Research Article | Vol. 30, No. 22 / 24 Oct 2022 / Optics Express 39663

Optics EXPRESS

Synthesis of a GS-16QAM signal for a simplified optical ISB system and direct detection with the K-means clustering algorithm

JUN MING,[1] JIANGNAN XIAO,[1,*] DONGYAN WU,[1] LEILEI WANG,[1] YE ZHOU,[1] AND LI ZHAO[2]

[1] University of Shanghai for Science and Technology, Shanghai 200093, China
[2] School of Information Science and Technology, Fudan University, Shanghai 200433, China
[*] jiangnanxiao@usst.edu.cn

Abstract: An independent sideband (ISB) is a promising scheme for short-reach and metro applications because of its high spectral efficiency, low complexity, and tolerance to chromatic dispersion. Here, we develop a signal synthesis scheme to further reduce the complexity of ISB direct-detection (DD) systems. Two lower-order quadrature amplitude modulation (QAM) sideband signals are generated digitally, then the left sideband (LSB) and right sideband (RSB) are modulated with regular quadrature phase shift keying (QPSK) and geometrically shaped shifted QPSK (GS-S-QPSK), respectively. Then, the two independent sideband signals are received in a single photodiode to synthesize a GS-16QAM signal. The LSB and RSB signals can be separated and demodulated by a digital signal process (DSP) instead of using two optical bandpass filters. The proposed scheme significantly reduces the complexity of the ISB-DD receiver, thus saving system cost. Three different GS-S-QPSK signals are evaluated, with the square GS-S-QPSK achieving the best bit error rate (BER) with its optimal shaping factor. Considering signal–signal beating interference, a sub-blind K-means clustering algorithm is used to improve the BER performance, and the results indicate that it can achieve a large received optical power (ROP) gain at a threshold of 3.8×10^{-3}.

1. Introduction

The immense growth of Internet traffic has been facilitated by a variety of path-breaking advances in the field of optical communication systems, and direct detection (DD) has become very popular owing to its simplicity and low cost [1–3]. To facilitate the development of short- and intermediate-reach transmission systems with high spectral efficiency and fast capacity, numerous studies concerning DD with single-sideband (SSB) modulations have been proposed in recent years. In one study [4], a 117-Gb/s Kramers–Kronig (KK) PS-256-QAM signal with spectral efficiency of 5.85 bit/s/Hz was successfully transmitted using a single in-phase and quadrature (IQ) modulator. Wang et al. demonstrated a 140-Gb/s PS-256QAM SSB orthogonal frequency-division multiplexing signal over a 25-km standard single-mode fiber (SSMF) transmission through KK detection [5]. In another research [6], a four-subcarrier SSB PS-256QAM signal with constant water filling was transmitted over a 40-km SSMF. SSB shows good tolerance to the power fading induced by dispersion; furthermore, SSB exhibits higher spectral efficiency than double-sideband modulation.

However, SSB modulation uses only one sideband to carry data. Independent sideband (ISB) modulation utilizes the full sideband by delivering two independent data over the left sideband (LSB) and right sideband (RSB), respectively, which can double the data rate. ISB modulation was later adopted to build bandwidth-efficient fiber-optic links. This is usually called optical independent sideband (O-ISB) modulation. The O-ISB signal can be generated using an optical IQ modulator driven by analog LSB and RSB image reject signals [7,8]. Zhang et al. [9]

#473283 https://doi.org/10.1364/OE.473283
Journal © 2022 Received 16 Aug 2022; revised 30 Sep 2022; accepted 4 Oct 2022; published 11 Oct 2022

Optics EXPRESS

proposed a twin SSB discrete multi-tone (DMT) signaling scheme that loads two independent SSB-DMT datasets on the LSB and RSB simultaneously. Chien et al. implemented a wavelength division multiplexing transmission of 16-PM-QPSK O-ISB channels over a 2100-km SMF-28 [10]. Zhang et al. [11] experimentally demonstrated KK detection of an ISB signal with mixed modulation formats by Brillouin-amplifying a weak pilot tone at the receiver side. With such compensation algorithms as the KK algorithm and signal–signal beating interference (SSBI) cancellation, the linear distortion induced by chromatic dispersion can be effectively combined, making ISB-DD attractive; thus, practical solutions for short- and intermediate-haul systems are expected.

In this study, a signal synthesis scheme is successfully implemented by using ISB modulation to further reduce the complexity of the ISB-DD receiver end. Two independent software-defined 4QAM sideband signals are first generated by digital single-sided up-conversion and optical IQ modulation, after which the 16QAM signal is synthesized by blending the LSB signal with the RSB signal in a single photodiode (PD), unlike in other studies [10,11], where the ISB signals are separated into two sideband signals using optical bandpass filters (OBPFs) and are detected by two PDs. Our scheme has some similarities with non-orthogonal multiple access (NOMA) schemes due to the overlap of frequency resources by different signals. However, the overlap of two sidebands occurs after photo-detection at the receiver and not by adding signals at the transmitter side as in traditional NOMA schemes. In the proposed scheme, the synthesized signals can be separated and demodulated by digital signal processing (DSP), which eliminates two OBPFs and one PD, thus eliminating the residual sideband spectral issues (caused by imperfections of the OBPFs) and the system cost is subsequently reduced. Furthermore, the proposed scheme effectively relaxes the resolution requirement of the DACs in the transmitter side, because lower-order QAM sideband signals are first generated in an electrical domain and then blended to synthesize a higher-order QAM signal in the PD, rather than generating and transmitting the higher-order signal directly. In addition, a low-computational-complexity sub-blind K-means clustering algorithm is proposed to combat the amplitude and phase shift caused by SSBI. The simulation results indicate that the K-means algorithm significantly improves BER performance, especially at high data rates and in long-haul transmission.

The remainder of this article is organized as follows. The principle of the generation of the synthesized 16-QAM signal and its de-mapping method are described in Section 2. The simulation setup for the O-ISB system is described in Section 3. The simulation results are analyzed in Section IV. Section V concludes the article.

2. Principle

2.1. Principle of the 16-QAM synthesis

Figure 1 illustrates the principle of the proposed 16QAM signal generation. First, two independent data are modulated in QPSK and staggered QPSK (S-QPSK) formats. Then, the generated radio frequency (RF) signals are up-converted to intermediate frequency (IF) signals by mixing with two carriers that are in the complex sinusoidal form: $\exp(-2\pi f_s t)$ and $\exp(2\pi f_s t)$. The two independent IF vector signals are LSB and RSB signals. The two sidebands can differ in their modulation formats, symbol rates, and filter roll-off factors. The two IF signals are then added to form a drive signal of the Mach–Zehnder modulator (MZM), which can be represented by Eq. (1), and the electric spectra of the drive signal are shown in Fig. 1(I).

$$\begin{aligned} E_{LSB}(t) &= A(t)\exp(j(\varphi_l(t) - \omega_l t)) \\ E_{LSB}(t) &= B(t)\exp(j(\varphi_r(t) - \omega_r t)) \\ E(t) &= E_{LSB}(t) + E_{RSB}(t). \end{aligned} \tag{1}$$

