IntechOpen

Applications of Digital Signal Processing through Practical Approach

Edited by Sudhakar Radhakrishnan

APPLICATIONS OF DIGITAL SIGNAL PROCESSING THROUGH PRACTICAL APPROACH

Edited by **Sudhakar Radhakrishnan**

Applications of Digital Signal Processing through Practical Approach
http://dx.doi.org/10.5772/59529
Edited by Sudhakar Radhakrishnan

Contributors

Baba Tatsuro, Jian Wang, Alan Willner, Hugo Guzmán, Federico Barrero, Mario Durán, Mario Bermúdez, Cristina Martín, Antonio J. R. Neves, Guo-Wei Lu, Olivier Romain, Julien Le Kernec

Notice

First published in Croatia, 2015 by INTECH d.o.o.
eBook (PDF) Published by IN TECH d.o.o.
Place and year of publication of eBook (PDF): Rijeka, 2019.
IntechOpen is the global imprint of IN TECH d.o.o.
Printed in Croatia

Legal deposit, Croatia: National and University Library in Zagreb

Additional hard and PDF copies can be obtained from orders@intechopen.com

Applications of Digital Signal Processing through Practical Approach
Edited by Sudhakar Radhakrishnan

p. cm.

ISBN 978-953-51-2190-9

eBook (PDF) ISBN 978-953-51-5764-9

Meet the editor

Dr. Sudhakar Radhakrishnan is currently the Professor and Head of the Department of Electronics and Communication Engineering, Dr. Mahalingam College of Engineering and Technology, Pollachi, India. He is an editorial board member of three international journals and a reviewer of nine international journals. He wrote a book chapter titled 'Wavelet-based Image Compression' in the book titled *Computational Intelligence Techniques in Handling Image Processing and Pattern Recognition* and wrote two books titled *Research Issues in Image Compression Using Wavelet Variants* and *Practicing Signals and Systems Laboratory Using MATLAB* published by Lambert Academic Publishing (LAP), Germany (2010). He edited a book titled *Effective Video Coding for Multimedia Applications* published by InTech, Croatia (2011). He has published 50 papers in international and national journals and conference proceedings. His areas of research include digital image processing, wavelet transforms and digital signal processing.

Contents

Preface

The rapid growth of microelectronics and digital computing has stimulated a significant growth in the area of digital signal processing (DSP). The concepts of DSP proliferated in many areas such as telecommunications, digital television, biomedical engineering, digital audio, and power conversion. DSP now seems to be a core for many new emerging digital applications and for the information society. Today's information revolution paves the way for the engineers in the areas of electronics, computer and communication engineering to think about DSP concepts.

Just a decade ago, digital signal processing was more of theory than practice. The only systems capable of doing signal processing were massive mainframes and supercomputers and, even then, much of the processing was not done in real time but off-line in batches. For example, seismic data were collected in the field, stored on magnetic tapes and then taken to a computing centre, where a mainframe might take hours or days to digest the information. The first practical real-time DSP systems emerged in the late 1970s and used bipolar 'bit-slice' components. The economics began to change in the early 1980s with the advent of single-chip metal-–xide semiconductor (MOS) DSPs.

Digital signal processors were invented to handle digital signal processing tasks and were available in a variety of applications like audio signal processing, audio and video compression, speech processing and recognition, digital image processing, digital communications, biomedicine, seismology and radar applications. Specific uses include speech transmission in mobile phones, seismic data processing, analysis of industrial processes, medical imaging such as computerized axial tomography (CAT) scans, MP3 compression and computer graphics.

Scope of the Book

Many books are available for understanding digital signal processing concepts. This book is an outcome of research done by various researchers and professors who have highly contributed to the field. This book would suit researchers in the field of digital signal processing.

Structure of the Book

The book contains six chapters divided into three sections. The reader is expected to know the fundamentals of digital signal processing, which are available in all the standard DSP books.

Section 1, consisting of three chapters, deals with applications of digital signal processing in communication engineering. Section 2 contains two chapters and describes the application of digital signal processing concepts in image processing. Section 3, consisting of a single chapter, focuses on the role of DSP in power conversion systems.

Acknowledgements

I thank the Almighty for showering His blessings and giving me the intelligence and energy to complete this work. My sincere thanks to the management of Dr. Mahalingam College of Engineering and Technology and Prof. C. Ramaswamy, Secretary, NIA Institutions, for their encouragement and patronage rendered to carry out this work. I am indebted to my wife Mrs. Vinitha Mohan. Her support, encouragement, quiet patience and unwavering love undeniably led me to the successful completion of the work. I am at a loss for words to thank my son S.V. Hemesh for putting up with my preoccupation, for his understanding and for the love he gave me. I am grateful to InTech publisher, especially the Publishing Process Manager Ms. Iva Lipović who constantly helped me in bringing this book to completion.

Dr. Sudhakar Radhakrishnan
Department of Electronics and Communication Engineering,
Dr. Mahalingam College of Engineering and Technology,
India

Section 1

Digital Signal Processing towards Communication Engineering

Chapter 1

Optical Signal Processing for High-Order Quadrature-Amplitude Modulation Formats

Guo-Wei Lu

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/61681

Abstract

In this book chapter, optical signal processing technology, including optical wavelength conversion, wavelength exchange and wavelength multicasting, for phase-noise-sensitive high-order quadrature-amplitude modulation (QAM) signals will be discussed. Due to the susceptibility of high-order QAM signals against phase noise, it is imperative to avoid the phase noise in the optical signal processing subsystems. To design high-performance optical signal processing subsystems, both linear and nonlinear phase noise and distortions are the main concerns in the system design. We will first investigate the effective monitoring approach to optimize the performance of wavelength conversion for avoiding undesired nonlinear phase noise and distortions, and then propose coherent pumping scheme to eliminate the linear phase noise from local pumps in order to realize pump-phase-noise-free wavelength conversion, wavelength exchange and multicasting for high-order QAM signals. All of the discussions are based on experimental investigation.

Keywords: Optical Signal Processing, Nonlinear Optics, Advanced Optical Modulation Formats, Quadrature Amplitude Modulation

1. Introduction

Recently, digital signal processing (DSP) is playing an increasingly important role in coherent detection for reconstructing the complex field of signal and compensating for the transmission impairments. It dramatically simplifies the reception of multi-level and multi-dimensional modulation formats such as high-order quadrature amplitude modulation (QAM), thus making high-order QAM become a promising and practical approach for achieving higher bit rate and higher spectral efficiency. However, optical signal processing is still highly desirable and appreciable in order to overcome the electronics bottlenecks, support the transparency and ultra-fast processing in future optical networks. As basic optical network functionalities, all-optical wavelength conversion, wavelength data exchange, and wavelength multicasting play important roles in the all-optical networks to enhance the re-configurability and non-blocking capacity, and facilitate the wavelength management in future transparent optical networks.

On the other hand, recently, lots of advanced modulation formats like single-carrier high-order QAM like 64QAM [1–5] or multi-carrier optical orthogonal frequency-division multiplexing (OFDM) have been introduced and realized in optical communications for enabling spectrally-efficient and ultra-fast optical transmissions. It is desirable to exploit optical signal processing schemes suitable for these advanced optical modulation formats. However, for these high-order QAM signals, the increasing number of states in the constellation makes the signal more sensitive to the intensity and phase noise. It is imperative to suppress phase noise in optical signal processing subsystems to allow compatibility phase-noise sensitive high-order QAM formats.

As one of the basic optical signal processing techniques, several all-optical wavelength conversion (AOWC) schemes have been demonstrated to realize AOWC functions of OFDM, 8ary phase-shift keying (8PSK), 16QAM, and 64QAM by using the second-order nonlinear effect in periodically-poled Lithium Niobate (PPLN) waveguide [6, 7], four-wave mixing (FWM) in highly-nonlinear fibers (HNLF) [8, 9], semiconductor optical amplifier (SOA) [10–12], or silicon waveguide. However, the implementation penalty of such subsystems varies from 2dB to 4dB at bit-error rate of 10^{-3} [9, 12], which is non-negligible for optical networks, especially when multiple wavelength conversion nodes are included in the networks. The distortions introduced in the AOWC mainly originate from: i) the phase noise from the pumps due to the finite laser linewidth, referred to as linear phase noise; and ii) other undesired nonlinear distortions or crosstalk co-existed in the nonlinear process, called as nonlinear phase noise or distortion. To suppress the linear phase noise from pumps, the straightforward way is to use narrow-linewidth lasers, such as external-cavity laser (ECL) or fiber laser (FL), as pump sources. However, it increases the implementation cost. On the other hand, since the nonlinear media in the sub-system is operated in the nonlinear operation region, expect the dominant nonlinear effect utilized for implementing optical signal processing functionalities, it is highly possible that other undesired nonlinear effects co-occur in this process, thus deteriorating the quality of the converted signal. For example, for the wavelength conversion based on the FWM in SOA, additional distortion from cross-gain modulation, cross-phase modulation (XPM), and self-phase modulation (SPM) may deteriorate the converted signal, while in the wavelength conversion based on FWM in HNLF, additional undesirable distortions are mainly from stimulated Brillouin scattering (SBS), SPM or XPM. High-order QAMs, especially going up to 32QAM, 64QAM or beyond, exhibit more sensitive to nonlinear phase noise like SPM or XPM [13]. Therefore, in order to realize a high-quality all-optical wavelength conversion (AOWC) sub-system for high-order QAMs, it is essential to optimize the system performance of AOWC through effective monitoring approach to suppress the distortion introduced by extra undesired nonlinear distortions. In this chapter, it is categorized into two parts. In the first part, the effective monitoring approach is discussed to avoid the undesired nonlinear phase noise and distortions in the optical signal processing subsystem to enable superior performance [14]. Then, a coherent pumping scheme is proposed and discussed in the second part to implement the pump-phase-noise-free wavelength conversion, wavelength exchange, and wavelength multicasting for high-order QAM signals. Figure 1 summarizes the main topics which will be discussed in this book chapter.

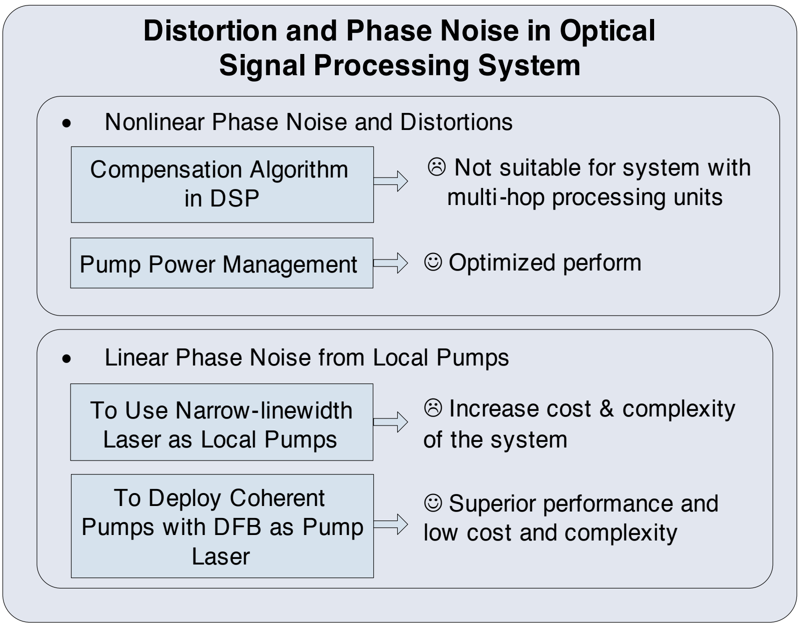


Figure 1. Topics to be discussed in this book chapter.

2. Performance optimization of wavelength conversion of high-order QAM signals

It is well-known that for high-order QAM signals, the increasing number of states in the constellation makes them more sensitive to the intensity and phase noise. Previously, power penalties of around 4 dB at 5Gbaud [12], and 2 dB at 21Gbaud [9] were experimentally demonstrated for the converted 64QAM at bit-error rate (BER) of 10^{-3}. As shown in Fig. 3, to implement the AOWC for high-order QAMs, a simple degenerate FWM in HNLF is deployed. Input QAM signal serves as probe, while a CW pump works as pump in AOWC. The phase of the converted signal follows the phase relationship: $\theta_{idler}=2\theta_{pump}-\theta_{probe}$, where θ_{idler}, θ_{pump}, and θ_{probe} are the phase of the idler, pump and probe, respectively. In order to implement an AOWC for QAM signals with minimal power penalty, the phase and intensity noise from both pump and probe should be suppressed. Since high-order QAM signals are sensitive to the phase noise in the system, to avoid the introduced linear phase noise from pump, it is preferred to deploy narrow linewidth light sources for the pump source. In the following experimental demonstration, a tunable external cavity laser (ECL) with the linewidth of around 100 kHz is employed as the light source of the input QAM signal (probe). On the other hand, two fiber lasers (FLs) with a linewidth of around 10 kHz are used as light sources for pump and local oscillator (LO) at the coherent receiver. Since a narrow-linewidth FL was deployed as pump source, the linear phase noise from pump was negligible.

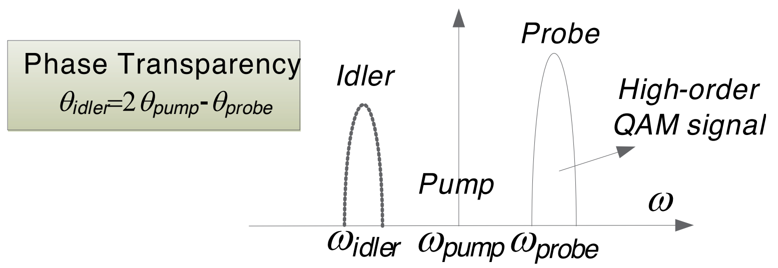


Figure 2. Operation principle of wavelength conversion using FWM in HNLF.

In the AOWC subsystem based on FWM in HNLF for high-order QAM signals, the main nonlinear distortions in the converted signal are mainly from the following sources:

1. SPM from the probe signal: Since the input QAM signal, i.e. the probe, exhibits multilevel in amplitude, in the nonlinear operation condition, the probe may experience SPM. The nonlinear phase noise will then be transferred to the converted signal through FWM and finally deteriorate the converted signal. Therefore, it is critical to manage the launched power of probe to avoid the degradation in the converted QAM signal caused by the probe-introduced SPM. However, it will sacrifice the conversion efficiency. There is a tradeoff between conversion efficiency and the quality of the converted signal in the performance optimization.

2. XPM from the pump signal: As discussed in [15,16], with limited optical signal-to-noise ratio (OSNR) in pump, the amplitude noise in pump may distort the converted signal by introducing nonlinear phase noise through XPM effect. In our experiment, a FL is used as the pump source. Thanks to the low relative intensity noise (RIN) of the FL, the OSNR of pump source is measured as around 57 dB, which avoids the pump-induced nonlinear phase noise.

3. SBS from the pump signal: In AOWC subsystems based on FWM in HNLF, SBS limits the conversion efficiency unless the pumps' linewidth is broadened to increase the SBS threshold. In an AOWC based on degenerate single-pump FWM, if intentionally applying phase dithering on the pump, it will deteriorate the converted QAM signals. Although it has been shown that the phase dithering could be compensated for at the coherent digital receiver by DSP [17], the applied phase dithering will be accumulated in the converted signal as distortions and be further transferred to the next node,

which is not suitable for multi-hop optical networks. In our experiment, thanks to the short length (150 m) and high nonlinearity (nonlinear coefficient: 18/W/km) of the deployed HNLF, the measured SBS threshold is around 24 dBm, which allows a high launching power even without applying additional phase dithering. However, the optimization of the pump power is required in order to avoid the SBS distortion in the pump.

As discussed above, the main undesired nonlinear components in the AOWC based on degenerate FWM in HNLF are from SPM of the input QAM (probe) and SBS of the CW pump. In order to eliminate these deleterious components in the converted signal, the launched pump and probe power should be well managed.

2.1. Experimental investigation

Figure 3 depicts the experimental setup used to achieve the AOWC of 36QAM and 64QAM through FWM in HNLF. Since high-order QAM signals are sensitive to the phase noise in the system, it is preferred to employ narrow-linewidth light sources in the experiment, especially for the pump source. Owing to the lack of instruments in the lab, in the experiment, a tunable ECL with a linewidth of around 100 kHz was deployed as a light source of the input QAM signal in the experiment, whereas two FLs with a linewidth of around 10 kHz worked as light sources for the pump and LO at the coherent receiver. To synthesize optical QAM signals, the light from the ECL, operating at 1551.38 nm, was modulated by a single in-phase/quadrature (IQ) modulator, which had a 3 dB bandwidth of around 25 GHz, and a 3.5 V half-wave voltage. Two de-correlated 6- or 8-level driving signals originating from pseudorandom binary sequence (PRBS) streams with a length of 2^{15}–1 from an arbitrary waveform generator (AWG) were used to drive the IQ modulator for generating optical 36QAM or 64QAM, respectively. After power amplification, the QAM signal was combined with amplified CW light at 1551.95 nm, and was then fed into a 150 m length of HNLF having an attenuation coefficient of 0.9 dB/km, a nonlinear coefficient of 18/W/km, a zero-dispersion wavelength of 1548 nm, and a dispersion slope of around 0.02 ps/nm^2/km. Note that, due to the inability to tune the wavelength of the FLs used in the experiment, wavelengths of the probe signal and pump could not be set for the optimum FWM efficiency. Nevertheless, owing to the high nonlinear effects and flat-dispersion-profile of the employed HNLF, the experimental results showed high conversion efficiency, which can ensure the superior performance of the converted signal. The produced idle signal at the wavelength of 1552.52 nm was filtered out and then led to the phase-diversity intradyne coherent receiver for the coherent detection and for BER measurement. The coherent receiver included an LO, a 90 degree optical hybrid device, and two balanced photo-detectors (PDs). After detection by the balanced PDs, the data was digitized at 50GSamples/s by employing a digital storage oscilloscope (Tektronix DP071254) which has the analog bandwidth of 12.5 GHz. The captured data was processed offline through the DSP that included compensation of skew, IQ imbalance, power, data resampling, linear equalization using the finite impulse response (FIR) filtering, carrier phase recovery, and the final hard-decision circuits. 89,285 symbols were used for the BER measurement.

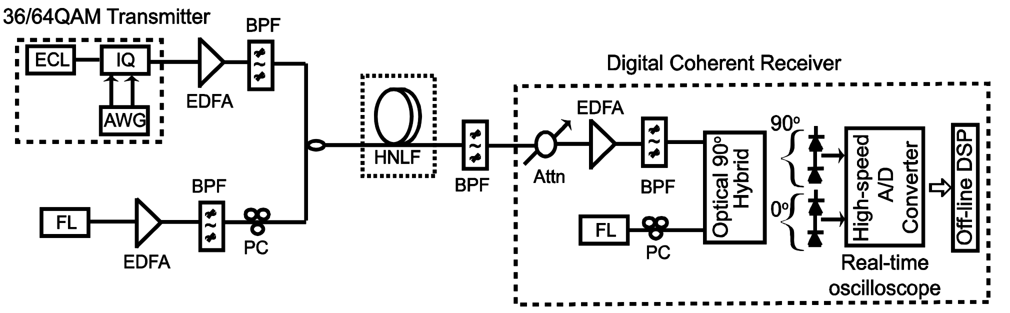


Figure 3. Experimental setup of the wavelength conversion of 36QAM and 64QAM signals.

2.1.1. AWOC of 36QAM

In order to eliminate possible deleterious components in the converted signal, the launched pump and probe power should be well managed. Figure 4(a) shows the measured EVMs and BERs at the received OSNR of around 25 dB when the probe power was tuned from 7 to 15 dBm and the pump power was fixed at around 20dBm. An improvement in both the EVMs and BERs of the converted 36QAM was observed with an increase in the probe power up to around 11 dBm. After the inflection point (around 11 dBm), both EVMs and BERs increased with the increase of the probe power, which was attributed to the SPM of probe in the nonlinear process. Therefore, we considered setting the probe power to around 11 dBm to avoid the SPM introduced in the probe. As previously mentioned, another main source of distortion is the SBS of the pump in AOWC. To optimize the pump power, we also measured the corresponding EVMs and BERs when the probe power was fixed at 11 dBm and the pump power was tuned from 15 dBm to 23 dBm (Fig. 4(b)). As the launched pump power increased, EVMs and BERs showed similar behavior. We found that it was better to operate the pump power in the range of 17.5–22 dBm. At the pump power of 15.4 dBm, the constellations were relatively noisy due to the low conversion efficiency. However, once the pump power was increased to 22.9 dBm, distortion from SBS started to appear in the measured constellation, acting mainly as intensity noise. To obtain the optimal performance, we set the pump power at 20 dBm in the AOWC of 36QAM. While monitoring the converted 36QAM, EVMs and BERs showed consistent behavior when tuning the probe and pump powers.

As we discussed previously, the optimal pump and probe power were 20 dBm and 11 dBm for the AOWC of 36QAM. The corresponding optical spectrum under the optimal condition is shown in Fig. 5(i), where a conversion efficiency of about −15 dB was obtained compared with the input probe power. Under the optimal operating condition, the BER performance was measured as the function of OSNR at 0.1 nm for both input and converted signals, and shown in Fig. 5(ii). For the input QAM signals, the power penalty of around 2 dB was obtained compared with theoretical BER measurement at the BER of 10^{-3}, which is better than the previously-reported QAM transmitters [2]. The power penalty is mainly owing to the imperfectness of the transmitter. With respect to the input QAM, a negligible power penalty (<0.3

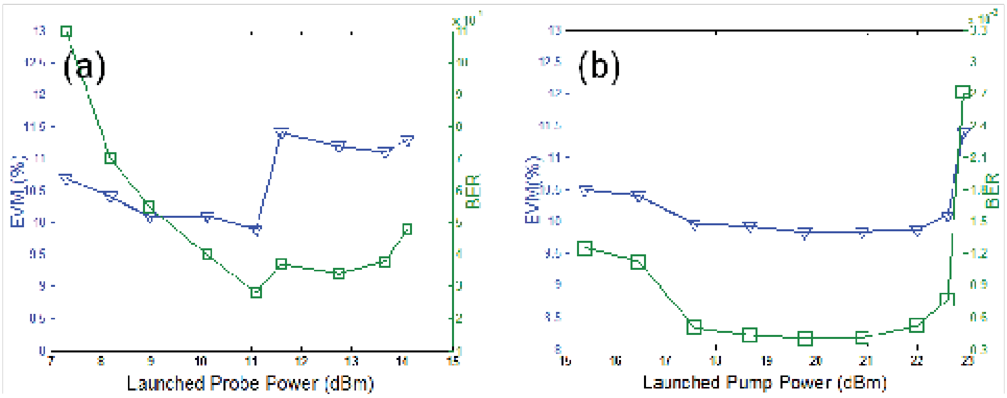


Figure 4. Measured EVM (triangles) and BER (squares) results of the converted 36QAM signals (**a**) when tuning probe power from 7 to 15 dBm, (**b**) when tuning pump power from 14 to 23 dBm.

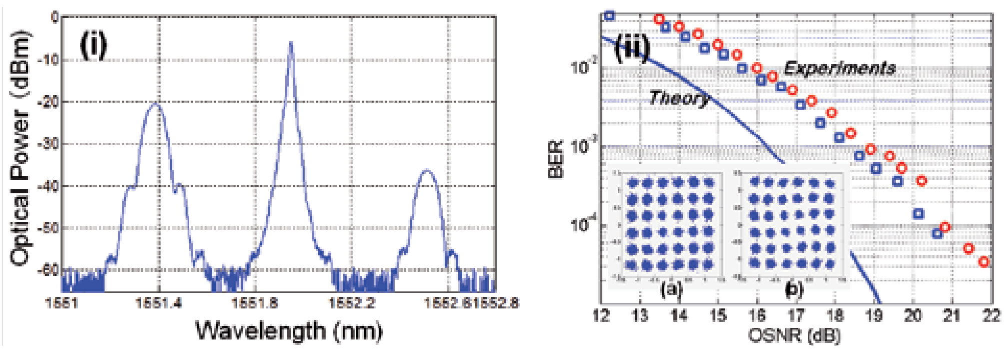


Figure 5. (**i**) Measured optical spectrum in the optimal condition, (**ii**) measured BER as function of the received OSNR (0.1 nm). Insets: (**a**) input and (**b**) converted 36QAM signals.

dB) was observed at a BER of 10^{-3}. The measured constellations of the input and converted 36QAM are shown in the insets of Fig. 5(ii), where the received OSNR was around 35 dB.

2.1.2. AWOC of 64QAM

To optimize the performance of AOWC for 64QAM, measurements similar to those described above were performed. Figure 6 (a) depicts the measured EVMs and BERs at the received 25 dB OSNR when the probe power was tuned from 7 to 15 dBm and the pump power was fixed at around 20 dBm. The increase in the launched probe power decreased the BER of the converted signal to around 12.2 dBm owing to the improved OSNR of the converted signals. When the probe power was increased further, the BER started to increase, attributed to the introduced SPM in the probe signal. The BER results with different probe powers suggested to operate the probe power in the range of 9 to 14 dBm. Furthermore, the measured constellations offered a more perceptive and precise approach for optimizing the performance. The EVMs with the various probe powers were calculated and are plotted in Fig. 6(a). With the increase of the probe power, both BER and EVM results show similar trends. However, according to the EVM and BER results, different optimum probe powers of around 9.2 dBm and 12.4 dBm were obtained, respectively. When the launched probe power was increased to around 12.4 dBm, SPM-induced distortion became visible in the constellation, causing the increase of EVM. However, the SPM-induced spiral rotation in the constellation happens to enlarge the symbol distance between symbols, thus decreasing the BER. Therefore, these results suggested that, to optimize the performance of AOWC, it would be effective to monitor the constellation or EVM, which gives a more intuitive and proper means to optimize the AOWC performance, in order to eliminate the extra undesired nonlinear phase noise introduced in the process.

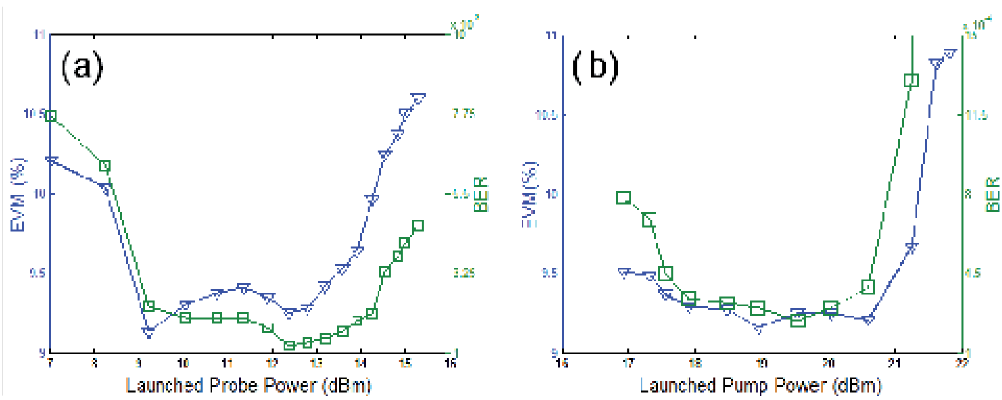


Figure 6. Measured EVM (triangles) and BER (squares) results of the converted 64QAM signals (**a**) when tuning probe power from 7 to 15 dBm, (**b**) when tuning pump power from 17 to 21 dBm.

For pump power optimization, the EVM and BER results were measured when the launched pump power was increased from 17 to 22 dBm, whereas the pump power was set at around 9 dBm. The measurement was done for optimizing the pump power and is shown in Fig. 6(b). Similar behavior was obtained for the measured EVM and BER values when the pump power was increased. In order to avoid the distortion owing to the SBS, we considered to set the pump power in the range of 17.5–20.5 dBm. It is clear that a high pump power was helpful for obtaining high conversion efficiency, therefore, resulted in a sufficient OSNR for the converted signal. Thus, in this experiment, the pump power was optimized to 20 dBm, which resulted in a conversion efficiency of about -15 dB and also ensured that there was no SBS distortion introduced for the converted signal. The distortion from SBS acted mainly as amplitude noise in the constellations, and became severe once the pump power was increased to more than 20.5 dBm.

To achieve the optimal performance of AOWC for 64QAM signal, the pump and probe power were set at 20 dBm and 9 dBm, respectively. A conversion efficiency of around −15 dB is obtained, as shown in Fig. 7(i). Under the optimized conditions, the BER performance as a function of OSNR was shown in Fig. 7(ii). The implementation penalty compared with the theoretical BER curve was around 2.8 dB for 64QAM, at a BER of 10^{-3}, which is much better

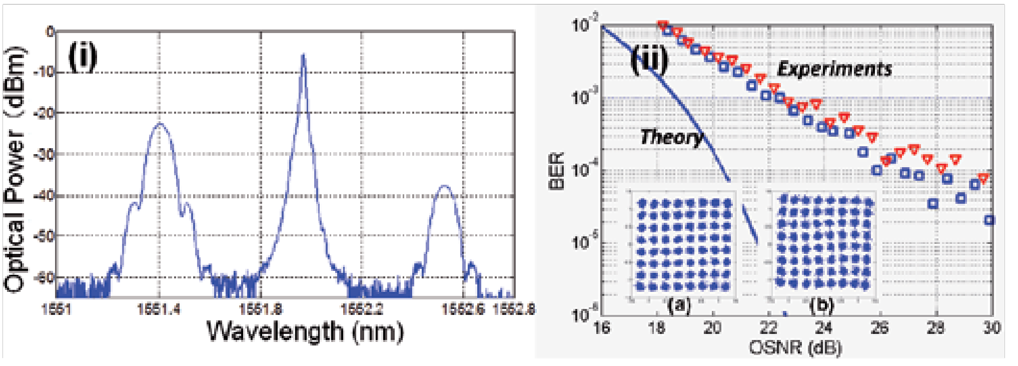


Figure 7. (**i**) Measured optical spectrum in the optimal condition, (**ii**) measured BER as function of the received OSNR (0.1 nm). Insets: (**a**) input and (**b**) converted 64QAM signals.

than those of the previously-reported 64QAM transmitters in [3–4]. Similar to that performance of 36QAM AOWC, a negligible power penalty of <0.3 dB was observed with the respect to the input signal at a BER of 10^{-3} after the conversion. The obtained constellations of the input and converted high-order QAMs at around 35 dB received OSNR and are shown in the insets of Fig. 7(ii).

2.2. Summary

We have experimentally demonstrated the AOWC of optical 10-Gbaud (50 Gbps) 36QAM and (60 Gbps) 64QAM through a degenerate FWM effect in HNLF with a power penalty of less than 0.3 dB at a BER of 10^{-3}. In order to optimize the AOWC performance, the converted high-order QAM signals were evaluated by measuring the BER and constellations, i.e., EVM. Since EVM showed higher sensitivity in the presence of nonlinear phase noise, the results suggested the effectiveness of optimizing the AOWC performance by monitoring EVM, rather than BER, especially for high-order QAM signals.

3. Pump-phase-noise-free optical signal processing

The previous session mainly focuses on how to avoid or suppress the nonlinear noise or distortion in optical signal processing. In this session, the focus is to exploit the approach to eliminate the linear phase noise from the local pumps deployed in optical signal processing subsystems. In optical signal processing subsystems, such as wavelength conversion, wavelength exchange or wavelength multicasting, it is inevitable to deploy local pump sources to realize the optical signal processing functionalities. As we discuss before, the linear phase noise from local pumps may introduce phase noise or distortion to the converted signal in optical signal processing subsystem. The most straightforward way is to deploy narrow-linewidth lasers as pump sources. However, it increases the implementation cost of the systems. We will present our proposed coherent pumping scheme. Thanks to the phase noise cancelling effect using this coherent pumping, it allows the use of low-cost distributed feedback (DFB) lasers as pump sources, and at the same time, ensures the superior performance since it is free of the phase noise from pumps. Here we will demonstrate several pump-phase-noise-free optical signal processing subsystems for high-order QAM signals, including: (a) pump-phase-noise-free wavelength conversion and wavelength exchange for high-order QAMs signals using cascaded second-order nonlinearities in PPLN [18, 19]; and (b) pump-phase-noise-free wavelength multicasting of QAM signals using FWM in HNLF [20].

3.1. Pump-phase-noise-free wavelength conversion and wavelength exchange in PPLN

Figure 8 depicts the operation principle of the pump-linewidth-free AOWC. It is based on cascaded second-order nonlinearity in PPLN. Two pumps at ω_{p1} and ω_{p2} are allocated at one side of quasi-phase-matching (QPM) wavelength of PPLN, whereas input signal at ω_1 is placed symmetrically with pump at ω_{p1} with respect to QPM wavelength. After AOWC, the input signal at ω_1 is shifted to the frequency ω_2, with $\omega_2=\omega_{p1}-\omega_{p2}+\omega_1$, where ω_{p1}, ω_{p2} and ω_1 are the frequencies of pump1, pump2, and the input signal, respectively. It is typically employed for performing the data exchange between the two input wavelengths [13], i.e. wavelength exchange. The frequencies ω_{p1} and ω_1 have to be arranged symmetrically with the respect to the PPLN's quasi-phase-matching (QPM) wavelength in order to satisfy the phase matching condition and for increasing the conversion efficiency. With the non-depletion assumption, linear mapping between the input and output relationship in complex amplitudes and phase are given by equations (1) and (2), respectively.

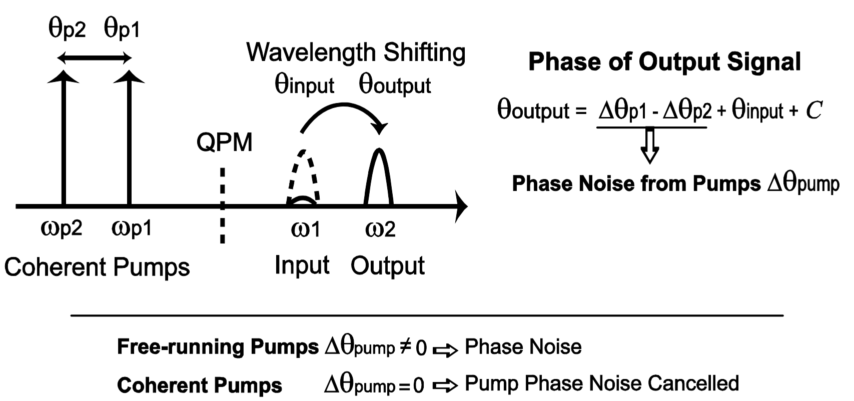


Figure 8. Operation principle of the pump phase-noise cancellation using coherent pumping.

$$A_{\omega_2} \propto A_{\omega_1} \cdot A^*_{\omega_{p2}} \cdot A_{\omega_{p1}} \tag{1}$$

$$\theta_{\text{output}} = \theta_{\text{input}} + \Delta\theta_{\text{p1}} - \Delta\theta_{\text{p2}} + C = \theta_{\text{input}} + \Delta\theta_{\text{pump}} + C \tag{2}$$

where θ_{output}, θ_{input}, $\Delta\theta_{p1}$, $\Delta\theta_{p2}$, and C are the phase of the converted and input signals, the phase noise from pump1 and pump2, and a constant term, respectively, and $\Delta\theta_{pump} = \Delta\theta_{p1} - \Delta\theta_{p2}$. Note that the phase information in each pump is transparently transferred to the converted signal as a subtraction term between them. In order to avoid additional phase noise introduced in the process, the phase noise term from pumps, $\Delta\theta_{pump}$, should be minimized. If the pumps are synthesized by a two-tone generator (TTG) from a single laser source, the phase noise from pumps is eliminated in the converted signal, i.e. $\Delta\theta_{pump}=0$. Hence, the wavelength conversion becomes free of the phase noise from pumps, allowing the use of lower cost lasers and at the same time ensuring a superior performance in terms of noise performance. The TTG may be constructed using either Mach-Zehnder modulators driven by a RF clock, or an optical frequency comb followed by an optical spectrum shaper. The two-tone spacing could vary from a fraction of nanometer to several nanometers, making it possible to cover a relative wide conversion range in the OWC. The TTG generated from a filtered optical frequency comb is more suitable and practical for the OWC based on HNLF.

3.1.1. Pump-phase-noise-free AOWC in PPLN

The experimental set-up is depicted in Fig. 6, showing OWC scheme of 16 and 64 QAM signals. A 5kHz linewidth FL at the wavelength of 1552.52 nm was deployed as the light source to minimize the phase noise from the input signal. And then the light was modulated by an in-phase/quadrature (IQ) modulator for generating QAM signals. The two de-correlated 4- or 8-level driving electronics derived from 10-Gbaud PRBS streams with the length of 2^{15}–1 were generated from an arbitrary waveform generator (AWG) to drive the IQ modulator, which has a $V\pi$ of 3.5 V and an optical bandwidth of around 25 GHz. Two different pump configurations were adopted for comparison. The two pumps were generated from a single laser source at the wavelength of 1548.08 nm using a TTG in the coherent pump configuration, which consisted of a high extinction-ratio (ER) optical modulator driven by a 25-GHz RF clock. The high-ER modulator was made up on the x-cut $LiNbO_3$ substrate with two embedded active trimmers in each arm and it has the extinction ratio of up to 60 dB. The two phase-correlated coherent pumps were obtained with the 50-GHz frequency separation with a >40-dB spurious suppression ratio. For the case of free-running pumping, two independent free-running lasers at the wavelengths of 1547.88 and 1548.28 nms were used as pumps with 50-GHz spacing. For each of the configurations, we tried either the 500-kHz linewidth ECLs or the 3.5-MHz linewidth DFBs as the laser sources for the pumps.

The optical spectra with and without pumps for wavelength conversion of 64QAM signals after the PPLN are shown in Fig. 10. Similar conversion efficiency (CE) and signal depletion (SD) were obtained for the both free-running pumps (ECL/DFB) and coherent two-tone pumps (ECL/DFB). Here, the CE is defined as the power ratio between the converted signal to the input signal after the PPLN. On the other hand, the SD is the power ratio of the input signal after the PPLN when the both pumps were switched OFF and ON, respectively. The total pump power launched into the PPLN was set to the maximum value of about 28.8 dBm (25.8 dBm for each pump) to maximize both CE and SD, where CE of -6.5 dB and SD of 25 dB were obtained with input signal power of 6 dBm.

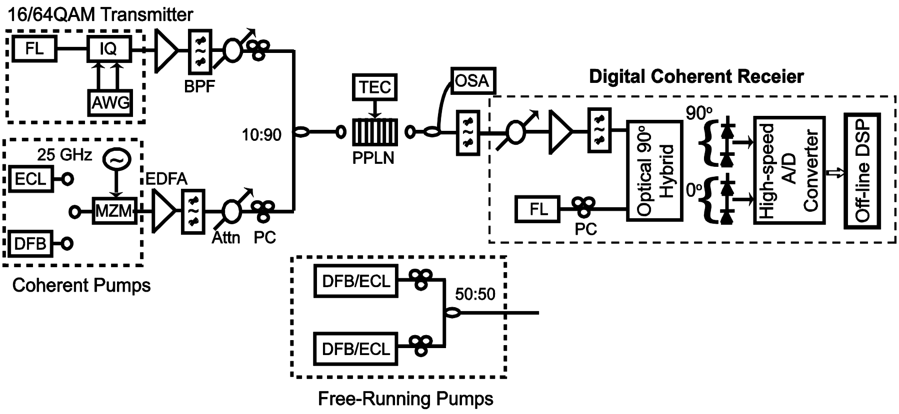


Figure 9. Experimental set-up for AOWC of 16QAM and 64QAM signals.

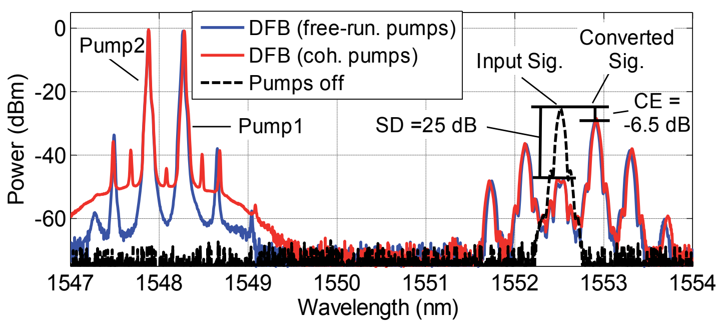


Figure 10. Optical spectra measured after PPLN when performing OWC of 64QAM with DFB pump lasers in both free-running and coherent configurations.

