Multilevel Modulation Coding For Four-Level Holographic Data Storage Systems

Chi Dinh Nguyen[1,2], Nghia Xuan Pham[3], Cuong Chinh Duong[1], and Luong Cong Nguyen[2,4]
[1]Faculty of Electrical and Electronic Engineering, Phenikaa University, Hanoi 12116, Vietnam
[2]Phenikaa Research and Technology Institute (PRATI), A&A Green Phoenix Group JSC, No.167 Hoang Ngan, Trung Hoa, Cau Giay, Hanoi 11313, Vietnam
[3]Faculty of Radio - Electronics Engineering, Le Quy Don Technical University, Hanoi 100000, Vietnam
[4]Faculty of Information Technology, Phenikaa University, Hanoi 12116, Vietnam
Email: {chi.nguyendinh*, cuong.duongchinh, luong.nguyencong}@phenikaa-uni.edu.vn, nghiapx@mta.edu.vn

***Abstract*—The principal target of this study is to enhance the performance of the 4-level holographic data storage (HDS) systems by using a well-suited modulation code scheme. The HDS system is one of the most promising approaches for storing massive digital data for the next-generation optical data storage systems. However, the performance of retrieving data in the HDS system is severely affected by fatal interference (FI) patterns. The proposed coding design aims to get rid of all FI patterns during the recording procedure with reasonable computational complexity.**

Keywords—multilevel coding, modulation coding, holographic storage, intersymbol interference

I. INTRODUCTION

The information storage systems continue to evolve to satisfy the ever-increasing demand for storing and retrieving a massive amount of process-needed information. Optical data storage systems have the advantages of long media life, high stability, low cost per bit, and "green" storage (i.e., saving energy). Owing to the advantages, optical data storage is suitable for both data centers and end-user storage applications. To further promote the use of optical data storage systems, it is necessary to expand data capacity. Besides of some potential candidates, such as glass data storage, dual-beam super-resolution optical storage, and fluorescent nanocrystal storage technology, the holographic data storage (HDS) is regarded as the most promising storage technology for the next-generation optical data storage systems [1]. For HDS systems, recording and retrieving information are conducted using two-dimensional (2-D) arrays (or pages). For two-level (binary) HDS systems, the user data is first converted into binary bits, and the bits are then translated into light and dark square pixels. They are then packed together into pages. For the recording process, the pixels are impressed onto a signal beam using the page formatting device, which is named spatial light modulator (SLM). The optical signal beam will be interfered with a reference beam inside the holographic material to produce holograms. As a result, the data is recorded by holograms. In order to obtain high storage densities, many holograms are superimposed in the same (or nearly the same) volume by multiplexing techniques [2]. For reading data, the media is illuminated by a suitable reference beam to reconstruct the encoded signal beam. Then a high-quality detector, named as a charge-coupled device (CCD), is used to detect the recorded information. The HDS systems provide unique advantages such as extremely short access time, fast input/output rate, or enormous search capability at a rate exceeding 100 Gb/s [3].

Besides its merits, however, there are several challenges to be faced, such as intersymbol interference (ISI), interpage interference (IPI), blur effect, and misalignment [3]. In particular, the effect of ISI is inherent in high-density storage systems, and the effect becomes extremely severe in multilevel storage systems. Since a single pixel can represent more than 1 bit, which means that a multilevel signal provides more information than a two-level signal, the multilevel HDS systems are new promising candidates for the next generation of ultra-high capacity optical storage systems [4]. Nevertheless, multilevel systems become more error-prone than binary ones. Many solutions have been proposed to overcome these challenges, in both recording and retrieving parts. Some solutions for the retrieving side may be mentioned as 2-D partial response maximum likelihood (2-D PRML) detection, approaches based on the 2-D soft-output Viterbi algorithm(2-D SOVA) [5-6]. The application of advanced coding schemes such as the low-density parity-check code and turbo codes leads to significant improvements in system performance. However, the coding techniques increase complexity in both encoder and decoder.

