IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 7, 2019, accepted March 27, 2019, date of publication April 9, 2019, date of current version April 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2909707

Ternary Trellis Coded Modulation

MAHMOUD ABDELAZIZ[1] **AND T. AARON GULLIVER**[2]**, (Senior Member, IEEE)**
[1]Department of Electrical Engineering, Military Technical College, Cairo, Egypt
[2]Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 2Y2, Canada

Corresponding author: T. Aaron Gulliver (agullive@ece.uvic.ca)

ABSTRACT Ternary trellis coded modulation (TTCM) is introduced. This combines triangular signal constellations with ternary convolutional codes. The performance of TTCM is presented and compared with binary trellis coded modulation (BTCM) which employs square quadrature amplitude modulation (SQAM) and binary convolutional coding. Ternary set partitioning (TSP) is used for TTCM, and binary-to-ternary conversion is determined that is suitable for TTCM. A construction for M-ary triangular QAM (TQAM) which is compatible with TSP is presented. The performance of BTCM is compared with that of TTCM.

INDEX TERMS Constellation diagram, modulation coding, convolutional codes.

I. INTRODUCTION

Ungerboeck proposed a technique called trellis coded modulation (TCM) which can efficiently provide high data rates over bandwidth limited channels [1]. TCM has been employed in optical communications [2] and cooperative communications [3], and is being considered for future 5G wireless systems [4]. To provide the redundancy necessary for coding, the signal constellation is expanded rather than increasing the bandwidth. TCM combines a trellis code with modulation so at the receiver decoding and demodulation are combined [1]. The performance of TCM depends on the minimum distance between constellation symbol sequences, which is called the free Euclidean distance. For the same transmit power, the free Euclidean distance with TCM can be larger than the minimum Euclidean distance between symbols of the uncoded (smaller) constellation, which results in a lower bit error rate (BER). Therefore, the goal is to maximize the free Euclidean distance.

The signal constellation plays an important role in the power efficiency and BER performance of a communication system. The power efficiency depends on the distances between points in the signal constellation and the origin. Square quadrature amplitude modulation (SQAM) was proposed in [5] and is commonly employed because of its simple structure [7]. Subsequently, several signal constellations such as triangular, pentagonal and hexagonal [6] were examined which have better power efficiency than SQAM. However, these constellations have higher detection complexity. The highest power efficiency constellations without a point at the origin were presented in [8]. However, the irregularity of these constellations makes detection very complex [8]. In [9], triangular QAM (TQAM) was introduced which has signal points from a triangular lattice. TQAM has a power efficiency which is close to optimum and higher than SQAM, but SQAM has a lower detection complexity. TQAM provides a good tradeoff between BER and detection complexity [9], and so is considered here for TCM.

Ungerboeck introduced a mapping for TCM that maximizes the minimum distance between symbol sequences [13]. This mapping labels the constellation symbols using set partitioning. Set partitioning initially divides the constellation symbols into two subsets. In subsequent levels, each subset is divided into two new subsets. The last level is reached when each subset has two symbols. This is called binary set partitioning (BSP) and is based on two rules [14]. First, all subsets have an equal number of constellation points. Second, coded symbols originating from the same state or merging into the same state in the trellis are assigned in the same subsets. With SQAM or phase shift keying (PSK), BSP increases the minimum Euclidean distance within the subsets in each partitioning level [15]. However, this property of BSP does not hold with TQAM because of the triangular structure of TQAM which results in symbols having more nearest neighbors on average than with SQAM. To solve this problem, ternary set partitioning (TSP) is introduced here. This solution results in an increase in the Euclidean distance within the subsets each partitioning level. Thus, in this paper TSP is employed with TQAM constellations.

TCM with TQAM uses TSP for the mapping, so the constellation symbols are labeled with trits, i.e. {0, 1, 2}. Therefore, ternary arithmetic and ternary logic must be employed.

The associate editor coordinating the review of this manuscript and approving it for publication was Cunhua Pan.

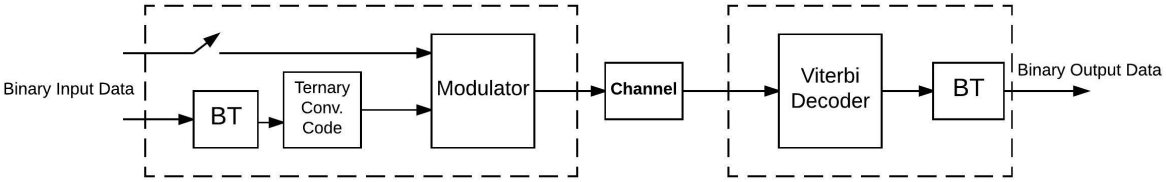

FIGURE 1. Block diagram of a ternary trellis coded modulation (TTCM) system.

Ternary systems have been shown to be faster than binary systems [16], [17]. Further, ternary systems can be implemented in hardware more efficiently than for any other integer base [17], and can provide higher data storage densities [17]. Binary TCM (BTCM) employs a binary convolutional code [1], but TCM with TQAM operates on a ternary alphabet so a ternary convolutional code should be employed. Ternary convolutional codes (TCCs) have been shown to provide better BER performance than binary convolutional codes (BCCs) for the same blocklength and complexity [18]. Here, TCM with a TCC is called ternary trellis coded modulation (TTCM).

Although ternary systems have several advantages over binary systems, commercial communication systems operate with binary data. Therefore, the performance of TTCM must be evaluated considering binary inputs and binary outputs. This requires that binary data be converted to ternary data (trits) before encoding. Then after TTCM decoding, the output trits must be converted to bits. Binary to ternary (BT) conversion was considered in [18], [19], [22]. With BT conversion, one trit error at the output may cause more than one bit error. Thus, the conversion should be designed to minimize these bit errors.

The block diagram of a TTCM system is shown in Fig. 1. First, binary input data is converted to ternary data and then encoded by the TCC. The encoded symbols along with uncoded binary or/and ternary data are mapped to symbols in the signal constellation. These symbols are then transmitted over the channel. In this paper, additive white Gaussian noise (AWGN) and Rayleigh fading channels are considered. At the receiver, the received signals are demodulated and decoded using a soft decision ternary Viterbi decoder [18], [20]. Finally, the resulting trits are converted to bits.

The contributions of this paper are as follows.

1) TCM is proposed which employs TQAM rather than SQAM or PSK. In addition, ternary set partitioning (TSP) is introduced and applied to TQAM.
2) The performance of TCM with TQAM and ternary convolutional coding (TTCM) is evaluated and compared with that of TCM with SQAM and binary convolutional coding (BTCM). In particular, a detailed comparison of BTCM with 16 SQAM (16 S-BTCM) and TTCM with 18 TQAM (18 TTCM) is presented to illustrate the advantages of TTCM. Several 18 TQAM signal constellations are considered for TTCM, and their TTCM performance is evaluated and compared to that of BTCM over AWGN and Rayleigh fading channels.
3) A new method is presented to construct a TQAM constellation which is suitable for TTCM which is called compatible TQAM (C-TQAM). Further, a technique for determining the best BT conversion for TTCM is given. The codes which provide the largest d_{Efree} for C-TTCM are determined, and the corresponding asymptotic coding gains (ACGs) are obtained.

The rest of the paper is organized as follows. In Section II, signal constellations are examined, and the advantages of triangular constellations are discussed. Section III introduces BTCM with SQAM and considers BTCM with TQAM. Section IV presents TSP for TQAM and the TTCM analysis. Further, the construction of TQAM constellations which are compatible with TSP (denoted C-TQAM) is presented and the corresponding BT conversion is investigated. Performance and simulation results are given in Section V to illustrate the advantages of TTCM over BTCM. Finally, some conclusions are given in Section VI.