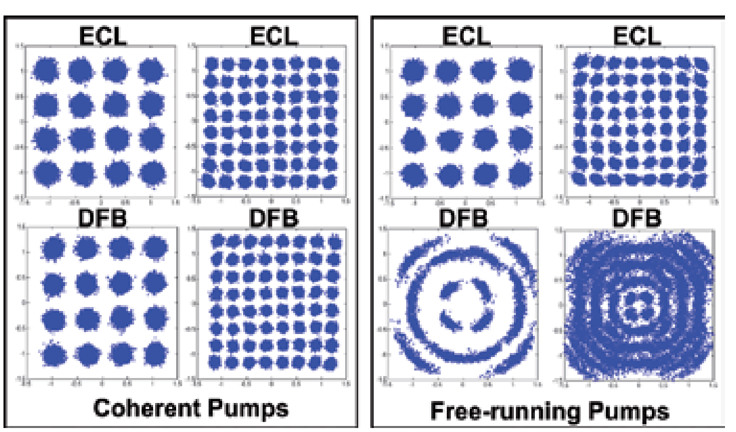


Figure 11. Measured QAM constellations using ECL and DFB pump lasers in coherent two-tone and free-running configurations (16QAM: OSNR=18 dB, 64QAM: OSNR =34 dB).

http://dx.doi.org/10.5772/61681

The constellations of the converted 16/64QAMs signals were re-constructed and observed with different pump lasers and pump configurations. As shown in Fig. 11, for either ECL or DFB pump laser, clear constellations are observed with coherent two-tone pumps. On the other ways, with the ECL pump lasers in free-running configuration, symbol rotation in phase starts to turn into obvious in the 64QAM constellation owing to the additional phase noise from the free-running ECL pumps. Furthermore, the presence of even larger pump phase noise causes clear spreading of the symbols around the unit circle for both formats with DFB free-running pumps, which is more severe for the higher amplitude symbols. From the measured BER curves, the results can also be confirmed as a function of optical signal-to-noise ratio (OSNR) at 0.1 nm for both input and converted 16/64QAM signals, as seen in Fig. 12. For both ECL and DFB pump lasers with coherent pump configuration, negligible power penalties of <0.1 dB for 16QAM and <0.3 dB for 64QAM at BER of 10^{-3} are observed with the respect of the input signal at 10Gbaud. Although we can get insignificant power penalty of <0.3 dB at BER of 10^{-3} for 16QAM with ECL as the pump laser, by increasing the modulation level to 64QAM, a 0.5 dB penalty at BER of 10^{-3} and an error floor at around 3×10^{-5} are observed in the case of free-running pumps. Owing to the strong phase noise with the free-running DFB pumps, even at >30 dB OSNR, a BER of around 10^{-2} is observed for 16QAM. The effectiveness of the pump-phase-noise removal in the OWC for high-order QAM with coherent two-tone pumps is verified by the BER and constellation results.

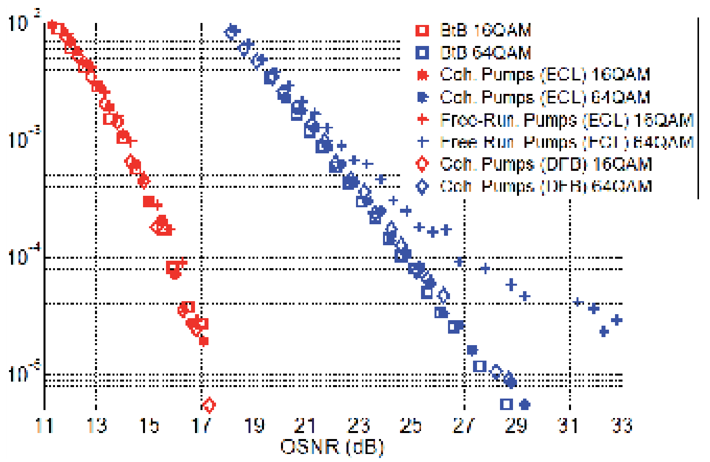


Figure 12. Measured BER vs. OSNR curves for 16/64QAM. Squares: back-to-back (BtB), stars: coherent pumps (ECL), crosses: free-running pumps (ECL), diamonds: coherent pumps (DFB).

3.1.2. Pump-phase-noise-free wavelength exchange in PPLN

Wavelength exchange is a kind of optical signal processing technique to realize bidirectional information swapping between different wavelengths. It consists of simultaneous signal depletion and wavelength conversion processes of two participated channel signals. Each of input signals is power consumed and its corresponding power is shifted to the other wavelength, finally realizing data exchange between two wavelengths in single device. So far, several works have been demonstrated through non-degenerate FWM in highly-nonlinear fiber [21–24] or cascaded second-order nonlinearities in PPLN waveguide [25–26].

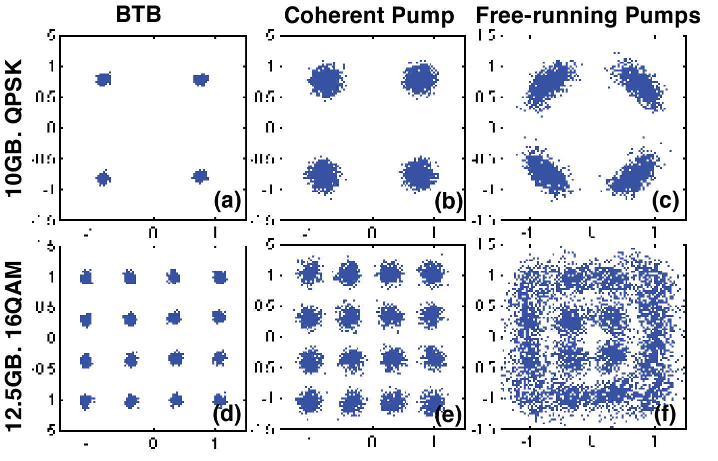


Figure 13. Measured constellations of input and converted signals after wavelength exchange (16QAM and QPSK).

Here, we apply the coherent pumping concept to wavelength exchange to demonstrate pump-phase-noise free wavelength exchange in PPLN. For experimental demonstration, an experimental setup similar to the one shown in Fig. 9 was deployed by adding another input signal. Two input signals modulated in 16QAM and QPSK, respectively were launched to PPLN as input signals for performing wavelength exchange. To evaluate the performance of wavelength exchange, BER and constellations were measured. The constellations of the swapped signals with different pump configurations are depicted in Fig. 13. With coherent pumps, clear constellations are observed for both QPSK and 16QAM. However, with DFB free-running pumps, the presence of pump phase noise causes clear spreading of the symbols around the unit circle with which is more severe for the higher amplitude symbols in 16QAM. It implies that with incoherent DFB pumps, the phase noise from pump severely deteriorates the swapped QAM signals. It can also be confirmed from the measured BER curves as a function of OSNR (0.1 nm) for both input and swapped signals. With coherent DFB pump, around 0.6 dB and 3 dB power penalties at BER of 10^{-3} were obtained for QPSK and 16QAM, respectively. As discussed above, this is mainly attributed to the crosstalk introduced by finite ER (20dB). However, in case of free-running pumps, although it was still possible to obtain BER curve for swapped QPSK, ~3.4-dB penalty and visible error-floor at BER of 5×10^{-4} were clearly observed. Due to the susceptibility of 16QAM against phase noise and crosstalk, it becomes impossible to obtain BER plot for the swapped 16QAM. This verifies the effectiveness of the elimination of the pump phase noise in the OWE for high-order QAM with coherent pumps.

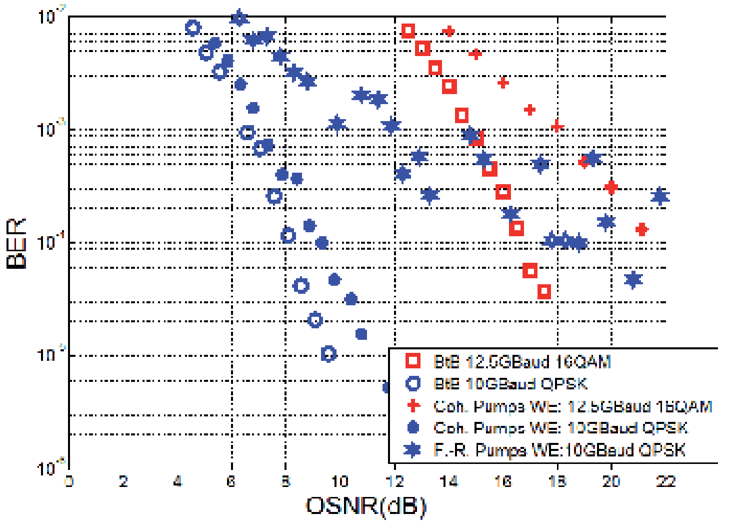


Figure 14. Measured BER vs. OSNR of the input and swapped QAM signals.

3.2. Pump-phase-noise-free wavelength multicasting of high-order QAM by FWM in HNLF

With the emergence of high-bandwidth point-to-multipoint applications such as high-definition Internet TV, big-data sharing, and data center migration, the need for wavelength multicasting has arisen recently to improve the network throughput and decrease the blocking probability in optical networks. Through multicasting, the network wavelength resources could be efficiently and flexibly managed in wavelength division multiplexing networks. Recently, it has also shown the application of wavelength multicasting in the all-optical spectrum defragmentation in elastic optical networks (EON) [27]. All-optical multicast through the nonlinearities in HNLF [28–30], SOA [31] and silicon nanowire waveguide [32] has been reported.

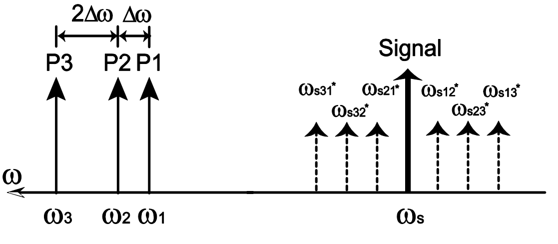


Figure 15. Operation principle of wavelength multicasting based on FWM in HNLF.

With the proposed coherent pumping, it is possible to achieve pump-phase-noise-free wavelength multicasting as well. Figure 15 illustrates the operation principle of the proposed pump-phase-noise-free wavelength multicasting scheme based on FWM with coherent multi-carrier pump. With the input three pumps at ω_1, ω_2, ω_3, and input signal at ω_s, seven multicasted channels, including the original input signal, are uniformly generated with a spacing of $\Delta\omega$. The frequency spacing settings of $\Delta\omega$ and $2\Delta\omega$ between ω_1 and ω_2, ω_2 and ω_3 could efficiently avoid the overlapping of spectrum among multicasted channels. It finally leads to a uniform frequency allocation of the multicasted signals alongside of the input signal with a spacing of $\Delta\omega$, which is important to realizing all-optical spectrum defragmentation [27]. The generated six components next to input signal are the non-degenerate FWM components with the frequencies of ω_{sij^*}, where i,j $\in$[1, 2, 3], i≠j, and * symbolizes the conjugate operation. The following equation shows the resultant phase in the multicasted signal at ω_{sij^*}:

$$\theta_{\text{output}} = \theta_{\text{input}} \pm \left(\Delta\theta_{\text{pi}} - \Delta\theta_{\text{pj}}\right) + C = \theta_{\text{input}} \pm \Delta\theta_{\text{pump}} + C \quad (3)$$

where $\Delta\theta_{pump}=\Delta\theta_{pi}-\Delta\theta_{pj}$, and θ_{output}, θ_{input}, $\Delta\theta_{pi}$, $\Delta\theta_{pj}$, and C are the phase of the output and input signals, the phase noise from pump i, j where i, j$\in$[1, 2, 3], and a constant term, respectively. When the pumps are coherent in phase, it is obvious that the phase noise from pumps are eliminated in the multicasted signals, i.e. $\Delta\theta_{pump}=0$. Therefore, the wavelength multicasting becomes tolerant against the phase noise from the pumps. Hence, lower-cost DFB lasers can be used as pump source. In practice, an optical comb with a spacing of $\Delta\omega$ could be employed to generate the coherent multi-carrier pump followed by an optical processor. The optical processor could ether be a liquid crystal on silicon (LCoS) device or cascaded band-pass and notch filters to select coherent carriers with desired spacing. The multicasting scale and the channel spacing of multicasted signals could be simply re-configured by programming the optical processor. The coherent pumping concept has been applied to wavelength conversion to remove the phase noise from the local pumps [18]. It is more beneficial and cost-effective when coherent pumping scheme is extended to multicasting with the flexible coherent multi-carrier pumping.

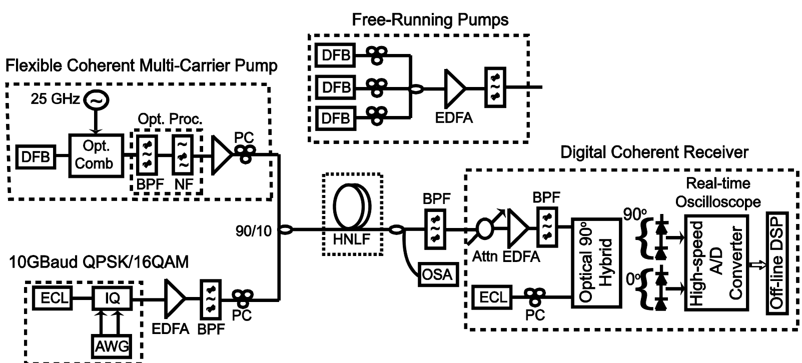


Figure 16. Experimental setup of pump-phase-noise-free 1 to 7 wavelength multicasting based FWM in HNLF.

To verify the proposed pump-phase-noise-free wavelength multicasting, a 1-to-7 multicasting experiment for QPSK and 16QAM signals was conducted with the setup shown in Fig. 16. Different from the setup shown in Fig. 9, a coherent multi-carrier pump is used as pump source, and a piece of highly-nonlinear fiber (HNLF) with length of 150 m is deployed as nonlinear media. The deployed HNLF has an attenuation coefficient of 0.9dB/km, a nonlinear coefficient of 18/W/km, a zero-dispersion wavelength of 1548 nm, a dispersion slope of around 0.02ps/nm^2/km and low β4 (2×10^{-56}s^4/m). Thanks to its high nonlinearity, a short length of HNLF (150 m) is sufficient to achieve the FWM-based wavelength multicasting. To retain the coherence of pumps, the short lengths, low and flat dispersion profile of the deployed HNLF are helpful to maintain the coherence of the pumps when propagating in HNLF. The constellations of the input and multicasted QPSK and16QAM signals with different pumping configurations are shown in Fig. 17. Even using DFB as pump laser, clear constellations are observed with coherent 3-carrier pumping. However, in the case of free-running DFB pumping, for the newly-produced components, clear symbol spreading around the unit circle occurred due to the phase noise from DFB pumps. It happens especially for the outer symbols with higher amplitude in 16QAM. The measured BER curves as function of OSNR (0.1 nm) is depicted in Fig. 18. For both QPSK and 16QAM, less than 0.8 dB power penalty was obtained at BER=10^{-3} for all of the seven multicasted signals with respect to the input signal with coherent pumping. On the other hand, owing to the strong phase noise transferred from the noisy pumps with free-running DFB pumping, error-floor at BER of 1×10^{-3} and 4×10^{-3} was observed for QPSK and 16QAM, respectively. The effectiveness of the elimination of the pump phase noise is verified in the multicasting for QAM signals with coherent multi-carrier pumping.

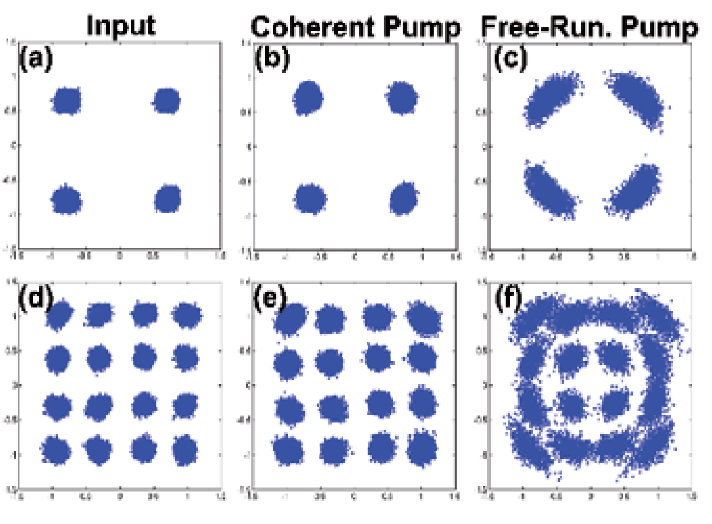


Figure 17. Measured constellations of input and converted signals with coherent pumping and free-running pumping schemes. QPSK: (a)~(c); 16QAM: (d)~(f).

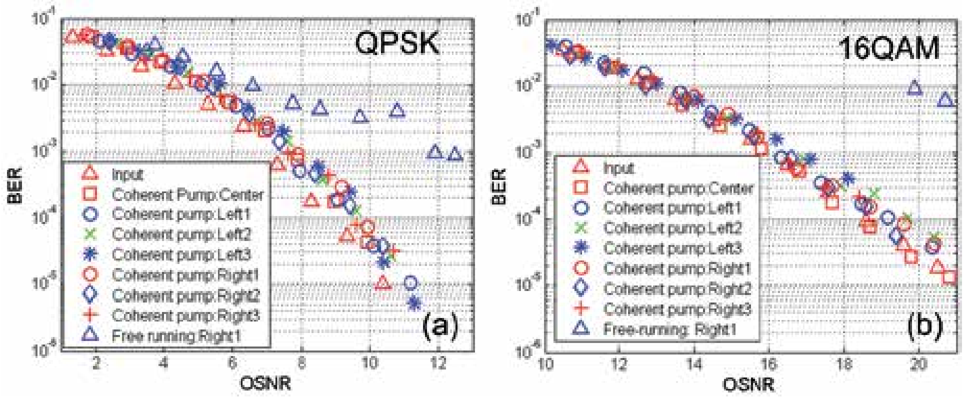


Figure 18. Measured BER vs. OSNR curves for (a) QPSK and (b) 16QAM multicasting systems.

3.3. Summary

In this section, in order to avoid the phase noise introduced from local pumps, coherent pumping concept has been proposed. Through experimental demonstration based on either cascaded second-order nonlinearities in PPLN or third-order nonlinearity in HNLF, we have successfully demonstrated that, even using low-cost noisy DFB lasers as pump source, the phase noise from local pumps could be effectively avoided in optical signal processing for high-order QAM signals, including wavelength conversion, wavelength exchange, and wavelength multicasting. However, in cases of free-running DFB pumps, it is impossible to obtain clear constellations for QAM signals, especially for 16QAM and 64QAM signals, which was significantly deteriorated by the large phase noise from DFB pumps.

4. Future works

To properly conduct optical signal processing for advanced high-order QAM signals, several issues have been addressed in this chapter. We also discussed proposed coherent pumping schemes for realizing the phase-noise-free optical signal processing for high-order QAM signals. For further study and investigation, the following aspects could be considered.

1. Phase-noise-free processing for multi-carrier high-order signals

Here, the study and investigation of phase-noise-free optical signal processing are mainly focusing on the single-carrier high-order modulation formats like high-order QAMs. It is also applicable to the multi-carrier high-order signals, such as coherent optical OFDM (CO-OFDM) with subcarriers modulated in high-order QAMs. Such QAM-CO-OFDM also suffers from the susceptibility against phase noise, especially for CO-OFDM with high-order QAM subcarrier modulations [33]. Therefore, it is highly desirable to further apply the coherent pumping concept to demonstrate the phase-noise-free processing for multi-carrier high-order signals in the near future.

2. Reconfigurable coherent optical multi-carrier

As we point out in the above sections, coherent optical multi-carrier could be produced by an optical comb followed by optical signal processing, which is usually an LCoS-based component. It is cost effective to share multi-carrier for multi-channel signal processing. However, it is still costly to include LCoS-based optical processor in optical signal processing subsystems. Thus, it is interesting to further develop cost-effective coherent multi-carrier with reconfigurable tone number and spacing. It will be one of key components for realizing reconfigurable optical signal processing in the future.

3. Other nonlinear media to realize optical signal processing

The experimental demonstration reported here is mainly focusing on HNLF and PPLN devices. Obviously, it could also be implemented in other nonlinear media like SOA especially quantum-dot SOA [34], and silicon waveguides [35].

5. Conclusion

Local pump lasers are indispensable for conducting optical signal processing for optical signals. For phase-noise-sensitive advanced modulation formats, the phase noise from local pumps are critical to be considered in order to realize superior optical signal processing. In this chapter, optical signal processing technology for high-order QAM signals has been discussed, with focus on wavelength conversion, wavelength exchange and wavelength multicasting for high-order QAM signals. To design high-performance optical signal processing subsystems, both linear and nonlinear phase noise and distortions are the main concerns in the system design. We first investigated the effective monitoring approach to optimize the performance of wavelength conversion for avoiding undesired nonlinear phase noise and distortions. Then, in the following sections, we discussed our proposed coherent pumping scheme to eliminate the linear phase noise from local pumps in order to realize pump-phase-noise-free wavelength conversion, wavelength exchange and multicasting for high-order QAM signals. Experimental demonstrations were present to verify the feasibility of the proposed coherent pumping schemes.

Acknowledgements

We acknowledge Takahide Sakamoto, Tetsuya Kawanishi, André Albuquerque, Benjamin J. Puttnam, Miguel Drummond, Rogério Nogueira, Atsushi Kanno, Satoshi Shinada, Naoya Wada for their collaborations, and the generous support of Grant-in-Aid for Scientific Research (C) (15K06033), and Grant-in-Aid for Young Scientist (A) (25709031) from the Ministry of Education, Science, Sports and Culture (MEXT), Japan.

Author details

Guo-Wei Lu*

Address all correspondence to: gordon.guoweilu@gmail.com

Tokai University, Japan

References

[1] X. Zhou and J. Yu, "Multi-level, multi-dimensional coding for high-speed and high spectral-efficiency optical transmission," J. Lightw. Technol., 27(16): 3641–3653, 2009.

[2] X. Zhou, J. Yu, M. F. Huang, Y. Shao, T. Wang, L. Nelson, P. D. Magill1, M. Birk, P. I. Borel, D. W. Peckham and R. Lingle "64-Tb/s, 8 b/s/Hz, PDM-36QAM transmission over 320 km using both pre- and post-transmission digital signal processing," J. Lightw. Technol., 29(4): 571–577, 2011.

[3] W. Peng, H. Takahashi, T. Tsuritani, and I. Morita, "DAC-free Generation and 320-km Transmission of 11.2-GBd PDM-64QAM Using a Single I/Q Modulator," in European Conference and Exhibition on Optical Communication, OSA Technical Digest (online) (Optical Society of America, 2012), paper We.1.C.3.

[4] A. Sano, T. Kobayashi, A. Matsuura, S. Yamamoto, S. Yamanaka, E. Yoshida, Y. Miyamoto, M. Matsui, M. Mizoguchi and T. Mizuno "100x120-Gb/s PDM 64-QAM transmission over 160 km using linewidth-tolerant pilotless digital coherent detection," European Conference in Optical Communications, 2010, paper PD2_4.

[5] G.-W. Lu, T. Sakamoto, and T. Kawanishi, "Flexible high-order QAM transmitter using tandem IQ modulators for generating 16/32/36/64-QAM with balanced complexity in electronics and optics," Opt. Express 21: 6213–6223, 2013.

[6] S. R. Nuccio, Z. Bakhtiari, O. F. Yilmaz, and A. Willner, "λ-Conversion of 160-Gbit/s PDM 16-QAM Using a Single Periodically-Poled Lithium Niobate Waveguide," in Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2011, OSA Technical Digest (CD) (Optical Society of America, 2011), paper OWG5.

[7] T. Richter, R. Nouroozi, H. Suche, W. Sohler, C. Schubert, "PPLN-Waveguide Based Tunable Wavelength Conversion of QAM Data Within the C-Band," IEEE Photonics Technology Letters, IEEE, 25(21): 2085–2088, Nov.1, 2013.

[8] X. Li, J. Yu, Z. Dong, and N. Chi, "Wavelength conversion of 544-Gbit/s dual-carrier PDM-16QAM signal based on the co-polarized dual-pump scheme," Opt. Express 20: 21324–21330, 2012.

[9] A. H. Gnauck, E. Myslivets, M. Dinu, B. P. P. Kuo, P. Winzer, R. Jopson, N. Alic, A. Konczykowska, F. Jorge, J. Dupuy, and S. Radic, "All-Optical Tunable Wavelength Shifting of a 128-Gbit/s 64-QAM Signal," in European Conference and Exhibition on Optical Communication, OSA Technical Digest (online) (Optical Society of America, 2012), paper Th.2.F.2.

[10] B. Filion, S. Amiralizadeh, A. T. Nguyen, L. A. Rusch, and S. LaRochelle, "Wideband Wavelength Conversion of 16 Gbaud 16-QAM Signals in a Semiconductor Optical Amplifier," in Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013, OSA Technical Digest (online) (Optical Society of America, 2013), paper OTh1C.5.

[11] G. Contestabile, Y. Yoshida, A. Maruta, and K. Kitayama, "100 nm-Bandwidth Positive-Efficiency Wavelength Conversion for m-PSK and m-QAM signals in QD-SOA," in Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013, OSA Technical Digest (online) (Optical Society of America, 2013), paper OTh1C.6.

[12] B. Filion, W. C. Ng, An. T. Nguyen, L. A. Rusch, and S. LaRochelle, "Wideband wavelength conversion of 16 Gbaud 16-QAM and 5 Gbaud 64-QAM signals in a semiconductor optical amplifier," Opt. Express 21: 19825–19833, 2013.

[13] E. Ip, A. P. T. Lau, D. J. F. Barros, and J. M. Kahn, "Coherent detection in optical fiber systems," Opt. Express 16: 753–791, 2008.

[14] Guo-Wei Lu, Takahide Sakamoto, and Tetsuya Kawanishi, "Wavelength conversion of optical 64QAM through FWM in HNLF and its performance optimization by constellation monitoring," Opt. Express 22: 15–22, 2014.

[15] S. Moro, A. Peric, N. Alic, B. Stossel, and S. Radic, "Phase noise in fiber-optic parametric amplifiers and converters and its impact on sensing and communication systems," Opt. Express 18: 21449–21460, 2010.

[16] R. Elschner, and L. Petermann, "Impact of Pump-Induced Nonlinear Phase Noise on Parametric Amplification and Wavelength Conversion of Phase-Modulated Signals," European Conference in Optical Communications, 2009, paper 3.3.4.

[17] T. Richter, R. Elschner, L. Molle, K. Petermann, and C. Schubert, "Coherent Receiver-Based Compensation of Phase Distortions Induced by Single-Pump HNLF-Based FWM Wavelength Converters," in Integrated Photonics Research, Silicon and Nanophotonics and Photonics in Switching, OSA Technical Digest (CD) (Optical Society of America, 2010), paper PWB2.

[18] G. Lu, A. Albuquerque, B. Puttnam, T. Sakamoto, M. Drummond, R. Nogueira, A. Kanno, S. Shinada, N. Wada, and T. Kawanishi, "Pump-linewidth-tolerant optical wavelength conversion for high-order QAM signals using coherent pumps," Opt. Express 22: 5067–5075 (2014).

[19] G. Lu, A. Albuquerque, B. Puttnam, T. Sakamoto, M. Drummond, R. Nogueira, A. Kanno, S. Shinada, N. Wada, and T. Kawanishi, “Pump-Linewidth-Tolerant Optical Data Exchange between 16QAM and QPSK with 50-GHz Channel-Spacing using Coherent DFB Pump,” in Proc. European Conference in Optical Communications, 2010, paper P.2.15.

[20] G. Lu, T. Sakamoto, and T. Kawanishi, "Pump-Phase-Noise-Tolerant Wavelength Multicasting for QAM Signals using Flexible Coherent Multi-Carrier Pump," in Optical Fiber Communication Conference, OSA Technical Digest (online) (Optical Society of America, 2015), paper M2E.2.

[21] K. Uesaka, Kenneth Kin-Yip Wong, M.E. Marhic, Leonid G. Kazovsky, "Wavelength exchange in a highly nonlinear dispersion-shifted fiber: theory and experiments," IEEE Journal of Selected Topics in Quantum Electronics 8(3): 560–568, 2002.

[22] C. H. Kwok, Bill P. P. Kuo, and Kenneth K. Wong, "Pulsed pump wavelength exchange for high speed signal de-multiplexing," Opt. Express 16: 10894–10899, 2008.

[23] Mengzhe Shen, Xing Xu, T.I. Yuk, Kenneth Kin-Yip Wong, "A 160-Gb/s OTDM Demultiplexer Based on Parametric Wavelength Exchange," IEEE Journal of Quantum Electronics 45(11): 1309–1316, 2009.

[24] X. Xu, M. Shen, T. I. Yuk, and K. K. Y. Wong, "Optical Time-Slot Swapping Based on Parametric Wavelength Exchange," in Asia Communications and Photonics Conference and Exhibition, Technical Digest (CD) (Optical Society of America, 2009), paper TuC1.

[25] Jian Wang, Zahra Bakhtiari, Omer F. Yilmaz, Scott Nuccio, Xiaoxia Wu, and Alan E. Willner, "10 Gbit/s tributary channel exchange of 160 Gbit/ssignals using periodically poled lithium niobate," Opt. Lett. 36: 630–632, 2011.

[26] Jian Wang, Zahra Bakhtiari, Scott R. Nuccio, Omer F. Yilmaz, Xiaoxia Wu, and Alan E. Willner, "Phase-transparent optical data exchange of 40 Gbit/s differential phase-shift keying signals," Opt. Lett. 35: 2979–2981; 2010.

[27] Yingying Xu, Juhao Li, Paikun Zhu, Bingli Guo, Yuanxiang Chen, Yucheng Zhong, Yan Wang, Zhangyuan Chen, and Yongqi He, “Demonstration of All-optical Inverse Multiplexing in Elastic Optical Networks,” in Optical Fiber Communication Conference (2014), paper Th1E.6,

[28] G.-W. Lu, K. S. Abedin, and T. Miyazaki, “DPSK multicast using multiple-pump FWM in Bismuths highly nonlinear fiber with high multicast efficiency,” Opt. Express 16(26): 21964–21970, 2008.

[29] C. Zhiyu, Y. Lianshan, P. Wei, L. Bin, Y. Anlin, G. Yinghui, and L. Ju Han, “One-to-nine multicasting of RZ-DPSK based on cascaded four-wave mixing in a highly nonlinear fiber without stimulated Brillouin scattering suppression,” IEEE Photon. Technol. Lett. 24(20): 1882–1885, 2012.

[30] M. P. Fok and C. Shu, "Multipump four-wave mixing in a photonic crystal fiber for 6 ×10 Gb/s wavelength multicasting of DPSK signals," IEEE Photon. Technol. Lett. 19(15): 1166–1168, 2007.

[31] M. Pu, H. Hu, H. Ji, M. Galili, L. K. Oxenløwe, P. Jeppesen, J. M. Hvam, and K. Yvind, "One-to-six WDM multicasting of DPSK signals based on dual-pump four-wave mixing in a silicon waveguide," Opt. Express 19(24): 24448–24453, 2011.

[32] W. Dawei, C. Tee-Hiang, Y. Yong-Kee, X. Zhaowen, W. Yixin, X. Gaoxi, and L. Jian-guo, "Performance comparison of using SOA and HNLF as FWM medium in a wavelength multicasting scheme with reduced polarization sensitivity," J. Lightwave Technol. 28(24): 3497–3505, 2010.

[33] W. Shief, "Maximum-likelihood phase and channel estimation for coherent optical OFDM," Photon. Technol. Lett., 20(8), 605, 2008.

[34] Jun Qin, Guo-Wei Lu, Takahide Sakamoto, Kouichi Akahane, Naokatsu Yamamoto, Danshi Wang, Cheng Wang, Hongxiang Wang, Min Zhang, Tetsuya Kawanishi, and Yuefeng Ji, "Simultaneous multichannel wavelength multicasting and XOR logic gate multicasting for three DPSK signals based on four-wave mixing in quantum-dot semiconductor optical amplifier," Opt. Express 22: 29413–29423, 2014.

[35] Mohammadreza Khorasaninejad and Simarjeet Singh Saini, "Silicon nanowire optical waveguide (SNOW)," Opt. Express 18: 23442–23457, 2010.

Chapter 2

High-Base Optical Signal Proccessing

Jian Wang and Alan E. Willner

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/61504

Abstract

Optical signal processing is a promising technique to enable fast data information processing in the optical domain. Traditional optical signal processing functions pay more attention to binary modulation formats (i.e., binary numbers) with single-bit information contained in one symbol. The ever-growing data traffic has propelled great success in high-speed optical signal transmission by using advanced multilevel modulation formats (i.e., high-base numbers), which encode multiple-bit information in one symbol with resultant enhanced transmission capacity and efficient spectrum usage. A valuable challenge would be to perform various optical signal processing functions for multilevel modulation formats, i.e., high-base optical signal processing. In this chapter, we review recent research works on high-base optical signal processing for multilevel modulation formats by exploiting degenerate and nondegenerate four-wave mixing in highly nonlinear fibers or silicon photonic devices. Grooming high-base optical signal processing functions including high-base wavelength conversion, high-base data exchange, high-base optical computing, and high-base optical coding/decoding are demonstrated. High-base optical signal processing may facilitate advanced data management and superior network performance.

Keywords: High-base optical signal processing, multilevel modulation format, four-wave mixing, wavelength conversion, data exchange, optical computing, coding/decoding

1. Introduction

The arrival of the era of big data has fuelled the increasing demand on both high-speed signal transmission and fast signal processing, which are known as two themes of great importance for optical communications. The advances in fiber-optic technologies have resulted in great success in delivering high-speed data signals in optical fiber transmission links [1-5]. The rapid development of photonics technologies has also promoted increasing interest for optical signal processing, which is regarded as a promising solution to facilitate high-speed signal processing in the optical domain and to eliminate complicated, inefficient, low-latency, and power-consuming optical-to-electrical-to-optical (O-E-O) conversions [6]. At network nodes of advanced photonic networks, different optical signal processing functions might be required to enable increased network flexibility and efficiency. Remarkably, nonlinear optics has offered great potential to develop optical signal processing in high-speed photonic networks using various optical nonlinearities [6-20]. Miscellaneous optical signal processing functions have been demonstrated, such as wavelength conversion, wavelength (de)multiplexing, wavelength multicasting, data exchange, add/drop, optical addressing, optical switching, optical logic gate, optical computing, optical format conversion, optical correlation, optical equalization, optical regeneration, tunable optical delay, optical coding/decoding, etc. [21-53]. These optical signal processing operations are enabled by exploiting different nonlinear effects in different nonlinear optical devices. The typical nonlinear effects include cross-gain modulation (XGM), self-phase modulation (SPM), cross-phase modulation (XPM), two-photon absorption (TPA), degenerate and nondegenerate four-wave mixing (FWM), second-harmonic generation (SHG), sum-frequency generation (SFG), difference-frequency generation (DFG), cascaded second-harmonic generation and difference-frequency generation (cSHG/DFG), and cascaded sum- and difference-frequency generation (cSFG/DFG). Typical nonlinear optical devices based on different platforms include semiconductor optical amplifiers (SOAs), highly nonlinear fibers (HNLFs), periodically poled lithium niobate (PPLN) waveguides, chalcogenide (As_2S_3) waveguides, silicon waveguides, and photonic crystal waveguides. It is noted that most of previous research efforts are dedicated to optical signal processing for binary modulation formats such as on–off keying (OOK), differential phase-shift keying (DPSK), and binary phase-shift keying (BPSK). Despite favorable operation performance achieved for binary optical signal processing, it suffers limited bitrate and low spectral efficiency since only single-bit information is carried by each symbol for binary modulation formats.

With the rapid growth of global broadband and mobile data traffic, high transmission capacity and high spectral efficiency are highly desirable. Fortunately, recent advances in multilevel modulation formats, coherent detection, and digital signal processing have led to tremendous increase in transmission capacity and spectral efficiency [54-63]. Beyond great progress in high-speed signal transmission, processing multilevel modulation formats in the optical domain could be another interesting topic compatible with superior network performance and advanced data management. Typically, multilevel modulation formats contain multiple bits in one symbol, e.g., 2, 3, and 4 bits in one symbol for quadrature phase-shift keying (QPSK), 8-ary phase-shift keying (8PSK), star 8-ary quadrature amplitude modulation (Star-8QAM), 16-ary phase-shift keying (16PSK), star 16-ary quadrature amplitude modulation (Star-16QAM), and square 16-ary quadrature amplitude modulation (Square-16QAM) (Fig. 1). Moreover, multiple points in the constellation plane can be used to represent high-base numbers, e.g., quaternary number for QPSK, octal numbers for 8PSK and Star-8QAM, and hexadecimal numbers for 16PSK, Star-16QAM and Square-16QAM (Fig. 1). Despite great success in transmission links using multilevel modulation formats [64-69], there have been relatively limited research efforts dedicated to their manipulation in the optical domain (i.e., high-base optical signal processing). In this scenario, a laudable goal would be to develop miscellaneous high-base optical signal processing functions for multilevel modulation formats

[70-86]. The aforementioned optical nonlinearities in various nonlinear optical devices would be promising candidates to facilitate grooming high-base optical signal processing operations.

In this chapter, we provide a comprehensive report of our recent research works on high-base optical signal processing for multilevel modulation formats by exploiting optical nonlinearities [71, 73-75, 77, 79, 80, 83, 85, 86]. The demonstrated high-base optical signal processing functions include wavelength conversion using degenerate FWM in a silicon waveguide [83], data exchange using degenerate/nondegenerate FWM in HNLFs or silicon-organic hybrid slot waveguides [71, 73, 74, 86], optical computing using degenerate/nondegenerate FWM in HNLFs or silicon–organic hybrid slot waveguides [75, 77, 80, 85], and optical coding/decoding using degenerate FWM in HNLFs [79].

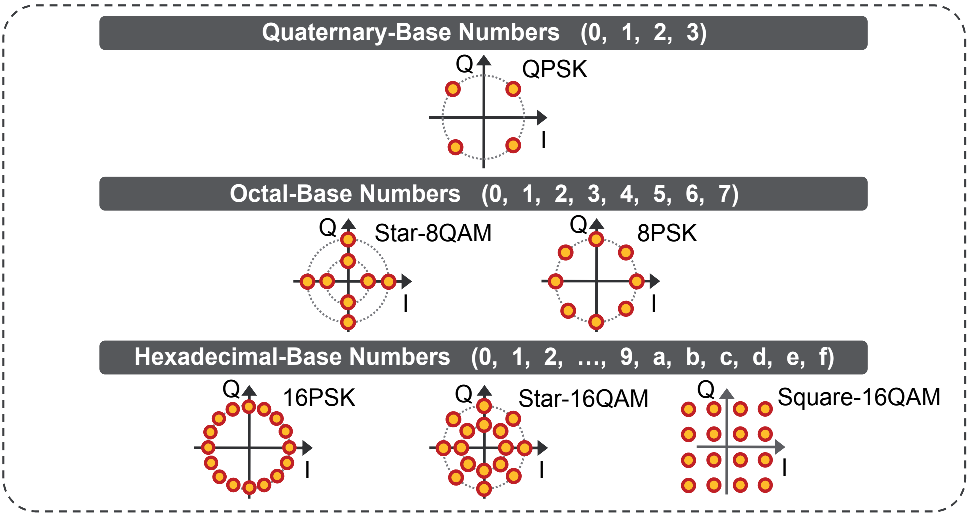


Figure 1. Schematic constellations of advanced multilevel modulation formats representing high-base (quaternary, octal, hexadecimal) numbers (QPSK, Star-8QAM/8PSK, 16PSK/Star-16QAM/Square-16QAM).

2. High-base wavelength conversion [83]

We demonstrate high-base all-optical wavelength conversions of multicarrier, multilevel modulation signals based on degenerate FWM in a silicon waveguide. Coherent multicarrier, multilevel modulations, i.e., orthogonal frequency-division multiplexing (OFDM) combined with advanced multilevel quadrature amplitude modulation (mQAM), are employed in the experiment.

Shown in Fig. 2(a) is the schematic cross section of a typical silicon waveguide. The calculated mode distribution using finite element method (FEM) is depicted in Fig. 2(b), from which one can see the tight light confinement in the top silicon region due to the high contrast index of the silicon waveguide. The measured scanning electron microscope (SEM) images of the fabricated silicon waveguide and grating coupling region are shown in Fig. 2(c) and (d). We fabricate the silicon waveguide on a silicon-on-insulator (SOI) wafer, on the top of which the silicon thickness is 340 nm with a 2-μm-thick buried oxide (BOX) layer. Using electron-beam lithography (EBL), followed by induced coupled plasma (ICP) etching, the desired silicon waveguide is formed for on-chip, high-base wavelength conversion.