Modulation code shaping the information sequence to match the characteristics of the physical recording medium is also a coding technique commonly used in storage systems, such as optical/magnetic recording systems and flash memories. Some multilevel modulation codes have been recently investigated for the HDS systems. For instance, Nguyen and Lee [7] proposed a 8/10 modulation code to avoid 2-D interference patterns for the HDS system. A 2/3 modulation code [8] was also proposed for the 4-level HDS system. The authors in [9] constructed a 4-level 4/6 modulation code with a trellis-based error correction capability. Recently, a 14/16 modulation code [10] was also shown. Its performance is degraded compared to the prior arts since it leaves some remaining fatal patterns aiming to get a higher code rate. Furthermore, a primary concern of these encoders is that they have a rather large output array size, which raises the complexity of encoding, detection, and decoding processes. For example, the above-mentioned 2/3 and 4/6 modulation codes must encode even and odd lines separately,

---

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.04-2019.307.

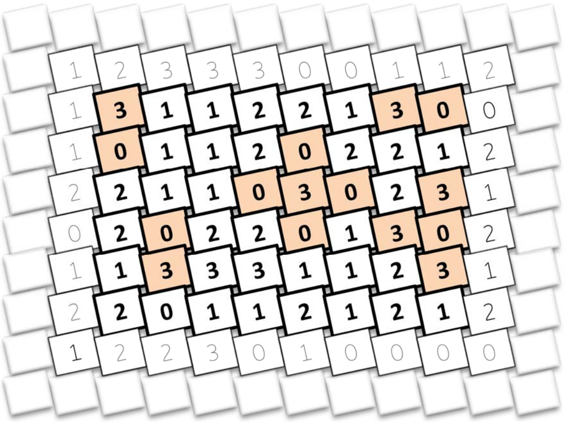
where the structure of codewords in even lines is reversed in the odd lines. The 14/16 modulation codes, whereas used relatively complex mapping tables. In this research, we present a new design of modulation coding, aiming to prevent the fatal patterns for the 4-level HDS systems with a reasonable complexity.

The rest of this paper is organized as follows. The design of the proposed modulation coding is presented in Section II. In Section III, we introduce a 4-level HDS system model for applying the proposed code. Simulation results and discussion are shown in Section IV. Finally, Section V concludes the paper.

II. DESIGN OF 4/6-RATE 4-ARY MODULATION CODE

A. Preliminaries

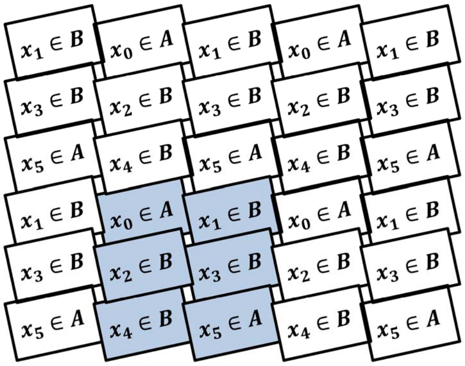
In general, multilevel systems can store more information than binary ones, with the same number of pixels on each page. For *M*-ary systems, one pixel represents for $\log_2 M$ bits. In this study, we consider one pixel represents 2 bits, which means a 4-level HDS system. Let us denote a pair of two bits of {00, 01, 10, 11} as 4-ary symbols {0, 1, 2, 3}. The performance of the HDS systems is significantly degraded as the smallest symbols (0s) are surrounded by the largest symbols (3s), or vice versa. The phenomenon is named as 2-D interference for the multilevel HDS system. Basically, the 2-D interference consists of some fatal interference (FI) patterns, i.e., (0,3) patterns. The FI patterns are illustrated in Fig. 1. To eliminate the effect of 2-D interference, a modulation coding scheme to prevent all the FI patterns is designed in this study.



1	2	3	3	3	0	0	1	1	2
1	3	1	1	2	2	1	3	0	0
1	0	1	1	2	0	2	2	1	2
2	2	1	1	0	3	0	2	3	1
0	2	0	2	2	0	1	3	0	2
1	1	3	3	3	1	1	2	3	1
2	2	0	1	1	2	1	2	1	2
1	2	2	3	0	1	0	0	0	0

Fig. 1. Fatal interference pattern and 2-D interference.

$$\boldsymbol{A} = \{\mathbf{0}, \mathbf{1}, \mathbf{2}, \mathbf{3}\} \qquad \boldsymbol{B} = \{\mathbf{1}, \mathbf{2}\}$$