Optics EXPRESS

where $E_{LSB}(t)$ and $E_{RSB}(t)$ represent LSB and RSB signals, respectively. Here, $A(t)$, $B(t)$;$\varphi_l(t)$, $\varphi_r(t)$; and $\omega_l(t)$, $\omega_r(t)$ denote the amplitudes, phases, and angular frequencies of the LSB and RSB signals, respectively. The drive signal $E(t)$ is separated digitally into a real part and an imaginary part, which drive the in-phase (I) and quadrature (Q) branches of the IQ modulator, respectively. The spectra of these two parts are illustrated in Figs. 1(II) and (III), respectively. The output continuous wave (CW) from the external cavity laser at a frequency of f_c can be represented as $E_{CW}\exp(jw_ct)$, and the output of the optical IQ modulator can be expressed as Eq. (2), where $\varphi_I(t)$ and $\varphi_Q(t)$ represent the phase differences of the I and Q branches of the IQ modulator, $E_{in}(t)$ and $E_{out}(t)$ denote the input and output of the IQ modulator, respectively, V_π denotes the half-wave voltage of the modulator, and Re and Im denote the real and imaginary components, respectively. In the proposed scheme, the IQ modulator consists of two parallel MZMs, which operate at the minimum transmission point. The direct current (DC) bias voltage of the PM is set to $V_\pi/2$ to ensure that the upper and lower branches of the IQ modulator have a fixed phase difference of $\pi/2$ to implement optical carrier suppression modulation. According to Eqs. (1) and (2), the output optical vector signal after the IQ modulator can be approximately defined as Eq. (3). We use the Jacobi-Anger expansion to rewrite (3) as an expansion of the Bessel series.

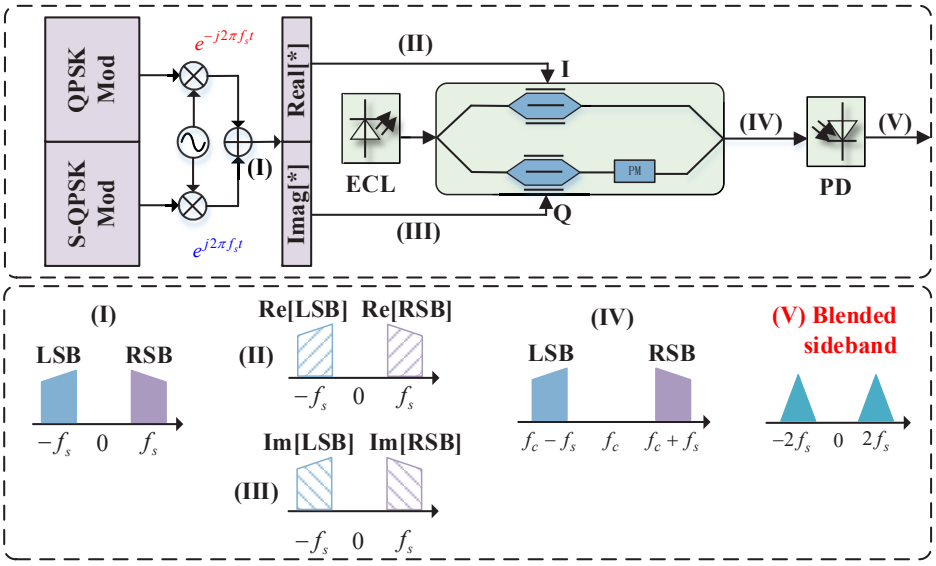

Fig. 1. Schematic of synthesized signal generation for ISB transmission. ECL: external cavity laser; PM: phase modulator; PD: photodiode; Re: real; Im: imaginary; (I) the electric spectra of the ISB signal; (II) the electric spectra of the real parts of the ISB signal; (III) the electric spectra of the imaginary parts of the ISB signal; (IV) the optical spectra of the signal after IQ Mod; (V) the electric spectra of the received synthesized signal after PD.

The first and second terms in Eq. (3) represent the LSB and RSB optical vector signals, respectively, where $J_{-1}(\cdot)$ and $J_1(\cdot)$ denote the first-order Bessel functions, m_l and m_r denote the modulation coefficients of the LSB and RSB signals, respectively, and m_l and m_r are related to

Optics EXPRESS

$A(t)$ and $B(t)$ in Eq. (1).

$$\begin{cases} E_{out}(t) = \dfrac{1}{2}E_{in}(t)\left(\cos(\dfrac{\varphi_I(t)}{2}) + j\sin\left(\dfrac{\varphi_Q(t)}{2}\right)\right) \\ \varphi_I(t) = \dfrac{Re[E(t)]}{V_\pi}\pi \\ \varphi_Q(t) = \dfrac{\mathrm{Im}[E(t)]}{V_\pi}\pi \end{cases} \tag{2}$$

$$\begin{cases} E_{IQ}(t) = E_{CW}[J_{-1}(m_l)\exp(j(\omega_c - \omega_l)t + j\varphi_l) \\ \qquad + J_1(m_r)\exp(j(\omega_c + \omega_r)t + j\varphi_r)] \\ m_l = \dfrac{\pi A(\mathrm{t})}{4V_\pi} \\ m_r = \dfrac{\pi B(\mathrm{t})}{4V_\pi} \end{cases} \tag{3}$$

Subsequently, the optical signal is detected by a single PD. According to the square law, the generated photocurrent can be expressed as Eq. (4). The first and second terms are DC components that can be blocked. The third term is the desired synthesized signal. The fourth term is located at the baseband, which can be neglected (or filtered). The fifth and sixth terms are the SSBI. They induce crosstalk and can be compensated for using equalization algorithms [12–16]. Figure 2 illustrates the ideal constellations of the $\pi/4$ phase-shifted QPSK signal in the LSB, S-QPSK signal in the RSB, and synthesized 16QAM signal.

$$i(t) = R\begin{Bmatrix} \dfrac{1}{2}A^2(t)J_{-1}^2(m_l) + \dfrac{1}{2}B^2(t)J_1^2(m_r) + \\ A(t)B(t)J_{-1}(m_l)J_1(m_r) \\ \cos[(\omega_l + \omega_r)t + (\varphi_l(t) + \varphi_r(t))] + \\ A(t)B(t)J_{-1}(m_l)J_1(m_r) \\ \cos[(\omega_l - \omega_r)t + (\varphi_l(t) - \varphi_r(t))] + \\ \dfrac{1}{2}A^2(t)J_{-1}^2(m_l)\cos[2(\omega_l t + \varphi_l(t))] + \\ \dfrac{1}{2}B^2(t)J_1^2(m_r)\cos[2(\omega_r t + \varphi_r(t))] \end{Bmatrix} \tag{4}$$

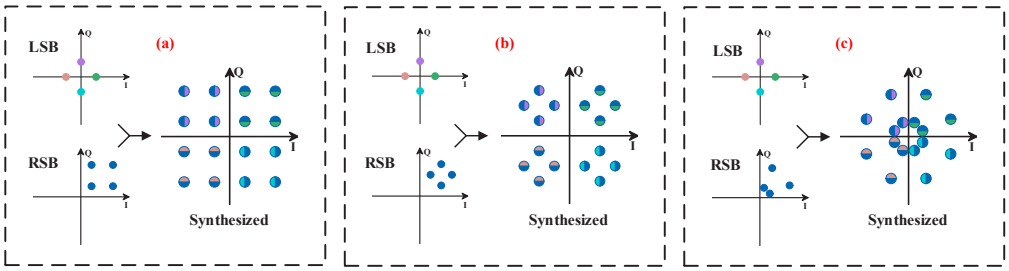

Fig. 2. Synthesized 16QAM constellation (a) square-16QAM (b) diamond-16QAM (c) circle-16QAM.