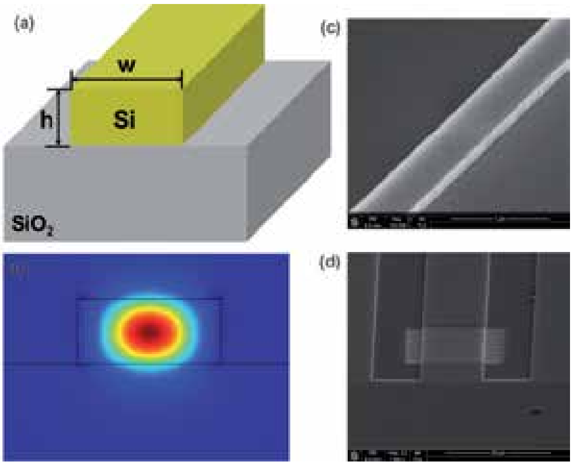


Figure 2. (a) Cross section and (b) calculated mode distribution of a typical silicon waveguide. (c)(d) Measured scanning electron microscope (SEM) images of the fabricated silicon waveguide and grating coupling region.

Figure 3 illustrates the wavelength conversion process based on degenerate FWM in a silicon waveguide. One OFDM m-QAM carrying data signal and one continuous-wave (CW) pump are launched into the silicon waveguide. When propagating along the silicon waveguide, pump photons are annihilated to create signal photons and newly converted idler photons through degenerate FWM process. At the output of the silicon waveguide, the converted idler takes the OFDM m-QAM data information carried by the input signal and the wavelength conversion from input signal to output idler is achieved. It is noted that the performance degradation of high-base wavelength conversion by degenerate FWM process can be ascribed to the accumulated phase noise transferred from the input pump and signal. Since the constellations of higher-order modulations (e.g., 16/32/64/128-QAM) inherently have a smaller phase noise tolerance due to the smaller spacing between adjacent constellation points, it is challengeable to realize high-base wavelength conversion of OFDM m-QAM signals, especially for higher-order modulations such as OFDM 16/32/64/128-QAM.

Shown in Fig. 4 is the experimental setup for high-base wavelength conversion of OFDM 16/32/64/128-QAM signals using a silicon waveguide. At the transmitter, an external cavity laser (ECL1) at 1563.849 nm is modulated by a single-polarization optical I/Q modulator. An arbitrary waveform generator (AWG) running at 10 GS/s sampling rate is used to produce the electrical OFDM m-QAM signal (m=16, 32, 64, 128). The transmitted OFDM signal is generated off-line from a data sequence of 2^{31}-1 pseudo random binary sequences (PRBS) and then mapped onto m-QAM constellation. The OFDM m-QAM signal is constructed by 82 subcar-

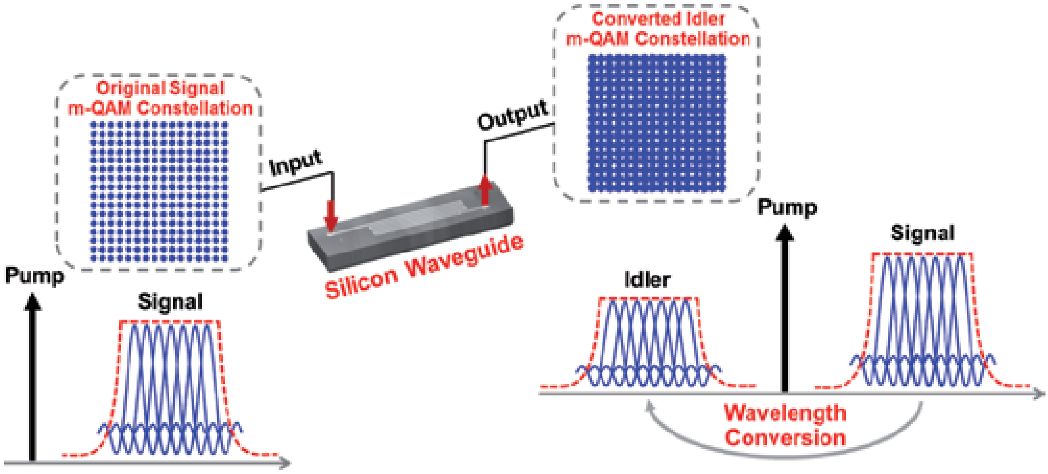


Figure 3. Illustration of high-base wavelength conversion of OFDM m-QAM signals based on degenerate FWM in a silicon waveguide.

riers, in which 78 subcarriers are used to carry the payloads with m-QAM signal, while 4 subcarriers are selected as the pilots with 4-QAM loading to estimate the phase noise. Another inverse fast Fourier transform (IFFT) with a size of 256 is used to convert the signal to time domain. No cyclic prefix (CP) is used as the signal passes through a system without dispersion-dominated devices. For the channel estimation, 10 training symbols are used for every 468 payload symbols in a manner of [A 0], where "A" denotes one OFDM m-QAM symbol. Another ECL (ECL2) employed as the pump is set at 1560.61 nm with a 6-dBm output power. Two polarization controllers (PC1, PC2) are used to adjust the polarization states of signal and pump, respectively. After the signal amplification by an erbium-doped fiber amplifier (EDFA1) with a maximum output power of 27 dBm and pump amplification by a second EDFA (EDFA2) with a maximum output power of 30 dBm, the signal and pump are combined with a wavelength selective switch (WSS) and then vertically coupled into the silicon waveguide, in which degenerate FWM process takes place to enable the wavelength conversion from the signal to the converted idler. In the experiment, the signal is amplified to 25.5 dBm by EDFA1 and the pump is amplified to 27 dBm by EDFA2. The WSS not only combines the amplified signal and pump together but also suppresses the amplified spontaneous emission (ASE) noise from two EDFAs. After the wavelength conversion, the signal, pump, and newly converted idler are vertically coupled out from the silicon waveguide. After the amplification by a third EDFA (EDFA3), the converted idler is filtered using a tunable optical filter (TOF) with a bandwidth of 0.4 nm. A variable optical attenuator (VOA) and one more EDFA (EDFA4) are employed to adjust the received optical signal-to-noise ratio (OSNR) for proper detection by the coherent receiver. At the receiver, the optical signal is first mixed with a local oscillator (LO) by an optical hybrid and detected by a typical balanced coherent receiver. The line width of the employed laser sources including ECL1, ECL2, and LO in the experiment is around 100 kHz. The obtained two radio frequency (RF) signals for the IQ components are sent into a Tektronix real-time digital oscilloscope acquired at 50 GS/s and processed off-line with a

MATLAB program. The offline digital processing of the received signal includes: 1) carrier frequency offset estimation and OFDM window synchronization; 2) fast Fourier transform (FFT); 3) channel estimation; 4) phase noise estimation (crucial to m-QAM signal); 5) constellation decision and bit-error rate (BER) calculation.

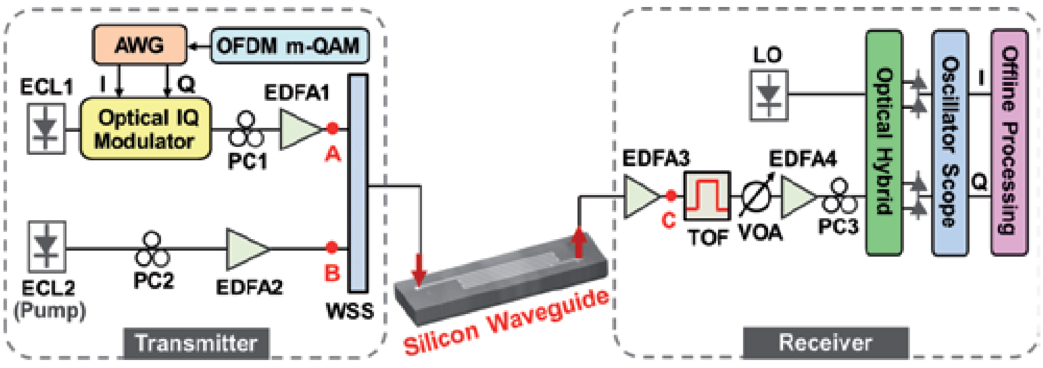


Figure 4. Experimental setup for high-base wavelength conversion of OFDM m-QAM signals using a silicon waveguide. ECL: external cavity laser; AWG: arbitrary waveform generator; PC: polarization controller; TOF: tunable optical filter; VOA: variable optical attenuator; LO: local oscillator; EDFA: erbium-doped fiber amplifier.

In order to characterize the performance of high-base wavelength conversion of OFDM m-QAM signals, we measure the BER curves as a function of received OSNR for back-to-back (B-to-B) and converted idler. Shown in Fig. 5(a)-(d) are measured BER performance for high-base wavelength conversions of OFDM 16-QAM, OFDM 32-QAM, OFDM 64-QAM, and OFDM 128-QAM, respectively. As shown in Fig. 5(a), for OFDM 16-QAM wavelength conversion the required OSNR at the 7% forward error correction (FEC) threshold (BER=1×10^{-3}) is 7.8 and 10.8 dB for the B-to-B signal and converted idler, respectively. The observed OSNR penalty is around 3 dB for OFDM 16-QAM wavelength conversion. Similarly, the received OSNR penalties of ~4 dB at 7% FEC threshold in Fig. 5(b), ~3.5 dB in Fig. 5(c) at 20% FEC threshold and ~4.5 dB in Fig. 5(d) at 20% FEC threshold are observed for high-base wavelength conversions of OFDM 32-QAM, OFDM 64-QAM, and OFDM 128-QAM, respectively. The right insets of Fig. 5(a)-(d) depict corresponding constellations of the B-to-B signals and converted idlers at the given OSNR values. One can see clear constellations of converted idlers, indicating favorable operation performance achieved for on-chip, high-base, all-optical wavelength conversions of multicarrier, multilevel modulation (OFDM 16/32/64/128-QAM) signals using a silicon waveguide.

3. High-base optical data exchange [71, 73, 74, 86]

We propose and demonstrate high-base all-optical data exchange of advanced multilevel modulation signals based on degenerate/nondegenerate FWM in HNLFs or silicon–organic hybrid slot waveguides.

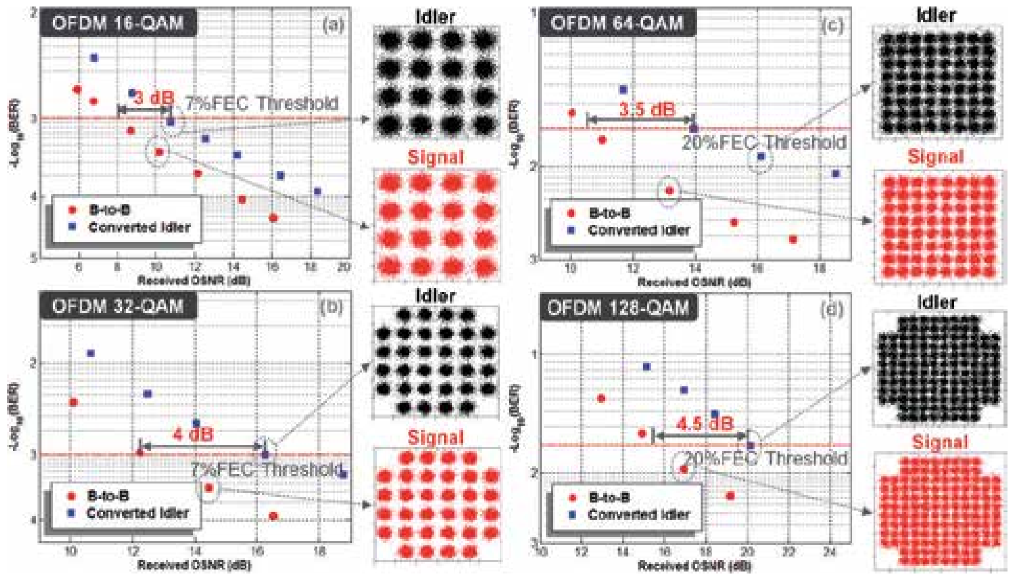


Figure 5. Measured BER versus received OSNR for high-base wavelength conversions of multicarrier, multilevel modulation signals. (a) OFDM 16-QAM. (b) OFDM 32-QAM. (c) OFDM 64-QAM. (d) OFDM 128-QAM.

We first demonstrate high-base optical data exchange of 100-Gbit/s return-to-zero differential QPSK (RZ-DQPSK) signals. The concept and principle for high-base optical data exchange of DQPSK modulation signals between two different wavelengths (S1:λ_{S1}, S2:λ_{S2}) are depicted in Fig. 6. The four-level phase information carried by two DQPSK signals at different wavelengths is swapped after the data exchange, as shown in Fig. 6(a). To perform high-base optical data exchange of DQPSK signals carrying phase information, the optical data exchange operation is expected to be phase transparent. Using the parametric depletion effect in a single HNLF, one may realize phase-transparent optical data exchange. Figure 6(b) depicts the principle of operation of parametric depletion. Two CW pumps (P1:λ_{P1}, P2:λ_{P2}) and signal 1 (S1:λ_{S1}) are fed into the HNLF. P1 and S1 are symmetrical about the zero-dispersion wavelength (ZDM) of HNLF. When propagating along the HNLF, the photons of P1 and S1 are annihilated to create the photons of P2 and S2 ($1/\lambda_{S2}+1/\lambda_{P2}=1/\lambda_{S1}+1/\lambda_{P1}$) by the nondegenerate FWM process. Thus, the parametric depletion of S1 is expected with its data information copied onto a newly generated S2. Similarly, the depletion of S2 accompanied by the creation of S1 is realized during the nondegenerate FWM process when sending two pumps and S2 into the HNLF. Figure 6(c) shows the principle of operation of optical data exchange. Two pumps and two signals are simultaneously launched into the HNLF. When P1(P2) and S1(S2) are almost symmetrical about the ZDW of HNLF, S1(S2) can be consumed to produce S2(S1) by appropriately adjusting the power of two pumps. As a consequence, one can implement optical data exchange between two signals (S1, S2).

Remarkably, under the nondepletion approximation and proper control of pump powers, one can easily derive linear relationships ($A_{S1}^{'} \propto A_{S2} \cdot A_{P2} \cdot A_{P1}^{*}$, $A_{S2}^{'} \propto A_{S1} \cdot A_{P1} \cdot A_{P2}^{*}$) of complex amplitudes between the output signals (A'_{S1}, A'_{S2}) and input signals and pumps (A_{S1}, A_{S2}, A_{P1}, A_{P2}). The linear complex amplitude relationships imply that nondegenerate FWM-based high-base data exchange has the characteristic of transparency to the modulation format including the phase transparency. We can further obtain the phase relationships of $\varphi_{S1}'=\varphi_{S2}+\varphi_{P2}-\varphi_{P1}$ and $\varphi_{S2}'=\varphi_{S1}+\varphi_{P1}-\varphi_{P2}$. It is worth noting that phase modulation is always applied to the pumps (φ_{P1}, φ_{P2}) to effectively suppress the stimulated Brillouin scattering (SBS) effect in HNLF. As a result, the pump power is efficiently utilized in the nondegenerate FWM process, which benefits the effective parametric depletion and data exchange. Remarkably, the pump phase transfer to the exchanged signals might cause serious trouble for the DQPSK data exchange. Fortunately, according to the deduced phase relationships, it is possible to cancel the pump phase transfer by applying the precisely identical phase modulation to the two pumps (i.e., φ_{P1}=φ_{P2}), which makes it possible to implement the high-base data exchange of DQPSK or other multilevel modulation signals containing phase information.

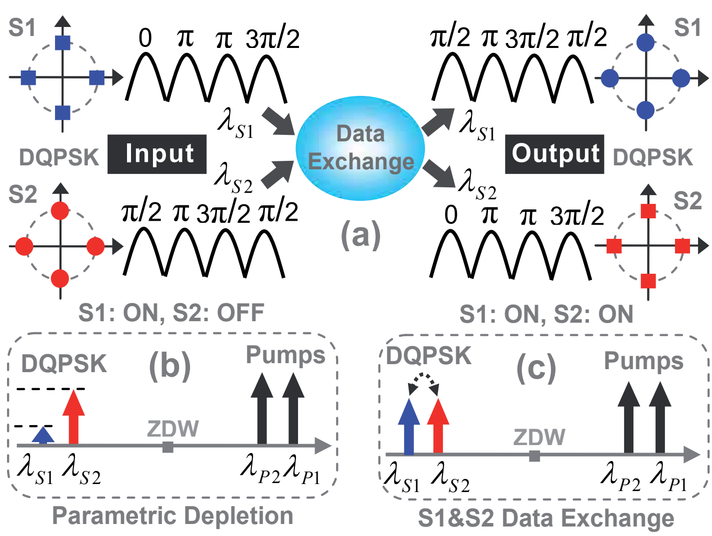


Figure 6. (a) Concept of high-base optical data exchange of DQPSK modulation signals. (b)(c) Principle of nondegenerate FWM-based parametric depletion and high-base optical data exchange.

In the experiment, two CW pumps (P1: 1564.4 nm, P2: 1558.6 nm) together with two 100-Gbit/s RZ-DQPSK signals (S1: 1539.4 nm, S2: 1545.4 nm) are coupled into a 1-km piece of HNLF with a nonlinear coefficient of 9.1 $W^{-1}\cdot km^{-1}$, a ZDW of ~1552 nm, and a fiber loss of 0.45 dB/km. The DQPSK optical data exchange is realized in the HNLF based on the parametric depletion effect of the nondegenerate FWM process. For the 100-Gbit/s DQPSK optical data exchange, shown in Fig. 7 are the measured temporal waveforms of demodulated channel I (Ch. I) and channel Q (Ch. Q). One can clearly see from Fig. 7 that after the nondegenerate FWM-based optical data exchange, the data information swapping between two 100-Gbit/s RZ-DQPSK signals is successfully implemented. Additionally, when looking at the temporal waveforms after wavelength conversion with only S1 or S2 present and the temporal waveforms after data exchange with both S1 and S2 present, the performance degradation of temporal waveforms after data exchange is observed with increased noise. Such phenomenon can be explained with the fact that the beating effect between the newly converted signal and original residual signal induces added noise.

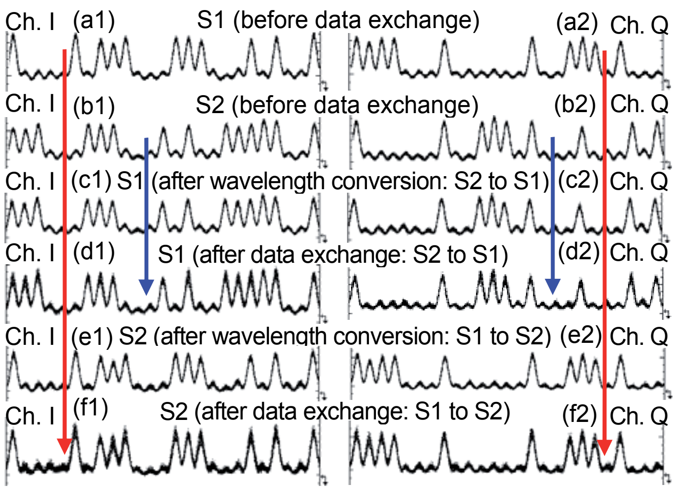


Figure 7. Measured temporal waveforms of demodulated channel I (Ch. I) and channel Q (Ch. Q) for high-base optical data exchange of 100-Gbit/s DQPSK signals. (a1)(a2) S1 is ON, P1 is OFF, and P2 is OFF. (b1)(b2) S2 is ON, P1 is OFF, and P2 is OFF. (c1)(c2) S2 to S1 wavelength conversion. S1 is OFF, S2 is ON, P1 is ON, and P2 is ON. (d1)(d2) S2 to S1 data exchange. S1 is ON, S2 is ON, P1 is ON, and P2 is ON. (e1)(e2) S1 to S2 wavelength conversion. S1 is ON, S2 is OFF, P1 is ON, and P2 is ON. (f1)(f2) S1 to S2 data exchange. S1 is ON, S2 is ON, P1 is ON, and P2 is ON

Shown in Fig. 8 is the measured BER performance and balanced eyes for high-base optical data exchange of 100-Gbit/s DQPSK signals. One can see from Fig. 8 that for wavelength conversion with only S1 or S2 and two pumps present, the power penalty is assessed to be less than 1.2 dB at a BER of 10^{-9}. In contrast, for data exchange with both two signals and two pumps present, the power penalty is measured to be less than 5 dB at a BER of 10^{-9}. It is expected that the extra power penalty of the high-base data exchange compared to the wavelength conversion could be due to the beating effect between the newly converted signal and the original residual signal.

We further investigate the tolerance of pump misalignment and the dynamic range of input signal power for the 100-Gbit/s RZ-DQPSK data exchange. Shown in Fig. 9 is the measured relative power penalty as a function of the pump misalignment. One can clearly see that the performance degradation of wavelength conversion and data exchange becomes severe when the pump misalignment is larger than +/-2 ps. Actually, under relatively large pump phase misalignment, the residual phase due to incomplete pump phase cancellation is transferred to the phase noise added to the wavelength converted signal and data exchanged signal, resulting in the degradation of operation performance. Under different pump phase misalignments, the measured typical balanced eyes of demodulated signals after data exchange are also shown in

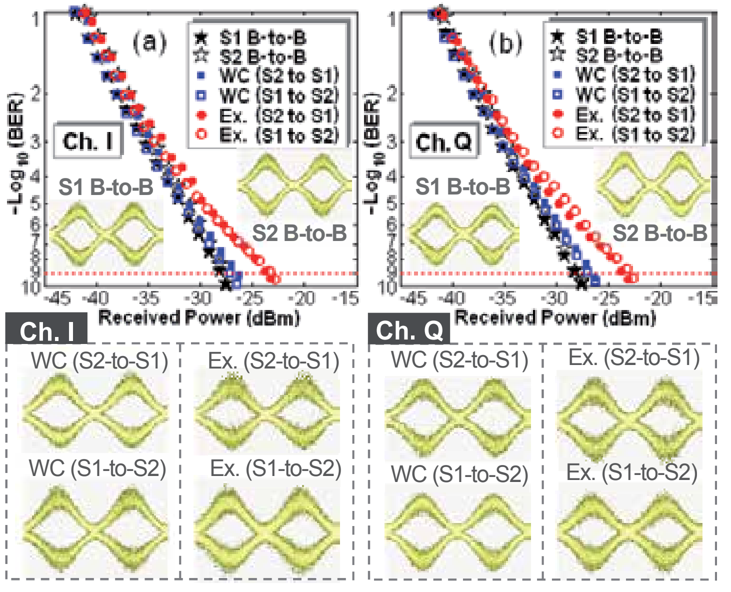


Figure 8. BER curves and balanced eyes for high-base data exchange of 100-Gbit/s DQPSK signals. (a) Ch. I. (b) Ch. Q.

the insets of Fig. 9. By comparing the balanced eyes shown in Fig. 8 with perfectly aligned two pumps, one can observe the performance degradation with added noise under pump phase misalignment of 3 ps and 4 ps. Especially, one can observe almost completely closed eyes of demodulated signals after data exchange under an even larger time misalignment of 10 ps between the two pumps. Consequently, precise time alignment between two pumps and resultant perfect pump phase cancellation is important and highly desired to obtain favorable operation performance for phase-transparent optical data exchange.

The measured received power versus the input signal power at a BER of 10^{-9} is shown in Fig. 10. Less than 3.5-dB fluctuation of the received power is observed at a BER of 10^{-9} when varying the input signal power from -12.0 to 8.1 dBm. Thus, the dynamic range of the input signal power is estimated to be around 20 dB for high-base optical data exchange of 100-Gbit/s RZ-DQPSK signals based on nondegenerated FWM process.

We then propose and demonstrate a simple alternative method to perform high-base data exchange between multichannel DQPSK signals using bidirectional degenerate FWM in a single HNLF accompanied by optical filtering. The concept and operation principle of multichannel, high-base optical data exchange is illustrated in Fig. 11. Four-channel DQPSK signals (S1-S4) and a single CW pump are used. Degenerate FWM process is employed. Note that four-channel DQPSK signals (S1-S4) are symmetrical about the CW pump. For multi-channel data exchange, one would expect to see simultaneous data information swapping between S1 and S4, S2 and S3. Generally speaking, for data exchange operation with two signals present, it is impossible to separate the newly converted signals from the original signals by unidirectional degenerate FWM process, so it is difficult to realize optical data

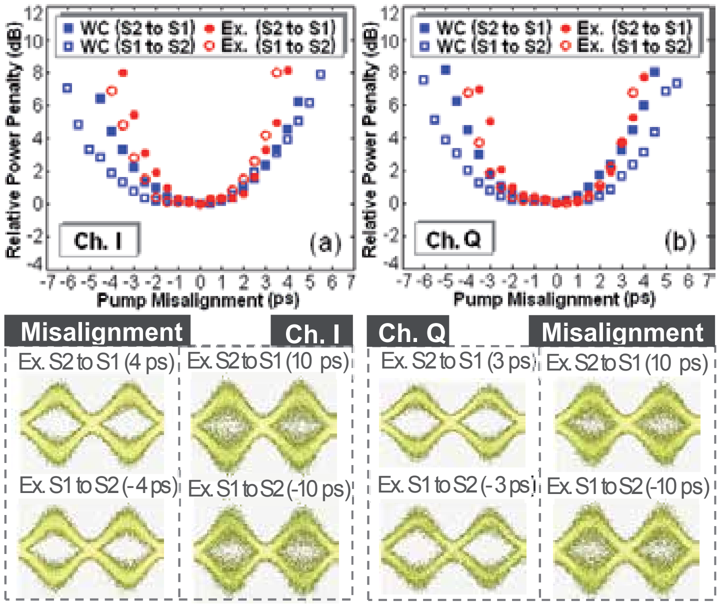


Figure 9. Impact of pump phase misalignment on the performance of high-base data exchange of 100-Gbit/s DQPSK signals. (a) Ch. I. (b) Ch. Q.

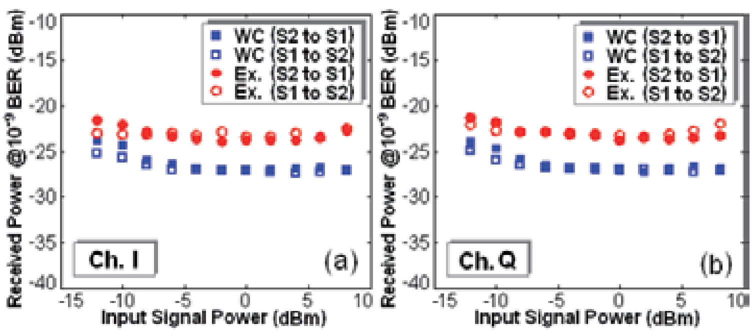


Figure 10. Dynamic range of input signal power for high-base data exchange of 100-Gbit/s DQPSK signals. (a) Ch. I. (b) Ch. Q.

exchange function based on unidirectional degenerate FWM in a single HNLF. We propose a possible solution by exploiting bidirectional degenerate FWM process in a single HNLF together with optical filtering. As illustrated in Fig. 11, taking four-channel optical data exchange as an example, there are four-channel DQPSK signals (S1-S4) at the input. 1) With optical filtering, S1 and S2 are selected and fed into the HNLF together with the CW pump from the left side. When propagating along the HNLF, S4 and S3 are generated by the degenerate FWM wavelength conversion process. After the generation of S4 and S3, the original S1, S2, and CW pump are suppressed, while the newly converted S4 and S3 are selected by optical filtering at the right side of HNLF. 2) At the same time, with optical filtering at the input, S3 and S4 are selected and sent into the HNLF together with CW pump from the right side. During the propagation through the HNLF, S2 and S1 are created by the degenerate FWM wavelength conversion process. After producing S2 and S1, the original S3, S4 and CW pump are removed, while the newly generated S2 and S1 are selected via optical filtering at the left side of HNLF. For the selected S4 and S3 (carrying data information of original S1 and S2) from the left side and selected S2 and S1 (carrying data information of original S3 and S4) from the right side of the HNLF, it is noted that data information carried by S1 and S4, S2 and S3 are swapped. As a result, by employing a single HNLF, exploiting bidirectional degenerate FWM process, and using optical filtering, simultaneous four-channel optical data exchange between S1 and S4 as well as S2 and S3 can be implemented. The combined S1-S4 from the left and right sides of the HNLF correspond to the output four-channel signals after optical data exchange. Remarkably, since the degenerate FWM process has distinct phase-conjugation property, for DQPSK signals the in-phase (Ch. I) and quadrature (Ch. Q) components are also swapped after the optical data exchange operation.

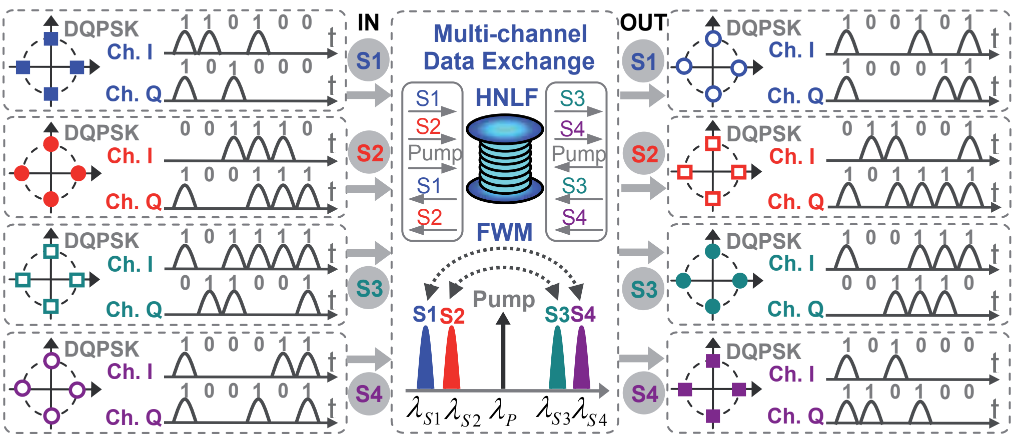


Figure 11. Concept and principle of simultaneous multichannel, high-base data exchange of DQPSK signals.

In the experiment, the bidirectional degenerate FWM in a single HNLF is enabled by a fiber loop mirror configuration, which consists of an HNLF with a length of 460 m, two optical band-pass filters, and optical fiber couplers. The typical parameters of the HNLF are as follows: ZDW: ~1556 nm; nonlinear coefficient: 20 $W^{-1}\cdot km^{-1}$; dispersion slope (S): ~0.026 $ps/nm^2/km$. Compared to the nondegenerate FWM-based data exchange with two pumps, single pump with its wavelength (1554.94 nm) close to the ZDW of HNLF is employed in the bidirectional degenerate FWM-based multichannel, high-base data exchange. ITU-grid-compatible four-channel 100-Gbit/s RZ-DQPSK signals (S1: 1546.12 nm, S2: 1547.72 nm, S3: 1562.23 nm, S4: 1563.86 nm) are employed for multichannel, high-base data exchange.

Shown in Fig. 12(a) is the measured spectrum of input four-channel, 100-Gbit/s RZ-DQPSK signals. S1(S2) and S4(S3) are symmetrical about the CW pump. The measured spectrum after four-channel optical data exchange with the CW pump ON is shown in Fig. 12(b) (solid blue line). For reference, the measured spectrum of residual signals with the CW pump OFF is also shown in Fig. 12(b) (dashed red line). It is expected that the residual signals are caused by the Rayleigh scattering in the HNLF. From Fig. 12(b), one can measure the extinction ratio of the newly exchanged signals to the residual signals to be 18.4 dB for S1, 19.5 dB for S2, 17 dB for S3, and 17 dB for S4, respectively.

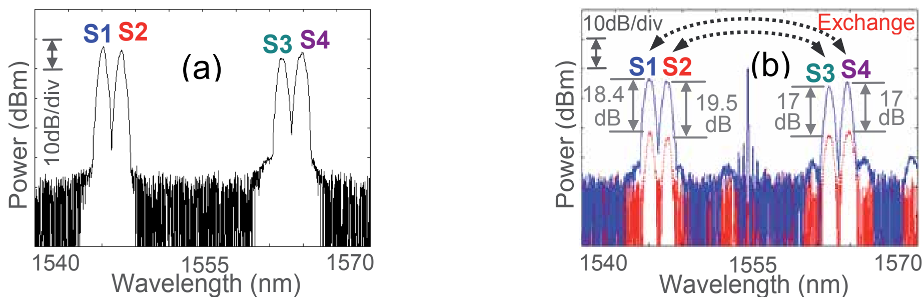


Figure 12. Spectra for four-channel, high-base data exchange of DQPSK signals. (a) Input four-channel, 100-Gbit/s RZ-DQPSK signals. (b) Spectra measured in the absence (dashed curve: Rayleigh scattering)/presence (solid curve: after data exchange) of CW pump.

Figure 13 further displays temporal waveforms and balanced eyes of demodulated in-phase (Ch. I) and quadrature (Ch. Q) components of 100-Gbit/s RZ-DQPSK signals before and after four-channel high-base optical data exchange. One can clearly confirm the successful implementation of simultaneous four-channel, 100-Gbit/s RZ-DQPSK optical data exchange between S1 and S4 as well as S2 and S3. Meanwhile, one can also see that for DQPSK signals, the Ch. I and Ch. Q components are swapped after optical data exchange, which is due to the optical phase-conjugation property of the degenerate FWM process.

Figure 14 plots the BER curves for four-channel, high-base data exchange of 100-Gbit/s RZ-DQPSK signals. Less than 4.7-dB power penalty is observed at a BER of 10^{-9}, which could be caused by the beating effect between the newly exchanged signals and the original residual signals.

By exploiting bidirectional degenerate FWM progress with a single pump in a single HNLF and employing liquid crystal on silicon (LCoS) technology in a double-pass configuration, we further propose a terabit-scale network grooming switch element, which can simultaneously perform multiple optical signal processing functions, e.g., high-base add/drop, high-base optical data exchange, and high-base power equalization. Using 23-channel, 100-Gbit/s RZ-DQPSK signals, we demonstrate reconfigurable 2.3-Tbit/s network grooming switch operation in the experiment. Remarkably, simultaneous implementation of all these high-base optical signal processing functions can potentially enhance the efficiency and flexibility of network management.

Shown in Fig. 15 is the concept and operation principle of the proposed high-base, multifunctional grooming switch element that could be used at the network nodes. When multiple

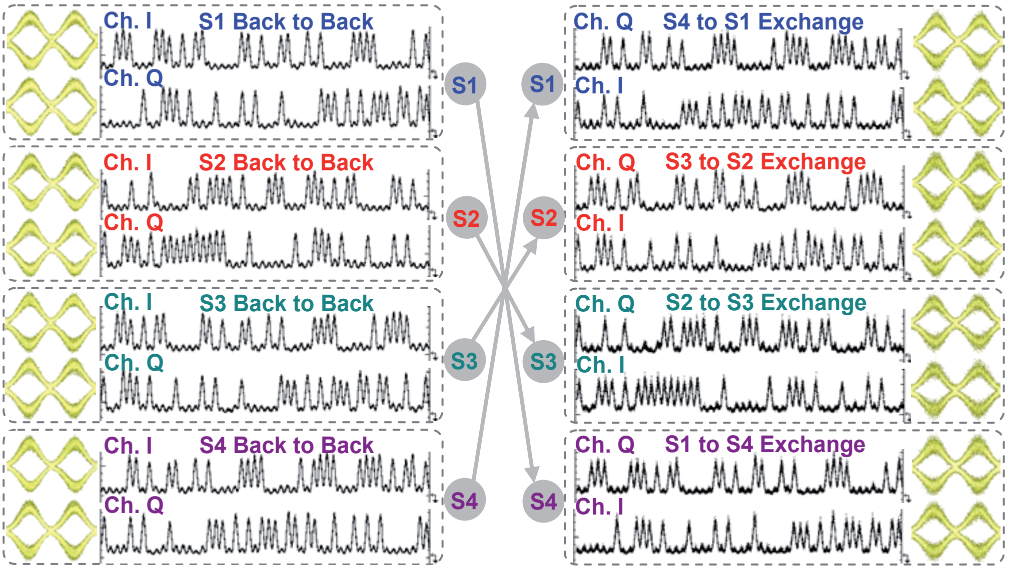


Figure 13. Waveforms and balanced eyes of demodulated in-phase (Ch. I) and quadrature (Ch. Q) components for four-channel, high-base data exchange of 100-Gbit/s DQPSK signals.

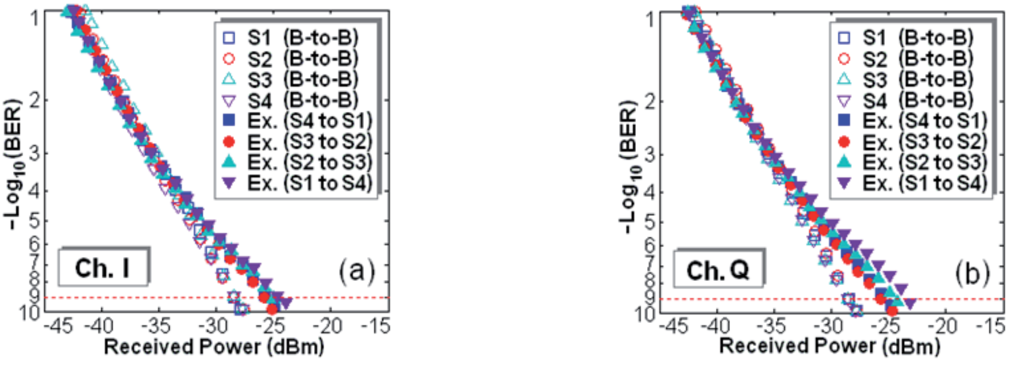


Figure 14. Measured BER curves for simultaneous four-channel, high-base data exchange of 100-Gbit/s DQPSK signals.

wavelength-division multiplexed (WDM) channels with unequalized power levels arrive at the network nodes, one would expect to flexibly manipulate these signals in the optical domain, in order to reduce the network latency and enhance the network efficiency. The typical favorable grooming optical signal processing functions are as follows: 1) optical data exchange between two or multiple channels of interest; 2) dropping of one or multiple channels of interest and adding of corresponding one or multiple channels with new data information; 3) power equalization for all the WDM channels. Moreover, it is also expected that these optical signal processing functions (optical data exchange, add/drop, power equalization) could be switchable, selective, and reconfigurable. For simplicity, shown in Fig. 15 is an example with

7-channel WDM signals. A wavelength selective switch (WSS) using a two-dimensional (2D) array of LCoS pixels is employed in the setup. The operation principle of the LCoS-based WSS is as follows. By changing the voltages loaded to the LCoS, one can adjust the phase retardance of each pixel of LCoS. The 2D LCoS array includes two axes with one horizontal wavelength axis and the other vertical displacement axis. The input 7-channel 100-Gbit/s DQPSK signals with unequalized power levels are sent to the port A of the input/output fiber array through a circulator. A diffraction grating collecting the input signals from port A then disperses different wavelength channels to different horizontal positions of the LCoS. Along the vertical direction, many pixels (~400 pixels) are covered due to the divergence of the light. The manipulation mechanism relies on the control of the LCoS. Since the phase shift of each pixel of LCoS can be adjusted by varying its applied voltage, it is possible to flexibly manipulate the phase front of the light through the control of the 2D array of LCoS pixels. By appropriately adjusting the independent pixel voltage, the propagation direction of different wavelength channels can be flexibly controlled, i.e., different wavelength channels can be delivered to different spatial positions at the output ports (e.g., S1 sent to port B, S4 and S5 sent to port C, S2 and S3 sent to port D, S6 and S7 sent to port E). Meanwhile, the power levels of different wavelength channels delivered to the desired fiber array ports (port B, port C, port D, port E) can be also adjusted. After separating and delivering different wavelength channels to different output fiber array ports together with flexible power control, various grooming optical signal processing functions can be carried out on these output fiber array ports: 1) high-base optical data exchange between port D and port E; 2) high-base wavelength add and drop at port B; 3) high-base power equalization of all wavelength channels. For the high-base optical data exchange between port D and port E, simultaneous multichannel, high-base optical data exchange between S2 and S7 and between S3 and S6 can be implemented by exploiting bidirectional degenerate FWM through a single HNLF. When compared to the similar optical data exchange scheme using degenerate FWM and employing optical band-pass filters to select desired wavelength channels, here the channel separation and selection are accomplished by LCoS. When compared to the optical data exchange approach using parametric depletion effect of nondegenerate FWM process with two pumps, here only single pump is employed in the setup. In particular, the simultaneous multichannel optical data exchange operation is switchable when employing the programmable LCoS. For the high-base wavelength add and drop, the S1 DQPSK signal is dropped at port B and a new S1 with updated data information is also added to port B through a circulator. For the high-base power equalization, the flexible attenuation control for all WDM channels is available by programming LCoS. Besides optical data exchange (S2 and S7, S3 and S6) and add/drop (S1) operations on the channels of interest, other channels (S4 and S5) without undergoing these operations should be kept and delivered back. A fiber loop structure could be employed at the port C. Remarkably, after multiple grooming optical signal processing operations, it is preferred that all the signals are sent back to the same input/output fiber array port A, which not only imports unequalized multiple WDM signals but also exports all the signals after the grooming switching. Such function can be implemented simply by running the LCoS device in a double-pass configuration assisted by use of some optical circulators. As shown in Fig. 15, if we consider the dashed boxes as a grooming switch unit based on HNLF and LCoS, it is actually a multifunctional, high-base grooming optical signal processing element with great reconfigurability. Simultaneous reconfigurable high-base add/drop, high-base optical data exchange, and high-base power equalization are implemented by exploiting bidirectional degenerate FWM in a single HNLF and double-pass programmable LCoS technology.

