model based on the following two assumptions.

Assumption 1: The partial derivative of the unknown function $f(\cdot)$ with respect to the upper and lower cutoff frequencies $\boldsymbol{u}(k), \boldsymbol{u}(k-1), ..., \boldsymbol{u}(k-n_s)$ are continuous, which is a typical constraint condition for general nonlinear systems.

Assumption 2: For any k, the MISO system (17) satisfies the generalized *Lipschitz* condition, i.e.,

$$|\Delta r(k)| \leq b\|\Delta \boldsymbol{u}(k)\|, \quad (18)$$

where $\Delta r(k) = r(k) - r(k-1)$, $\Delta \boldsymbol{u}(k) = \boldsymbol{u}(k) - \boldsymbol{u}(k-1), \Delta \boldsymbol{u}(k) \neq 0$, and b is a positive constant. This assumption imposes an upper bound limitation on the change rate of the average-to-peak power ratio relative to the change rate in the upper and lower cut-off frequencies.

Proposition 1: For any nonlinear MISO system (17) that satisfies *Assumptions 1* and *2* for all k, there must exist a time-varying parameter $\boldsymbol{\varphi}(k)$, such that system (17) can be transformed into the following equivalent dynamic linearization data model [32], [35]:

$$\Delta r(k) = \boldsymbol{\varphi}(k)\Delta \boldsymbol{u}(k), \quad (19)$$

where $\boldsymbol{\varphi}(k) = [\varphi_1(k), \varphi_2(k)]$ is called PPD vector at the k-th measurement, and $\|\boldsymbol{\varphi}(k)\| \leq b$ is bounded for all k.

Proof : Please refer to Appendix A. ■

D. Data-Driven BPF Design for Symbol Rate Estimation

Since the PPD parameter $\boldsymbol{\varphi}(k)$ in equation (56) is unknown, the modified projection algorithm [30] is used to estimate the parameter $\boldsymbol{\varphi}(k)$ in the design. From the perspective of eliminating steady-state deviation and ensuring the MISO system stability, we choose a criterion function for the PPD estimation as:

$$\begin{aligned} J(\hat{\boldsymbol{\varphi}}(k)) = &\|\Delta r(k-1) - \hat{\boldsymbol{\varphi}}(k)\Delta \boldsymbol{u}(k-1)\|^2 \\ &+ \mu(k-1)\|\hat{\boldsymbol{\varphi}}(k) - \hat{\boldsymbol{\varphi}}(k-1)\|^2, \end{aligned} \quad (20)$$

where $\hat{\boldsymbol{\varphi}}(k)$ is the estimation value of $\boldsymbol{\varphi}(k)$ and $\mu(k-1) \geq \mu_{\min} > 0$ is a variable penalty factor that restricts the excessive variation in the PPD parameter estimation. Taking the partial derivative of $J(\hat{\boldsymbol{\varphi}}(k))$ with respect to $\hat{\boldsymbol{\varphi}}(k)$:

$$\begin{aligned} \frac{\partial J(\hat{\boldsymbol{\varphi}}(k))}{\partial \hat{\boldsymbol{\varphi}}(k)} = &2[\hat{\boldsymbol{\varphi}}(k)\Delta \boldsymbol{u}(k-1) - \Delta r(k-1)]\Delta \boldsymbol{u}(k-1)^T \\ &+ 2\mu(k-1)[\hat{\boldsymbol{\varphi}}(k) - \hat{\boldsymbol{\varphi}}(k-1)]. \end{aligned} \quad (21)$$

Solving for the optimal condition $\frac{\partial J(\hat{\boldsymbol{\varphi}}(k))}{\partial \hat{\boldsymbol{\varphi}}(k)} = 0$, then:

$$\begin{aligned} \hat{\boldsymbol{\varphi}}(k) = &\left[\mu(k)\hat{\boldsymbol{\varphi}}(k-1) + \Delta r(k-1)\Delta \boldsymbol{u}(k-1)^T\right] \\ &\times \left[\Delta \boldsymbol{u}(k-1)\Delta \boldsymbol{u}(k-1)^T + \mu(k)\boldsymbol{I}\right]^{-1}. \end{aligned} \quad (22)$$

To avoid matrix inversion operation, we apply the *Matrix Inversion Lemma* [37] to simplify (22), and then obtain

$$\begin{aligned} \hat{\boldsymbol{\varphi}}(k) = &\hat{\boldsymbol{\varphi}}(k-1) \\ &+ \frac{\eta[\Delta r(k-1) - \hat{\boldsymbol{\varphi}}(k-1)\Delta u(k-1)]\Delta u(k-1)^T}{\mu(k-1) + \|\Delta u(k-1)\|^2}, \end{aligned} \quad (23)$$

where η is the step-length factor.

Proof : Please refer to Appendix C. ■

From the above equation, the parameter $\hat{\boldsymbol{\varphi}}(k)$ is related to the cut-off frequencies and the average-to-peak power ratio till time instant $k-1$. This implies that $\hat{\boldsymbol{\varphi}}(k)$ can be regard as a slowly time-varying parameter.

To make the estimation parameter $\hat{\boldsymbol{\varphi}}(k)$ have a stronger time-varying tracking ability and ensure $\Delta \boldsymbol{u}(k) \neq 0$, we apply the following constraints

$$\begin{aligned} \hat{\varphi}_i(k) = \hat{\varphi}_i(1), \text{ if } sign(\hat{\varphi}_i(k)) \neq sign(\hat{\varphi}_i(1)), \ i = 1,2 \\ \hat{\boldsymbol{\varphi}}(k) = \hat{\boldsymbol{\varphi}}(1), \text{ if } \|\hat{\boldsymbol{\varphi}}(k)\| \leq \sigma \text{ or } \|\boldsymbol{\Delta} u(k)\|^2 \leq \sigma, \end{aligned} \quad (24)$$

where $\hat{\boldsymbol{\varphi}}(1)$ is the initial value of the PPD parameter vector, σ is a small positive constant to reset $\|\hat{\boldsymbol{\varphi}}(k)\|$ if the magnitude of $\|\hat{\boldsymbol{\varphi}}(k)\|$ or $\|\boldsymbol{\Delta} u(k)\|$ is too small. This reset scheme can strengthen the tracking ability of the estimated algorithm.

Let r^* be the target average-to-peak power ratio. Our objective is to find the upper and lower cut-off frequencies $\boldsymbol{u}(k)$, at which the resulting $r(k)$ is close to r^*. Define a cost function as

$$J(\boldsymbol{u}(k)) = \|r^* - r(k)\|^2 + \lambda(k-1)\,\|\boldsymbol{u}(k) - \boldsymbol{u}(k-1)\|^2, \tag{25}$$

where $\lambda(k-1) > 0$ is a variable penalty factor used to prevent excessive variation in estimating $\boldsymbol{u}(k)$. Combining (23) and (56) gives

$$r(k) = r(k-1) + \hat{\varphi}(k)\Delta\boldsymbol{u}(k). \tag{26}$$

By substituting (26) into (25), the partial derivative of $J(\boldsymbol{u}(k))$ with respect to $\boldsymbol{u}(k)$ is

$$\frac{\partial J(\boldsymbol{u}(k))}{\partial \boldsymbol{u}(k)} = 2\left[\hat{\varphi}(k)\Delta\boldsymbol{u}(k) + r(k-1) - r^*\right]\hat{\varphi}(k)^T + 2\lambda(k-1)\Delta\boldsymbol{u}(k). \tag{27}$$

Solving the optimal condition $\frac{\partial J(\boldsymbol{u}(k))}{\partial \boldsymbol{u}(k)} = 0$ gives

$$\boldsymbol{u}(k) = \boldsymbol{u}(k-1) + \frac{\rho\hat{\varphi}(k)^T}{\lambda(k-1) + \|\hat{\varphi}(k)\|^2}\left[r^* - r(k-1)\right], \tag{28}$$

where ρ is the step-length factor.

E. Optimization and Convergence Analysis

Note that the penalty factors $\lambda(k)$ and $\mu(k)$ restrict the variation of $\boldsymbol{u}(k)$ and $\hat{\varphi}(k)$, respectively. They can also reduce the system steady-state error, and prevent the abnormal situation where the denominator in (23) or (28) is zero. The penalty factors should be appropriately chosen to ensure the system stability and the tracking capability [30], [36]. In this paper, we apply the gradient descent algorithm [38] to iteratively optimize the penalty factors. Therefore, $\lambda(k)$ becomes

$$\lambda(k) = \lambda(k) - \beta_1\left[\frac{\partial J(\boldsymbol{u}(k))}{\partial \boldsymbol{u}(k)}\right]^T\left[\frac{\partial \boldsymbol{u}(k)}{\partial \lambda(k-1)}\right], \tag{29}$$

where β_1 is the learning rate, $\frac{\partial J(\boldsymbol{u}(k))}{\partial \boldsymbol{u}(k)}$ was previously calculated in equation (27). From (28), $\frac{\partial \boldsymbol{u}(k)}{\partial \lambda(k-1)}$ is given by

$$\frac{\partial \boldsymbol{u}(k)}{\partial \lambda(k-1)} = -\frac{\rho\hat{\varphi}(k)^T(r^* - r(k-1))}{\left[\lambda(k-1) + \|\hat{\varphi}(k)\|^2\right]^2}. \tag{30}$$