The $\pi/4$ phase shift of the QPSK signal in the LSB is not indispensable for the generation of the synthesized 16QAM signal; it makes the synthesized 16QAM constellation appear regular.

Optics EXPRESS

The relationship between the QPSK signal, S-QPSK signal, and corresponding synthesized 16QAM signal can be expressed as Eq. (5).

$$i(t) = \underbrace{RA(t)B(t)J_{-1}(m_l)J_1(m_r)}_{Amplitude} * \cos\{(\omega_l + \omega_r) + \underbrace{(\varphi_l(t) + \varphi_r(t))}_{Phase}\} \tag{5}$$

Here, $RA(t)B(t)J_{-1}(m_l)J_1(m_r)$ represents the amplitude of the synthesized signal, where R denotes the photoelectric conversion efficiency, $(\omega_l + \omega_r)$ represents the sum of the frequencies of the LSB and RSB signals, and $(\varphi_l(t) + \varphi_r(t))$ represents the sum of the phases. The Fourier components of Eq. (5) are given by $(\omega_l + \omega_r)$ and $-(\omega_l + \omega_r)$.

2.2. De-mapping method and GS design

As illustrated in Section 2.1, the amplitude of the 16QAM constellation is the product of the amplitudes of the QPSK constellation and S-QPSK constellation, while the phase of the 16QAM constellation is the sum of the phases of the QPSK constellation and S-QPSK constellation. In short, the amplitude and phase of the 16QAM constellation are a combination of the amplitudes and phases of the QPSK and S-QPSK constellations. Herein, the set of denial symbols $\{0, 1, 2, 3\}$ is used to represent the constellation points of the QPSK and S-QPSK signals, and $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}$ is used for the synthesized 16QAM signal.

The relationship between the symbol labeling of the QPSK constellation, square-S-QPSK constellation, and square-16QAM is shown in Fig. 3, which shows that each constellation point of the square-16QAM can be exclusively demodulated into one point of the QPSK signal and one point of the S-QPSK signal. For example, the symbol "0" of square-16QAM (Gray coded) can be exclusively demodulated into "0" of QPSK and "2" of S-QPSK. All combinations of symbols of square-16QAM are shown in Table 1. Furthermore, Fig. 3 shows that the symbol of QPSK can be determined by distinguishing the quadrant to which it belongs, and the constellation points of square-16QAM can be considered as the rotation version (around the origin point) of S-QPSK constellation points, with rotated phases of 0, π/2, π, and 3π/2.

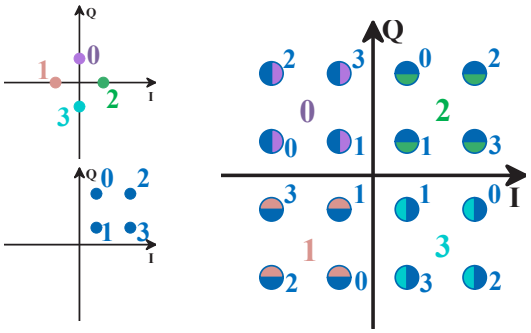

Fig. 3. Symbol labeling of the synthesized square-16QAM.

Similarly, Fig. 4 depicts the Gray-coded symbol labeling of the diamond-S-QPSK and circle-S-QPSK. The symbol labeling of the corresponding diamond-16QAM and circle-16QAM is similar to that for square-16QAM.

Next, three types of S-QPSK constellation design are presented to evaluate the performances of different geometric shaping (GS) schemes for the S-QPSK signal. The GS technique uses a non-uniformly spaced constellation with equally probable symbols, which has been proven to

Optics EXPRESS

Table 1. Symbol combinations for QPSK, S-QPSK, and square-16QAM

Square-16QAM	QPSK	S-QPSK
0	0	2
1	0	0
2	1	2
3	1	3
4	0	3
5	0	1
6	1	0
7	1	1
8	2	2
9	2	3
10	3	2
11	3	0
12	2	0
13	2	1
14	3	3
15	3	1

Fig. 4. Symbol labeling of the diamond-S-QPSK (left) and circle-S-QPSK (right).

offer a signal-to-noise ratio (SNR) gain for an additive white Gaussian noise (AWGN) channel; however, the synthesis of GS constellations has received little attention. In this study, the minimum Euclidean distance (MED) was considered the major metric for different S-QPSK constellations. Initially, the MEDs of different S-QPSK constellations are evaluated (Table 2). For a fair comparison, the powers of all S-QPSK constellations are normalized. As shown in Table 2, the square-S-QPSK with the largest MED is considered to be the most promising GS scheme for the S-QPSK signal in the proposed system.

Table 2. MEDs of differently shaped S-QPSK constellations

S-QPSK	MED
Square	0.45
Diamond	0.42
Circle	0.3

To evaluate the performances of the three types of GS-S-QPSK constellations further, a mathematical model is established to quantify the GS design.

$$X_{16QAM} = X_{QPSK} \otimes \varepsilon X_{S-QPSK} \tag{6}$$

where X_{QPSK}, X_{S-QPSK}, and X_{16QAM} represent the constellations of QPSK, S-QPSK, and synthesized 16QAM signals, respectively, $\otimes$ is the procedure of blending (see Section 2.1), and denotes

Optics EXPRESS

the shaping factor of the GS-S-QPSK constellation (the mechanism of ε is shown in Fig. 5). The "2" (see the mapping in Figs. 3 and 4) of square-S-QPSK is fixed, and the rest is closer to "2" when ε increases. The "0, 1, 2, and 3" of diamond-S-QPSK gather closer to their centroid when ε is larger. In addition, the "0, 2" of circle-S- QPSK is fixed, and the radius of "1, 3" is larger when ε increases. For a fair comparison, the powers of all the S-QPSK signals are normalized, and the QPSK signal is fixed.