II. SIGNAL CONSTELLATIONS

The structure and shape of the signal constellation have a significant impact on the modulation power efficiency which can be expressed as [21]

$$\eta_P = \frac{d_{min}}{E_b}, \tag{1}$$

where E_b is the average bit energy and d_{min} is the minimum Euclidean distance between constellation symbols. A better power efficiency means a lower E_b and/or larger d_{min}. Signal constellations with points close to the origin will have a low E_b, but a small distance between symbols. Thus, maximizing the power efficiency of a signal constellation is a tradeoff between minimizing E_b and maximizing d_{min}. A larger d_{min} results in a lower BER, so the constellation should have a large d_{min}. Consider the 16 PSK and 16 SQAM constellations shown in Figs. 2(a) and 2(b), respectively. With 16 PSK, $d_{min} = 2\sqrt{E_S}\sin(\frac{\pi}{16})$, while for 16 SQAM, $d_{min} = 2\sqrt{\frac{E_S}{10}}$, where $E_S = E_b \log_2 M$ [21]. For the same value of E_S, the distance between 16 SQAM symbols is 1.62 times larger than 16 PSK symbols. This implies that for the same BER, 16 SQAM requires a signal to noise ratio (SNR) $20\log_{10}(1.62) = 4.19$ dB less than 16 PSK [21].

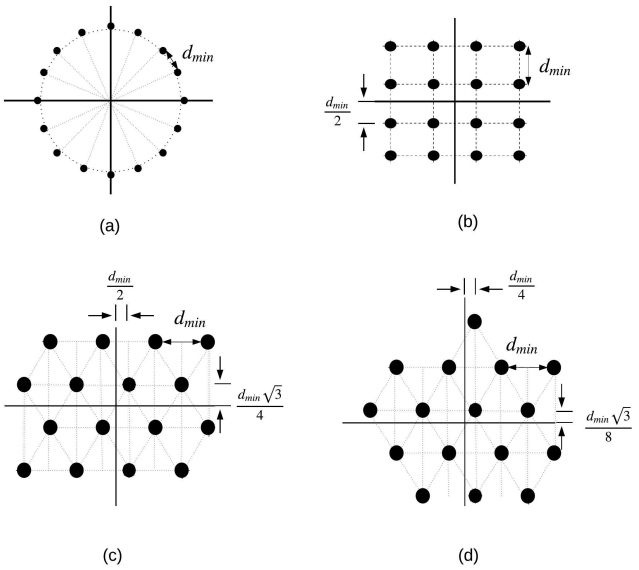

FIGURE 2. The (a) 16 PSK, (b) 16 SQAM, (c) 16 R-TQAM, and (d) 16 I-TQAM signal constellations.

It has been shown that TQAM, which has constellation points from a triangular lattice, provides a larger d_{min} than SQAM, which has points from a square lattice [9], [23]. Thus for the same average symbol energy and number of constellation points (called the modulation order M), TQAM has better BER performance than SQAM [9]. Fig. 2 shows the constellations of 16 TQAM, 16 SQAM and 16 PSK. For TQAM, the points on the edge of the constellation have at least two nearest neighbors (NNs), while the middle points have six NNs. Further, these points do not lie on lines parallel to the axes. Therefore for a given M, TQAM has a higher detection complexity than SQAM because it has a more complex structure and a larger average number of NNs [12]. For example, the average number of nearest neighbors for 16 SQAM is $NN = 3$, but for 16 TQAM it is $NN = 4.125$. The ML detection complexity of TQAM was considered in [9], [11], and shown to be higher than that of SQAM for all M.

Regular TQAM (R-TQAM) and irregular TQAM (I-TQAM) were introduced in [9], [11] and the constellations for $M = 16$ are shown in Figs. 2(c) and 2(d), respectively. R-TQAM has a symmetric constellation about the origin, while the I-TQAM constellation is more irregular. The power efficiency of I-TQAM is higher than that of R-TQAM, but it has a higher detection complexity [11]. For 16 R-TQAM, $d_{min} = 2\sqrt{\frac{E_S}{9}}$, while for 16 I-TQAM, $d_{min} = 2\sqrt{\frac{E_S}{8.75}}$, which is 1.054 and 1.069 times that of 16 SQAM for the same E_S, respectively [9], [11]. Therefore for the same BER, 16 R-TQAM and 16 I-TQAM require an SNR $20\log_{10}(1.054) = 0.46$ dB and $20\log_{10}(1.069) = 0.58$ dB less than 16 SQAM, respectively. Further, 16 R-TQAM has $NN = 4.123$ and 16 I-TQAM has $NN = 4.3125$, so the detection complexity of I-TQAM is higher. Table 1 gives the values of d_{min} for SQAM and TQAM with $4 \leq M \leq 265$.

TABLE 1. d_{min} for SQAM, R-TQAM, and I-TQAM.

M	d_{min} SQAM	R-TQAM	I-TQAM
8	$2\sqrt{\frac{E_S}{6}}$	$2\sqrt{\frac{E_S}{5}}$	$2\sqrt{\frac{E_S}{4.5}}$
16	$2\sqrt{\frac{E_S}{10}}$	$2\sqrt{\frac{E_S}{9}}$	$2\sqrt{\frac{E_S}{8.75}}$
32	$2\sqrt{\frac{E_S}{20}}$	$2\sqrt{\frac{E_S}{17.75}}$	$2\sqrt{\frac{E_S}{16.625}}$
64	$2\sqrt{\frac{E_S}{42}}$	$2\sqrt{\frac{E_S}{37}}$	$2\sqrt{\frac{E_S}{35.25}}$
128	$2\sqrt{\frac{E_S}{82}}$	$2\sqrt{\frac{E_S}{72}}$	$2\sqrt{\frac{E_S}{70.54}}$
256	$2\sqrt{\frac{E_S}{170}}$	$2\sqrt{\frac{E_S}{149}}$	$2\sqrt{\frac{E_S}{141.01}}$

For the same E_S, SQAM has the smallest d_{min} and I-TQAM has the largest d_{min} for all values of M.

The second parameter that affects the signal constellation is the number of bit differences between adjacent symbols. A smaller number of differences results in better BER performance. With a Gray mapping, SQAM constellations have one bit difference between adjacent symbols. A Gray mapping is not possible with TQAM constellations because of the number of nearest neighbors. The Gray penalty G_P is the average number of bit differences between a symbol and its nearest neighbors, so $G_P = 1$ for a Gray mapping. G_P for R-TQAM and I-TQAM for $M = 16$ and 64 was given in [9], [11]. G_P is greater than 1 for both R-TQAM and I-TQAM, but it is larger with I-TQAM. However, TQAM has better BER performance than SQAM [9], [11]. Given the results in this section, TCM with TQAM is considered in this paper.

III. BINARY TRELLIS CODED MODULATION

Ungerboeck introduced binary TCM (BTCM), which combines a binary convolutional code (BCC) with M-ary SQAM or M-ary PSK [1]. The advantage of BTCM arises from the fact that the coding and modulation are combined. The performance of BTCM is determined by the asymptomatic coding gain (ACG) which is a function of the free Euclidean distance d_{Efree}. The free Euclidean distance is the shortest distance between symbol sequences so that

$$d_{Efree}^2 = \sum_{i=0}^{N-1} d_i^2, \tag{2}$$

and N is the length of the symbol sequence that diverges from a state in the trellis and then joins that state as shown in Fig. 3, and d_i is the distance between the symbol sequences at trellis stage i.