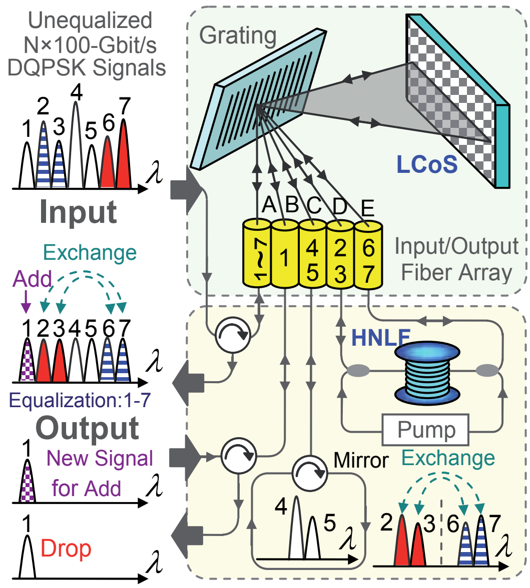


Figure 15. Concept and principle of LCoS+HNLF-based multifunctional, high-base grooming switch (add/drop, data exchange, power equalization).

Similar operation principle is adopted for reconfigurable 2.3-Tbit/s network grooming switch with 23x100-Gbit/s RZ-DQPSK channels. In the experiment, ITU-grid-compatible 23 wavelength channels (from S1: 1531.12 nm to S23: 1566.31 nm) each carrying 100-Gbit/s RZ-DQPSK modulation signal with a channel spacing of 200 GHz are utilized. A 520-m piece of HNLF with a ZDW of ~1555 nm and a nonlinear coefficient (γ) of 20 $W^{-1}\cdot km^{-1}$ is employed. The single pump wavelength is set to be 1555.75 nm for bidirectional degenerate FWM.

Figure 16 shows the measured optical spectrum and balanced eyes for input unequalized 23 wavelength channels each carrying a 100-Gbit/s RZ-DQPSK signal. The observed power fluctuation of all 23 wavelength channels is assessed to be around 9.1 dB. The insets of Fig. 16 depict measured typical balanced eyes for the demodulated in-phase (Ch. I) and quadrature (Ch. Q) components of 100-Gbit/s RZ-DQPSK signals.

We first perform 2.3-Tbit/s grooming switch with single-channel, high-base add/drop and two-channel high-base optical data exchange. The measured optical spectrum together with typical balanced eyes for 100-Gbit/s RZ-DQPSK signals after the multifunctional, high-base grooming switch is shown in Fig. 17. Three high-base grooming optical signal processing functions are implemented as follows: 1) high-base optical data exchange between S12 and S21; 2) high-base

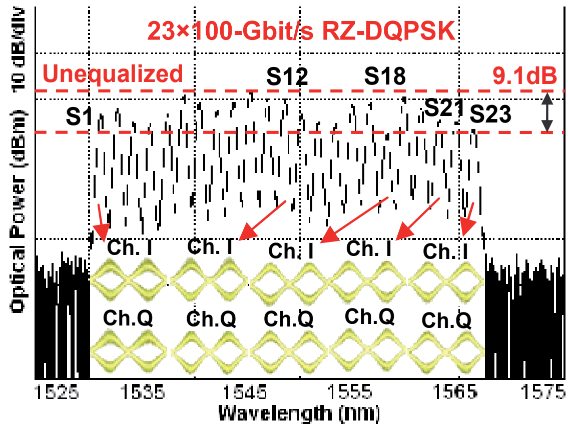


Figure 16. Measured optical spectrum and balanced eyes for input unequalized 23-channel 100-Gbit/s RZ-DQPSK signals.

dropping of the original S18 and high-base adding of new S18 with updated data information; 3) high-base power equalization for all 23-channel 100-Gbit/s RZ-DQPSK signals (power fluctuation: <1 dB). We also measure power penalties at a BER of 10^{-9} as shown in Fig. 18 for the multichannel, multifunctional grooming switch.

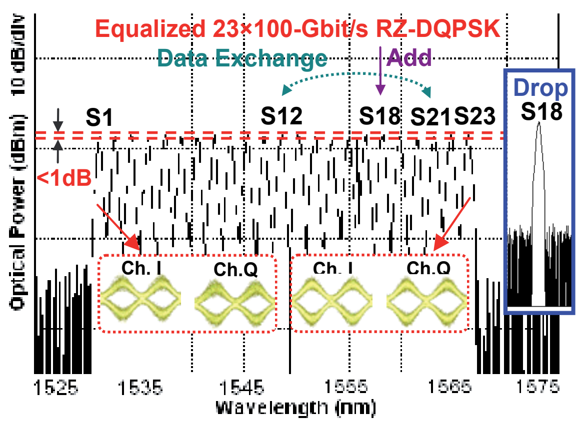


Figure 17. Measured optical spectrum and balanced eyes for 100-Gbit/s RZ-DQPSK signals after multifunctional, high-base grooming switch (high-base optical data exchange between S12 and S21; high-base add/drop for S18; high-base power equalization for all 23 wavelength channels S1-S23).

Due to the programmable LCoS employed in the configuration, the proposed multichannel, multifunctional grooming switch is reconfigurable. For instance, one can perform switchable simultaneous multichannel optical data exchange simply by changing the wavelength channels of interest sent to the fiber array port D and port E.

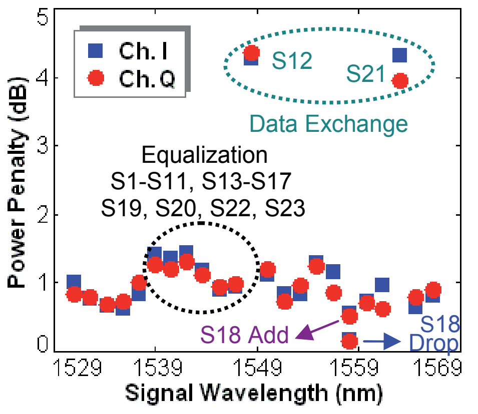


Figure 18. Measured power penalties at a BER of 10^{-9} for the multichannel, multifunctional high-base grooming switch (high-base optical data exchange between S12 and S21; high-base add/drop for S18; high-base power equalization for all 23 wavelength channels S1-S23).

We also demonstrate 2.3-Tbit/s grooming switch with two-channel add/drop and six-channel optical data exchange. Shown in Fig. 19 is the measured optical spectrum and typical balanced eyes for 100-Gbit/s RZ-DQPSK signals after the multifunctional, high-base grooming switch: 1) simultaneous six-channel, high-base optical data exchange between S10 and S23, S11 and S22, S12 and S21; 2) simultaneous two-channel, high-base dropping of the original S6 and S7 and high-base adding of new S6 and S7 with updated data information; 3) high-base power equalization with power fluctuation less than 1 dB for all 23 wavelength channels. Shown in the inset of Fig. 19 is the measured optical spectrum of dropped two wavelength channels of S6 and S7. Figure 20 plots the measured BER performance for simultaneous multichannel, high-base optical data exchange and high-base add/drop. The observed power penalties are assessed to be less than 1.2 dB for two-channel high-base add, 0.5 dB for two-channel high-base drop, and 5 dB for six-channel high-base optical data exchange at a BER of 10^{-9}.

In addition to high-base data exchange based on degenerate/nondegenerate FWM in HNLFs, we also propose and simulate ultrahigh-speed high-base data exchange using nondegenerate FWM in a silicon–organic hybrid slot waveguide. The working principle is also based on the parametric depletion effect of nondegenerate FWM as in an HNLF. The designed silicon–organic hybrid slot waveguide offers tight light confinement, enhanced nonlinearity, and negligible TPA and free-carrier absorption (FCA). Using nonlinear coupled-mode equations under the slowly varying envelope approximation and taking full consideration of group-velocity mismatching (GVM), group-velocity dispersion (GVD), TPA, FCA, and free-carrier dispersion (FCD), the proposed silicon–organic hybrid slot waveguide based high-base data exchange is simulated. In the following simulations, two 640 Gbaud 2^{13}-1 pseudorandom binary sequence (PRBS) 16-QAM/64-QAM signals (λ_{SA}: 1542 nm, λ_{SB}: 1544 nm) and two pumps (λ_{P1}: 1548 nm, λ_{P2}: 1550 nm) are sent into a 17-mm-long silicon–organic hybrid slot waveguide, in which 16-QAM/64-QAM data exchange is realized based on the nondegenerate FWM process. Note that the high-speed 640 Gbaud 16-QAM/64-QAM signal could be optical time-

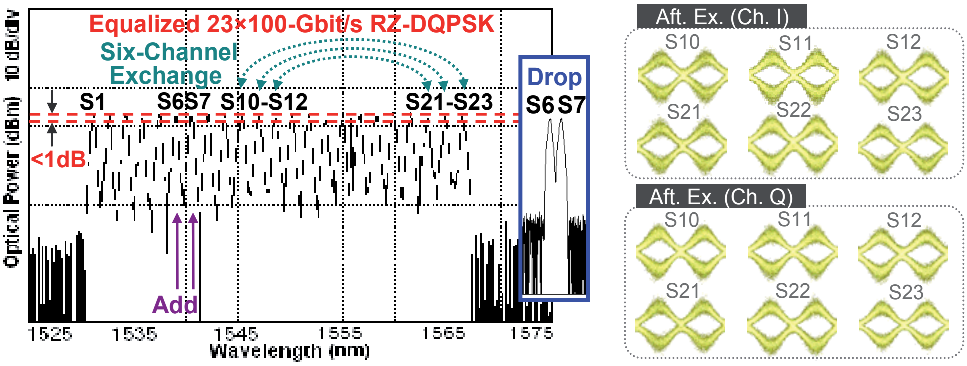


Figure 19. Measured optical spectrum and balanced eyes for 100-Gbit/s RZ-DQPSK signals after multichannel, multi-functional high-base grooming switch (simultaneous six-channel, high-base optical data exchange between S10 and S23, S11 and S22, S12 and S21; simultaneous two-channel, high-base add/drop for S6 and S7; high-base power equalization for all 23 wavelength channels S1-S23).

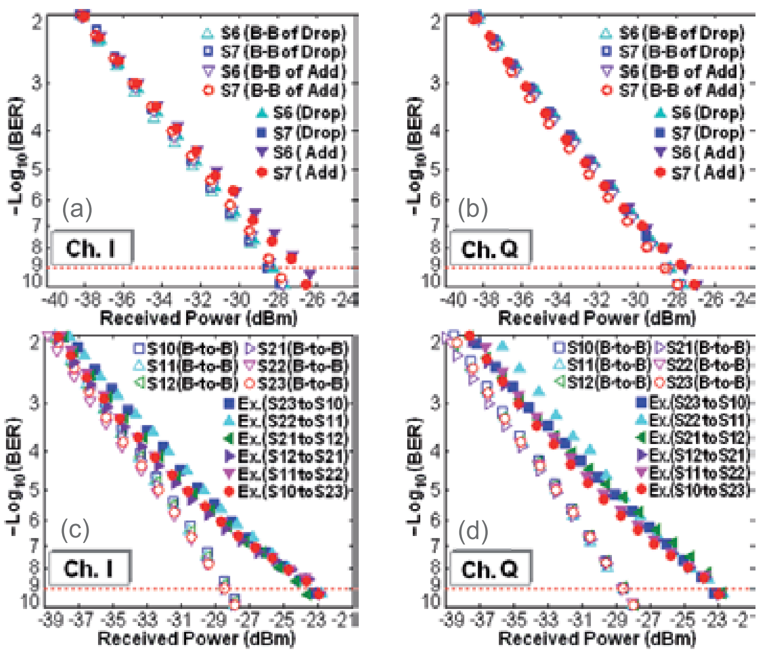


Figure 20. Measured BER performance for (a)(b) simultaneous two-channel, high-base add/drop (S6 and S7) and (c)(d) simultaneous six-channel, high-base optical data exchange between S10 and S23, S11 and S22, S12 and S21.

division multiplexed (OTDM) signal from 64 low-speed 10 Gbaud tributaries in practical applications.

The obtained results (symbol sequences) for high-base optical data exchange of 640 Gbaud (2.56 Tbit/s) 16-QAM signals are shown in Fig. 21. One can easily confirm the successful realization of the proposed high-base optical data exchange of 16-QAM signals by comparing the 10 symbol sequences for two signals (SA, SB) before optical data exchange (Bef. Ex.) and after optical data exchange (Aft. Ex.). Figure 22 shows simulated constellations for high-base optical data exchange of 16-QAM signals. For a signal-to-noise ratio (SNR) of 10 dB the error vector magnitude (EVM) is also assessed in Fig. 22. The simulated EVM and BER performance versus SNR for high-base optical data exchange of 640 Gbaud (2.56 Tbit/s) 16-QAM signals is shown in Fig. 23(a) and (b). For reference we also plot in Fig. 23(b) the theoretical 16-QAM BER curve. By comparing the simulated BER curves of two signals before and after optical data exchange, one can see negligible SNR penalty induced by the high-base optical data exchange operation at a BER of $2x10^{-3}$, which is the enhanced forward error correction (EFEC) threshold.

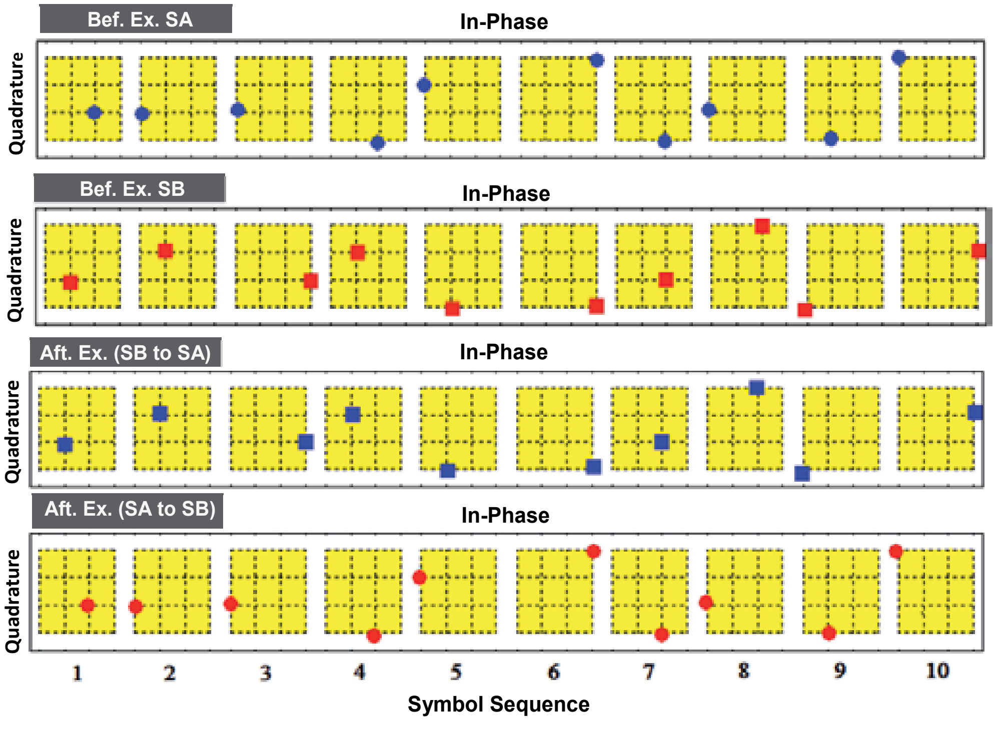


Figure 21. Simulated symbol sequences for high-base optical data exchange of 640 Gbaud (2.56 Tbit/s) 16-QAM signals.

We further simulate high-base optical data exchange of 640 Gbaud (3.84 Tbit/s) 64-QAM signals. The obtained results (symbol sequences) for high-base optical data exchange of 640 Gbaud (3.84 Tbit/s) 64-QAM signals are shown in Fig. 24. One can also confirm the successful implementation of the proposed high-base optical data exchange of 64-QAM signals by comparing the 10 symbol sequences for two signals (SA, SB) before optical data exchange (Bef. Ex.) and after optical data exchange (Aft. Ex.).

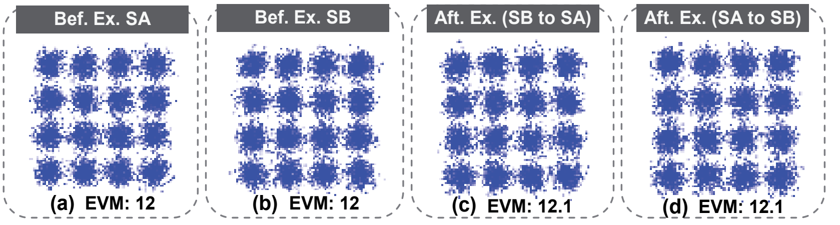


Figure 22. Simulated constellations of (a)(b) input and (c)(d) output signals for high-base optical data exchange of 640 Gbaud (2.56 Tbit/s) 16-QAM signals.

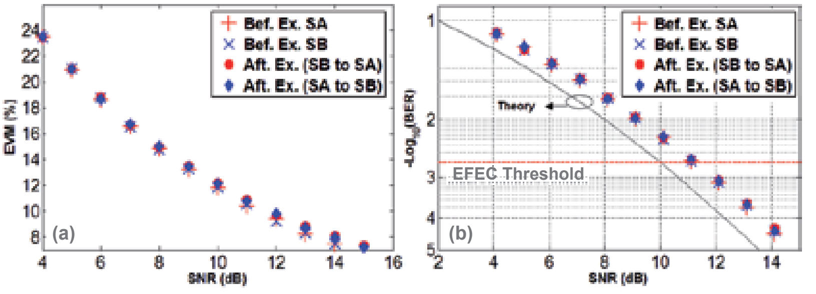


Figure 23. Simulated (a) EVM and (b) BER versus SNR for high-base optical data exchange of 640 Gbaud (2.56 Tbit/s) 16-QAM signals.

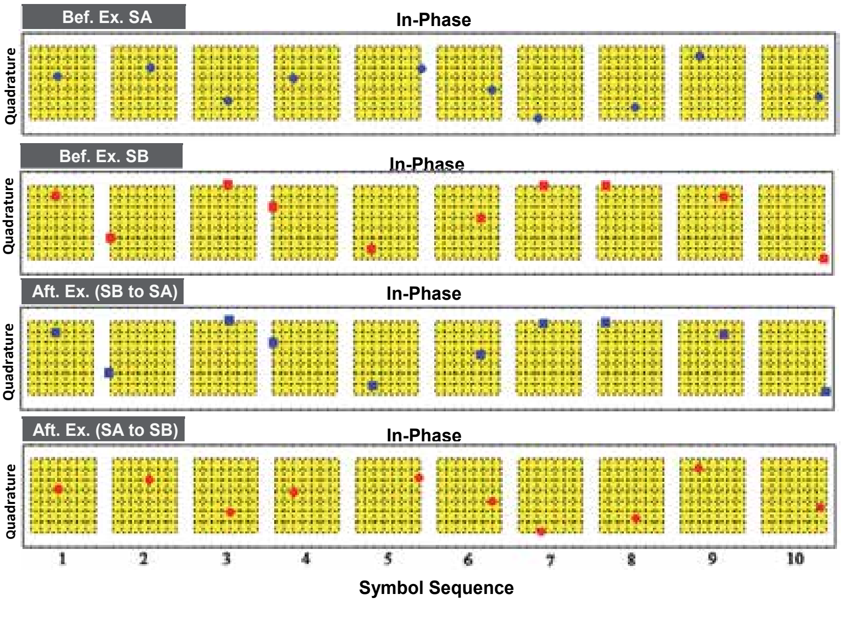


Figure 24. Simulated symbol sequences for high-base optical data exchange of 640 Gbaud (3.84 Tbit/s) 64-QAM signals.

Figure 25 shows simulated constellations for high-base optical data exchange of 64-QAM signals. For an SNR of 14 dB the EVM is also evaluated in Fig. 25. The simulated EVM and BER performance versus SNR for high-base optical data exchange of 640 Gbaud (2.56 Tbit/s) 64-QAM signals is shown in Fig. 26(a) and (b). For reference we also plot in Fig. 26(b) the theoretical 64-QAM BER curve. By comparing the simulated BER curves of two signals before and after optical data exchange, one can see that the SNR penalty induced by the high-base optical data exchange operation is assessed to be less than 2 dB at a BER of 2×10^{-3} which is the EFEC threshold.

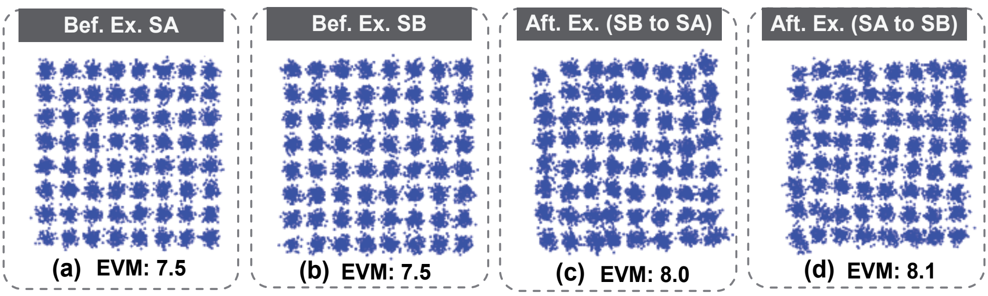


Figure 25. Simulated constellations of (a)(b) input and (c)(d) output signals for high-base optical data exchange of 640 Gbaud (3.84 Tbit/s) 64-QAM signals.

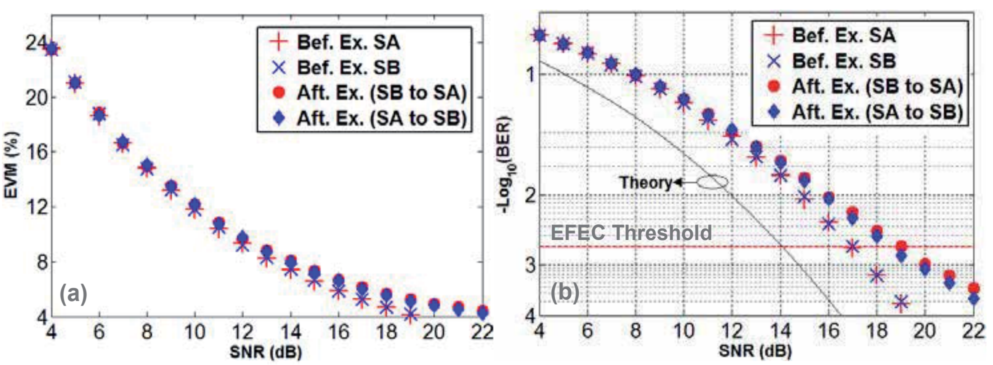


Figure 26. Simulated (a) EVM and (b) BER versus SNR for high-base optical data exchange of 640 Gbaud (3.84 Tbit/s) 64-QAM signals.

4. High-base optical computing [75, 77, 80, 85]

We propose and demonstrate high-base optical computing of advanced multilevel modulation signals based on degenerate/nondegenerate FWM in HNLFs or silicon–organic hybrid slot waveguides.

We first demonstrate high-speed two-input high-base optical computing (addition/subtraction/complement/doubling) of quaternary numbers using optical nonlinearities and DQPSK signals.

The concept and principle of operation of quaternary optical computing are shown in Fig. 27. As depicted in Fig. 27(a), DQPSK modulation signals have four-phase levels, i.e., 0, $\pi/2$, π, $3\pi/2$, which can be used to represent quaternary numbers, i.e., 0, 1, 2, 3. For two input signals A and B carrying quaternary numbers, it is expected that multiple outputs carrying different quaternary optical computing results could be achieved by employing a single nonlinear device. As depicted in Fig. 27(b), one can exploit three nondegenerate FWM processes and three degenerate FWM processes in a single HNLF with low and flat dispersion to implement simultaneous multiple quaternary optical computing functions. When launching signal A, signal B, and one CW pump into the HNLF, six converted idlers can be obtained with three idlers (idler 1-3) produced by three nondegenerate FWM processes and the other three idlers (idler 4-6) created by three degenerate FWM processes. For the six idlers generated by six FWM processes, one can derive the electrical field (E) and optical phase (Φ) relationships under the nondepletion approximation expressed as $E_{i1} \propto E_A\ E_B\ E_{CW}^*$, $\Phi_{i1}=\Phi_A+\Phi_B-\Phi_{CW}$ (1), $E_{i2} \propto E_A\ E_B^*\ E_{CW}$, $\Phi_{i2}=\Phi_A-\Phi_B+\Phi_{CW}$ (2), $E_{i3} \propto E_A^*\ E_B\ E_{CW}$, $\Phi_{i3}=\Phi_B-\Phi_A+\Phi_{CW}$ (3), $E_{i4} \propto E_{CW}\ E_{CW}\ E_A^*$, $\Phi_{i3}=2\Phi_{CW}-\Phi_A$ (4), $E_{i5} \propto E_{CW}\ E_{CW}\ E_B^*$, $\Phi_{i5}=2\Phi_{CW}-\Phi_B$ (5), $E_{i6} \propto E_B\ E_B\ E_{CW}^*$, $\Phi_{i6}=2\Phi_B-\Phi_{CW}$ (6). Remarkably, since optical phase has a periodicity of 2π due to its phase wrap characteristic, one can clearly see from Eqs. (1)-(6) that the six converted idlers actually take modulo 4 functions of quaternary optical computing, i.e., idler 1 for quaternary addition (A+B), idler 2 for quaternary subtraction (A-B), idler 3 for quaternary subtraction (B-A), idler 4 for quaternary complement (-A), idler 5 for quaternary complement (-B), and idler 6 for quaternary doubling (2B).

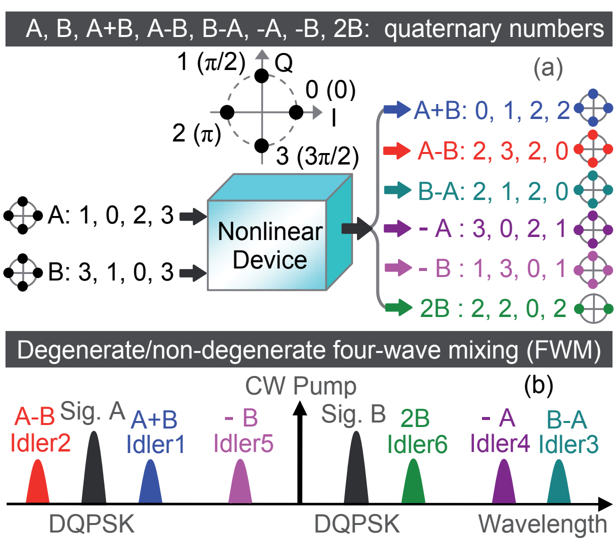


Figure 27. (a) Concept and (b) principle of two-input high-base optical computing (quaternary addition/subtraction/complement/doubling) using a single nonlinear device and DQPSK signals.

Shown in Fig. 28 are measured spectra. One CW pump (1553.2 nm) and two 100-Gbit/s 2^7-1 RZ-DQPSK signals (A: 1546.6 nm, B: 1555.5 nm) are fed into a 460-m-long HNLF. The ZDW, dispersion slope (S) and nonlinear coefficient (γ) of the HNLF are ~1556 nm, ~0.026 ps/nm^2/km, and 20 W^{-1} km^{-1}, respectively. The employed HNLF has low and flat dispersion, which benefits simultaneous multiple FWM processes. As a consequence, it is possible to simultaneously generate six idlers (idler 1: 1544.3 nm, idler 2: 1548.9 nm, idler 3: 1562.2 nm, idler 4: 1559.9 nm, idler 5: 1550.9 nm, idler 6: 1557.7 nm) corresponding to simultaneous addition (A+B), subtraction (A-B, B-A), complement (-A, -B), and doubling (2B) of quaternary numbers (A, B).

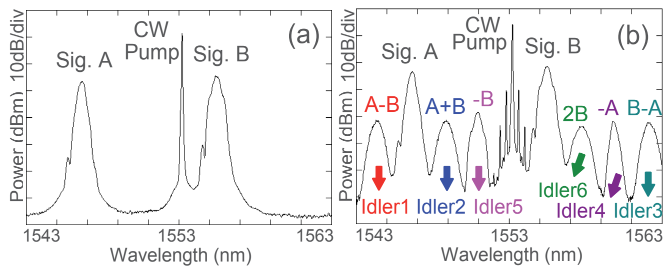


Figure 28. Measured spectra (a) before HNLF and (b) after HNLF for 50-Gbaud two-input quaternary optical computing (addition, subtraction, complement, doubling).

In order to confirm the quaternary optical computing (addition, subtraction, complement, doubling), the waveforms and balanced eyes of the demodulated in-phase (Ch. I) and quadrature (Ch. Q) components of two-input 100-Gbit/s RZ-DQPSK signals and six converted idlers by multiple FWM processes are recorded. A 50-GHz delay-line interferometer (DLI) is used to demodulate 100-Gbit/s RZ-DQPSK. A relative delay of 20 ps is introduced between the two arms of the 50-GHz DLI. Remarkably, quaternary numbers can be represented by the combination of Ch. I and Ch. Q (i.e., 00: ‘0’, 01: ‘1’, 11: ‘2’, 10: ‘3’). By exploiting multiple degenerate and nondegenerate FWM processes, one can clearly see from Figs. 29 and 30 that simultaneous 50-Gbaud quaternary optical computing of addition (A+B), dual-directional subtraction (A-B, B-A), complement (-A, -B), and doubling (2B) are successfully implemented with 100-Gbit/s DQPSK signals.

The BER performance of the quaternary optical computing is characterized as shown in Fig. 31. The measured power penalty at a BER of 10^{-9} is less than 4 dB for addition (A+B), 3 dB for subtraction (A-B, B-A), 2 dB for complement (-A, -B), and 3.1 dB for doubling (2B), respectively. Remarkably, one can see that the quaternary addition, subtraction, and doubling show relatively large power penalties compared to the quaternary complement. Such interesting phenomenon can be briefly explained as follows. According to the relationships of electrical fields, the distortions of input signals are transferred into converted idlers (i.e., computing results). Actually, the degradations of quaternary addition/subtraction, complement, and doubling are respectively induced by the accumulated distortions from signal A and signal B, distortion from single signal B, and twice distortions from signal B. Additionally, the BER curves of two-output signals from the HNLF are also plotted in Fig. 31(c) and (d) for reference.

One can clearly see that the two signals suffer negligible performance degradations during high-base arithmetical operations.

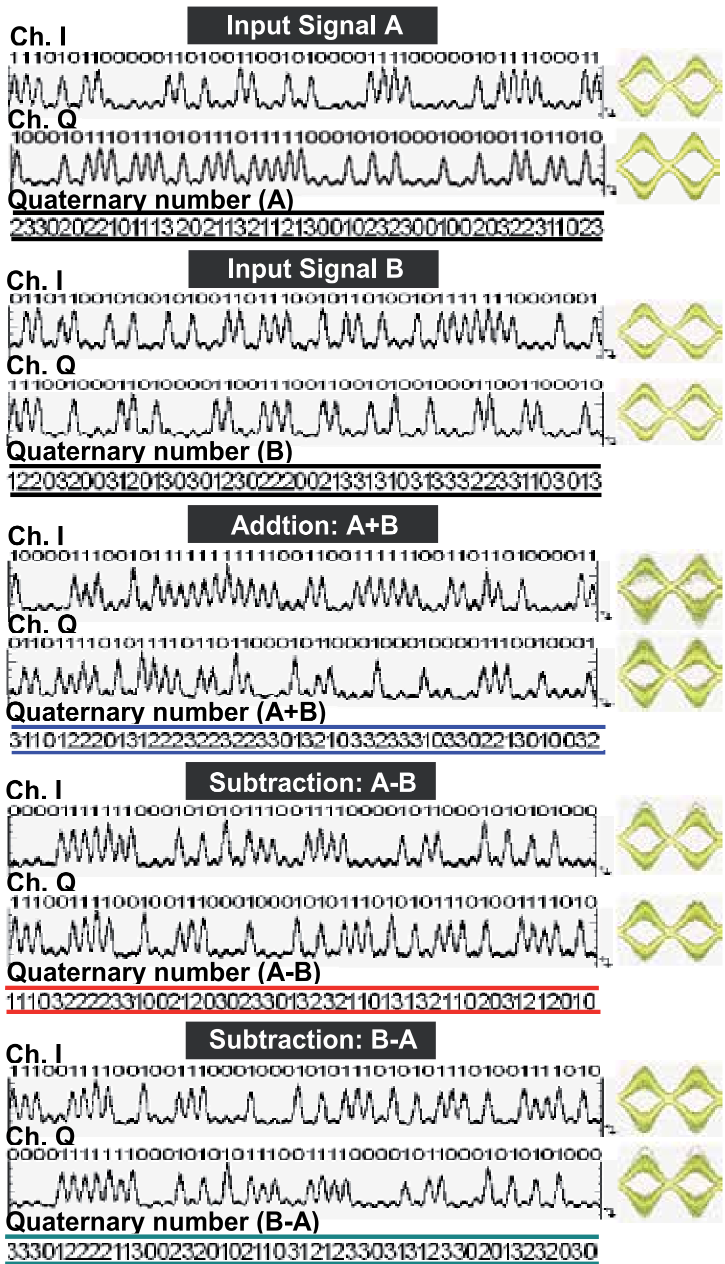


Figure 29. Demodulated waveforms and balanced eyes for 50-Gbaud two-input quaternary addition and dual-directional subtraction using 100-Gbit/s DQPSK signals.

Shown in Fig. 32 are measured constellations for input/output signals and output computing results. An optical complex spectrum analyzer (APEX AP2440A) is employed in the experiment. One can clearly see from Fig. 32 that the quaternary addition (A+B), quaternary subtraction (A-B, B-A), and quaternary complement (-A, -B) have four-phase levels (0, $\pi/2$, π, $3\pi/2$) while the quaternary doubling (2B) has two-phase levels (0, π).

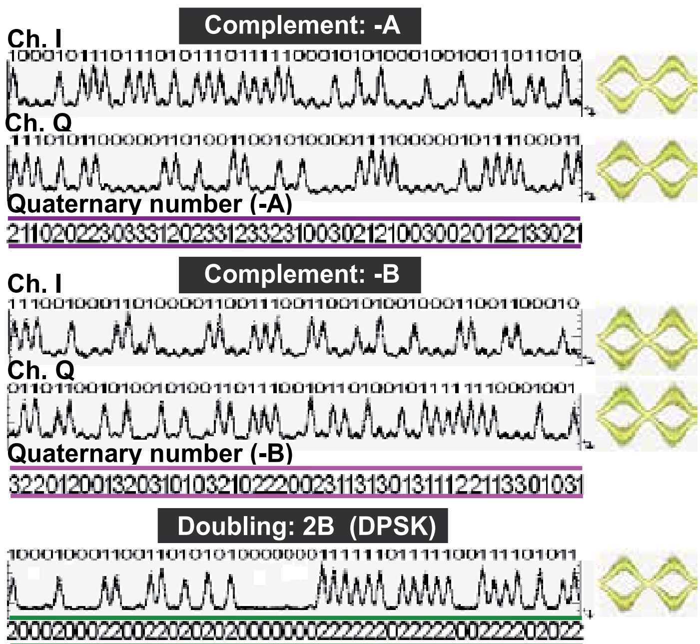


Figure 30. Demodulated waveforms and balanced eyes for 50-Gbaud quaternary complement and doubling using 100-Gbit/s DQPSK signals.

We then demonstrate high-speed three-input high-base optical computing (addition and subtraction) of quaternary numbers using multiple nondegenerate FWM processes in a single HNLF and DQPSK signals. Figure 33 illustrates the concept and operation principle.

Shown in Fig. 34 are measured spectra for 50-Gbaud three-input quaternary optical computing (addition, subtraction). Figure 34(a) depicts the spectrum for degenerate FWM, which enables the conversion from C to –C (i.e., quaternary complement). In the experiment, the wavelengths of CW pump, input signal C (Sig. C) and converted signal (–Sig. C) are 1552.0, 1548.7, and 1555.5 nm, respectively. Figure 34(b) shows the typical spectrum for three-input quaternary optical computing, i.e., quaternary hybrid addition and subtraction (A+B-C, A+C-B, B+C-A). In the experiment, the wavelengths of three input 100-Gbit/s RZ-DQPSK signals (A, B, C) are 1546.6 (Sig. A), 1553.2 (Sig. B), and 1555.5 nm (Sig. C), respectively. It is clearly shown that three converted idlers, i.e., idler 1 at 1544.3 nm, idler 2 at 1548.9 nm, and idler 3 at 1562.2 nm, are generated by three nondegenerate FWM processes. Actually, idler 1, idler 2, and idler 3 correspond to A+B-C, A+C-B, and B+C-A, respectively. Figure 34(c) displays the spectrum for

http://dx.doi.org/10.5772/61504

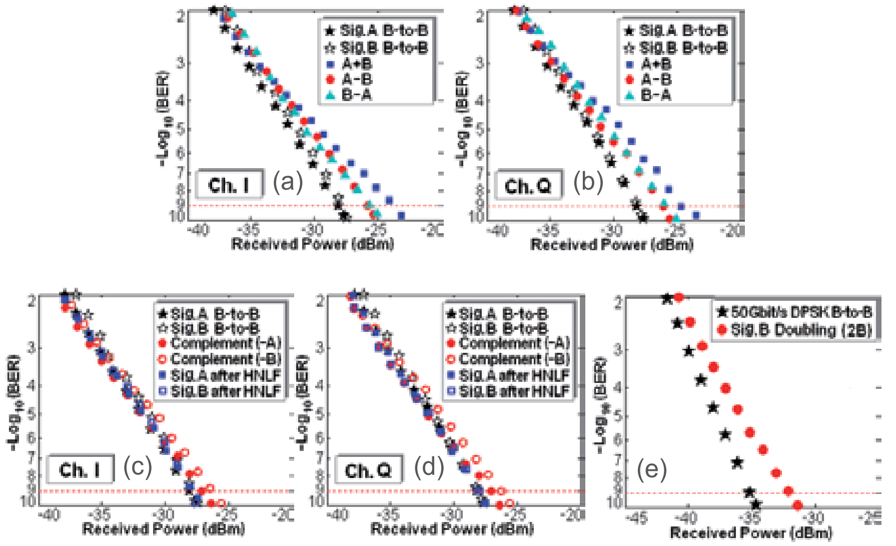


Figure 31. Measured BER curves for input/output signals (A, B), quaternary addition (A+B), dual-directional subtraction (A-B, B-A), complement (-A, -B), and doubling (2B).

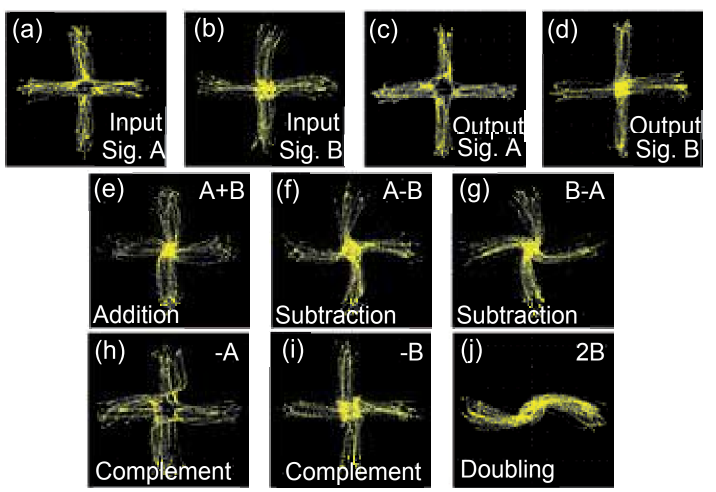


Figure 32. Measured constellations for 50-Gbaud two-input quaternary optical computing (addition, dual-directional subtraction, complement, doubling) with 100-Gbit/s DQPSK signals.

three-input quaternary addition of A+B+C. In the experiment, the converted signal (-Sig. C) by degenerate FWM shown in Fig. 34(a) is selected and used as the input signal shown in Fig. 34(b), i.e., -Sig. C is employed instead of Sig. C as shown in Fig. 34(c). After the nondegenerate FWM process, the converted idler 1 carrying quaternary addition result of A+B+C is obtained.

