CLAIMS

1. A data processing method, comprising:

performing probabilistic constellation shaping, PCS, processing on a first bit set in k bits, to obtain a second bit set, wherein k is an integer greater than 1;

performing forward error correction, FEC, encoding on the second bit set and a third bit set in the k bits excluding the first bit set, to obtain a fourth bit set, wherein the fourth bit set comprises m_0 first bit subsets, each of the first bit subsets comprises F_0 bits, m_0 is an integer greater than 1, and F_0 is an even number greater than 1; and

performing first interleaving processing on the fourth bit set to obtain a fifth bit set, wherein the fifth bit set comprises m_0 second bit subsets, each of the second bit subsets comprises F_0 bits, $F_0/2$ bits in each of the second bit subsets are from the second bit set, and the other $F_0/2$ bits in each of the second bit subsets are from the third bit set and/or parity bits of the FEC encoding.

2. The method according to claim 1, wherein the fifth bit set comprises $m_0 \times F_0$ bits, and each of the second bit subsets comprises F_0 consecutive bits.

3. The method according to claim 1, wherein each of the second bit subsets is presented as a square block.

4. The method according to any one of claims 1 to 3, wherein F_0 bits in each of the second bit subsets are distributed in F_1 rows and F_1 columns, $F_1/2$ bits of F_1 bits in each row of each of the second bit subsets are from the second bit set, the second bit set comprises $m_0 \times F_0/2$ bits, and the other $F_1/2$ bits of the F_1 bits in each row of each of the second bit subsets are from the third bit set and/or the parity bits of the FEC encoding.

5. The method according to claim 4, wherein $F_0 = 256$ and $F_1 = 16$.

6. The method according to claim 4 or 5, wherein $F_0 = 256$, $F_1 = 16$, and a bit in row i_2 and column $\left(\left\lfloor \frac{j_2}{2} \right\rfloor \times 4 + j_2\%2 + i_2 + 2\right)\%16$ of each of the second bit subsets is from the second bit set, wherein $0 \le i_2 < 16$ and $0 \le j_2 < 8$; or

a bit in row i_2 and column $(j_2 \times 2 - i_2\%2 + 1)$ of each of the second bit subsets is from the second bit set, wherein $0 \le i_2 < 16$ and $0 \le j_2 < 8$, wherein

$\lfloor a \rfloor$ represents performing a floor operation on a, and $b\%c$ represents b modulo c.

7. The method according to claim 4 or 5, wherein $F_0 = 256$, $F_1 = 16$, and the bit in each of the second bit subsets satisfies one or more of the following:

bits in row 0 and columns 2, 3, 6, 7, 10, 11, 14, and 15 of each of the second bit subsets are from the second bit set;

bits in row 1 and columns 0, 3, 4, 7, 8, 11, 12, and 15 of each of the second bit subsets are from the second bit set;

bits in row 2 and columns 0, 1, 4, 5, 8, 9, 12, and 13 of each of the second bit subsets are from the second bit set;

bits in row 3 and columns 1, 2, 5, 6, 9, 10, 13, and 14 of each of the second bit subsets are from the second bit set;

bits in row 4 and columns 2, 3, 6, 7, 10, 11, 14, and 15 of each of the second bit subsets are from the second bit set;

bits in row 5 and columns 0, 3, 4, 7, 8, 11, 12, and 15 of each of the second bit subsets are from the second bit set;

bits in row 6 and columns 0, 1, 4, 5, 8, 9, 12, and 13 of each of the second bit subsets are from the second bit set;

bits in row 7 and columns 1, 2, 5, 6, 9, 10, 13, and 14 of each of the second bit subsets are from the second bit set;

bits in row 8 and columns 2, 3, 6, 7, 10, 11, 14, and 15 of each of the second bit subsets are from the second bit set;

bits in row 9 and columns 0, 3, 4, 7, 8, 11, 12, and 15 of each of the second bit subsets are from the second bit set;

bits in row 10 and columns 0, 1, 4, 5, 8, 9, 12, and 13 of each of the second bit subsets are from the second bit set;

bits in row 11 and columns 1, 2, 5, 6, 9, 10, 13, and 14 of each of the second bit subsets are from the second bit set;

bits in row 12 and columns 2, 3, 6, 7, 10, 11, 14, and 15 of each of the second bit subsets are from the second bit set;

bits in row 13 and columns 0, 3, 4, 7, 8, 11, 12, and 15 of each of the second bit subsets are from the second bit set;

bits in row 14 and columns 0, 1, 4, 5, 8, 9, 12, and 13 of each of the second bit subsets are from the second bit set; and

bits in row 15 and columns 1, 2, 5, 6, 9, 10, 13, and 14 of each of the second bit subsets are from the second bit set.