$x_1 \in B$	$x_0 \in A$	$x_1 \in B$	$x_0 \in A$	$x_1 \in B$
$x_3 \in B$	$x_2 \in B$	$x_3 \in B$	$x_2 \in B$	$x_3 \in B$
$x_5 \in A$	$x_4 \in B$	$x_5 \in A$	$x_4 \in B$	$x_5 \in A$
$x_1 \in B$	$x_0 \in A$	$x_1 \in B$	$x_0 \in A$	$x_1 \in B$
$x_3 \in B$	$x_2 \in B$	$x_3 \in B$	$x_2 \in B$	$x_3 \in B$
$x_5 \in A$	$x_4 \in B$	$x_5 \in A$	$x_4 \in B$	$x_5 \in A$

Fig. 2. The structure of the proposed output codeword (highlight) arranged on a part of the page.

B. Encoding and decoding procedure

The proposed modulation encoder converts a 4-symbol sequence of user data into a 6-symbol codeword of 3 × 2, and the codewords are then sequentially arranged into a page. Let $\boldsymbol{c_6}$ denote an encoder's output array as follows, $\boldsymbol{c_6} \equiv [\boldsymbol{x_0}\ \boldsymbol{x_1};\ \boldsymbol{x_2}\ \boldsymbol{x_3};\ \boldsymbol{x_4}\ \boldsymbol{x_5}]$. Moreover, the set {0,1,2,3} is denoted by $\boldsymbol{A}$, and the set {1,2} is denoted by $\boldsymbol{B}$. The basic idea of the proposed modulation code is to alleviate the whole of FI patterns on a page. To do this, our approach is to limit the appearance of 0 and 3 symbols at the encoder's output. In other words, a well-designed output array of the encoder is to limit the appearance of the largest and smallest symbols, while ensuring a reasonable encoding ratio. The structure of the output codeword of the proposed encoder is shown in Fig. 2. Thereby, the $\boldsymbol{x_0}$ and $\boldsymbol{x_5}$ symbols can take any value in $\boldsymbol{A}$, and the other symbols take values in $\boldsymbol{B}$. The total number of possible codewords at the output of the encoder under this architecture would be 256, i.e., $256 = 4 \times 2 \times 2 \times 2 \times 2 \times 4$. It perfectly matches the total possibilities of the encoder input. The output values of the encoder can be calculated as follows, where a_i denotes the user symbol sequence,

$$\boxed{\begin{aligned} \boldsymbol{x_0} &= a_0 \\ \boldsymbol{x_1} &= 1 + \mathrm{mod}(a_1/2\,, 2) \\ \boldsymbol{x_2} &= 1 + \mathrm{mod}(a_1, 2) \\ \boldsymbol{x_3} &= 1 + \mathrm{mod}(a_2/2\,, 2) \\ \boldsymbol{x_4} &= 1 + \mathrm{mod}(a_2, 2) \\ \boldsymbol{x_5} &= a_3 \end{aligned}}$$

An encoder of the proposed 4/6 modulation code is illustrated in Fig. 3, where mapping the user data symbols of a_1 and a_2 to the codeword symbols is carried out by functions *f*(*a*) and *g*(*a*).

A SOVA detector detects the received signal. The output of the detector is then is sent to the demodulator. The signal is demodulated by quantizing the SOVA output values, and then de-mapping 1:1 to obtain the values $a_0, a_1, a_2,$ and a_3.

III. THE HDS SYSTEM MODEL

A continuous point spread function (PSF) [11] is used to model the HDS channel herein as follows,

$$h(x, y) = \frac{1}{\sigma_b^2} \mathrm{sinc}^2 \left(\frac{x - m_x}{\sigma_b}, \frac{y - m_y}{\sigma_b} \right) \qquad (1)$$

where σ_b is the grade of the blur; m_x and m_y are the *x* and *y* axis misalignments, respectively. A discrete PSF is defined as follows,

$$h[p, q] = \int_{q-1/2}^{q+1/2} \int_{p-1/2}^{p+1/2} \frac{1}{\sigma_b^2} \mathrm{sinc}^2 \left(\frac{x - m_x}{\sigma_b}, \frac{y - m_y}{\sigma_b} \right) dxdy \quad (2)$$