To verify the successful realization of three-input quaternary optical computing (addition, subtraction), the waveforms and balanced eye diagrams of the demodulated in-phase (Ch. I)

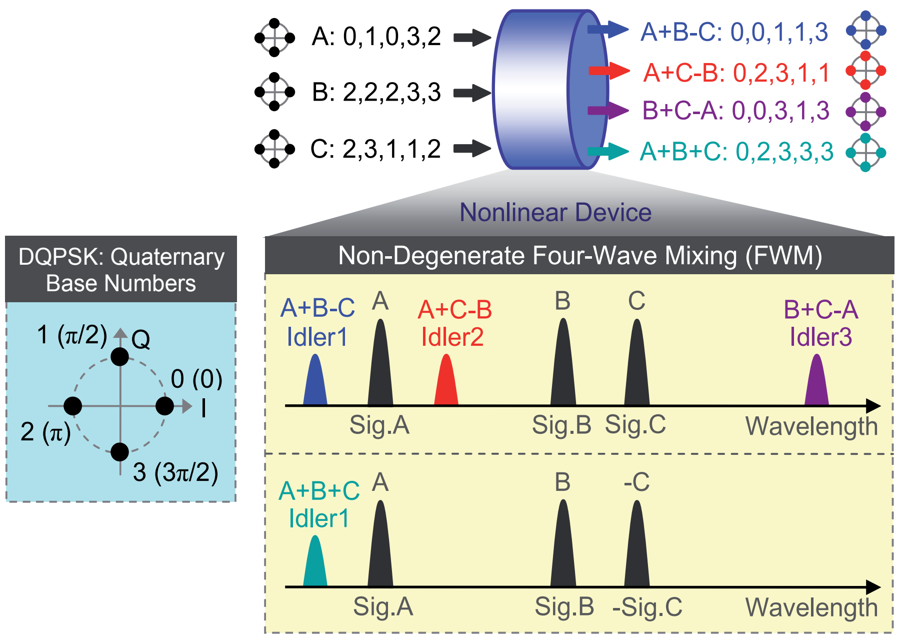


Figure 33. Concept and principle of three-input (A, B, C) optical quaternary addition and subtraction (A+B-C, A+C-B, B+C-A, A+B+C) using nondegenerate FWM and DQPSK signals.

and quadrature (Ch. Q) components of three-input 100-Gbit/s RZ-DQPSK signals and three-output converted idlers by nondegenerate FWM processes are recorded. Figure 35 depicts the measured sequences of input signals and converted idlers. It is clearly shown that the degenerate FWM process enables 50-Gbaud conversion from C to –C (i.e., quaternary complement) and three nondegenerate FWM processes perform three-input quaternary optical computing, i.e., hybrid quaternary addition and subtraction (A+B-C, A+C-B, B+C-A, A+B+C).

We measure the BER curves as shown in Fig. 36 for 50-Gbaud three-input quaternary optical computing (A+B-C, A+C-B, B+C-A). It is shown from Figs. 36(a) and (b) that the power penalties at a BER of 10^{-9} of three-input quaternary optical computing (A+B-C, A+C-B, B+C-A) are measured to be less than 6 dB. Shown in Fig. 37 are the measured BER curves for 50-Gbaud conversion from C to –C (i.e., quaternary complement) and 50-Gbaud three-input quaternary addition (A+B+C). The observed power penalty is negligible for the conversion from C to –C. For the quaternary addition of A+B+C, the power penalty at a BER of 10^{-9} is assessed to be less than 6 dB. Similar to two-input quaternary optical computing, it is believed that the performance degradations of three-input quaternary optical computing (i.e., quaternary hybrid addition and subtraction of A+B-C, A+C-B, B+C-A, and A+B+C) are mainly caused by accumulated distortions originated from three-input signals (A, B, C or –C). Such phenomenon can be explained according to the electrical field and linear optical phase relationships of nondegenerate FWM processes. Shown in Fig. 36(c)(d) and Fig. 37(a)(b) are measured BER curves for three output signals (A, B, C or –C) from HNLF after three-input quaternary optical

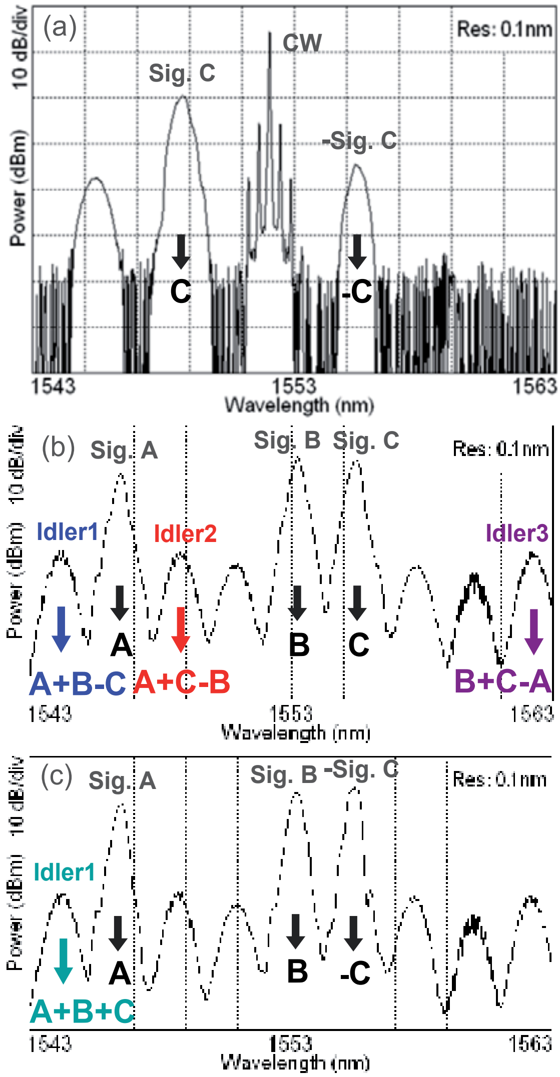


Figure 34. Measured spectra for 50-Gbaud three-input quaternary optical computing (addition, subtraction). (a) Degenerate FWM process (C to –C conversion); (b) three-input quaternary hybrid addition and subtraction (idler 1: A+B-C, idler 2: A+C-B, idler 3: B+C-A) by degenerate FWM process; (c) three-input quaternary addition (idler: A+B+C) by degenerate FWM process.

computing. For the three signals during the three-input quaternary optical computing operations, no significant performance degradations are observed in the experiment.

We also measure the constellation diagrams for three-input/output 100-Gbit/s RZ-DQPSK signals (A, B, C/-C) and six converted idlers corresponding to quaternary hybrid addition and subtraction of A+B-C, A+C-B, B+C-A, and A+B+C. An optical complex spectrum analyzer (APEX AP2440A) is employed in the experiment. From Fig. 38 one can clearly observe four-phase levels (i.e., 0, π/2, π, 3π/2) of all input/output signal and output idlers. These four-phase levels can represent quaternary base numbers.

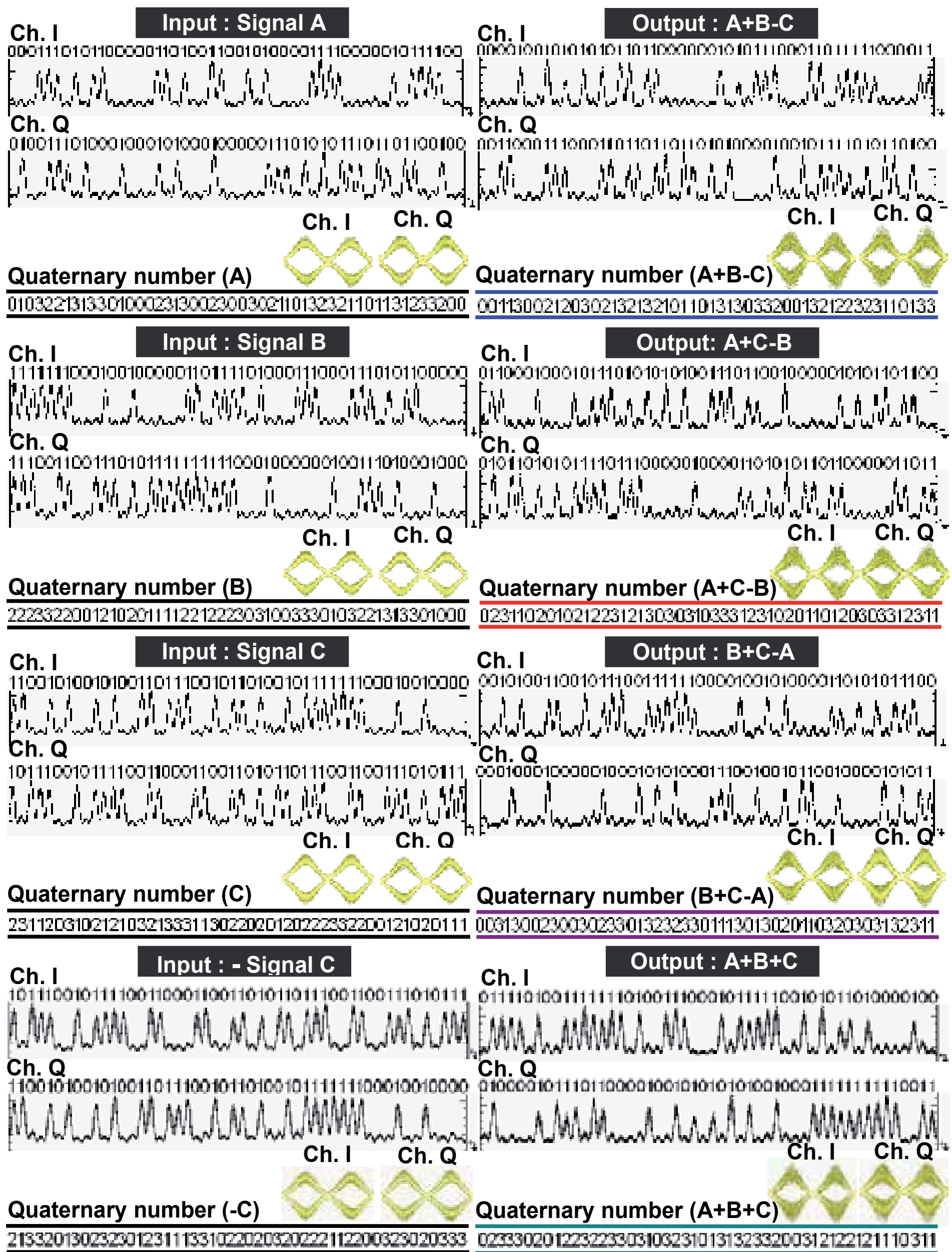


Figure 35. Demodulated temporal waveforms and balanced eye diagrams for 50-Gbaud three-input (A, B, C) quaternary addition and subtraction (A+B-C, A+C-B, B+C-A, A+B+C).

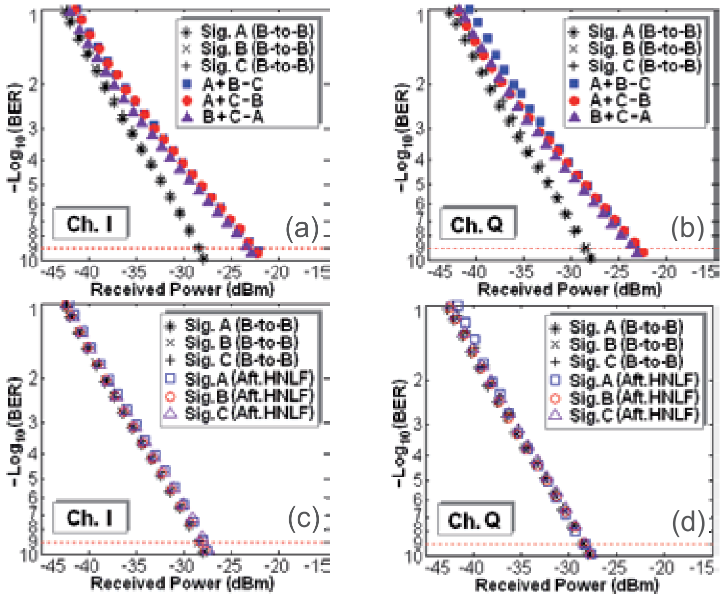


Figure 36. Measured BER curves for 50-Gbaud three-input quaternary optical computing. (a)(b) Hybrid addition and subtraction of A+B-C, A+C-B, and B+C-A. (c)(d) Output signals (Sig. A, Sig. B, Sig. C) from HNLF. (a)(c) Ch. I. (b)(d) Ch. Q. B-to-B: back-to-back.

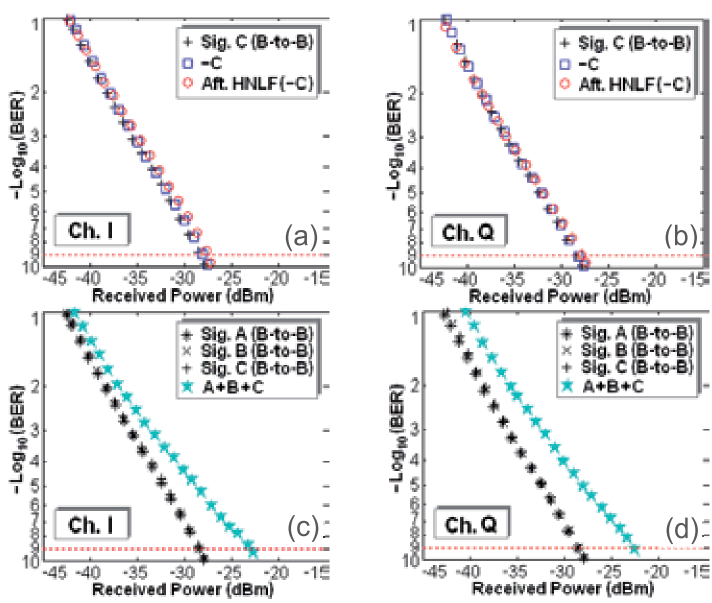


Figure 37. Measured BER curves for 50-Gbaud conversion from C to –C and three-input quaternary addition of A+B +C. (a)(b) Conversion from C to –C. (c)(d) A+B+C. (a)(c) Ch. I. (b)(d) Ch. Q. B-to-B: back-to-back.

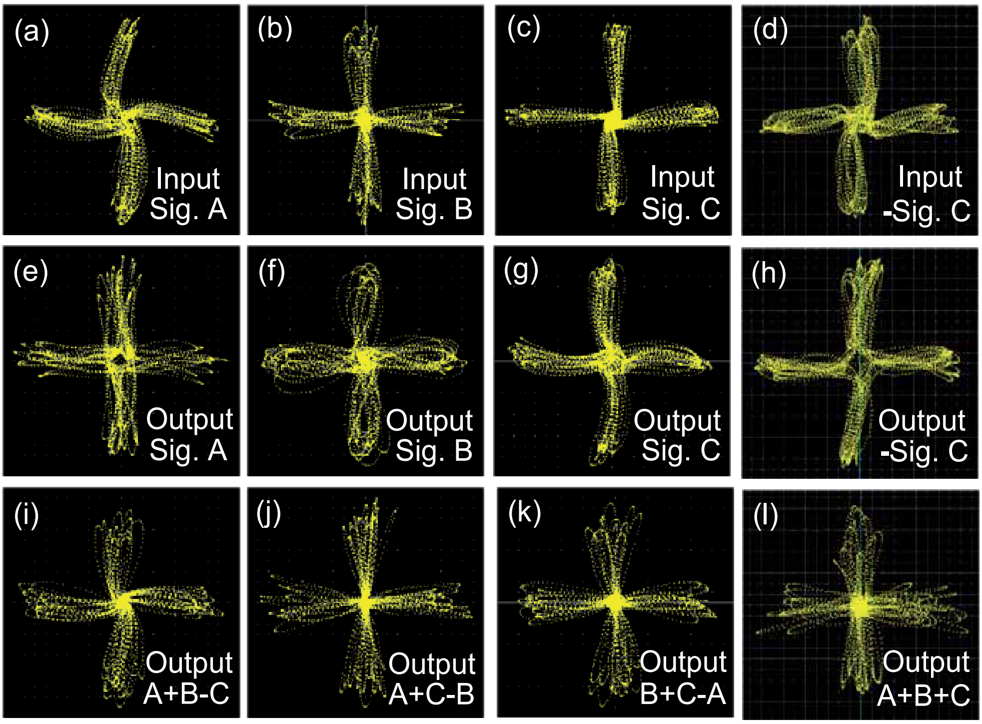


Figure 38. Measured constellation diagrams for 50-Gbaud three-input quaternary addition and subtraction.

In addition to two-/three-input high-base optical computing based on degenerate/nondegenerate FWM in HNLFs, we also propose and simulate three-input high-base optical computing (hexadecimal addition and subtraction) in a single silicon–organic hybrid slot waveguide based on nondegenerate FWM processes.

Shown in Fig. 39(a) is the schematic 3D structure of the proposed silicon–organic hybrid slot waveguide. It has a sandwich structure formed by a low-refractive-index PTS [polymer poly (bis para-toluene sulfonate) of 2, 4-hexadiyne-1,6 diol] layer inserted between two high-refractive-index silicon layers. The cladding of the structure is air. The substrate is silicon dioxide. In the designed silicon–organic hybrid slot waveguide, the waveguide width is W=250 nm, the upper silicon height is Hu=180 nm, the lower silicon height is Hl=180 nm, and the slot height is Hs=25 nm. We plot in Fig. 39(b)-(d) the quasi-TM mode distribution together with its normalized power density along x and y directions. It is clearly shown that the mode is highly confined in the nanoscale nonlinear organic slot region (i.e., tight light confinement). As a consequence, high nonlinearity and instantaneous Kerr response are achievable without impairments by TPA and FCA. Using finite-element method, we assess the effective mode area and nonlinearity to be $7.7\text{x}10^{-14}$ m^2 and 5500 $w^{-1}m^{-1}$, which can potentially facilitate efficient high-base optical signal processing (e.g., hexadecimal addition/subtraction). Figure 40 illustrates the operation principle which is similar to that in HNLFs. Instead of using DQPSK for quaternary optical computing, here 16PSK signals are used to achieve hexadecimal optical computing.

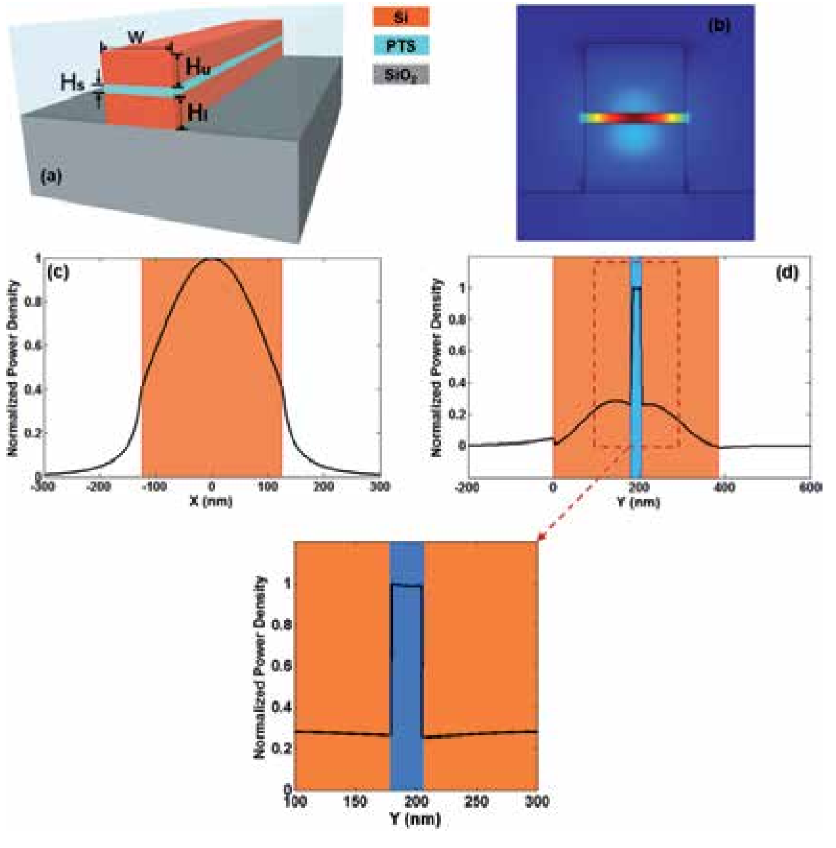


Figure 39. (a) 3D structure, (b) mode distribution, (c)(d) normalized power density along x and y directions of a silicon–organic hybrid slot waveguide.

In the following simulations, three 40-Gbaud 2^{13}-1 PRBS 16-PSK signals (λ_A: 1546 nm, λ_B: 1552 nm, λ_C: 1550 nm) are adopted. A 1-mm-long silicon–organic hybrid slot waveguide is employed. Figure 41 shows simulation results for three-input 40-Gbaud (160-Gbit/s) hexadecimal addition/subtraction. Twenty-symbol sequences are plotted in Fig. 41, which confirms the successful implementation of three-input hexadecimal addition/subtraction (A+B-C, A+C-B, B+C-A, A+B+C, A-B-C, B-A-C). The constellations are also shown in Fig. 42 with assessed EVM under an OSNR of 28 dB for input signals. The observed degradation of EVM for hexadecimal addition/subtraction can be ascribed to the accumulated noise from input 16-PSK signals and impairments from nonlinear interactions inside the silicon–organic hybrid slot waveguide. We further investigate the EVM of input signals and output idlers against the OSNR of input signals as shown in Fig. 43(a) and (b). The EVM penalties are assessed to be less than 4.5 for hexadecimal addition/subtraction under an OSNR of 28 dB.

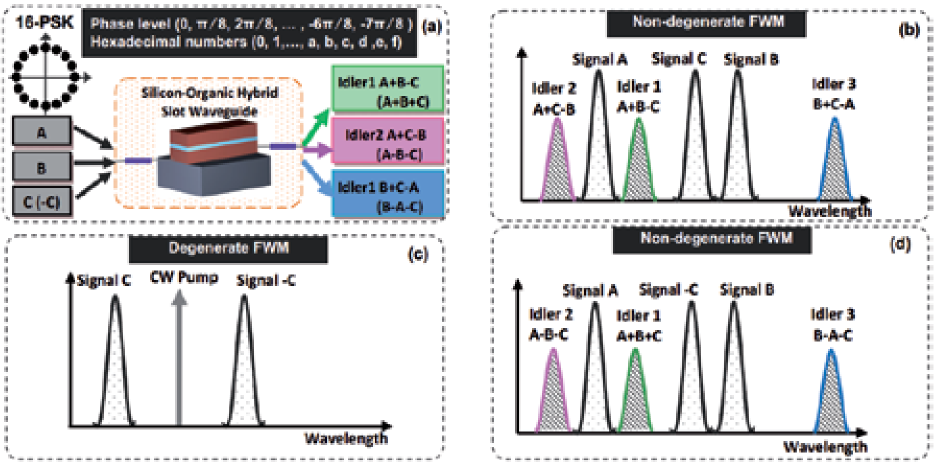


Figure 40. (a) Concept and (b)(c)(d) operation principle of three-input hexadecimal addition/subtraction.

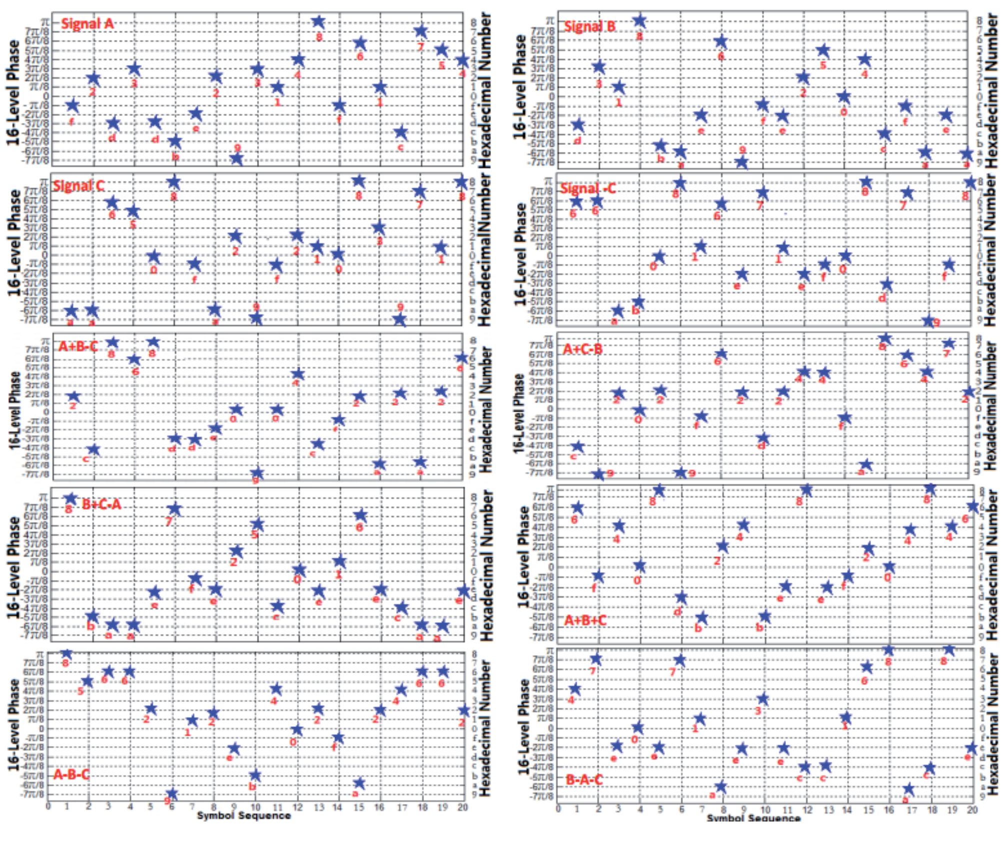


Figure 41. Simulated symbol sequence for three-input 40-Gbaud (160-Gbit/s) hexadecimal addition/subtraction using silicon–organic hybrid slot waveguide.

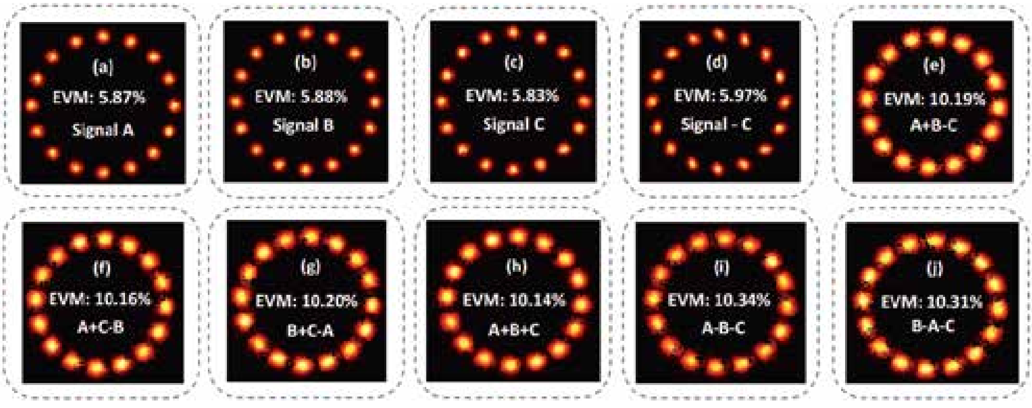


Figure 42. Simulated constellations for three-input 40-Gbaud (160-Gbit/s) hexadecimal addition/subtraction using silicon–organic hybrid slot waveguide.

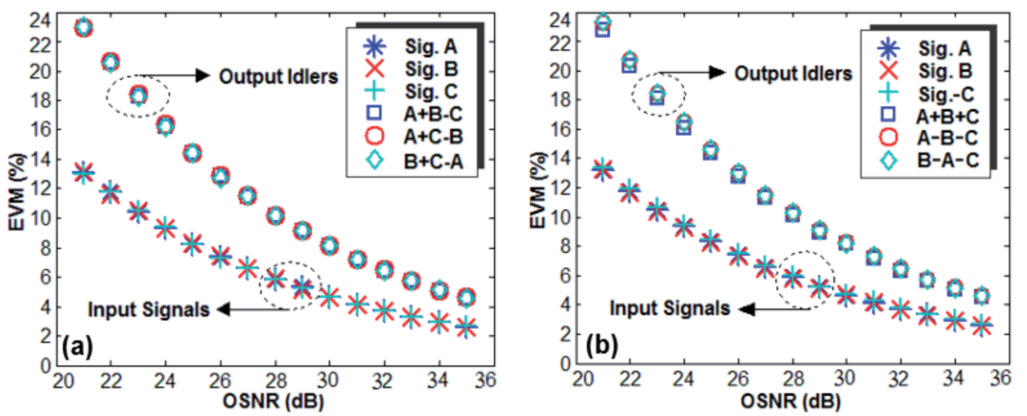


Figure 43. Simulated EVM vs. OSNR for 40-Gbaud (160-Gbit/s) hexadecimal addition/subtraction using silicon–organic hybrid slot waveguide.

5. High-base coding/decoding [79]

We propose and demonstrate high-base optical coding/decoding of advanced multilevel modulation signals based on degenerate FWM in HNLFs.

Figure 44 illustrates the concept and principle of the proposed symbol-wise hexadecimal coding/decoding using degenerate FWM and 16-QAM signals. Symbol-wise hexadecimal coding/decoding can be regarded as the constellation manipulation in the I/Q plane. The pump, original signal, coded signal and decoded signal are denoted by Ki, Pi, Ci, and Di, respectively. In the symbol-wise hexadecimal coding/decoding the pump can be CW or phase modulated. Illustrated in Fig. 44(a1) and (b1) are the symbol-wise hexadecimal coding, in which a CW or phase-modulated pump (Ki) and a 16-QAM signal (Pi) are launched into a nonlinear device such as HNLF to take part in the nonlinear interaction such as degenerate FWM process. 16-QAM signal can represent a hexadecimal number. When propagating along the HNLF, the degenerate FWM process generates a coded signal (Ci). Note that the electrical field ($E_{C\,i}$) of the coded signal (Ci) satisfies the relationship of $E_{C\,i} \propto E_{K\,i}^{2} \cdot E_{P\,i}^{*}$. From the electrical fields a linear phase relationship of $\Phi_{C\,i}=2\Phi_{K\,i}-\Phi_{P\,i}$ is achieved, i.e., twice the pump phase modulation ($2\Phi_{K\,i}$) and the conjugated phase of the original signal ($-\Phi_{P\,i}$) contribute together to the phase of the coded signal. Consequently, the coding algorithm simultaneously relies on the pump phase modulation and degenerate-FWM-induced phase conjugation. For the CW pump-assisted symbol-wise hexadecimal coding as shown in Fig. 44(a1), all constellation points in the I/Q plane are moved to their symmetrical positions with respect to the I-axis because of the phase conjugation property of degenerate FWM. Actually, hexadecimal code conversion from one number to another is achieved simply by conjugated degenerate FWM process. For the symbol-wise hexadecimal coding exploiting a phase-modulated pump, i.e. (0, π/4) phase modulation, as illustrated in Fig. 44(b1), all constellation points are mapped symmetrically with respect to the I-axis. Meanwhile, the pump phase modulation also introduces additional symbol-varying coding. When the constellation point of 16-QAM in one symbol meets the π/4 pump phase modulation, it will rotate in a counter-clockwise direction by π/2. As a result, the coding algorithm becomes $\Phi_{C\,i}=2\Phi_{K\,i}-\Phi_{P\,i}$ which determines the rule of hexadecimal coding.

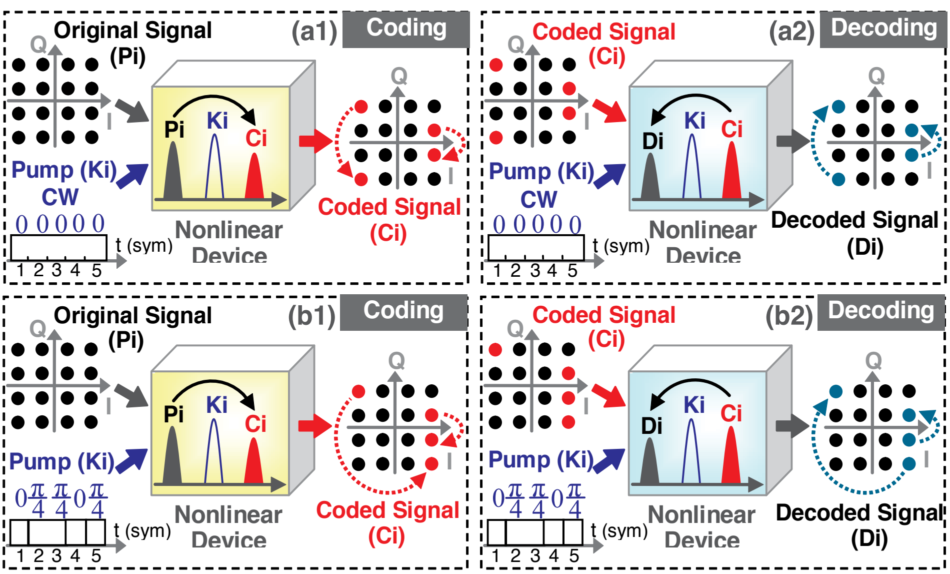


Figure 44. Concept and operation principle of variable symbol-wise hexadecimal coding/decoding by use of optical nonlinearity and 16-QAM. (a1)(a2) Symbol-wise hexadecimal coding/decoding assisted by CW pump; (b1)(b2) symbol-wise hexadecimal coding/decoding assisted by (0, π/4) phase-modulated pump; (a1)(b1) Coding; (a2)(b2) Decoding.

Figure 44(a2) and (b2) illustrate the symbol-wise hexadecimal decoding. The pump (Ki) and the coded signal (Ci) are fed into another nonlinear device such as HNLF to participate in the nonlinear interaction such as degenerate FWM process which generates the decoded signal (Di). It is noted that the electrical field of the decoded signal (Di) satisfies the relationship of $E_{D\,i} \propto E_{K\,i}^{2} \cdot E_{C\,i}^{*} \propto E_{K\,i}^{2} \cdot E_{K\,i}^{*2} \cdot E_{P\,i} = |E_{K\,i}|^{4} \cdot E_{P\,i} \propto E_{P\,i}$. Thus, the phase of the decoded signal (Di) meets the relationship of $\Phi_{D\,i} = 2\Phi_{K\,i} - \Phi_{C\,i} = 2\Phi_{K\,i} - (2\Phi_{K\,i} - \Phi_{P\,i}) = \Phi_{P\,i}$. As a consequence, the decoded signal (Di) recovers the original signal (Pi) after the decoding process. The decoding algorithm is determined by $\Phi_{D\,i} = 2\Phi_{K\,i} - \Phi_{C\,i} = 2\Phi_{K\,i} - (2\Phi_{K\,i} - \Phi_{P\,i}) = \Phi_{P\,i}$. Remarkably, the decoding algorithm corresponds to the constellation manipulation in the complex plane. The concept and principle shown in Fig. 44 indicate that the constellation of a 16-QAM signal can be manipulated by employing optical nonlinearity, which enables the symbol-wise hexadecimal coding/decoding. Moreover, exploiting a CW or (0, π/4) phase-modulated pump can facilitate optical variable symbol-wise hexadecimal coding/decoding assisted by optical nonlinearity.

Shown in Fig. 45 is the experimental setup for the proposed optical symbol-wise hexadecimal coding/decoding. A 10-Gbaud (40-Gbit/s) 16-QAM signal is prepared via the vector addition of two copies of QPSK signal using an I/Q QPSK modulator, polarization controllers (PCs), a tunable differential group delay (DGD) element, and a polarizer (Pol.). A 10-Gbit/s phase-modulated pump with (0, π/4) binary phase modulation, which is synchronized with the 10-Gbaud 16-QAM signal, is provided by employing a phase modulator (PM) driven by PRBS patterns. Note that the PM is not utilized for the CW pump-assisted hexadecimal coding/decoding. For the hexadecimal coding process, the 16-QAM signal (Pi) and the CW/phase-modulated pump (Ki) are launched into a 460-m piece of HNLF. The ZDW, dispersion slope (S) and nonlinear coefficient (γ) of the HNLF employed in the experiment are ~1556 nm, ~0.026 ps/nm^2/km, and 20 W^{-1}·km^{-1}, respectively. When the 16-QAM signal (Pi) and the CW/phase-modulated pump (Ki) propagate along the HNLF, a coded signal (Ci) is generated by degenerate FWM process. The coded signal (Ci) takes the result of hexadecimal coding. For the hexadecimal decoding process, the coded signal (Ci) and the CW/phase-modulated pump (Ki) are fed into another 520-m piece of HNLF which has a ZDW of ~1555 nm, S of ~0.026 ps/nm^2/km, and γ of 20 W^{-1}·km^{-1}. When the coded signal (Ci) and the CW/phase-modulated pump transmit through the HNLF, a decoded signal (Di) is obtained by degenerate FWM process. The decoded signal (Di) recovers the original signal corresponding to hexadecimal decoding. In the experimental setup, BPFs at the output of HNLFs are employed to suppress unwanted frequency components and pick up coded/decoded signals. For coherent detection of 16-QAM signals, an optical modulation analyzer (Agilent N4391A) and a digital phosphor oscilloscope (Tektronix DPO72004) with a 50-Gs/s sample rate and a 20-GHz electrical bandwidth are employed in the experiment.

The measured spectra for optical variable symbol-wise hexadecimal coding/decoding are shown in Fig. 46. Both, CW pump and (0, π/4) phase-modulated pump are employed in the experiment. The original signal (Pi), pump (Ki), coded signal (Ci), and decoded signal (Di) have wavelengths of 1557.0, 1555.6, 1554.2, and 1557.0 nm, respectively. We set the power of the original signal for coding and the coded signal for decoding to be around 10.8 dBm. For

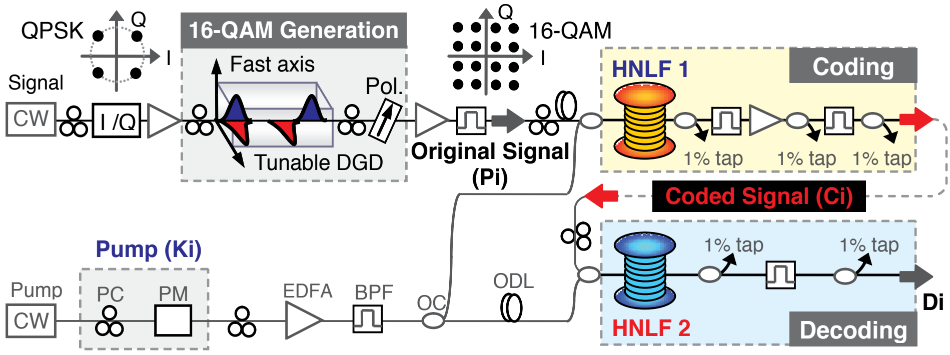


Figure 45. Experimental setup for high-base coding/decoding. Degenerate FWM in HNLF, 16-QAM signal and CW/phase-modulated pumps are employed to enable symbol-wise hexadecimal coding/decoding. QPSK: quadrature phase-shift keying; QAM: quadrature amplitude modulation; HNLF: highly nonlinear fiber; CW: continuous-wave; PC: polarization controller; EDFA: erbium-doped fiber amplifier; DGD: differential group delay; Pol.: polarizer; BPF: band-pass filter; ODL: tunable optical delay line; PM: phase modulator; OC: optical coupler.

the symbol-wise hexadecimal coding/decoding using a CW pump, the power of CW pump is ~12.8 dBm. The conversion efficiency is assessed to be about -15.4 dB for the symbol-wise hexadecimal coding while -14.9 dB for the symbol-wise hexadecimal decoding. For the symbol-wise hexadecimal coding/decoding using a (0, π/4) phase-modulated pump, the power of the (0, π/4) phase-modulated pump is ~9.8 dBm. The symbol-wise hexadecimal coding has a conversion efficiency of about -20.9 dB, while the symbol-wise hexadecimal decoding shows a conversion efficiency of around -19.1 dB.

