Improved Error Correction for Visible Light Communications

by

Elie Ngomseu Mambou

A dissertation submitted to the Faculty of Engineering
in partial fulfillment of the requirements for the degree of

Master of Science

in

Electrical and Computer Engineering

at the

McGill University

Supervisor
Prof. Warren J. Gross

November 4, 2020

Abstract

The number of mobile device subscribers worldwide grows at an exponential rate every year. This generates a congestion in the radio frequency (RF) bandwidth that may create unreliable communication. Therefore alternative means of communications should be investigated to sustain the RF channel. Visible light communication (VLC) which is a subset of optical wireless communication (OWC) could be a potential complement to RF channels. VLC is defined as a short-range optical wireless communication through light-emitting diode (LED) source. This technology could decongest the RF channel in the sense that VLC could cover indoor wireless communication where short distance transmissions are widely used while RF channels apply mainly for outdoor communication. There are many advantages about the VLC channel including a low cost of deployment, higher security, lower interference, bigger and unregulated bandwidth compared to RF systems. However, the VLC field presents several challenges such as light beam flickering and dimming. In other words, how to achieve reliable communication that supports different light intensities while avoiding light blinking perceptible to human eyes? This thesis aims to bring some elements to address those challenges. In VLC channels, the flickering mitigation is most often imposed by the use of DC-free run-length limited (RLL) codes such 1B2B, 4B6B and 8B10B codes; due to their low error correction performance, a forward error correction (FEC) code is concatenated with RLL codes to improve the channel reliability. In this thesis, we present a new class of 4B6B codes, the most efficient and reliable in terms of error correction performance. This code was obtained from a permutation of the conventional 4B6B code based on some distance constraints. Furthermore, a flicker-free FEC coding using Knuth's balancing algorithm is implemented; this is a non-RLL technique which achieves flicker-free communication using the famous Knuth's algorithm introduced by Donald Knuth in 1986. Finally, we provide an efficient dimming scheme based on non-DC free codes which does not make use of puncturing, interleaving, compensation symbols (CS) insertion as some state-of-the-art (SOA) methods, therefore reduces the decoding complexity while improving the spectral efficiency.

Abrégé

Le nombre d'abonnés à des appareils mobiles dans le monde croît à un rythme exponentiel chaque année; cela génère une congestion dans la bande passante des radiofréquences (RF) qui peut conduire une communication peu fiable. Par conséquent, d'autres moyens de communication devraient être étudiés pour renforcer le canal RF. La communication par lumière visible (VLC) qui est un sous-ensemble de la communication optique sans fil (OWC) pourrait être un complément potentiel aux canaux RF. VLC est défini comme une communication sans fil optique à courte portée via une source à diodes électroluminescentes (LED). Cette technologie peut décongestionner le canal RF de plusieurs façons, par exemple, elle pourrait être utilisée pour la communication sans fil intérieure où les transmissions à courte distance sont prédominantes tandis que les canaux RF seraient dédiés aux communications extérieures. Le canal VLC présente de nombreux avantages, notamment un faible coût de déploiement, une sécurité plus élevée, des interférences plus faibles, une bande passante plus grande et non rgulée par rapport aux systèmes RF. Cependant, le domaine VLC présente plusieurs défis tels que le scintillement et la gradation du faisceau lumineux. En d'autres termes, comment parvenir à une communication fiable supportant différentes intensités lumineuses tout en évitant le clignotement de la lumière perceptible par les yeux humains? Cette thèse vise à apporter quelques éléments pour relever ces défis. Dans les canaux VLC, l'atténuation du scintillement est le plus souvent imposée par l'utilisation de codes à longueur limitée (RLL) balancés tels que les codes 1B2B, 4B6B et 8B10B; en raison de leurs faibles performances de correction d'erreur, un code correcteur d'erreur (FEC) est concaténé avec des codes RLL pour améliorer la fiabilité du canal. Dans cette thèse, nous présentons une nouvelle classe de code 4B6B, les plus efficaces et fiables en termes de performances de correction d'erreurs; ce code a été obtenu à partir d'une permutation du code 4B6B conventionnel et en fixant certaines propriétés de distance comme contraintes. En outre, nous fournissons un schéma de gradation efficace basé sur des codes à poids variable et non balancés qui n'utilisent pas l'insertion de symboles de compensation (CS), de poinçonnage, d'entrelacement, comme certaines méthodes de l'état de l'art (SOA), réduisent donc la complexité tout en conser-

vant une efficacité spectrale plus élevée. Enfin, un codage FEC sans scintillement utilisant l'algorithme d'équilibrage de Knuth est implmenté; il s'agit d'une technique sans code RLL qui permet une communication sans scintillement en utilisant le célèbre algorithme de Knuth créé par Donald Knuth en 1986.

Acknowledgements

I would like to express my gratitude and appreciation to Prof. Warren Gross for his academic support and guidance.

I want to extend my words of appreciation to the members of the integrated systems for information processing (ISIP) lab at McGill University for their useful discussion and collaboration that made a great impact on this current work. I would like to particularly thank Ali Hashemi for introducing me to ISIP group; Thibaud Tonnellier for his collaboration and C++ proficiency that brought a nice flavour to my research topic; Nghia Doan for his machine learning in channel coding knowledge; Harsh Aurora for his great neighbourhood and discussion in the lab. My sincere gratitude goes also to Prof. Steve Hranilovic for reviewing this thesis and providing valuable comments. Moreover, I need to address special mentions to the following collaborators: Furkan Ercan, Harsh Aurora, Arash Ardakani, Jerry Li, Adam Cavatassi, Mohsin Abbas and Amir Ardakani.

Lastly but not least, I want to acknowledge my parents Augustin and Yolande, my siblings Herve, Fabrice, Corine and Monique, my friends, my beloved Tatiana and my little child Kimia, for encouragements and trust.

Contents

Publications

(1) E. Ngomseu Mambou, T. Tonnellier, S. A. Hashemi and W. J. Gross, “Efficient Flicker-Free FEC Codes Using Knuth’s Balancing Algorithm for VLC,” *2019 IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, HI, USA, 2019, pp. 1-6.

(2) E. Ngomseu Mambou, T. Tonnellier and W. J. Gross, “Improved DC-free run-length limited 4B6B codes for concatenated schemes,” to be submitted *IEEE Photonics Letters*, 2020.

(3) E. Ngomseu Mambou, T. Tonnellier and W. J. Gross, “Efficient Dimming Scheme based on Non-DC Free RLL Codes for VLC,” to be submitted *IEEE Photonics Letters*, 2020.

(4) N. Doan, S. A. Hashemi, E. N. Mambou, T. Tonnellier and W. J. Gross, “Neural Belief Propagation Decoding of CRC-Polar Concatenated Codes,” *2019 IEEE International Conference on Communications (ICC)*, Shanghai, China, 2019, pp. 1-6.

(5) W. J. Gross, N. Doan, E. Ngomseu Mambou and S. Ali Hashemi, “Deep Learning Techniques for Decoding Polar Codes,” in *Machine Learning for Future Wireless Communications*, John Wiley & Sons, Ltd, 2020, ch. 15, pp. 287 –301.

(1) is covered in Chapter 4; and (2) in Chapter 3.

List of Figures

List of Tables

List of Abbreviations

BEC	Binary erasure channel
BSC	Binary symmetric channel
CD	Compact disc
DC	Direct current
DSV	Digital sum variation
ECC	Error correcting codes
FEC	Forward error correction
FL	Fixed length prefix method
GF	Galois field
GSM	Global system for mobile communications
IC	Integrated circuit
KA	Knuth's balancing algorithm
OWC	Optical wireless communications
RLL	Run length limited
VL	Variable length prefix method
VLC	Visible light communication
VLSI	Very-large scale integration

1 Introduction

1.1 Overview

The genesis of visible light communication dates back to 1880 when Alexander Graham Bell and his assistant Charles Sumner invented the photophone at Bell's laboratory, Washington, D.C.; the photophone been a telecommunication medium that transmits speech through modulated sunlight over hundreds of meters [7]. In 1948, the world of communication technologies embarked on a new journey with the publication of Shannon's article entitled *"A mathematical theory of communication"* [8]; this established a possibility of achieving a noiseless communication. Since then, research have been conducted with the aim of realising an error-free transmission of information from a sender to a receiver over a medium. VLC works have resumed extensively in beginning of the 21st century following the apparition of the 3G wireless network. This technology has drawn attention from several research structures such as Nakagawa Laboratory at Keio University since 2000 [9–11]; the communication and information technology research center (CICTR) at Pennsylvania state University that proposed a cascaded power line communication PLC-VLC channel for broadband access indoor applications in 2006 [12]; Siemens and Heinrich Hertz institute in Berlin implemented a 500 Mbits/s transmission using a single white LED over 5 meters in 2010 [13]; the company LVX was the first to commercialize the technology in November 2010 to the city St. Cloud, Minnesota [14]; in July 2011, the term *LiFi* (light fidelity) by analogy to *WiFi* was introduced by Harald Haas at TED talk in Edinburgh where he demonstrated a high-definition (HD) video transmission through LED lamp [15]. Many other works have framed interesting directions in the VLC field [16–19]. With the huge interest brought by the VLC technology, the transmission in this field was regularised by the Institute of Electrical and Electronics Engineers (IEEE) working group IEEE 802.15.7; this standard organised the physical layer (PHY) of the VLC channel into PHY I, PHY II and PHY III with specific modulation and coding schemes for each of these layers [1]. PHY I and II use on-off keying (OOK) and variable pulse position modulation (VPPM) as modulators and operates at 11.67–266.6 kbit/s and 1.25–96 Mbit/s respectively. In PHY III, the colour shift keying (CSK) modulation is recommended as multiple sources are required and operates at 12–96 Mbit/s. Over the past 10 years, several works have been dedicated to achieve a more reliable VLC channel with the use of modern codes such as turbo codes [20], low-density parity check (LDPC) codes [21]; and recently polar codes were introduced by Erdal Arıkan as the first class of FEC codes to asymptotically achieve the channel capacity of binary memoryless channel (BMC) with infinite code length [22].

1.2 Summary of Contributions

below is the summary of contributions achieved within the scope of this thesis.

Most Efficient DC-free 4B6B RLL Code

A new class of 4B6B code is introduced as the most efficient class of DC-free 4B6B RLL codes in terms of error correction capabilities for a serially concatenated architecture. This was obtained from a permutation of the 16 codewords from the conventional 4B6B code based on some established metrics. The designed 4B6B code presents some similar properties with the conventional one such as the spectral efficiency of 0.67 bit/s/Hz, the maximum run-length of 4. The run-length refers to the number of consecutive zeros or ones within a codeword. However, the proposed 4B6B code outer-performs the conventional one as well as most state-of-the-art schemes in terms of error correction performance. This achievement holds for any linear FEC code used as outer code in concatenation with the proposed 4B6B code. The proposed scheme can also be extended to other DC-free RLL codes such as 6B8B or 8B10B.

Efficient Flicker-free VLC Channel using Knuth's Balancing Algorithm

In VLC, the flickering is mitigated mostly by the use of DC-free RLL codes, which uses look-up tables and attenuate the spectral efficiency of the channel. An efficient construction of flicker-free channel is presented using the Knuth's balancing algorithm in replacement of the DC-free RLL codes. This technique does not make use of look-up tables; the low redundancy of the Knuth's algorithm is exploited to provide high transmission data rates at a reduced complexity. Knuth's balancing algorithm consists of flipping bits within a binary sequence until the balance state is reached; that is the state where a sequence has the same number of zeros and ones. The proposed scheme works as follows: the payload is encoded by the outer FEC code then balanced via Knuth's algorithm, the generated prefix is encoded and balanced again to obtain an overall DC-free transmitted codeword. The proposed scheme achieves a better performance than most state-of-the-art ones.

Efficient Dimming Technique based on Variable Weight RLL Code
Conventional dimming techniques are making use of puncturing, interleaving and compensation symbols (CS) insertion which increases the decoding complexity, attenuate the spectral efficiency and degrade the channel. New families of non-DC free RLL codes with variable weights 1B3B, 4B7B, 6B9B and 8B11B were designed based on certain properties for different dimming ratios. This enables a wider dimming range with a better channel reliability compared to the conventional 4B6B and some existing methods; and less redundancy which implies less power dissipation. Usually, the best channel performance is found when the dimming ratio of 50% imposed by the DC-free RLL codes is used. However the proposed 4B7B code at dimming ratios of 28.5/71.4% outer-performs the original 4B6B.

1.3 Thesis Outline

The rest of this thesis is structured as follows: Chapter 2 provides the necessary preliminary knowledge related to VLC channel coding, this includes modulation techniques, transmission blocks, impairment sources, FEC codes used, etc., as per VLC standards. In Chapter 3, a new class of DC-free RLL codes with the case study of 4B6B was introduced; this 4B6B code gives the most reliable VLC channel in terms of error correction capabilities for a serially concatenated architecture. For the span of Chapter 4, an efficient construction of flicker-free VLC channel is presented where Knuth's balancing algorithm is applied on the encoded payload and the resulting prefix is likely protected. This study is sustained by various analysis including the redundancy study, RLL characteristics, error correction performance and complexity report. Chapter 5 proposes an efficient dimming scheme based on variable weight non-DC free RLL codes; new families of RLL codes are introduced with different dimming ratios; the code construction is explained and simulations with comparisons to some existing methods are provided as well. Finally, the thesis is concluded in Chapter 6 with a summary of each chapter followed by some future research directions to improve the present work.

2 Literature Review

In this section, we investigate some background material related to coding theory in visible light communication.

2.1 Overview on Constraint Coding

The constrained coding aims to build a channel that imposes some constraints to sequence codes for a reliable transmission of codewords [23]. Fig. 2.1 shows the block diagram of a constrained channel. It can be observed that a concatenation of error correction codes (ECC) and constrained codes (CS) is used in this type of channel; this is because the constrained coding present poor error correction capabilities and therefore a forward error correction (FEC) code is usually used to improve the performance of the channel.

FEC and CS codes are both channel coding techniques but fundamentally different; FEC coding consists of improving the accuracy of decoded sequences in other words, it aims to lower to probability of errors occurring during transmission of data; mostly linear codes are used for this task. On the other hand, CS coding encodes data such that it has properties that are closely matching the time or frequency-domain constraints of the channel; non-linear codes are commonly employed based on encoding and decoding through hand-crafted look-up tables. They exist several type of constraints depending on applications and the most common are constant-weight and run-length limited; constant-weight constraint imposes a fixed weight on the transmitted codeword; the balanced code

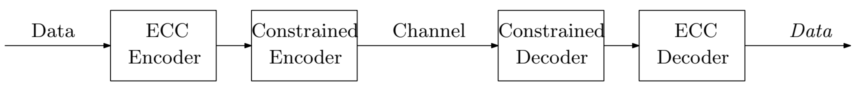


Figure 2.1: Block diagram of a constrained channel

constraint is a special case of the constant-weight one where the number of 1's and 0's of the transmitted code must be the same; the run-length limited (RLL) codes impose some constraints on the runs of 0's and 1's; for example the (d, k) RLL code constraint sets a minimum of d and the maximum of k number of consecutive 0's; the DC-free RLL codes αBβB maps each α coded bits into a codeword of size β via look-up tables.

Constrained coding has found applications in several fields such as wireless energy harvesting [24, 25], optical and magnetic recording [26, 27], DNA-based storage [28, 29], power line communication [30, 31], fiber optics [32, 33], solid state memories [34], visible light communication [35,36] and etc. Furthermore the use of deep learning has significantly improve state-of-the-art techniques in constrained coding [37, 38].

2.2 Overview on Visible light communication

VLC is a subset of optical wireless communications, where the data transmission and reception medium uses visible light spectrum with frequency between 350 and 750 Terahertz equivalent to wavelengths 380-780 nano-meters. The LED current is easily modulated and can transmit signals at around 500 Mbit/s. This is done through *intensity modulation / direct detection (IM/DD)* which is the principle of modulating the intensity of a light source in optical wireless communications and detecting intensity fluctuations at the receiver. In low data rate communications, an Ethernet speed of 10 Mbit/s is achievable at a distance of 1-2 kilometres [39]. In VLC technology, LEDs are usually used to provide illumination and optical wireless access. Illumination from LEDs provides relatively high radiant flux and low modulation bandwidths. Considering typical illumination environment of around 400 lx (unit of luminance, measuring luminous flux per unit area), a SISO VLC channel operates at high SNR and low bandwidth [44]. The scope of this thesis is mainly limited to SISO VLC channels considering IM/DD.

As shown in Fig. 2.2, the VLC and radio frequency (RF) belong to the big umbrella of communication technologies. VLC is a short range data transmission through modulated light beam. In VLC, light emitting diodes (LEDs) are most used as their current intensity is easily modulated. Optical wireless communications can play a major role into decongesting RF systems and could be part of further communication standards as presented in Fig. 2.3 (this image was inspired from [40]).

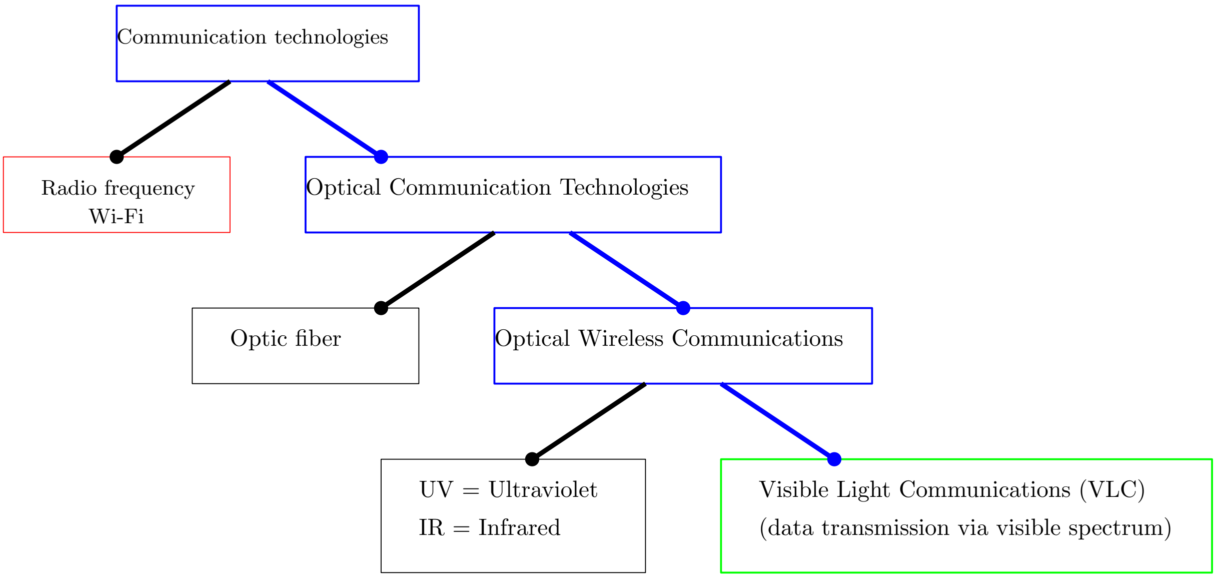


Figure 2.2: VLC classification

It can be observed that the VLC could be used for indoor communication where the range of transmission is relatively short. Due to some limitations on Ultraviolet (UV) or Infrared (IR) such as monochromatic spectrum, small bandwidth or low date rate, VLC tends to be one of the most performant optical wireless communication tool. The visible light spectrum ranges from 380 to 780 THz.This provides ten thousand times bigger and unregulated bandwidth with lower interference compared to RF systems. Furthermore, the cost for deployment of VLC is affordable as transceivers are made of LEDs and photodetectors (PD), it provides higher security for indoor applications in the sense that lighting is bounded by walls and can easily be delimited within a specific area.

The term *Li-Fi* by analogy to *Wi-Fi* was first proposed by Harald Haas [41], where he stated that Li-Fi could replaced indoor Wi-Fi especially in area like hospitals where the interference could perturb X-rays or other equipments. Li-Fi could also replace Wi-Fi for indoor airplane to avoid interference with cockpit communication. Li-Fi could also improve several fields such as underwater communication, robotics, intelligent LED traffic light systems, camera surveillance, face recognition, vehicle-to-vehicle communication and etc. In Li-Fi technology network access points are replaced by smart LED lighting.

However, VLC technology presents some challenges from its physical layer aspect, that are summarised into flickering, dimming and communication throughput. To avoid human visual perception of light flickering, the frequency of the clock rate should greater than

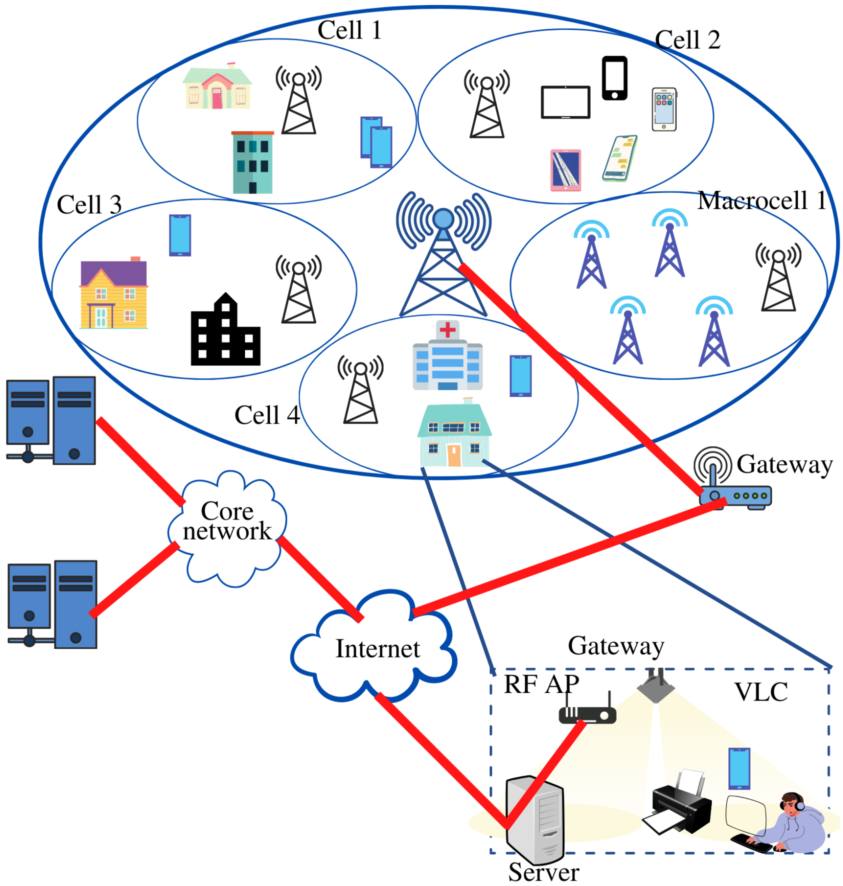


Figure 2.3: 5G cellular network architecture

200 Hz. This corresponds to the maximum flickering time period (MFTP) of 5 ms, that is the maximum duration that the light intensity can vary without being perceptible by the human eye [42]. The date rate gets reduced under dimming condition, this attenuates the transmission of data and save power. This average power needs compensation due to dimming.

2.3 Introduction to Coding Theory in VLC

The basic transmission architecture in VLC is presented in Fig. 2.4. It consists of a transmitter and a receiver as most communication channel. The transmitter encodes and modulates the payload then send the outcome through LED; at the receiver, the signal is captured by a photodetector (PD) and then an altered version of the signal is obtained

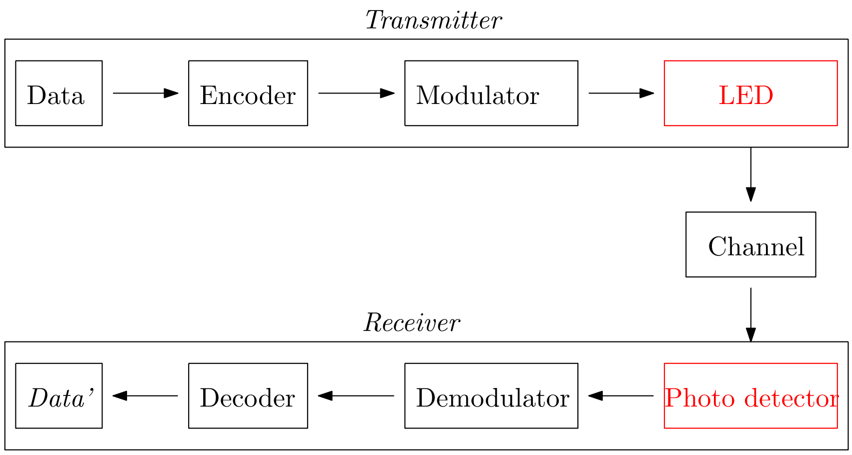


Figure 2.4: Basic transmission architecture in VLC

Figure 2.5: Line-of-sight and non line-of-sight VLC link

after demodulation and decoding.

2.3.1 Types of VLC channels

Single input - single output (SISO) and multiple input - multiple outputs (MIMO) are the two main type of VLC channels encountered in practice.

2.3.1.1 SISO VLC channel

The SISO channel is subdivided into the line-of-sight (LOS) and non line-of-sight (NLOS) links as shown in Fig. 2.5.