In addition, $\lambda(k) > \lambda_{\min} > 0$. Similarly, $\mu(k)$ is given by

$$\mu(k) = \mu(k) - \beta_2\left[\frac{\partial J(\hat{\varphi}(k))}{\partial \hat{\varphi}(k)}\right]\left[\frac{\partial \hat{\varphi}(k)}{\partial \mu(k-1)}\right]^T, \tag{31}$$

where β_2 is the learning rate, $\frac{\partial J(\hat{\varphi}(k))}{\partial \hat{\varphi}(k)}$ was previously calculated in equation (21). From (23), $\frac{\partial \hat{\varphi}(k)}{\partial \mu(k-1)}$ is given by

$$\frac{\partial \hat{\varphi}(k)}{\partial \mu(k-1)} = \frac{-\eta\left[\Delta r(k-1) - \hat{\varphi}(k-1)\Delta\boldsymbol{u}(k-1)\right]\Delta\boldsymbol{u}(k-1)^T}{\left[\mu(k-1) + \|\Delta\boldsymbol{u}(k-1)\|^2\right]^2}. \tag{32}$$

In this proposed algorithm, the upper and lower cut-off frequencies $\boldsymbol{u}(k)$ of BPF are continuously optimized based on data-driven method and gradient descent method. In each iteration, the data-driven algorithm can continuously adjust the BPF upper and lower cut-off frequencies $\boldsymbol{u}(k)$ according to (23) and (28). Then, $r(k)$ for given $\boldsymbol{u}(k)$ is calculated and used to update the upper and lower cut-off frequencies to obtain $\boldsymbol{u}(k+1)$. The iteration is stopped until the smallest $r(k)$ is found or the number of iterations reaches a predetermined threshold. During iterations, the estimation parameter $\hat{\varphi}(k)$ maintains a constant direction, and thus makes both $u_1(k)$ and $u_2(k)$ move towards the spectral peak of the complex envelope. Note that the cut-off frequencies must be kept as $u_2(k) > r_c(k) > u_1(k)$ in each iteration. We also show the upper and lower cut-off frequencies as the iteration proceeds in Fig. 5. It can be seen that the upper and lower cut-off frequencies first quickly approach the spectral peak and finally converge to a stable value. In this way, an optimal BPF is generated based on the obtained optimal upper and lower cut-off frequencies. The procedure is summarized in Algorithm 1. Next, the obtained optimal BPF is used to process the squared complex envelope of the same set of signals received. Finally, the spectral peaks in the complex envelope spectrum are extracted as the estimated symbol rate spectral lines to achieve the symbol rate estimate.

Proposition 2: For nonlinear system (17), if it satisfies *Assumption 1* and *Assumption 2*, and r^* is a constant, then there

Algorithm 1 Data-Driven BPF Design for Symbol Rate Estimation

```
Input:
  • Target average-to-peak power ratio r*.
  • Small positive constant σ.
  • Maximum iterations number t_max.
Initialize φ̂(1), r(1), η, ρ, β2, β1, μ(1), λ(1), u(1), Δu(1).
for t < t_max do
    φ̂(k+1) = φ̂(k) + η[Δr(k) − φ̂(k)Δu(k)]Δu(k)^T / (μ(k) + ||Δu(k)||^2)
    if sign(φ_i(k+1)) ≠ sign(φ_i(1)), i = 1,2 then
        φ̂_i(k+1) = φ̂_i(1)
    end if
    if ||φ̂(k+1)|| ≤ σ or ||Δu(k)||^2 ≤ σ then
        φ̂(k+1) = φ̂(1)
    end if
    u(k+1) = u(k) + ρφ̂(k+1)^T / (λ(k) + ||φ̂(k+1)||^2) [r* − r(k)]
    Calculate r_c(k+1) = [u2(k+1) + u1(k+1)]/2
    if u1(k+1) > u2(k+1) or u2(k+1) < r_c(k+1) or u1(k+1) > r_c(k+1)
     then
        u1(k+1) = u1(k), u2(k+1) = u2(k)
    end if
   Store u(k+1) and compute r(k+1)
    λ(k+1) = λ(k) − β1 [∂J(u(k+1))/∂u(k+1)]^T [∂u(k+1)/∂λ(k)]
    μ(k+1) = μ(k) − β2 [∂J(φ̂(k+1))/∂φ̂(k+1)] [∂φ̂(k+1)/∂μ(k)]^T
    if r(k+1) > r(k) then
        break
    end if
    t = t+1, k = k+1
end for
Find u(k+1) corresponding to the minimum r(k+1).
Generate the optimal BPF.
```

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3114678, IEEE Transactions on Wireless Communications

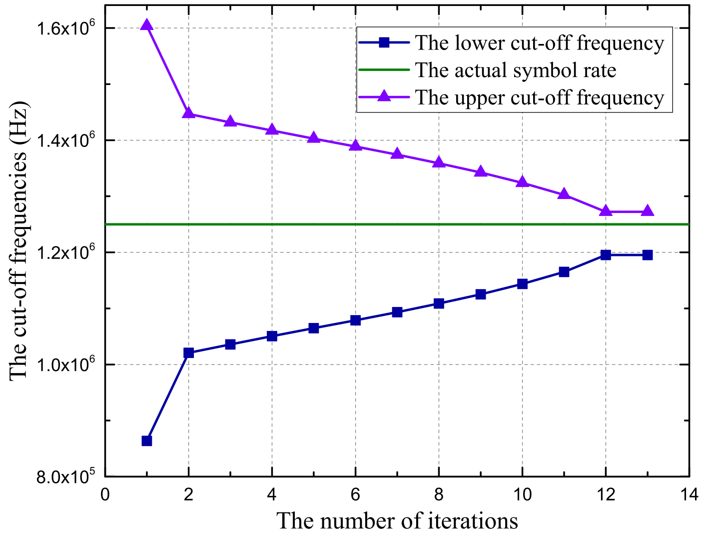

Fig. 5: The upper and lower cut-off frequencies as the iteration proceeds.

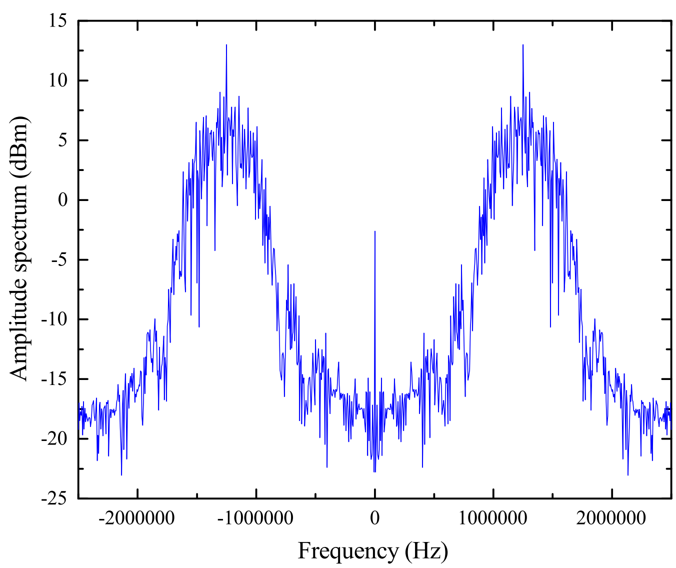

Fig. 6: Complex envelope spectrum obtained using the data-driven BPF.

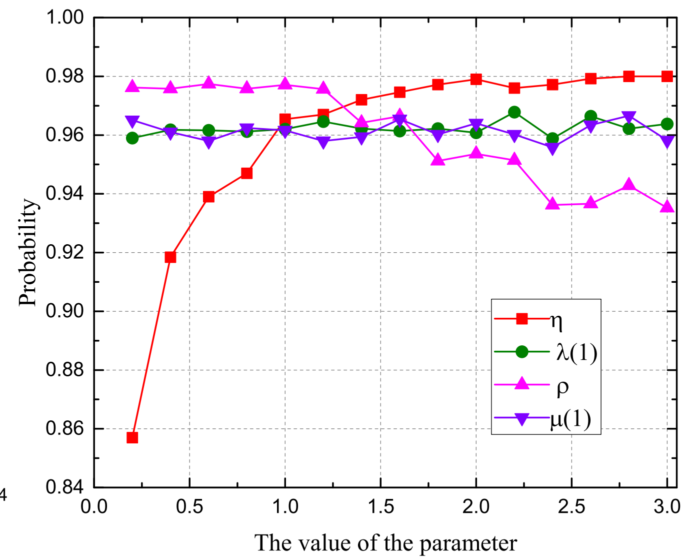

Fig. 7: The symbol rate estimation success probability of the proposed algorithm under different η, ρ, $\lambda(1)$ and $\mu(1)$.

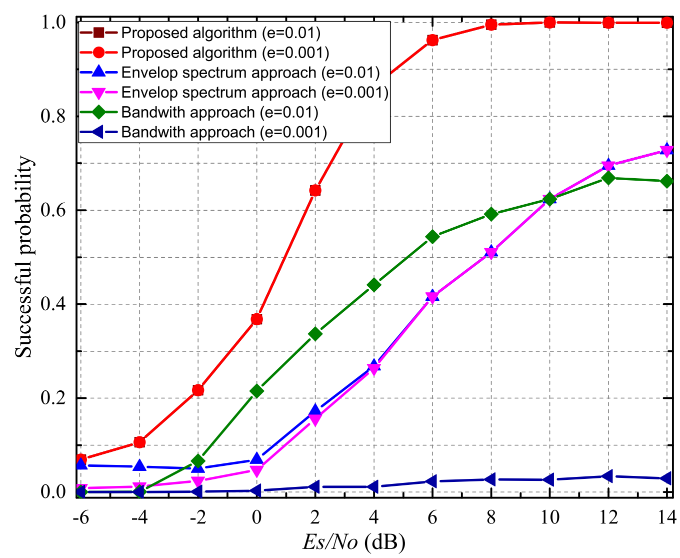

Fig. 8: Effective estimation probability of 16-PSK signals at different Es/No.

exists $\lambda(k) > \lambda_{\min} > 0$ that guarantees 1) $\lim_{k\to\infty} |r^* - r(k)| = 0$ and 2) $\{\boldsymbol{u}(k)\}$ and $\{r(k)\}$ are bounded sequences. ■

Proof : Refer to Appendix B.