The asymptotic coding gain (ACG) of TCM is defined as [1]

$$\gamma = 10\log_{10}\left(\frac{d_{Efree}^2/E_S}{d_{min}^2/E_S'}\right), \tag{3}$$

where E_S and E_S' are the average symbol energy for coded and uncoded symbols, and d_{min} is the minimum Euclidean distance for the uncoded signal constellation. As E_S and E_S' are fixed, the goal with TCM is to maximize d_{Efree}.

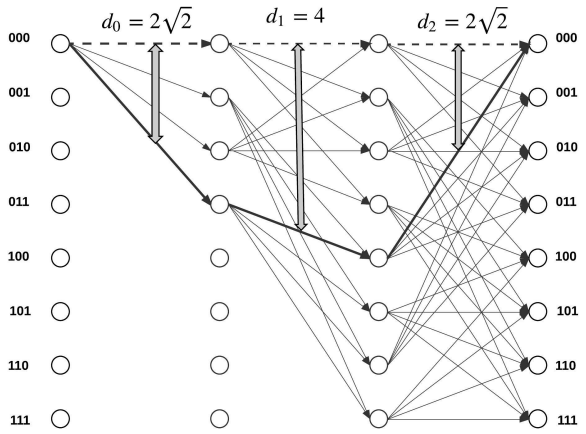

FIGURE 3. The binary trellis diagram for 16 S-BTCM.

The value of d_{Efree} depends on the mapping of constellation symbols and the code employed. Ungerboeck mapped the constellation symbols using a technique called set partitioning. Here, this is denoted as binary set partitioning (BSP). BSP divides the constellation symbols into two equal subsets, and the symbols in each subset are assigned the same binary label. The partitioning continues until there are only two symbols in each subset. The number of partition levels is L, where 2^{L+1} is the number of constellation symbols. For example, 16 SQAM has three BSP levels as shown in Fig. 4. The minimum distance between constellation symbols in the subsets is denoted by $d^{(l)}$, where l is the partition level. Note that $d^{(l)}$ increases as the level increases, which indicates that BSP is compatible with SQAM, i.e. in Fig. 4 $d^{(1)} = 2\sqrt{2}$, $d^{(2)} = 4$ and $d^{(3)} = 4\sqrt{2}$.

A BCC is used in BTCM to increase d_{Efree} by increasing the distance between sequences of constellation symbols. This code is represented by three parameters (n, k, m) where k and n are the number of input and output data streams

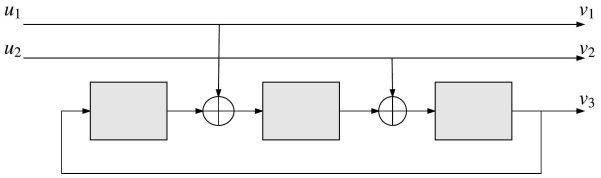

FIGURE 5. A systematic binary feedback convolutional code with $R = \frac{2}{3}$ and $m = 3$.

respectively, and m is the encoder memory length. The code rate of a BCC is $R = \frac{k}{n}$. A systematic feedback BCC is used in BTCM with $n = k + 1$ [1], [14]. Thus, a BCC(3, 2, 3) code is used with 16 S-BTCM as shown in Fig. 5, where $\oplus$ denotes modulo 2 addition. The generator sequences for this code are $G = [(0100), (0010), (1001)]$, where a 1 indicates a connection among the inputs, modulo 2 adders and/or memory elements, and a 0 indicates no connections. For example, the first sequence indicates a connection between the first input and the first modulo 2 adder, while the second sequence indicates a connection between the second input and the second modulo 2 adder. Each time step, 3 coded bits are combined with an uncoded bit and mapped to a 16 SQAM symbol. This is the best BCC for BTCM with 16 SQAM (16 S-BTCM) as it provides the maximum d_{Efree} [1]. The trellis for this coding is shown in Fig. 3 and has

$$\frac{d^2_{Efree}}{E_S} = \frac{(2\sqrt{2})^2}{10} + \frac{(2)^2}{10} + \frac{(2\sqrt{2})^2}{10} = 2. \tag{4}$$

The corresponding ACG over uncoded 8 PSK [1] is

$$\gamma = 10\log_{10}\left(\frac{20/10}{0.586/1}\right) = 5.33 \text{ dB}. \tag{5}$$

Now BTCM is considered with TQAM (T-BTCM), so that BSP is used for mapping the constellation symbols. Many symbols in a TQAM constellation have six neighbors, so it is not possible to have $d^{(1)} > d_{min}$. However in subsequent

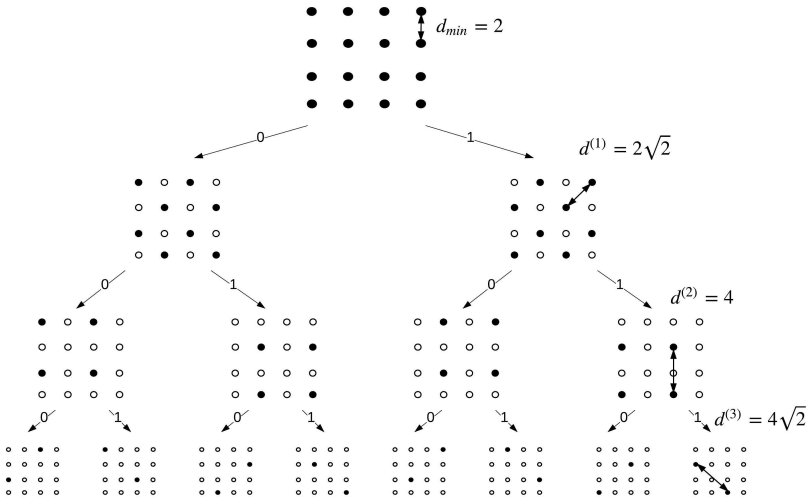

FIGURE 4. Binary set partitioning for 16 SQAM.

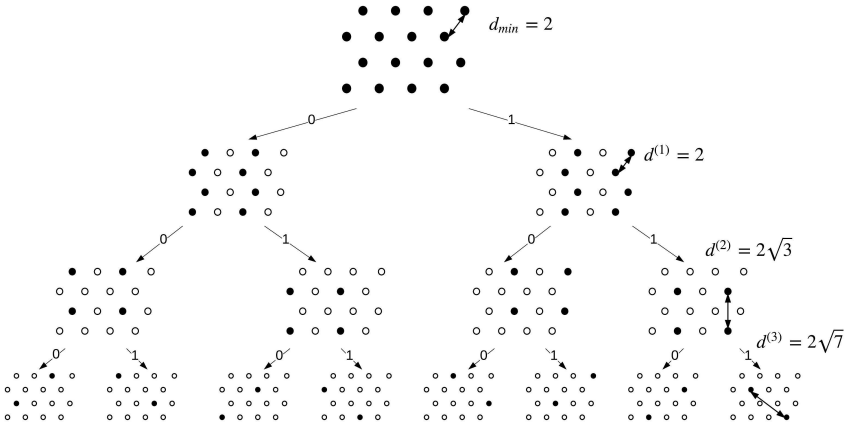

FIGURE 6. Binary set partitioning (BSP) for 16 R-TQAM.

levels, $d(l)$ can be increased as the number of NNs is reduced. For example, Fig. 6 shows BSP for 16 TQAM, and in the first level $d^{(1)} = d_{min} = 2$. Thus, BSP is not compatible with TQAM. In subsequent levels, the distances between constellation symbols in a subset increases to $d^{(2)} = 2\sqrt{3}$ and $d^{(3)} = 2\sqrt{7}$. The BCC that provides the largest d_{Efree} for 16 T-BTCM was obtained using a search as in [1]. For 16 T-BTCM, this code gives

$$\begin{aligned}\frac{d_{Efree}^2}{E_S} &= \frac{1}{E_S}(d^{(1)^2} + d^{(0)^2} + d^{(1)^2}),\\ &= \frac{2^2}{9} + \frac{2^2}{9} + \frac{2^2}{9} = 1.33, \end{aligned} \tag{6}$$

and the corresponding ACG over uncoded 8 PSK is

$$\gamma = 10\log_{10}\left(\frac{12/9}{0.586/1}\right) = 3.57\text{dB}. \tag{7}$$

From the previous results, the ACG of 16 T-BTCM is 1.76 dB less than the ACG of 16 S-BTCM, which confirms that BSP is not compatible with TQAM. Therefore, a suitable set partitioning for TQAM is introduced in the next section.