Figure 47 depicts observed constellation diagrams and in-phase (I) and quadrature (Q) components for optical variable symbol-wise hexadecimal coding/decoding. Figure 47(a) shows the 10-Gbaud 16-QAM signal corresponding to the back-to-back (B-B) case. The EVM is measured to be 5.5%rms. The 16 constellation points can be clearly seen in the complex I/Q plane. Note that hexadecimal numbers can be represented by these 16 constellation points. For the symbol-wise hexadecimal coding/decoding using a CW pump, the phase-conjugated degenerate FWM process determines the coding and decoding algorithms to be ($\Phi_{C\,i}=-\Phi_{P\,i}$) and ($-(-\Phi_{P\,i})=\Phi_{P\,i}$), respectively. The constellations in the complex I/Q plane are manipulated following the coding and decoding algorithms. Figure 47(b) and (c) show the constellation diagrams of coded signal with an EVM of 6.3%rms and decoded signal with an EVM of 6.4%rms, respectively. For the symbol-wise hexadecimal coding/decoding using a phase-modulated pump, a (0, π/4) pump phase modulation with an EVM of 5.0%rms is employed in the experiment, as shown in Fig. 47(d). The constellation diagrams of the coded signal with an EVM of 7.8%rms and decoded signal with an EVM of 6.4%rms are shown in Fig. 47(e) and (f). The constellation manipulation in the complex I/Q plane follows the coding algorithm ($\Phi_{C\,i}=2\Phi_{K\,i}-\Phi_{P\,i}$) for the symbol-wise hexadecimal coding process and decoding algorithm ($2\Phi_{K\,i}-\Phi_{C\,i}=2\Phi_{K\,i}-(2\Phi_{K\,i}-\Phi_{P\,i})=\Phi_{P\,i}$) for the symbol-wise hexadecimal decoding process. Remarkably, for phase-modulated pump-assisted symbol-wise hexadecimal coding/decoding, the pump phase modulation and phase conjugation of degenerate FWM contribute together to the coding and decoding algorithms.

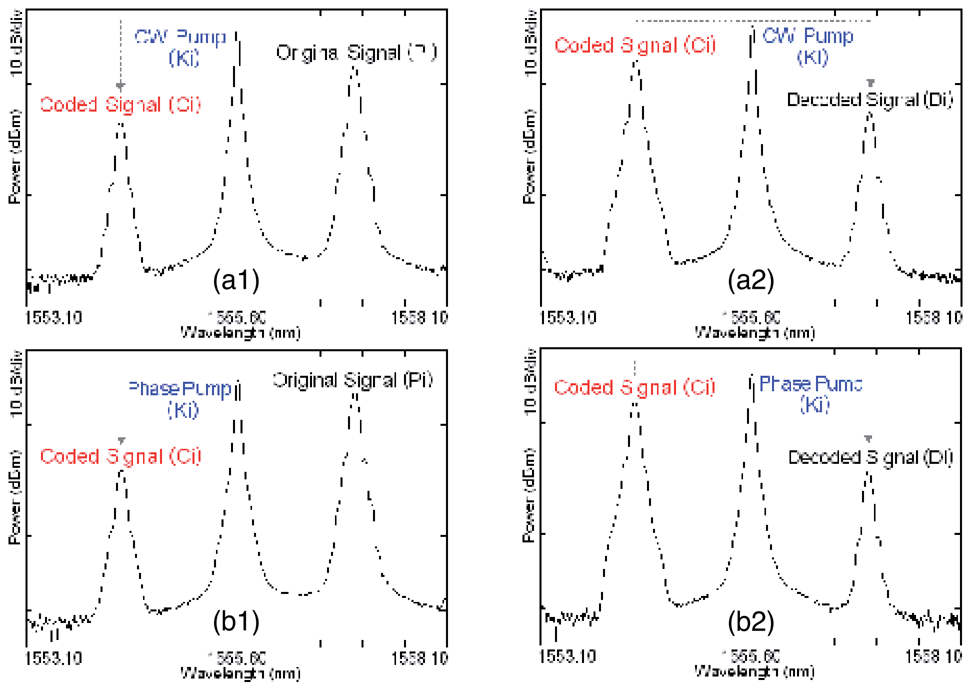


Figure 46. Measured spectra for high-base coding/decoding. Degenerate FWM in HNLF, 16-QAM signal and CW/phase-modulated pumps are employed to enable symbol-wise hexadecimal coding/decoding. (a1)(a2) Symbol-wise hexadecimal coding/decoding using a CW pump; (b1)(b2) symbol-wise hexadecimal coding/decoding using a (0, π/4) phase-modulated pump; (a1)(b1) symbol-wise hexadecimal coding; (a2)(b2) symbol-wise hexadecimal decoding.

To confirm the implementation of optical variable symbol-wise hexadecimal coding/decoding, the complex amplitudes (i.e., in-phase and quadrature components) of symbol sequence for different signals are recorded in the experiment. As shown in Fig. 48, for symbol-wise hexadecimal coding/decoding using a CW pump, by comparing the symbol sequence of coded signal and original signal, one can clearly see that all the constellation points in the complex I/Q plane are mapped to their symmetrical positions with respect to the I-axis. This constellation manipulation is determined by the coding algorithm of CW pump-assisted hexadecimal coding. Additionally, by comparing the symbol sequence of decoded signal and original signal one can confirm that the decoded signal recovers the original signal.

As shown in Fig. 49, for symbol-wise hexadecimal coding/decoding using a (0, π/4) phase-modulated pump, the corresponding coding algorithm manipulates the constellation points in the complex I/Q plane as follows. All the constellation points in the complex I/Q plane are first flipped to their symmetrical points with respect to the I-axis. Then, a counter-clockwise rotation of π/2 is introduced to the constellation points, which meet the pump phase modulation of π/4. One can expect enhanced security for the symbol-wise hexadecimal coding using a phase-modulated pump owing to the added coding algorithm contribution from the pump. When compared to the symbol-wise hexadecimal coding using a CW pump, the phase-modulated pump-assisted symbol-wise hexadecimal coding is not so straightforward.

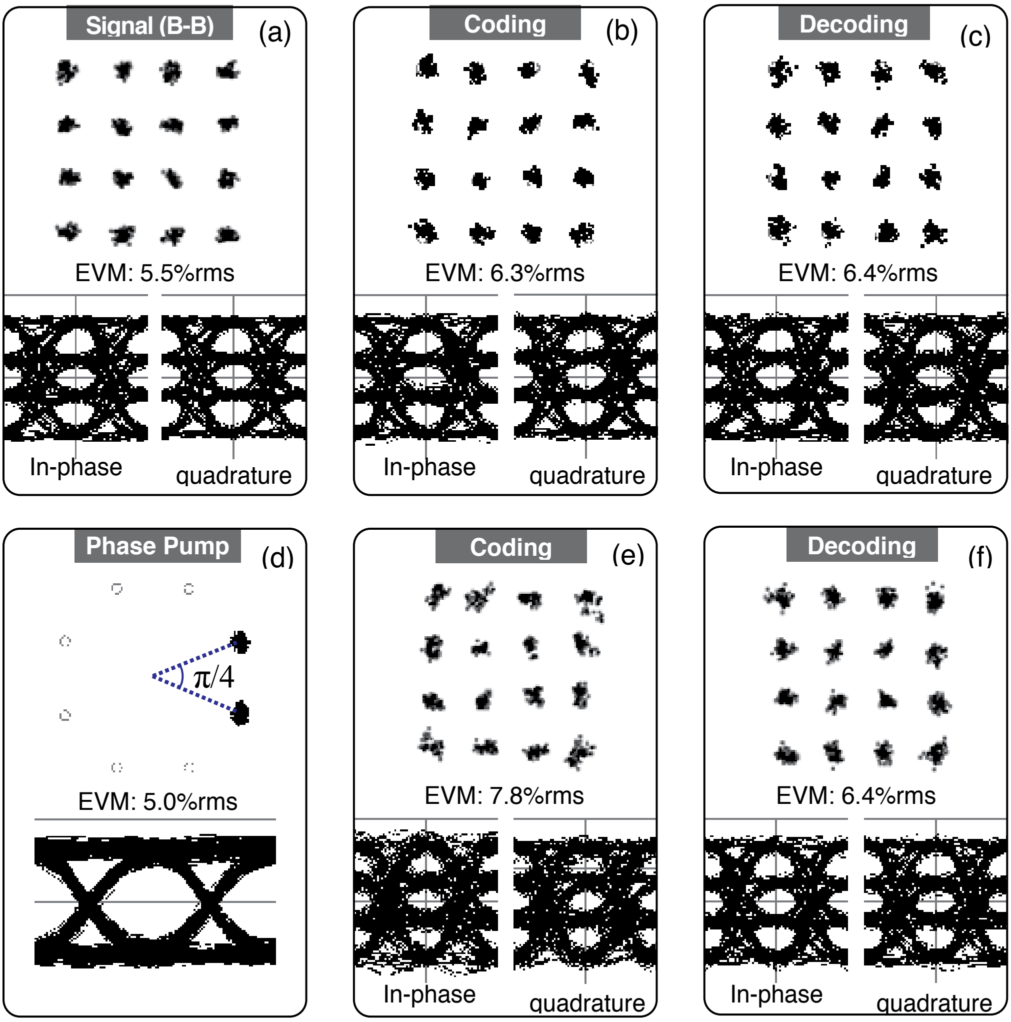


Figure 47. Measured constellation diagrams and in-phase (I) and quadrature (Q) components for high-base coding/decoding. Degenerate FWM in HNLF, 16-QAM signal, and CW/phase-modulated pumps are employed to enable symbol-wise hexadecimal coding/decoding. (a) Back-to-back (B-B) 16-QAM signal; (b) coded signal using a CW pump; (c) decoded signal using CW pump; (d) (0, π/4) phase-modulated pump; (e) coded signal using a (0, π/4) phase-modulated pump; (f) decoded signal using a (0, π/4) phase-modulated pump.

Nevertheless, the hexadecimal coding process is still verified from Fig. 49, i.e., the symbol sequence relationship of coded signal and original signal follows the coding algorithm of (0, π/4) phase-modulated pump-assisted symbol-wise hexadecimal coding. In addition, for the symbol-wise hexadecimal decoding process, the decoded signal recovers the information carried by the original signal. From the obtained results as shown in Figs. 48 and 49, one can clearly confirm the successful realization of 10-Gbaud optical variable symbol-wise hexadecimal coding/decoding by exploiting degenerate FWM in HNLF, 16-QAM signal, and CW/phase-modulated pumps.

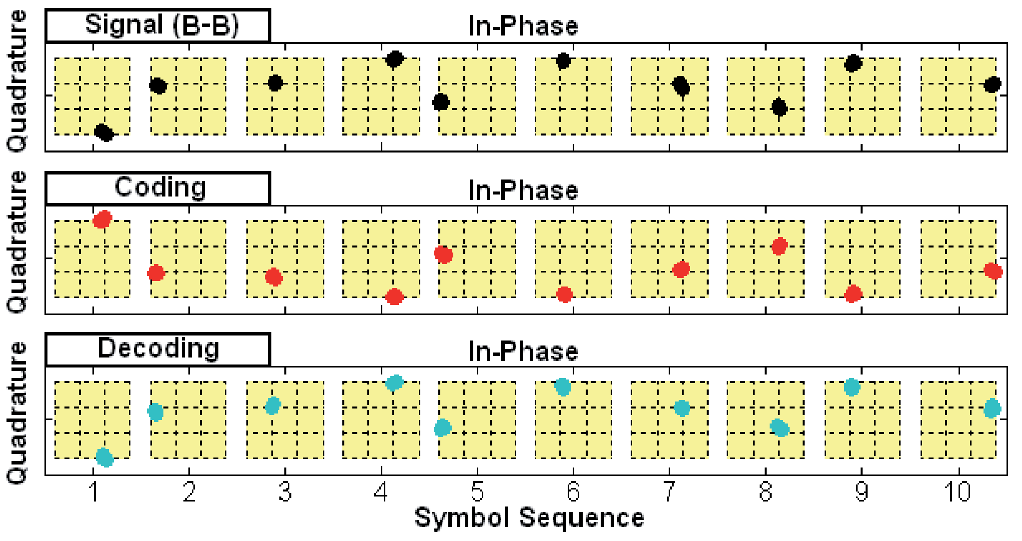


Figure 48. Measured complex amplitudes (i.e., in-phase and quadrature components) of symbol sequence for optical symbol-wise hexadecimal coding/decoding using a CW pump.

The BER performance is characterized for CW/phase-modulated pump-assisted optical variable symbol-wise hexadecimal coding/decoding. Shown in Fig. 50(a) are measured BER curves for the symbol-wise hexadecimal coding/decoding using a CW pump. OSNR penalty is used for performance evaluation defined by the ratio of the received OSNR of the coded signal to that of the back-to-back (B-B) signal. The measured OSNR penalty at a BER of 2e-3 is ~0.6 dB for CW pump-assisted symbol-wise hexadecimal coding. The measured OSNR penalty at a BER of 2e-3 for CW pump-assisted symbol-wise hexadecimal decoding, i.e., the ratio of the received OSNR of the decoded signal to that of the B-B signal, is around 1.1 dB. Shown in Fig. 50(b) are measured BER curves for the symbol-wise hexadecimal coding/decoding using a (0, π/4) phase-modulated pump. From Fig. 50(b) one can see that the OSNR penalty at a BER of 2e-3 is measured to be ~1.2 dB for symbol-wise hexadecimal coding process and ~0.9 dB for symbol-wise hexadecimal decoding process, respectively.

We study the BER performance of symbol-wise hexadecimal coding/decoding as a function of the pump phase modulation depth. Figure 51(a) and (b) show measured results for symbol-wise hexadecimal coding and decoding, respectively. The OSNR is fixed around 20 dB. For the symbol-wise hexadecimal coding process as shown in Fig. 51(a), the coding operation performance is sensitive to the pump phase modulation depth. In contrast, for the symbol-wise hexadecimal decoding process as shown in Fig. 51(b), the decoding operation performance changes slightly. Such interesting phenomenon can be briefly explained as follows. For the symbol-wise hexadecimal coding process with the coding algorithm of$\Phi_{C\,i}=2\Phi_{K\,i}-\Phi_{P\,i}$, twice phase modulation of the pump is added to the coded signal. As a result, any change of the pump phase modulation depth and resultant offset from π/4 pump phase modulation can cause the deviation of the constellation points of 16-QAM from their standard positions. Thus, the coding performance is degraded for symbol-wise hexadecimal coding process. To maintain

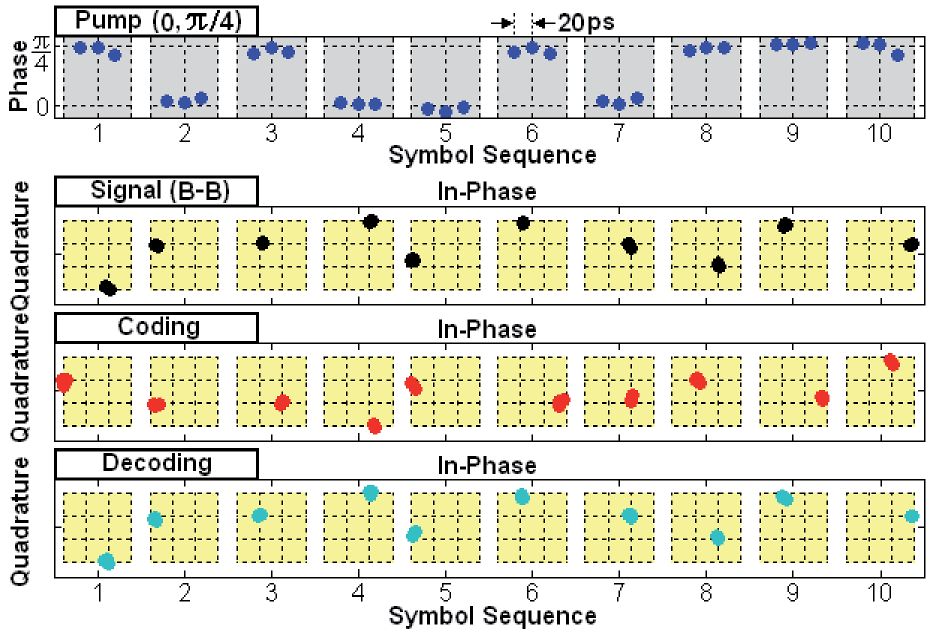


Figure 49. Measured complex amplitudes of symbol sequence for optical symbol-wise hexadecimal coding/decoding using a phase-modulated pump. A binary phase modulation of (0, π/4) is applied to the pump.

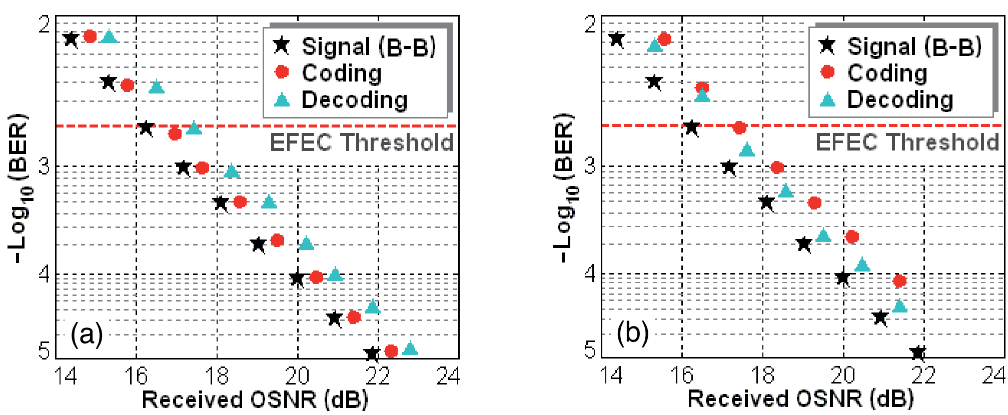


Figure 50. Measured BER curves for optical variable symbol-wise hexadecimal coding/decoding. (a) CW pump; (b) (0, π/4) phase-modulated pump.

the BER below 2e-3 (EFEC threshold), the tolerance of the pump phase modulation offset is assessed to be about 0.023π, as shown in Fig. 51(a). For the symbol-wise hexadecimal decoding process with the decoding algorithm of $2\Phi_{K\,i}-\Phi_{C\,i}=2\Phi_{K\,i}-(2\Phi_{K\,i}-\Phi_{P\,i})=\Phi_{P\,i}$ algorithms, it is easy to understand that the BER performance of the decoded signal is independent on the pump phase modulation, i.e., insensitive to the modulation depth of the pump as shown in Fig. 51(b).

http://dx.doi.org/10.5772/61504

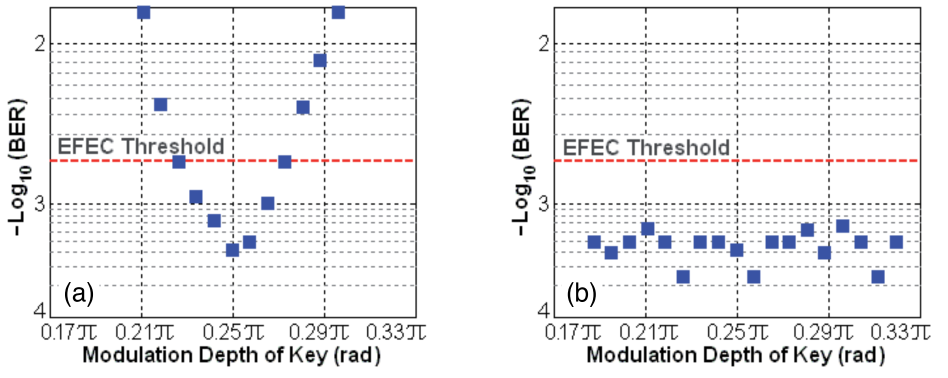


Figure 51. Measured dependence of BER performance on the phase modulation depth of pump. (a) symbol-wise hexadecimal coding; (b) symbol-wise hexadecimal decoding.

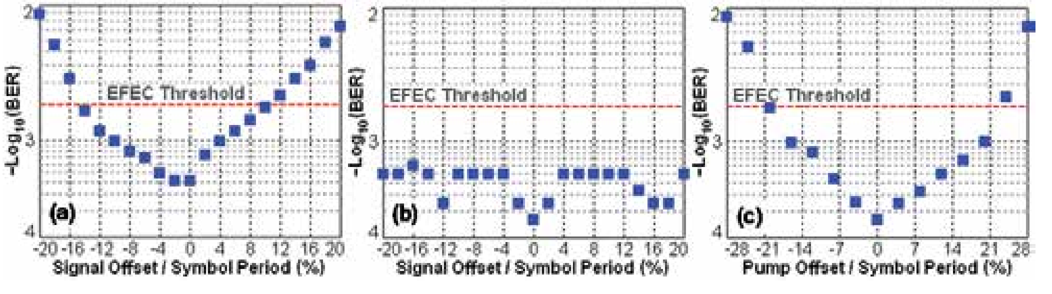


Figure 52. Measured BER performance of symbol-wise hexadecimal coding/decoding versus signal/pump offset in the time domain. (a) Measured BER of coding as a function of the offset in the time domain between the signal and the pump for coding. The pump for decoding is not involved; (b) Measured BER of decoding as a function of the offset in the time domain between the signal and the pump for decoding. The pump for decoding is aligned to the pump for coding; (c) Measured BER of decoding as a function of the offset in the time domain between the pump for decoding and the pump for coding. The pump for coding is aligned to the signal.

We further evaluate the BER performance of symbol-wise hexadecimal coding/decoding versus the signal offset and pump offset in the time domain, as shown in Fig. 52. In the experiment, the OSNR is fixed around 20 dB. Figure 52(a) depicts the measured BER of symbol-wise hexadecimal coding as a function of the offset in the time domain between the signal and the pump for coding. Note that the pump for decoding is not involved. It is shown that the coding is sensitive to the signal offset from the pump. This is predictable according to the coding algorithm of $\Phi_{C\,i}=2\Phi_{K\,i}-\Phi_{P\,i}$. To keep the BER below 2e-3 (EFEC threshold), the tolerance of the relative signal offset to the symbol period is measured to be about 10%. Figure 52(b) plots the measured BER of symbol-wise hexadecimal decoding as a function of the offset in the time domain between the signal and the pump for decoding. The pump for decoding is aligned to the pump for coding. One can clearly see that the BER performance is insensitive to the signal offset in the time domain. This is easy to understand based on the decoding algorithm of $2\Phi_{K\,i}-\Phi_{C\,i}=2\Phi_{K\,i}-(2\Phi_{K\,i}-\Phi_{P\,i})=\Phi_{P\,i}$. Figure 52(c) shows measured BER of symbol-wise hexadecimal decoding as a function of the offset in the time domain between the pump for decoding and that for coding. The pump for coding is aligned to the signal. It can be clearly seen that the performance of decoding process is dependent on the offset in the time domain between the pump for decoding and that for coding. To maintain the BER below 2e-3 (EFEC threshold), the tolerance of the relative pump offset to the symbol period is assessed to be about 20%.

6. Conclusion

In this chapter, we have reviewed recent research efforts toward high-base optical signal processing by adopting multilevel modulation signals and exploiting optical nonlinearities.

1. High-Base Wavelength Conversion: On-chip, high-base, all-optical wavelength conversion of multicarrier, multilevel modulation signals has been demonstrated using degenerate FWM in a silicon waveguide and OFDM m-QAM signals. Impressive operation performance of on-chip 3.2 Gbaud/s OFDM 16/32/64/128-QAM wavelength conversion has been achieved in the experiment.
2. High-Base Optical Data Exchange: Phase-transparent, high-base optical data exchange between two 100-Gbit/s DQPSK signals has been demonstrated using the parametric depletion effect of nondegenerate FWM in an HNLF. Simultaneous multichannel data exchange has been proposed and demonstrated using bidirectional degenerate FWM in a single HNLF. Moreover, a reconfigurable Tbit/s network switching element using double-pass LCoS technology accompanied by bidirectional degenerate FWM in a single HNLF has been proposed. 2.3-Tbit/s multifunctional grooming switch has been demonstrated in the experiment, performing simultaneous selective high-base add/drop, high-base switchable data exchange, and high-base power equalization, for ITU-grid-compatible 23-channel 100-Gbit/s RZ-DQPSK signals. Additionally, ultrahigh-speed high-base optical data exchange of 640 Gbaud (2.56 Tbit/s) 16-QAM and 640 Gbaud (3.84 Tbit/s) 64-QAM signals has been proposed and simulated by exploiting non-degenerate FWM in a silicon-organic hybrid slot waveguide.
3. High-Base Optical Computing: By adopting 100-Gbit/s two-input RZ-DQPSK signals (A, B) and exploiting three degenerate FWM processes and three nondegenerate FWM processes in an HNLF, simultaneous 50-Gbaud two-input quaternary addition (A+B), dual-directional subtraction (A-B, B-A), complement (-A, -B), and doubling (2B) have been demonstrated in the experiment. By employing 100-Gbit/s three-input RZ-DQPSK signals (A, B, C/-C) and three nondegenerate FWM processes in an HNLF, 50-Gbaud three-input quaternary hybrid addition and subtraction (A+B-C, A+C-B, B+C-A, A+B+C) have been demonstrated in the experiment. Furthermore, three-input (A, B, C) 40-Gbaud (160-Gbit/s) optical hexadecimal addition/subtraction (A+B-C, A+C-B, B+C-A, A+B+C, A-B-C, B-A-C) has also been proposed and simulated based on nondegenerate FWM in a silicon–organic hybrid slot waveguide.
4. High-Base Optical Coding/Decoding: By exploiting degenerate FWM in an HNLF and adopting 16-QAM signal, 10-Gbaud optical variable symbol-wise hexadecimal coding/

http://dx.doi.org/10.5772/61504

decoding assisted by a CW pump or a phase-modulated pump has been demonstrated in the experiment. The former takes the coding through the phase conjugation of degenerate FWM, and the latter offers enhanced coding via the combined contributions from the phase modulation of the pump and the phase-conjugated FWM.

Beyond high-base wavelength conversion, data exchange, optical computing, and optical coding/decoding based on degenerate/nondegenerate FWM in HNLFs or silicon waveguides, with future improvements, other different optical nonlinearities on various nonlinear optical device platforms would also be employed to flexibly manipulate the amplitude and phase information of advanced multilevel modulation signals, which might open diverse interesting applications in robust high-base optical signal processing.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC) under grants 61222502, 11574001, 11274131, and 61077051; the Program for New Century Excellent Talents in University (NCET-11-0182); the National Basic Research Program of China (973 Program) under grant 2014CB340004; the Wuhan Science and Technology Plan Project under grant 2014070404010201; the Fundamental Research Funds for the Central Universities (HUST) under grants 2012YQ008 and 2013ZZGH003. The authors thank the Center of Micro-Fabrication and Characterization (CMFC) of WNLO for the support in the manufacturing process of silicon waveguides. The authors also thank the facility support of the Center for Nanoscale Characterization and Devices of WNLO.

Author details

Jian Wang[1*] and Alan E. Willner[2]

*Address all correspondence to: jwang@hust.edu.cn

1 Wuhan National Laboratory for Optoelectronics, School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan, Hubei, China

2 Department of Electrical Engineering, University of Southern California, Los Angeles, California, USA

References

[1] P. Kaminow, T. Li, and A. E. Willner, *Optical Fiber Telecommunications VIB Systems and Networks*, Sixth Edition (Elsevier, 2013).

[2] P. P. Mitra and J. B. Stark, "Nonlinear limits to the information capacity of optical fibre communications," *Nature* 411, 1027–1030 (2001).

[3] R. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightwave Technol*. 28, 662–701 (2010).

[4] D. J. Richardson, J. M. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres," *Nature Photon*. 7, 354–362 (2013).

[5] P. J. Winzer, "Making spatial multiplexing a reality," Nature Photon. 8, 345–348 (2014).

[6] M. Saruwatari, "All-optical signal processing for terabit/second optical transmission," *IEEE J. Sel. Top. Quantum Electron*. 6, 1363–1374 (2000).

[7] A. E. Willner, S. Khaleghi, M. R. Chitgarha, and O. F. Yilmaz, "All-optical signal processing," *J. Lightwave Technol.* 32, 660–680 (2014).

[8] D. Cotter, R. J. Manning, K. J. Blow, A. D. Ellis, A. E. Kelly, D. Nesset, I. D. Phillips, A. J. Poustie, and D. C. Rogers, "Nonlinear optics for high-speed digital information processing," *Science* 286, 1523–1528 (1999).

[9] S. Watanabe and F. Futami, "All-optical signal processing using highly nonlinear optical fibers," *IEICE Trans. Electron.* E84-B, 1179–1189 (2001).

[10] S. Radic and C. J. McKinstrie, "Optical amplification and signal processing in highly nonlinear optical fiber," *IEICE Trans. Electron.* E88-C, 859–869 (2005).

[11] M. D. Pelusi, V. G. Ta'eed, L. B. Fu, E. Mägi, M. R. E. Lamont, S. Madden, D.-Y. Choi, D. A. P. Bulla, B. Luther-Davies, and B. J. Eggleton, "Applications of highly-nonlinear chalcogenide glass devices tailored for high-speed all-optical signal processing," *IEEE J. Sel. Top. Quantum Electron*. 14, 529–539 (2008).

[12] M. D. Pelusi, F. Luan, E. Magi, M. R. E. Lamont, D. J. Moss, B. J. Eggleton, J. S. Sanghera, L. B. Shaw, and I. D. Aggarwal, "High bit rate all-optical signal processing in fiber photonic wire," *Opt. Express*. 16, 11506–11512 (2008).

[13] C. Langrock, S. Kumar, J. E. McGeehan, A. E. Willner, and M. M. Fejer, "All-optical signal processing using $\chi^{(2)}$ nonlinearities in guided-wave devices," *J. Lightwave Technol*. 24, 2579–2592 (2006).

[14] R. Salem, M. A. Foster, A. C. Turner, D. F. Geraghty, M. Lipson, and A. L. Gaeta, "Signal regeneration using low-power four-wave mixing on silicon chip," *Nature Photon*. 2, 35–38 (2008).

[15] C. Koos, P. Vorreau, T. Vallaitis, P. Dumon, W. Bogaerts, R. Baets, B. Esembeson, I. Biaggio, T. Michinobu, F. Diederich, W. Freude, and J. Leuthold, "All-optical high-speed signal processing with silicon-organic hybrid slot waveguides," *Nature Photon*. 3, 216–219 (2009).

[16] K. Nozaki, T. Tanabe, A. Shinya, S. Matsuo, T. Sato, H. Taniyama, and M. Notomi, "Sub-femtojoule all-optical switching using a photonic-crystal nanocavity," *Nature Photon*. 4, 477–483 (2010).

[17] J. Leuthold, C. Koos, and W. Freude, "Nonlinear silicon photonics," *Nature Photon*. 4, 535–544 (2010).

[18] A. E. Willner, O. F. Yilmaz, J. Wang, X. X. Wu, A. Bogoni, L. Zhang, and S. R. Nuccio, "Optically efficient nonlinear signal processing," *IEEE J. Sel. Top. Quantum Electron*. 17, 320–332 (2011).

[19] L. K. Oxenløwe, H. Ji, M. Galili, M. Pu, H. Hu, H. C. H. Mulvad, K. Yvind, J. M. Hvam, A. T. Clausen, and P. Jeppesen, "Silicon photonics for signal processing of Tbit/s serial data signals," *IEEE J. Sel. Top. Quantum Electron*. 18, 996–1005 (2011).

[20] B. J. Eggleton, T. D. Vo, R. Pant, J. Schr, M. D. Pelusi, D. Yong Choi, S. J. Madden, and B. Luther-Davies, "Photonic chip based ultrafast optical processing based on high nonlinearity dispersion engineered chalcogenide waveguides," *Laser Photon. Rev*. 6, 97–114 (2012).

[21] S. J. B. Yoo, "Wavelength conversion technologies for WDM network applications," *J. Lightwave Technol*. 14, 955–966 (1999).

[22] H. Hu, H. Ji, M. Galili, M. Pu, C. Peucheret, H. C. H. Mulvad, K. Yvind, J. M. Hvam, P. Jeppesen, and L. K. Oxenløwe, "Ultra-high-speed wavelength conversion in a silicon photonic chip," *Opt. Express.* 19, 19886–19894 (2011).

[23] C.-S. Bres, N. Alic, E. Myslivets, and S. Radic, "Scalable multicasting in one-pump parametric amplifier," *J. Lightwave Technol*. 27, 356–363 (2009).

[24] X. Wu, A. Bogoni, O. F. Yilmaz, S. R. Nuccio, J. Wang, and A. E. Willner, "8-fold 40–320 Gbit/s phase-coherent multiplexing and 320-40 Gbit/s demultiplexing using highly nonlinear fibers," *Opt. Lett*. 35, 1896–1898 (2010).

[25] A. Biberman, B. G. Lee, A. C. Turner-Foster, M. A. Foster, M. Lipson, A. L. Gaeta, and K. Bergman, "Wavelength multicasting in silicon photonic nanowires," *Opt. Express* 18, 18047–18055 (2010).

[26] O. F. Yilmaz, J. Wang, S. Khaleghi, X. Wang, S. R. Nuccio, X. Wu, and A. E. Willner, "Preconversion phase modulation of input differential phase-shift-keying signals for wavelength conversion and multicasting applications using phase-modulated pumps," *Opt. Lett*. 36, 731–733 (2011).

[27] K. Uesaka, K. K.-Y. Wong, M. E. Marhic, and L. G. Kazovsky, "Wavelength exchange in a highly nonlinear dispersion-shifted fiber: Theory and experiments," *IEEE J. Sel. Top. Quantum Electron*. 8, 560–568 (2002).

[28] R. W. L. Fung, Henry K. Y. Cheung, and K. K. Y. Wong, "Widely tunable wavelength exchange in anomalous-dispersion regime," *IEEE Photon. Technol. Lett*. 19, 1846–1848 (2007).

[29] J. Wang and Q. Sun, "Theoretical analysis of power swapping in quadratic nonlinear medium," *Appl. Phys. Lett*. 96, 081108 (2010).

[30] J. Wang, S. R. Nuccio, X. Wu, O. F. Yilmaz, L. Zhang, I. Fazal, J. Y. Yang, Y. Yue, and A. E. Willner, "40 Gbit/s optical data exchange between wavelength-division-multiplexed channels using a periodically poled lithium niobate waveguide," *Opt. Lett*. 35, 1067–1069 (2010).

[31] J. Wang, Z. Bakhtiari, S. R. Nuccio, O. F. Yilmaz, X. Wu, and A. E. Willner, "Phase-transparent optical data exchange of 40 Gbit/s differential phase-shift keying signals," *Opt. Lett*. 35, 2979–2981 (2010).

[32] C. Schubert, C. Schmidt, S. Ferber, R. Ludwig, and H. G. Weber, "Error free all-optical add-drop multiplexing at 160 Gbit/s," *Electron. Lett*. 39, 1074–1076 (2003).

[33] H. Ji, M. Pu, H. Hu, M. Galili, L. Oxenløwe, K. Yvind, J. M. Hvam, and P. Jeppesen, "Optical waveform sampling and error-free demultiplexing of 1.28 Tb/s serial data in a nanoengineered silicon waveguide," *J. Lightwave Technol*. 29, 426–431 (2011).

[34] A. Bogoni, X. Wu, S. R. Nuccio, J. Wang, Z. Bakhtiari, and A. E. Willner, "Photonic 640-Gb/s reconfigurable OTDM add-drop multiplexer based on pump depletion in a single PPLN waveguide," *IEEE J. Sel. Top. Quantum Electron*. 18, 709–716 (2012).

[35] J. Wang, H. Fu, D. Geng, and A. E. Willner, "Single-PPLN-assisted wavelength-time-selective switching dropping swapping for 100-GHz-spaced WDM signals," *Opt. Express.* 21, 3756–3774 (2013).

[36] J. Wang, J. Q. Sun, and Q. Z. Sun, "Single-PPLN-based simultaneous half-adder, half-subtracter, and OR logic gate: proposal and simulation," *Opt. Express*. 15, 1690–1699 (2007)

[37] J. Wang, J. Q. Sun, X. L. Zhang, D. X. Huang, and M. M. Fejer, "Ultrafast all-optical three-input boolean XOR operation for differential phase-shift keying signals using periodically poled lithium niobate," *Opt. Lett*. 33, 1419–1421 (2008).

[38] A. Bogoni, X. Wu, I. Fazal, and A. E. Willner, "160 Gb/s time domain channel extraction/insertion and all-optical logic operations exploiting a single PPLN waveguide," *J. Lightwave Technol*. 27, 4221–4227 (2009).

[39] A. Bogoni, X. Wu, I. Fazal, and A. E. Willner, "320 Gbit/s photonic processing based on sum/difference frequency generation and pump depletion in a single PPLN waveguide," *Opt. Lett*. 34, 1825–1827 (2009).

[40] A. Bogoni, X. Wu, Z. Bakhtiari, S. Nuccio, and A. E. Willner, "640 Gbit/s photonic logic gates," *Opt. Lett*. 35, 3955–3957 (2010).

[41] J. Wang, J. Sun, Q. Sun, D. Wang, M. Zhou, X. Zhang, D. Huang, M. M. Fejer, "All-optical format conversion using a periodically poled lithium niobate waveguide and a reflective semiconductor optical amplifier," *Appl. Phys. Lett*. 91, 051107 (2007).

[42] J. Wang, J. Sun, X. Zhang, D. Huang, and M. M. Fejer, "All-optical format conversions using periodically poled lithium niobate waveguides," *IEEE J. Quantum Electron*. 45, 195–205 (2009).

[43] J. Wang, J. Sun, X. Zhang, D. Huang, and M. M. Fejer, "Optical phase erasure and its application to format conversion through cascaded second-order processes in periodically poled lithium niobate," *Opt. Lett*. 33, 1804–1806 (2008).

[44] J. Wang, Q. Z. Sun, and J. Q. Sun, "All-optical 40 Gbit/s CSRZ-DPSK logic XOR gate and format conversion using four-wave mixing," *Opt. Express*. 17, 12555–12563 (2009).

[45] M. S. Rasras, I. Kang, M. Dinu, J. Jaques, N. Dutta, A. Piccirilli, M. A. Cappuzzo, E. Y. Chen, L. T. Gomez, A. Wong-Foy, S. Cabot, G. S. Johnson, L. Buhl, and S. S. Patel, "A programmable 8-bit optical correlator filter for optical bit pattern recognition," *IEEE Photon. Technol. Lett*. 20, 694–696 (2008).

[46] K. Croussore and G. Li, "Phase and amplitude regeneration of differential phase-shift keyed signals using phase-sensitive amplification," *IEEE J. Sel. Top. Quantum Electron*. 14, 648–658 (2008).

[47] R. Slavík, F. Parmigiani, J. Kakande, C. LundstrÖm, M. SjÖdin, P. A. Andrekson, R. Weerasuriya, S. Sygletos, A. D. Ellis, L. GrÜner-Nielsen, D. Jakobsen, S. Herstrøm, R. Phelan, J. O'Gorman, A. Bogris, D. Syvridis, S. Dasgupta, P. Petropoulos, and D. J. Richardson, "Alloptical phase and amplitude regenerator for next-generation telecommunications systems," *Nature Photon*. 4, 690–695 (2010).

[48] Z. Tong, C. Lundström, P. A. Andrekson, C. J. McKinstrie, M. Karlsson, D. J. Blessing, E. Tipsuwannakul, B. J. Puttnam, H. Toda, and L. Grüner-Nielsen, "Towards ultrasensitive optical links enabled by low-noise phase-sensitive amplifiers," *Nature Photon*. 5, 430–436 (2011).

[49] O. F. Yilmaz, S. R. Nuccio, X. Wu, and A. E. Willner, "40-Gb/s optical packet buffer using conversion/dispersion-based delays," *J. Lightwave Technol*. 28, 616–623 (2010).

[50] S. R. Nuccio, O. F. Yilmaz, X. Wu, and A. E. Willner, "Fine tuning of conversion/dispersion based optical delays with a 1 pm tunable laser using cascaded acousto-optic mixing," *Opt. Lett*. 35, 523–525 (2010).

[51] S. R. Nuccio, O. F. Yilmaz, X. Wang, H. Huang, J. Wang, X. Wu, and A. E. Willner, "Higher-order dispersion compensation to enable a 3.6 μs wavelength-maintaining delay of a 100 Gb/s DQPSK signal," *Opt. Lett*. 35, 2985–2987 (2010).

[52] Y. Dai, Y. Okawachi, A. C. Turner-Foster, M. Lipson, A. L. Gaeta, and C. Xu, "Ultra-long continuously tunable parametric delays via a cascading discrete stage," *Opt. Express*. 18, 333–339 (2010).

[53] N. Alic, E. Myslivets, S. Moro, B. P.-P. Kuo, R. M. Jopson, C. J. McKinstrie, and S. Radic, "Microsecond parametric optical delays," *J. Lightwave Technol*. 28, 448–455 (2010).

[54] P. J. Winzer and R. Essiambre, "Advanced optical modulation formats," *Proc. IEEE*. 94, 952–985 (2006)

[55] P. J. Winzer and R. Essiambre, "Advanced modulation formats for high-capacity optical transport networks," *J. Lightwave Technol*. 24, 4711–4728 (2006).