Fig. 6 shows the complex envelope spectrum of 16-PSK signal processed by the optimal BPF, where the number of symbols is 200, the roll-off factor of the pulse shaping filter 0.35, the Es/No is 8 dB and the algorithm parameters $\hat{\varphi}(1)$, η, ρ, β_1, β_2, $\mu(1)$, $\lambda(1)$ set properly as detailed in Section IV below. Note that the optimal BPF filtered out the unnecessary noise around the symbol rate spectral lines, and thus improves the symbol rate estimation accuracy.

Here, we assess the computational complexity of the proposed symbol rate estimation method. The complexity includes two parts: 3dB bandwidth estimation and the data-driven BPF algorithm. The 3dB bandwidth estimation process estimates the received signal amplitude spectrum $X(f)$ and uses center of gravity method [15] to estimate the carrier frequency. Let N be the numbers of data samples used for estimating the estimated 3dB bandwidth B_w, the overall complexity of the bandwidth estimation method is $O(Nlog_2N)$. The data driven BPF only involves the operation of addition, subtraction, multiplication, division and circulation. Therefore, the overall complexity of the proposed algorithm is $O(Nlog_2N)$.

IV. Experimental Results and Analysis

This section evaluates the proposed estimator.

A. Simulation Results

We first compare the data-driven BPF algorithm with the traditional envelop spectrum method and the bandwidth method. Taking both 16-PSK and 16-QAM as examples, we assess the performance under various SNRs, roll-off factors, and numbers of data symbols. In the simulation experiment, the sampling frequency f_s is 5MHz, the symbol rate is 1.25Msym/s and the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3114678, IEEE Transactions on Wireless Communications

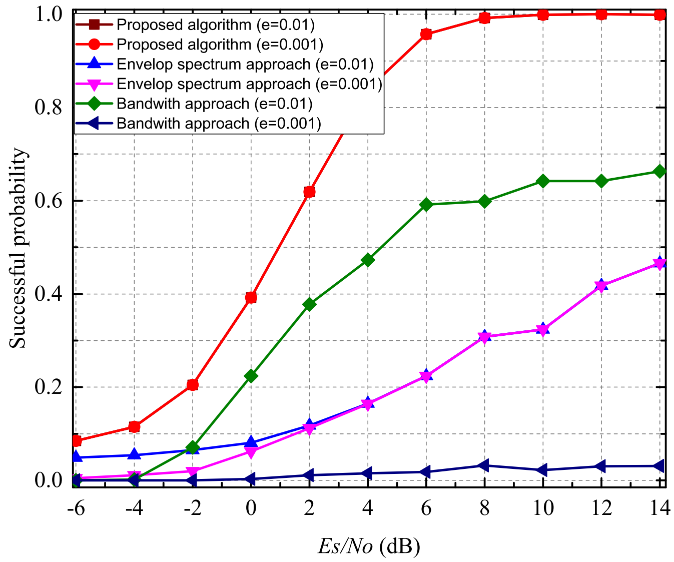

Fig. 9: Effective estimation probability of 16-QAM signals at different Es/No.

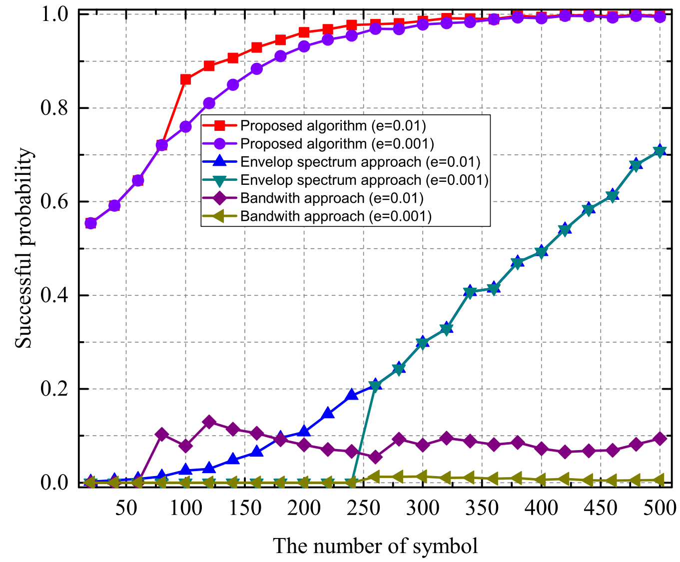

Fig. 11: Effective estimation probability of 16-QAM signals with different number of symbols.

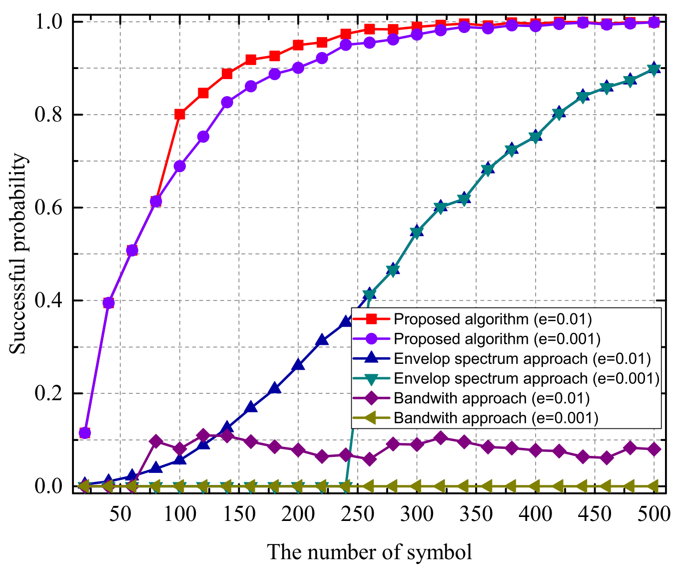

Fig. 10: Effective estimation probability of 16-PSK signals with different number of symbols.

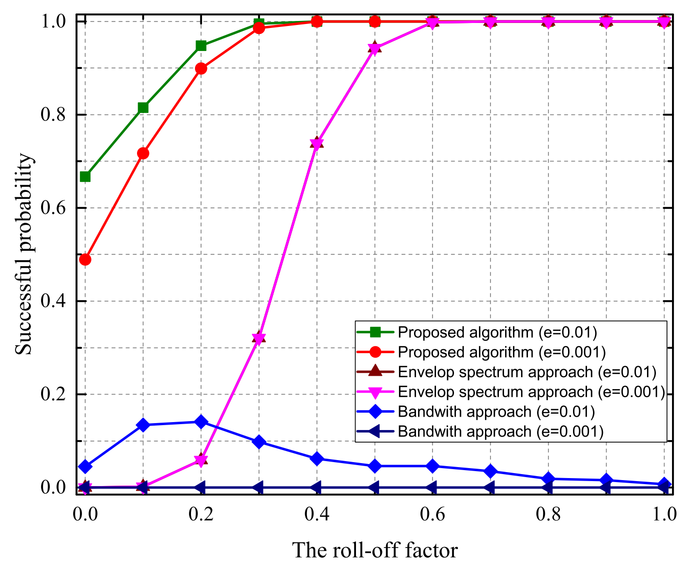

Fig. 12: Effective estimation probability of 16-PSK signals at different roll-off factor.

pulse filter span is 6. The channel is assumed to be additive white Gaussian noise (AWGN). Unless otherwise stated, the roll-off factor is 0.35, the number of symbols is 300, and Es/No is 8 dB.

The initialization parameters may affect the performance of the proposed data-driven symbol rate estimation algorithm. The step size factors η and ρ are used to make the control algorithm more general and more flexible. To quantify the influence of the parameters η, ρ, $\lambda(1)$ and $\mu(1)$ on the performance of symbol rate estimation, we compute the symbol rate estimation success probability of the proposed scheme under 5000 trials, where the tolerable estimation error e is equal to 0.001. We compare the performance of different initialization parameters on the symbol rate estimation success probability in Fig. 7. According to [30]–[32], the initial interval of these parameters is set to $(0,3]$. In the experiment, we choose a 16-QAM modulated signal, the SNR is set to 8, and the number of symbols of the signal is 300. Through 5000 Monte Carlo experiments, we found that the success probability of signal symbol rate estimation is above 0.94 when $\lambda(1)$, $\mu(1)$, η and ρ are in the interval of $[0.6,3]$. At the same time, the different $\lambda(1)$ and $\mu(1)$ in the interval [0.2,3] have little effect on the symbol rate estimation success probability. Therefore, according to the experiment results, we set $\lambda(1)=2$, $\mu(1)=1$, $\eta=1$, $\rho=1$. As for the choice of learning factors β_1 and β_2, if the learning factors are too small, the algorithm iteration speed will be slow and the algorithm optimization time will become long. Refer to the paper [39] and after trial debugging, we set the learning factors β_1 and β_2 to 0.05. In addition, $\hat{\varphi}(1)$ is the initial value of the PPD parameter $\hat{\varphi}(k)$, the PPD parameters can be optimized iteratively by (23) in the revised manuscript. Refer to [30], [40] and after experimental

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3114678, IEEE Transactions on Wireless Communications

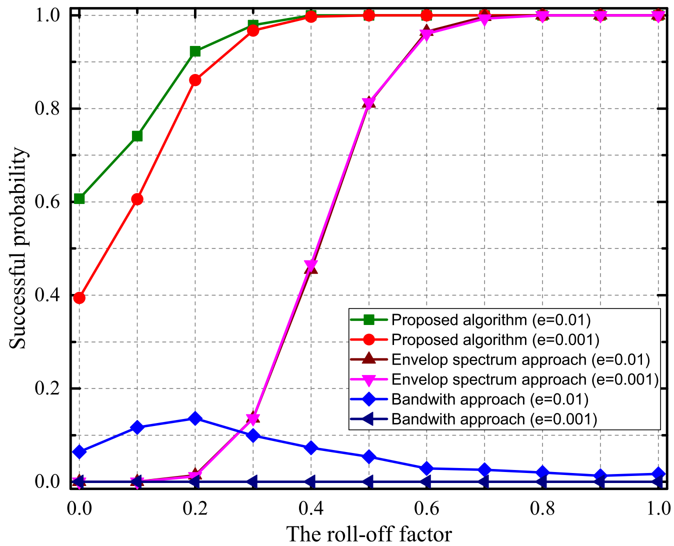

Fig. 13: Effective estimation probability of 16-QAM signals at different roll-off factor.