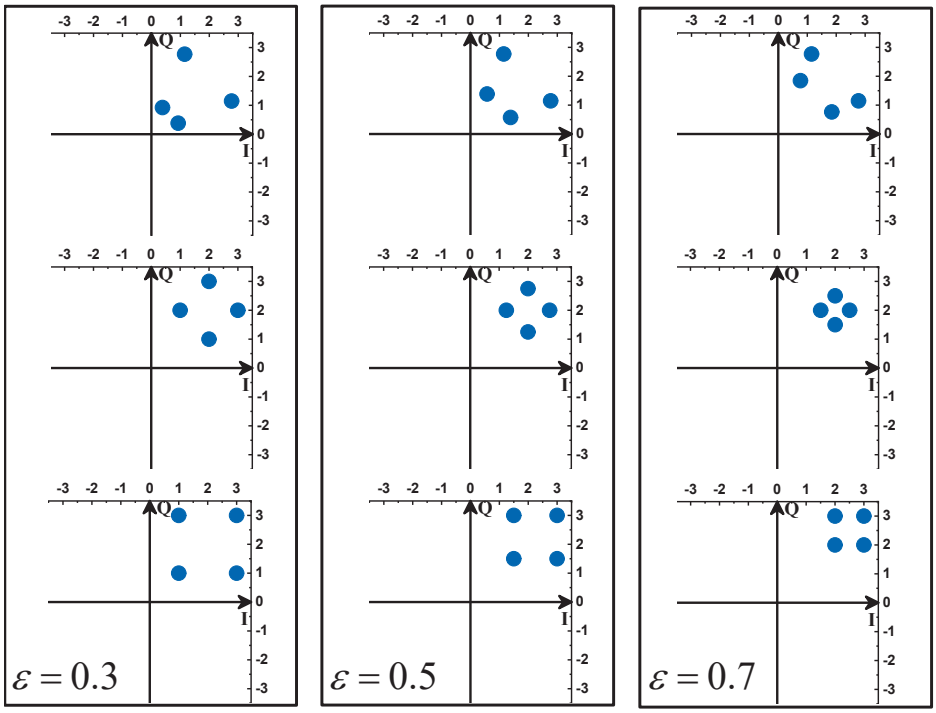

Fig. 5. Constellations of the different types of GS-S-QPSK.

The BERs of the different GS-16QAM constellations with optimal shaping coefficients simulated in an AWGN channel are presented in Fig. 6, which shows that the square synthesized 16QAM signal achieves the best BER performance, as indicated by the MED result mentioned previously. It is believed that a larger MED and thorough Gray coding ensures the best BER performance of the square-16QAM.

As Fig. 7 shows, the BER performance of the synthesized 16QAM constellation is evaluated with different shaping factors ε in the AWGN channel, where the SNRs are set at 8, 10, and 12 dB. The estimated 16QAM constellations are shown when the shaping factor ε is set at 0.2, 0.35, and 0.55, and the SNR is set to 12 dB. Figure 7 reveals that the BERs first decrease and then increase with the increase in ε, which is consistent with the theoretical analysis of MED. The best BER performance is achieved when ε is equal to 0.35 in all SNR cases.

2.3. *Decision with the K-means clustering algorithm*

Recently, the K-means clustering algorithm has become popular in the field of optical communications [17–20]. It has several applications that are not just confined to nonlinearity equalization, modulation format identification, phase recovery, and data detection. In this study, the K-means clustering algorithm is used to address the generalized distortion of the constellation by using the centroids of the converged clusters as the standard decision constellation points instead of

Optics EXPRESS

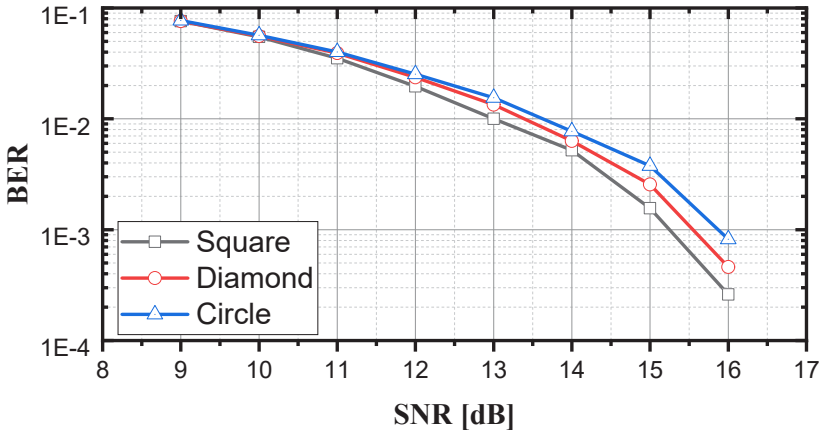

Fig. 6. BER versus SNR for different GS-shaped synthesized 16QAM constellations.

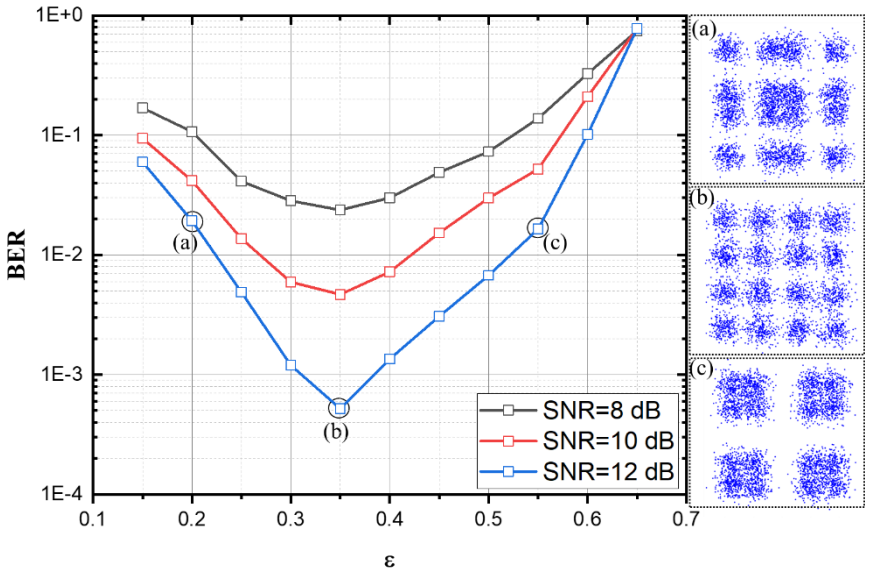

Fig. 7. BER versus different shaping factors: (a) $\varepsilon = 0.2$, SNR = 12 dB, (b) $\varepsilon = 0.35$, SNR = 12 dB, and (c) $\varepsilon = 0.55$, SNR = 12 dB.

Optics EXPRESS

the theoretically calculated points. The distortion considered is mainly induced by the SSBI illustrated by the fifth and sixth terms of Eq. (4).

Generally, the conventional unsupervised K-means algorithm must determine the number of clusters before executing the iterative procedure. Then, a set of initial centroids is used to calculate the Euclidean distances between each constellation point and all centroids, which requires an enormous amount of computational resources, especially for high-order modulation formats. In the next iteration, the centroids of each cluster are updated until the clusters converge. The initial centroids must be partitioned well to approximate clusters in the sparse constellation space. The traditional K-means algorithm usually uses randomly picked initial centroids, which results in few chances to find appropriate initial points. In addition, the k-means algorithm normally uses a non-convex squared error evaluation function as the clustering criterion function, which results in a deviation from the searching range of the global optimal solution. Thus, it is critical to locate the centroids accurately and quickly.

A sub-blind K-means clustering algorithm is used to reduce calculation complexity and data redundancy. The tracking of centroids of each cluster is more accurate compared with the traditional K-means algorithm, and the optimal centroids are updated by a new method rather than iteration. The procedures of the sub-blind K-means clustering algorithm include the following steps.

(a) The value of the clusters is set at 16, and the theoretical constellation points of the received 16QAM signal are used as the initial centroids.