[56] P. J. Winzer, "Modulation and multiplexing in optical communication systems," *IEEE LEOS Newsletter*. 23, 4–10 (2009).

[57] P. J. Winzer, "High-spectral-efficiency optical modulation formats," *J. Lightwave Technol*. 30, 3824–3835 (2012).

[58] X. Zhou and J. Yu, "Multi-level, multi-dimensional coding for high-speed and high spectral-efficiency optical transmission," *J. Lightwave Technol*. 27, 3641–3653 (2009).

[59] X. Zhou, "Digital signal processing for coherent multi-level modulation formats," *Chin. Opt. Lett*. 8, 863–870 (2010).

[60] J. Yu and X. Zhou, "Multilevel modulations and digital coherent detection," *Opt. Fiber Technol*. 15, 197–208 (2009).

[61] J. Yu and X. Zhou, "Ultra-high-capacity DWDM transmission system for 100 G and beyond," *IEEE Commun. Mag*. 48, S56–S64 (2010).

[62] J. Yu, "Beyond 100G Ethernet," *IEEE Commun. Mag*. 48, 26-30 (2010).

[63] S. J. Savory, "Digital coherent optical receivers: algorithms and subsystems," *IEEE J. Sel. Top. Quantum Electron*. 16, 1164–1179 (2010).

[64] P. J. Winzer, G. Raybon, H. Song, A. Adamiecki, S. Corteselli, A. H. Gnauck, D. A. Fishman, C. R. Doerr, S. Chandrasekhar, L. L. Buhl, T. J. Xia, G. Wellbrock, W. Lee, B. Basch, T. Kawanishi, K. Higuma, and Y. Painchaud, "100-Gb/s DQPSK transmission: from laboratory experiments to field trials," *J. Lightwave Technol*. 26, 3388–3402 (2008).

[65] X. Zhou, J. Yu, D. Qian, T. Wang, G. Zhang, and P. D. Magill, "High-spectral-efficiency 114-Gb/s transmission using polmux-RZ-8PSK modulation format and single-ended digital coherent detection technique," *J. Lightwave Technol*. 27, 146–152 (2009).

[66] P. J. Winzer, A. H. Gnauck, C. R. Doerr, M. Magarini, and L. L. Buhl, "Spectrally efficient long-haul optical networking using 112-Gb/s polarization-multiplexed 16-QAM," *J. Lightwave Technol*. 28, 547–556 (2010).

[67] A. H. Gnauck, P. J. Winzer, S. Chandrasekhar, X. Liu, B. Zhu, and D. W. Peckham, "Spectrally efficient long-haul WDM transmission using 224-Gb/s polarization-multiplexed 16-QAM," *J. Lightwave Technol*. 29, 373–377 (2011).

[68] X. Liu, S. Chandrasekhar, X. Chen, P. J. Winzer, Y. Pan, T. F. Taunay, B. Zhu, M. Fishteyn, M. F. Yan, J. M. Fini, E.M. Monberg, and F.V. Dimarcello, "1.12-Tb/s 32-QAM-OFDM superchannel with 8.6-b/s/Hz intrachannel spectral efficiency and space-division multiplexed transmission with 60-b/s/Hz aggregate spectral efficiency," *Opt. Express*. 19, B958–B964 (2011).

[69] D. Qian, M.-F. Huang, E. Ip, Y.-K. Huang, Y. Shao, J. Hu, and T. Wang, "High capacity/spectral efficiency 101.7-Tb/s WDM transmission using PDM-128QAM-OFDM over 165-km SSMF within C- and L-bands," *J. Lightwave Technol*. 30, 1540–1548 (2012).

[70] G.-W. Lu and T. Miyazaki, "Optical phase erasure based on FWM in HNLF enabling format conversion from 320-Gb/s RZ-DQPSK to 160-Gb/s RZ-DPSK," *Opt. Express*. 17(16), 13346–13353 (2009).

[71] J. Wang, S. R. Nuccio, H. Huang, X. Wang, J.-Y. Yang, and A. E. Willner, "Optical data exchange of 100-Gbit/s DQPSK signals," *Opt. Express*. 18, 23740–23745 (2010).

[72] G.-W. Lu, E. Tipsuwannakul, T. Miyazaki, C. Lundström, M. Karlsson, P. A. Andrekson, "Format conversion of optical multilevel signals using FWM-based optical phase erasure," *J. Lightwave Technol*. 29(16), 2460–2466 (2011).

[73] J. Wang, H. Huang, X. Wang, J.-Y. Yang, and A. E. Willner, "Multi-channel 100-Gbit/s DQPSK data exchange using bidirectional degenerate four-wave mixing," *Opt. Express*. 19, 3332–3338 (2011).

[74] J. Wang, H. Huang, X. Wang, J.-Y. Yang, and A. E. Willner, "Reconfigurable 2.3-Tbit/s DQPSK simultaneous add/drop, data exchange and equalization using double-pass LCoS and bidirectional HNLF," *Opt. Express*. 19, 18246–18252 (2011).

[75] J. Wang, J.-Y. Yang, X. Wu, O. F. Yilmaz, S. R. Nuccio, and A. E. Willner, "40-Gbaud/s (120-Gbit/s) octal and 10-Gbaud/s (40-Gbit/s) hexadecimal simultaneous addition and subtraction using 8PSK/16PSK and highly nonlinear fiber," in *Proc. Opt. Fiber Commun. Conf., Mar*. 2011, pp. 1–3.

[76] E. Lazzeri, A. Malacarne, G. Serafino, and A. Bogoni, "Optical XOR for error detection and coding of QPSK I and Q components in PPLN waveguide," *IEEE Photon. Technol. Lett*. 24, 2258–2261 (2012).

[77] J. Wang, S. R. Nuccio, J.-Y. Yang, X. X. Wu, A. Bogoni, and A. E. Willner, "High-speed addition/subtraction/complement/doubling of quaternary numbers using optical nonlinearities and DQPSK signals," *Opt. Lett*. 37, 1139–1141 (2012).

[78] A. Malacarne, E. Lazzeri, V. Vercesi, M. Scaffardi, and A. Bogoni, "Colorless all-optical sum and subtraction of phases for phase-shift keying signals based on a periodically poled lithium niobate waveguide," *Opt. Lett*. 37, 3831–3833 (2012).

[79] J. Wang, J.-Y.Yang, X. X. Wu, and A. E. Willner, "Optical hexadecimal coding/decoding using 16-QAM signal and FWM in HNLFs," *J. Lightwave Technol*. 30, 2890–2900 (2012).

[80] J. Wang, J.-Y. Yang, H. Huang, and A. E. Willner, "Three-input optical addition and subtraction of quaternary base numbers," *Opt. Express*. 21, 488–499 (2013).

[81] A. Bogris, "All-optical demultiplexing of 16-QAM signals into QPSK tributaries using four-level optical phase quantizers," *Opt. Lett*. 39, 1775–1778 (2014).

[82] T. Umeki, O. Tadanaga, M. Asobe, Y. Miyamoto, and H. Takenouchi, "First demonstration of high-order QAM signal amplification in PPLN-based phase sensitive amplifier," *Opt. Express*. 22, 2473–2482 (2014).

[83] C. Li, C. C. Gui, X. Xiao, Q. Yang, S. Yu, and J. Wang, "On-chip all-optical wavelength conversion of multicarrier, multilevel modulation (OFDM m-QAM) signals using a silicon waveguide," *Opt. Lett*. 39, 4583-4586 (2014).

[84] R. Adams, M. Spasojevic, M. Chagnon, M. Malekiha, J. Li, D. V. Plant, and L. R. Chen, "Wavelength conversion of 28 GBaud 16-QAM signals based on four-wave mixing in a silicon nanowire," *Opt. Express*. 22, 4083–4090 (2014).

[85] C. C. Gui and J. Wang, "Silicon-organic hybrid slot waveguide based three-input multicasted optical hexadecimal addition/subtraction," *Sci. Rep*. 4, 7491 (2014).

[86] C. C. Gui and J. Wang, "Optical data exchange of m-QAM signals using a silicon-organic hybrid slot waveguide: proposal and simulation," Opt. Express. 22, 24796-24807 (2014).

Multitones' Performance for Ultra Wideband Software Defined Radar

Julien Le Kernec and Olivier Romain

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/60804

Abstract

This chapter proposes and tests an approach for an unbiased study of radar waveforms' performances. Through an empirical performance analysis, the performances of Chirp and Multitones are compared with both simulations and measurements. An ultra wideband software defined radar prototype was designed and the prototype has performances comparable to the state of the art in software defined radar. The study looks at peak-to-mean-envelope power ratio, spectrum efficiency, and pulse compression as independent waveform criteria. The experimental results are consistent with the simulations. The study shows that a minimum of 10 bits resolution for the AD/DA converters is required to obtain near-optimum performances.

Keywords: Software Defined Radar, OFDM, Empirical Modelling, Chirp

1. Introduction

In the past few decades, analogue circuits have been replaced by digital circuits. This evolution has permitted the use of purely digital waveforms (such as Multitones which have numerous commercial applications in the wireless communication industry – such as wireless LAN [1]) which present numerous advantages (i.e., increased data throughput, robustness against fading). To date, Multitones have seldom been implemented in operational radar.

Operational radar predominantly uses the linear frequency modulated pulse (also known as Chirp) and has been routinely used since the late 1940s [2]. The relatively slow adoption of Multitones in radar applications can be explained by a variety of factors. For example, it is unlikely that a technology will advance to marketable applications unless there is demand for them. Lately, the use of a Unmanned Airborne Vehicles for military operations over urban areas are required to simultaneously perform radar sensing and remotely communicate data to a base station. This cannot be achieved with traditional Chirp. Consequently, there have been increased research efforts in integrating telecommunication waveforms such as Multitones into radar applications.

The constant developments in ADC/DAC, digital signal processors, signal synthesis/digitization, and component's instantaneous bandwidth allow digital platforms to process ultra wideband (UWB) signals. In radar applications, UWB signals enable finer slant range resolution for target identification and the implementation of waveform/spectrum diversity. Those recent technological developments constitute the foundation of software-defined radar, which can dynamically reconfigure its digital signal processing and adapt the frequency of converter. Such radar is inherently multifunction switching from operating mode to another (surveillance, tracking, imaging, and telecommunications).

Multitones will only be widely adopted when its capabilities match the specific task's requirements. The successful integration and subsequent widespread use in operational systems depends solely on that condition. In other words, without a viable commercial application, the development of a technology is unlikely to succeed.

Considering the capabilities of Multitones and/or OFDM signals with respect to classical radar waveforms, the second half of the introduction provides an overview of the literature on the subject.

Refs. [3-5] concern the communication aspect of multi-carriers in radar, leaving radar performances with multicarrier signals aside. A comparison of performances is found in terms of detection in Ref. [6]. The authors compared single carrier and multicarrier radar systems in simulations. They found that for target detection in radar based on multicarrier modulation, the required constant false alarm rate detection threshold is lower than for a single carrier radar system with polyphase codes.

In Ref. [2], it is shown that trains of diverse Multitone pulses coded in phase and amplitude yielded near thumbtack ambiguity functions. These ambiguity functions do not suffer from range-Doppler coupling as Chirp does. In Ref. [7], a near thumbtack ambiguity function is obtained using random spread tone agility. In both cases, this ambiguity function comes at the cost of a higher pedestal level.

Finally, new processing capabilities are emerging using the Multitones' structure such as Doppler resolution while using agility [7]. This particular feature cannot be performed with classic radar waveforms while using agility. In Ref. [8], the Doppler ambiguity is resolved over one pulse train.

For those reasons, Multitones are foreseen as a viable prospect for the future digital software defined radar. In order to improve power amplifier efficiency, Peak-to-Mean Envelope Power Ratio (PMEPR) reduction schemes (phase/amplitude modulation) are overlaid on Multitones. This signal can be a composite of independent bands for separate processing in multimode

http://dx.doi.org/10.5772/60804

scenarios [9]. Also in the presence of frequency selective fading, Multitones can still ensure successful detection of the target [10]. The waveform/spectrum agility is essential for stealthy operations to evade jamming and spectrum reuse with radar networks [11].

Based on these studies of the performances of Multitones compared to classic radar waveforms, Multitones show a great potential for new radar advances. However, it is important to note that most of these results come from simulations. For a real evaluation of the potential performances of Multitones for radar systems, experimental validations are required. Hence the simulations presented in section 4 will be compared to experimental results in section 5 based on an experimental setup that is described in part 3.5. Also Multitones need to be compared to a reference in radar applications: the linear frequency modulated (LFM) signal aka Chirp [2].

For this chapter, the focus will be on the performances of both Multitones and Chirp with respect to quantization. The underlying issue of implementation is the effect of the radio frequencies (RF) equipment on radar performances. The DAC and ADC converters determine the radar's dynamic range and thus contribute greatly to detection capabilities. On the subject of Multitones and quantification, the literature focuses on telecommunications [3, 12, 13], so there is no evaluation of radar performance. Note that to the best of the authors' knowledge, the literature is lacking on this particular subject for radar applications. The following section will review the state of the art for Multitones' performances in radar with an emphasis on the isuse of quantization.

2. State of the art

In the radar community, one of the main goals is to improve detection to see further and with a higher sensitivity. The smallest received power depends on receiver sensitivity which is closely related to the ADC resolution: In Ref. [14], the rationale behind investigating various linearization techniques was to increase radar receiver dynamic ranges for the detection of small targets with a highly cluttered background. To determine the best ADC resolution for a given application, the effects of quantization on performances must be investigated. The novel approach adopted in this chapter is to study Multitones performance for radar applications only, using unbiased tools in simulation and experimentation. Multitones will also be compared to a signal of reference in the radar community. This will position the Multitones' performance with respect to reference known to the community. The quantization process will allow determination of the limits of utilization of a given hardware with respect to performance requirements.

In Ref. [12], the author conducted a survey of ADC performances ranging from the 1970s to present day and extracted possible trends in ADC evolution forward to 2020. It is reported that most recent designs benefit from scaled device geometries and higher bandwidth, but suffer in dynamic range and sampling linearity due to reduced supply voltages and available swing. The available swing is most likely the cause for ADC saturation noise floor around -160 dB observed in the survey. It also shows that for a given ENOB, the sampling rates are entering or already are in a saturation phase and it is speculated that the improvement of state of the art sampling rates will be lower than 5 times by 2020. Now looking at the evolution of ENOB with respect to sampling frequencies, the projections show that ADCs with over 1 GS/s have not entered the saturation phase yet. The survey also shows that the main efforts in ADC research now focused more on power efficiency rather than SNDR/SNR to reduce the ADC Figure-of-Merit.

Practical use of ADCs are plagued by many physical limitations such as quantization in time and amplitude, aliasing, clock jitter aperture jitter, thermal noise, non-linear distortions (INL, DNL), etc. Some of the physical limitations can be partially compensated using oversampling. However with the high-end of wideband ADCs (e.g., Tekmicro announced a 2-channel digitizer with 5 GS/s and 10 bits resolution with 3 GHz of instantaneous bandwidth on the Proteus V6 [15] equipped with the EV10AQ190 [16] from E2V) oversampling is not an option and even if possible would be prohibitively expensive. Jitter (clock and jitter) is well known to severely reduce the achievable signal-to-noise ratio (SNR) [17].

Regarding the use of Multitones in telecommunications, common measures of spurious free dynamic range, total harmonic distortion, signal to noise and distortion and effective number of bits are defined for one tone or two tones only and the definitions used for some of the metrics are not unified. In the literature, clip correction post-processing allows the relaxation of ADC resolution constraints to improve packet error rate at the cost of an increased complexity in processing [3]. The second allows bit error rate improvements in the presence of narrowband interference [13]. In Ref. [18], the ADC resolution of multi-band and pulsed-OFDM ultra wideband systems (IEEE 802.15.3a) is derived using simulation results. They show that 4-bit resolution is enough to obtain a bit error rate with respect to SNR performances quasi-identical to the ideal case with infinite resolution.

Working on relaxing ADC requirements with digital post-processing, to compensate for the impairments of hardware ("Dirty RF") and to increase the performance of telecommunication, is a very active research field. Given the projections in Ref. [12], the ADCs' non-linearities are increasing with the reduction of voltage swing, maximum SNR capabilities for wideband digitizers are not improving or maybe will worsen, digital enhancements are going to be required especially in radar to maintain current levels of sensitivity and detection.

To the best of the authors' knowledge, the literature is mostly investigating performances for telecommunications and not for radar performances, also very few experimental results were found. Before trying to improve performances, these performances for radar have to be established and in this chapter the quantization process is investigated.

3. Empirical approach for the evaluation of the radar performances

In order to compare different waveforms on equal grounds, they have to be compared on waveform-independent criteria. Also to further this concept, the simulated processes and the experimental test bench should be identical to evaluate the performances without bias.

3.1. Waveform independent criteria

Several characteristics were chosen to determine the optimum operating point: power efficiency, peak to mean envelope power ratio (PMEPR), and pulse compression characteristics. The combination of both PMEPR and power efficiency gives information on the effective average power in the signal useful bandwidth. These criteria allow the evaluation of detection range at the radar system level. Besides, a high PMEPR may reduce the average transmitted power [6] thus reducing the detection range. At the ADC level, the maximum input power determines the maximum SNR after digitization. In Ref. [19], it was shown SNR decreases as the PMEPR increases, so the PMEPR will set to the maximum achievable SNR without clipping. In radar, the pulse compression is used to evaluate the radar detection capabilities [20]. The detection is realized using a matched filter. The characteristics of interest for this study are the spatial resolution and the contrast; these are measured with the characteristics of the main lobe and the side lobes.

These parameters will allow determining the respective performances of any waveforms. PMEPR, power efficiency, and pulse compression will allow determining the detection capabilities for each waveform. Others could be used to get a more accurate picture of the performances. Nonetheless, these criteria are sufficient for a first performance evaluation.

3.2. Simulated processes

For the study the data will be filtered to simulate a 1 GHz bandwidth to match the ADC's Nyquist band used for the experiment. The quantization process and the Nyquist band chosen for simulations are the same as the equipment employed for the experiment the Neptune VXS II digitizer [15]. The encoded value on n bits, $n \in [2, 24]$, is floored to the nearest signed integer. Thus the quantized values range from $[-2^{n-1}; 2^{n-1}-1]$. The model adopted is perfect quantization.

The minimum number of useful bits required to reach near nominal theoretical values with respect to PMEPR, power efficiency, and pulse compression performances will be assessed in order to evaluate the ADC characteristics required to maximize the radar system detection capabilities.

The simulation process also matches the quantization schemes adopted for the experimental radar system which is presented in the following section.

3.3. Design approach for an unbiased experimental study

In order to unbiasedly compare different waveforms, it is essential that waveform-independent criteria are used. Further, to evaluate the performances without bias, the simulated processes and the experimental test bench should be identical. The maximum detection range and pulse compression in range profile can be used as a first step to evaluate radar waveform performances.

To compare the different waveforms, it is not sufficient to simply examine simulation results; and thus this comparison should be experimentally validated. It is therefore necessary to develop a software defined radar prototype that can test the waveforms under study without any bias. The novel approach is to compare the studied waveforms on the same platform to remove any bias. In this paper, simulations and measurements are designed to provide the basis for an unbiased study of the radar waveforms.

It should be noted that the radar prototype should be designed prior to the simulations, this way the characteristics of the prototype can then be fed to the simulator for a subsequent and direct comparison between simulated and experimental results.

3.4. Experimental design

3.4.1. Design of RF system

A few constraints were established for the test bench design. The first step was to optimize the instantaneous bandwidth to maximize the radar spatial resolution. To perform as well as state of the art radar prototypes [4, 19, 21], the bandwidth should be greater than 500 MHz. The radar should support any type of waveform with no changes to the RF frontend. These two requirements ensure an unbiased study of various waveforms on the same prototype. Also a reference channel is implemented to compensate for some of the circuit transfer function. This constraint is a special feature that is not normally implemented in operational radar systems but does allow refreshing the matched filter dynamically to compensate for any fluctuations in transfer function especially with power amplifiers.

Due to spatial constraints on the experimental grounds, a maximum of 50 m in slant range is achievable. Consequently, the architecture must be bi-static and emit in continuous wave to allow for pulse compression gain greater than 20 dB.

Two architectures are proposed as candidates for the implementation: frequency-interleaved and parallel. The frequency-interleaved architecture is inspired from the prototype in Ref. [19]. It is investigated because it reduces the number of components and the number of ADC channels. The parallel architecture is derived from the frequency interleaved architecture. Although more components are required, it has a potential for more versatile usage.

3.5. Parallel architecture

A synoptic of the parallel architecture is shown in Figures 1.a, 1.b, and 1.c. The signal is directly synthesized in intermediate frequencies (IF) ranging from 1 to 2 GHz and a low pass filter removes the mirror image. The IF signal is up-converted in radio frequencies (RF) ranging from 9.9 to 10.9 GHz by FLO1 = 8.9 GHz, and a band pass filter removes the mirror image. For short-range applications, the signal can be amplified by a low noise amplifier; and for longer ranges, a power amplifier can be used. At the output of the amplifier stage, a 20dB directional coupler splits the signal: the coupled output feeds the signal to the reference channel and the direct path is connected to the transmitter antenna feed. The backscattered signal is received by the second antenna which is connected to the test channel. The received signal travels through a low noise amplifier and a band pass filter removes the mirror image before down-conversion by FLO1 = 8.9 GHz. The signal in the reference channel is attenuated and down-

http://dx.doi.org/10.5772/60804

converted by FLO2 = FLO1 = 8.9 GHz. In both the reference and the test channels, the signals are band pass filtered to avoid aliasing and are then amplified prior to digitization.

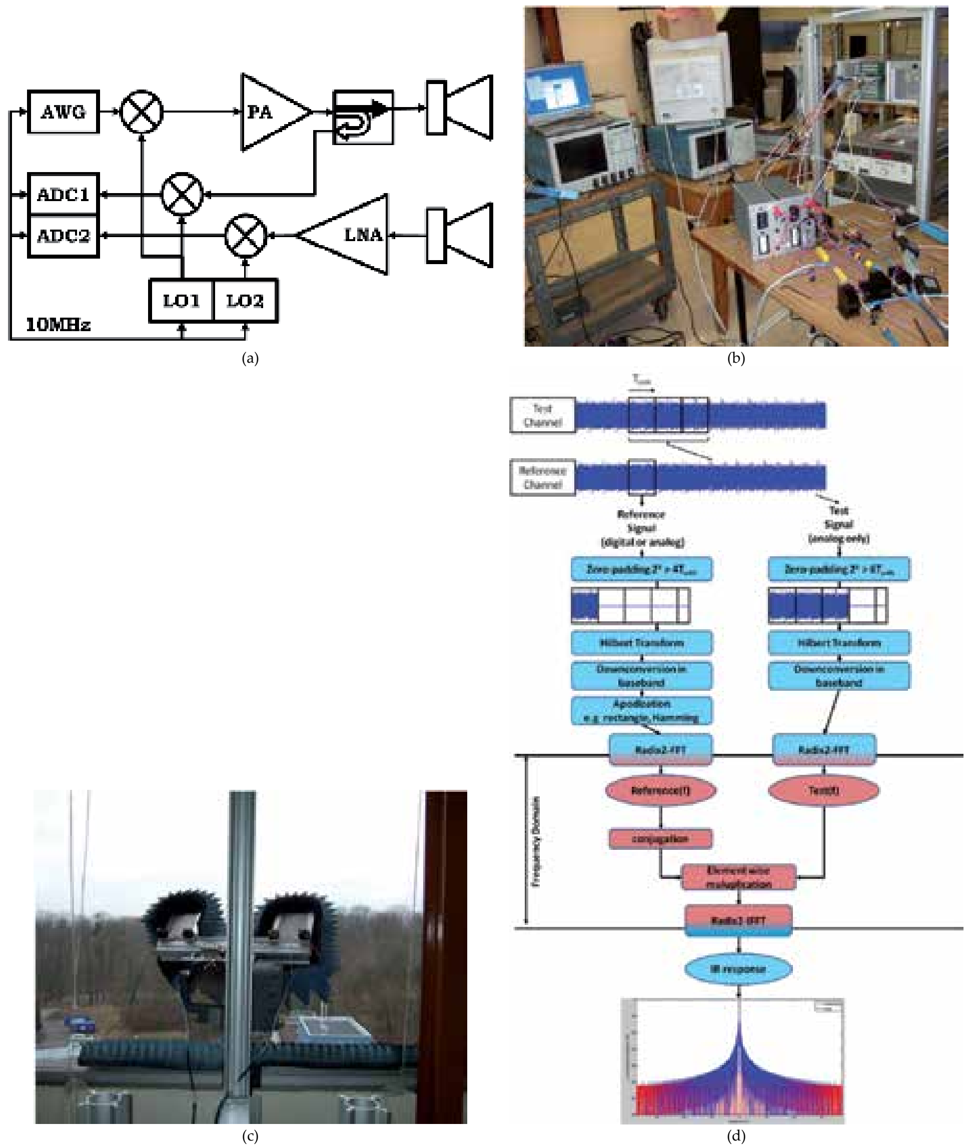


Figure 1. (a): Schematic of parallel architecture. (b): Experimental test bench system overview. (c): Lab experimental test bench set-up. (d): Pulse compression algorithm radix-2 FFT for parallel architecture

A generic algorithm (Figure 1.d) was devised according to the architectures' characteristics, and with the objective to compare waveforms. The algorithm is implemented to process any kind of waveforms. This allows comparing two distinct signals on waveform-independent criteria. Radar systems use pulse compression in order to "see" the targets within the antenna beam. Matched filtering was chosen to process the data and the algorithm was modified to reduce the processing power required using radix-2 FFT.

This section presented the performance criteria, the simulation processes, and the radar system for an unbiased comparison of different waveforms. The next section will present the simulation results.

4. Waveform simulations

The radar emits in continuous wave and the waveforms will cover the bandwidths of 1 MHz, 10 MHz, 150 MHz, and 800 MHz, and pulse repetition period (PRP) of 500 ns, 5 μs, 50 μs, 500 μs, and 1 ms. Each bandwidth value will be tested with every PRP values. It cannot be done in one case as 500 ns pulse already produces 2 MHz instantaneous bandwidth, thus the combination 1 MHz with 500 ns is not possible. The IF sampling frequency is 2 GS/s and the IF frequency range is centred on 1.5 GHz, the signal instantaneous bandwidth varies from 1 MHz to 800 MHz.

The studied signals are the Newman Phase Coded [22] Multitones and the linear Chirp. The latter is a reference in the radar community and will be used as reference to evaluate the performances of Multitones.

A multitude of phase codes exist to reduce PMEPR for Multitones such as Reed–Muller with complementary Golay codes, bi-phase codes, Newman phase codes, etc. For radar application, Doppler tolerance is important to detect moving targets and avoid the multiplication of filters to process the data, Newman phase-codes [22] were chosen because they are easy to implement, the PMEPR reduction is sufficient and it is Doppler tolerant. Furthermore this code is compatible with any vector size. Other codes – an overview of coding techniques can be found in [23] – may be more efficient but Newman phase-codes were chosen because they fit the requirements for radar applications; the aim is to evaluate the contribution of Multitones for radar, not to optimize the waveform phase code. Also note that Multitones need to respect constraints at the generation and digitization to avoid intermodulation interference.

4.1. Simulated results of the performance criteria

In this section, the simulated results for the PMEPR, the power efficiency, and the pulse compression will be presented. Note that the errors or differences express the variations between quantized with respect to perfect performance criteria.

4.1.1. PMEPR

The effects of quantization on the nominal value of PMEPR are now evaluated through simulations for all bandwidth-time configurations of Chirps and Multitones under test. The Chirp's PMEPR increases along with bandwidth, starting at 3.01dB @ 1MHz and going up to 4.22dB @ 800MHz. The increase in PMEPR for wideband Chirp (800 MHz) is explained by the

http://dx.doi.org/10.5772/60804

filter used to ensure a 1 GHz receiver bandwidth, cutting off the edges of the infinite Chirp spectrum. This effectively increases the Chirp's PMEPR by creating peaks in time domain. The PMEPR for Multitones is in the range 5.44 dB to 5.65 dB which matches the expected PMEPR reduction for Newman phase codes. Comparing both Chirp and Multitones, their difference in PMEPR reduces as bandwidth increases. The difference ranges from 1.5 dB @800 MHz to 2.5 dB @1 MHz. As the signal bandwidth reaches the order of the receiver bandwidth, the difference between PMEPR reduces. Using the radar equation, the maximum detection range for Chirp with respect to Multitones will be up to 15% greater in narrowband and up to 9% greater in wideband. The simulation results show that from 4 bits, the PMEPRs are at most 0.1 dB away from their nominal values which is negligible. Thus with respect to PMEPR, the minimum resolution required is 4 bits.

4.1.2. Power efficiency

The effects of the quantization process on the nominal value of power efficiency are now evaluated through simulations for all bandwidth-time configurations of Chirps and Multitones under test. The power efficiencies of both waveforms increase as the bandwidth-time product increases. The relative error on power efficiencies between both Chirp and Multitones decreases as the bandwidth-time product increases. Multitones have higher power efficiency than Chirp but the error is lower than 2% which is negligible. Thus both waveforms are equivalent regarding power efficiencies.

A minimum of 10 bits is necessary to get within 5% of the nominal power efficiencies for every signal configuration. With lower bit resolution, Chirp is more power efficient than Multitones. So in case of low bandwidth-time product and low bit-resolution, Chirp has a higher efficiency by up to 12%.

4.1.3. Pulse compression

If the bit resolution is not sufficient, the pedestal level of the pulse compression increases, although the characteristics of the main lobe and second side lobes are not affected. In order to dimension the digital radar DA/AD converters in single target scenarios, the highest bandwidth-time product should be set, in order to determine the required number of bits to obtain a pulse compression close to the nominal value. Considering a relative mean error of -40 dB and relative max error of -27.5 dB acceptable, the results showed that Chirp requires 14 bits resolution and Multitones 15 bits resolution to meet the acceptable error level for all signal configurations. Since the test bench only has up to 10 bits resolution, the quantization noise for any waveforms increases by 6 dB for every missing bit. The extra bit required for Multitones is related to PMEPR: The Multitones are hindered compared to constant envelope signals such as Chirp, explaining the need for an extra bit to reach the same relative mean error. Increasing the number of bits further than the minimum requirements reduces the noise on the curve; the distance compression pedestal remains unchanged. When using a measured reference, the noise floor will be raised by 6 dB if the minimum number of bits is not respected. However, the transfer function is corrected since the signal comes from the radar system.

4.2. Simulated system level performances

The average power in the useful bandwidth is determined by combining the results of PMEPR and power efficiency from the simulations at 10 bits for quantization. The difference in average power between Chirp and Multitones is in the range 1.18 dB to 2.55 dB. The difference in average power shows that Chirp will have 7% to 16% higher detection range compared to Multitones. In terms of consumption, the Chirp should be more efficient than Multitone signals at the amplifier and ADC level. Especially if the system has a low bit-resolution and is narrowband, Chirp should be favoured over Multitones. On the difference in average power between both waveforms, the result showed that as the signal bandwidth reached the order of the receiver bandwidth, the gap in power was reduced. Note that the simulations were realized with a constant receiver bandwidth of 1 GHz for all bandwidth configurations. On operational radar systems, the receiver bandwidth should be matched with the signal bandwidth to reduce noise power and avoid interferers to maximize the SNR. Extrapolating from the results at 800 MHz, with a receiver bandwidth matched to the signal bandwidth, the difference in average power would be around 1 dB between Chirp and Multitones, resulting in detection range difference around 7%.

Concerning pulse compression with respect to quantization and saturation, Multitones and Chirp have the same characteristics for main lobe and side lobes. Chirp displays a better contrast than Multitones, but the difference is of the order of a couple of dBs.

The analysis revealed that given 10 bit resolution, any waveform reached their nominal values in terms of PMEPR and power efficiencies. Manufacturers of state of the art converters announce DAC AWG7122C [24] at 12GS/s with 10 bit resolution and ADC Proteus V6 [15] at 5GS/s with 10 bit resolution or Calypso V6 [15] at 3.6 GS/s with 12 bit resolution. This means that direct synthesis of signals up to X band and digitization of signals up to S band and part of C band is possible with nominal values of PMEPR and power efficiencies.

The error on pulse compression depends on the bandwidth-time product. For a set error on compression, Multitones need an extra bit in resolution to reach the set value. Depending on the chosen emission band, requiring an extra bit resolution on state of the art AD/DA converters will either result in increased AD/DA converter consumption or in a reduced sampling frequency.

The simulations were indeed basic using perfect quantization process. The simulations were performed without any noise, jitter, or any complex models. This allowed determining a base for the experimental tests. If the experimental results are not satisfactory, then the simulations will go through more complex modelling to approach realistic conditions. However, simple simulations were chosen to reduce time to experiment and get a feel of the processes at work.

5. Experimental results

In this section, the experimental results extracted from the measurements on the radar system will be analysed and compared to the simulated results. The measurements were done on a

http://dx.doi.org/10.5772/60804

trihedral corner reflector located 27 m away from the radar test bench. The results will be presented as for simulations starting with PMEPR, then power efficiency, and finally pulse compression.

5.1. PMEPR

From Figure 2, the measured PMEPR for Multitones and Chirp are consistent with simulations on the closed-loop DAC-filter-ADC experiment, with a difference between measured and simulated values ranging from -0.19 dB to 0.8 dB. The PMEPR for Multitones is in the range [5dB; 6dB]. As for Chirp, PMEPR increases as the signal bandwidth grows closer to the receiver instantaneous bandwidth. The differences in PMEPR between both waveforms are within the range [1.5dB; 2.5dB].

From simulation results, it was determined that 4 bits were sufficient to reach the nominal value of PMEPR. On this experiment, upgrading the resolution from 8 to 10 bits only affected the result on PMEPR by 0.15 dB, which is negligible. This confirms the hypothesis on bit resolution for PMEPR.

In this experiment, the anti-aliasing filter was wider than the first Nyquist band and some of the frequency contents from the first and third Nyquist band leak into the second Nyquist band, thus the recorded signals can be distorted. Also, the gain is not flat over the full bandwidth. This might have contributed to the PMEPR degradation. However, the simulated and measured results on PMEPR match, and this was not predictable a priori.

5.2. Power efficiency

The measured power efficiency is within 10% of the expected value and its general behaviour is consistent with simulations. Also, the difference between 8 and 10 bits resolutions is at most 0.62%, against 10% in simulation. So, this indicates that changing the DAC resolution from 8 to 10 bits for this experiment has little impact on this feature. This confirms the idea that 8 to 10 bit resolution is sufficient to get near nominal values for power efficiency.

Figure 3 displays the measured spectrum of Chirp and Multitones for 1 MHz and 800 MHz. It illustrates in the frequency domain the unevenness of the gain response of the closed loop DAC-filter-ADC experiment. Some unwanted signals are visible in the narrowband case, which reduces the power efficiency of the narrowband signals, explaining the error. However, these are also present in WB case, but since they are buried in the useful bandwidth, they do not affect power efficiency.

Since we are in closed loop, the unwanted signals come from the test bench. This means that with a radar platform with a receiver bandwidth adapted and a fine tuning to have a clean spectrum, the power efficiencies in narrowband would match the simulated values. Thus, extrapolating from the wideband case on this performance criterion, the measurement results are coherent with expected values, and this was not foreseeable before experimental testing.

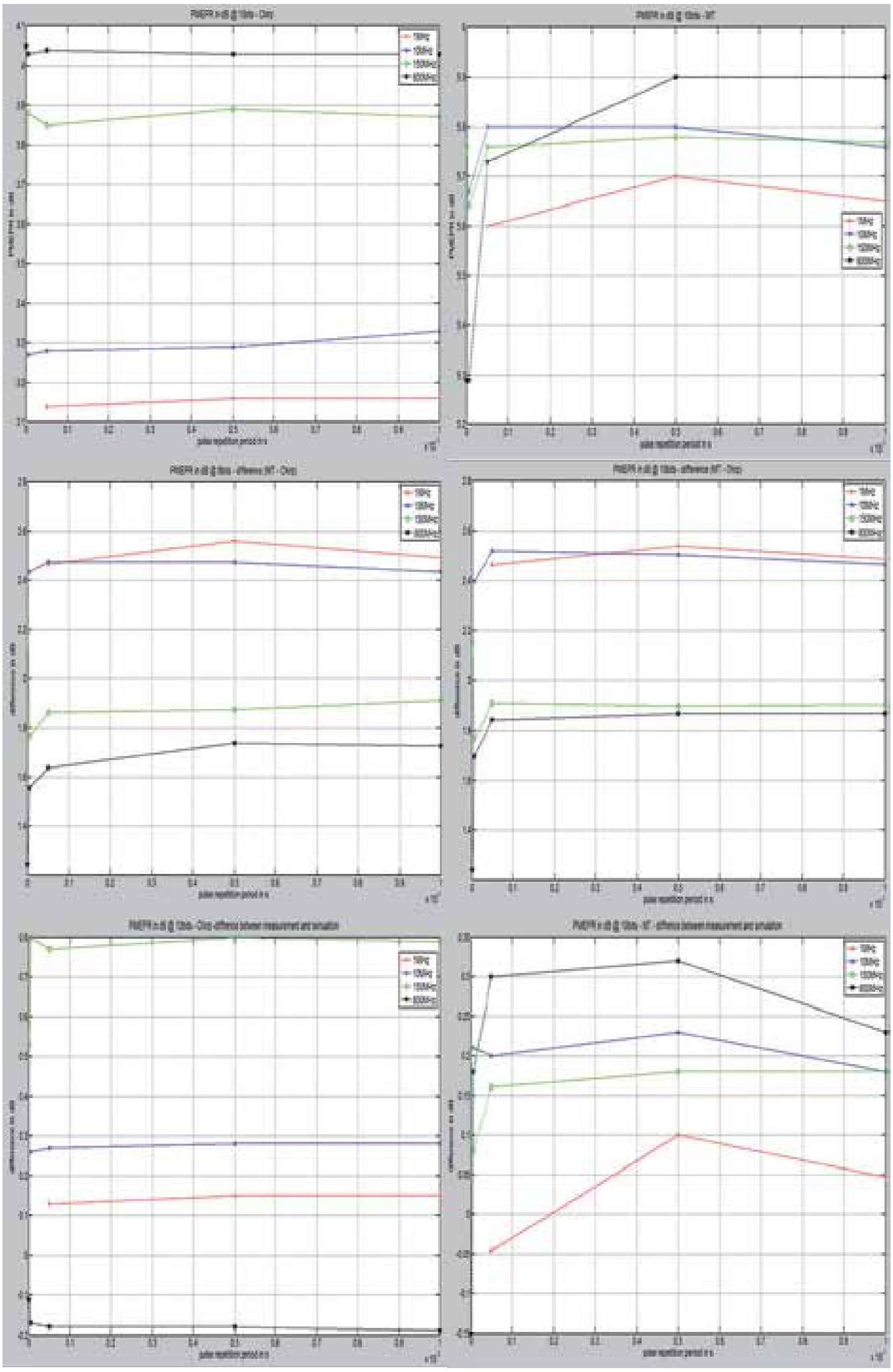


Figure 2. Top: PMEPR @ 10 bits for Chirp and Multitones; middle: difference between Multitones and Chirp @ 8 and 10 bits; bottom: difference between measurement and simulation @ 10 bits for Chirp and Multitones

http://dx.doi.org/10.5772/60804

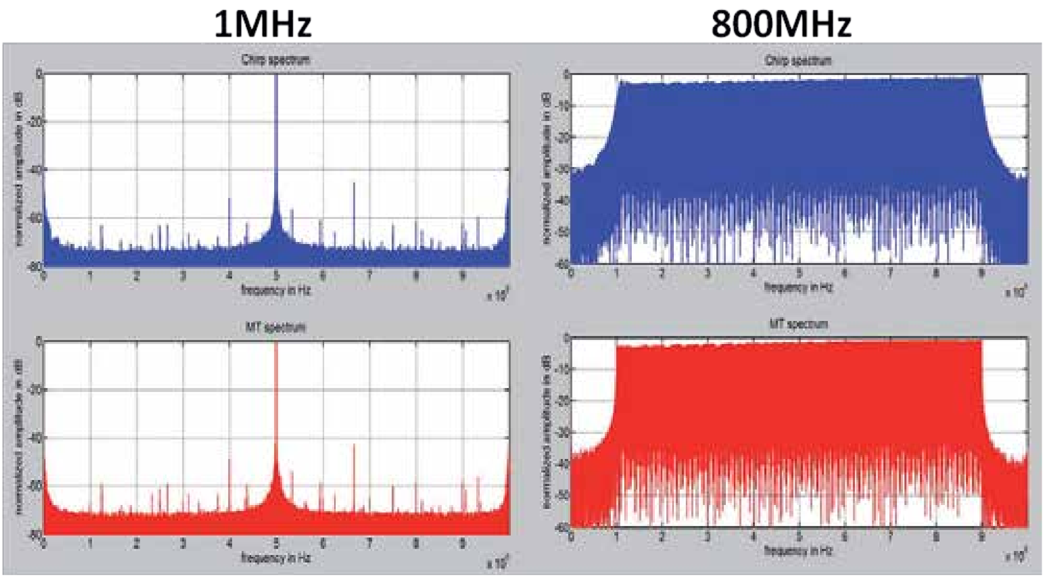


Figure 3. Measured spectrum of Chirp and Multitones at 1 MHz and 800 MHz

5.3. Pulse compression DAC-ADC measurements

The pulse compression was performed with a digital replica of the tested signals. The digital replica is a band-pass sampled version of the generated waveform. This generated waveform is sampled @ 10 GHz and the digital replica @ 2 GHz. The right hand side of the pulse compression presents reflections that are buried when the data is raw, but appear clearer when Hamming windowing is applied. The higher the bandwidth is, the more visible the circuit imperfections are, as shown in the figure. Indeed, problems with standing wave ratios cause uneven second side lobes @ 800 MHz, thus the second side lobes' characteristics will be exploited only for signal bandwidth, from 1 MHz to 150 MHz.