The difference between the two links is that, the LOS VLC link presents a direct line connection between the LED and the PD whereas in the NLOS link, there is an obstacle between the sender and the receiver which makes the signal to be attenuated due to path

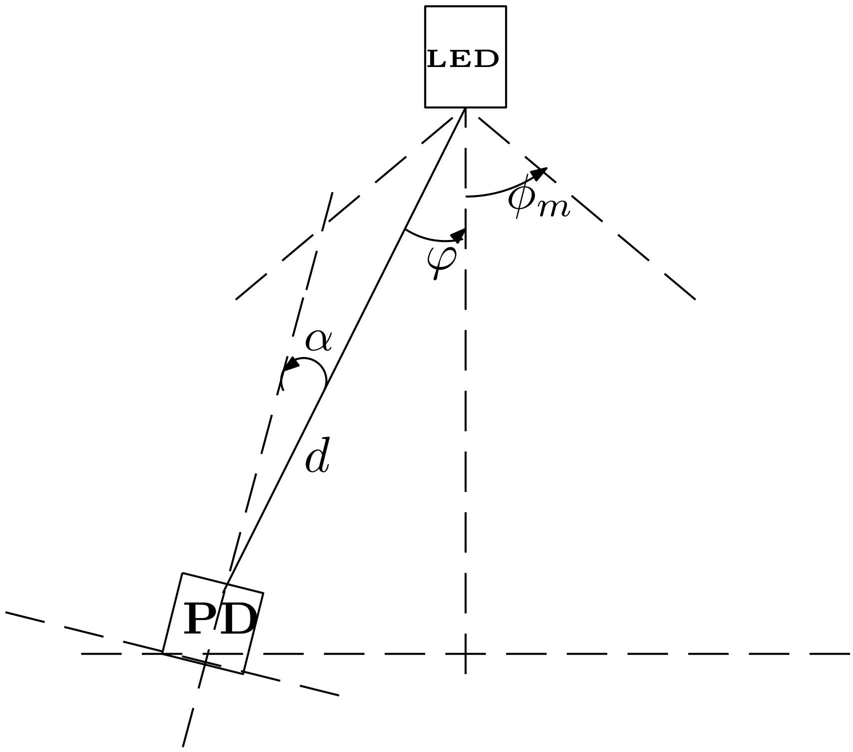


Figure 2.6: non-direct LOS (ndLOS) VLC link

deviation. The received signal $\boldsymbol{r}_i$ from the i^{th} LED is given as follows:

$$\boldsymbol{r}_i = \mathbf{H}\boldsymbol{s}_i + \boldsymbol{w}_i, \tag{2.1}$$

where $\boldsymbol{s}_i$ is the transmitted signal, $\mathbf{H}$ the channel response and $\boldsymbol{w}_i$ the channel noise. The capacity of a SISO VLC channel is obtained as:

$$C_{SISO} = \log_2 \left(1 + \frac{g^2 P_t}{\sigma^2 B}\right), \tag{2.2}$$

P_t represents the transmitted power independent from illumination, B the bandwidth, σ^2 the overall noise variance, g the channel gain. The term $\frac{g^2 P_t}{\sigma^2 B}$ denotes the SNR of the channel.

There are two subcategories of LOS VLC link namely direct LOS (dLOS) and non-direct LOS (ndLOS) as shown in Fig. 2.6. In the case of dLOS, there is a perpendicular line to both LED and PD planes, $\varphi = 0$ and for ndLOS, $\varphi \neq 0$. The channel gain of the

LOS link is estimated as follows [43]:

$$g(LOS) = \Big[\frac{(\xi+1)A}{2\pi d^2}\Big]\cos^{\xi}(\varphi)T_f(\alpha)g(\alpha)\cos(\alpha), \tag{2.3}$$

such that $0 \leq \varphi \leq \phi_m$, $T_f(\alpha)$ denotes the transmission filter, $g(\alpha)$ is the concentration gain, d represents the minimum distance between the LED and the PD. $g(LOS) = 0$ for $\varphi \geq \phi_m$. α, φ, ϕ_m, and d are indicated on Fig. 2.6. The Lambertian order, $\xi = -\ln 2/\ln(\Theta_{0.5})$, where $\Theta_{0.5}$ is the angle representing half the received power. The radiant intensity $\kappa(\varphi)$ of the Lambertian transmitter [44] is given by $\kappa(\varphi) = \frac{\xi+1}{2\pi}\cos^{\xi}(\varphi)$.

The average optical power generated by a single LED is expressed

$$\mathbf{P}_{LED} = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\boldsymbol{p}_{LED}(t)dt \tag{2.4}$$

And the received power in the case of LOS VLC equals

$$\mathbf{P}_{PD} = \mathbf{P}_{LED} \times g(LOS). \tag{2.5}$$

For the NLOS VLC link [45,46], the channel response is derived as

$$\mathbf{H}_{NLOS} = \sum_{k=0}^{\infty} h^{(k)}, \tag{2.6}$$

where $h^{(k)}$ denotes the impulse response of rays undergoing the k^{th} path. The received signal for the NLOS link $\boldsymbol{r}_i$ is as follows:

$$\boldsymbol{r}_i = \mathbf{H}_{NLOS}\boldsymbol{s}_i + \boldsymbol{w}_i = \rho\mathbf{H}_{LOS}\boldsymbol{s}_i + \boldsymbol{w}_i, \tag{2.7}$$

with ρ being the coefficient characterising the NLOS link.

2.3.1.2 MIMO VLC channel

In general, many VLC channels are used to increase the transmitted power signal. Fig. 2.7 shows the configuration of a multi-users VLC channels where N LEDs are used as transmitters and M PDs are receivers [39].

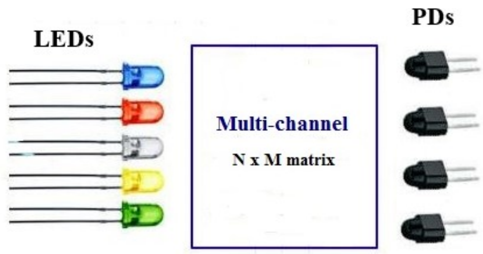


Figure 2.7: Multi-users VLC channels

The channel response of such a MIMO link is estimated as

$$\mathbf{H}_{mult} = \begin{bmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,n} \\ h_{2,1} & h_{2,2} & \dots & h_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{m,1} & h_{m,2} & \dots & h_{m,n} \end{bmatrix}. \tag{2.8}$$

$h_{i,i}$ denotes the front-end gain between the i^{th} LED and corresponding PD. In case that no crosstalk exists in the channel, $\mathbf{H}_{mult}$ becomes a diagonal matrix.

The capacity of a MIMO VLC channel is given by

$$C_{mimo} = \Gamma C_{SISO} \tag{2.9}$$

where $\Gamma = minimum(n, m)$.

In case of RGB-LEDs (multi-wavelength), the transfer matrix $\mathbf{H}_{3x3}$ is obtained as

$$\mathbf{H}_{3\times3} = \begin{bmatrix} h_{r,r} & h_{r,g} & h_{r,b} \\ h_{g,r} & h_{g,g} & h_{g,b} \\ h_{b,r} & h_{b,g} & h_{b,b} \end{bmatrix} \tag{2.10}$$

For RGBA-LEDs which are RGB-LEDs augmented with an amber chip to compensate for white light [47], the bandwidth is wider and the transfer matrix is expressed as

$$\mathbf{H}_{4\times 4} = \begin{bmatrix} h_{r,r} & h_{r,g} & h_{r,b} & h_{r,a} \\ h_{g,r} & h_{g,g} & h_{g,b} & h_{g,a} \\ h_{b,r} & h_{b,g} & h_{b,b} & h_{b,a} \\ h_{a,r} & h_{a,g} & h_{a,b} & h_{a,a} \end{bmatrix} \tag{2.11}$$

subscripts r, g, b and a represent colours red, green, blue and amber respectively.

2.3.1.3 Noise in VLC channels

Various factors can influence the VLC system such the sunlight, fluorescent or incandescent lighting. The receiver also produces some noise due to the contact with photons. This noise can be categorised into shot and thermal noise. The shot noise is generated from the random absorption of photons into electron holes while the thermal noise arises from the motion charge carriers. The shot noise is also called Poisson noise as it follows a Poisson distribution. The variance of the shot noise, σ_{shot} is composed by current fluctuations of light and dark photons, and it is derived as follows [39, 48]:

$$\sigma^2_{shot} = \sigma^2_{light} + \sigma^2_{dark} = B[2qI_{bg}\zeta + 2qP_r\gamma R_r], \tag{2.12}$$

where q is the electric charge, B the bandwidth, R_r the responsivity, P_r the average received power, I_{bg} the background current, ζ a noise bandwidth factor.

And the thermal noise variance [39] is computed as:

$$\sigma^2_{th} = \frac{8\pi K T_k}{G}\eta A\zeta B^2 + \frac{16\pi K T_k \Gamma}{g_m}\eta^2 A^2 \zeta_{th} B^3. \tag{2.13}$$

K refers to the Boltzmann constant, T_k the absolute temperature, A the detector area, η the channel noise factor of the field effect transistor, G the open loop voltage.

The total noise variance is given by the sum of variances for the shot, the thermal noise and the interference, $\sigma^2_t = \sigma^2_{shot} + \sigma^2_{th} + \sigma^2_{other}$, where σ^2_{other} is the variance from the noise generated by interference which includes external sources of light (induced and background noise).

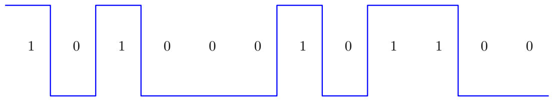


Figure 2.8: NRZ OOK

In VLC systems, the signal-to-noise ratio (SNR) is assumed to be white and Gaussian [48]. The SNR in VLC channel was derived in [10] as:

$$SNR = \frac{\left[\eta s(t) \otimes h(t)\right]^2}{\sigma_n^2} = \frac{\mathbf{R}_r^2\mathbf{P}_r^2}{\sigma_{shot}^2 + \sigma_{th}^2 + \sigma_{other}^2} \tag{2.14}$$

In case the inter-symbol interference (ISI) is considered within the channel, the previous SNR is attenuated and resulting into the following equation [49]:

$$SNR = \frac{\mathbf{R}_r^2\mathbf{P}_r^2}{\sigma_{shot}^2 + \sigma_{th}^2 + \mathbf{R}_r^2\mathbf{P}_{r,ISI}^2} \tag{2.15}$$

2.3.2 Modulation techniques in VLC

2.3.2.1 On-Off Keying (OOK) modulation

This is special case of the amplitude shift keying (ASK). It represents digital data as presence or absence of carrier wave. In this case, 1 denotes the presence of light and 0, absence of light. Fig. 2.8 presents a non return to zero (NRZ) OOK. A line of consecutive zeros or ones will create an unbalanced intensity in transmission; this is also called light flickering.

2.3.2.2 Variable pulse position modulation (VPPM)

This is a variation of the pulse position modulation (PPM). The PPM is a signal modulation used for both analog and digital signal transmissions, in which M message bits are encoded by transmitting a single pulse in 1 of 2^M possible require time-shifts.

The data mapping for VPPM is presented in Table 2.1 where the logical data '$0'$ or '$1'$ are mapped to high or low state depending on the dimming value.

Table 2.1: Data mapping for VPPM mode

Logical value	PHY value d is the VPPM duty cycle ($0.1 \leq d \leq 0.9$)	
0	High	$0 \leq t \leq dT$
	Low	$dT \leq t \leq T$
1	Low	$0 \leq t \leq (1-d)T$
	High	$(1-d)T \leq t \leq T$

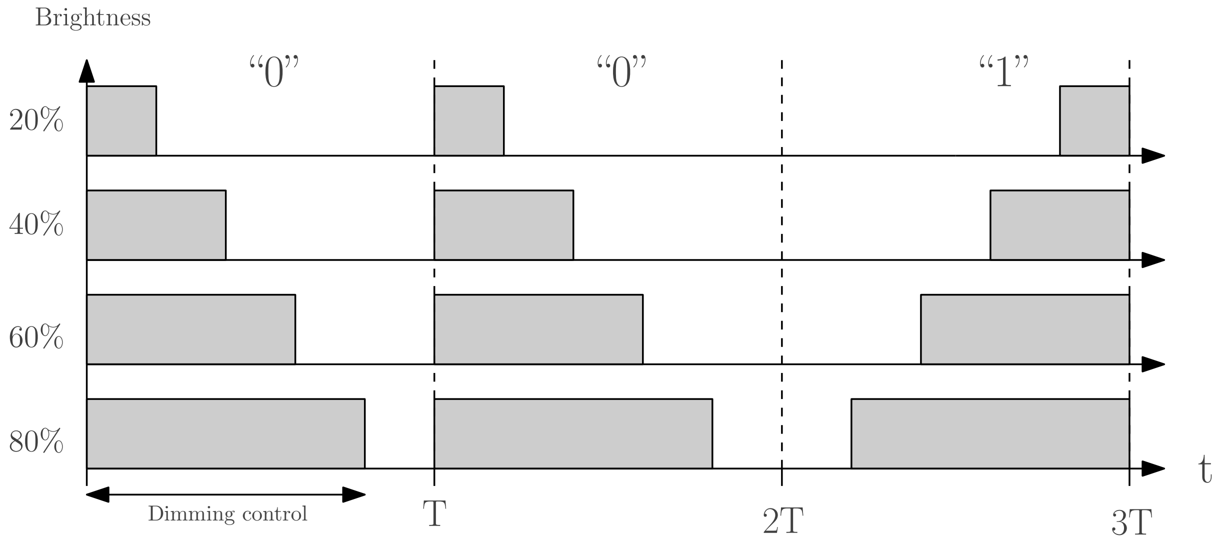


Figure 2.9: Schematic mechanism for VPPM modulation

The VPPM consists of varying the duty cycle to enable communication in case of light dimming. Fig. 2.9 depicts the mechanism of the VPPM modulation where the duty cycle of the period T_b can be adjust to support illumination while conveying data with dimming control. The VPPM modulation scheme can support simultaneously the illumination with dimming control while conveying data. VPPM uses characteristics of 2-PPM for non-flicker and pulse width modulation (PWM) for dimming control and full brightness.

2.3.2.3 Colour shift keying (CSK)

The CSK modulation scheme transmits data imperceptibly though the variation of the colour emitted by red, green and blue LEDs. Data symbols are mapped to colours pro-

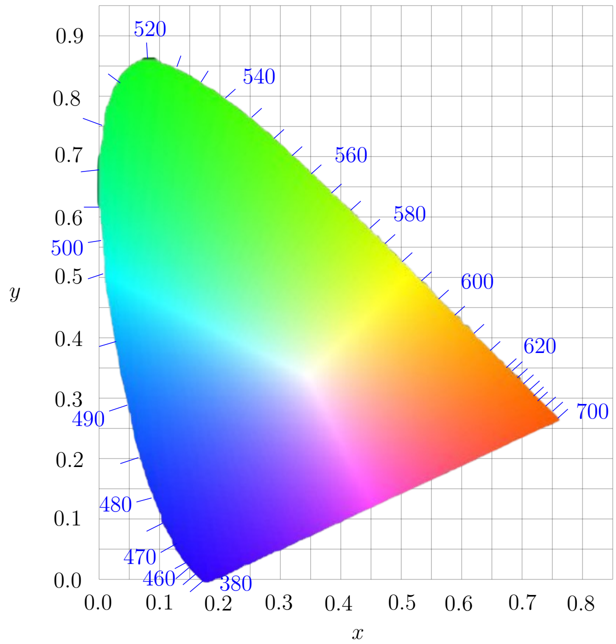


Figure 2.10: Color space

duced by RGB-LEDs.

If $\mathbf{S}_i = \{S_1, S_2, \ldots, S_N\}$ is the input symbol vector of size N; the CSK aims to map $\mathbf{S}_i$ into a corresponding colour constellation $\mathbf{C}_i = \{C_1, C_2, \ldots, C_N\}$. The CSK modulation represents any incoming symbol colour from the color space as shown in Fig. 2.10 into the 2D constellation triangle as a function of red, green and blue colours as depicted in Fig. 2.11. The constellation colour space will decompose each C_j, $j \in \{1, 2, \ldots, N\}$ into 3 coordinates: R_j, G_j and B_j. Different current values are applied to each RGB-LED to produce the colour C_k giving that the sum of R_j, G_j and B_j coordinates adds up to 1 [50, 51].

C_j could be identified through its xy coordinates defined as follows:

$$x_k = R_j x_r + G_j x_g + B_j x_b$$
$$y_k = R_j y_r + G_j y_g + B_j y_b$$

where $\{(x_r, y_r), (x_g, y_g), (x_b, y_b)\}$ are the xy coordinates of colours red, green and blue respectively.

Other modulations used in VLC channels include frequency shift keying (FSK), orthogonal frequency division multiplexing (OFDM) and optical spacial modulation (OSM)

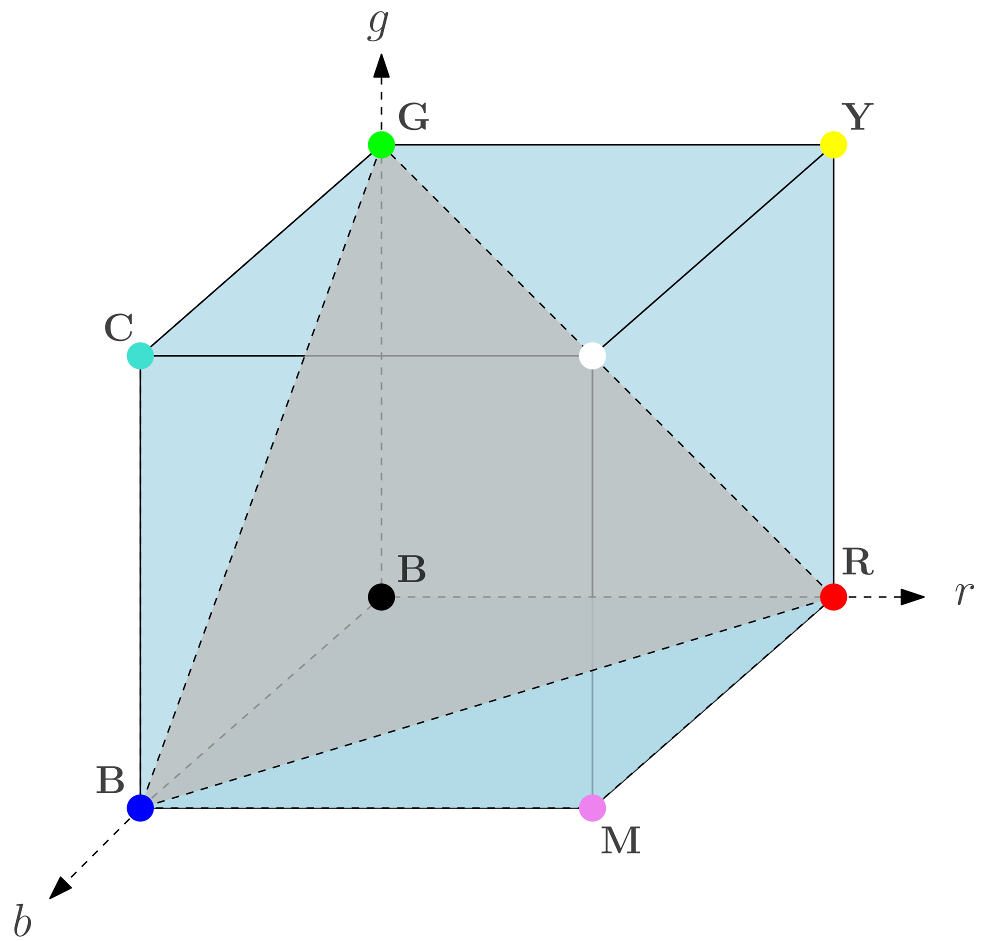


Figure 2.11: RGB plan and constellation triangle

for MIMO systems.

2.4 Run-Length Limited (RLL) Codes for VLC

Conventionally, the use of RLL codes impose a balanced flicker-free transmission. Manchester, 4B6B and 8B10B codes are the most used RLL codes in VLC [1]. A αBβB RLL code maps all length α words to β DC-free codewords with $(\alpha, \beta) \in \mathcal{N}$. There are $\binom{\beta}{\frac{\beta}{2}} = \frac{\beta!}{\frac{\beta}{2}!\frac{\beta}{2}!}$ balanced codewords of length β.

- The Manchester codes also called 1B2B codes encode the bit $0->01$ and the bit $1->10$.

Table 2.2: The 4B6B code as per VLC standards [1]

Input	Output	Dec Output	Input	Output	Dec Output
0000	001110	14	1000	011001	25
0001	001101	13	1001	011010	26
0010	010011	19	1010	011100	28
0011	010110	22	1011	110001	49
0100	010101	21	1100	110010	50
0101	100011	35	1101	101001	41
0110	100110	38	1110	101010	42
0111	100101	37	1111	101100	44

Table 2.3: The 3B4B code [2]

Input		$RD = -1$	$RD = +1$	Input		$RD = -1$	$RD = +1$
	HGF	fghj			HGF	fghj	
D.x.0	000	1011	0100	K.x.0	000	1011	0100
D.x.1	001	1001		K.x.1	001	0110	1001
D.x.2	010	0101		K.x.2	010	1010	0101
D.x.3	011	1100	0011	K.x.3	011	1100	0011
D.x.4	100	1101	0010	K.x.4	100	1101	0010
D.x.5	101	1010		K.x.5	101	0101	1010
D.x.6	110	0110		K.x.6	110	1001	0110
D.x.P7	111	1110	0001				
D.x.A7	111	0111	1000	K.x.7	111	0111	1000

- Table 2.2 presents the 4B6B code as defined in VLC standards [1], where 'dec output' represent the decimal notation of the binary outputs.

- The 8B10B code is a concatenation of 5B6B with 3B4B, this is because the number of balanced codewords of size 10, $\binom{10}{5} = 252 < 2^8$. The running disparity (RD) refers to the difference between the number of ones and zeros in a transmitted codeword. The 3B4B and 5B6B codes are defined based on two states RD: -1 and $+1$ such that their concatenation produces a balanced 8B10B code with RD of 0 [2].

  Tables 2.3 and 2.4 presents the 3B4B and 5B6B codes respectively that compose the 8B10B code used in VLC [2].

Table 2.4: The 5B6B code [2]

Input		$RD = -1$	$RD = +1$	Input		$RD = -1$	$RD = +1$
	EDCBA	abcdei			EDCBA	abcdei	
D.00	00000	100111	011000	D.16	10000	001011	100100
D.01	00001	011101	100010	D.17	10001	100011	
D.02	00010	101101	010010	D.18	10010	010011	
D.03	00011	110001		D.19	10011	110010	
D.04	00100	110101	001010	D.20	10100	001011	
D.05	00101	101001		D.21	10101	101010	
D.06	00110	011001		D.22	10110	011010	
D.07	00111	111000	000111	D.23	10111	111010	000101
D.08	01000	111001	000110	D.24	11000	110011	001100
D.09	01001	100101		D.25	11001	100110	
D.10	01010	010101		D.26	11010	010110	
D.11	01011	110100		D.27	11011	110110	001001
D.12	01100	001101		D.28	11100	001110	
D.13	01101	101100		D.29	11101	101110	010001
D.14	01110	011100		D.30	11110	011110	100001
D.15	01111	010111	101000	D.31	11111	101011	010100
				K.28	11100	001111	110000

2.5 Polar Codes for VLC

There is no direct relation between polar codes and VLC systems. However, VLC systems make use of FEC codes to strengthen the error correction performance of RLL codes. Throughout this thesis, we chose polar codes as FEC codes mainly because it is easier to estimate the theoretical error performance of such code given a specific channel. Polar codes are the first class of codes that explicitly achieves the Shannon capacity. They were introduced by Erdal Arıkan [52]. The construction of polar codes relies on many recursive combinations of short kernel code which converts a single channel into virtual channels. The reliability of data bits becomes higher as the number of recursions increases, while the rest of bits also known as *frozen bits* become very unreliable. This phenomenon is referred to as channel polarization.

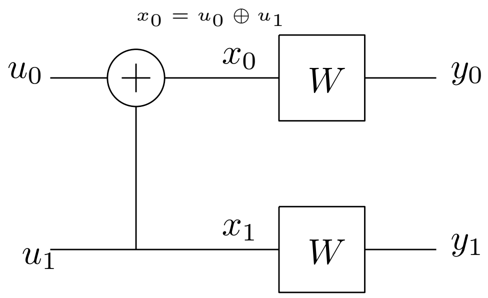


Figure 2.12: Channel combining for $N = 2$

2.5.1 Channel polarization

The channel capacity C is such that $0 \leq C \leq 1$. When a codeword is encoded into N channel users, N is also called the block length, the channel converts asymptotically C fraction of the N bit-channels as noiseless ($C \to 1$) and the remaining $(1 - C)$ fraction as highly noisy $C \to 0$.

This process of channel polarization is happening in two phases:

2.5.1.1 Channel combining

This is a recursive way to build the channel vector W_N where $N = 2^n$ from an input vector $\boldsymbol{u} = u_0 u_1 \dots u_{K-1}$ of length K. The initial operation is the XOR between the two first information bits as presented in Fig. 2.12.

This is a linear mapping: $\{x_0, x_1\} \mapsto \{y_0, y_1\}$. Eq. (2.16) presents the transition probability of the channel combining for two channel-users as shown in Fig. 2.12.

$$W_2(y_0, y_1|u_0, u_1) = W(y_0|u_0 \oplus u_1)W(y_1|u_1). \tag{2.16}$$

Fig. 2.13 presents the channel combining for $N = 8$ with $n = 3$ levels denoted by c_i, $1 \leq i \leq n$.

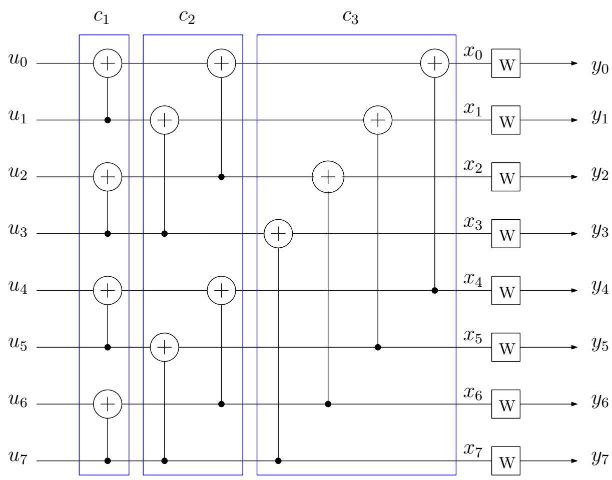


Figure 2.13: Channel combining for $N = 8$

2.5.1.2 Channel Splitting

This phase consists of splitting the channel vector W_N back to N channels $W_N^{(i)}$: $\{0,1\} \mapsto \mathcal{Y}_N \times \{0,1\}^i$, where $0 \leq i \leq N$.

The channels W^- and W^+ have the following transition probabilities:

$$W^-(y_0, y_1|u_0) = \frac{1}{2}\sum_{u_1} W(y_0|u_0 \oplus u_1)W(y_1|u_1). \tag{2.17}$$

$$W^+(y_0, y_1, u_1|u_0) = \frac{1}{2}W(y_0|u_0 \oplus u_1)W(y_1|u_1). \tag{2.18}$$

The channel splitting described above was done for $N = 2$. This technique can be extended for any $N = 2^n$. For instance $N = 4$, the channel vector W_4 can be split into 4 copies of the channel: W^{++}, W^{+-}, W^{-+} and W^{--}. Similarly for $N = 8$, W_8 can be separated into 8 different channels: W^{+++}, W^{++-}, W^{+-+}, W^{+--}, W^{-++},W^{-+-}, W^{--+} and W^{---}.