IV. TERNARY TRELLIS CODED MODULATION

A. TERNARY SET PARTITIONING

Set partitioning with TQAM should divide the constellation into more than two groups, so here three groups are considered which is called ternary set partitioning (TSP). As in [14], the number of constellation symbols in each subset should be equal, so the modulation order cannot be a power-of-two. Therefore, TSP can be employed if $M = 2^j \times 3^{t+1}$, where t is the number of coded and uncoded trits and j is the number of uncoded bits. For $j > 0$, TSP continues until the number of constellation symbols in each subset is 2 symbols, while for $j = 0$, the subsets in the last TSP level have 3 symbols. The number of set partitioning levels is $j + t$. For example, for $M = 18 = 2 \times 3^2$, the number of levels is 2, and each symbol corresponds to two trits and one bit. In addition, the subsets in the last level have 2 symbols. For $M = 27 = 3^3$, the number of TSP levels is 2, so the subsets in the last level have 3 symbols.

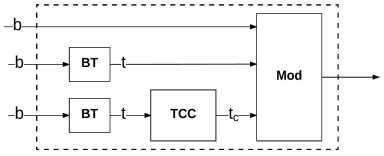

FIGURE 7. Block diagram for a TCM encoder which employs a TCC with uncoded bits and trits.

A ternary convolutional code is required with TSP as shown in Fig. 7. Here there are two possible cases, with or without uncoded bits and/or trits. For example, TCC(2, 1, 2) is employed with 27 TQAM so there are two coded trits and an uncoded trit. Conversely, 36 TQAM employs TCC(2, 1, 2) with two coded trits and two uncoded bits, and both TSP and BSP are used. BSP is used last because the uncoded bits have less protection, and the largest distance between constellation symbols is in the last level of partitioning [14]. In the next section, the performance of ternary trellis coded modulation (TTCM) with TSP is analyzed.

B. TTCM WITH 18 TQAM

A block diagram of the proposed ternary trellis coded modulation system is shown in Fig. 1. The TTCM encoder consists of a BT converter, TCC encoder, and modulator, while the decoder has a soft decision Viterbi decoder and BT converter. TTCM is embedded in a binary communication system where the input and output are binary data. First the binary data is converted to trits using BT conversion. These trits are then encoded by the TCC. The coded trits with uncoded bits and/or trits are mapped to a constellation symbol, and this symbol is labeled using TSP (and BSP if there are uncoded bits).

The binary to ternary (BT) conversion employed is important because it affects the BER performance [18]. At the decoder output, one trit error causes one or more bit errors because of this conversion. The goal is to minimize the aver-

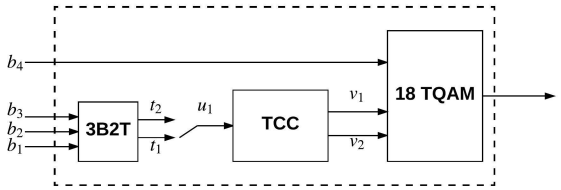

FIGURE 8. TTCM with 18 TQAM and a ternary convolutional code (TCC) using 3B2T conversion.

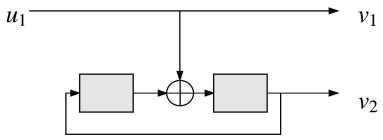

FIGURE 9. The best TCC for TTCM with 18 TQAM.

age number of bit errors due to a trit error, denoted e_{av}. In [18], three bit to two trit conversion (3B2T) was shown to provide the smallest value of e_{av}, and so 3B2T conversion is used here. For $M = 2^j \times 3^{t+1}$, each constellation symbol represents j bits and t trits. However, as the input to the system is binary, the number of bits per symbol with 3B2T conversion is

$$I = t \times \frac{3}{2} + j. \tag{8}$$

For example, with $M = 2^2 \times 3^2$, $I = 1 \times (1.5) + 2 = 3.5$ bits/symbol, and with $M = 2^0 \times 3^3 = 27$, $I = 2 \times (1.5) + 0 = 3$ bits/symbol.

To evaluate the performance of TTCM and compare it with that of BTCM, both 18 TQAM and 16 SQAM are considered in this section. The 18 TTCM block diagram is shown in Fig. 8. First, 3 bits are converted to 2 trits using 3B2T conversion, and the trits are input to the TCC. A ternary convolutional code TCC(2, 1, 2) is employed and the best code obtained using the search method in [1] is shown in Fig. 9. In this case, $\oplus$ denotes a modulo 3 adder. This code has

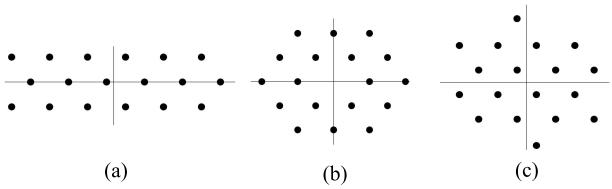

FIGURE 10. The constellations for (a) 18 R-TQAM, (b) 18 H-TQAM, and (c) 18 I-TQAM.

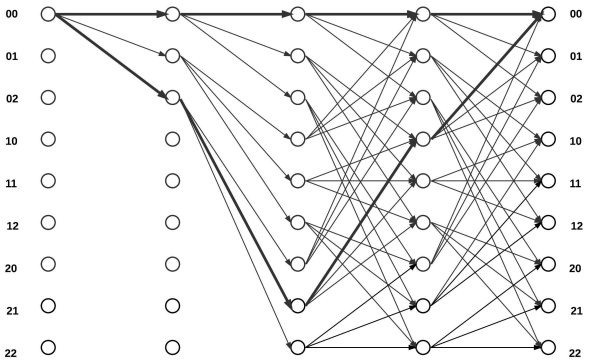

FIGURE 11. Trellis diagram for TCC(2,1,2).

generator sequences $G = [(010), (101)]$. The two encoded trits are combined with an uncoded bit and mapped to an 18 TTCM symbol. Therefore, there are $I = 2.5$ bits/symbol. In this section, regular, hexagonal and irregular 18 TQAM are considered, and the constellations are shown in Fig. 10. The structure of the constellation affects d_{Efree}, so the performance of TTCM will differ for these constellations.