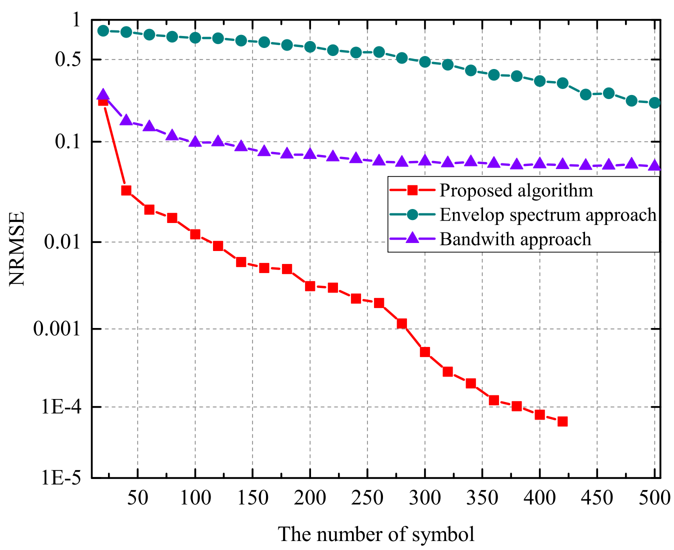

Fig. 14: NRMSE of 16-PSK signals with different number of symbols.

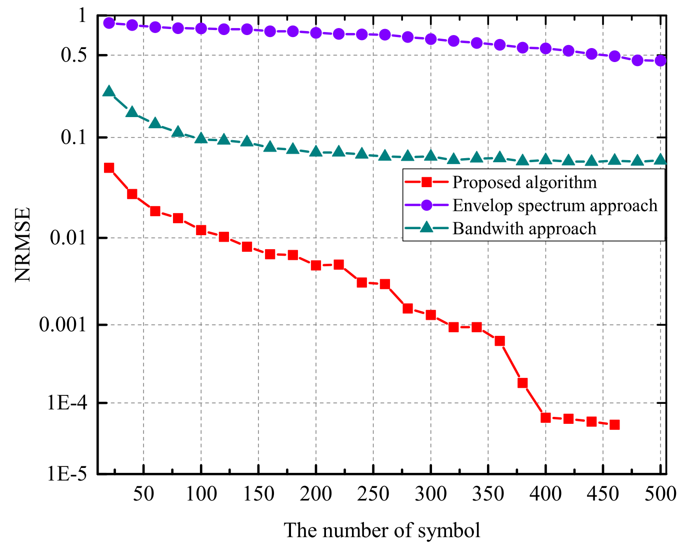

Fig. 15: NRMSE of 16-QAM signals with different number of symbols.

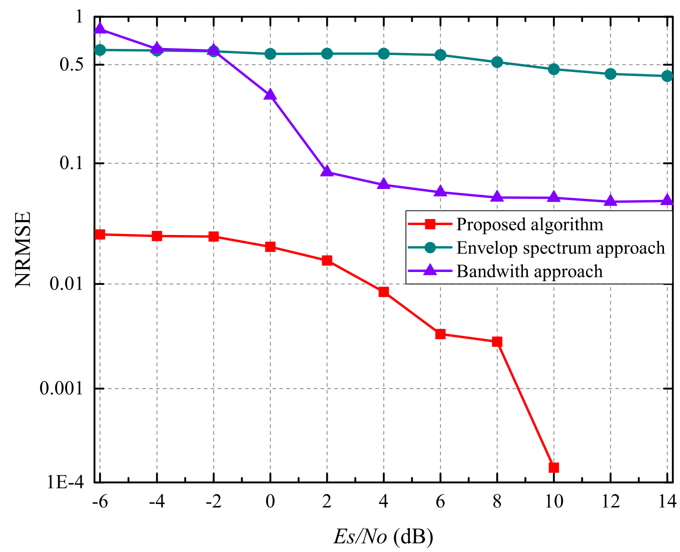

Fig. 16: NRMSE of 16-PSK signals at different Es/No.

debugging, then the other data driven algorithm parameters are set as: $\hat{\varphi}(1) = [-0.35, 0.35]$, $\Delta \boldsymbol{u}(1) = [0,0]^T$, $r(1) = 0.15$, $t_{\max} = 50$, $r* = 0.08$, $\sigma = 10^{-5}$.

Define a tolerable estimation error as $e(e > 0)$, which is the maximum difference allowed between the estimation and the true value of the symbol rate R_s. If the estimation error is smaller than e, we consider the estimation successful; otherwise the estimation fails. According to [18] [20], we consider two tolerable levels $e = 0.01$ and $e = 0.001$. In the numerical examples below, every probability is obtained from 5000 times Monte Carlo simulation.

In the first experiment, we compare the effective estimation probability of the proposed algorithm with traditional envelop spectrum approach and the bandwidth approach at different Es/No and modulation methods. As is shown in Fig. 8 and Fig. 9, the proposed data-driven BPF algorithm can achieve approximately 100% estimation probability when Es/No is around 8 dB for both modulation methods. In contrast, the traditional complex envelope spectrum method achieves merely 50% and 31% effective estimation probabilities for 16-PSK and 16-QAM, respectively, where Es/No is around 8 dB. Even if Es/No increases to 14 dB, the traditional envelop spectrum method can only achieve 72% accuracy for 16-PSK and less than 48% for 16-QAM signals, respectively. The effective estimation success rate of the bandwidth estimation method is not high under the two modulation methods. Even if Es/No is 14 dB, the effective estimation probability achieved by the bandwidth method is less than 70%. The results demonstrate the superiority of the proposed data-driven BPF scheme over the traditional method and bandwidth method, especially in the low SNR region.

In the second experiment, we vary the number of symbols and compare the proposed data-driven BPF algorithm with the traditional envelop spectrum method and the bandwidth method. Fig. 10 and Fig. 11 compares the proposed data-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3114678, IEEE Transactions on Wireless Communications

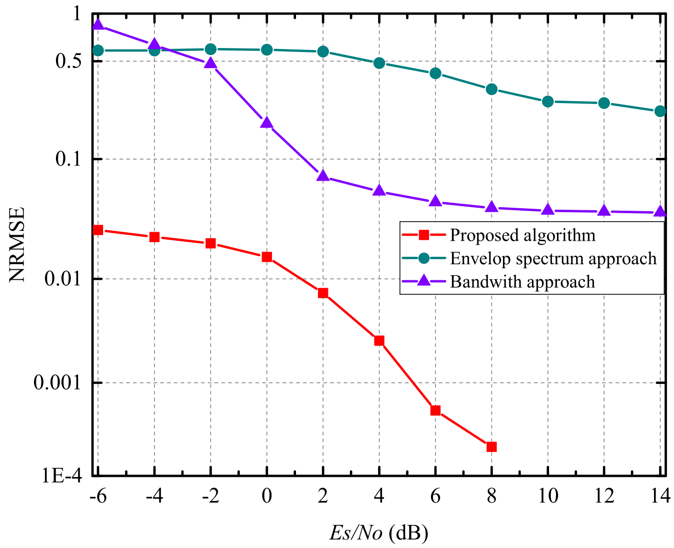

Fig. 17: NRMSE of 16-QAM signals at different Es/No.

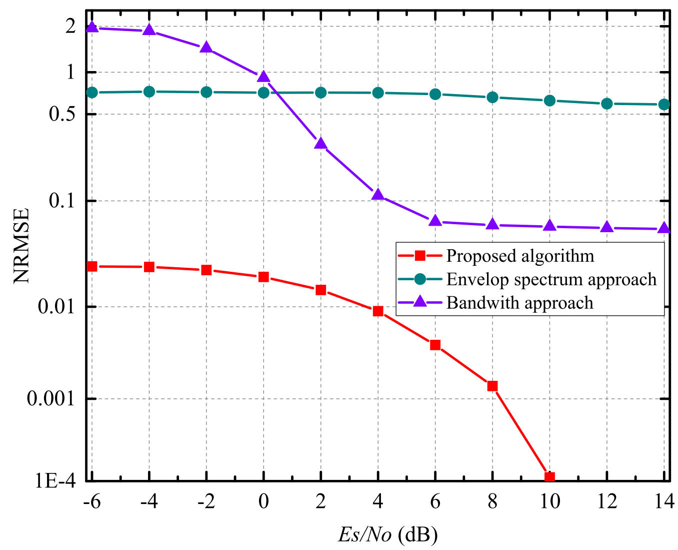

Fig. 19: NRMSE of 64-QAM signals at different Es/No.

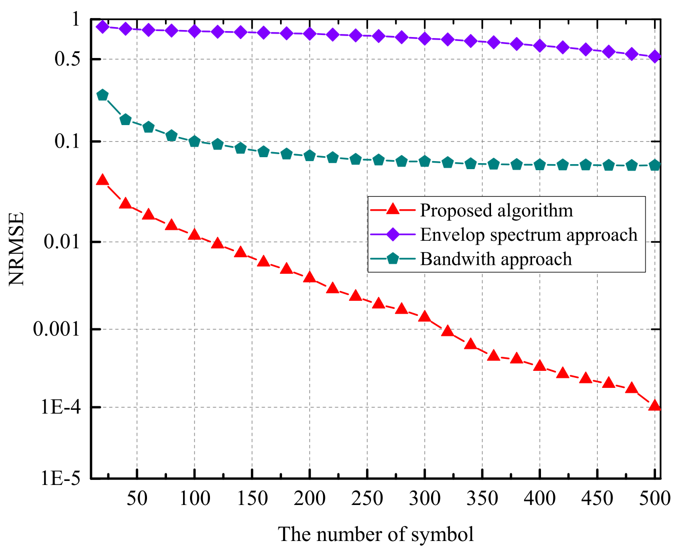

Fig. 18: NRMSE of 64-QAM signals with different number of symbols.