2.5.2 Construction of Polar codes

Assuming a source information vector $\boldsymbol{u}$ of length K to be encoded into $\boldsymbol{x}$ of N channel-users. The code construction consists of ranking the indices of N polarized channels

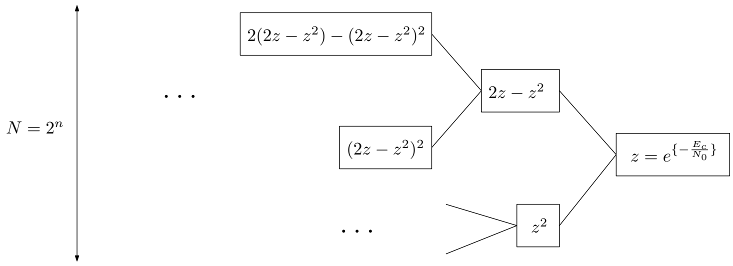


Figure 2.14: Polar code construction

according to the reliability criteria; then assigning the bit informations to the most reliable K indices and set the remaining $N - K$ indices as *frozen bits*. The *frozen bits* are set to "0" and represent the indices of the worst reliable channel-users.

In [52], the code construction based on Bhattacharyya parameters was presented using a Monte-Carlo simulation. Tal and Vardy [53] described an approach based on channel quantization. In [54], a code construction was done through the density evolution Gaussian approximation (DEGA) technique.

The Bhattacharyya parameter has an important role when constructing polar codes. It was demonstrated in [52] (Appendix-D) that a code construction can be performed recursively from the following initial functions $\{z, z\} \mapsto \{2z - z^2, z^2\}$ at each level due to channel polarization, where z is the Bhattacharyya parameter.

Fig. 2.14 presents the polar code construction for AWGN channel through a tree where root corresponds to the design SNR which is set as the Bhattacharyya parameter, $z = e^{-\frac{E_c}{N_0}}$. This identifies the polar code as *non-universal* under successive cancellation (SC) decoding because they change according to the design SNR.

2.5.3 Encoding of Polar codes

The encoding consists of a transformation of the source information vector $\boldsymbol{u}^{K-1}$ into the encoded codeword $\boldsymbol{x}^{N-1}$; this can be done with matrices. The basic transformation for $N = 2$ is provided by F.

$$F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

The $N \times N$ matrix is constructed recursively using the Kronecker power of F (denoted by $\otimes$) to obtain, $F^{\otimes \log N}$;

$$F^{\otimes \log N} = F \otimes F^{\otimes \log N-1}.$$

Let $\mathcal{B}$ be the set of $N-1$ indices; and $\mathcal{I}$ be the set of reversed bit indices in $\mathcal{B}$, $\mathcal{B} = bitrev(\mathcal{I})$, $bitrev(b_0 b_1 \dots b_{N-1}) \equiv b_{n-1} \dots b_1 b_0$. An illustration of the bit reversal function is presented in Table 2.5 for $N = 8$. The complementary set of $\mathcal{I}^c$ is called frozen bit indices. Let $\boldsymbol{d}^{N-1}$ be the carrier of the information vector $\boldsymbol{u}^{K-1}$ and defined as follow:

$$\boldsymbol{d}^{N-1} : \begin{cases} \boldsymbol{d}_{\mathcal{I}^c} = 0 \\ \boldsymbol{d}_{\mathcal{I}} = \boldsymbol{u} \end{cases}$$

The encoded vector is $\boldsymbol{x}^{N-1} = F^{\otimes \log N} \boldsymbol{d}^{N-1}$. This encoding scheme is similar to most linear codes with a generator matrix G_N corresponding to $F^{\otimes \log N}$, such that: $\boldsymbol{x}^{N-1} = G_N \boldsymbol{d}^{N-1}$.

For $N = 8$ as in Fig. 2.12, the kernel matrix is presented as follow:

$$F^{\otimes 3} = G_8 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The encoded codeword $\boldsymbol{x}^{N-1}$ is non-systematic, that is the information bit sequence appear altered in the encoded version of the vector.

Example 2.1 For parameters $N = 8$, $K = 5$, $E_c/N_0 = 0.5$, the constructed codeword is $\boldsymbol{u}^{N-1} = u_7 u_3 u_5 u_6 u_1 \underline{u_2 u_4 u_0}$ and $\mathcal{I} = \{1, 3, 5, 6, 7\}$. If we assume the information source

Table 2.5: Bit reversal illustration for $N = 8$

$\mathcal{B}$	bit representation	$bitrev\{\mathcal{B}\}$	$\mathcal{I}$
0	000	000	0
1	001	100	4
2	010	010	2
3	011	110	6
4	100	001	1
5	101	101	5
6	110	011	3
7	111	111	7

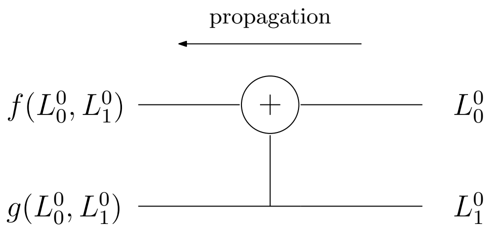


Figure 2.15: SC polar decoder for a single node ($N = 2$)

$\boldsymbol{u} = 11011$, then the encoding is done as follow:

$$\begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \underline{0} \\ 1 \\ \underline{0} \\ 1 \\ \underline{0} \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

So the encoded vector is $\boldsymbol{x}^{N-1} = 01100101$. □

2.5.4 Decoding of polar codes

The successive cancellation (SC) decoding is a basic and recursive decoder [52]. It make uses of two-way algorithm based on N received LLR values. SC decoding is a sequential decoder contrarily to belief propagation decoding which is done in parallel.

Fig. 2.15 presents a basic node of the SC decoding with two LLR values. Those LLR values are propagated from right to left following the two functions: f and g define as follow:

$$\begin{pmatrix} l_j^i \\ l_{j+1}^i \end{pmatrix} \mapsto \begin{pmatrix} f(l_j^i, l_{j+1}^i) \\ g(l_j^i, l_{j+1}^i) \end{pmatrix} = \begin{pmatrix} \frac{l_j^i l_{j+1}^i + 1}{l_j^i l_{j+1}^i} \\ l_j^i l_{j+1}^i \text{ or } l_{j+1}^i / l_j^i \end{pmatrix} \tag{2.19}$$

l_j^i corresponds to the LR value of x_j at the level i, with $0 \leq j \leq N-1$ and $0 \leq i \leq \log N$. The g-function depends on f; it is based on the bit decision from upper branch: $0 \mapsto l_j^i l_{j+1}^i$ and $1 \mapsto l_{j+1}^i / l_j^i$.

By transposing in the *log*-domain, Eq. 2.19 becomes

$$\begin{aligned} \begin{pmatrix} L_j^i \\ L_{j+1}^i \end{pmatrix} &\mapsto \begin{pmatrix} \log f(e^{L_j^i}, e^{L_{j+1}^i}) \\ \log g(e^{L_j^i}, e^{L_{j+1}^i}) \end{pmatrix} \\ &= \begin{pmatrix} \log \left(\frac{1+e^{L_j^i + L_{j+1}^i}}{e^{L_j^i} + e^{L_{j+1}^i}} \right) \\ L_j^i L_{j+1}^i \text{ or } L_{j+1}^i / L_j^i \end{pmatrix} \\ &\approx \begin{pmatrix} \text{sign}(L_j^i)\text{sign}(L_{j+1}^i)\min\{|L_j^i|, |L_{j+1}^i|\} \\ L_j^i + (-1)^j L_{j+1}^i \end{pmatrix} \end{aligned} \tag{2.20}$$

Similarly, L_j^i corresponds to the LLR value of x_j at the level i, with $0 \leq j \leq N-1$ and $0 \leq i \leq \log N$. Fig. 2.16 shows the SC decoding structure for $N = 4$; because of the sequential nature of the decoder, the LLR values are propagated to the right and once a hard decision is made (according to Eq. (2.20)) for the first bit, then this decision gets *broadcast* back (from left to right) to estimate other bits, as shown in red on the same figure. Fig. 2.17 is a simple representation of the calculation of the first bit reliability through a tree structure.

The BER (bit-error rate) and FER (frame error rate) are some metrics to appreciate the error probability P_E of the channel while decoding the source information. The BER is the amount of error received bits divided by the total number of bits whereas the FER is the number of corrupted frames out of the total number of received frames through a channel. Those are estimate for long time interval and high number of bits error. The bit error probability is the expectation value of the BER. A union bound on the BER of

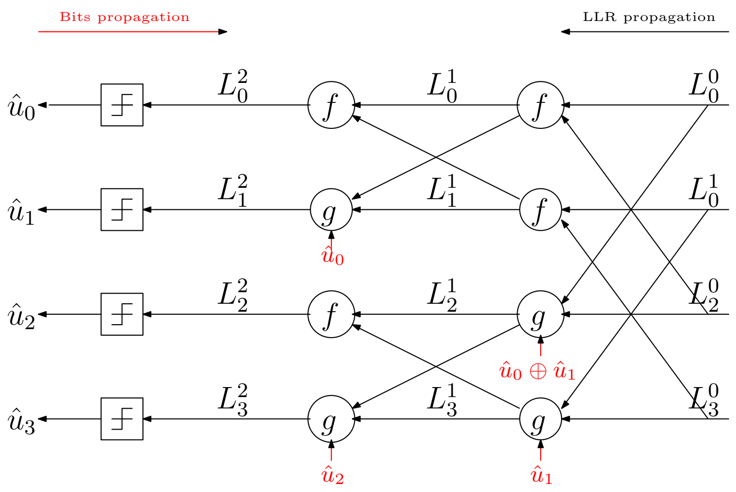


Figure 2.16: SC polar decoder for $N = 4$

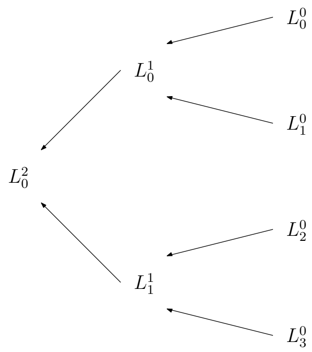


Figure 2.17: SC polar tree for the first bit where $N = 4$

polar codes was provided in [52] (3, Proposition 2) as follow:

$$P_{BER} \leq \sum_{i \in \mathcal{I}} P_E(y_i|x_i). \tag{2.21}$$

where $P_E(y_i|x_i)$ is the error probability of the channel on a specific user.

3 Improved DC-free Run Length Limited 4B6B Codes for Concatenated Schemes

In this chapter, we introduce a new class of improved DC-free 4B6B codes in terms of error correction performance for serially concatenated coding techniques. This class of code is made of the same codewords contained in the original 4B6B code as per IEEE 802.15.7 standard for visible light communication (VLC).

3.1 Introduction

In VLC, the use of DC-free RLL codes enforce the mitigation of the flickering during communication by ensuring a constant average illumination. However, due to the poor error correction ability of these RLL codes, a forward error correction scheme is always used as outer codes. Moreover, the design of good RLL codes become important in order to achieve a reliable VLC channel.

There are several works from the literature based on designing RLL codes for VLC. In [55], a FEC-aware design of RLL codes is presented where the frozen indexes of the polar codes are pre-determine by the generator matrix together with channel selection methods. This polar codes construction provides patterns of different *a priori* probabilities in RLL decoding and improve the bit error rate (BER) results. But, for high-rate codes where the number of free frozen bits positions is reduced, this technique becomes almost obsolete.

In [4], a DC-free 5B10B RLL code was introduced for concatenated schemes. The code improves the error correction performance of a VLC channel compared to 4B6B and Manchester codes. However, its spectral efficiency is as low as that of Manchester code and its decoding computational complexity is much larger compared to existing RLL codes.

A modification of the 4B6B code into a 3B6B code was presented in [56]. This code has been designed for recording systems [57] and was improved for VLC systems. A trellis code was designed to maximise the code distance profile while enhancing the spectral efficiency and the error correction performance fat the brightness level of 50%. Similarly, a low complexity phase code was proposed as RLL code in serial concatenation with convolutional codes; the extrinsic information transfer (EXIT) was exploited to analyse the iterative decoding of the system based on trellis diagram study. Bounds for concatenated VLC system were derived with non-linear RLL codes as inner codes.

A class of rate $(n-1)/n$ RLL codes was introduced in [3]. This class aims to generate $(0,k)$ RLL codes with $k = [3,7]$, $(0,k)$ is a (d,k) sequence having the minimum and maximum number of consecutive zeros being 0 and k respectively. The error correction performance of this code increases with larger run-length k; this could considerably attenuate the flicker-free property of the light during transmission.

An enhanced Miller code denoted by *eMiller* code was shown in [58]. Although the fact that it slightly improves the error correction capability compared to some RLL codes, its spectral efficiency is as low as that of the Manchester code. Furthermore, the decoding of this code involves a high complexity that increases with larger constraint lengths compared to the traditional Viterbi algorithm.

3.2 Background

Fig. 3.1 represents the block diagram for a flicker-free VLC system. The input data $\boldsymbol{u} \triangleq [u_0, u_1, \ldots, u_{K-1}]$ is encoded by an outer FEC code then by an inner DC-free RLL code. The resulting codeword having a dimming ratio of 50% is modulated via on-off keying (OOK) as $\boldsymbol{y} \triangleq [y_0, y_1, \ldots, y_{N-1}]$ and then broadcast by a LED light into a VLC channel, modelled in this case as an AWGN channel. At the receiver, the signal $\boldsymbol{r}$ is absorbed by a photodetector (PD), demodulated and then the corrupted version of the input data is recovered after the RLL and FEC decoding. The rate for an error control code and RLL line code for given VLC channel with a given dimming constraint is governed by the IEEE 802.15.7 standards for VLC [1] where modulations, clock rates and data rates are recommended depending on the PHY types (PHY I, PHY II, PHY III) as shown in Table 3.1.

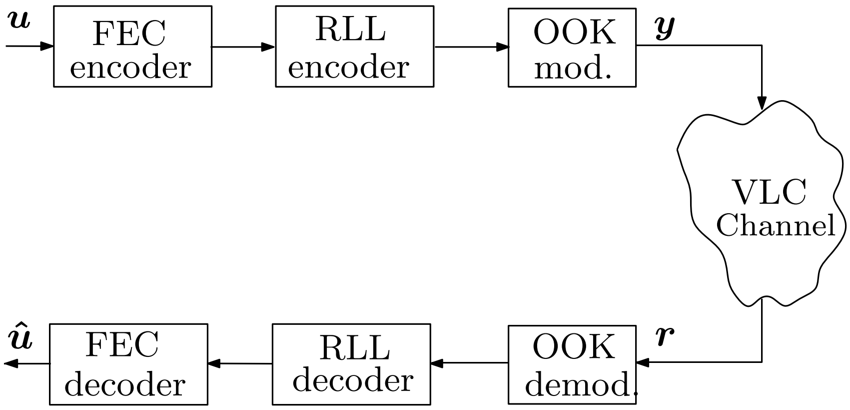


Figure 3.1: Block diagram of the VLC system.

Table 3.1: Three physical layers in IEEE 802.15.7 standard [1].

PHY types	Modulation	Clock rates	Data rates
PHY-I	OOK & VPPM	200/400 kHz	11.67–266.6 kb/s
PHY-II	OOK & VPPM	$\leq$ 120 MHz	1.25–96 Mb/s
PHY-III	CSK	12/24 MHz	12–96 Mb/s

Table 3.2: Flicker mitigation for various modulation modes [1].

Flicker mitigation	Data transmission (intra-frame flicker)	Idle or RX periods (inter-frame flicker)
OOK modulation	Dimmed OOK mode, RLL code	Idle/visibility patterns
VPPM modulation	VPPM guarantee no intra-frame flicker, RLL code	
CSK modulation	constant average power across multiple light sources, scrambler, high optical clock rates (MHz)	

It can be observed the OOK modulation is recommended for PHY-I and PHY-II. For the sake of uniformity, this modulation is used throughout the simulations conducted in this thesis. Table 3.2 shows the flicker mitigation for different modulation modes. For the case of OOK, a dimmed OOK mode allows communication at various light dimming values while the RLL code ensures a flicker-free communication at a dimming value of 50%.

Tables 3.3, 3.4 and 3.5 depict the operating modes for PHY-I, II and III respectively. The operating mode consists of associating an outer FEC code to an inner code in function of the data and optical clock rate for a specific modulation.

Table 3.3: PHY-I Operating modes [1].

Modulation	RLL code	Optical clock rate	FEC		Data rate
			Outer code (RS)	Inner code (CC)	
OOK	Manchester	200 kHz	(15,7)	1/4	11.67 kb/s
			(15,7)	1/3	24.44 kb/s
			(15,7)	2/3	48.89 kb/s
			(15,7)	none	73.3 kb/s
			none	none	100 kb/s
VPPM	4B6B	400 kHz	(15,2)	none	35.56 kb/s
			(15,4)	none	71.11 kb/s
			(15,7)	none	124.4 kb/s
			none	none	266.6 kb/s

Table 3.4: PHY-II Operating modes [1].

Modulation	RLL code	Optical clock rate	FEC	Data rate
VPPM	4B6B	3.75 MHz	$RS(64, 32)$	1.25 Mb/s
			$RS(160, 128)$	2 Mb/s
		7.5 MHz	$RS(64, 32)$	2.5 Mb/s
			$RS(160, 128)$	4 Mb/s
			none	5 Mb/s
OOK	8B10B	15 MHz	$RS(64, 32)$	6 Mb/s
			$RS(160, 128)$	9.6 Mb/s
		30 MHz	$RS(64, 32)$	6 Mb/s
			$RS(160, 128)$	19.2 Mb/s
		60 MHz	$RS(64, 32)$	24 Mb/s
			$RS(160, 128)$	38.4 Mb/s
		120 MHz	$RS(64, 32)$	48 Mb/s
			$RS(160, 128)$	76.8 Mb/s
			none	96 Mb/s

Table 3.5: PHY-III Operating modes [1].

Modulation	Optical clock rate	FEC	Data rate
4-CSK	12 MHz	$RS(64, 32)$	12 Mb/s
8-CSK		$RS(64, 32)$	18 Mb/s
4-CSK	24 MHz	$RS(64, 32)$	24 Mb/s
8-CSK		$RS(64, 32)$	36 Mb/s
16-CSK		$RS(64, 32)$	48 Mb/s
8-CSK		none	72 Mb/s
16-CSK		none	96 Mb/s

Several DC-free RLL codes have been proposed from the literature. Fig. 3.2 shows

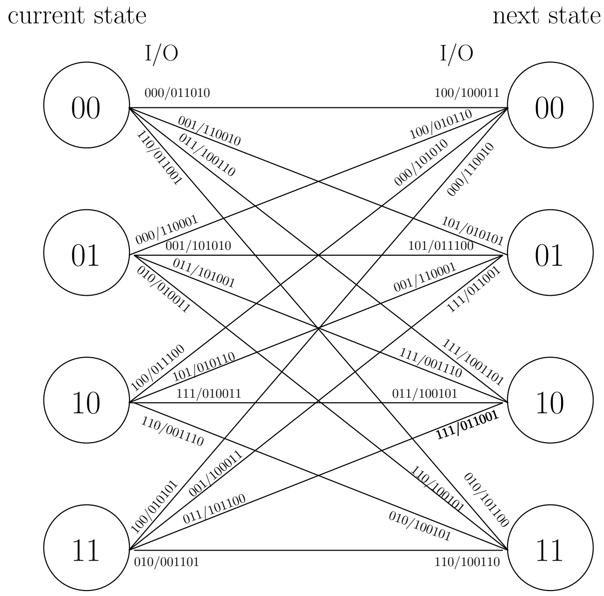


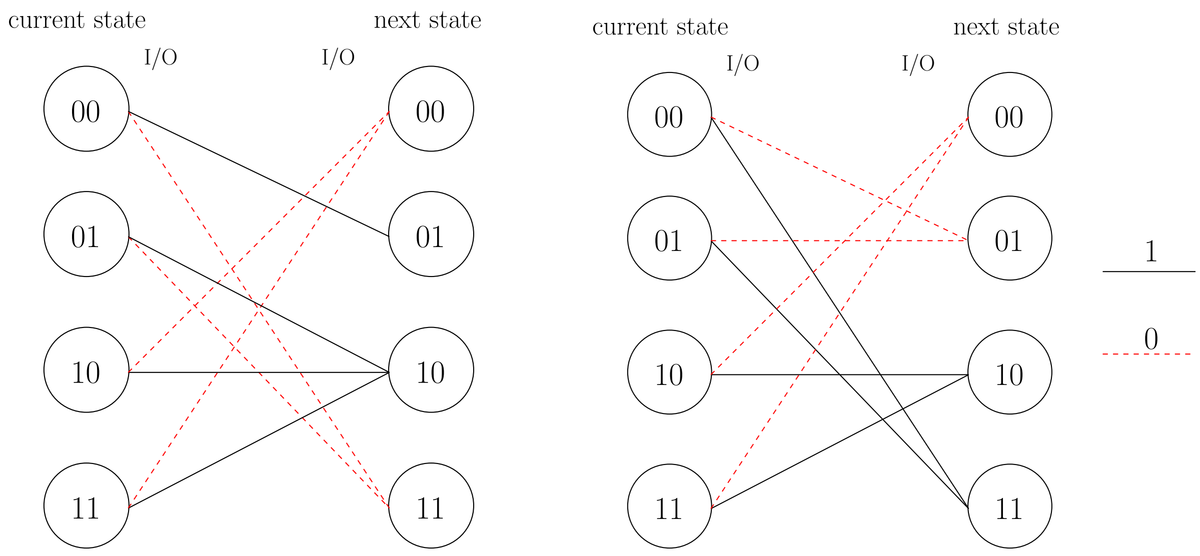
Figure 3.2: Trellis diagram of the 3B6B code.

the trellis diagram of the 3B6B code used in the scheme [56]. Every two branches are represented with one transition line and two labels on the right and left side. The 3B6B code has been designed in [57] for magnetic recording application.

Fig. 3.3a shows the Miller code. In [58] an enhanced version of Miller code denoted as eMiller code was proposed. Fig 3.3b shows the eMiller code derived from the Miller code. A decoder was also proposed for the eMiller code based on the modification performed on the trellis at the cost of a slightly higher complexity.

A class of $(0, k)$ RLL codes was introduced in [3] to build flicker-free VLC channel. The scheme was described through a 2-bit input and 3-bit output code, but it can be extended to rate $(n-1)/n$ coding schemes with $n > 2$. Table 3.6 shows the state transition structure of $(0, 4)$ 2/3 RLL code as defined in [3], where $k = 4$ represents the maximum run-length in the code.

A 5B10B code was introduced in [4] to improve the error correction performance of VLC channels. Table 3.7 shows the 5B10 code as per [4].



(a) Miller code

(b) eMiller code

Figure 3.3: Miller vs. eMiller code.

Table 3.6: Trellis code for 2/3 RLL code [3].

Current state	Input bits	Next state	Output bits
$S_0(000)$	00	$S_2(010)$	010
$S_2(010)$	01	$S_3(011)$	011
$S_4(100)$	10	$S_6(110)$	110
$S_6(110)$	11	$S_7(111)$	111
$S_1(001)$	00	$S_0(000)$	000
$S_3(011)$	01	$S_1(001)$	001
$S_5(101)$	10	$S_4(100)$	100
$S_7(111)$	11	$S_5(101)$	101

3.3 Proposed DC-free Run Length Limited Codes

It was shown in [59] that the weight distribution of a code can be represented by the *weight enumerating function* (WEF) that is expressed as $A(D) = \sum_{d=0}^{\beta} A_d D^d$ where A_d is the number of codewords with weight d. The *input-output weight enumerating function* (IOWEF), $B(W, D)$ of the 4B6B code maps the WEF of the 4-bit code input to the 6-bit code output:

$$B(W, D) \triangleq \sum_{w=0}^{\alpha} \sum_{d=0}^{\beta} B_{w,d} W^w D^d. \tag{3.1}$$

Table 3.7: 5B10B code [4].

Dec	Input bits	Code	Dec	Input bits	Code
0	00000	1100110001	16	10000	0111010001
1	00001	1110001001	17	10001	0101111000
2	00010	1110010010	18	10010	0101100011
3	00011	0100011011	19	10011	0110101010
4	00100	1101000101	20	10100	0110110100
5	00101	1100011100	21	10101	0100101101
6	00110	1100100110	22	10110	0101010110
7	00111	1101001010	23	10111	0111001100
8	01000	1001010011	24	11000	1001101001
9	01001	1011011000	25	11001	0010111001
10	01010	1010100011	26	11010	0011110010
11	01011	1000111010	27	11011	0011001011
12	01100	1001110100	28	11100	0011100101
13	01101	1010010101	29	11101	0001011101
14	01110	1011000110	30	11110	0001101110
15	01111	1010101100	31	11111	0010011110

Because of the non-linearity of the code, an average is performed over $B_{w,d}$ values as shown in [56] with $B_{w,d}$ being the number of codewords with weight d derived from words of weight w; α and β are the lengths of the input and output code respectively.

The traditional 4B6B code has the following IOWEF:

$$\begin{aligned} B(W,D) = {} & D^0(W^0) \\ & + D^2(1.5W^1 + 3.625W^2 + 1.5W^3 + 0.375W^4) \\ & + D^4(2.5W^1 + 2.250W^2 + 1.75W^3 + 0.5W^4) \\ & + D^6(0.125W^2 + 0.75W^3 + 0.125W^4). \end{aligned} \tag{3.2}$$

Table 3.8 presents the distance profile between words of the 4-bit codebook. Given 2 words $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ within this set, the weight is computed as $w = sum(\boldsymbol{u}_1 \ominus_2 \boldsymbol{u}_2)$, where $\ominus_2$ denotes the binary subtraction using 1's complement, sum, the algebraic sum of all bit and $w \in \{0, 1, 2, 3, 4\}$.

Table 3.9 shows the distance profile between codewords of the 6-bit codebook for the original 4B6B. Given two codewords from this set, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, the distance is defined as $d = sum(\boldsymbol{x}_1 \oplus_2 \boldsymbol{x}_2)$, where $\oplus_2$ denotes the binary addition and $d \in \{0, 2, 4, 6\}$.