Fig. 12 shows the ternary set partitioning for 18 R-TQAM. Note that $d^{(l)}$ increases with each level. For this constellation, if $d_{min} = 2$ then $E_S = 13.92$. The trellis diagram for the TCC in Fig. 9 is given in Fig. 11. This shows that four transitions are required to diverge and then merge with the same state,

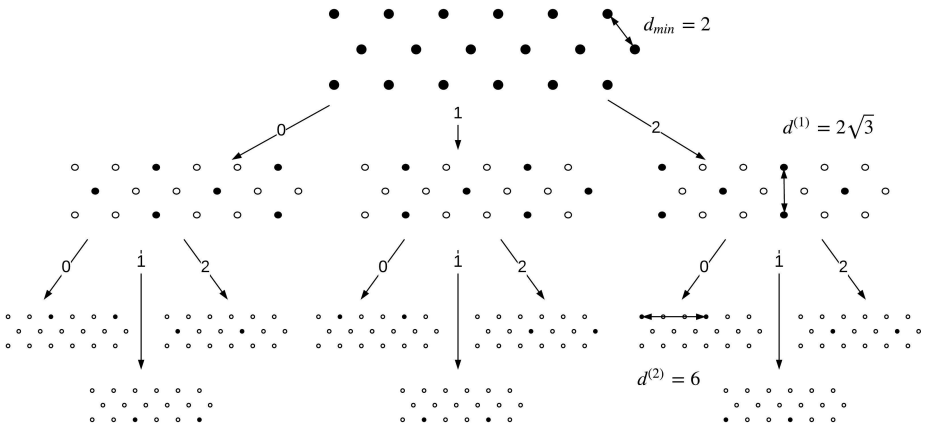

FIGURE 12. Ternary set partitioning for 18 R-TQAM.

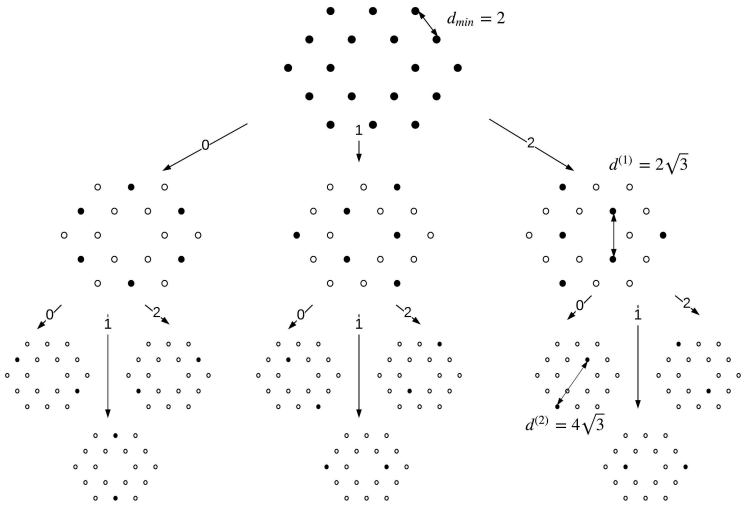

FIGURE 13. Ternary set partitioning for 18 H-TQAM.

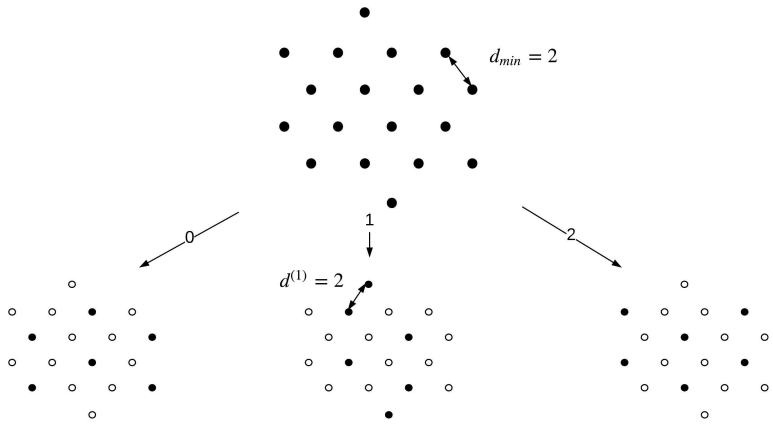

FIGURE 14. Ternary set partitioning for 18 I-TQAM.

so that for TTCM with 18 R-TQAM (18 R-TTCM)

$$\frac{d^2_{Efree}}{E_S} = \frac{(2\sqrt{3})^2 + 2^2 + (2\sqrt{3})^2 + 2^2}{13.92} = 2.30. \quad (9)$$

For comparison purposes, the ACG of 18 TTCM over uncoded 8 PSK is employed. However, each 18 TTCM symbol represents only $I = 2.5$ uncoded bits whereas uncoded 8 PSK represents 3 bits, which represents a loss of

$$10\log_{10}\left(\frac{3}{2.5}\right) = 0.792 \text{ dB}. \quad (10)$$

Thus, the ACG of 18 R-TTCM over uncoded 8 PSK is

$$\gamma = 10\log_{10}\left(\frac{2.30}{0.586}\right) - 0.792 = 5.14 \text{ dB}. \quad (11)$$

H-TQAM has constellation points from a triangular lattice in a hexagonal shape as shown in Fig. 10(b). Fig. (13) shows the TSP for 18 H-TQAM. The average symbol energy of 18 H-TQAM is $E_S = 10.67$ for $d_{min} = 2$, which is lower than with R-TQAM. TTCM with 18 H-TQAM (18 H-TTCM) gives

$$\frac{d^2_{Efree}}{E_S} = \frac{(2\sqrt{3})^2 + 2^2 + (2\sqrt{3})^2 + 2^2}{10.67} = 3.00, \quad (12)$$

and the corresponding ACG over 8 PSK is

$$\gamma = 10\log_{10}\left(\frac{3}{0.586}\right) - 0.792 = 6.30 \text{ dB}. \quad (13)$$

This is 1.17 dB better than R-TTCM because of the lower value of E_S.

The 18 I-TQAM constellation shown in Fig. 10(c) has $E_S = 10.17$, which is lower than with R-TQAM and H-TQAM. Fig. 14 shows the TSP for 18 I-TQAM, and illustrates that 18 I-TQAM is not compatible with TSP

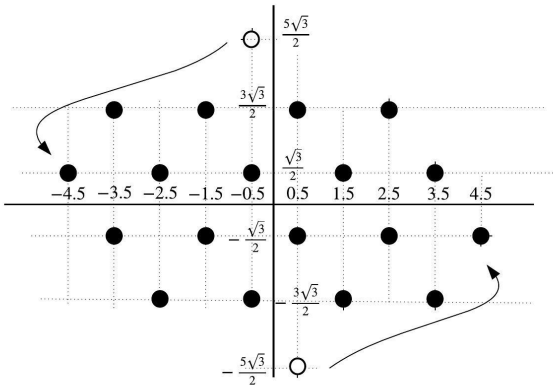

FIGURE 15. Modifying 18 I-TQAM to be compatible with TSP, denoted 18 C-TQAM.

as $d_{min} = d^{(1)}$. TTCM with 18 I-TQAM (18 I-TTCM) gives

$$\frac{d^2_{Efree}}{E_S} = \frac{2^2 + 2^2 + 2^2 + 2^2}{10.17} = 1.57, \tag{14}$$

and the corresponding ACG over 8 PSK is

$$\gamma = 10 \log_{10}\left(\frac{1.57}{0.586}\right) - 0.792 = 3.50 \text{ dB}. \tag{15}$$

I-TQAM provides the lowest ACG for the three 18 TQAM constellations, and is also 1.04 dB lower than 16 S-BTCM because of the incompatibility with TSP. To make 18 I-TQAM compatible with TSP, some of the constellation points must be relocated. Although E_S for this new constellation will be larger, the TTCM ACG will be lower. Fig. 15 shows the relocation of two 18 I-TQAM constellation points. This new constellation is compatible with TSP, and so is called compatible 18 TQAM, (18 C-TQAM). For $d_{min} = 2$, $E_S = 10.33$, which is lower than with R-TQAM and H-TQAM, but higher than with 18 I-TQAM. Fig. 16 shows the TSP for 18 C-TQAM. TTCM with 18 C-TQAM (18 C-TTCM) gives

$$\frac{d^2_{Efree}}{E_S} = \frac{(2\sqrt{3})^2 + 2^2 + (2\sqrt{3})^2 + 2^2}{10.33} = 3.09, \tag{16}$$

and the corresponding ACG over 8 PSK is

$$\gamma = 10 \log_{10}\left(\frac{3.09}{0.586}\right) - 0.792 = 6.43 \text{ dB}. \tag{17}$$

As expected, C-TTCM provides the best ACG because E_S for C-TQAM is lower than with R-TQAM and H-TQAM, and it is compatible with TSP.