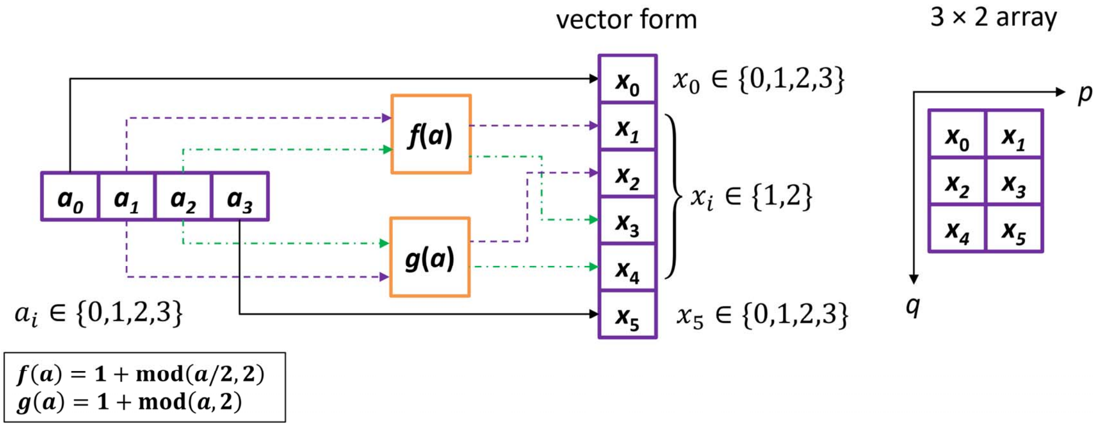


Fig. 3. An encoder of the proposed 4/6 modulation code

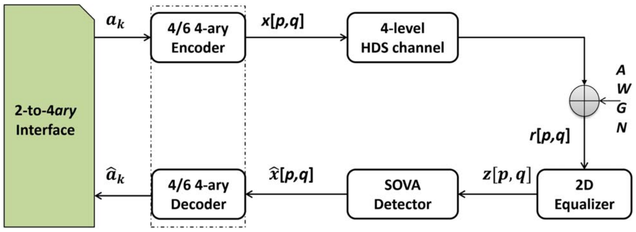


Fig. 4. An system model.

where the linear fill factor of the CCD is equal to 1, and we consider the range of discrete PSF is 5 × 5.

A block diagram of the system model is illustrated in Fig. 4. The input data $a_k, a_k \in \boldsymbol{A}$, are converted into a 2-D array by the proposed modulation encoder, then the signal is passed through the HDS channel. The received signal $r[p,q]$ can be expressed by the following,

$$r[p,q] = x[p,q] \otimes h[p,q] + n[p,q] \qquad (3)$$

where $x[p,q]$ is the output of the proposed encoder, the operator $\otimes$ represents a 2-D convolution, and $n[p,q]$ is an AWGN element. An equalizer is used to equalize the channel output to a partial response (PR) target of [0 1 0; 1 3 1; 0 1 0] in this research. The equalized signal is detected by the SOVA detector [12], and then decoded by the demodulator. We set σ_w^2 as the noise power, and the system signal-to-noise ratio (SNR) is defined as,

$$SNR = 10\log_{10}\frac{1}{\sigma_w^2} \qquad (4)$$

IV. SIMULATION RESULTS AND DISCUSSION

We used 1000 pages, page of 1024 × 1024 pixels, for simulation in this work. We analyze the histograms of the received intensities of the proposed modulation code. Fig. 5 shows the histograms of the raw data and modulated data. As can be observed, the symbol distributions of the raw data are relatively uniform. Those of the proposed code, whereas, are distinct, wherein the distributions of 1 and 2 symbols are superior to 0 and 3 symbols. This indicates that the encoding

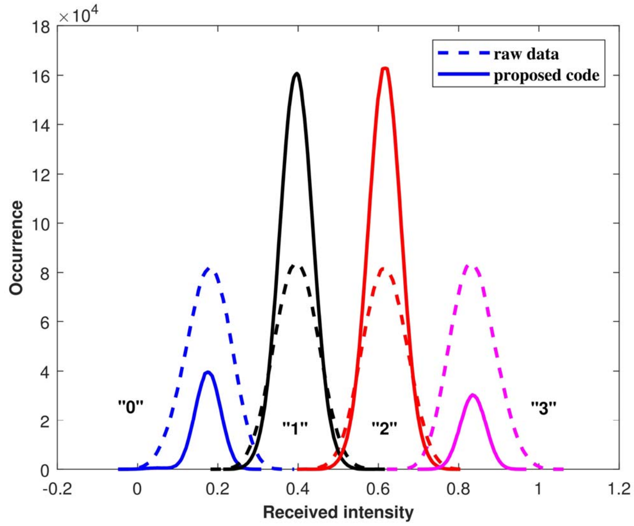


Fig. 5. Distribution of the symbols.