Table 3.8: 4B6B code weight profile of the 4-bit codebook

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0		3	3	2	3	2	2	1	3	2	2	1	2	1	1	0
1	3		2	3	2	3	1	2	2	3	1	2	1	2	0	1
2	3	2		3	2	1	3	2	2	1	3	2	1	0	2	1
3	2	3	3		1	2	2	3	1	2	2	3	0	1	1	2
4	3	2	2	1		3	3	2	2	1	1	0	3	2	2	1
5	2	3	1	2	3		2	3	1	2	0	1	2	3	1	2
6	2	1	3	2	3	2		3	1	0	2	1	2	1	3	2
7	1	2	2	3	2	3	3		0	1	1	2	1	2	2	3
8	3	2	2	1	2	1	1	0		3	3	2	3	2	2	1
9	2	3	1	2	1	2	0	1	3		2	3	2	3	1	2
10	2	1	3	2	1	0	2	1	3	2		3	2	1	3	2
11	1	2	2	3	0	1	1	2	2	3	3		1	2	2	3
12	2	1	1	0	3	2	2	1	3	2	2	1		3	3	2
13	1	2	0	1	2	3	1	2	2	3	1	2	3		2	3
14	1	0	2	1	2	1	3	2	2	1	3	2	3	2		3
15	0	1	1	2	1	2	2	3	1	2	2	3	2	3	3	

Table 3.9: 4B6B code distance profile of the 6-bit codebook

	14	13	19	22	21	35	38	37	25	26	28	49	50	41	42	44
14		2	4	2	4	4	2	4	4	2	2	6	4	4	2	2
13	2		4	4	2	4	4	2	2	4	2	4	6	2	4	2
19	4	4		2	2	2	4	4	2	2	4	2	2	4	4	6
22	2	4	2		2	4	2	4	4	2	2	4	2	6	4	4
21	4	2	2	2		4	4	2	2	4	2	2	4	4	6	4
35	4	4	2	4	4		2	2	4	4	6	2	2	2	2	4
38	2	4	4	2	4	2		2	6	4	4	4	2	4	2	2
37	4	2	4	4	2	2	2		4	6	4	2	4	2	4	2
25	4	2	2	4	2	4	6	4		2	2	2	4	2	4	4
26	2	4	2	2	4	4	4	6	2		2	4	2	4	2	4
28	2	2	4	2	2	6	4	4	2	2		4	4	4	4	2
49	6	4	2	4	2	2	4	2	2	4	4		2	2	4	4
50	4	6	2	2	4	2	2	4	4	2	4	2		4	2	4
41	4	2	4	6	4	2	4	2	2	4	4	2	4		2	2
42	2	4	4	4	6	2	2	4	4	2	4	4	2	2		2
44	2	2	6	4	4	4	2	2	4	4	2	4	4	2	2	

The same codewords from the 4B6B code were used to obtain the proposed 4B6B code. This is to avoid using the codewords dedicated for synchronization and also to preserve the run-length property amongst the codewords.

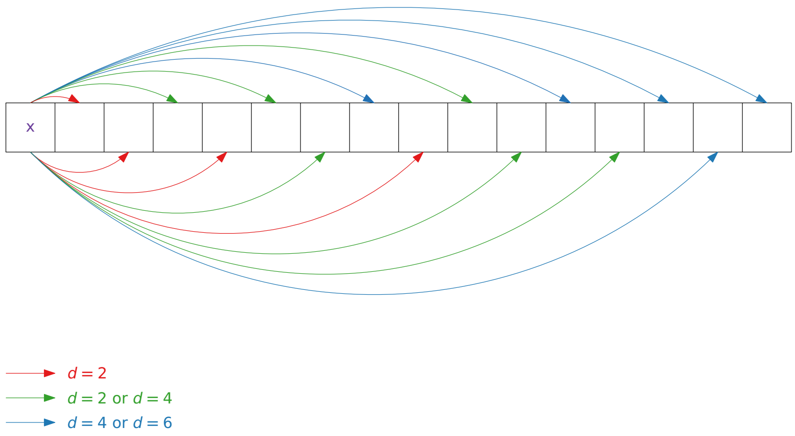


Figure 3.4: Graphical representation of distance constraints for codeword 1.

But, the change in order of codewords within the codebook influences the error correction performance which is directly linked to the IOWEF of a codebook. There are $16! = 2.09228 \times 10^{13}$ different possible permutations, having different distance profiles. We define the metric M^d to find the most efficient code in terms of error correction performance.

$$M^d = \sum_w (A_w \times w) \text{ where } A_w = \sum_{d=0}^{\beta} B_{w,d} D^d. \tag{3.3}$$

This metric gets lower values when close words are mapped to close codewords. This boils down into minimising M^2 then M^4, and finally M^6.

This task was solved through a backtracking algorithm; Knuth's algorithm X [60] was used to produced an efficient implementation of this algorithm. Some constraints are imposed to get lower values of M^d. Each codeword has 7 codewords distant by a Hamming distance of 2, 7 other codewords distant by a Hamming distance of 4, and 1 codeword with a distance of 6. In addition, each word considers 4 words distant by 1, 6 words distant by 2, 6 words distant by 3, and 1 word distant by 1. Thus, words at a distance of 1 must map to codewords that have a distance of 2 between each other, and that the word distant to 4 maps to a codeword distant to 6. For the others locations, the constraint can be relaxed with the codewords being distant from 2 or 4. The different constraints for the every codeword from 1 to 16 are depicted in Figs. 3.4 to 3.19 where 'X' represents the current position of the codeword subject to distance constraints.

These graphical representation inform on how to arrange codewords from the original

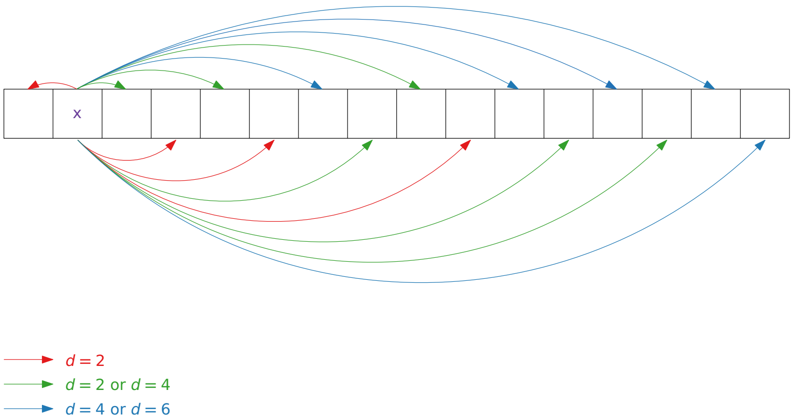


Figure 3.5: Graphical representation of distance constraints for codeword 2.

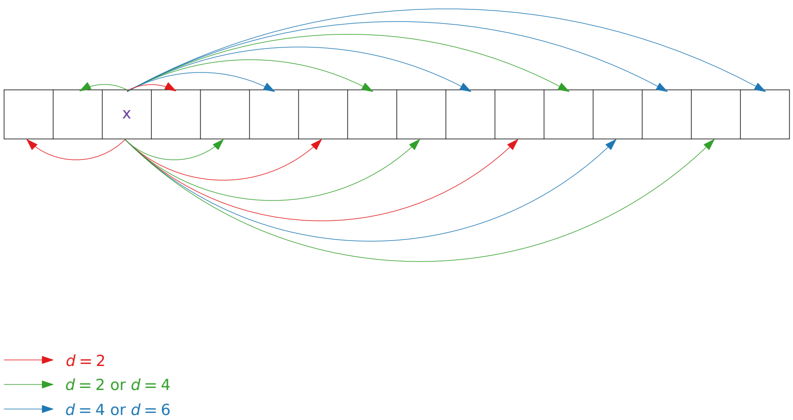


Figure 3.6: Graphical representation of distance constraints for codeword 3.

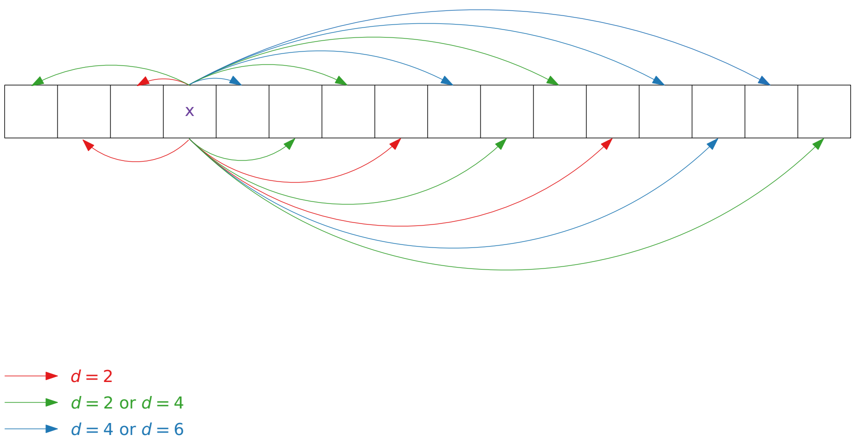


Figure 3.7: Graphical representation of distance constraints for codeword 4.

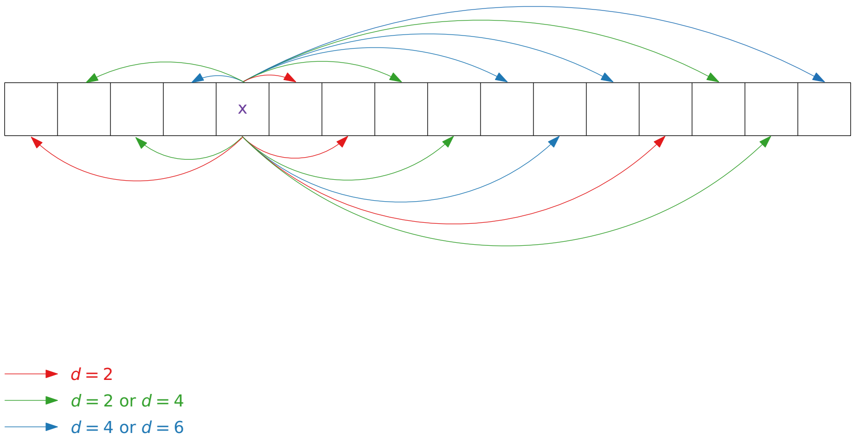


Figure 3.8: Graphical representation of distance constraints for codeword 5.

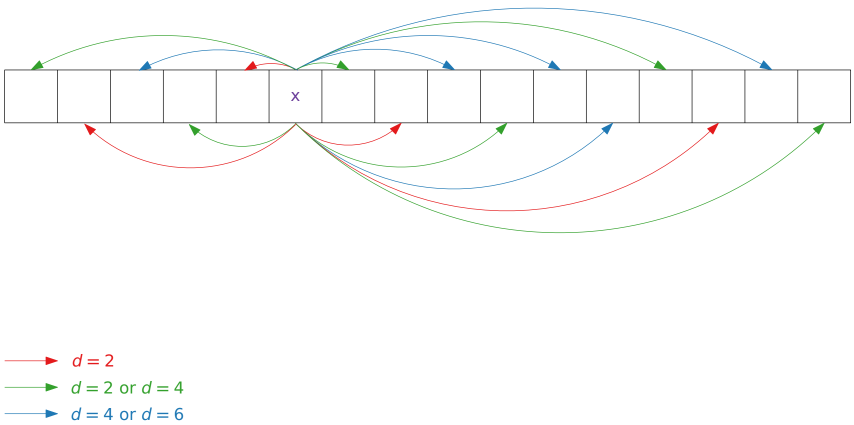


Figure 3.9: Graphical representation of distance constraints for codeword 6.

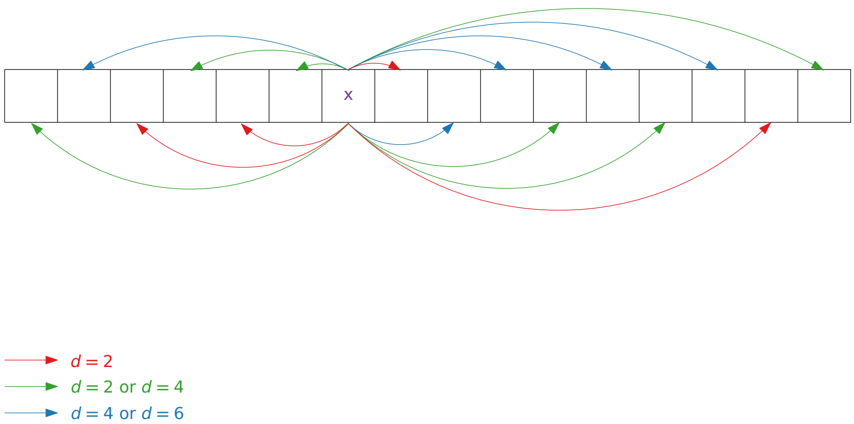


Figure 3.10: Graphical representation of distance constraints for codeword 7.

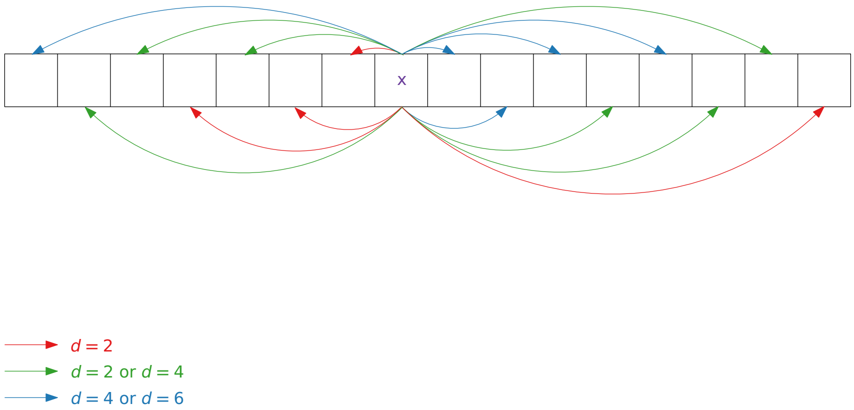


Figure 3.11: Graphical representation of distance constraints for codeword 8.

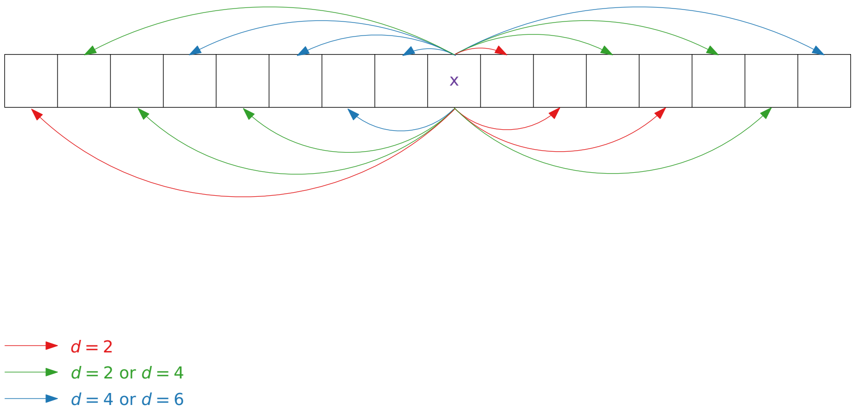


Figure 3.12: Graphical representation of distance constraints for codeword 9.

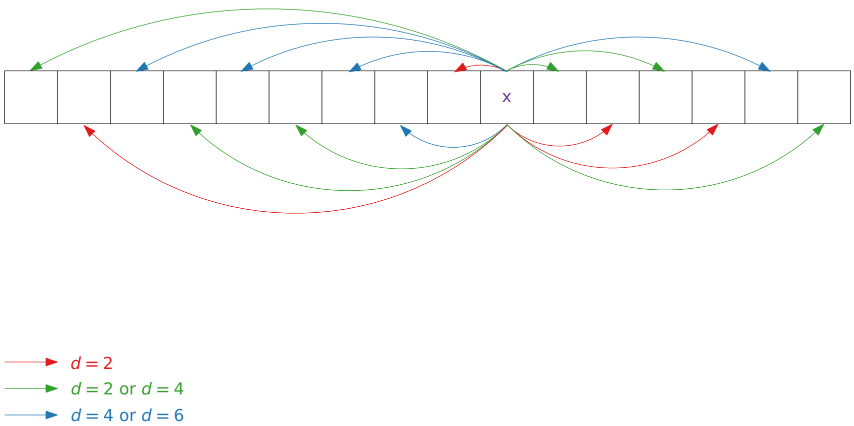


Figure 3.13: Graphical representation of distance constraints for codeword 10.

4B6B codebook in such a way to fulfil the distance constrained in order to obtain the minimal M^d. 768 codebooks that satisfied the aforementioned constraints were obtained.

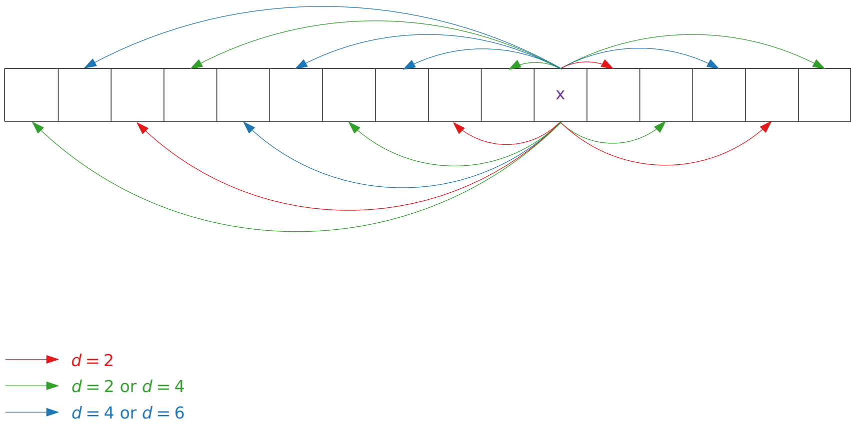


Figure 3.14: Graphical representation of distance constraints for codeword 11.

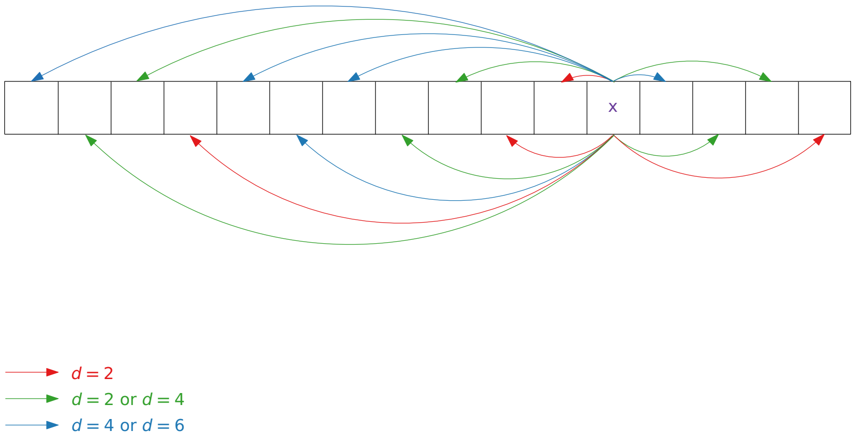


Figure 3.15: Graphical representation of distance constraints for codeword 12.

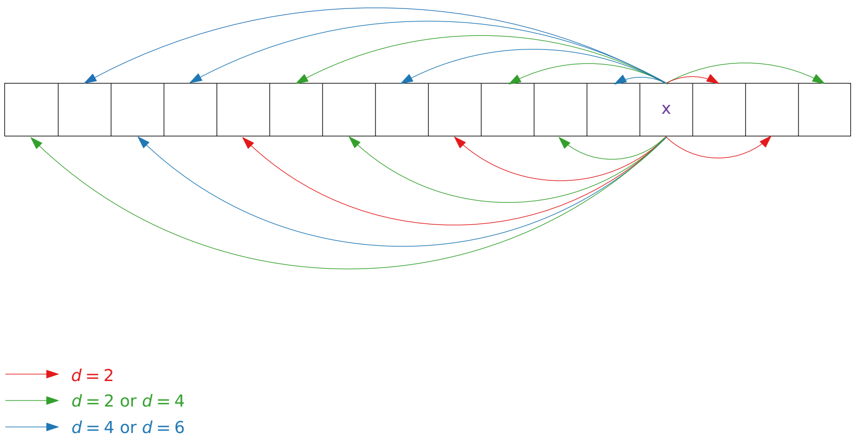


Figure 3.16: Graphical representation of distance constraints for codeword 13.

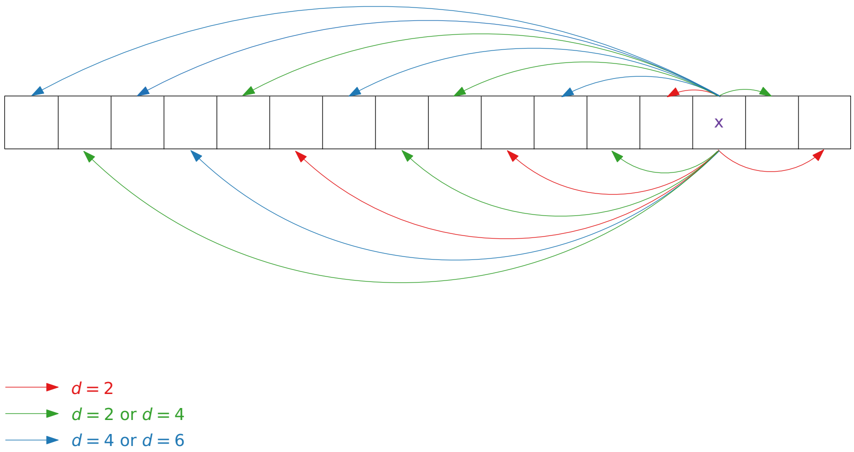


Figure 3.17: Graphical representation of distance constraints for codeword 14.

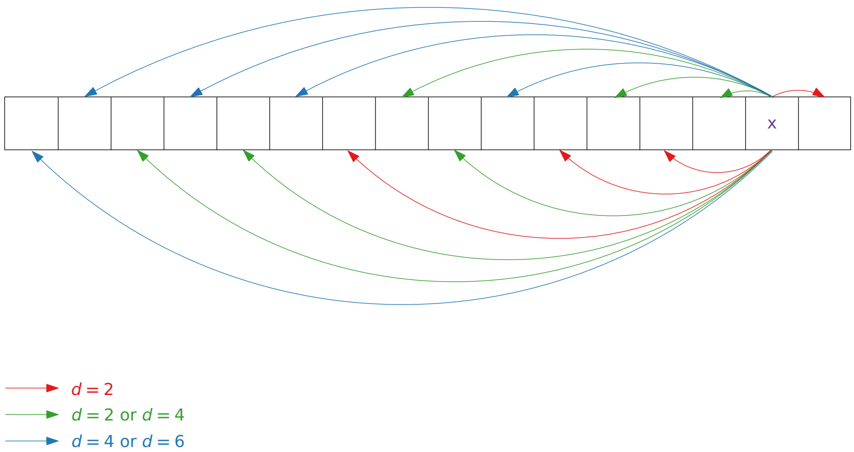


Figure 3.18: Graphical representation of distance constraints for codeword 15.

All these codes have the same IOWEF described as:

$$\begin{aligned} B(W, D) &= D^0(W^0) + D^2(4W^1 + 3W^2) \\ &\quad + D^4(3W^2 + 4W^3) + D^6(W^4). \end{aligned} \tag{3.4}$$

The profile of the most efficient 4B6B code corresponds to the minimum metric of $M^2 = 10$. It can be observed that (3.4) has fewer components compared to (3.2). A brute-force search was performed to confirm that this is indeed the minimum metric. Table 3.10 presents one of the 768 codes with the minimum metric. It was also verified that this metric leads to the best decoding performance for concatenated schemes. The proposed 4B6B code result from optimizing the mapping between the same set of input

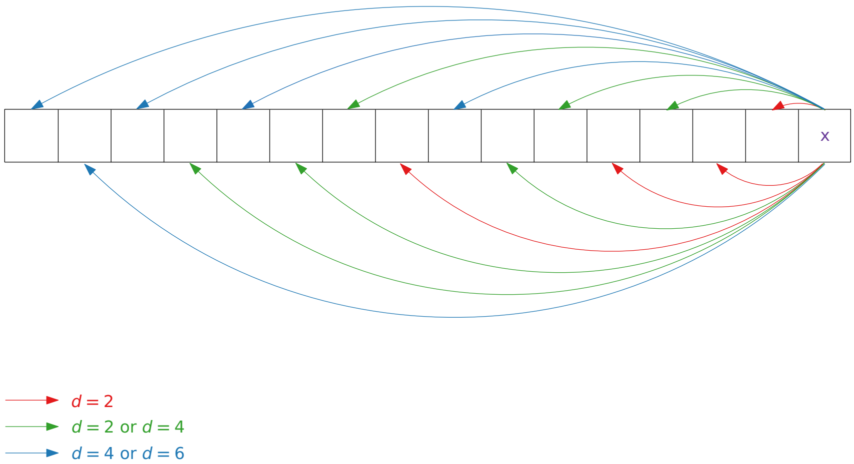


Figure 3.19: Graphical representation of distance constraints for codeword 16.

and output sequences from the conventional 4B6B RLL code. Therefore the spectral efficiency of 0.67 bit/s/Hz as well as the maximum run-length of $k = 4$ remain the same for both codes. The original and the proposed 4B6B codes can be viewed as (d, k) RLL codes where $d = 0$ and $k = 4$.

Since there exist 20 balanced codewords of length 6 and only 16 are used for the 4B6B code, the M^2 metric could be further improved by considering codewords different from the original 4B6B codebook. This can surely improve the error correction performance of the proposed 4B6B code but the flicker-free property will be impacted.

3.4 Results and Analysis

3.4.1 Error Correction Performance

For simulation purposes, an AWGN channel was considered with the OOK modulation. The analytical performance of the BER and FER for the uncoded proposed 4B6B are highlighted in (3.5) and (3.6) respectively [59],

$$BER = 0.6 \times \sum_{w} w \times \frac{A_w}{K} \times erfc\Big(\sqrt{d \times R \times E_b/N_0}\Big) \tag{3.5}$$

Table 3.10: Original vs. proposed 4B6B code.

Input	Original 4B6B	Improved 4B6B
0000	001110	001101
0001	001101	010101
0010	010011	011001
0011	010110	010011
0100	010101	011100
0101	100011	010110
0110	100110	011010
0111	100101	110010
1000	011001	101100
1001	011010	100101
1010	011100	101001
1011	110001	101001
1011	110001	110001
1100	110010	001110
1101	101001	100110
1110	101010	101010
1111	101100	100011

$$FER = 0.6 \times \sum_{d} A_d \times \frac{A_w}{K} \times erfc\left(\sqrt{d \times R \times E_b/N_0}\right) \tag{3.6}$$

where $erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt$ is the *complementary error function*; R is the code rate, E_b/N_0, the energy per bit to noise power spectral density ratio and K the payload length.