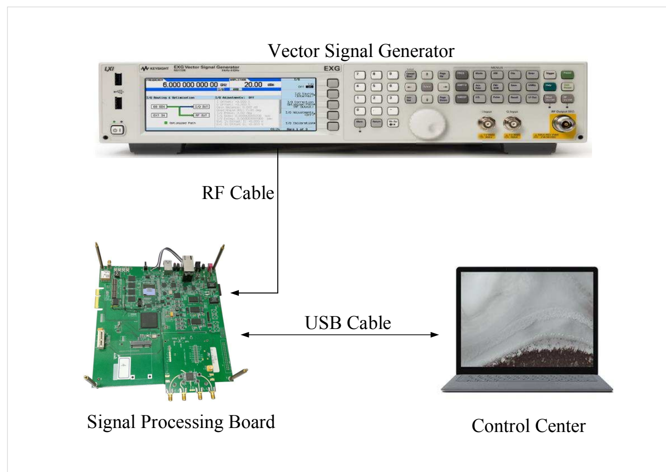

Fig. 20: Real-world experiment setup for symbol rate estimation.

driven algorithm against the envelope spectrum method and the bandwidth method in terms of effective estimation probability for two modulation methods. It can be seen that the 16-PSK signal can achieve an estimation success rate of 90% with only 200 symbols using the proposed algorithm, as compared to 500 symbols using the traditional complex envelope method. Similarly, to achieve 90% effective estimation probability, the 16-QAM signal needs only 180 symbols using the proposed algorithm, while the traditional complex envelope method requires more than 500 symbols. The symbol rate estimation success rate of the bandwidth method for 16-PSK and 16-QAM modulation modes with number of symbols less than 500 are all lower than 15%. Moreover, when the number of signal symbols is less than or equal to 240, $e \leq 0.001$, the proposed algorithm can achieve over 95% effective estimation probability while the traditional complex envelope spectrum method fails most of the time.

In Fig. 12 and Fig. 13, we study the effect of the roll-off factor. As seen, the proposed algorithm achieves close to 100% estimation probability at 0.4 roll-off factor, while the traditional complex envelope spectrum method requires 0.6 and 0.7 roll-off-factor for 16-PSK and 16-QAM respectively. It is well known that the smaller the roll-off factor, the more serious the inter-symbol interference of the signal. However, when the roll-off factor is small, the proposed algorithm still has a high symbol rate estimation success rate, which shows that the proposed algorithm has strong anti-interference ability.

We also use the normalized root mean square error (N-RMSE) to evaluate the performance of the proposed symbol rate estimation algorithm. Define

$$NRMSE = \sqrt{\sum_{i=1}^{N} \left(R_s - \hat{R}_s(i)\right)^2 \Big/ NR_s^2}, \qquad (33)$$

where $\hat{R}_s(i)$ is the i-th estimated symbol rate and R_s is the actual symbol rate.

Fig. 14 and Fig. 15 illustrate the NRMSE of the symbol rate estimation for 16-PSK/16-QAM signals, respectively. The

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3114678, IEEE Transactions on Wireless Communications

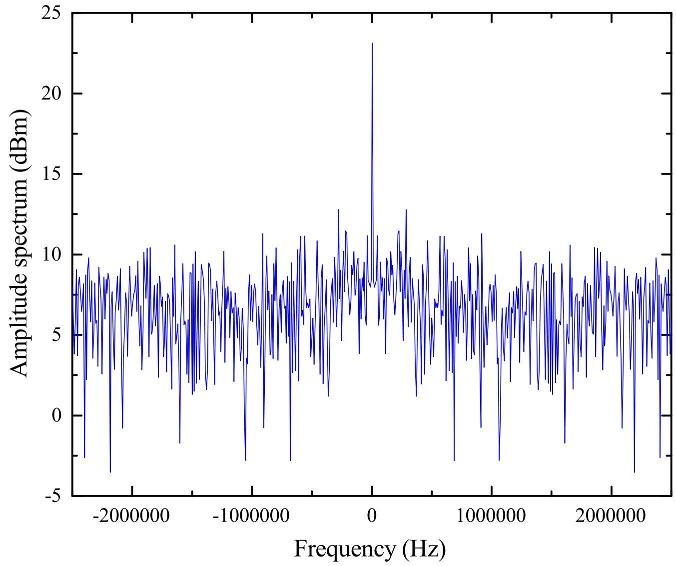

Fig. 21: The complex envelope spectrum of the obtained data

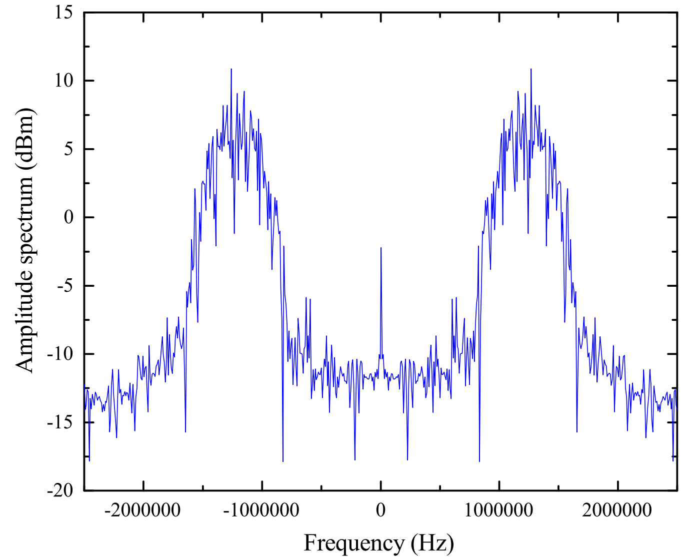

Fig. 22: The complex envelope spectrum of the obtained data after the BPF.

proposed symbol rate estimation algorithm significantly outperforms the traditional complex envelope spectrum method and bandwidth method. Specifically, for 16-PSK, the NRMSE of the complex envelope spectrum method is higher than 0.2 when the number of symbols is less than 500, the bandwidth method is higher than 0.08 when the number of symbols is less than 500, while that of the proposed algorithm is close to 0.001 when the number of symbols is less than 280. For 16-QAM, the NRMSE of the complex envelope spectrum method is higher than 0.4 when the number of symbols is less than 500, the bandwidth approach is higher than 0.08 when the number of symbols is less than 500, while that of the proposed algorithm is close to 0.001 when the number of symbols is less than 300. Fig. 16 and Fig. 17 show symbol rate estimation for 16-PSK/16-QAM under various Es/No. When Es/No is greater than 4 dB, the NRMSE of the proposed algorithm is is less than 0.01, but that of the traditional method is greater than 0.5, the NRMSE of the bandwidth method is about 0.09.

To evaluate the performance of our proposed algorithm on other M-ary modulated signals, we conducted simulation experiments on 64-QAM modulation signals in Fig. 18 and Fig. 19. The experiment results show that the proposed data-driven algorithm can also be applied to other higher order modulation signals. Fig. 19 illustrate the NRMSE of the symbol rate estimation for 64-QAM signal under various Es/No. The result shows that the performance of the proposed data-driven algorithm significantly outperforms the traditional complex envelope spectrum method and 3dB bandwidth method. Under the same low Es/No, the NRMSE of the proposed symbol rate estimation algorithm is smaller than that of the other two algorithms. When Es/No is greater than 4, the NRMSE of the proposed symbol rate estimation algorithm is less than 0.01. At the same time when the number of symbols is more than 320, Fig. 18 shows the NRMSE of the proposed symbol rate estimation algorithm is less than 0.001. Both figures show the effectiveness of the proposed algorithm in estimating the symbol rate of higher-order modulation methods.

B. Hardware Experiment

The experiment setup is as follows. We use an N5172B EXG series RF vector signal generator to generate the 16-PSK signal at a carrier frequency of 2.4GHz with an output power of -50 dBm. The signal is sampled by an AD9361 RF front-end controlled by a Xilinx Artix-7 FPGA AC701 development board, as shown in Fig. 20. After down-converting it, we collect 584 data samples to evaluate the performance of the symbol rate estimation. The sampling rate of the down-converted signal is 10MHz. The spectrum of the traditional complex envelope spectrum method is depicted in Fig. 21, where the real symbol rate spectral lines are buried under noise and difficult to identify. When we use the square complex envelope method to estimate the symbol rate of the obtained down-converted signal, the estimated value of symbol rate is 0.2863Mhz, and its normalized estimation error is 77.1%. This shows that the square complex envelope method is completely failed at this time, and it cannot complete the symbol rate estimation with a small number of symbols. By using the 3dB bandwidth method to estimate the symbol rate of the processed signal, we can get the estimated symbol rate of 1.130Mhz, which has a normalized estimation error of 9.6%. The error is also very large. In comparison, as shown in Fig. 22, after the signal passes through the data-driven BPF, the symbol rate spectral lines become clearly visible. The normalized estimation error between the estimated symbol rate $\hat{R}_s$=1.2521MHz and the actual symbol rate R_s=1.25MHz is 0.168%, and is within the acceptance range.

V. Conclusion

This paper has developed a new symbol rate estimation approach for wireless communication systems at low SNR and with a small number of symbols. We proposed a data-driven BPF design scheme that can continuously optimize the upper and lower cut-off frequencies based on the measured average-to-peak power ratio of the complex envelope spectrum, until the average-to-peak power ratio converges to the minimum.