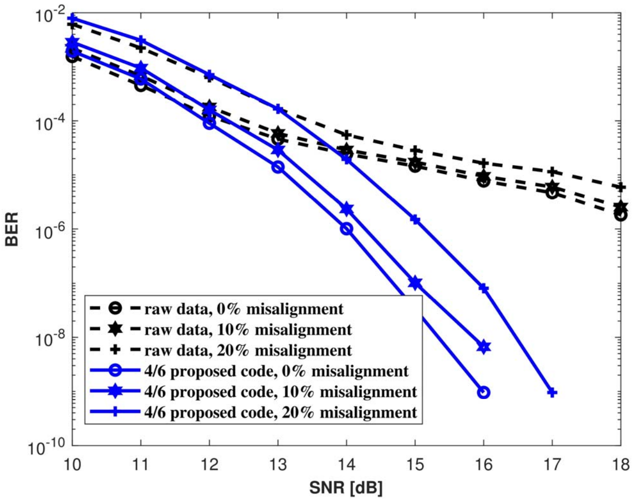


Fig. 6. BER performance at the SOVA detector output.

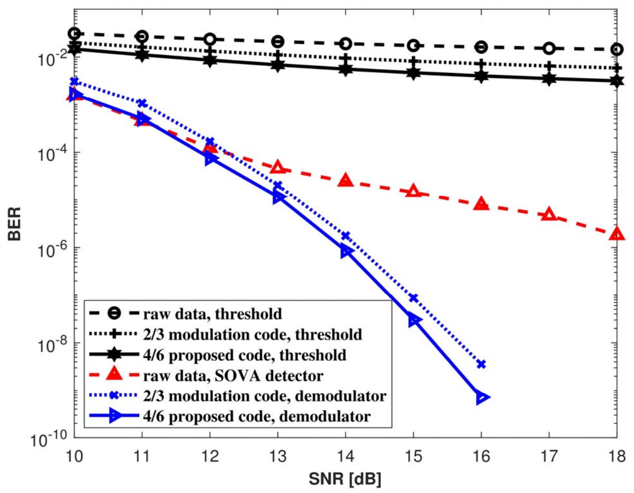


Fig. 7. BER performance comparison.

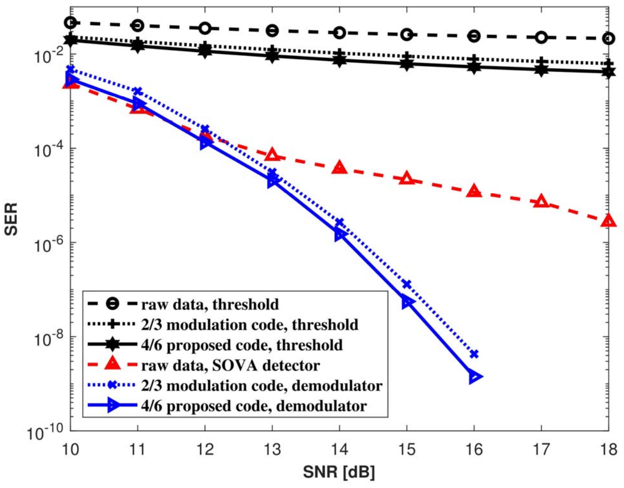


Fig. 8. SER performance comparison.

procedure is accomplished as desired. The level of overlap between symbols reveals the effect of 2-D ISI in the systems. The overlap area is greatly reduced when using the proposed code. As a result, the proposed code is efficient for mitigating the 2-D interference for the HDS system.

Fig. 6 shows the bit-error-rate (BER) performance comparison between the proposed code and the raw data under the misalignment effect after the SOVA detection. First, it is observed that the performance of the proposed code is weak at low SNR (< 12 dB herein). This result is entirely consistent with the results reported in [8]. It can be explained because the detector used is for the binary signal only [8, 12]. Better performance can be expected if a new detector is well designed for the multilevel signals for HDS systems. As shown in the figure, the proposed code provides an outstanding performance compared to the raw data. For instance, at SNR of 15 dB, the proposed code obtains about 10^{-8} BER, whereas the raw data can only reach approximately 10^{-5} BER under 0% misalignment. Under the effect of 20% misalignment, for 17 dB, the proposed code and the raw data deliver about 10^{-9} BER and 10^{-5} BER, respectively.