(b) The Euclidean distance of the received constellation points to every four centroids that belong to its quadrant is calculated. Then, each constellation point is assigned to the corresponding cluster according to the nearest Euclidean distance.

(c) The centroids are updated by calculating the mean of the 16 clusters obtained.

(d) Based on the centroids obtained in step (c), the Euclidean distance of the received constellation points to each centroid is recalculated. Then, each constellation point is classified by the nearest Euclidean distance.

(e) Finally, 16 optimal centroids are obtained, and the received constellation points are assigned to the 16 different clusters according to the nearest Euclidean distance.

The sub-blind k-means clustering algorithm has the following advantages. (1) The centroids can be quickly located because the theoretical constellation points are close to the optimal centroids. (2) The computational complexity is significantly reduced because the calculation iterations are reduced and the points are pre-classified by quadrants.

3. Simulation setup

The simulation setup of the proposed scheme is shown in Fig. 8. At the transmitter end, two independent bit sequences were generated by a random number generator, and one sequence was mapped into the LSB by regular QPSK modulation with an overall phase shift of $\pi/4$.The other sequence was mapped into the RSB by square-S-QPSK, diamond-S-QPSK, or circle-S-QPSK modulation. The two baseband signals were filtered with a 0.2 roll-off root-raised-cosine (RRC) filter, and then 2× oversampling was applied. Next, the baseband signals were up-converted to the desired LSB and RSB bands at −16 GHz and +16 GHz, respectively. Then, the LSB and RSB signals were added to form a drive signal; the drive signal was divided into real and imaginary components, which then drive the I and Q branches of the IQ modulator, respectively. A CW light wave at 1548.706 nm was generated from an external cavity laser with a linewidth of less than 100 kHz, with 16-dBm light power. The DC bias voltages of the two MZMs were set to

Optics EXPRESS

enable the MZM to work at the minimum output optical power, and the DC bias voltage of the PM was set such that the PM obtains a $\pi/2$ phase shift.

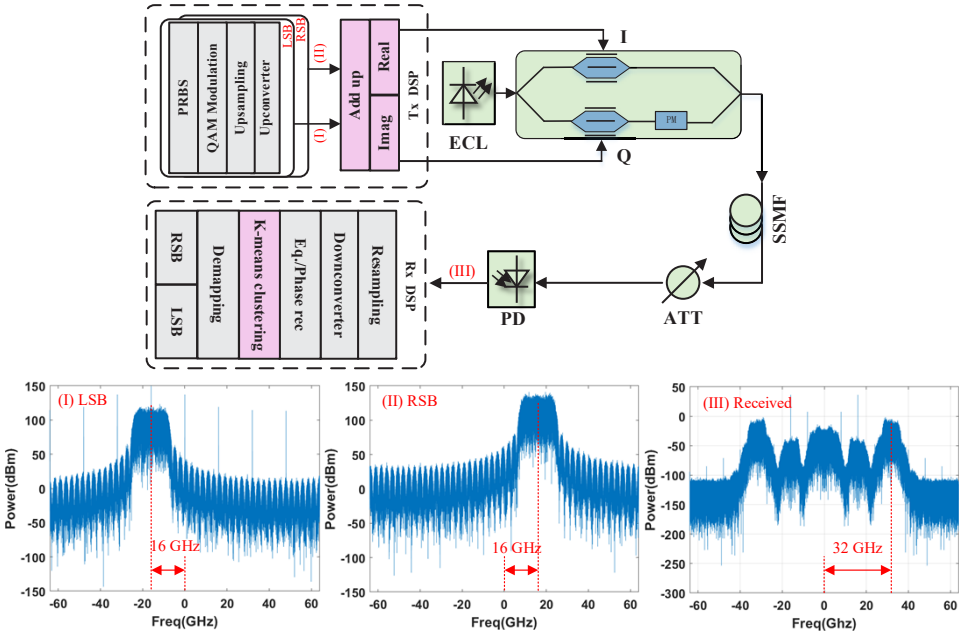

Fig. 8. Simulation system diagram. ECL: external cavity laser; SSMF: standard single-mode fiber; ATT: attenuator; PD: photodiode; Eq.: equalization; Phase rec: phase recovery. (I) the electric spectra of the LSB signal; (II) the electric spectra of the RSB signal; (III) the electric spectra of the received signal after PD.

After transmission through an SSMF, a variable optical attenuator was deployed to control the ROP into the PD, and then the optical signals were captured by direct detection with a single PD. Next, the captured signals were equalized and recovered by DSP algorithms, including resampling, downconverter, and channel equalization using a cascaded multi-modulus algorithm [21]. A blind phase search [22] carrier phase recovery algorithm was used to compensate for the phase shift, and a K-means clustering algorithm was deployed to obtain decision constellation points. Then, the received synthesized 16QAM signal was de-mapped with the decision constellation supplied by the K-means clustering algorithm. As mentioned in Section 2.1, the QPSK and S-QPSK signals were separated from the 16QAM signal, and the BERs of the overall 16QAM, QPSK, and S-QPSK signals were then calculated.

4. Simulation results and discussion

Initially, the performance of the proposed system is evaluated with different shaping factors. The square-S-QPSK signal and corresponding synthesized 16QAM signal are used in this assessment, and the data rates of the LSB and RSB signals are both set to 16Gbaud. The signal is transmitted through a 2-km SSMF. In Fig. 9, the simulation results of the measured BERs versus different ε values of the square-S-QPSK are presented, and the received 16QAM constellations (after the DSP procedure) are shown when the shaping factor ε is set to 0.27, 0.40, and 0.49, and the ROP is -2 dBm. The BERs confirm the theoretical analysis, where, for each curve, the BER initially decreases and then increases. The best BER performance is achieved when the shaping factor ε is 0.4, regardless of the ROP case. Furthermore, the optimal ε in Fig. 9 is different from that in Fig. 7. This is mainly because of the difference in the transmission channel.

Optics EXPRESS

Moreover, considering the 7% pre-FEC BER threshold of 3.8×10^{-3}, the range of the value of ε is approximately 0.22 to 0.53.

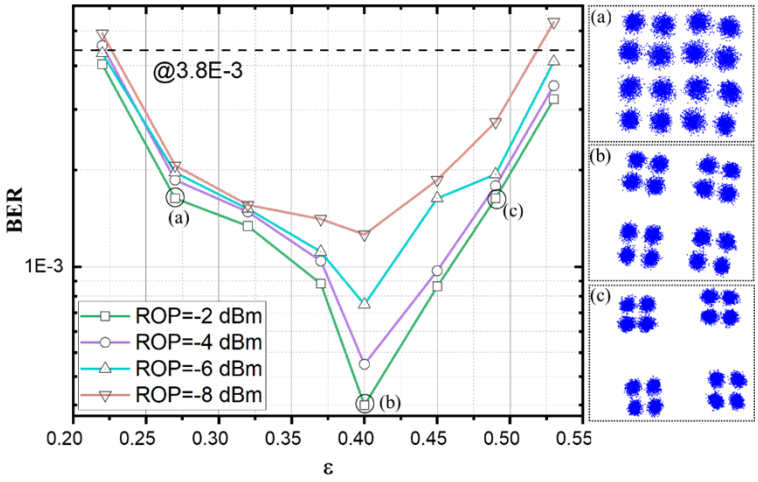

Fig. 9. BER versus different values of ε: (a), (b), and (c) are the constellations when the ROP is −2 dBm and ε is 0.27, 0.40 and 0.49, respectively.