In fig. 3.20, the comparison in bit-error rate and frame-error rate is presented between the original and the proposed 4B6B code. Both codes were decoded through the maximum likelihood (ML) decoding. These simulated curves were validated against a number of 10000 frame errors for each E_b/N_0 value. For FER graphs, the original and the proposed 4B6B codes are perfectly similar. In terms of BER, it can be observed that the proposed 4B6B code performs present a gain of at least 1.5 dB compared to the original 4B6B code for a BER higher than 10^{-1}. The gain between these two codes becomes narrow as SNR increases. Even though the FER performance is similar for both the proposed and the original 4B6B codes, the BER metric matters the most in the case of a concatenated scheme.

Fig.3.21 presents the BER and FER of some uncoded DC-free RLL codes for a transmitted length of $N = 30$. The Maximum Likelihood (ML) was considered for decoding block codes and the Viterbi algorithm (VA) for trellis based codes. For fair comparison,

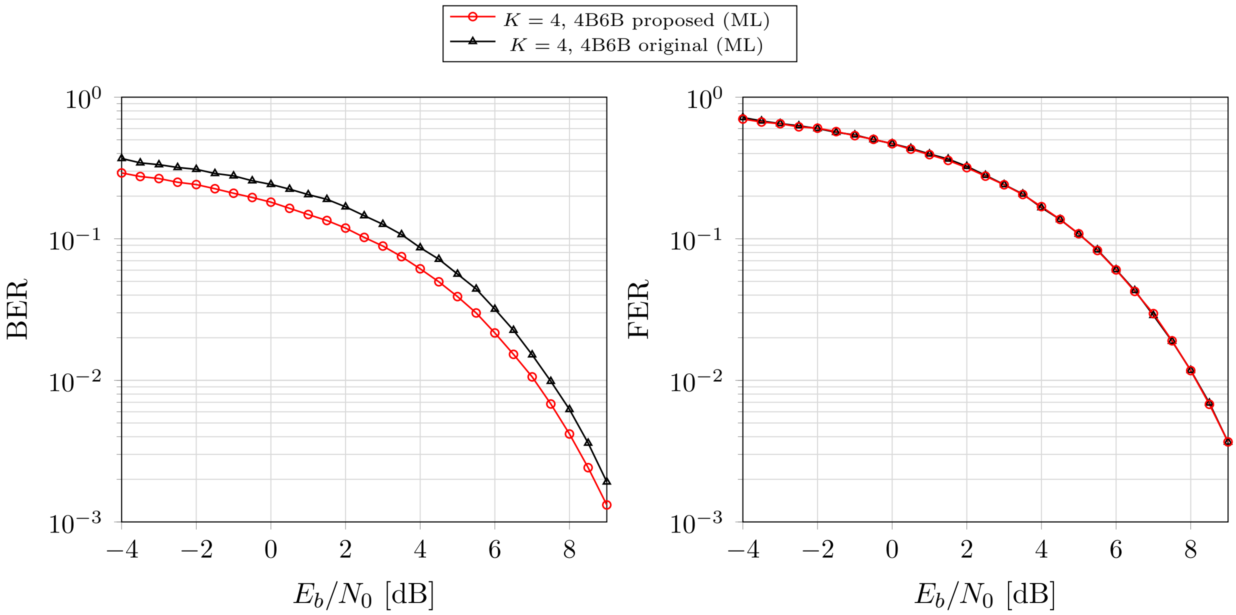


Figure 3.20: Comparison of uncoded 4B6B codes for a transmitted code length of $N = 6$.

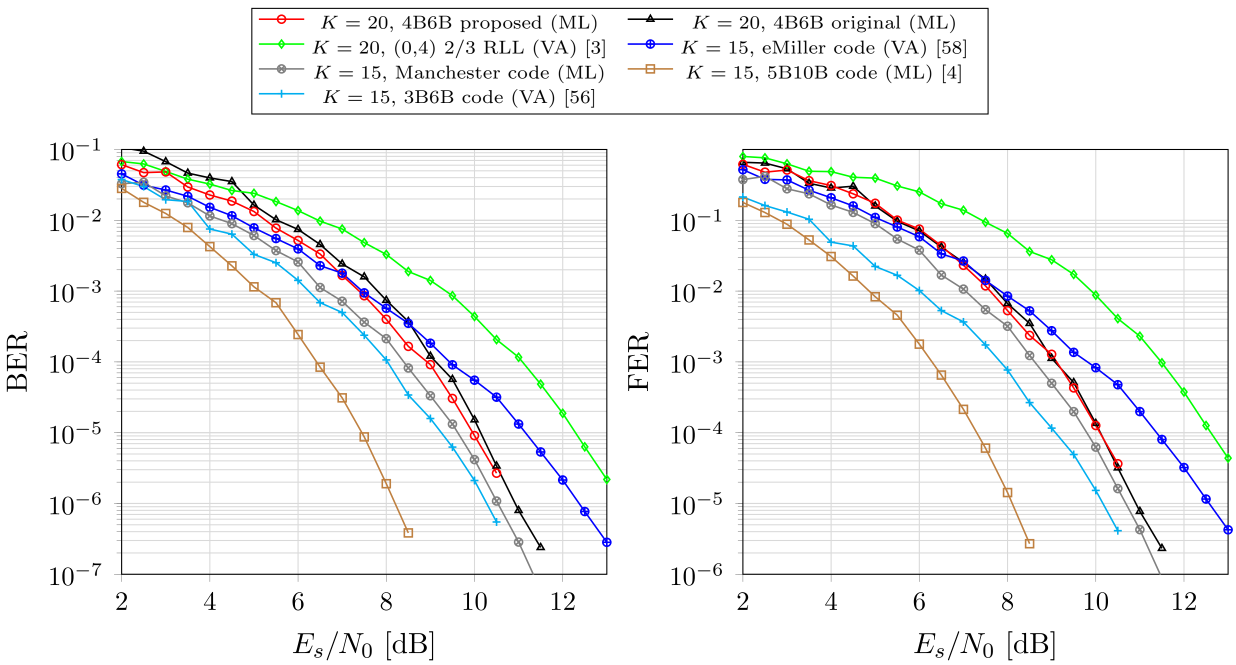


Figure 3.21: Comparison of uncoded 4B6B codes for a transmitted code length of $N = 30$.

we plotted BER and FER vs Es/No, because the normalization Es/No will adjust for extra energy used to send parity symbols and then overall, all codewords will be transmitted with the same amount of energy. For FER graphs, the original and the proposed 4B6B codes are similar at 10^{-4} and present gains (dB) of 1.5, and 2.5 compared to eMiller code [58] and $(0,4)$ 2/3 RLL code [3] respectively. The proposed 4B6B code was compared to the $(0,4)$ 2/3 RLL code [3] because they present the same run-length of 4. In this study, only two spectral efficiencies were considered, 2/3 bit/s/Hz for schemes [3] and 4B6B codes, and 1/2 bit/s/Hz for schemes [4, 56, 58] and the Manchester code. However, the 5B10B code [4], the 3B6B code [56] and the Manchester code outperforms the 4B6B codes at a FER of 10^{-4} by 2.5, 1 and 0.25 respectively. This gain is explained by the fact that these codes have a smaller spectral efficiency than 4B6B code, so they have more parity bits to exploit during the decoding. In terms of BER, at 10^{-5}, the performance of the proposed 4B6B code is slightly better than the original 4B6B code. The proposed work presents gains (dB) of 1 and 2.2 against eMiller code [58] and $(0,4)$ 2/3 RLL code [3] respectively. On the other hand, the 5B10B code [4], the 3B6B code and the Manchester code outperforms the proposed scheme at a BER of 10^{-5} by 2.25, 0.6 and 0.3 dB respectively.

For the FEC-coded RLL schemes, the channel was built as described in Fig. 3.1. Fig. 3.22 shows the performance of various polar coded RLL schemes. The length-matching and the frozen set scheme make use of the standardized 5G polar code [61]. Polar codes are decoded with the successive cancellation (SC) algorithm, trellis based codes are decoded with the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [62] and 4B6B, 5B10B and Manchester codes are decoded with an a posteriori probability (APP) decoder. BCJR algorithm is also an APP decoder.

All referenced schemes were designed with a transmitted length of $N \approx 192$ with overall rates of 1/3. The plots were validated against a number of 100 frame errors for each E_b/N_0 value. For the BER performance, the proposed scheme performs better than all referenced schemes from an SNR of 6 dB. At 10^{-4}, the following gains in dB are recorded 0.3, 0.4, 0.9, and 1.4 against $PC(95,64)+$ 5B10B [4] and $PC(96,64)+$ eMiller [58], $PC(96,64)+$ Manchester, $PC(128,64)+$(0,4) 2/3 RLL [3], and $PC(128,64)+$ conventional 4B6B are recorded respectively. Similarly for the FER, our proposed scheme performance is higher than that of all referenced schemes from SNR of 5.5 dB. At 3.10^{-4}, the following gains (dB) are observed: 0.4, 1 and 1.5 compared to $PC(95,64)+$ 5B10B or $PC(96,64)+$ eMiller or $PC(96,64)+$ Manchester, $PC(128,64)+$(0,4) 2/3 RLL, and $PC(128,64)+$ conventional

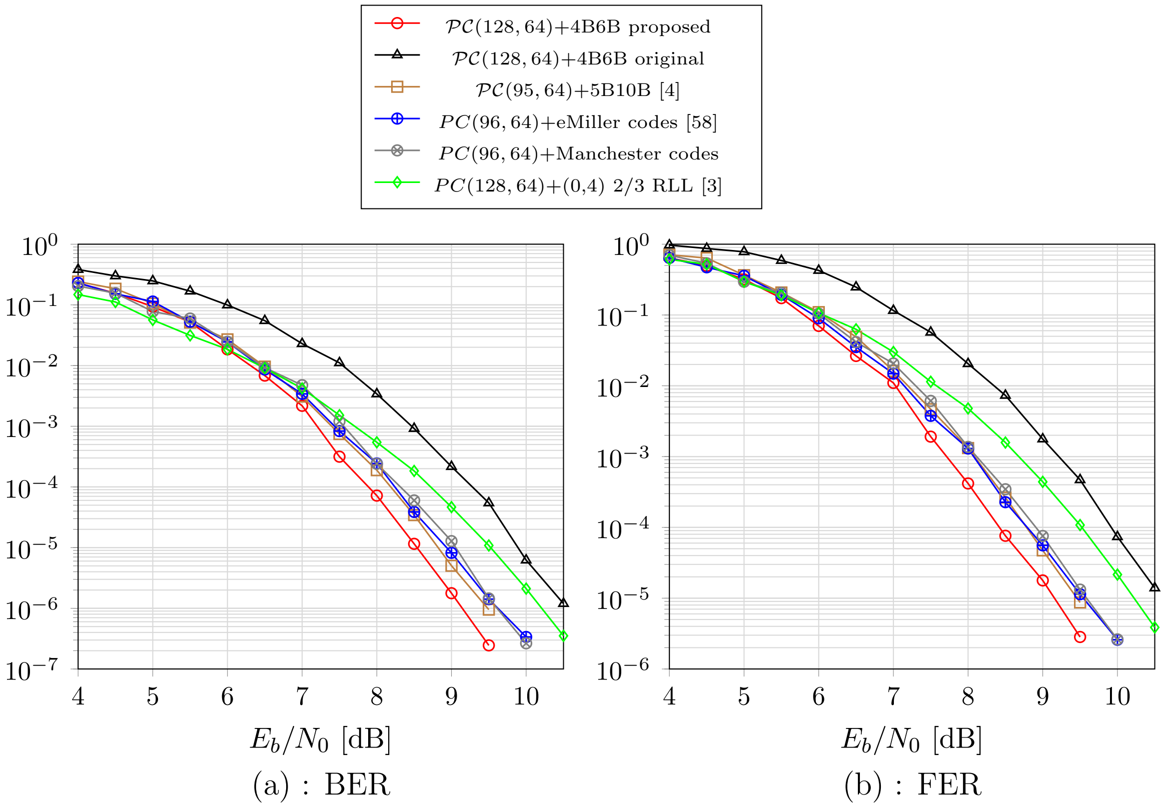


Figure 3.22: Comparison of polar coded RLL codes for a transmitted code length of $N \approx 192$.

4B6B respectively.

The proposed scheme is made of the same codewords as in the original 4B6B code, therefore the same run-length is obtained as well as the computational complexity while the error correction performance is improved. Moreover, the decoding of the 4B6B code can be performed in parallel while trellis based schemes might experience a channel degradation in case of parallelization.

On the other hand, the Viterbi algorithm does not decode well on short length codes compared to ML decoding. In Fig. 3.23, the performances of various convolutional coded RLL codes are plotted. A rate 1/2 convolutional code with the generator $(13, 15)_8$ is considered. The proposed 4B6B codes outer-performs all referenced schemes for both BER and FER analysis. In the case of BER, gains of 0.4, 1.5 and 1.75 dB are observed over $\mathcal{CC}(128, 64)$+ 4B6B original, $\mathcal{CC}(96, 48)$+ 5B10B code [4] and $\mathcal{CC}(128, 64)$+ eMiller code [3] at 10^{-4}. For the FER performance, at 2.10^{-3}, we noted gains of 0.4, 1.4 and 2

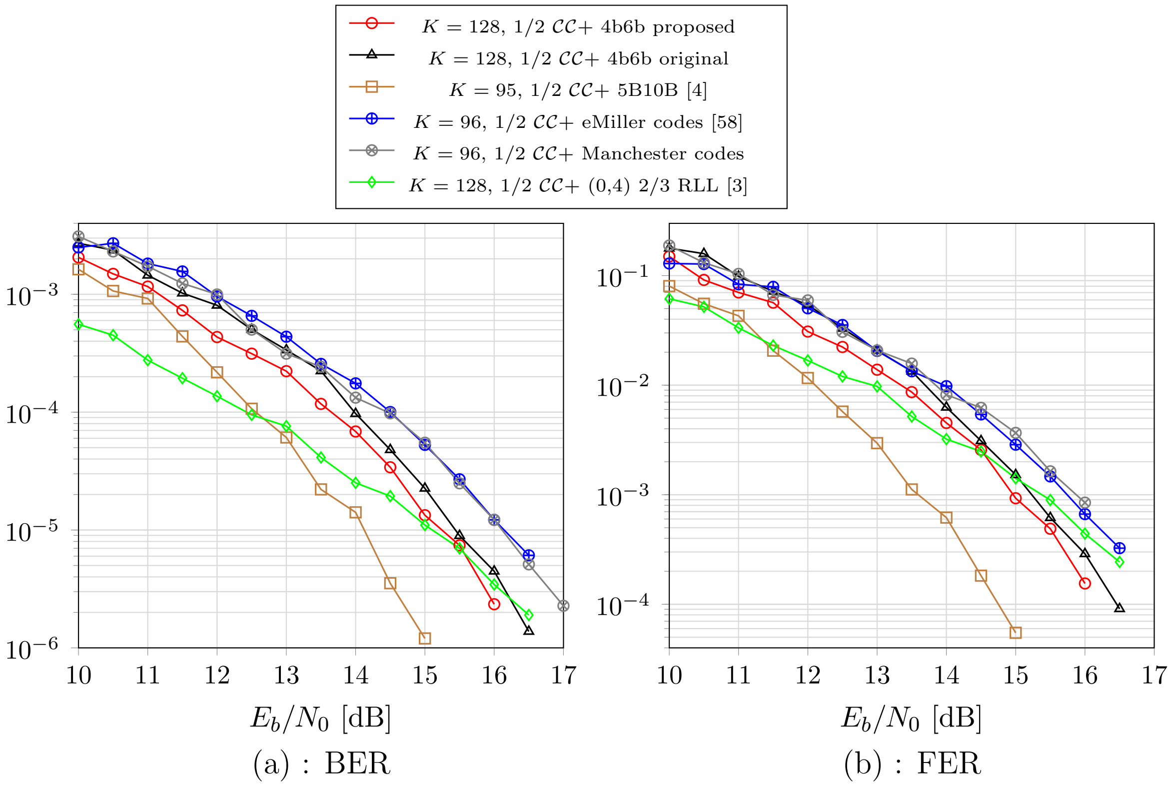


Figure 3.23: Comparison of convolutional coded RLL codes for a transmitted code length of $N = 384$.

Table 3.11: Complexity of decoding DC-free RLL codes.

	# of operations per bit
4B6B code, (APP)	27
5B10B code [4], (APP)	62
$(0,4)$ 2/3 RLL code [3], (BCJR)	122
eMiller code [58], (BCJR)	28
Manchester code, (APP)	1

dB compared to $\mathcal{CC}(128,64)+$ 4B6B original, [4] and [3] respectively.

3.4.2 Computational complexity analysis

Concerning the evaluation of the computation complexity of the proposed scheme compared to the state-of-the-art, we only focus on the complexity of the decoding RLL based on soft inputs and soft outputs (SISO) decoding.

The ML decoding provides the most likely codeword amongst all possible 2^n, n being the data size. This is done by computing the probabilities of all 2^n codewords and then marginal likelihoods for different bit is obtained by the ratio between the sum of probabilities of the $2^n/2$ codewords. For the Viterbi algorithm, the path with the best metric determines the ML codeword. The metric corresponds to the likelihood of the path which is equivalent to the Euclidean distance (soft detection) in AWGN channels between the 2^k inputs; k is memory size of the codes. In case of merging two or more paths, the survivor path is the one with the best metric.

It takes only one subtraction operation to decode the Manchester code.

Coming to the 4B6B code, in order to perform the APP decoding of each 6-bit block using the *max-log* approximation, three steps are required.

- The computation of the probability of each possible codeword is done. It takes 8 additions to determine the first 8 probabilities where each codeword is made of 6 data requiring 5 adders. Additionally, the inversion operation is performed on each of the first 8 probabilities to obtain the remaining 8 ones; this takes 1 adder for each inversion. A subtotal of 48 adders are needed to complete this first step.
- The marginal probability is calculated for each bit. The maximum of 8 data is determined through 7 comparators; given that 8 different maximum values are computed, we obtain a subtotal of 56 comparisons.
- Finally, the 4 logarithmic likelihood ratios (LLRs) are obtained by 4 subtractions.

In total, it takes 108 operations to decode a 6-bit block from a 4B6B code. Then, 27 operations are required per decoded bit. By applying the same reasoning for the 5B10B code, a total of 309 operations are required to decode each 10-bit block, with 62 operations for each decoded bit.

On the other hand, the decoding of trellis code is done per trellis section. The branch and two state (forward and backward) metrics are determined for each section. The addition of these 3 metrics for each potential value of information bit or tuple of information bits provides the output decoding.

The decoding of the $(0, k)$ 2/3 RLL code via *max-log* approximation consists of the following steps:

- The branch metric counts 8 possible codewords transmitted. 4 of them are opposite of 4 others. 3 LLRs should be computed via 2 adders, then 8 adders for each section.

- For the forward metric, all combinations of previous forward and branch metric coming to a node are obtained. It takes 4 adders to combine metrics from the 4 incoming branches; then their maximum is determined through the use of 3 comparators leading to a subtotal of 7 operations to obtain 1 forward metric or 56 operations for the 8 states per trellis section.

- For the backward metric, also 56 operations are required similar to the forward metric. The only difference is that the link between nodes are different from the previous case.

- About the LLR outputs, one trellis section has 32 paths, then 32 sums are performed from 3 inputs giving 64 adders. Then, the maximum of 16 bits is obtained using 15 comparators, for a subtotal of 60 comparators for the 4 cases: when the first bit is 0 or 1, and when the second bit is also 0 or 1.

Therefore, the BCJR decoding of $(0, k)$ 2/3 RLL requires a total of 244 or 122 operations for 2 and 1 decoded bits respectively. Similarly, the BCJR decoding of the eMiller code takes 56 operations for 2 decoded bits.

Table 3.11 depicts the summary on the complexity of decoding DC-free RLL codes. This classification is based on the number of operations required to decode a single bit through APP or BCJR algorithm. Every 2-input operation such as comparison, subtraction or addition is regarded as equivalent.

Although, the Manchester code presents the smallest complexity in terms of the number of operations required per decoded bit, its spectral efficiency is only 0.5 bit/s/Hz against 0.67 bit/s/Hz for the 4B6B code; moreover, its error correction performance is lower than that of the proposed 4B6B code as shown in Fig. 3.22 and Fig. 3.23. The 5B10B code [4] has the higher complexity amongst the referenced schemes, which is 2.3 times larger than that of the proposed scheme. However, though the error correction performance of the 5B10B code is attractive as presented in Fig. 3.23, its complexity is very large and increases with the code length. Additionally, the decoding complexity of the 4B6B code is 4.51 and 1.03 times lower than that of the $(0, 4)$ 2/3 RLL code [3] and the eMiller code [58] respectively. Therefore the proposed scheme shows a good trade-off be-

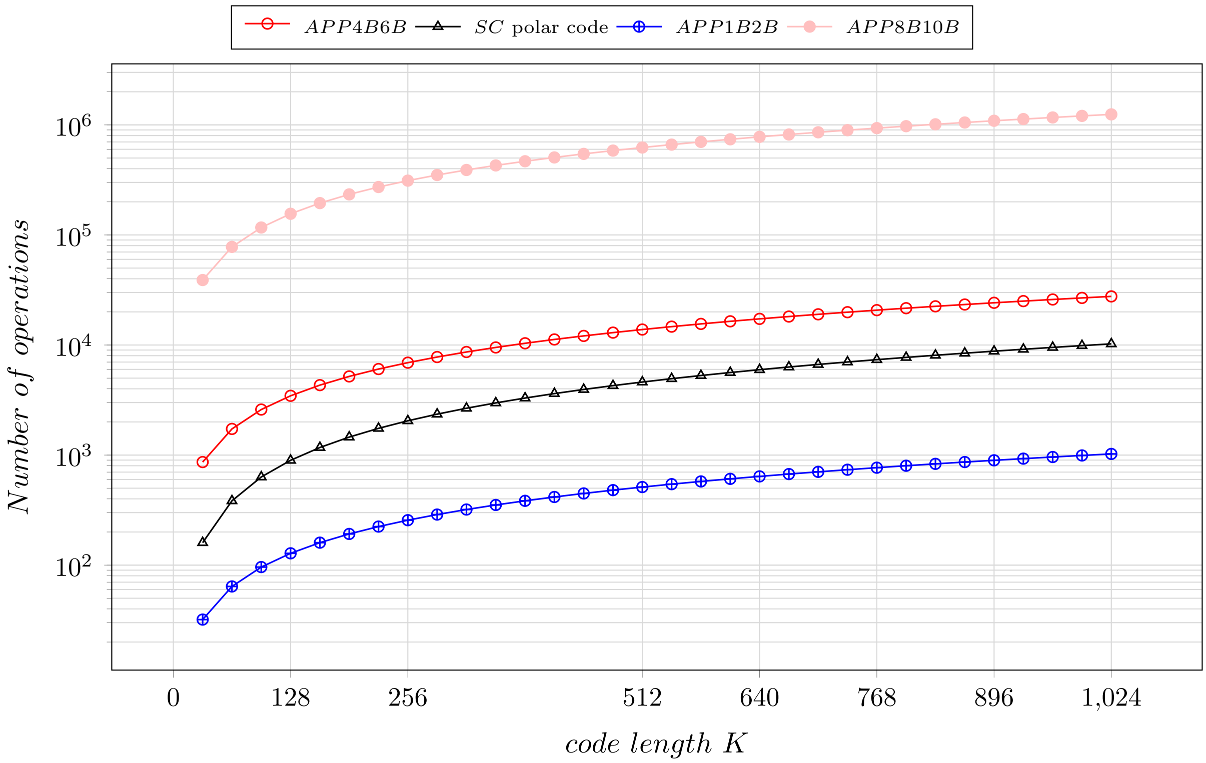


Figure 3.24: Complexity comparison of RLL codes versus SC decoding.

tween decoding complexity and error correction performance compared to state-of-the-art schemes.

Fig. 3.24 shows the complexity comparison of RLL codes versus SC decoding in function of the code length. The SC decoding presents a complexity of $N \log_2 N$ operations, where N is the code length. The Manchester decoding has the lowest complexity. It can be observed that the complexity decoding of 4B6B and 8B10B code is much larger than that of the SC decoding. This can be explained by the fact that RLL codes are non-linear codes which can only be decoded through APP where an exhaustive search is conducted within a codebook.

4 Efficient Flicker-Free FEC Codes Using Knuth's Balancing Algorithm for VLC

The mitigation of flickering in VLC channels depends mostly on DC-free run-length limited (RLL) codes according to literature. Similarly the scheme presented in the previous chapter was also based on RLL codes. However, DC-free RLL codes present some inconveniences such as the use of lookup tables which are memory consuming especially for large length RLL codes like the 8B10B codes which bring an exponential computational decoding complexity. Additionally, the use of RLL codes attenuates the spectral efficiency of the channel by adding unnecessary redundancy.

In this chapter, an efficient construction of flicker-free forward error correction (FEC) codes is proposed to address the issue of channel flickering. This scheme does not make use of lookup tables as most conventional methods. Moreover, the proposed scheme has a decoding complexity that is lower than that of other FEC-coded RLL schemes. Also, the redundancy added is less than referenced methods, this leads to an improvement of the transmission spectral efficiency.

4.1 Introduction

According to IEEE 802.15.7 standards for VLC, the flicker-free property should be enforce in VLC channels by the use of DC-free RLL codes such as 1B2B, 4B6B and 8B10B codes [1]. Since RLL codes are non-linear codes, they are decoded using *a postperiori probability* (APP) decoding which is soft-input soft-output (SISO) decoding that demands relatively high computational complexity.

Due to poor error correction performance of uncoded RLL schemes, FEC codes are usually as outer codes. Modern FEC codes like Reed-Solomon and convolutional codes were proposed in the standards to build efficient FEC-coded RLL schemes for VLC.

However, other FEC codes have been used in VLC systems such as low-density parity-check (LDPC) codes [63], turbo codes [64] and more recently polar codes [65, 66].

In [65], polar codes were used in VLC channels to improve error correction performance. It was reported that the polar encoded of a payload at a rate of half leads averagely to balanced and flicker-free codewords given that the input probability for bit 1 or 0 is half. *Balanced* refers to a state of a codeword having the same number of ones and zeros. In this case, the use of RLL codes to enforce flicker-free progression was omitted. Although, this scheme increases the spectral efficiency over the channel, the balanced state (dimming ratio of 50%) nor the flicker-free property are guaranteed.

Another way to increase the spectral efficiency of a VLC channel is the puncturing of the FEC code in order to reduce the redundancy produced during encoding. In [66, 67], the puncturing of polar codes was used to increase the spectral efficiency as outer codes coupled with RLL codes to enforce flicker-free channel with a dimming ratio of 50%. However, other values of the dimming ratio can be obtained by inserting compensation symbols (CS) at the tail of the polar coded RLL codes. The puncturing is a good tool to achieve rate-matching in VLC systems but, it increases the decoding complexity of the FEC codes. Also the use of CS demands an interleaver in order to impose flicker-free property within CS. The interleaving operation also increases the decoding complexity even though it might slightly improve the error correction performance.