TABLE 2. Comparison between 18 TTCM and 16 S-BTCM.

	E_S	$\frac{d^2_{Efree}}{E_S}$	ACG over 8 PSK
18 R-TTCM	13.92	2.30	5.14
18 H-TTCM	10.67	3.00	6.30
18 I-TTCM	10.17	1.57	3.50
18 C-TTCM	10.33	3.09	6.43
16 S-BTCM	10.00	2.00	5.33

Table 2 gives the values of E_S, d^2_{Efree}/E_S, and ACG for regular, irregular, hexagonal and compatible 18 TTCM, and 16 S-BTCM. From the table, 18 H-TTCM and C-TTCM have a better ACG than 16 S-BTCM by $6.30-5.33 = 0.97$ dB and $6.43-5.33 = 1.10$ dB, respectively. Thus, C-TTCM provides the highest ACG. The construction of M-ary C-TTCM is considered in the next section.

C. M-ARY C-TTCM

The I-TQAM constellations presented in [11] provide the lowest values of E_S, but are not compatible with TSP. The I-TQAM constellations for small M can easily be modified to obtain compatible constellations, but for large M this can be difficult. Therefore, a construction technique for compatible

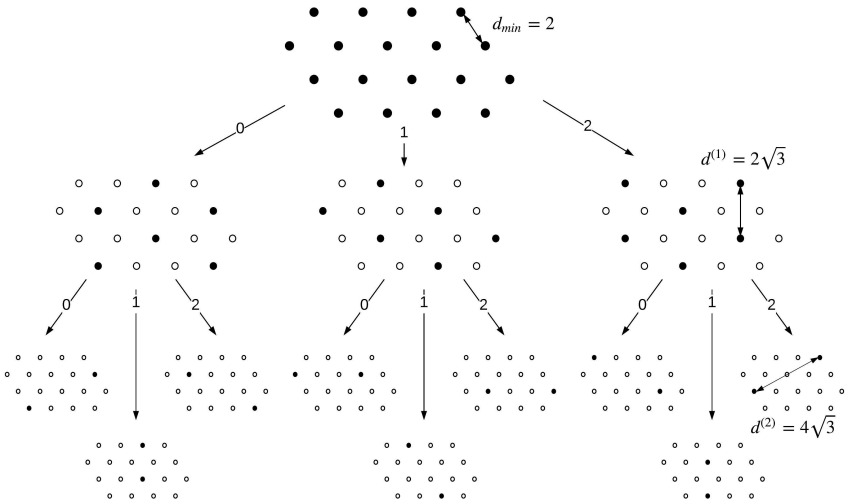

FIGURE 16. Ternary set partitioning for 18 C-TQAM.

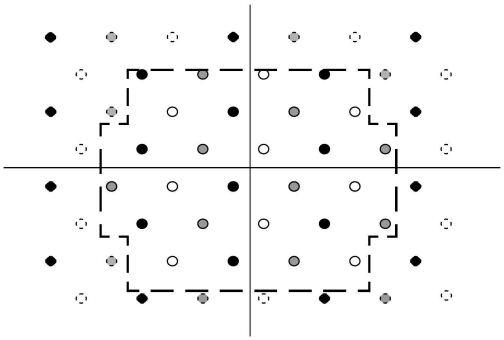

FIGURE 17. The 27 C-TQAM constellation which is compatible with TSP.

TQAM (C-TQAM) is introduced here and is given in the following steps.

1) Construct a rectangle of points from a triangular lattice with the origin at the center. The rectangle dimensions are the two integers closest to $\sqrt{2M}$, with the largest equal to the number of rows and the smallest equal to the number of columns. In particular, the number of rows is $\lceil\sqrt{2M}\rceil$ and the number of columns is $\lfloor\sqrt{2M}\rfloor$.
2) Divide the points into three groups as follows.
   a) The first group is selected by choosing the first of every three points in the odd numbered rows, and the second of every three points in the even numbered rows.
   b) The second group is selected by choosing the second of every three points in the odd numbered rows, and the third of every three points in the even numbered rows.
   c) The third group is selected by choosing the third of every three points in the odd numbered rows, and the first of every three points in the even numbered rows.
3) The $\frac{M}{3}$ constellation points in each group closest to the origin are selected, and the remaining points are deleted.
4) The selected points constitute the M-ary C-TQAM constellation that is compatible with TSP and has the lowest value of E_S.

For example, Fig. 17 shows the 27 C-TQAM constellation. In the construction, a rectangle of $2 \times 27 = 54$ points is used, so the number of rows is 8 and the number of points in each row is 7. The colors in the figure represent the three groups. The $\frac{27}{3} = 9$ points in each group closest to the origin are selected (shown within the dashed line), and the remaining points are deleted. This construction ensures that $d^{(1)} > d_{min}$. The resulting constellations for $M = 27$, 36, 54 and 108 are shown in Fig. 18. The constellation of 36 H-TQAM in [10] is compatible with TSP but $E_s = 30.85$ for $d_{min} = 2$. Applying the above construction technique for 36 C-TQAM gives $E_s = 20.33$ for $d_{min} = 2$ as shown in Fig. 18. Therefore, 36 C-TQAM has a higher power efficiency than 36 H-TQAM in [10].

TABLE 3. Parameters for M-ary C-TTCM with TCC(2, 1, m).

R	M	m	E_S	$\frac{d^2_{Efree}}{E_S}$	G	I
$\frac{1}{2}$	9	1	4.92	1.62	[(01), (11)]	1.5
		2		2.02	[(010), (101)]	
	18	2	10.33	3.16	[(010), (101)]	2.5
		3		3.81	[(0120), (2021)]	
	27	2	15.00	3.20	[(010), (101)]	3
		3		3.92	[(0110), (2101)]	
		4		4.25	[(01000), (11002)]	
	36	2	20.33	3.52	[(010), (101)]	3.5
		3		4.33	[(0110), (1011)]	
		4		4.52	[(01000), (12011)]	

The TCC code rate is related to the TQAM modulation order M. Tables 3 and 4 present the parameters for C-TTCM with $R = \frac{1}{2}$ and $R = \frac{2}{3}$, respectively. The best TCCs were found using the search method in [1] and considering the unequal error probabilities of the trits. The average symbol energy was calculated using the constellations shown in Fig. 18 with $d_{min} = 2$. The corresponding values of d_{Efree} are given in the tables.

D. BINARY TO TERNARY CONVERSION IN TTCM

3B2T conversion maps every string of 3 bits to a unique string of 2 trits. The goal is to have a low value of e_{av} considering d_{Efree} rather than d_{min} as in [18]. A TTCM decoding error occurs when a wrong trellis path is chosen. Therefore, e_{av} can be reduced by minimizing the number of trit differences between branches leaving a state.

TCC(2, 1, 2) is considered here to illustrate the steps in creating the 3B2T conversion that provides the lowest e_{av}. The trellis diagram of TCC(2, 1, 2) is shown in Fig. 11. For 18 C-TTCM, the lines between states represent two parallel branches, so each state has six output branches. The outputs of the first state are (000, 001, 100, 101, 200, 201), where the first two digits are encoded trits and the last is an uncoded bit. The focus here is on trit errors, so the uncoded bit is ignored, in which case there are only three branches leaving a state given by (00, 10, 20). These ternary strings are mapped to binary strings so that the average bit difference is lowest. For example, mapping to (000, 001, 101) results in an average bit difference of 1.33 bits.