The optimal BPF is used to process the signal's complex envelope and extract spectral peaks as the symbol rate estimate. The numerical and experimental results demonstrate that the proposed data-driven BPF design scheme can effectively improve the accuracy of symbol rate estimation at low SNR and with a small number of symbols.

Appendix A
Proof of Proposition 1

proof : From the dynamic linearization data model (17), we have:

$$\begin{aligned}\Delta r(k) &= r(k)-r(k-1)\\ &= f(r(k-1),...,r(k-n_p),\boldsymbol{u}(k),...,\boldsymbol{u}(k-n_s))\\ &\quad - f(r(k-2),...,r(k-n_p-1),\boldsymbol{u}(k-1),...,\boldsymbol{u}(k-n_s-1))\\ &= f(r(k-1),r(k-2),...,r(k-n_p),\boldsymbol{u}(k),\boldsymbol{u}(k-1),\\ &\quad ...,\boldsymbol{u}(k-n_s))-f(r(k-1),r(k-2),...,r(k-n_p),\boldsymbol{u}(k-1),\\ &\quad \boldsymbol{u}(k-1),...,\boldsymbol{u}(k-n_s))+f(r(k-1),r(k-2),...,\\ &\quad r(k-n_p),\boldsymbol{u}(k-1),\boldsymbol{u}(k-1),...,\boldsymbol{u}(k-n_s))-f(r(k-2),\\ &\quad ...,r(k-n_p),\boldsymbol{u}(k-1),\boldsymbol{u}(k-2),...,\boldsymbol{u}(k-n_s-1)).\end{aligned} \tag{34}$$

Let

$$\begin{aligned}\psi(k) =& f(r(k-1),r(k-2),...,r(k-n_p),\boldsymbol{u}(k-1),\\ &\boldsymbol{u}(k-1),...,\boldsymbol{u}(k-n_s))\\ &-f(r(k-2),r(k-3),...,r(k-n_p-1),\boldsymbol{u}(k-1),\\ &\boldsymbol{u}(k-2),...,\boldsymbol{u}(k-n_s-1)).\end{aligned} \tag{35}$$

By the differential mean value theorem [41] and *Assumption 1*, (34) can be rewritten as

$$\Delta r(k)=\frac{\partial f^*}{\partial \boldsymbol{u}(k)}\Delta\boldsymbol{u}(k)+\psi(k), \tag{36}$$

where

$$\frac{\partial f^*}{\partial \boldsymbol{u}(k)}=\left[\frac{\partial f^*}{\partial u_1(k)},\frac{\partial f^*}{\partial u_2(k)}\right], \tag{37}$$

and $\partial f^*/\partial u_i(k), i\in 1,2$ is the partial derivative of $f(\cdot)$ with respect to the cut-off frequency $u_i(k)$ at some point between in the interval $[u_i(k),u_i(k-1)]$. For each fixed k, we apply the following equation

$$\psi(k)=\boldsymbol{\xi}(k)\Delta\boldsymbol{u}(k). \tag{38}$$

Since $\|\Delta\boldsymbol{u}(k)\|\neq 0$, (38) must have at least one solution $\boldsymbol{\xi}^*(k)$ for each k. Let

$$\varphi(k)=\frac{\partial f^*}{\partial \boldsymbol{u}(k)}+\boldsymbol{\xi}^*(k). \tag{39}$$

Based on (39) and (38), (36) can be written as

$$\Delta r(k)=\varphi(k)\Delta\boldsymbol{u}(k), \tag{40}$$

where $\varphi(k)$ is called the PPD parameter vector at the k-th iteration. Therefore, we have $\|\varphi(k)\|\leq b$ based on *Assumption 2*. ■

Appendix B
Proof of Proposition 2

Proof : This proof includes two steps. The first step is to prove the boundness of the estimated PPD parameter and the second is to prove the convergence and stability of the proposed algorithm.

$\hat{\varphi}(k)$ is bounded if it satisfied one of three cases: $\|\hat{\varphi}(k)\|\leq\sigma$ or $\|\Delta u(k)\|^2\leq\sigma$ or $sign(\hat{\varphi}_i(k))\neq sign(\hat{\varphi}_i(1)$. Otherwise, define the PPD parameter vector estimation error as $\tilde{\varphi}(k)=\hat{\varphi}(k)-\varphi(k)$ and subtract $\varphi(k)$ from both sides of equation (23), then

$$\begin{aligned}\tilde{\varphi}(k) &= \hat{\varphi}(k-1)-\varphi(k-1)+\varphi(k-1)-\varphi(k)\\ &\quad+\frac{\eta\left[\varphi(k-1)\Delta\boldsymbol{u}(k-1)-\hat{\varphi}(k-1)\Delta\boldsymbol{u}(k-1)\right]\Delta\boldsymbol{u}(k-1)^T}{\mu(k-1)+\|\Delta\boldsymbol{u}(k-1)\|^2}\\ &= \varphi(k-1)-\varphi(k)+\tilde{\varphi}(k-1)\\ &\quad-\frac{\eta\tilde{\varphi}(k-1)\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T}{\mu(k-1)+\|\Delta\boldsymbol{u}(k-1)\|^2}.\end{aligned} \tag{41}$$

From *Proposition 1*, we have $\|\varphi(k-1)-\varphi(k)\|\leq 2b$ since $\|\varphi(k)\|\leq b$. Take the norms on both sides of formula (41), and then we can get

$$\|\tilde{\varphi}(k)\|\leq\left\|\tilde{\varphi}(k-1)\left[\boldsymbol{I}-\frac{\eta\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T}{\mu(k-1)+\|\Delta\boldsymbol{u}(k-1)\|^2}\right]\right\|+2b. \tag{42}$$

By squaring the first term of right side of formula (42), we have

$$\begin{aligned}&\left\|\tilde{\varphi}(k-1)\left[\boldsymbol{I}-\frac{\eta\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T}{\mu(k-1)+\|\Delta\boldsymbol{u}(k-1)\|^2}\right]\right\|^2\\ &=\|\tilde{\varphi}(k-1)\|^2\\ &+\left[\frac{\eta\|\Delta\boldsymbol{u}(k-1)\|^2}{\mu(k-1)+\|\Delta\boldsymbol{u}(k-1)\|^2}-2\right]\times\frac{\eta\|\tilde{\varphi}(k-1)\Delta\boldsymbol{u}(k-1)\|^2}{\mu(k-1)+\|\Delta\boldsymbol{u}(k-1)\|^2}.\end{aligned} \tag{43}$$

Since $0<\eta\leq 2$, $\mu(k-1)\geq\mu_{\min}>0$, we have

$$\left[-2+\frac{\eta\|\Delta\boldsymbol{u}(k-1)\|^2}{\mu(k-1)+\|\Delta\boldsymbol{u}(k-1)\|^2}\right]<0. \tag{44}$$

From (43) and (44), we can deduce that there exists $0<d_1<1$, such that

$$\left\|\tilde{\varphi}(k-1)\left[\boldsymbol{I}-\frac{\eta\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T}{\mu(k-1)+\|\Delta\boldsymbol{u}(k-1)\|^2}\right]\right\|\leq d_1\|\tilde{\varphi}(k-1)\|. \tag{45}$$

Please note that the value of d_1 is not required here as long as we know its existence. Substituting (45) into (42) yields

$$\begin{aligned}\|\tilde{\varphi}(k)\| &\leq d_1\|\tilde{\varphi}(k-1)\|+2b\leq d_1(d_1\|\tilde{\varphi}(k-2)\|+2b)+2b\\ &\leq ...<d_1^{k-1}\|\tilde{\varphi}(1)\|+\frac{2b}{1-d_1}.\end{aligned} \tag{46}$$

The inequality (46) means that $\tilde{\varphi}(k)$ is bounded. According to *Proposition 1*, $\|\varphi(k)\|<b$, thus $\hat{\varphi}(k)$ is bounded.

In the second step, we prove that $\{r(k)\}$ is bounded. Define

$$e(k)=r^*-r(k). \tag{47}$$

By equations (26) and (28), we can get the system tracking error

$$\begin{aligned} e(k) &= r^* - [r(k-1) + \hat{\varphi}(k)\Delta \boldsymbol{u}(k)] \\ &= r^* - r(k-1) - \frac{\rho\hat{\varphi}(k)\hat{\varphi}(k)^T[r^* - r(k-1)]}{\lambda(k-1) + \|\hat{\varphi}(k)\|^2}. \end{aligned} \tag{48}$$

Take the absolute value of both sides of (48)

$$\begin{aligned} |e(k)| &= |r^* - r(k-1) - \hat{\varphi}(k)\Delta \boldsymbol{u}(k)| \\ &= |e(k-1)| \left| 1 - \frac{\rho\hat{\varphi}(k)\hat{\varphi}(k)^T}{\lambda(k-1) + \|\hat{\varphi}(k)\|^2} \right|. \end{aligned} \tag{49}$$

Since $\hat{\varphi}(k)\hat{\varphi}(k)^T = \|\hat{\varphi}(k)\|^2 > 0$ and $0 < \rho \leq 1$, there exists a constant $0 < c < 1$ such that

$$0 < 1 - \frac{\rho\hat{\varphi}(k)\hat{\varphi}(k)^T}{\lambda(k-1) + \|\hat{\varphi}(k)\|^2} \leq c < 1. \tag{50}$$

From (49) and (50), we have

$$|e(k)| \leq c\,|e(k-1)| \leq c^2\,|e(k-2)| \leq ... c^{k-1}\,|e(1)|. \tag{51}$$

This guarantees that as k increases, the system tracking error $e(k)$ decreases exponentially. This implied that the output error of the MISO system is bounded. As r^* is a constant, $\{r(k)\}$ is also bounded.