In Fig. 10, the BER performance of the square-16QAM signal and the corresponding QPSK and the square-S-QPSK signal is shown. The results reveal that the BER of the QPSK signal decreases to zero when ε is 0.27, yet the BERs of square-16QAM and square-S-QPSK are quite similar to each other. The constellations in Fig. 9 distinctly explain this. The constellation points move away from the boundary of the quadrant when ε increases; therefore, the BER of the QPSK signal decreases quickly as ε increases.

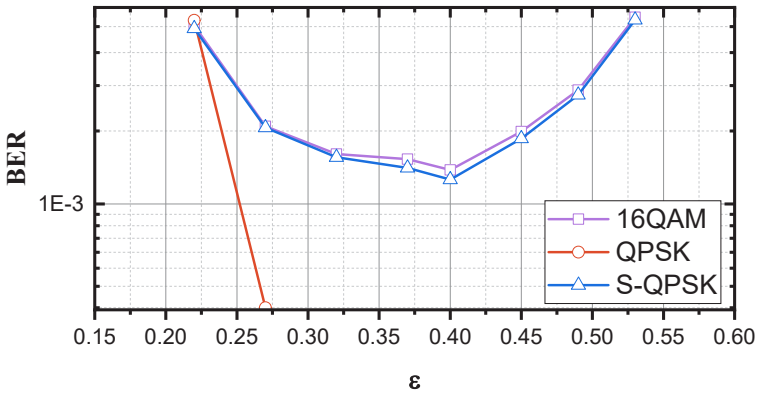

Fig. 10. BER versus ε for square-16QAM, QPSK, and square-S-QPSK signal when ROP is −8 dBm.

The detailed constellations in Fig. 9 show that the received signal still suffers from phase distortion and IQ imbalance, even after processing by the compensation algorithm of the DSP procedure. Therefore, the performance of the sub-blind K-means algorithm mentioned in Section 2.3 is evaluated. The square-S-QPSK scheme is used, and the value of ε is set at 0.4.

Figure 11 shows that the BER of each curve decreases with an increase in ROP. For the black curve, the BER decreases quickly with an increase in the ROP. However, for the red curve (without the K-means algorithm), the BER decreases more smoothly with increasing ROP. For

Optics EXPRESS

the green and blue curves, the BERs decrease more slowly compared with the black and red curves. The results of these four curves indicate that using the centroids of the converged clusters as the standard decision constellation points yields a much better BER performance when the ROP is increased, and a larger BER gain is achieved when the data rate is higher.

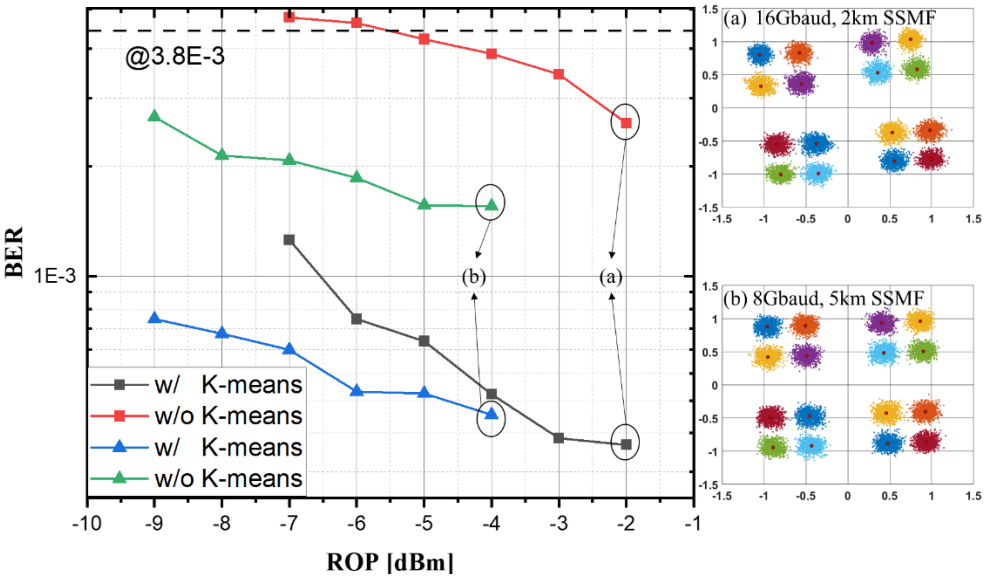

Fig. 11. BER versus ROP with K-means and without K-means: black and red curves are for a 16-Gbaud data rate through 2-km SSMF; the green and blue curves are for an 8-Gbaud data rate through 5-km SSMF.

Moreover, one can understand the excellent performance of the K-means algorithm intuitively from Figs. 11(a) and 11(b), where the large red points represent the centroids of each cluster, and each constellation point can be accurately distinguished by the clusters. The grid can be seen as the decision boundary for the conventional hard decision, and some clusters are divided by the grid, indicating that many constellation points are misjudged and finally cause a BER performance loss. The proposed sub-blind K-means clustering algorithm significantly improves the BER performance, particularly when the signal suffers from a severe IQ imbalance.

The BER performances of the three different GS-shaped S-QPSK schemes are compared, and the results are shown in Fig. 12. In this evaluation, the data rate is set to 16Gbaud, and the signal is transmitted through a 2-km SSMF. Because the optimal shaping coefficients for the three schemes are different, the transmission power is normalized for a fair comparison. The results in Fig. 12 show that the square-S-QPSK has the best BER performance, which is similar to the results in Fig. 6, where both the red and black curves are below the 7% pre-FEC BER threshold of 3.8×10^{-3} in all ROP cases. However, the blue curve is far from the other curves, and the BER is less than 3.8×10^{-3} only when the ROP is −2 dBm. The constellation points in Fig. 10(a) are more scattered than those in Figs. 11(b) and 11(c), indicating a BER performance loss. Furthermore, as mentioned in Section 2.2, the synthesized 16QAM signal can be demodulated into a QPSK signal and an S-QPSK signal, and the QPSK symbol can be determined by distinguishing the quadrant of the synthesized 16QAM constellation. For the square-16QAM and diamond-16QAM signals, the quadrant decision can be easily made with zero error. However, for the circle-16QAM signal, there are numerous constellation points located at the boundary of the quadrant, which reduces the BER performance.