The difference between the three ternary blocks is one trit, so choosing an incorrect branch results in one trit error, or from the given mapping 1.33 bit errors on average. The same occurs for the fourth and seventh states in this example. Following a similar procedure for the second, fifth, and eighth states, (01, 11, 21) is mapped to (111, 011, 010). The average difference between these ternary strings is one trit, and the average bit difference is 1.33 bits. For the third, sixth, and ninth states, (02, 12, 22) is mapped to the remaining binary strings giving (110, 100, 100), where the last binary string is repeated. A difference between these ternary strings of one trit results in an average difference between the mapped binary strings of 0.67 bit. Thus, the average number

TABLE 4. Parameters for *M*-ary C-TTCM with TCC(3, 2, *m*).

R	M	m	E_S	$\frac{d^2_{Efree}}{E_S}$	G	I
$\frac{2}{3}$	27	2	15.00	2.75	[(010), (001), (101)]	3
		3		3.33	[(0100), (0010), (1201)]	
		4		3.62	[(01000), (01100), (12101)]	
	54	2	28.47	3.25	[(010), (001), (101)]	4
		3		3.55	[(0100), (0010), (2102)]	
		4		4.02	[(01000), (02100), (12201)]	
	81	3	49.90	3.67	[(0100), (0010), (1201)]	4.5
		4		4.22	[(01000), (02100), (12201)]	
	108	3	62.77	4.22	[(0100), (0010), (1201)]	5
		4		4.84	[(01000), (02100), (12201)]	

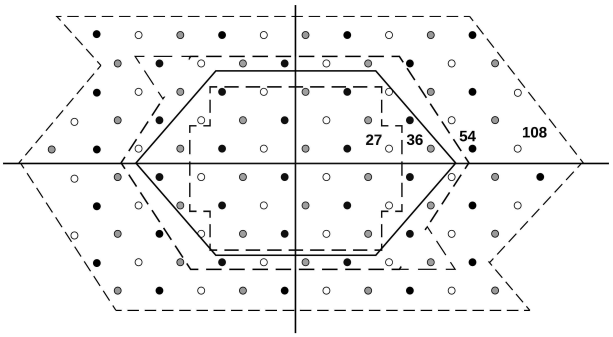

FIGURE 18. The C-TQAM constellations for $M = 27$, 36, 54 and 108.

of bit errors due to a trit error for this 3B2T conversion in TTCM is

$$e_{av} = \frac{6(1.33) + 3(0.67)}{9} = 1.11 \text{ bits/trit error.} \quad (18)$$

This 3B2T conversion is shown in Table 5. Based on an exhaustive search, it has the lowest average number of bit errors for BT conversion with 18 C-TTCM using TCC(2, 1, 2).

The 3B2T conversion given above is used for coded trits. For uncoded trits, the 3B2T conversion in [18] is optimal and has $e_{av} = 1.55$. For example, 27 C-TTCM with TCC(2, 1, 2) has two coded trits and one uncoded trit. Therefore, the first two trits are mapped using Table 5, while the uncoded trit is mapped using the table in [18]. Thus for the three trits, $e_{av} = \frac{1.55+2(1.11)}{3} = 1.26$ bits/trit error. Note that the output trits from 3B2T have unequal probabilities [24]. In particular, the probabilities of 0 and 1 are 0.375 while the probability of 2 is 0.25. These probabilities were used in determining the BER results.

V. PERFORMANCE RESULTS

In this section, the probability of bit error for 16 S-BTCM and 18 TTCM is evaluated over additive white Gaussian noise (AWGN) and Rayleigh fading channels. Monte Carlo simulation is employed with 10^8 bits for each value of E_b/N_0 where E_b is the energy per bit given by $E_b = E_s/I$ and N_0 is the noise power spectral density. $I = 3$ for 16 S-BTCM

TABLE 5. Optimal three bit to two trit conversion.

Binary Block Input	Ternary Block	Binary Block Output
000	00	000
001	10	001
101	20	101
111	01	111
011	11	011
010	21	010
110	02	110
100	12	100
xxx	22	100

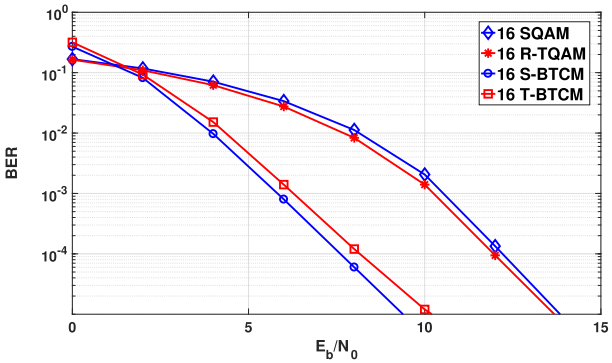

FIGURE 19. Bit error rates for 16 S-BTCM and 16 T-BTCM over an AWGN channel.

while $I = 2.5$ for 18 TTCM. For a fair comparison, the values of E_b for 16 S-BTCM and 18 TTCM should be the same. This is done by reducing d_{min} of 16 S-BTCM and 18 TTCM by $\sqrt{\frac{I}{E_S}}$ [21]. Therefore from Tables 2 and 3, d_{min} becomes 1.095, 0.848, 0.968, and 0.984 for 16 SQAM, 18 R-TQAM, 18 H-TQAM, and 18 C-TQAM, respectively. These scaled values of d_{min} affect both $d^2_E free$ and E_S, so the resulting values of $\frac{d^2_E free}{E_S}$ and γ are the same as in Table 2.

The bit error rate (BER) over an AWGN channel for 16 S-BTCM and 16 T-BTCM, and uncoded 16 SQAM and 16 TQAM, is shown in Fig. 19. The BER performance of uncoded 16 TQAM is better than uncoded 16 SQAM by 0.43 dB. However, 16 S-BTCM is better than 16 T-BTCM by 1.42 dB at BER $= 10^{-5}$, which shows that BSP is not compatible with TQAM.

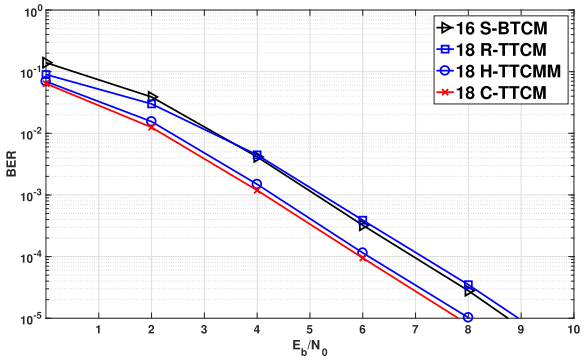

FIGURE 20. Bit error rates for TTCM with 18 R-TQAM, 18 C-TQAM, and 18 H-TQAM, and 16 S-BTCM over an AWGN channel.

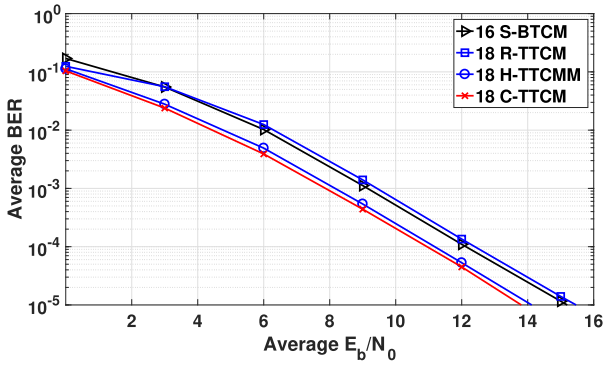

FIGURE 21. Bit error rates for TTCM with 18 R-TQAM, 18 C-TQAM, and 18 H-TQAM, and 16 S-BTCM over a Rayleigh fading channel.