From (28) and $\lambda(k-1) + \|\hat{\varphi}(k)\|^2 \geq 2\sqrt{\lambda_{\min}\|\hat{\varphi}(k)\|^2}$,

$$\begin{aligned} \|\Delta\boldsymbol{u}(k)\| &= \left\| \frac{\rho\hat{\varphi}(k)^T[r^* - r(k-1)]}{\lambda(k-1) + \|\hat{\varphi}(k)\|^2} \right\| \\ &\leq \frac{\left\|\rho\hat{\varphi}(k)^T\right\| |e(k)|}{\left\|\lambda(k-1) + \|\hat{\varphi}(k)\|^2\right\|} \\ &\leq \frac{\left\|\rho\hat{\varphi}(k)^T\right\| |e(k)|}{\left\|2\sqrt{\lambda_{\min}\|\hat{\varphi}(k)\|^2}\right\|}. \end{aligned} \tag{52}$$

Let

$$\frac{\left\|\rho\hat{\varphi}(k)^T\right\|}{\left\|2\sqrt{\lambda_{\min}\|\hat{\varphi}(k)\|^2}\right\|} = P, \tag{53}$$

where P is a bounded constant, then

$$\begin{aligned} \|\boldsymbol{u}(k)\| &= \|\boldsymbol{u}(k) - \boldsymbol{u}(k-1) + \boldsymbol{u}(k-1) - ... - \boldsymbol{u}(1) + \boldsymbol{u}(1)\| \\ &\leq \|\boldsymbol{u}(k) - \boldsymbol{u}(k-1)\| + ... + \|\boldsymbol{u}(2) - \boldsymbol{u}(1)\| + \|\boldsymbol{u}(1)\| \\ &= \|\Delta\boldsymbol{u}(k)\| + \|\Delta\boldsymbol{u}(k-1)\| + ... + \|\Delta\boldsymbol{u}(2)\| + \|\boldsymbol{u}(1)\| \\ &\leq pc^{k-1}\,|e(1)| + pc^{k-2}\,|e(1)| + ... + pc\,|e(1)| + \|\boldsymbol{u}(1)\| \\ &< \frac{pc}{1-c}\,|e(1)| + \|\boldsymbol{u}(1)\|. \end{aligned} \tag{54}$$

Therefore, the sequence $\{\boldsymbol{u}(k)\}$ is bounded. ■

APPENDIX C

Matrix Inversion Lemma [37]: $\mathbf{A}$ is a positive-definite matrix, $\mathbf{B}$ is an $\mathrm{n} \times r$ matrix, then the inverse of the matrix $\mathbf{A}$ added to a block of dyads (represented as $\mathbf{B}\mathbf{B}^{\mathbf{H}}$) can be represented as:

$$(\mathbf{A} + \mathbf{B}\mathbf{B}^{\mathbf{H}})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{B}^{\mathbf{H}}\mathbf{A}^{-1}\mathbf{B} + \mathbf{I}^{-1})^{-1}\mathbf{B}^{\mathbf{H}}\mathbf{A}^{-1}, \tag{55}$$

where superscript $\mathbf{H}$ denotes the complex conjugate transpose operation. According to equations (20) and (21) and the optimization condition $\frac{\partial J(\hat{\varphi}(k))}{\partial\hat{\varphi}(k)} = 0$, we can get

$$\hat{\varphi}(k) \times \boldsymbol{\Theta} = [\mu(k-1)\hat{\varphi}(k-1) + \Delta r(k-1)\Delta\boldsymbol{u}(k-1)^T], \tag{56}$$

where $\boldsymbol{\Theta} = [\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T + \mu(k-1)\mathbf{I}]$. then

$$\begin{aligned} \hat{\varphi}(k) =& [\mu(k-1)\hat{\varphi}(k-1) + \Delta r(k-1)\Delta\boldsymbol{u}(k-1)^T] \times \boldsymbol{\Theta}^{-1} \\ =& \mu(k-1)\hat{\varphi}(k-1) \times \boldsymbol{\Theta}^{-1} \\ &+ \hat{\varphi}(k-1)\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T \times \boldsymbol{\Theta}^{-1} \\ &- \hat{\varphi}(k-1)\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T \times \boldsymbol{\Theta}^{-1} \\ &+ \Delta r(k-1)\Delta\boldsymbol{u}(k-1)^T \times \boldsymbol{\Theta}^{-1} \\ =& \hat{\varphi}(k-1)[\mu(k-1)\mathbf{I} + \Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T] \times \boldsymbol{\Theta}^{-1} \\ &+ [\Delta r(k-1) - \hat{\varphi}(k-1)\Delta\boldsymbol{u}(k-1)]\Delta\boldsymbol{u}(k-1)^T \times \boldsymbol{\Theta}^{-1} \\ =& \hat{\varphi}(k-1) \\ &+ [\Delta r(k-1) - \hat{\varphi}(k-1)\Delta\boldsymbol{u}(k-1)]\Delta\boldsymbol{u}(k-1)^T \times \boldsymbol{\Theta}^{-1}. \end{aligned} \tag{57}$$

Let $\mathbf{A} = \mu(k-1)\mathbf{I}$, $\mathbf{B} = \Delta\boldsymbol{u}(k-1)$. Since $\Delta\boldsymbol{u}(k-1)$ is a real vector, $\mathbf{B}^{\mathbf{H}} = \Delta\boldsymbol{u}(k-1)^T$, then $\mathbf{A}^{-1} = \frac{\mathbf{I}}{\mu(k-1)}$, so

$$\begin{aligned} \mathbf{B}^{\mathbf{H}}\mathbf{A}^{-1}\mathbf{B} + \mathbf{I}^{-1} &= \Delta\boldsymbol{u}(k-1)^T\frac{\mathbf{I}}{\mu(k-1)}\Delta\boldsymbol{u}(k-1) + \mathbf{I} \\ &= \frac{\|\Delta\boldsymbol{u}(k-1)\|^2 + \mu(k-1)}{\mu(k-1)}. \end{aligned} \tag{58}$$

Applying the *Matrix Inversion Lemma* to $[\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T + \mu(k-1)\mathbf{I}]^{-1}$ and combining equation (58), we can get

$$\begin{aligned} &[\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T + \mu(k-1)\mathbf{I}]^{-1} \\ =& \frac{\mathbf{I}}{\mu(k-1)} - \left[\frac{\|\Delta\boldsymbol{u}(k-1)\|^2 + \mu(k-1)}{\mu(k-1)}\right]^{-1} \frac{\Delta\boldsymbol{u}(k-1)}{\mu(k-1)}\frac{\Delta\boldsymbol{u}(k-1)^T}{\mu(k-1)} \\ =& \frac{\mathbf{I}}{\mu(k-1)} - \frac{\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T}{\mu(k-1) \times [\|\Delta\boldsymbol{u}(k-1)\|^2 + \mu(k-1)]} \\ =& \frac{[\|\Delta\boldsymbol{u}(k-1)\|^2 + \mu(k-1)]\mathbf{I}}{\mu(k-1) \times [\|\Delta\boldsymbol{u}(k-1)\|^2 + \mu(k-1)]} \\ &- \frac{\Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T}{\mu(k-1) \times [\|\Delta\boldsymbol{u}(k-1)\|^2 + \mu(k-1)]} \\ =& \frac{\|\Delta\boldsymbol{u}(k-1)\|^2\mathbf{I} + \mu(k-1)\mathbf{I} - \Delta\boldsymbol{u}(k-1)\Delta\boldsymbol{u}(k-1)^T}{\mu(k-1) \times [\|\Delta\boldsymbol{u}(k-1)\|^2 + \mu(k-1)]}. \end{aligned} \tag{59}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3114678, IEEE Transactions on Wireless Communications

Substituting (59) into (57), and we get

$$\begin{aligned}\hat{\varphi}(k) =&\hat{\varphi}(k-1)+[\Delta r(k-1)-\hat{\varphi}(k-1)\Delta \boldsymbol{u}(k-1)]\Delta \boldsymbol{u}(k-1)^T \\ &\times \frac{[\|\Delta \boldsymbol{u}(k-1)\|^2 \mathbf{I}+\mu(k-1)\mathbf{I}-\Delta \boldsymbol{u}(k-1)\Delta \boldsymbol{u}(k-1)^T]}{\mu(k-1)\times[\|\Delta \boldsymbol{u}(k-1)\|^2+\mu(k-1)]} \\ =&\hat{\varphi}(k-1)+[\Delta r(k-1)-\hat{\varphi}(k-1)\Delta \boldsymbol{u}(k-1)] \\ &\times \frac{[\|\Delta \boldsymbol{u}(k-1)\|^2 \Delta \boldsymbol{u}(k-1)^T+\mu(k-1)\Delta \boldsymbol{u}(k-1)^T-\boldsymbol{U}]}{\mu(k-1)\times[\|\Delta \boldsymbol{u}(k-1)\|^2+\mu(k-1)]} \\ =&\hat{\varphi}(k-1)+[\Delta r(k-1)-\hat{\varphi}(k-1)\Delta \boldsymbol{u}(k-1)] \\ &\times \frac{[\mu(k-1)\Delta \boldsymbol{u}(k-1)^T]}{\mu(k-1)\times[\|\Delta \boldsymbol{u}(k-1)\|^2+\mu(k-1)]} \\ =&\hat{\varphi}(k-1)+\frac{[\Delta r(k-1)-\hat{\varphi}(k-1)\Delta \boldsymbol{u}(k-1)]\Delta \boldsymbol{u}(k-1)^T}{[\|\Delta \boldsymbol{u}(k-1)\|^2+\mu(k-1)]},\end{aligned} \tag{60}$$

where $\boldsymbol{U}=\Delta \boldsymbol{u}(k-1)^T\Delta \boldsymbol{u}(k-1)\Delta \boldsymbol{u}(k-1)^T$. To make the control algorithm more more general and more flexible, a step factor $\eta \in (0,2]$ is added, so

$$\begin{aligned}\hat{\varphi}(k) =&\hat{\varphi}(k-1) \\ &+\frac{\eta[\Delta r(k-1)-\hat{\varphi}(k-1)\Delta \boldsymbol{u}(k-1)]\Delta \boldsymbol{u}(k-1)^T}{\mu(k-1)+\|\Delta \boldsymbol{u}(k-1)\|^2}.\end{aligned} \tag{61}$$

■

REFERENCES

[1] J. L. Yang, W. J. Yang, and G. Q. Xu, "Blind estimation of carrier frequency and symbol rate based on cyclic spectrum density," *Procedia Engineer.*, vol. 29, no. 4, p. 514ĺC519, 2012.