We seek to address the problem of decoding complexity which is also tied to the use of non-linear RLL codes in VLC. We propose an efficient construction of flicker-free FEC codes based on Knuth's balancing algorithm [68]. Knuth's parallel balancing algorithm is very well known and used in constrained coding especially in fields such as magnetic and optical recording devices, detection and correction of unidirectional errors, cable transmissions, and noise attenuation in VLSI systems [26, 69]. Knuth's balancing algorithm is exploited to provides a low redundancy in the proposed scheme, this will improve the spectral efficiency. Also there is no need for RLL codes based on lookup tables that are memory consuming coupled with high decoding complexity. Simulations show that the proposed scheme is better than referenced ones in terms of error correction performance. For an information block length of 256, the proposed scheme is 1.8 dB and 0.9 dB better than that of the regular schemes at the bit error rate of 10^{-6} for a rate of 0.44 and 0.23, respectively.

4.2 Preliminaries

The famous Knuth's parallel balancing algorithm [68] consisting of flipping the first e bits of the binary sequence $\boldsymbol{u} \triangleq [u_0, u_1, \ldots, u_{N-1}]$ of length N (for even N) where $1 \leq e \leq N$. The index e corresponding to the so called balancing point is appended at the tail as the prefix or suffix $\boldsymbol{p}$ of the Knuth encoded codeword $\boldsymbol{x}$. The concatenated codeword $\boldsymbol{xp}$ is sent through the channel. At the receiver, the prefix $\boldsymbol{p}$ corresponding to the balancing point is used to recover the original sequence by flipping back all bits up to the balancing point. If we consider the bipolar alphabet, $\mathbb{A}^2 \in \{-1, 1\}$, the disparity of $\boldsymbol{x}$ is defined as $d(\boldsymbol{x}) = \sum_{i=1}^{N} x_i$, that is the algebraic sum of all bits in $\boldsymbol{x}$. Then, $\boldsymbol{x}$ is called balanced only if $d(\boldsymbol{x}) = 0$.

The *running digital sum* (RDS) computes a partial disparity $d_j(\boldsymbol{x})$ of a sequence over j bits of $\boldsymbol{x}$ where $1 \leq j \leq N$, $d_j(\boldsymbol{x}) = \sum_{i=1}^{j} x_i$, for $i \leq N$. The Knuth's algorithm boils down to splitting the sequence $\boldsymbol{u}$ into two segments where the first one has all bits flipped and the second remains unchanged, $d(\boldsymbol{x}) = -\sum_{i=1}^{e} x_i + \sum_{i=e+1}^{N} x_i$. A balanced codeword will always be achieved by using Knuth's algorithm. This is verified by the fact that a balancing point e is always existing as $d_{j+1}(\boldsymbol{x}) = d_j(\boldsymbol{x}) \pm 2$. The length of the prefix $\boldsymbol{p}$ is $\log_2 N$. The redundancy of the full set of balanced codewords is such that $H = N - \log_2 \binom{N}{N/2}$. The redundancy generated by Knuth's algorithm is almost twice larger than H as N gets larger.

The balancing process according to Knuth's algorithm is shown in (4.1) and (4.2) where the balancing of sequences 10111111 and 100001 is considered respectively.

$$101111 \rightarrow \mathbf{0}01111 \rightarrow \mathbf{01}1111 \rightarrow \mathbf{010}111 \rightarrow \mathbf{0100}11. \tag{4.1}$$

$$100001 \rightarrow \mathbf{0}00001 \rightarrow \mathbf{01}0001 \rightarrow \mathbf{011}001. \tag{4.2}$$

The bold symbols are the flipped bits. The balancing points are 4 (100) and 3 (011) in (4.1) and (4.2) respectively.

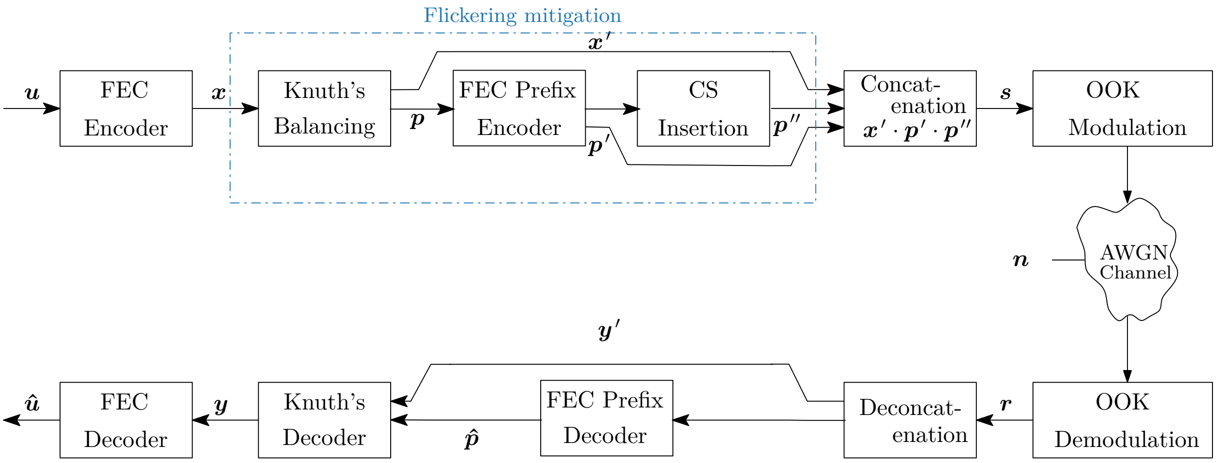


Figure 4.1: The proposed FEC coding scheme.

4.3 Proposed Efficient Flicker-Free FEC code scheme

The proposed scheme aims to produces a flicker-free VLC channel at the dimming ratio of exactly 50%, which corresponds to the average brightness of LEDs.

4.3.1 Scheme description

Fig. 4.1 shows the block diagram of the proposed scheme to achieve efficient flicker-free FEC code. On the transmitter side, the payload $\boldsymbol{u}$ is encoded by a FEC code which produces $\boldsymbol{x}$, then Knuth's encoder is used as inner code to balance $\boldsymbol{x}$ into $\boldsymbol{x}'\boldsymbol{p}$. Because the prefix $\boldsymbol{p}$ is not encoded, a second FEC code is applied only on $\boldsymbol{p}$ which gives $\boldsymbol{p}'$. A insertion of CS is performed in order to generate a balanced encoded prefix $\boldsymbol{p}'\boldsymbol{p}''$. This is done by appending $\boldsymbol{p}'$ to its 1's complement $\boldsymbol{p}''$. Finally the transmitted codeword is the concatenation of the FEC-encoded Knuth codeword $\boldsymbol{x}'$ with the encoded prefix $\boldsymbol{p}'$ and its 1's complement $\boldsymbol{p}''$, $\boldsymbol{x}'\boldsymbol{p}'\boldsymbol{p}''$.

The resulting codeword $\boldsymbol{s}$ is then OOK-modulated. The length of $\boldsymbol{s}$ is such that $S = N + 2p'$, where p' is the length of the encoded Knuth's prefix $\boldsymbol{p}'$ and N, the length of the FEC encoded payload $\boldsymbol{x}$. The rate of this channel is $R = K/(N + 2p')$. The modulated codeword is sent trough a VLC channel modeled as additive white Gaussian noise (AWGN) channel. At the receiver, the sequence $\boldsymbol{r} = \boldsymbol{s} + \boldsymbol{n}$ is received where

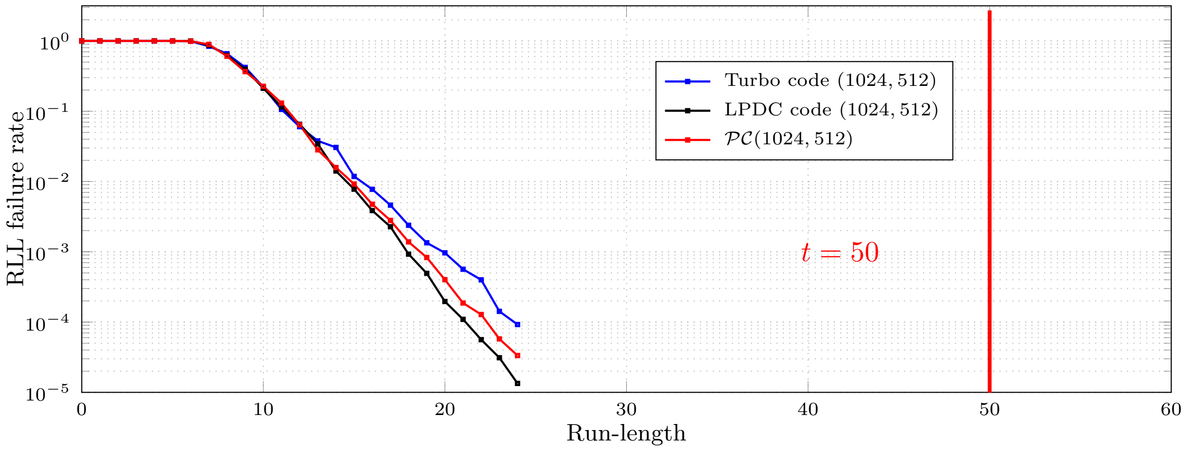


Figure 4.2: Run-length performance comparison for various codes with $R = 0.5$.

$\boldsymbol{n}$ represents the channel noise. Then the de-concatenation operation is performed to decompose $\boldsymbol{r}$ into $\boldsymbol{y}'$ which is the corresponding FEC-coded Knuth's sequence, and $\boldsymbol{p}'$ with $\boldsymbol{p}''$. A soft-input soft-output decision as $p'_i - p''_i$, where p'_i and p''_i are elements of $\boldsymbol{p}'$ and $\boldsymbol{p}''$ respectively. This gives the corrupted version of the encoded Knuth's prefix which goes through FEC decoding then a hard decision is done to obtain the associated index e of the balancing point back. The Knuth's decoder is performed where the sequence $\boldsymbol{y}'$ is flipped back up to e. Finally, another FEC decoding is done to obtain the decoded payload sequence.

Any linear FEC code can be used in the proposed scheme. Fig. 4.2 shows the run-length performance of various codes at a rate of half. Polar, turbo and LDPC codes were considered in this comparison. The run-length performance test consists of estimate the RLL failure rate that is the probability to have codewords with the largest run-length under a threshold. The lenght code of 1024 was set for all referenced codes. Each run-length $l \in \{0, 1, \ldots, 20\}$ was validated against 100 frame errors. The fact that all referenced codes have a decreasing RLL failure rate means that the scheme works for those FEC codes.

Using the lowest optical clock rate of 200 kHz, with the MFTP of 5 ms corresponding to the eye-safe frequency of 200 Hz. We obtain the corresponding run-length as

$$L \times \frac{1}{20010^3} < \frac{1}{200} \rightarrow L = 1000.$$

For a threshold of 20 times less than L, $t = 1000/20 = 50$.

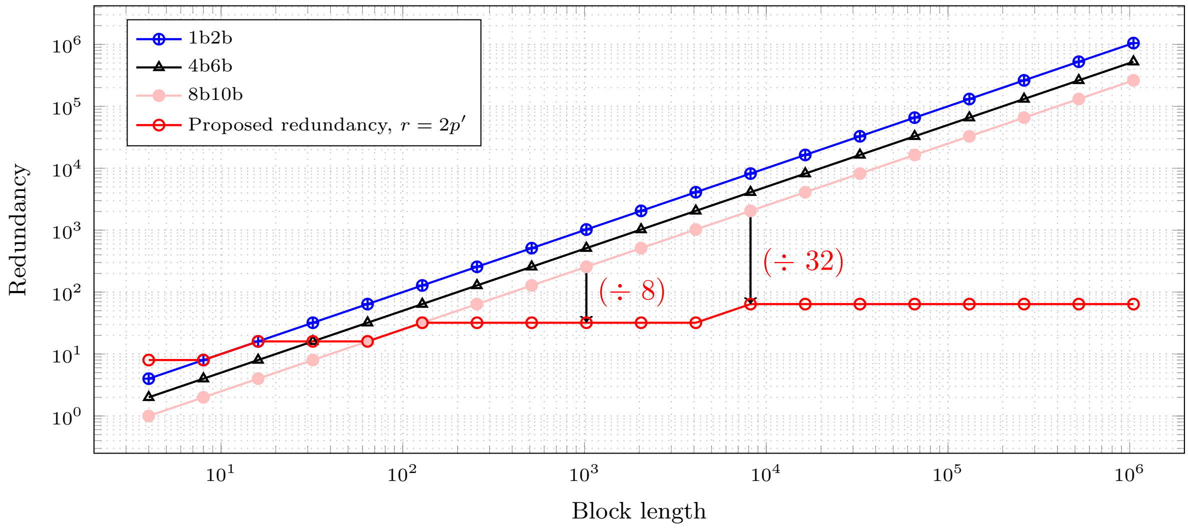


Figure 4.3: Redundancy performance comparison.

4.3.2 Redundancy study

In the objective of limiting power dissipation from LED source transmission in VLC, the amount of additional bits during communication should be reduced. Like stated earlier on, the use on RLL codes as recommended by the VLC standards, brings more redundancy in the channel and then reduces its spectral efficiency. One of the main achievement of our scheme is that a redundancy similar to that of the Knuth's balancing algorithm is obtained.

Fig. 4.3 presents the redundancy versus the block length of the proposed scheme versus 1B2B, 4B6B and 8B10B codes. The 1b2b, 4b6b, and 8b10b RLL codes present redundancies of N, $N/2$, and $N/4$ respectively and the proposed scheme based on Knuth's algorithm has a redundancy of $r = 2p'$. It can be observed that the proposed scheme redundancy plot grows logarithmically whereas those from RLL codes grow linearly. At block lengths of 10^3 and 10^4, the proposed scheme redundancy is 8 and 32 times respectively less than that of the 8B10B code.

4.3.3 RLL analysis

In this section, we study the property of the run-length characteristic of the concatenation of polar encoded Knuth's sequence.

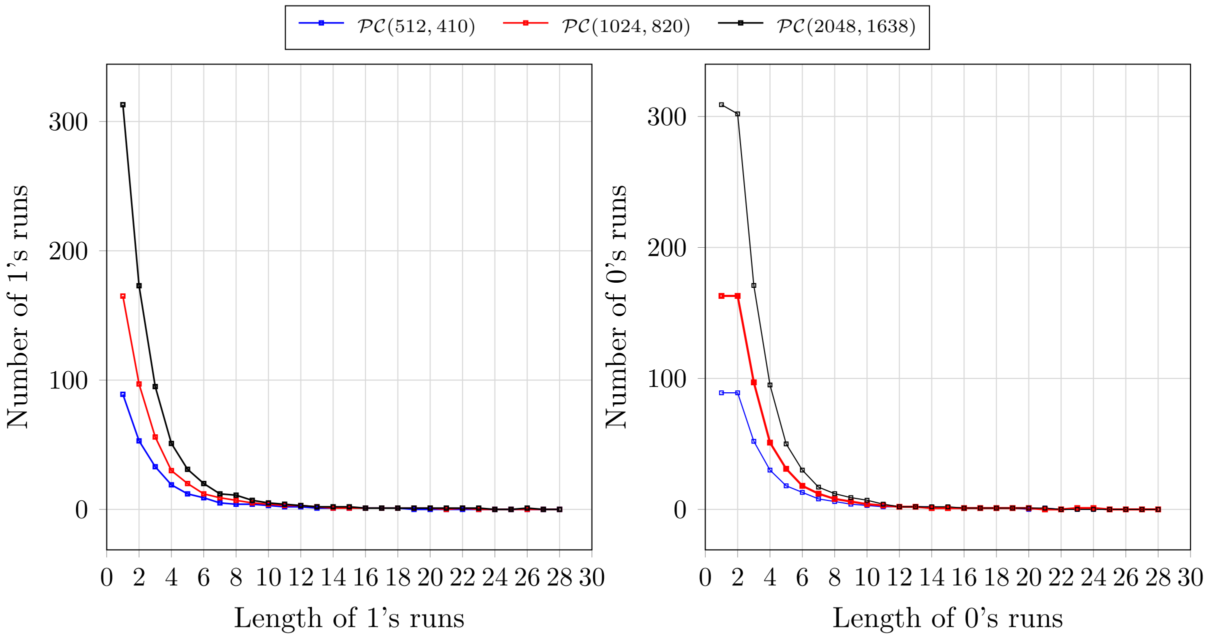


Figure 4.4: Distribution of RLLs at $R = 0.8$.

Fig. 4.4 presents the distribution of run-lengths for sequence encoded sequentially via polar coding and Knuth's balancing algorithm. The information was generated based the equal probability of input bit 1 and 0. The polar code lengths of 512, 1024 and 2048 were considered at a code rate of 0.8. Our analysis is based on the codeword $\boldsymbol{s}$ as per Fig. 4.1. The number runs of ones and zeros were statistically recorded against their lengths as per Fig. 4.4.

For each polar code $PC(512, 410)$, $PC(1024, 820)$ and $PC(2048, 1638)$ of rate 0.8, around 90% of each codeword has a run-length $l < 8$. The highest run of $l = 28$ was obtained for $PC(2048, 1638)$. Even by considering the highest run-length, the flicker-free property of the channel is not perturbed as the human eye does not perceive the light flickering corresponding at that run-length.

In Fig. 4.5, the RLL failure rate is estimated against the run-length for the same polar codes as in Fig. 4.4. The RLL failure rate decreases with large run-lengths, this is in accordance with results from Fig. 4.4. It is observed that the RLL failure rate for the highest run-length of $l = 28$ as reported in Fig. 4.4 is around 10^{-7}.

The lowest optical clock of 200 kHz for VLC systems corresponds to a time period

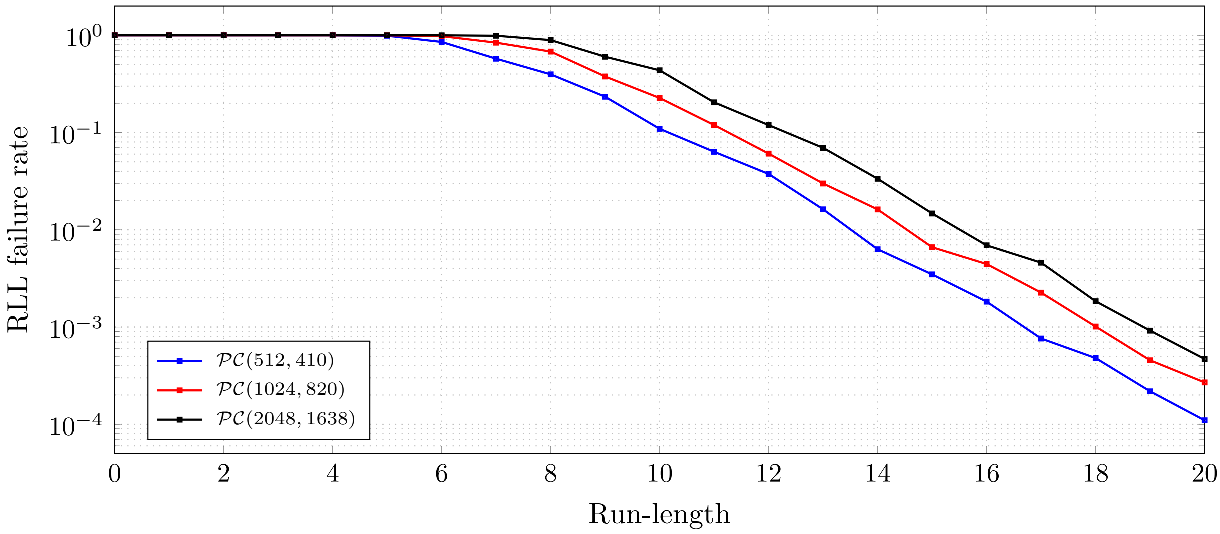


Figure 4.5: Run-length performance at $R = 0.8$.

of $28 \times 1/200000 = 0.14$ ms. This is equivalent to a switching frequency of $1/0.14\text{ms} = 7143$ Hz. The time period of 0.14 ms is very far below the MFTP of 5 ms. In other words, the switching frequency of 7143 Hz is much higher than the eye-safe clock rate of 200 Hz.

Considering the use of a 15 MHz LED bandwidth as recommended in VLC standards, we obtain a switching period of $28/(15 \times 10^6) \approx 1.87$ μs which corresponds to a frequency of $1/(1.87 \times 10^{-6}) \approx 535.7$ MHz.

4.3.4 Error correction performance

This section informs on the error correction capability of the proposed scheme which comprises the theoretical FER and the simulated FER and BER results.

4.3.4.1 Theoretical FER

The proposed scheme is the concatenation of an outer FEC code with Knuth's code. The correct decoding of the payload depends a lot on the prefix decoding. That is, in case an error happened on the prefix, the overall outcome will be wrong independent of the FEC coding. For the scope of this thesis, both decoders for the inner and outer codes are based on polar code successive cancellation (SC) decoding. However, any liner FEC code used would confirm the results in this report. The polar code was chosen because the theoretical prediction of SC decoding is achievable which is not the case for most other linear codes. This allows us to optimize the code rate of the inner code for a better decoding performance.

The theoretical FER of the proposed scheme in case of polar coding used both as inner and outer codes and decoded through SC algorithm is given by:

$$FER = \Big(1 - \prod_{i \in \mathcal{I}_2}(1 - \boldsymbol{Q}_i)\Big) + \prod_{i \in \mathcal{I}_2}(1 - \boldsymbol{Q}_i) \cdot \Big(1 - \prod_{i \in \mathcal{I}_1}(1 - \boldsymbol{Q}_i)\Big), \tag{4.3}$$

with $\mathcal{I}_1$ and $\mathcal{I}_2$ being the information sets respectively for the main polar code and for the polar code for the prefix encoding.

4.3.4.2 Simulation results

The robustness of the proposed scheme relies a lot of the protection of the prefix generated by the Knuth's algorithm. Figs. 4.6 and 4.7 presents two situations where the impact on the importance of prefix coding is highlighted.

In Fig. 4.6, a $PC(128, 64)$ of rate half was considered as FEC outer code, and the prefix of 7-bit was encoded through 3 different ways: $PC(32, 7)$, $PC(16, 7)$ and $PC(8, 7)$. It can be observed that the $PC(128, 64) + PC(32, 7)$ and $PC(128, 64) + PC(16, 7)$ have a quasi-similar performance for BER and FER plots. This shows that the performance of the prefix coding does not depend on the amount of parity bits involved in the coding process. So it is advised to use the $PC(16, 7)$ instead of $PC(32, 7)$ in order to improve the spectral efficiency throughout. However, the $PC(128, 64) + PC(8, 7)$ has a lower performance than other referenced schemes. At 10^{-3}, gains of 1 dB and 1.5 dB were recorded between the $PC(128, 64) + PC(32, 7)$ or $PC(128, 64) + PC(16, 7)$ against the $PC(128, 64) + PC(8, 7)$ respectively in terms of BER and FER.

In Fig. 4.7, a $PC(64, 48)$ was considered as outer FEC code with a rate of 3/4. The 6-bit prefix was encoded through 3 different ways as in Fig. 4.6 that is $PC(32, 6)$, $PC(16, 6)$ and $PC(8, 6)$. It can be observed that they all have the same performance both in FER and BER. Therefore, it is appropriate to choose the $PC(8, 6)$ in order to maintain

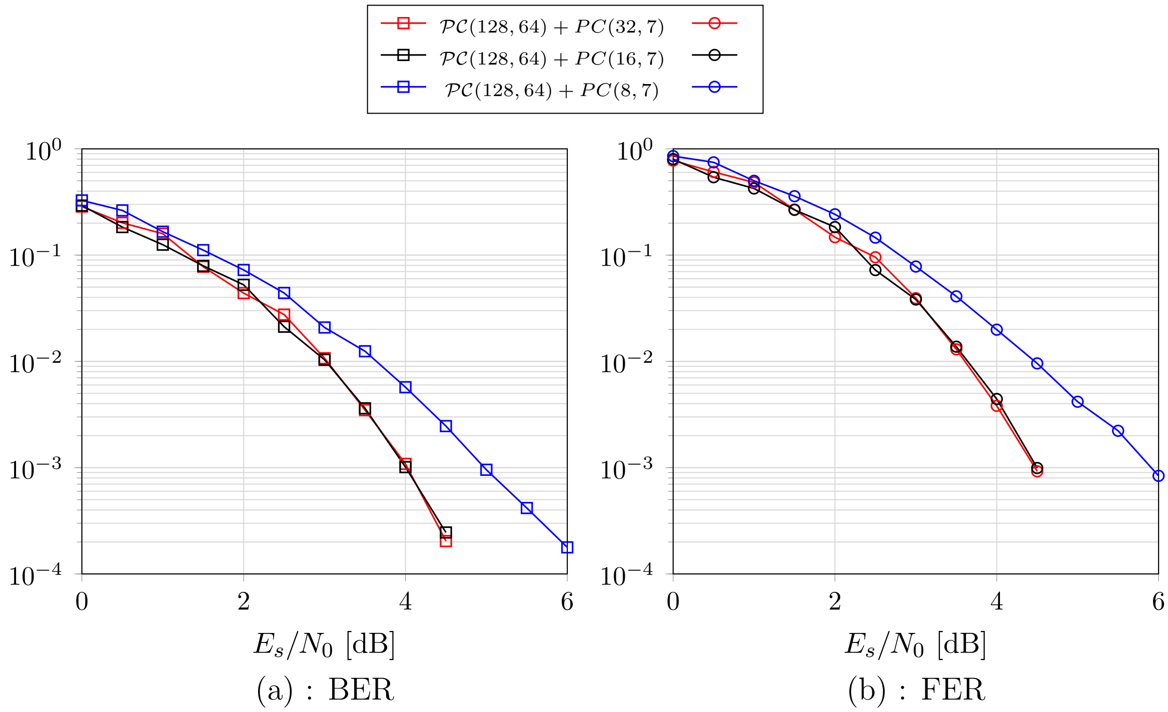


Figure 4.6: Impact of prefix coding, $PC(128, 64)$, rate of 1/2.