Fig. 20 presents the BER for 18 TTCM and 16 S-BTCM. 18 C-TTCM is better than 16 S-BTCM by 0.93 dB at BER $= 10^{-5}$, while 18 H-TTCM is better than 16 B-STCM by 0.81 dB. However, the performance of 16 S-BTCM is better than that of 18 R-TTCM by 0.12 dB. Therefore, for the same average bit energy, 18 C-TTCM and 18 H-TTCM provide better BER performance than 16 S-BTCM. Fig. 21 presents the BER for 18 TTCM and 16 S-BTCM over a frequency-flat Rayleigh fading channel. The average SNR is $\bar{E}_b/N_0$ where $\bar{E}_b = \bar{E}_S/I$ and $\bar{E}_S = \mathrm{E}\{E_S\}$. These results show that 18 C-TTCM and 18 H-TTCM are better than 16 S-BTCM by 0.91 dB and 0.79 dB, respectively, at BER $= 10^{-5}$. Thus, the performance of 18 TTCM is closer to that of 16 S-BTCM in Rayleigh fading channels compared to AWGN channels.

VI. CONCLUSION

A new form of coded modulation was introduced called ternary trellis coded modulation (TTCM). This employs triangular quadrature modulation (TQAM) which provides better performance than square QAM. It was shown that binary set partitioning (BSP) is not compatible with TQAM, so ternary set partitioning (TSP) was employed. The advantage of TTCM over binary TCM (BTCM) with SQAM was illustrated. A new class of TQAM constellations called compatible TQAM (C-TQAM) was introduced that is compatible with TSP and has a low average symbol energy. As communication systems have binary inputs and outputs, a new binary to ternary (BT) conversion technique was proposed for TTCM. For the same average bit energy, results were presented which show that the bit error rate (BER) with 18 C-TTCM and 18 H-TTCM is better than that with 16 S-BTCM in both AWGN and Rayleigh fading channels. Further, TTCM with C-TQAM provides the best performance compared with the TQAM constellations in the literature, namely regular TQAM (R-TQAM), irregular TQAM (I-TQAM), and hexagonal TQAM (H-TQAM).

REFERENCES

[1] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.

[2] S. Alreesh, C. Schmidt-Langhorst, R. Emmerich, P. W. Berenguer, C. Schubert, and J. K. Fischer, "Four-dimensional trellis coded modulation for flexible optical communications," *J. Lightw. Technol.*, vol. 35, no. 2, pp. 152–158, Jan. 2017.

[3] H. Sun, S. X. Ng, C. Dong, and L. Hanzo, "Decode-and-forward cooperation-aided triple-layer turbo-trellis-coded hierarchical modulation," *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1136–1148, Apr. 2015.

[4] B. Di, L. Song, and Y. Li, "Trellis coded modulation for non-orthogonal multiple access systems: Design, challenges, and opportunities," *IEEE Wireless Commun.*, vol. 25, no. 2, pp. 68–74, Apr. 2018.

[5] C. N. Campopiano and B. G. Glazer, "A coherent digital amplitude and phase modulation scheme," *IRE Trans. Commun. Syst.*, vol. 10, no. 1, pp. 90–95, Mar. 1962.

[6] M. Simon and J. Smith, "Hexagonal multiple phase-and-amplitude-shift-keyed signal sets," *IEEE Trans. Commun.*, vol. 21, no. 10, pp. 1108–1115, Oct. 1973.

[7] C. Cahn, "Combined digital phase and amplitude modulation communication systems," *IRE Trans. Commun. Syst.*, vol. 8, no. 3, pp. 150–155, Sep. 1960.

[8] G. Foschini, R. Gitlin, and S. Weinstein, "Optimization of two-dimensional signal constellations in the presence of Gaussian noise," *IEEE Trans. Commun.*, vol. 22, no. 1, pp. 28–38, Jan. 1974.

[9] S. J. Park, "Triangular quadrature amplitude modulation," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 292–294, Apr. 2007.

[10] M. Tanahashi and H. Ochiai, "A multilevel coded modulation approach for hexagonal signal constellation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 10, pp. 4993–4997, Oct. 2009.

[11] S.-J. Park and M.-K. Byeon, "Irregularly distributed triangular quadrature amplitude modulation," in *Proc. IEEE 19th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Cannes, France, Sep. 2008, pp. 1–5.

[12] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. Expected complexity," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2806–2818, Aug. 2005.

[13] L.-F. Wei, "Trellis-coded modulation with multidimensional constellations," *IEEE Trans. Inf. Theory*, vol. 33, no. 4, pp. 483–501, Jul. 1987.

[14] S. S. Pietrobon, R. H. Deng, A. Lafanechere, G. Ungerboeck, and D. J. Costello, "Trellis-coded multidimensional phase modulation," *IEEE Trans. Inf. Theory*, vol. 36, no. 1, pp. 63–89, Jan. 1990.

[15] R. Wesel, X. Liu, J. Cioffi, and C. Komninakis, "Constellation labeling for linear encoders," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2417–2431, Sep. 2001.

[16] T. N. Rajashekhara and I.-S. E. Chen, "A fast adder design using signed-digit numbers and ternary logic," in *Proc. IEEE Tech. Conf. Southern Tier*, Binghamton, NY, USA, Apr. 1990, pp. 187–194.

[17] B. Hayes, "Third base," *Amer. Scientist*, vol. 89, no. 6, pp. 490–494, Nov./Dec. 2001.

[18] M. Abdelaziz and T. A. Gulliver, "Ternary convolutional codes for ternary phase shift keying," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1709–1712, Sep. 2016.

[19] N. Ekanayake and T. T. Tjhung, "On ternary phase-shift keyed signaling," *IEEE Trans. Inf. Theory*, vol. 28, no. 4, pp. 658–660, Jul. 1982.

[20] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York, NY, USA: McGraw-Hill, 1979.

[21] U. Madhow, *Fundamentals of Digital Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[22] T. Koike and S. Yoshida, "Space-time trellis-coded ternary PSK for mobile communications," *Electron. Lett.*, vol. 40, no. 16, pp. 1011–1012, Aug. 2004.
[23] J. Lee, D. Yoon, and K. Cho, "Error performance analysis of M-ary θ-QAM," *IEEE Trans. Veh. Techol.*, vol. 61, no. 3, pp. 1423–1427, Mar. 2012.
[24] M. Abdelaziz and T. A. Gulliver, "Triangular constellations for adaptive modulation," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 756–766, Feb. 2018.

MAHMOUD ABDELAZIZ received the M.Sc. degree in electrical engineering from the Military Technical College, Egypt, in 2011, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 2017. He is currently a Lecturer with the Department of Electrical Engineering, Military Technical College. His research interests include satellite communications, mobile communications, communication theory, and channel coding.

T. AARON GULLIVER received the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 1989. From 1989 to 1991, he was a Defence Scientist with Defence Research Establishment Ottawa, Ottawa, ON, Canada. He has held academic appointments with Carleton University, Ottawa, and the University of Canterbury, Christchurch, New Zealand. He joined the University of Victoria, in 1999, where he is currently a Professor with the Department of Electrical and Computer Engineering. His research interests include information theory and communication theory, algebraic coding theory, multicarrier systems, smart grids, intelligent networks, cryptography, and security. In 2002, he became a Fellow of the Engineering Institute of Canada. In 2012, he was elected as a Fellow of the Canadian Academy of Engineering. From 2007 to 2012, he was an Editor, and from 2012 to 2017, an Area Editor, of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.