[2] C. Shen and T. Ye, "Improving symbol rate estimation accuracy by bandpass filter bank," in *2019 IEEE 19th Int. Conf. Commun. Technol. (ICCT)*, 2019, pp. 203–209.

[3] B. Suesser-Rechberger and W. Gappmair, "New results on symbol rate estimation in digital satellite receivers," in *2018 11th Int. Symposium. Commun. Systems, Networks Digital Signal Process. (CSNDSP)*, 2018, pp. 1–6.

[4] B. Suesser Rechberger and W. Gappmair, "Asymptotic CRLB for blind symbol rate estimation in digital satellite receivers," in *2020 12th Int. Symposium. Commun. Systems, Networks and Digital Signal Process. (CSNDSP)*, 2020, pp. 1–6.

[5] S. Majhi and T. S. Ho, "Blind symbol-rate estimation and test bed implementation of linearly modulated signals," *IEEE Trans. Veh. Technol.*, vol. 64, no. 3, pp. 954–963, March 2015.

[6] Q. Zhang, O. A. Dobre, Y. A. Eldemerdash, S. Rajan, and R. Inkol, "Second-order cyclostationarity of BT-SCLD signals: Theoretical developments and applications to signal classification and blind parameter estimation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1501–1511, 2013.

[7] M. Gao, Y. Li, O. A. Dobre, and N. Al-Dhahir, "Blind identification of SFBC-OFDM signals based on the central limit theorem," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3500–3514, 2019.

[8] H. Wu, M. Saquib, and Z. Yun, "Novel automatic modulation classification using cumulant features for communications via multipath channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3098–3105, 2008.

[9] H. Wymeersch and M. Moeneclaey, "Blind symbol rate detection for low-complexity multi-rate receivers," in *2005 IEEE 61st Vehicular Technol. Conf.*, vol. 2, 2005, pp. 1171–1175 Vol. 2.

[10] C. Mosquera, S. Scalise, and R. Ĺópez-Valcarce, "Non-data-aided symbol rate estimation of linearly modulated signals," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 664–674, 2008.

[11] S. Majhi, M. Kumar, and W. Xiang, "Implementation and measurement of blind wireless receiver for single carrier systems," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 8, pp. 1965–1975, 2017.

[12] W. G. M. Flohberger and O. Koudelka, "Blind demodulation framework for satellite signals," in *Proc. 3rd Int. ICST Conf. Personal Satellite Services*, pp. 236–249.

[13] B. Suesser-Rechberger and W. Gappmair, "Digital video broadcasting (DVB); second generation framing structure channel coding and modulation systems for broadcasting interactive services news gathering and other broadband satellite applications; part ii: S2-extensions (S2-X)," in *ETSI Standard*, 2014.

[14] Z. Yu, Y. Q. Shi, and W. Su, "Symbol-rate estimation based on filter bank," in *2005 IEEE Int. Symposium. Circuits Sys.*, May 2005, pp. 1437–1440 Vol. 2.

[15] M. S. J.G. Proakis, *Digital Communications, 5th edition*. McGraw-Hill, New York, 2011.

[16] X. Jun, W. Fu-ping, and W. Zan-ji, "The improvement of symbol rate estimation by the wavelet transform," in *Proceedings. 2005 Int. Conf. Commun. Circuits Syst., 2005.*, vol. 1, May 2005, pp. 100–103 Vol. 1.

[17] P. Ciblat, P. Loubaton, E. Serpedin, and G. B. Giannakis, "Asymptotic analysis of blind cyclic correlation-based symbol-rate estimators," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1922–1934, July 2002.

[18] M. Flohberger, W. Kogler, W. Gappmair, and O. Koudelka, "Symbol rate estimation with inverse fourier transforms," in *2006 Int. Workshop on Satellite and Space Commun.*, Sep. 2006, pp. 110–113.

[19] M. El Hadi Lakhdari, H. Wang, and M. Bouchou, "Hybrid blind symbol rate estimation for linearly modulated signals," in *2017 IEEE 17th Int. Conf. on Commun. Technol. (ICCT)*, Oct 2017, pp. 287–292.

[20] H. Xu, Y. Zhou, and Z. Huang, "Blind roll-off factor and symbol rate estimation using IFFT and least squares estimator," in *2007 Int. Conf. Wireless Commun. Netw. Mobile Comput.*, Sep. 2007, pp. 1052–1055.

[21] N. M. Blachman and S. H. Mousavinezhad, "The spectrum of the square of a synchronous random pulse train," *IEEE Trans. Commun.*, vol. 38, no. 1, pp. 13–17, 1990.

[22] A. A. Kuchumov, V. I. Lipatkin, and E. M. Lobov, "Advanced algorithms for automatic symbol rate estimation of M-FSK and M-PSK signals," in *2017 Sys. Signal Synchronization, Generating and Proces. Telecommun. (SINKHROINFO)*, July 2017, pp. 1–4.

[23] S. Jing, J. Hall, Y. R. Zheng, and C. Xiao, "Signal detection for underwater IoT devices with long and sparse channels," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 6664–6675, 2020.

[24] Q. Yang, H. Wang, T. Zheng, Z. Han, and M. H. Lee, "Wireless powered asynchronous backscatter networks with sporadic short packets: Performance analysis and optimization," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 984–997, April 2018.

[25] P. Walk, P. Jung, B. Hassibi, and H. Jafarkhani, "MOCZ for blind short-packet communication: Practical aspects," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6675–6692, 2020.

[26] P. Walk, P. Jung, and B. Hassibi, "MOCZ for blind short-packet communication: Basic principles," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5080–5097, 2019.

[27] O. A. Adeladan and J. M. Shea, "Interference mitigation with partially coherent demodulation in a slow frequency-hopping spread-spectrum system," in *MILCOM 2009 - 2009 IEEE Military Commun. Conf.*, Oct 2009, pp. 1–7.

[28] C. B. Ha and H. K. Song, "Signal detection scheme based on adaptive ensemble deep learning model," *IEEE Access*, vol. PP, pp. 1–1, 2018.

[29] M. S. Chaudhari and S. Majhi, "Automated symbol rate estimation over frequency-selective fading channel by using deep neural network," in *2020 IEEE Int. Conf. Advanced Networks and Telecommun. Sys. (ANTS)*, 2020.

[30] Z. Hou and S. Jin, "A novel data-driven control approach for a class of discrete-time nonlinear systems," *IEEE Trans. Control Syst. Technol.*, vol. 19, no. 6, pp. 1549–1558, Nov 2011.

[31] L. Xie, O. Elnahas, Q. Zhao, and Z. Quan, "Data-driven RF transmit power calibration for wireless communication systems," *IEEE Wireless Commun. Lett.*, vol. PP, no. 99, pp. 1–1, 2020.

[32] Y. Ma, X. Wang, Z. Quan, and H. V. Poor, "Data-driven measurement of receiver sensitivity in wireless communication systems," *IEEE Trans. Commun.*, vol. PP, no. 99, pp. 1–1, 2019.

[33] X. Quan, X. Dou, Z. Wu, M. Hu, and A. Q. Huang, "Complex-coefficient complex-variable filter for grid synchronization based on linear quadratic regulation," *IEEE Trans. Ind. Informat.*, vol. 14, no. 5, pp. 1824–1834, 2018.

[34] J. Chen, Y. Ma, J. Cai, L. Zhou, Z. Bao, and W. Che, "Novel frequency-agile bandpass filter with wide tuning range and spurious suppression," *IEEE Trans. Ind. Electron.*, vol. 62, no. 10, pp. 6428–6435, Oct 2015.

[35] D. Xu, B. Jiang, and P. Shi, "A novel model-free adaptive control design for multivariable industrial processes," *IEEE Trans. Ind. Electron.*, vol. 61, no. 11, pp. 6391–6398, 2014.

[36] Z. Hou and S. Jin, "Data-driven model-free adaptive control for a class of MIMO nonlinear discrete-time systems," *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2173–2188, 2011.

[37] K. Kohno, Y. Inouye, and M. Kawamoto, “A matrix pseudo-inversion lemma for positive semidefinite hermitian matrices and its application to adaptive blind deconvolution of mimo systems,” *IEEE Trans. Circuits Sys. I, Reg. Papers*, vol. 55, no. 1, pp. 424–435, 2008.

[38] D. P. Mandic, A. I. Hanna, and M. Razaz, “A normalized gradient descent algorithm for nonlinear adaptive filters using a gradient adaptive step size,” *IEEE Signal Proces. Lett.*, vol. 8, no. 11, pp. 295–297, 2001.

[39] S. L. Goh and D. P. Mandic, “Stochastic gradient-adaptive complex-valued nonlinear neural adaptive filters with a gradient-adaptive step size,” *IEEE Trans. Neural Netw.*, vol. 18, no. 5, pp. 1511–1516, 2007.

[40] Z. Hou and S. Jin, “Data-driven model-free adaptive control for a class of MIMO nonlinear discrete-time systems,” *IEEE Trans. Neural Netw.*, vol. 22, no. 12, pp. 2173–2188, 2011.

[41] Donghui Li, “On asymptotic properties for the median point of cauchy mean-value theorem,” in *2011 Int. Conf. Multimedia Technol.*, 2011, pp. 5943–5945.