Optics EXPRESS

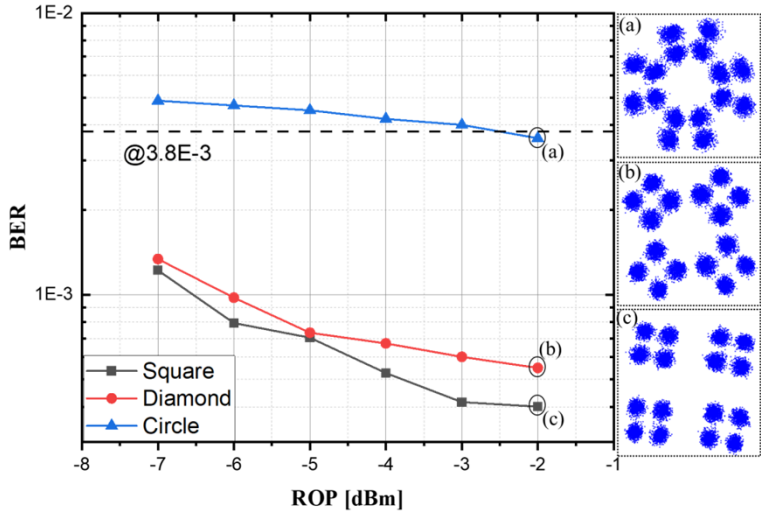

Fig. 12. BER versus ROP for different GS-shaped synthesized 16QAM constellations.

Next, we evaluate the BER and error vector magnitude (EVM) performance of conventional DSB-16QAM, SSB-16QAM, ISB-16QAM, and single-carrier 16QAM (SC-16QAM) methods. For a fair comparison, the data rate, and transmission distance of each method are set to 2Gbaud, and 2 km SSMF, respectively. Figure 13 (I) gives the BER characteristics of the evaluated schemes. In the DSB case, the BER performance is dramatically degraded due to the weak dispersion tolerance compared to the SSB and ISB methods. Compared to the SSB case, a large BER improvement is observed in the ISB case, which can be attributed to the lower-order modulation format, since the lower-order modulation format can achieve a larger Euclidean distance at the same Tx power. The SC-16QAM is transmitted in a traditional coherent system, it achieves a much better BER performance than the proposed ISB scheme, due to its complex coherent receiver and an extra ECL. Compared to the SC-16QAM case, the ISB scheme makes a tradeoff between the performance and cost. Figure 13 (II) gives the EVM characteristics of the evaluated schemes, which we can conclude the similar performance comparison.

In our proposed scheme, the bias voltage of the IQ modulator is set to implement optical suppressed-carrier modulation, its performance is related to the carrier to sideband ratio (CSR) [23], and we thus give the results of achievable BER for an ROP of -10dBm versus CSR for ISB and SSB signals in Fig. 14 (a). The results show that the optimal CSR for DSB scheme, SSB scheme and ISB scheme are 4 dB, 2 dB and 6 dB, respectively. Compared to the SSB signals with 2 dB CSR, the BER performance of our ISB signals with 6 dB CSR is much better. Even with a non-optimal CSR (2 dB), our scheme still outperforms the SSB scheme with optimal CSR. And DSB modulation performance is the worst due to chromatic dispersion. Figure 14 (b) gives the measured BER versus optical signal-to-noise ratio (OSNR) for ISB-32QAM and ISB-64QAM signals, note that the ISB-32QAM and ISB-64QAM signals can be synthesized by changing the S-QPSK signal (see Section 2.2) into S-8QAM and S-16QAM signal, respectively. We can observe that all tested cases can reach the 2.4E-2 limit when OSNR is larger than 12 dB, the results prove that higher order modulation formats are practical for our scheme.

Optics EXPRESS

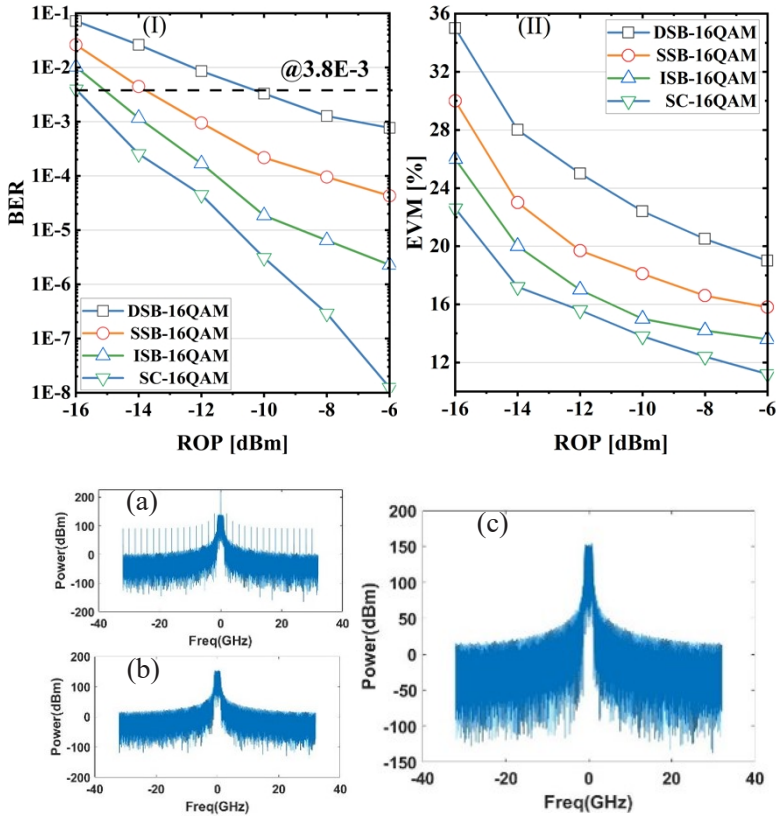

Fig. 13. (I) BER versus ROP for DSB, SSB, ISB, and SC-16QAM signal transmissions; (II) EVM versus ROP for DSB, SSB, ISB, and SC-16QAM signal transmissions; (a) baseband signal for LSB of ISB-16QAM; (b) baseband signal for RSB of ISB-16QAM; (c) baseband signal for DSB, SSB, and SC-16QAM.

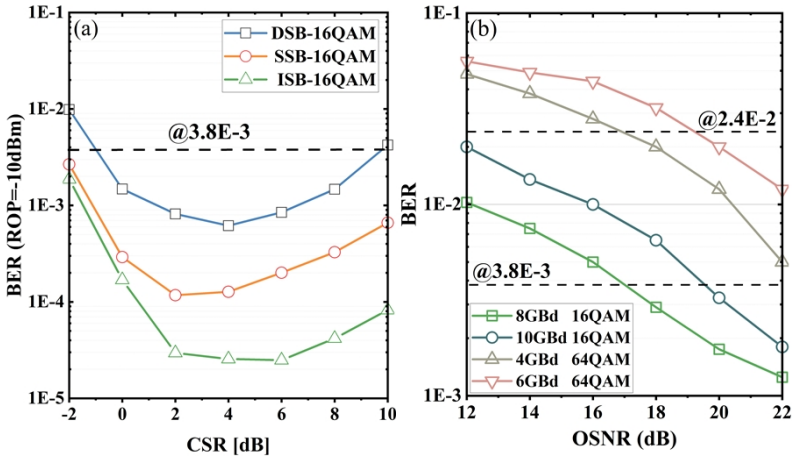

Fig. 14. (a) BER versus CSR for DSB, SSB and ISB-16QAM signals; (b) BER versus OSNR for ISB-32QAM and ISB-64QAM signals.