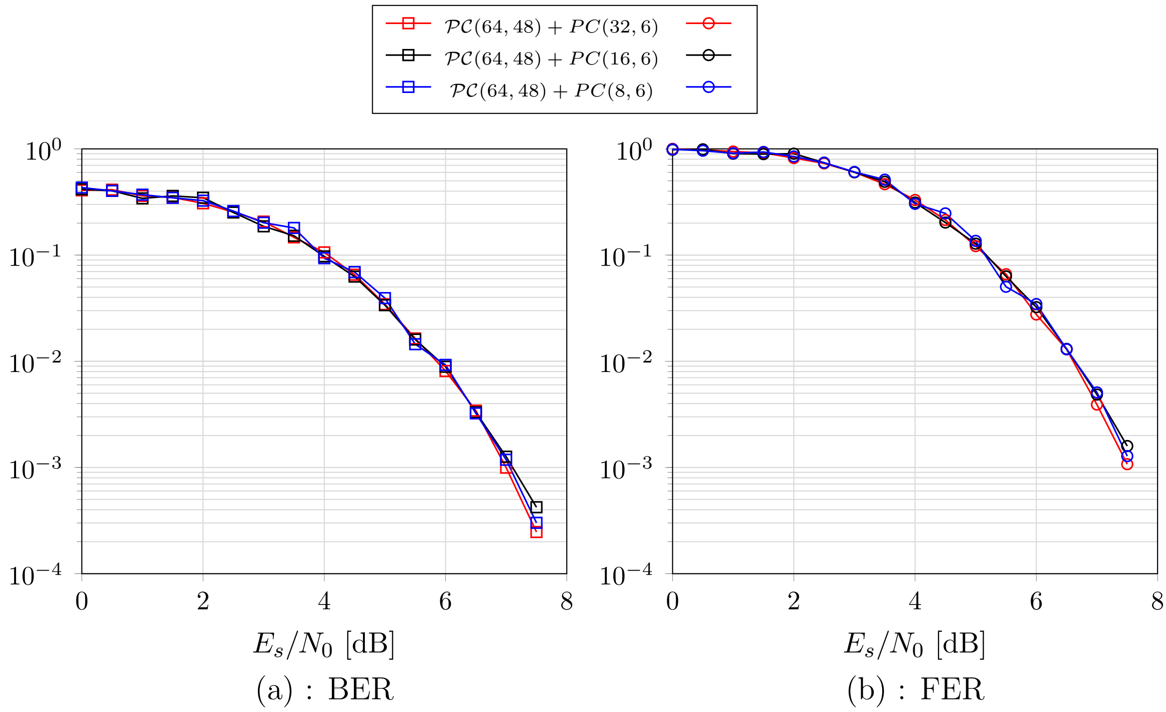


Figure 4.7: Impact of prefix coding, $PC(64, 48)$, rate of 3/4.

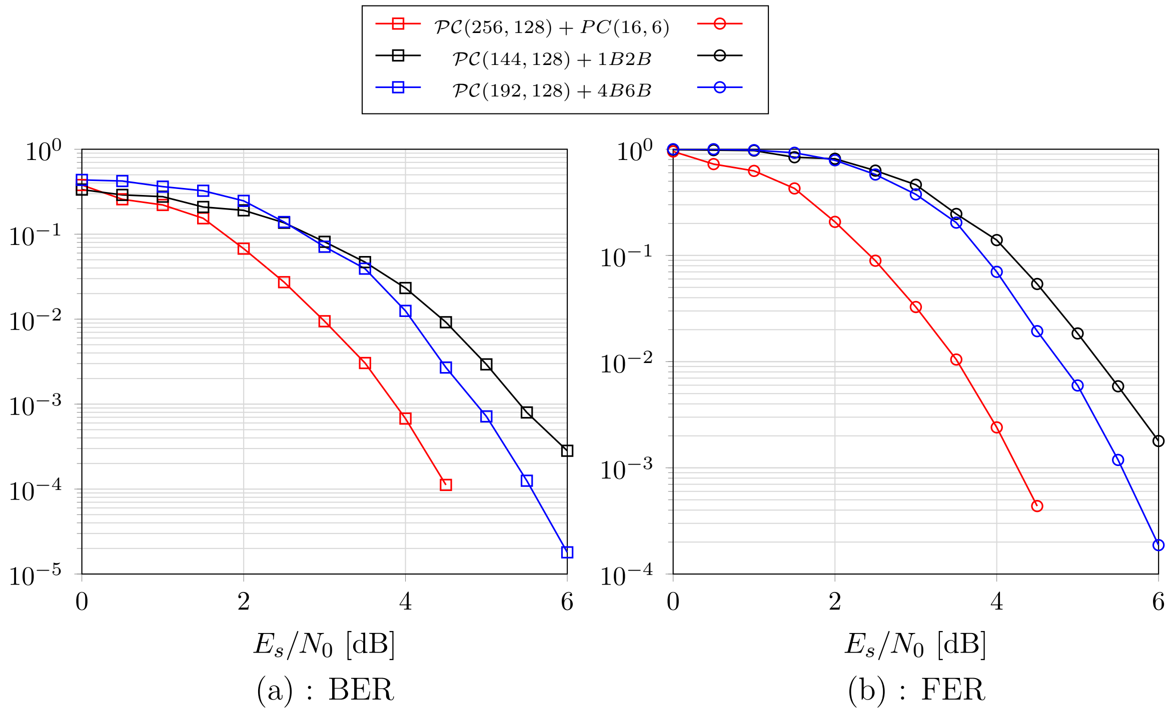


Figure 4.8: Comparison of various schemes at 50% dimming ratio and rate 4/9.

a relatively high spectral efficiency.

Fig. 4.8 shows the comparison in FER and BER of the proposed scheme against the a polar code concatenated with the 1B2B and 4B6B code distinctly, at a rate of 4/9 for a 50% dimming ratio. The soft-input soft-output APP algorithm was used for the decoding the 1B2B and 4B6B codes. At the BER of 10^{-3}, the proposed scheme presents gains in dB of 1 and 1.5 against the $PC(192,128)+4B6B$ and the $PC(144,128)+1B2B$ respectively. And in terms of FER, at 10^{-2}, the proposed scheme has gains in dB of 1.2 and 1.9 against the $PC(192,128)+4B6B$ and the $PC(144,128)+1B2B$ respectively.

The use of RLL codes reduces significantly the spectral efficiency in VLC channel. The 1B2B and 4B6B codes have rates of 1/2 and 2/3 respectively whereas Knuth's algorithm has a rate of $N/(N+\log_2 N)$ which converges to 1 as N get larger.

Table 4.1: Number of operations required for decoding RLL codes and the proposed scheme for $K = 128$ as per Fig. 4.8.

	$R = 4/9$
1b2b	144
4b6b	7776
Proposed	336

4.3.5 Computational complexity

In section, we evaluate the computational complexity brought by the proposed scheme and compare it to that of the referenced schemes. This evaluation is done under the assumption that, the outer FEC code which is the polar code in our case, takes soft information. Therefore the decoding complexity of the proposed scheme is mainly computed from Knuth's balancing process.

The decoding of the 1B2B or Manchester code takes N subtractions, where N is the length of the block code. The decoding of the 4B6B code was detailed in chapter 3, section 3.4. It is based on the APP decoding which is performed through 2 main stages. In the first stage, the 16 probabilities for each potential codeword is calculated. This requires 8 additions to obtain the 8 first probabilities where each, is computed by adding 6 data using 5 adders. The remaining 8 probabilities are obtained by considering the inverse on the first 8. The second stage consists of estimating marginal probabilities for each bit.

The computational complexity of the proposed scheme can be evaluated via 3 steps.

- Firstly, the Manchester code which was used to balanced the encoded prefix, is decoded through p' subtractions.
- Then, the polar code used to protect the prefix $\boldsymbol{p}$ is decoded, this requires $M \log_2 M$ operations, where $M = 2^{\lceil \log_2 p' \rceil}$.
- Finally, the decoding of Knuth's algorithm is performed. this is done by flipping th obtained LLR values up to the block length N. This step takes up to N distinct comparison operations.

Table 4.1 presents the number of operations required for the decoding of some RLL codes versus the proposed scheme for $K = 128$ as per Fig. 4.8. We treated any 2-input operations such as addition, subtraction, comparison or maximum as equivalent. It can be observed that the 4B6B code requires the largest number of operations. This decoder might be impractical in a soft-input soft-output decoding situation due to the high level of complexity involved. On the other hand, the 1B2B code presents the smallest number of operations but it reduces significantly the spectral efficiency of the channel as low as below 50% versus 66.6% for the 4B6B code. The decoding complexity of the proposed scheme is 2.33 times higher than that of the 1B2B code.

Overall, the proposed scheme can be practically feasible based on the relatively low complexity involved in the decoding process. Also, it presents a good trade-off between the error correction performance and the spectral efficiency optimization.

5 Efficient Dimming Scheme based on Non-DC Free RLL Codes for VLC

A reliable VLC channel should include dimming control which aims to maintain a flicker-free communication at several levels of lighting brightness. According to some existing works, the dimming control can be imposed in the VLC channel by the use of compensation symbols (CS) insertion. This are purely redundant bits of ones or zeros appended to the transmitted codeword in order to adjust its overall weight. Then, an interleaver is recommended to create a flicker-free channel by scrambling bits to avoid long runs of zeros or ones. Since the insertion of CS deteriorates the spectral efficiency, puncturing is often used on the outer FEC code to increase the data throughput.

In this chapter, we propose a new family of non-DC free RLL codes to achieve dimming control in VLC. We focused our study on the family of the 4B7B code which can provide dimming ratios of 2/7 and 5/7. The proposed 4B7B code presents a good balance between the error correction performance and the spectral efficiency optimization for VLC channels.

5.1 Introduction

Designing reliable VLC channels embedded with dimming control remains a challenge in VLC. The IEEE 802.15.7 standards for VLC recommends the use of run-length limited (RLL) codes coupled with the insertion of compensation symbols (CS) in on-off keying (OOK) modulation. However, several schemes have been proposed from the literature to achieve dimmable VLC channels.

A family of rate-compatible punctured convolutional (RCPC) codes obtained through a wise code search was presented to achieve wide range of brightness [70]. A dimming range of $[10\% - 90\%]$ was experimented, but the performance significantly decreases as the dimming factor diverges from 50% due to the increase number of punctured bits. Moreover, this scheme relies on the searching of good puncturing patterns which increases its decoding complexity.

In [63], quasi cyclic (QC) low density parity check bits (LDPC) codes were proposed as FEC codes to control dimming in VLC channel. It was reported that this scheme presents a low complexity the encoding and decoding processes as well as an easy hardware implementation. Also, it makes use of an interleaver to mitigate flickering and the CS combined with RLL codes for dimming control.

In [67], polar codes as used as outer FEC codes concatenated with RLL codes to produces a 50% dimming ratio and flicker-free channel. Then, CS and puncturing techniques are involved to perform the dimming control within the VLC channel. In [66], the same scheme as [67] was proposed but only the CS insertion was recommended to achieve a dimming target as puncturing degrades the error correction performance of the channel although it improves on the spectral efficiency.

A scheme similar to [67] was also presented in [65], but without the use of RLL codes. This scheme is based under the assumption that polar encoded codewords have a dimming ratio averagely around 50% given that the input probability of bit 1 or 0 is half. However, puncturing and CS insertion was used to achieve dimming control.

A 4B5B code was proposed in [5] for dimming control. However, it does not provide constant dimming as the weight amongst codewords are different. It is crucial to observe consistency in the brightness of the LEDs during communication.

In [6], a method was presented based on CS addition using nearest neighbour. This scheme uses the CS as parity bits that contribute on improving error correction performance. The nearest neighbour CS addition is based on 4B6B code. The length of the CS is obtained from the target dimming and the CS are integrated in the 4B6B code using some Hamming distance properties rules.

In general, many schemes [6,65–67] from literature, achieves a dimmable VLC channel through encoding the input data by a modern FEC code then, a RLL code is applied to obtain a 50% dimming. Furthermore, the insertion of CS coupled with some puncturing techniques are performed to achieve a certain dimming ratio.

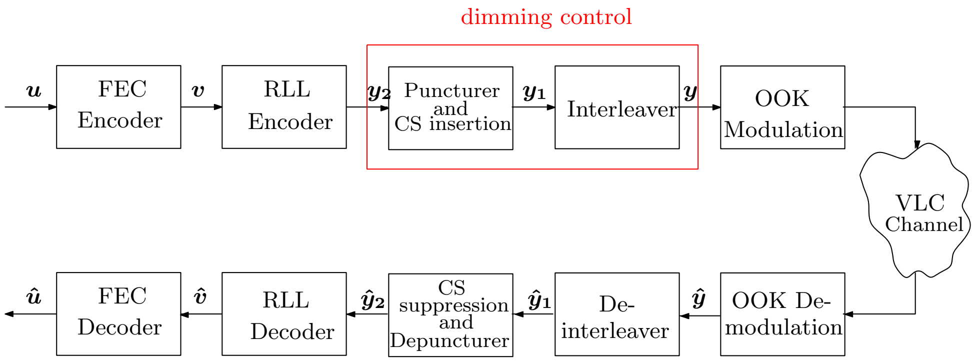


Figure 5.1: Conventional dimming control system in VLC

But these schemes presents many drawbacks such as the decrease of the channel spectral efficiency due to high redundancy brought by CS insertion. A low error correction performance due to the use of puncturing that increases decoding complexity; also the CS are useless redundancy that are thrown away during the decoding process.

A new family of RLL based on various weights is presented in this chapter. A wide range of dimming ratio values could be achieve while improving on the channel error correction performance and the spectral efficiency compared to existing schemes.

The remaining of the chapter is outlined as followed. Section 5.2 presents some preliminaries and previous works related to the current topic. The proposed dimming scheme is described in section 5.3, followed by results and discussion in section 5.4.

5.2 Preliminaries

Fig. 5.1 shows the conventional method of achieving dimming control in VLC channels. This block diagram is almost similar to that the flicker-free VLC channel described in chapter 4. The RLL block is used to enforce a dimming ratio of 50% on the FEC-encoded data. The *a posteriori probability* (APP) algorithm is used for the decoding of RLL codes which it is a soft-input soft-output decoding as RLL codes are non-linear codes. The dimming control is made of puncturing bits and CS insertion combined with the interleaving operation. The puncturing or shortening techniques can be used for rate-matching which improves the channel spectral efficiency while the CS insertion is performed to enforce a certain dimming ratio in the transmitted codeword. It is assumed

Table 5.1: $4B5B$ code [5].

Hex	Input bits	Codewords
0	0000	01001
1	0001	01010
2	0010	01011
3	0011	01101
4	0100	01110
5	0101	01111
6	0110	10011
7	0111	10101
8	1000	10110
9	1001	10111
A	1010	11001
B	1011	11010
C	1100	11011
D	1101	11101
E	1110	11110
F	1111	11111

that the dimming ratio is set at a upper layer and that the sender as well as the receiver are aware about that value.

In [5], the following 4B5B code was proposed as presented in Table 5.1. It can be observed that the codewords have different weight values. This produces a non-uniform illumination, that is a variable dimming for every transmission. This is not desired in VLC channel where a dimming control should be enforced in transmission. Therefore, although this 4B5B code presents some attractive error correction performance, it does not achieve dimming control.

Let l be the length of CS, γ the dimming value, N_2 the length of the FEC-coded RLL $\boldsymbol{y}_2$ as per Fig. 5.1. The length of the transmitted codeword is $N = N_2 + l$.

- If $\gamma > 50\%$, $l = \frac{N_2(\gamma-0.5)}{1-\gamma}$. CS are composed of all 1s.
- If $\gamma < 50\%$, $l = \frac{N_2(0.5-\gamma)}{\gamma}$. CS are composed of all 0s.

In [6], an improved RLL code was proposed based on the 4B6B code. Given l is the length of CS which all 1s or 0s codewords depending on the dimming value. Let $N = \lceil l/2 \rceil$, l is splitting into two sequences of lengths $S = \lceil N/2 \rceil$ and $T = N - S$. CS

Table 5.2: RLL codes for different γ [6].

Input	4B6B ($\gamma = 0.5$)	4B8B ($\gamma = 0.625$)	4B12B ($\gamma = 0.75$)
0000	001110	11001110	110011101111
0001	001101	00110111	001101111111
0010	010011	11010011	111111010011
0011	010110	01011011	010111111110
0100	010101	11010101	011101110111
0101	100011	10001111	101111110011
0110	100110	11100110	111110011011
0111	100101	10010111	111011011101
1000	011001	11011001	111101100111
1001	011010	01101011	110111101110
1010	011100	01110011	011111111100
1011	110001	11110001	111100011111
1100	110010	11001011	110010111111
1101	101001	10100111	101011111101
1110	101010	11101010	101110111011
1111	101100	11101100	111111101100

are inserted into the existing 4B6B code in such a way to obtain the largest Hamming distance between complementary codewords pairs as described (5.1).

$$\begin{array}{ccccccc}
T & 00 & 11 & 10 & S & & \\
00 & 11 & 01 & T & S & & \\
T & S & 01 & 00 & 11 & & \\
01 & 01 & T & S & 10 & & \\
T-1 & 01 & P & 01 & P & 01 & S-1 \\
10 & S & T & 00 & 11 & & \\
S & 10 & 01 & 10 & T & & \\
S-1 & 10 & P & 01 & P & 01 & T-1
\end{array} \tag{5.1}$$

(5.1) shows the encoding rulesthe first 8 4-bit words, the other 8 are obtained by complementing all the previous codewords except T, P and S. Also, for $N = 1$, N CS is appended on either side alternatively.

Table 5.2 presents some RLL codes proposed in [6] based on the conventional 4B6B codes for dimming values of 0.625 and 0.75.

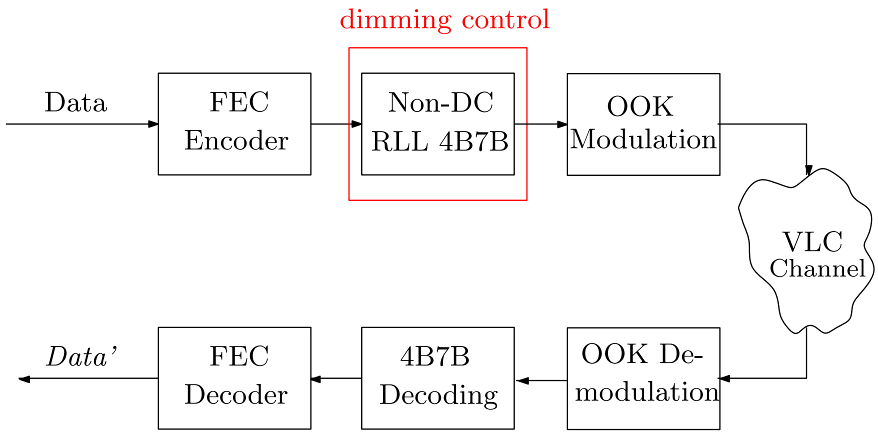


Figure 5.2: Proposed dimming control system for VLC

5.3 Proposed dimming scheme

Our proposed scheme is described by Fig. 5.2. The use of CS neither puncturing or interleaving is used in the proposed scheme as it is the case for conventional dimming schemes. Our approach is in line with that from [6] where the generated redundancy is used in the RLL decoding to improve on the error correction performance instead of the insertion of CS.

Our scheme consists of generating new families of non-DC RLL codes based on existing ones. As depicted in Fig. 5.3, from the conventional DC-free RLL code at the left of figure: 1B2B, 4B6B, 6B8B and 8B10B. We add one parity bit on each of these codes to propose: 1B3B, 4B7B, 6B9B and 8B11B codes respectively.

The Pascal's triangle shows the weights distribution of binary codebooks. For example, the weight distribution of a 6-bit codebook is highlighted at the 7^{th} row. It reads that, there are 20 codewords of weight 4, then 15 with weights 3 or 5, 6 of weights 2 or 6 and 1 with weights 1 or7, for a cardinality of $2^6 = 64$.

In Fig. 5.3, the numbers in brackets indicate weights that can be achieved. For instance, the 4B7B code with weights distribution as per line 8 on Fig. 5.3 can generate codebooks of weights 2, 3, 4 and 5.

Table 5.3 shows the families of proposed codes with corresponding dimming ratios. One can observe that the spectral efficiency of these codes are slightly greater than those from conventional RLL codes.

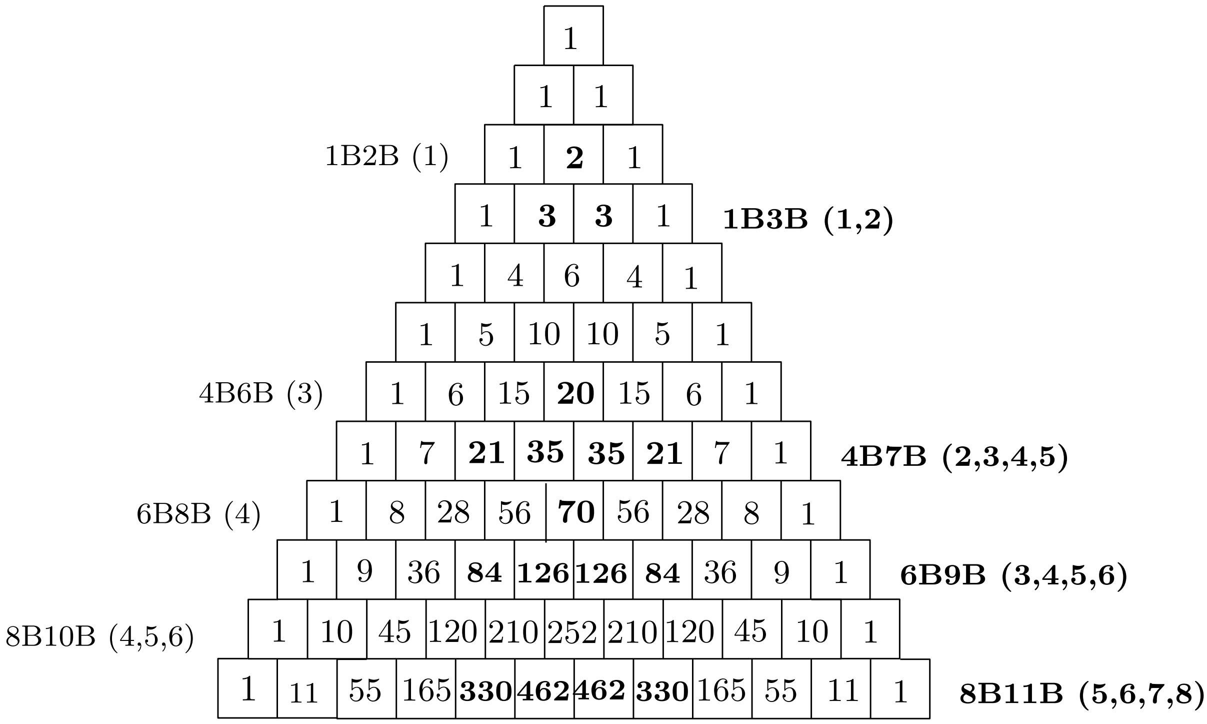


Figure 5.3: Pascal's triangle

Table 5.3: Summary of proposed codes with dimming ratios.

Code	Weights	Dimming (%)	Spectral Efficiency (bit/s/Hz)
$1B3B$	$(1,2)$	$\{33.3, 66.6\}$	0.33
$4B7B$	$(2,3,4,5)$	$\{28.5, 42.8, 57.1, 71, 4\}$	0.57
$6B9B$	$(3,4,5,6)$	$\{33.3, 44.4, 55.6, 66.7\}$	0.67
$8B11B$	$(5,6,7,8)$	$\{45.5, 54.5, 63.6, 72.7\}$	0.73

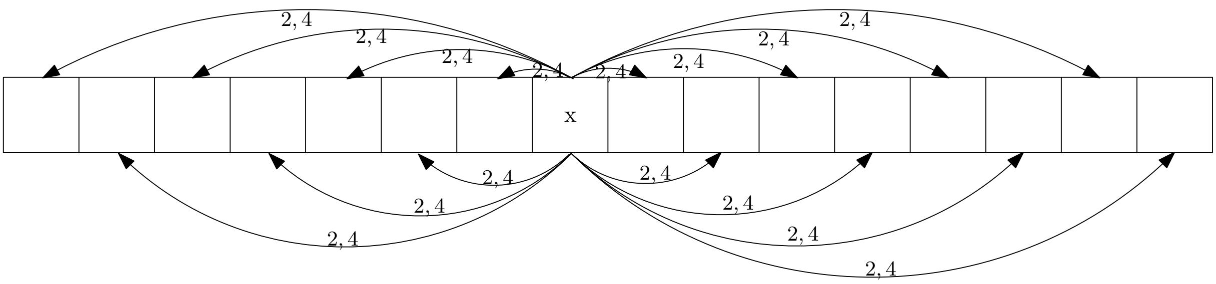


Figure 5.4: Graphical representation of distance constraints of the proposed code.

For the scope of this study, we focus on the 4B7B code with weights 2 and 5, denoted by $4B7B(2)$ and $4B7B(5)$.

Table 5.4: Proposed $4B7B(2)$ code.

4B	7B	4B	7B	4B	7B	4B	7B
0000	0001001	0100	0010001	1000	1000001	1100	0100001
0001	0001010	0101	0010010	1001	1000010	1101	0100010
0010	0001100	0110	0010100	1010	1000100	1110	0100100
0011	0011000	0111	0110000	1011	1001000	1111	0101000

There are 21 codewords of the $4B7B$ code with weight 2, that are S_2 = {3, 6, 9, 10, 12, 17, 18, 20, 24, 33, 34, 36, 40, 48, 65, 66, 68, 72, 80, 96}; and also 21 with weight 5, S_5 = {31, 47, 55, 59, 61, 62, 79, 87, 91, 93, 94, 103, 107, 109, 110, 115, 117, 118, 121, 122, 124,}. Amongst the 21 codewords from S_2 and S_5, 16 are chosen in each set in such a way to have a maximum run-length lesser or equal to 6. The following codebooks were obtained; for weight 2, $P_2 = \{9, 10, 12, 17, 18, 20, 24, 33, 34, 36, 40, 48, 65, 66, 68, 72\}$; for weight 5, $P_5 = \{31, 47, 55, 59, 61, 62, 79, 87, 91, 93, 94, 103, 107, 109, 110, 115, 117, 118, 121, 122, 124\}$.

Furthermore, there are $16! = 2.09228 \times 10^{13}$ different codebooks that can be obtained from codewords from P_2 and P_5. The optimal $4B7B(2)$ and $4B7B(5)$ codes in terms of error correction performance can be obtained through a search based on distance constraints amongst codewords.