The BER and symbol-error-rate (SER) performances of the proposed code compared to the raw data and the 2/3 modulation code are shown in Figs.7 and 8, respectively. The simulation is carried out at 0% misalignment and the blur grade of 1.0. Threshold detection is also deployed for the channel output. As shown in both figures, the performance of the proposed code is the most effective, followed by those of the 2/3 modulation code and the random signal, respectively. Results in Figs. 7 and 8 show that the performance of the proposed code is better than the 2/3 modulation code. This can be explained by the difference in the designed structure of the proposed code's encoder output. According to [8], while keeping 0 and 3 symbols with low distributions, the distribution difference of the 1 and 2 symbols is quite high. In other words, the design reduces the balance across the page. The proposed code, as shown in Fig. 5, generates a quite-similar distribution for 1 and 2 symbols. As a result, the full page's intensity becomes more balanced, thus minimizing interference between symbols.

V. CONCLUSION

In this study, we present a simple but effective technique for designing a 4/6 4-ary modulation code. Accordingly, the proposed code limits the occurrence of 0 and 3 symbols, which constitute the fatal interference pattern, thereby preventing the 2-D interference. At the same time, the code makes the page's intensity more balanced. The simulation results show that the proposed code delivers better performances than the raw signal, even at 20% misalignment. The designed code with simpler encoding and decoding procedures also outperforms the previously suggested 2/3 modulation code.

ACKNOWLEDGMENT

The authors would like to express special thanks to anonymous reviewers for their suggestions and comments.

REFERENCES

[1] K. Kim, S. Kim, G. Koo, M. Seo, and S. Kim, “Decision feedback equalizer for holographic data storage,” Appl. Opt. vol. 57, pp. 4056-4066, May 2018.

[2] K. Curtis, L. Dhar, A. Hill, W. Wilson, and M. Ayres, Holographic Data Storage: From Theory to Practical Systems. John Wiley & Sons: Chichester, UK, 2010.

[3] “light fantastic,” Electron. Lett., vol. 52, no. 17, pp. 1422-1422, Aug. 2016.

[4] T. Nobukawa and T. Nomura, “Multilevel recording of complex amplitude data pages in a holographic data storage system using digital holography,” Opt. Express, vol. 24, pp. 21001-21011, Sep. 2016.

[5] G. Kong and S. Choi, “An ant colony optimization aided two-dimensional soft output Viterbi algorithm detection scheme for holographic data storage systems,” Jpn. J. Appl. Phys., vol. 58, pp. SKKD03-1-SKKD03-5, Sep. 2019

[6] C. D. Nguyen and J. Lee, “Iterative detection supported by extrinsic information for holographic data storage systems,” Electron. Lett., vol. 51, pp. 1857-1859, Nov. 2015.

[7] C. D. Nguyen and J. Lee, “8/10 two-dimensional modulation code for holographic data storage systems,” Electron. Lett., vol. 52, pp. 710-712, Apr. 2016.

[8] S. Y. Kim and J. Lee, “A Simple 2/3 Modulation Code for Multi-Level Holographic Data Storage,” Jpn. J. Appl. Phys., vol. 52, pp. 09LE04, Sep. 2013.

[9] S. Jeong and J. Lee, “4-Level 4/6 Modulation Code with Trellis Encoding on Holographic Data Storage,” J. I. E. I. E., vol. 54, pp. 853-858, Sep. 2015. [in Korean]

[10] J. Kim, S. Jeong and J. Lee, “4-ary 14/16 modulation code for reducing two-dimensional inter-symbol interference,” IET Comm., vol. 14, pp. 1335-1339, May 2020.

[11] M. Keskinoz and B.V.K.V. Kumar “Efficient modeling of volume holographic storage channels (VHSC),” Proc. SPIE, vol. 4090, pp. 205-210, Sep. 2000

[12] C. D. Nguyen and J. Lee, “Iterative detection supported by extrinsic information for holographic data storage systems,” Electron. Lett., vol. 51, no. 23, pp. 1857-1859, Nov. 2015.