8. The method according to any one of claims 1 to 7, wherein $F_0 = 256$, and a total of $m_1 \times 256$ bits in m_1 first bit subsets in the fourth bit set are from the second bit set, wherein $m_0=m_1 \times 2$.

9. The method according to any one of claims 1 to 8, wherein $F_0 = 256$, $m_0 \times 256$ bits in the fourth bit set are distributed in 32 rows and $m_0 \times 8$ columns, $m_0 \times 256$ bits in the fifth bit set are distributed in 32 rows and $m_0 \times 8$ columns, and $m_0 \times 8$ bits in row r_1 in the fifth bit set are from $m_0 \times 8$ bits in row r_0 in the fourth bit set, wherein $0 \leq r_0 < 32$ and $0 \leq r_1 <$

32.

10. The method according to any one of claims 1 to 9, wherein the performing FEC encoding on the second bit set and a third bit set in the k bits excluding the first bit set, to obtain a fourth bit set comprises:

performing pre-encoding interleaving processing on the second bit set and the third bit set to obtain a sixth bit set, wherein a quantity of bits in the sixth bit set is equal to a sum of a quantity of bits in the second bit set and a quantity of bits in the third bit set, the sixth bit set comprises m_3 third bit subsets, the m_3 third bit subsets are distributed in two rows and m_4 columns, $m_3 = 2 \times m_4$, and m_3 is an integer greater than 1 and less than m_0, wherein some third bit subsets each comprises F_0 bits, the other third bit subsets each comprises F_2 bits, and F_2 is an even number greater than 1 and less than F_0; and

performing FEC encoding on the sixth bit set to obtain the fourth bit set.

11. The method according to claim 10, wherein m_3=14, m_4=7, F_0=256, F_2=240, each third bit subset in column 0 to column 5 comprises bits in 16 rows and 16 columns, each third bit subset in column 6 comprises bits in 16 rows and 15 columns, $F_0/2$ bits in each third bit subset in column 0 to column 4 are from the second bit set, the other $F_0/2$ bits in each third bit subset in column 0 to column 4 are from the third bit set, 9 bits in each row of each third bit subset in column 5 are from the second bit set, the other 7 bits in each row of each third bit subset in column 5 are from the third bit set, and F_2 bits in each third bit subset in column 6 are from the second bit set.

12. The method according to any one of claims 9 to 11, wherein $r_0 = r_1$.

13. The method according to any one of claims 1 to 12, wherein m_0=12 or 16.

14. The method according to any one of claims 1 to 13, wherein after the performing interleaving processing on the fourth bit set to obtain a fifth bit set, the method further comprises:

performing second interleaving processing on every two first bit streams in L first bit streams to obtain a total of L/2 second bit streams, wherein each first bit stream comprises a plurality of fifth bit sets, and L is an even number greater than 0;

performing symbol mapping and polarization distribution on the L/2 second bit streams to obtain one dual-polarization symbol stream, wherein t bits are mapped to one dual-polarization symbol through the symbol mapping and the polarization distribution, and t is an integer greater than 0; and

performing digital signal processing, DSP, framing on the dual-polarization symbol stream.

15. The method according to claim 14, wherein an amplitude bit in the dual-polarization symbol is from the second bit set.

16. The method according to claim 15, wherein t=8, and the dual-polarization symbol is a dual-polarization, DP-16QAM, symbol.

17. The method according to claim 16, wherein the performing DSP framing on the dual-polarization symbol stream comprises:

performing framing processing on every 172032 dual-polarization DP-16QAM symbols to obtain one DSP super-frame, wherein the super-frame comprises 175104 dual-polarization symbols.

18. The method according to any one of claims 1 to 17, wherein a quantity of bits in the first bit set is an integer multiple of 2, 4, 8, or 16, the second bit set comprises 2048 bits, the third bit set comprises 1504 bits, and the fourth bit set comprises 4096 bits.

19. A data processing apparatus, comprising a processor and a transceiver, wherein the processor is configured to perform the method according to any one of claims 1 to 18, and the transceiver is configured to perform data sending and receiving operations.

20. A communication system, comprising a transmitter data processing apparatus and a receiver data processing apparatus, wherein the transmitter data processing apparatus is configured to perform the method according to any one of claims 1 to 18, and the transmitter data processing apparatus is configured to send a data stream to the receiver data processing apparatus.