The same method based on [59] and [60] are used as in chapter 3 to find those optimal codebooks. The aim is to get the lowest metric M^d as defined in (3.3). A distance of 2 or 4 between codewords should be observed while mapping them to 4-bit words at distance of 1, 2 or 3 distinctly. Fig. 5.4 presents a graphical representation of the constraints to be imposed amongst codewords in order to obtain the optimal $4B7B(2)$ and $4B7B(5)$ codes.

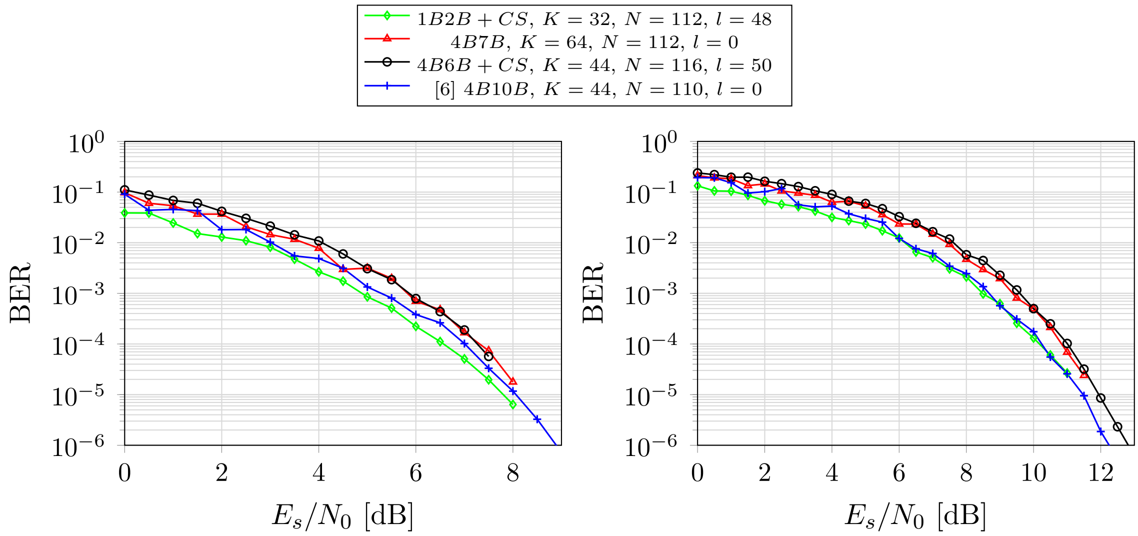
The optimal $4B7B(2)$ and $4B7B(5)$ codes in terms of error correction for concatenated schemes have a similar IOWEF as per defined in (3.1) given by

$$\begin{aligned} B(W,D) = {} & D^0(W^0) \\ & + D^2(4W^1 + 2.625W^2 + 0.625W^3 + 0.125W^4) \\ & + D^4(3.375W^2 + 3.375W^3 + 0.875W^4). \end{aligned} \tag{5.2}$$

Tables 5.4 and 5.5 presents the optimal $4B7B(2)$ and $4B7B(5)$ codes for concatenated schemes based on error correction performance. The minimum metric $M^2 = 11.625$ was

Table 5.5: Proposed $4B7B(5)$ code.

4B	7B	4B	7B	4B	7B	4B	7B
0000	0110111	0100	1010111	1000	1100111	1100	1001111
0001	0111011	0101	1011011	1001	1110011	1101	1101011
0010	0111101	0110	1011101	1010	1110101	1110	1101101
0011	0111110	0111	1011110	1011	1110110	1111	1101110



(a) : dim. $2/7 \approx 28.5\%$

(b) : dim. $5/7 \approx 71.4\%$

Figure 5.5: Uncoded RLL codes comparison, $N \approx 112$.

obtained for both codes. This metric can be further reduced if we considered the relaxation on the length-run property.

5.4 Simulations and Results

In this section, we present the simulations of the proposed scheme against some conventional ones. Furthermore, their decoding complexity is discussed.

The simulation aims to compare uncoded RLL codes. The proposed schemes namely $4B7B(2)$ and $4B7B(5)$ codes are compared against the 1B2B, 4B6B and 4B10B [6] codes for dimming values of about 30% and 70% as depicted in Fig. 5.5. The transmitted length is set as $N \approx 112$. The CS insertion was included in 1B2B and 4B6B codes for

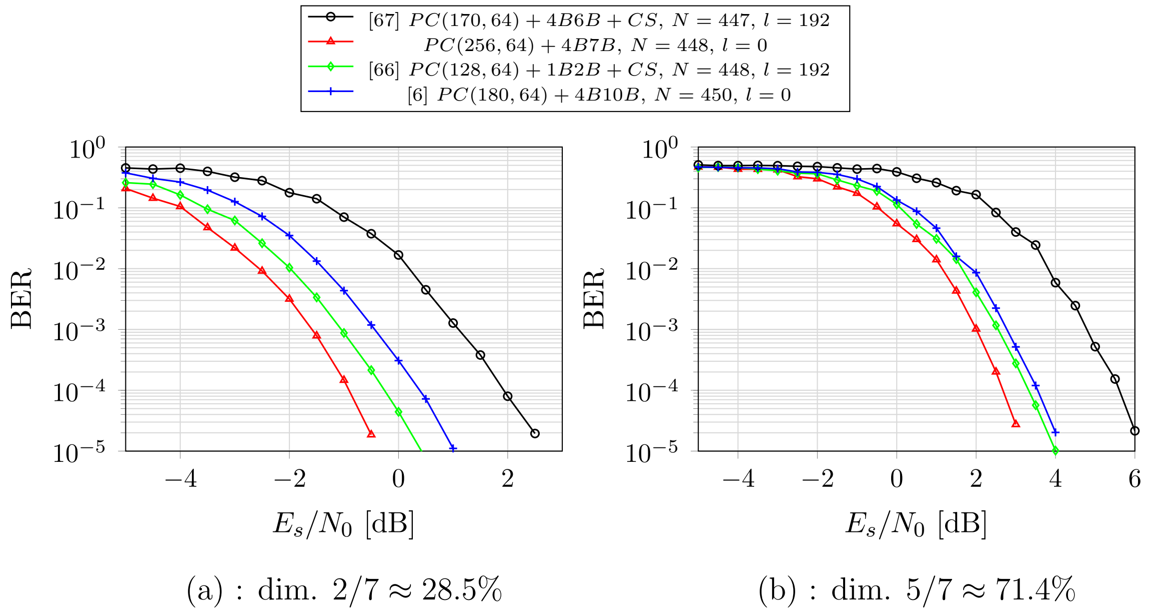


Figure 5.6: Polar encoded RLL codes comparison, $N =$ and rate of $1/7$.

rate-matching purposes. The *a posteriori probability* algorithm is used to decode RLL codes. The BER was derived in function of the energy per symbol E_s/N_0 for the sake of fair comparison. The normalization on energy per symbol aims to adjust for extra energy used to send parity symbols as the referenced schemes have slightly different transmission rates. It can be observed that the 4B6B code and the proposed code present the same performance for BER below 10^{-3} at dimming values of 28.5% and 71.4%. Similarly, the 1B2B and the 4B10B [6] codes have almost the same BER performance for both dimming values. These are slightly better than the proposed 4B7B codes. At BER of 10^{-4}, there is a 0.7 dB gap between the 4B10B code [6] and the proposed codes for both dimming values.

The FEC coded RLL codes were simulated to observe the impact of the proposed codes in the concatenated architecture. Fig. 5.6 shows the comparison amongst various polar encoded RLL codes for dimming values of about 30% and 70%. Note that polar codes were decoded through the successive cancellation (SC) algorithm. The proposed $4B7B(2)$ and $4B7B(5)$ codes outperforms referenced schemes for both dimming values. For the 30% dimming value, gains in dB of 0.8, 1.3 and 3 are observed between the proposed scheme, $PC(256,64)+4B7B(2)$ and schemes $PC(128,64)+1B2B+CS$ [66], $PC(180,64)+4B10B$ [6] and $PC(170,64)+4B6B+CS$ [67] respectively at a BER of 10^{-4}. In the case of 70% dimming value, gains in dB of 0.6, 0.7 and 2.6 are recorded

Table 5.6: Complexity of decoding DC-free RLL codes.

APP decoding	# of operations per bit
4B6B code [1]	27
4B7B code	29
4B10B code [6]	35
1B2B code [1]	1
8B10B [1]	33345

between the proposed scheme $PC(256, 64) + 4B7B(5)$ against [6, 66, 67] respectively at a BER of 10^{-4}.

The proposed 4B7B codes offers attractive error correction performance while mitigating the channel flickering at a relatively low frequency. Table 5.6 shows the decoding complexity of the proposed 4B7B code against other RLL codes. All 2-input operations such as addition, multiplication or comparison are regarded as equivalent. The complexity is computed in terms of the number of operations required to decode those codes. The decoding complexity of the proposed 4B7B code is almost similar to that of the existing 4B6B code; moreover, it is 1.2 times less than that of [6] and 115.34 times less complex than the decoding of the 8B10B code as per IEEE VLC standards [1].

6 Conclusion

The rationale that motivated this research was based on investigating new methods to achieve efficient flicker-free and dimmable VLC channels. From chapter 3 to 6, we have attempted to bring some elements of response to this rationale.

6.1 Achievements

In chapter 3, a new class of improved 4B6B codes was introduced. Some metrics based distance profile performance was proposed in order to obtain the most efficient 4B6B code in terms of error correction capability for concatenated schemes. This was a simple permutation of the conventional 4B6B code, therefore they both have the same run-length properties, spectral efficiencies as well as decoding complexities. It was reported that the proposed 4B6B code outperforms state-of-the-art techniques based on BER and FER for concatenated schemes. Considering the case of polar coded RLL codes, gains of at least 0.3 dB were observed at a BER of 10^{-4}. Moreover, a 1.4 dB gain was recorded between the proposed and the standardized 4B6B codes.

In chapter 4, a scheme was presented to generate flicker-free FEC codes for VLC channels at a dimming ratio of 50%. This method was based on the Knuth's parallel balancing algorithm where the generated prefix is encoded as inner codes instead of conventional RLL codes. It was reported that a good protection of Knuth's prefix improve the overall BER performance. Also, theoretical results were derived to confirm the simulation. The proposed scheme does not make use of RLL codes and reduces significantly the amount of redundancy brought by RLL codes; in other words, the spectral efficiency of the channel is increased. It is well known that the use of some RLL codes are memory consuming coupled with complex APP decoding as RLL codes are not linear codes. Furthermore, this scheme is attractive as it offers a competitive error correction performance and decoding complexity compared to most existing schemes especially RLL based schemes. The polar code was used FEC codes for the scope of this work because, it offers the possibility of easily derive the theoretical FER values which is not straightforward for some other linear codes. The proposed scheme is flexible in the sense that any linear FEC code can be verified the reported results.

In chapter 5, a scheme was introduced to achieve flicker-free dimming for VLC channels. Conventional ways of performing dimming in VLC channels include puncturing or shortening, CS insertion, interleaving, DC-free RLL coding which present some drawbacks. The puncturing aims to reduce the redundancy generated by the outer FEC code; this will affect the error performance performance of the FEC code. The CS insertion significantly deteriorates the spectral efficiency of the channel by adding run of zeros or ones to reach the dimming ratio target which is set at a upper layer. The interleaving scrambles the channel in order to break long runs of ones or zeros which create flickering. However, the use of interleaving increases the decoding complexity of the channel. Most conventional dimming schemes in VLC make use of DC-free RLL codes to ensure a 50% dimming ratio prior to achieve the dimming ratio target. This is because, the FEC-coded payload does generate codewords with variable dimming ratios, but since the fixed-length coding is considered, the length of the transmitted codeword should be fixed and known at the decoder. New families of non-DC free RLL codes were introduced where the 4B7B codes of weight 2 and 5 were investigated as study case for dimming ratios of about 30% and 70%. It was reported that the proposed codes outperforms the referenced schemes based on FER and BER comparisons. Given a polar coded RLL code, gains of at least 0.6 dB were reported at a BER of 10^{-4} for both dimming values. Also a gain of about 3 dB was observed between the polar coded 4B7B against 4B6B with CS for both dimming values at a BER of 10^{-4}.

6.2 Future Work

In chapter 3, the 4B6B code was optimised to present outstanding error correction performance for concatenated schemes. This involved defining metrics based on distance profile. A similar study could be perform on the 8B10B code which is recommended as per VLC standards. It is well known that the conventional 8B10B code is mainly used in magnetic recording devices, therefore some work have to be conducted for its adaptability in VLC. Although the 8B10B code brings a high decoding complexity, it significantly improves the spectral efficiency of the channel by adding only 2 redundant bits for 8-bit information. This study would aim to obtain the most optimal 8B10B code for concatenated schemes based on some predefined distance metrics.

In chapter 4, The encoded prefix from Knuth's algorithm is balanced through a Manchester or 1B2B code, which includes a 50% data rate in the encoded prefix. The balancing of the encoded prefix could be further improved by considering high-rate RLL codes such as 4B6B or 8B10B codes. This will further improve the channel spectral efficiency at the cost of a slightly increased decoding complexity. Also, the polar code was the unique FEC code considered in this study based on the successive cancellation (SC) decoding. It could be on great interest to improve the protection of the prefix by considering most sophisticated decoders such as SC list or SC flip. Moreover, the insertion of cyclic redundant check (CRC) bits in the prefix encoding could also significantly improves the error correction performance of this scheme.

In chapter 5, the efficient dimming scheme was investigated only for two fixed dimming ratios that are 30% and 70% based on the $4B7B(2)$ and $4B7B(5)$ codes respectively. However, an efficient dimmable VLC channel should be flexible in the sense that a large range of dimming values must be achievable. A large family of codes was proposed but not explored. The proposed $4B7B$ codes can also achieve weights of 3 and 4 or dimming ratios of 43% and 57% respectively. Furthermore, $1B3B$, $6B9B$ and $8B11B$ codes achieve dimming values in percentage (%) of $\{33, 67\}$, $\{33, 44, 56, 67\}$ and $\{45, 55, 64, 73\}$ respectively. Therefore a more flexible dimmable VLC channel could be build by combining different codes. A short redundancy could be appended at the tail of the transmitted codeword to inform of the codes used as well as their orders. It is assumed that dimming values are set an upper layer and are both known at the sender and receiver.

References

[1] "IEEE standard for local and metropolitan area networks–part 15.7: Short-range wireless optical communication using visible light," *IEEE Std 802.15.7-2011*, pp. 1–309, Sept 2011.

[2] A. X. Widmer and P. A. Franaszek, "A DC-balanced, partitioned-block, 8B/10B transmission code," *IBM J. Research Dev.*, vol. 27, no. 5, pp. 440–451, Sept. 1983.

[3] Z. Li, H. Yu, B. Shan, D. Zou, and S. Li, "New run-length limited codes in OnOff keying visible light communication systems," *IEEE Wireless Communications Letters*, vol. 9, no. 2, pp. 148–151, 2020.

[4] V. A. Reguera, "New RLL code with improved error performance for visible light communication," *arXiv preprint arXiv:1910.10079*, 2019.

[5] K. Kim and al., "Appropriate RLL coding scheme for effective dimming control in VLC," *Electronics Letters*, vol. 52, no. 19, pp. 1622–1624, Sept. 2016.

[6] T. Uday, A. Kumar, and L. Natarajan, "Improved run length limited codes for VLC using dimming control compensation symbols," in *2018 10th International Conference on Communication Systems Networks (COMSNETS)*, 2018, pp. 457–460.

[7] A. G. Bell, "Apparatus for signalling and communicating, called photophone," U.S. Patent 235 199, Dec. 7, 1880.

[8] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423, Jul. 1948.

[9] Y. Tanaka, S. Haruyama, and M. Nakagawa, "Wireless optical transmissions with white colored LED for wireless home links," in *11th IEEE International Symposium on Personal Indoor and Mobile Radio Communications. PIMRC 2000. Proceedings (Cat. No.00TH8525)*, London, UK, 18–21, 2000, pp. 1325–1329.

[10] T. Komine and M. Nakagawa, "Integrated system of white LED visible light communication and power-line communication," *IEEE Transactions on Consumer Electronics*, vol. 49, no. 1, pp. 71–79, Feb. 2003.

[11] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using led lights," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 100–107, 2004.

[12] M. Kavehrad and P. Amirshahi, "Hybrid MV-LV power lines and white light emitting diodes for triple-play broadband access communications," *IEC Comprehensive Report on Achieving the Triple Play: Technologies and Business Models for Success*, pp. 167–178, 2006.

[13] Siemens, "500 megabits/second with white LED light," Press Release, January 2010, https://web.archive.org/web/20120929044003/http://www.siemens.com/innovation/en/news/2010/500-megabits-second-with-white-led-light.htm.

[14] TMCNET News, "St. cloud first to sign on for new technology," Press Release, November 2010, https://www.tmcnet.com/usubmit/2010/11/19/5148608.htm,.

[15] Harald Haas, TEDGlobal 2011, "Wireless data from every light bulb," Press Release, July 2011, https://www.ted.com/talks/harald_haas_wireless_data_from_every_light_bulb.

[16] J. Grubor, K.-D. L. S. Randel, and J. W. Walewski, "broadband information broadcasting using LED-based interior lighting," *Journal of lightware technology*, vol. 26, no. 24, pp. 3883–3892, Dec. 2008.

[17] L. Zeng, D. C. O'Brien, H. L. Minh, G. E. Faulkner, K. Lee, D. Jung, Y. Oh, and E. T. Won, "High data rate multiple input multiple output (MIMO) optical wireless communications using white LED lighting," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 9, pp. 1654–1662, 2009.

[18] A. H. Azhar, T. Tran, and D. O'Brien, "Demonstration of high-speed data transmission using MIMO-OFDM visible light communications," in *2010 IEEE Globecom Workshops*, Miami, FL, USA, Dec.6–10, 2010, pp. 1052–1056.

[19] W. Kim, C. Bae, S. Jeon, S. Pyun, and D. Cho, "Efficient resource allocation for rapid link recovery and visibility in visible-light local area networks," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 524–531, 2010.

[20] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *Proceedings of ICC '93 - IEEE International Conference on Communications*, vol. 2, 1993, pp. 1064–1070,.

[21] R. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, pp. 21–28, 1962.

[22] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3051–3073, 2009.

[23] B. Marcus, R. Roth, and P. Siegel, *An introduction to coding for constrained systems*, 5th ed., Oct 2001.

[24] C. Pielli, . Stefanovi, P. Popovski, and M. Zorzi, "Joint compression, channel coding, and retransmission for data fidelity with energy harvesting," *IEEE Transactions on Communications*, vol. 66, no. 4, pp. 1425–1439, 2018.

[25] M. Calvo-Fullana, J. Matamoros, and C. Antn-Haro, "Reconstruction of correlated sources with energy harvesting constraints in delay-constrained and delay-tolerant communication scenarios," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1974–1986, 2017.

[26] K. A. Immink, *Codes for Mass Data Storage Systems*, 2nd ed. The Netherlands: Shannon Foundation Publishers, 2004.

[27] K. A. S. Immink, "Spectral null codes," *IEEE Transactions on Magnetics*, vol. 26, no. 2, pp. 1130–1135, 1990.

[28] D. Limbachiya, M. K. Gupta, and V. Aggarwal, "Family of constrained codes for archival DNA data storage," *IEEE Communications Letters*, vol. 22, no. 10, pp. 1972–1975, 2018.

[29] W. Song, K. Cai, M. Zhang, and C. Yuen, "Codes with run-length and GC-content constraints for DNA-based data storage," *IEEE Communications Letters*, vol. 22, no. 10, pp. 2004–2007, 2018.

[30] H. Ferreira, L. Lampe, J. Newbury, and T. Swart, *Power Line Communications: Theory and Applications for Narrowband and Broadband Communications over Power Lines*. Wiley, 2011. [Online]. Available: https://books.google.ca/books?id=QGdJcfGYrr4C

[31] K. Ouahada, H. C. Ferreira, A. J. H. Vinck, and W. A. Clarke, "Combined higher order spectral nulls codes and OFDM modulation for power line communications," in *2006 IEEE International Symposium on Power Line Communications and Its Applications*, 2006, pp. 122–127.

[32] I. B. Djordjevic and B. Vasic, "Constrained coding techniques for the suppression of intrachannel nonlinear effects in high-speed optical transmission," *Journal of Lightwave Technology*, vol. 24, no. 1, pp. 411–419, 2006.

[33] T. Koike-Akino, K. Kojima, D. S. Millar, K. Parsons, T. Yoshida, and T. Sugihara, "Pareto optimization of adaptive modulation and coding set in nonlinear fiber-optic systems," *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 1041–1049, 2017.

[34] S. Takeyama and H. Kamabe, "Code rate of constrained code for phase change memories," in *2014 International Symposium on Information Theory and its Applications*, 2014, pp. 620–624.

[35] Z. Yu, R. J. Baxley, and G. T. Zhou, "Brightness control in dynamic range constrained visible light OFDM systems," in *2014 23rd Wireless and Optical Communication Conference (WOCC)*, 2014, pp. 1–5.

[36] G. Zheng, C. Gong, and Z. Xu, "Constrained partial group decoding with maxmin fairness for multi-color multi-user visible light communication," *IEEE Transactions on Communications*, vol. 67, no. 12, pp. 8573–8584, 2019.

[37] C. Cao, D. Li, and I. Fair, "Deep learning-based decoding of constrained sequence codes," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2532–2543, 2019.

[38] C. Zou and F. Yang, "Dimming-aware deep learning approach for OOK-based visible light communication," *Journal of Lightwave Technology*, pp. 1–1, 2020.

[39] A. R. Ndjiongue, H. C. Ferreira, and T. Ngatched, "Visible light communications (VLC) technology," *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–15, June 2015.

[40] A. Gupta and R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, 2015.

[41] H. Haas, "High-speed wireless networking using visible light," *SPIE Newsroom*, no. 4, 2013.

[42] IEEE, "IEEE recommended practices for modulating current in high brightness LEDs for mitigating health risks to viewers," *IEEE Standards 179-2015*, vol. 15, no. 8, pp. 1–80, June 2015.

[43] R. Roberts and Z. Xu, "Update on VLC link budget work," *IEEE 802.15.7 VLC Standard Meeting*, 2009.

[44] P. M. Butala, H. Elgala, and T. D. C. Little, "SVD-VLC: A novel capacity maximizing VLC MIMO system architecture under illumination constraints," in *2013 IEEE Globecom Workshops (GC Wkshps)*, 2013, pp. 1087–1092.

[45] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, no. 2, pp. 265–298, 1997.

[46] M. Biagi, T. Borogovac, and T. D. C. Little, "Adaptive receiver for indoor visible light communications," *Journal of Lightwave Technology*, vol. 31, no. 23, pp. 3676–3686, 2013.

[47] Y. Wang, L. Tao, X. Huang, J. Shi, and N. Chi, "8-Gb/s RGBY LED-based WDM VLC system employing high-order CAP modulation and hybrid post equalizer," *IEEE Photonics Journal*, vol. 7, no. 6, pp. 1–7, 2015.

[48] K. Cui, G. Chen, Z. Xu, and R. D. Roberts, "Line-of-sight visible light communication system design and demonstration," in *2010 7th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP 2010)*, 2010, pp. 621–625.

[49] Trong-Hop Do and Myungsik Yoo, "Received power and SNR optimization for visible light communication system," in *2012 Fourth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2012, pp. 6–7.

[50] E. Monteiro and S. Hranilovic, "Constellation design for color-shift keying using interior point methods," in *2012 IEEE Globecom Workshops*, 2012, pp. 1224–1228.

[51] R. J. Drost and B. M. Sadler, "Constellation design for color-shift keying using billiards algorithms," in *2010 IEEE Globecom Workshops*, 2010, pp. 980–984.

[52] E. Arıkan, "Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.

[53] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6562–6582, 2013.

[54] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Trans. on Commun.*, vol. 60, no. 11, pp. 3221–3227, Nov. 2012.

[55] H. Wang and S. Kim, "Design of polar codes for run-length limited codes in visible light communications," *IEEE Photonics Technology Letters*, vol. 31, no. 1, pp. 27–30, 2019.

[56] Z. Dhannoon and R. F. H. Fischer, "Performance of run-length-limited codes in visible-light communications," in *SCC; 12th International ITG Conference on Systems, Communications and Coding*, 2019, pp. 1–6.

[57] H. Ferreira, "The synthesis of magnetic recording trellis codes with good hamming distance properties," *IEEE Transactions on Magnetics*, vol. 21, no. 5, pp. 1356–1358, 1985.

[58] X. Lu and J. Li, "New Miller codes for run-length control in visible light communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1798–1810, 2018.

[59] S. Benedetto and E. Biglieri, *Principles of Digital Transmission: With Wireless Applications*. Kluwer Academic/Plenum Publishers, 1999.

[60] D. E. Knuth, "Dancing links," *arXiv preprint arXiv:cs/0011047*, 2000.

[61] 3GPP, "NR; Multiplexing and Channel Coding (Release 15)," http://www.3gpp.org/DynaReport/38-series.htm, Tech. Rep. TS 38.212 V15.2.0 (2018-06), Jan. 2018.

[62] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, 1974.

[63] S. Kim, "Adaptive FEC codes suitable for variable dimming values in visible light communication," *IEEE Photonics Technology Letters*, vol. 27, no. 9, pp. 967–969, May 2015.

[64] S. H. Lee and J. K. Kwon, "Turbo code-based error correction scheme for dimmable visible light communication systems," *IEEE Photonics Technology Letters*, vol. 24, no. 17, pp. 1463–1465, Sept. 2012.

[65] J. Fang, Z. Che, and X. Yu, "An efficient flicker-free FEC coding scheme for dimmable visible light communication based on polar codes," vol. 9, no. 3, June 2017.

[66] H. Wang and S. Kim, "Dimming control systems with polar codes in visible light communication," *IEEE Photonics Technology Letters*, vol. 29, no. 19, pp. 1651–1654, Aug. 2017.

[67] ——, "Adaptive puncturing method for dimming in visible light communication with polar codes," *IEEE Photonics Technology Letters*, vol. 30, no. 20, pp. 1780–1783, Oct. 2018.

[68] D. E. Knuth, "Efficient balanced codes," *IEEE Transactions on Information Theory*, vol. 32, no. 1, pp. 51–53, Jan. 1986.

[69] E. N. Mambou and T. G. Swart, "A construction for balancing non-binary sequences based on gray code prefixes," *IEEE Transactions on Information Theory*, vol. 64, no. 8, pp. 5961–5969, Aug. 2018.

[70] J. Kim and H. Park, "A coding scheme for visible light communication with wide dimming range," *IEEE Photonics Technology Letters*, vol. 26, no. 5, pp. 465–468, Mar. 2014.