Optics EXPRESS

5. Conclusion

A synthesized signaling scheme is proposed to further reduce the complexity and cost of the ISB-DD system. In the proposed scheme, a 16QAM signal is synthesized by the beating of a QPSK signal in the LSB with an S-QPSK signal in the RSB. The LSB and RSB signals can be separated and demodulated from the synthesized 16QAM signal by a DSP approach, instead of being divided by OBPFs, as in the conventional SSB/DSB DD system. The DAC resolution requirement is significantly reduced by the lower order of the transmitted signal, which reduces the system cost. Three different geometrically shaped S-QPSK schemes are evaluated. It is found that square-S-QPSK achieves the best BER performance, which benefits from the highest MED. Furthermore, a sub-blind K-means clustering algorithm with reduced computational complexity is used to combat the phase distortion and IQ imbalance of the constellation, the results prove that the sub-blind K-means algorithm significantly improves the BER performance. The proposed system is available for higher-order modulation formats, and the probabilistic constellation shaping technique can be adopted in future research. We believe this research can provide a complementary choice for systems requiring a simpler structure and lower cost.

Funding. National Key Research and Development Program of China (2018YFB1801503); National Natural Science Foundation of China (61675048); The Research Project of Key Laboratory for Information Science of ElectromagneticWaves (MoE) (EMW201911).

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this article are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. X. Li, Y. Xu, and J. Yu, "Single-sideband W-band photonic vector millimeter-wave signal generation by one single I/Q modulator," Opt. Lett. **41**(18), 4162–4165 (2016).
2. X. Li, J. Xiao, and J. Yu, "Long-distance wireless mm-wave signal delivery at W-band," J. Lightwave Technol. **34**(2), 661–668 (2016).
3. Y. Wang, J. Yu, N. Chi, and G.-K. Chang, "Experimental demonstration of 120-Gb/s Nyquist PAM8-SCFDE for short-reach optical communication," IEEE Photonics J. **7**(4), 1–5 (2015).
4. Y. Wei, Y. Zhou, C. Liu, K. Wang, J. Zhang, F. Wang, J. Ding, and J. Yu, "SSB single carrier and multicarrier in C-band FSO transmission with KK receiver," J. Lightwave Technol. **38**(18), 5000–5007 (2020).
5. K. Wang, Y. Wei, M. Zhao, W. Zhou, and J. Yu, "140-Gb/s PS-256-QAM transmission in an OFDM system using Kramers–Kronig detection," IEEE Photonics Technol. Lett. **31**(17), 1405–1408 (2019).
6. X. Li, Z. Xing, M. S. Alam, M. E. Mousa-Pasandi, M. O'Sullivan, and D. V. Plant, "Demonstration of C-Band Amplifier-Free 100 Gb/s/λ Direct-Detection Links Beyond 40-km SMF Using a High-Power SSB Transmitter," J. Lightwave Technol. **38**, 1 (2020).
7. A. Narasimha, X. Meng, M. Wu, and E. Yablonovitch, "Tansingle-sidebandband modulation scheme for doubling spectral efficiencyanalogogue fibre links," Electron. Lett. **36**(13), 1135–1136 (2000).
8. P. Saghari, S. M. Nezam, A. Sahin, and A. Willner, "Polarization-state-rotation and filtering receiver for bandwidth-efficient tandem single sideband systems," in *Optical Fiber Communication Conference*, (Optical Society of America, 2004), FN5.
9. L. Zhang, T. Zuo, Q. Zhang, E. Zhou, G. N. Liu, and X. Xu, "Transmission of 112-Gb/s+ DMT over 80-km SMF Enabled by Twin-SSB Technique at 1550 nm," in *2015 European Conference on Optical Communication (ECOC)*, (IEEE, 2015), 1–3.
10. H. C. Chien, Z. Jia, J. Zhang, Z. Dong, and J. Yu, "Optical independent-sideband modulation for bandwidth-economic coherent transmission," Opt. Express **22**(8), 9465–9470 (2014).
11. Q. Zhang and C. Shu, "Carrier regeneration assisted Kramers-Kronig detection of an independent sideband signal," in *CLEO: Science and Innovations*, (Optical Society of America, 2018), STu3C. 2.
12. A. Mecozzi, C. Antonelli, and M. Shtaif, "Kramers–Kronig coherent receiver," Optica **3**(11), 1220–1227 (2016).
13. X. Zhang, J. Li, and Z. Li, "SSBI Cancellation Method for IMDD-OFDM System with a Single Photodiode," in *PIERS Proceedings*, 2014.
14. Z. Li, M. S. Erkılınç, K. Shi, E. Sillekens, L. Galdino, B. C. Thomsen, P. Bayvel, and R. I. Killey, "SSBI mitigation and the Kramers–Kronig scheme in single-sideband direct-detection transmission with receiver-based electronic dispersion compensation," J. Lightwave Technol. **35**(10), 1887–1893 (2017).
15. C. Ju, N. Liu, X. Chen, and Z. Zhang, "SSBI mitigation in A-RF-tone-based VSSB-OFDM system with a frequency-domain Volterra series equalizer," J. Lightwave Technol. **33**(23), 4997–5006 (2015).

Optics EXPRESS

16. X. Wang, J. Yu, Z. Cao, J. Xiao, and L. Chen, "SSBI mitigation at 60 GHz OFDM-ROF system based on optimization of training sequence," Opt. Express **19**(9), 8839–8846 (2011).
17. J. Zhang, W. Chen, M. Gao, and G. Shen, "K-means-clustering-based fiber nonlinearity equalization techniques for 64-QAM coherent optical communication system," Opt. Express **25**(22), 27570–27580 (2017).
18. X. Wu and N. Chi, "The phase estimation of geometric shaping 8-QAM modulations based on K-means clustering in underwater visible light communication," Opt. Commun. **444**, 147–153 (2019).
19. Q. Zhang, Y. Yang, C. Guo, X. Zhou, Y. Yao, A. P. T. Lau, and C. Lu, "Accurate BER estimation scheme based on K-means clustering assisted Gaussian approach for arbitrary modulation format," J. Lightwave Technol. **38**(8), 2152–2157 (2020).
20. C. Zhendong, J. Weining, X. Changbo, and L. Min, "Modulation recognition based on constellation diagram for M-QAM signals," in *2013 IEEE 11th International Conference on Electronic Measurement & Instruments*, (IEEE, 2013), 70–74.
21. W. Zhou, L. Zhao, J. Zhang, K. Wang, J. Yu, Y.-W. Chen, S. Shen, R.-K. Shiu, and G.-K. Chang, "135-GHz D-band 60-Gbps PAM-8 wireless transmission employing a joint DNN equalizer with BP and CMMA," J. Lightwave Technol. **38**(14), 3592–3601 (2020).
22. D. A. Mello, F. A. Barbosa, and J. D. Reis, "Interplay of probabilistic shaping and the blind phase search algorithm," J. Lightwave Technol. **36**(22), 5096–5105 (2018).
23. J. Ye, A. Li, X. Chen, L. Yan, and W. Shieh, "Performance improvement in radio-over-fiber transmission using pre-distorted double-sideband signals," in *Optical Fiber Communication Conference*, (Optica Publishing Group, 2013), OW1D. 2.