Tables 1 and 2 and Figure 4 show the measured impulse responses: main characteristics and differences/errors between measurements and simulations.

Bandwidth	1 MHz	10 MHz	150 MHz	800 MHz
Main lobe 3 dB width	133m	13.3m	0.9m	0.15/0.225m
Side lobe amplitudes	-13.3dB	-13.2dB	-13.3dB	-19.9dB/-10dB
Side lobe positions	±215m	±21.5m	±1.425m	±0.3m

Table 1. Main characteristics of the pulse compression with respect to bandwidth

In Table 2, the large errors for 3 dB main lobe width and side lobes positions at 800 MHz are caused by sample speck and perturbations induced by standing wave ratios in the circuit.

Otherwise, the other signals from 1 MHz to 150 MHz are within 3.1% of expected values, for 3 dB main lobe width and side lobe positions, and the difference in side lobes amplitudes are lower than 0.3 dB. Also both waveforms are equivalent on pulse compression. These results are really close to the expected values and the matching performances indicate good quality regarding the test bench.

Bandwidth	1 MHz	10 MHz	150 MHz	800 MHz
Main lobe 3 dB width error	<1.9%	<1.8%	<2.3%	<37%
Side lobe amplitudes difference	<0.3dB	<0.3dB	<0.3dB	-7dB / 3dB
Side lobe positions	<0.7%	<1.7%	<3.1%	<67%

Table 2. Relative error on main lobe width and side lobes' characteristics between measurements and simulations

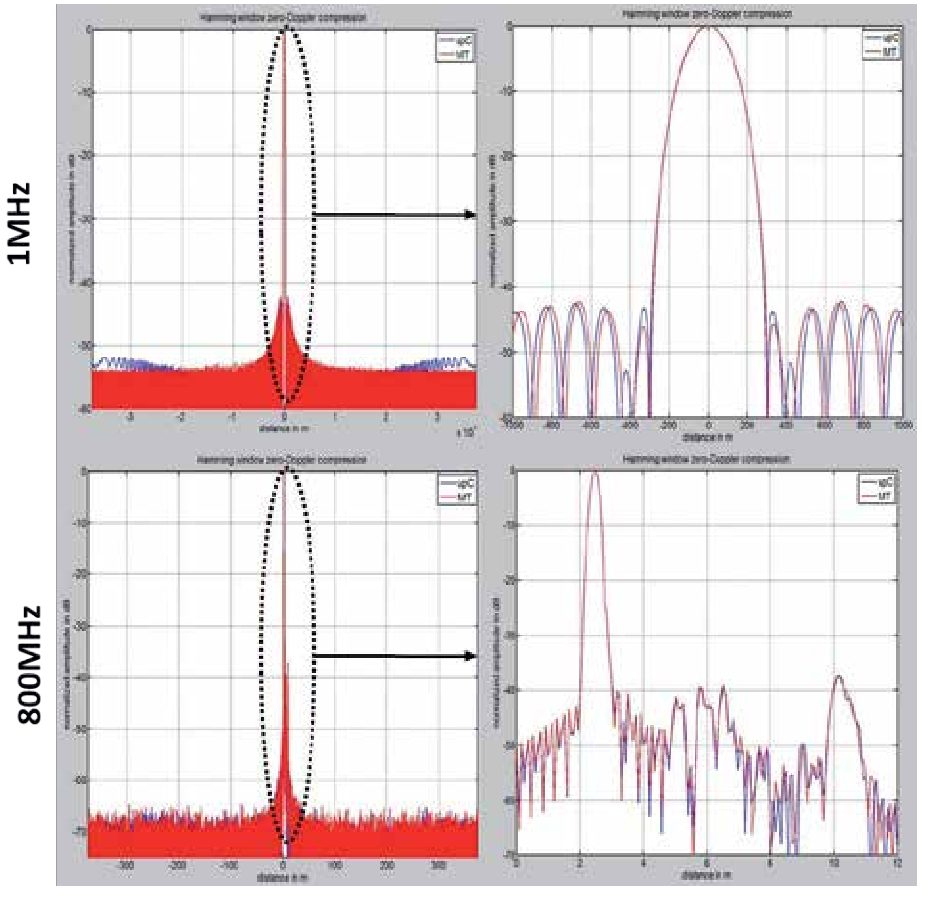


Figure 4. Compression in distance of Chirp and Multitones with (B 1 MHz, PRP 500 us) and (B 800 MHz, PRP 5 us) with Hamming window

http://dx.doi.org/10.5772/60804

The pulse compression displays large errors at 800 MHz caused by reflections in the circuit. From the results obtained for the other signals, reducing the standing wave ratios in the circuit would result in a good match between expected and measured performances at 800 MHz. In other words, imperfections in the circuit can be overlooked for narrowband systems as it only affects the pulse compression by fractions of dBs. As the bandwidth increases, the imperfections cause impairments and are visible in the distance compression. For radar systems, these reflection levels need to be reduced below target detection thresholds to avoid causing false alarms. Also, in the presence of two targets close from one another, one big target and one small, the reflection level may mask the smaller target, thus they should be kept below the desired contrast.

Furthermore, increasing the bandwidth allows locating smaller targets; however, a greater care has to be put to system reflections, as the sources of those reflections appear in the pulse compression. The upside is that with a high bandwidth, the sources of reflections can be more accurately located in the circuit.

Concerning Figure 4, the reflections in the circuit create secondary peaks that change the results on the error. Thus, this formula will not be experimentally validated.

5.4. Synthesis

The closed-loop DAC-filter-ADC measurements were remarkably close to the performance criteria's expected values. This allowed confirming the stability of PMEPR and power efficiency with bit resolution of 8 to 10 bits. This proves that the equipment used to perform the closed-loop experiment closely matches the simulation results obtained using perfect quantization process. These experiments showed that the digitizer technology was mature and that jitter is negligible. Thus simulation for high performance digitizers need not model the jitter. With state of the art digitizers, the expected performances in simulations will be the obtained performances in measurement.

5.4.1. Experiment on static targets: Stability measurements

This experiment used the whole radar system on a trihedral corner reflector located 27 m from the antennas and allowed determining the stability on the peak response of the compression in amplitude and phase over one pulse. The worst-case results are displayed in Table 3 and the evolution of stability over 16 ms. The measurements on stability were obtained using a digital replica and a measured replica. The difference in stability between the two methods is lower than 0.7 dB on the mean and minimum stability with respect to relative error, thus both methods are equivalent. Overall, the relative error in amplitude and phase is about -40dB in mean value and -30dB in minimum value. Both waveforms perform with equivalent performances with respect to stability. Thus, stability depends mostly on hardware rather than waveform.

This measurement of -40dB in stability shows the robustness of the system to clock drift. Note that stability measurements usually remain stable for a set period of time and then degrade with clock drift. Here two hypotheses can be considered: either the set time has not been reached, or the clock is stable. The latter is actually the most plausible, as the sampling clock for the ADC was generated using the DAC, thus when the clock drifts in the DAC, it drifts accordingly in the ADC. Moreover the aperture jitters of the converters are lower than 200 fs, compared to a 500 ns sampling period which is excellent. Finally the mean value found in measurement is of the order of the predicted -42.3dB in RMS quantization noise floor, established for the Neptune VXS2 [15], with a sine wave @ 0 dBFS based on ENOB + losses.

Relative error		**1 MHz**	**10 MHz**	**150 MHz**	**800 MHz**
Raw	Mean	-41.6dB	-40.1dB	-38.6	-39
	Max	-32.11dB	-27dB	-28.8	-27.9

Table 3. Worst case relative error on stability with respect to bandwidth with digital replica

5.4.2. Synthesis

The experiments proved that the measurements matched the results obtained using perfect quantization. This indicates the degree of accuracy of the AD/DA converters (AWG7102 (86) and Neptune VXS 2 (74)) used in this experiment, which had aperture jitters lower than 200 fs. This accuracy is confirmed from the stability measurements, with a mean relative error on peak response subtraction of -40dB. With state of the art converters from 2006, the simple simulation results allowed accurate predictions of the PMEPR, power and efficiency, and compression performances. Future converters will have improved performances compared to that. This means that more complex modelling of jitter effect is unnecessary in that case. The requirements on bit resolution for radar systems could be dimensioned using this simple simulation process, rather than complex modelling.

In radar systems, the receiver bandwidth is matched to the signal bandwidth. This cuts off some of the Chirp spectrum, thus raising its PMEPR, and effectively reduces the gap in average power between both waveforms. Given unbound spectrum and linear properties, the average power difference between Chirp and Multitones is about 2.5 dB. When considering the receiver bandwidth matched to the signal bandwidth, this difference drops to about 1 dB. It is common in a radar system using Chirp to widen the receiver bandwidth to keep good signal properties and avoid spectrum clipping. Multitones could actually allow slightly reducing receiver bandwidth to slightly improve the SNR level, or use the full receiver bandwidth to slightly improve the spatial resolution. In any case, the conclusion of these measurements is that Chirp and Multitones have equivalent performances. The Chirp's maximum detection range is extended by 7% with respect to Multitones' maximum detection range. Also the maximum achievable SNR using the full ADC dynamic range would be about 1 dB higher for Chirp than for Multitones, thus improving a little detection performances and consumption at the ADC.

The outcome of the experiments is that Multitones are neck and neck with Chirp when the receiver bandwidth was equal to the signal bandwidth. The experiment on quantization allowed determining that the converter technology is reliable and accurate. This was demonstrated by the good agreement between measured and simulated results as well as the platform mean stability of -40dB.

http://dx.doi.org/10.5772/60804

6. Conclusions

In order to answer the issue on the contributions of Multitones to UWB software defined radar, an operational reconfigurable ultra wideband radar platform was developed. It supports any kind of waveforms and has 800 MHz instantaneous bandwidth on each ADC channel and 1.6 GHz tuning range. The mean stability is -40dB. The contribution of Multitones to UWB software defined radar is on performances, indeed Multitones displayed equivalent performances compared to Chirp. The detection range is at most 7% higher for Chirp than for Multitones. However Multitones allow more flexibility and thus enable the software defined radar development. Indeed with Multitones, it opens the path toward multifunction, spectrum insertion, sub-band independence, and signal diversity. On the effects of RF components on radar performances, it was demonstrated that simple simulations are sufficient to predict system performances. The AD/DA converters technology is now mature enough for radar applications. And for the performances criteria that were set a minimum of 10 bits resolution are necessary to get nominal performances. Higher resolution improves pedestal error on the impulse response.

7. Perspectives

The use of Multitones is mainly dealing with linearization [14] and performances for radio applications [4-5, 9-10, 12-13, 17-18, 23]. The results mostly come from simulations and were not experimentally validated. When looking into the impact of RF equipment on multicarrier signals, another key component stands out: the transmitter amplifier [5, 14]. The saturation effect will need to be studied to determine the best operating point to maximize radar detection capabilities. Concerning the spectrum insertion, the effect of notched spectrum on performances should be studied.

In the long term, a few technological limitations should be overcome before the implementation of a UWB software defined radar. Research must be pursued in digital architectures, truly adaptive RF components, and antenna arrays and digital beam forming.

Author details

Julien Le Kernec[1*] and Olivier Romain[2]

*Address all correspondence to: Julien.lekernec@nottingham.edu.cn

1 University of Nottingham Ningbo, Ningbo, China

2 University of Cergy Pontoise, Laboratoire ETIS UMR8051, ENSEA, France

References

[1] Terry J., Heiskala J., OFDM Wireless LANs: A Theoretical and Practical Guide, Sams Publishing, 2001 – ISBN: 978-0672321573

[2] Levanon N., Mozeson E., Radar Signals, John Wiley & Sons, 2004, ISBN: 978-0471473787

[3] Rietman R., Linnartz J., De Vries E., Clip Correction in Wireless LAN Receivers, EuWiT - European Wireless Technology, Amsterdam, Netherlands, 2008, p174-177 – ISBN: 978-2-87487-008-8

[4] Garmatyuk D., Kauffman K., Radar and data communication fusion with UWB-OFDM software-defined system, Proc. IEEE International Conference on Ultra-Wideband ICUWB, Vancouver, Canada, 2009, p454-458 – DOI: 10.1109/ICUWB.2009.5288748

[5] Dixon B.J., Pollard R.D., Iezekiel S., A discussion of the effects of amplifier back-off on OFDM, High Frequency Postgraduate Student Colloquium, Leeds, UK, 1999, p14-19 – DOI: 10.1109/HFPSC.1999.809271

[6] Prasad N. N. S. S. R. K., Shameem V., Desai U. B., Merchant S. N., Improvement in target detection performance of pulse coded Doppler radar based on multicarrier modulation with fast fourier transform (FFT)., IEE Proceedings -Radar, Sonar and Navigation, 2004, 151:11-17 – DOI: 10.1049/ip-rsn:20040119

[7] Lellouch G., Tran P., Pribic R., Van Genderen P., OFDM waveforms for frequency agility and opportunities for Doppler processing in radar, RADAR '08 - Proceedings of IEEE Radar Conference, Rome, Italy, 2008, p1-6 - DOI: 10.1109/RADAR.2008.4720798

[8] Tigrek R., De Heiji W.J.A., Van Genderen P., Solving Doppler ambiguity by Doppler sensitive pulse compression using multi-carrier waveform, EuRAD - Proceedings of European Radar Conference, Amsterdam, Netherlands, 2008, p 72-75 – ISBN: 978-2-87487-009-5.

[9] Garmatyuk D., Simulated imaging performance of UWB SAR based on OFDM, Proc. of IEEE International Conference on Ultra-Wideband, Waltham, MA, USA, 2006, p237-242 – DOI: 10.1109/ICU.2006.281556.

[10] Nikookar H., Prasad R., Waveshaping of multicarrier signal for data transmission over wireless channels, IEEE 6th International Conference on Universal Personal Communications Record, San Diego, CA, USA, 1997, p173-177 – DOI: 10.1109/ICUPC.1997.625530

[11] Levanon N., Train of diverse multifrequency radar pulses, Proc. of IEEE Radar Conference, Atlanta, GA, USA, 2001, p93-98.– DOI: 10.1109/NRC.2001.922958

[12] Jonsson B.E., A survey of A/D-Converter performance evolution, Electronics, 17th IEEE International Conference on Circuits, and Systems (ICECS), Athens, Greece, 2010, p766-769 – DOI: 10.1109/ICECS.2010.5724625

[13] Shi K., Zhou Y., Kelleci B., Fischer T. S. E., Karsilayan A., Impacts of narrowband interference on OFDM-UWB receivers: Analysis and Mitigation, IEEE Transactions on Signal Processing, 2007, 55, 3:1118-1128 – DOI: 10.1109/TSP.2006.887153

[14] Harker B., Dobrosavljevic Z., Craney E., Chadwick A., Hayward S., Penney R., Cole M., Belchert R., Dynamic range enhancements for radars and RF systems, 5th EMRS DTC Technical Conference, Edinburgh, 2008.

[15] Tekmicro [Internet] 2015 Available from: http://www.tekmicro.com/products/digitizers.cfm [Accessed: 2014-09-09]

[16] e2v EV10AQ190A [Internet] 2013 Available from: http://www.e2v.com/shared/content/resources/File/documents/broadband-data-converters/1084B.pdf [Accessed 2014-09-09]

[17] Walden R.H., Analog-to-digital converter survey and analysis, IEEE Journal on Selected Areas in Communications, 1999, 17, 14:539-550, – DOI: 10.1109/49.761034.

[18] Saberinia E., Tewfik A., Chang K.-C., Sobelman G., Analog to digital converter resolution of multi-band OFDM and pulsed-OFDM ultra wideband systems, 1st International Symposium on Control, Communications and Signal Processing, 2004, p787-790 – DOI: 10.1109/ISCCSP.2004.1296563

[19] Paichard Y., Microwave Camera for Multi-Dimensional Analysis of the RCS of Time-Varying targets [thesis], Paris: University of Paris 11, 2007

[20] Darricau J., physique et théorie du radar, édition Sodipe, 2005

[21] Jankiraman M., Wessels B.J., Van Genderen P., Pandora multifrequency FMCW/SFCW Radar, Proceedings of the IEEE International Radar Conference, Alexandria, VA, USA, 2000, p35-41–- DOI: 10.1109/RADAR.2000.851801

[22] Levanon N., Waveform analysis and design, Proceedings of IEEE Radar Conference, Rome, Italy, 2008, p1 - DOI: 10.1109/RADAR.2008.4721150

[23] Han S., Lee J., An overview of peak-to-average reduction techniques for multicarrier transmission, Modulation, IEEE Journal on Wireless Communications, 2005, 12, 2:56-65 – DOI:10.1109/MWC.2005.1421929

[24] Tektronix [Internet] 2015 Available from : http://www.tek.com/datasheet/awg7000-arbitrary-waveform-generator [Accessed on 2013-09-09]

Section 2

Application of Digital Signal Processing Concepts towards Image Processing

Application of DSP Concept for Ultrasound Doppler Image Processing System

Baba Tatsuro

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/61164

Abstract

Blood-flow measurements using Doppler ultrasound system are popular in ultrasonic diagnoses. But the blood-flow measurement inside the heart is difficult. There are many reasons behind it. The deep range and fast blood-flow are difficult to measure because of limitation of acoustic velocity. Moreover, strong heart valve signals mix into the blood-flow signal. Against such difficulties, the statistics mathematical model was applied to analyze many clinical data sets. The system identification method based on the mathematical model could realize a new blood-flow measurement system that has ultrasound Doppler information as input and electrocardiogram as output.

Keywords: ARX model, system identification, blood-flow measurement, gap-filling, STFT, Doppler ultrasound system

1. Introduction

It has been more than 70 years since Doppler ultrasound technology was born [1]. In the meantime, in the field of medical blood-flow measurement, various diagnostic methods and diagnostic indices were proposed and had been standardized. By re-focusing on the empirical data from these diagnoses, the relationship between biosignals that many medical doctors had accumulated was revealed. In this chapter, new applications of statistical diagnosis methods and imaging technology are proposed.

Medical Doppler ultrasound systems are commonly used for various diagnostic applications, including examination of cardiac and abdomen. Figure 1 is the example of diagnostic image of a left ventricle inflow. The upper part of Fig. 1 (B-mode image) shows the tomogram of echo, and the lower part of Fig. 1 (D-mode image) shows the spectrum Doppler image. The D-mode image shows the blood-flow velocity at the mitral valve on the B-mode image. In the D-mode image, the horizontal axis is time and the vertical axis is the blood-flow velocity which corresponds to Doppler-shift frequency. The waveform displayed in the lower part of D-mode image is an electrocardiogram (ECG). The amplitude of echo reflected from tissue constructs B-mode image, and the Doppler-shift signal from blood-flow constructs D-mode image.

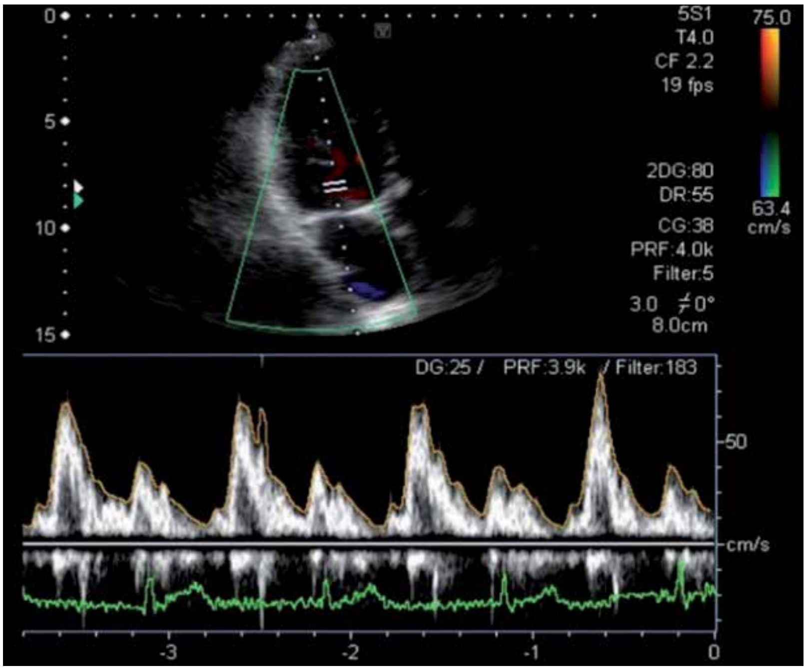


Figure 1. Diagnostic image of Doppler ultrasound system.

2. A new approach to the Doppler diagnostic: ARX model and biosignals

2.1. Accumulated medical database

Currently, time-sharing blood-flow measurement in Doppler ultrasound system appeared. But the time-sharing systems have many problems caused by acoustic velocity range limitation. To address these problems, the mathematical models based on system identification methods were proposed in this chapter. One of the system identification models has ECG as input and has short time Fourier transform (STFT) image parameters as outputs. Based on this model, a new gap-filling system introduced in Section 3 was developed. It can fill the 100 ms gap.

Doppler ultrasound diagnoses of the left ventricle inflow and outflow are very helpful. The diagnostic techniques using the peak velocity waveform (the envelope trace of the Doppler spectra) are the standards of blood-flow measurement. Synchronizing with the systolic phase and diastolic phase that ECG shows, the mimetic diagrams of the outflow from an aortic valve and the inflow from a mitral valve are shown in Fig. 2. The features of these waveforms are measured and evaluated, and they are used as the standards of cardiac disease diagnoses [2]. New diagnostic technology that applies causal relationship between biosignals (here, they are an ECG and a Doppler trace waveform) introduced in Section 4 was developed. Furthermore, many medical doctors make standards of causal relationship between these biosignals over the time of 30 years or more. They are suitable to be applied to the statistical models.

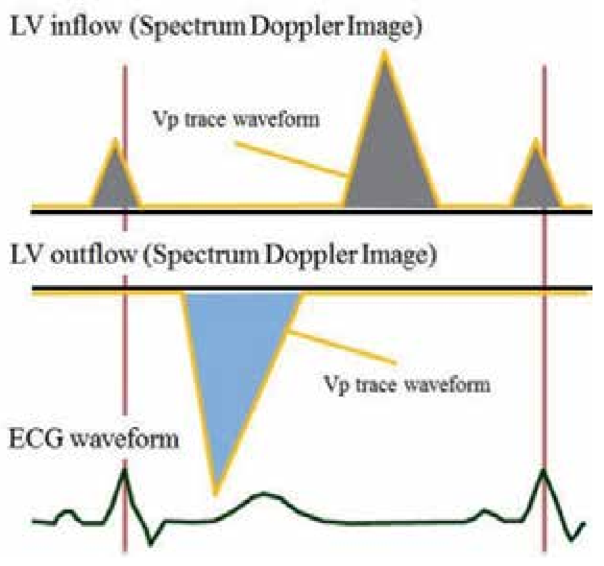


Figure 2. Heartbeat indices and their measurement guidance.

2.2. Mathematical models used in system identification

In order to express the causal relationship between biosignals, the mathematical model that consists of an input, an output, and a black-box is suitable. The black-box model is shown in Fig. 3. Since the input x and the output y assume multi-inputs and multi-outputs (MIMO), they are expressed as vectors. Moreover, in order to take a time response into consideration, two linear partial differentiation equations (1) and (2) express the model using the state variable u.

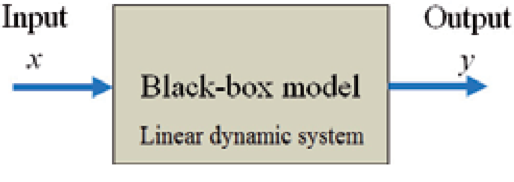


Figure 3. Black-box model and linear dynamic system.

$$\dot{\vec{u}}(t) = \vec{A} \cdot \vec{u}(t) + \vec{B} \cdot \vec{x}(t) \tag{1}$$

$$\vec{y}(t) = \vec{C} \cdot \vec{u}(t) + \vec{D} \cdot \vec{x}(t) \tag{2}$$

ECG, the Doppler waveform and the spectrum Doppler image were used for the causal relationship analyses. In Section 3, the mathematical model based on ECG and Doppler imaging parameters was used. In Section 4, the mathematical model based on ECG and Doppler trace waveform was used. Figure 4 shows the expression of ARX model frequently used in these investigations.

$$\begin{aligned} \vec{y}(t) = {} & a_0 && : \text{Offset(DC) compoment} \\ & + \sum_{i=1}^{p} a_i \cdot \vec{y}(t-i) && : \text{Feedback compoment} \\ & + \sum_{j=1}^{q} b_j \cdot \vec{x}(t-j) && : \text{External input} \\ & + \vec{e}(t) && : \text{Noise(residual) compoment} \end{aligned}$$

Figure 4. Numerical formula of ARX model.

2.3. System identification based on ECG and Doppler waveform

When ECG and the Doppler waveform are applied to ARX model, the system is expressed as the model of Fig. 5. Two coefficient sequences a_i and b_j which determine the system response are calculated using a statistical method. This is the system identification using ARX model.

$$\hat{V}_p(n) = a_0 + \sum_{i=1}^{p} a_i \cdot V_p(n-i) + \sum_{j=1}^{q} b_j \cdot ECG(n-j) + noise(n)$$

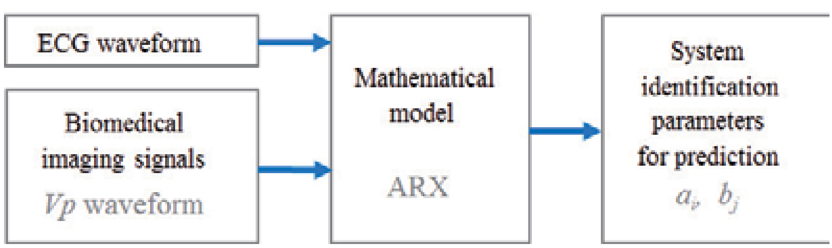


Figure 5. ARX model based on ECG and the Doppler waveform.

The physical meaning of ARX model can be explained with IIR type digital filter. Figure 6 is a digital filter with two inputs: *noise*(n) and *ECG*. Since *noise*(n) is random, there is no time

http://dx.doi.org/10.5772/61164

causal relationship, but *ECG* has FIR ingredients. V_p (Doppler waveform) has IIR ingredients of its feedbacks.

By using system identification, the causal relationship between *ECG* and V_p was summarized in two coefficient sequences. These coefficient sequences are equivalent to the feedback coefficient sequence and feed-forward coefficient sequence of the digital filter. The waveform of ECG resulting from the pulsation of a heart can be explained by how it is related to the flow velocity expressed by the Doppler waveform. It is also the same expression as IIR digital filter.

Thus the response of black-box model can be presumed by system identification using statistical data. Noise rejection technology of the Doppler waveform and gap-filling technology of Doppler image based on system identification are introduced. Moreover, possibilities such as technology that complements the lack part of ECG and automatic diagnostic technology (computer aided diagnosis [CAD]) using statistical data will be introduced to another opportunity.

$$\hat{V}_p(n) = a_0 + \sum_{i=1}^{p} a_i \cdot V_p\,(n-i) + \sum_{j=1}^{q} b_j \cdot ECG(n-j) + noise(n)$$

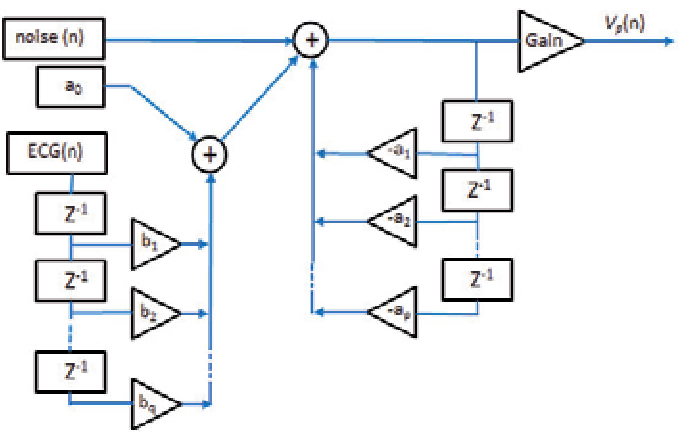


Figure 6. ARX model denoted by digital filter expression.

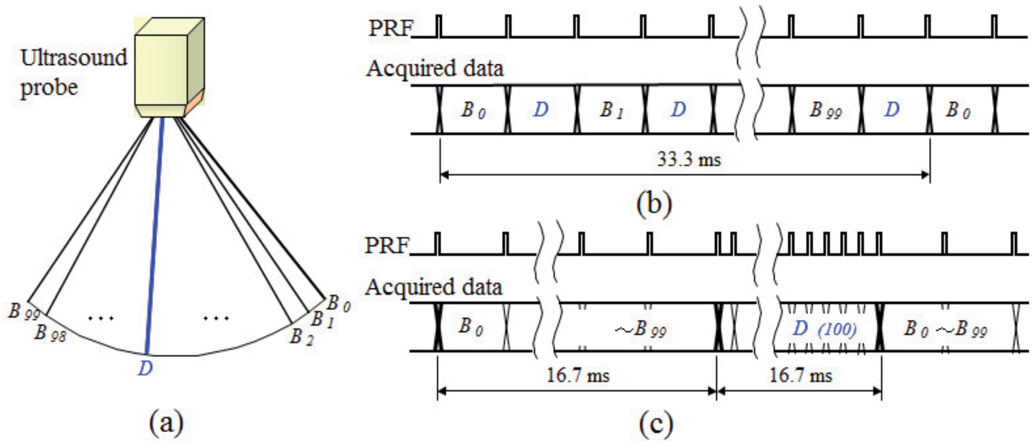
3. Application of system identification to a new gap-filling algorithm

3.1. Limitation of ultrasound scans by acoustic velocity

Doppler velocity range limitation caused by transmit pulse repetition frequency (PRF) is a serious matter with blood-flow measurement. There exists a trade-off between depth range and velocity range.

In order to display the B-mode image and the D-mode image simultaneously in real time, a time-sharing scanning is needed. Normally, PRF of approximately 4 kHz is employed, taking the propagation time of acoustic wave and the attenuation in the living body into account. Use of higher PRF has many advantages. For example, the scanning line density is increased, and as a result B-mode images with higher azimuth resolution can be obtained. In addition, the velocity range of D-mode is expanded. On the other hand, the higher PRF reduces the imaging depth range. Therefore, information concerning deeper regions cannot be obtained. So PRF control is complicated, especially with time-sharing of B-mode scanning and D-mode scanning.

In current Doppler ultrasound systems, it is difficult to optimize both the D-mode image quality and the B-mode image quality simultaneously. A new Doppler gap-filling algorithm based on ARX model was developed, which had ECG as an external input, for detecting the high-speed blood-flow in heart, carotid arteries, etc., and for generating high-quality D-mode images. The conventional gap-filling algorithm of D-mode image suffers from various problems such as presence of noise or artifacts and poor reproducibility in the rapid velocity change.



(a) Ultrasound beam locations in simultaneous scanning, (b) interleave scanning: PRF of B-mode is 6 kHz and PRF of D-mode is 6 kHz, (c) segment scanning: PRF of B-mode is 6 kHz and PRF of D-mode is 12 kHz.

Figure 7. Examples of the interleave scanning and the segment scanning.

Examples of time-sharing transmission/reception of the interleave scanning and the segment scanning are shown in Fig. 7. The B-mode images (100 beams, 6 kHz PRF, approximately 7 cm depth) with a frame rate of 30 Hz are obtained in both Fig. 7(b) and (c). However, the velocity ranges obtained simultaneously in D-mode differ. The velocity range of D-mode is set to 6 kHz in the interleave scanning shown in Fig. 7(b). But the velocity range of D-mode is set to 12 kHz in the segment scanning shown in Fig. 7(c). Since both B-mode and D-mode become discontinuous in the case of segment scanning, the gap-filling algorithm is needed in D-mode signal processing and interpolation processing is needed in D-mode image processing. Moreover, when the gap-filling algorithm is applied, both quality of the image and the audio are degraded. On the other hand, since the PRFs can be set to B-mode and D-mode independently, the Doppler velocity range can be expanded.

3.2. Conventional gap-filling algorithm

Simultaneous real-time display of B-mode and D-mode by the segment scanning has been used for many years. The gap-filling algorithm fills in the gaps of IQ signal (shown in Fig. 8). The gap is filled with the predicted waveform that is generated based on an autoregressive (AR) model. Recently, many improved gap-filling algorithms have been reported [3]. For example, in one method, the gaps are filled in from both time directions; in another method, narrow-band noise is used as a source of signal; and in another method, an autoregressive moving average (ARMA) model is used.

Figure.8 shows a Doppler ultrasound system with the conventional gap-filling algorithm. The received beam is generated in digital beam former (DBF). The output of DBF is sent to the envelope detection processing in which echo intensity is detected. Then the echo intensity signal is sent to the B-mode image processing, and then displayed as the B-mode image. In the spectrum Doppler processing section, the Doppler-shift signal is detected by quadrature detection. Since the detected signal contains low-velocity and high-power components called clutter from tissues (vessel walls etc.), the wall filter rejects the clutters except for blood-flow components. The gap-filling algorithm interpolates the gaps of the D-mode image and the Doppler audio caused by the segment scanning.

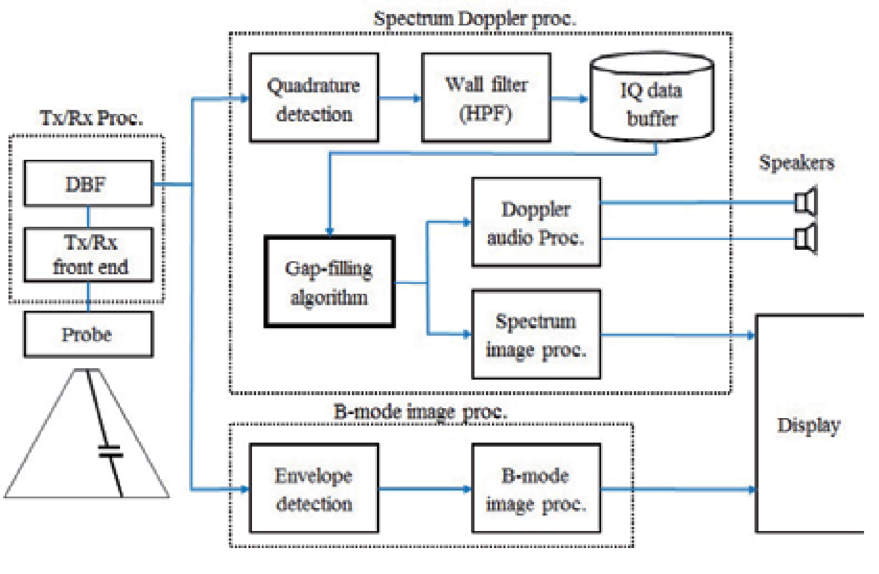


Figure 8. Doppler ultrasound system with conventional gap-filling algorithm.

Figure 9 shows the details of gap-filling algorithm. It shows system identification and linear prediction based on AR model. Figure 10 shows the timing chart of its signal-processing shown in Fig. 9. To estimate the output for Gap(B), band limiting is applied to a white noise source. Prediction is performed in both forward and reverse directions, and blending is performed with $W_1(t)$ and $W_2(t)$ in order to improve continuity. $W_1(t)$ and $W_2(t)$ are weighing functions used for blending actual waveform and predicted waveform. Predicted waveforms of two directions fill in the Gap(B). Because the rapid audio response is important, generally only forward prediction is used. A part of the time sequential data before the gap is used to calculate forward prediction coefficient series $Af(p)$ and bandwidth (BW) of the residual noise. Using $Af(p)$ and BW obtained from the data immediately before Gap(B), the IQ signals are estimated and outputted.

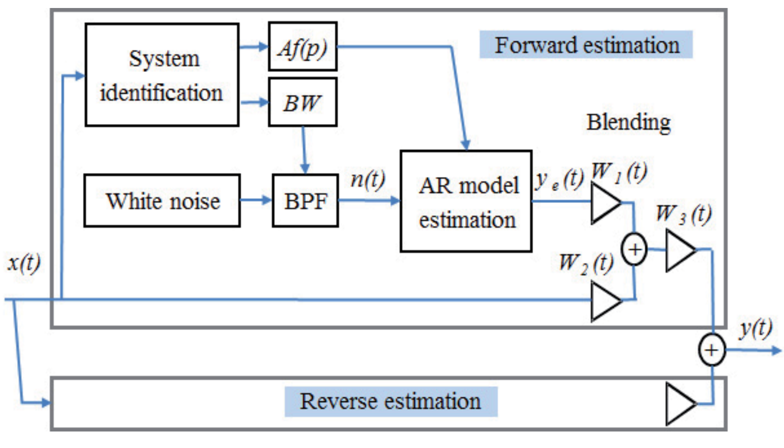


Figure 9. System identification and prediction based on AR model.

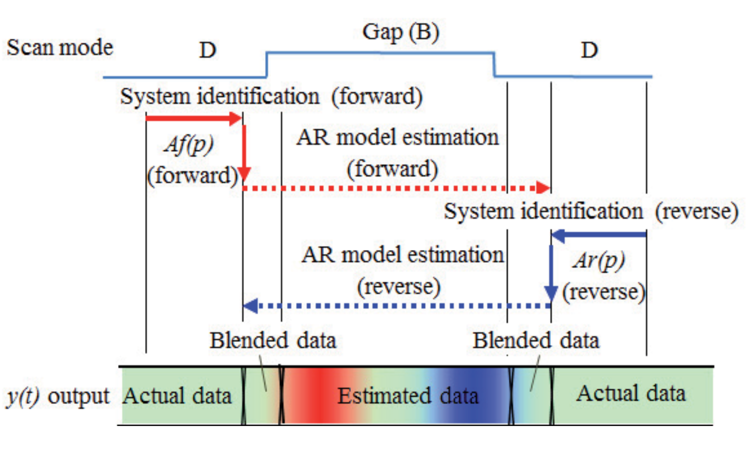
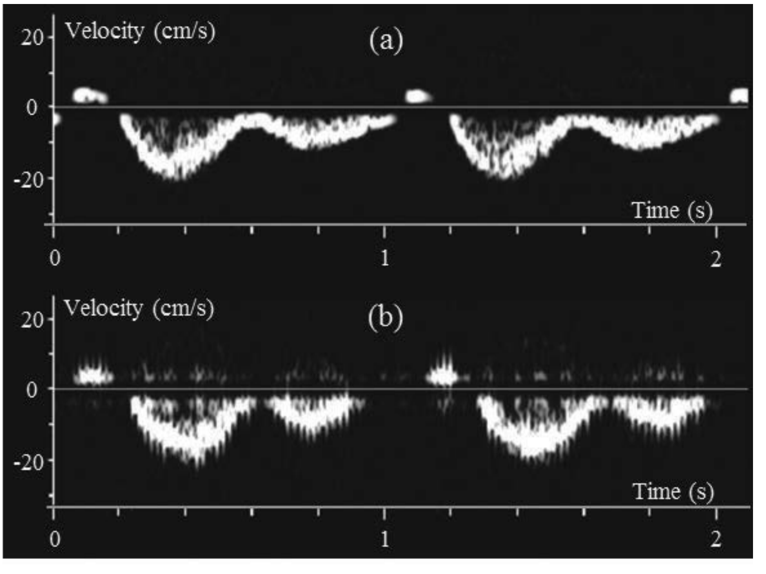


Figure 10. Timing chart of gap-filling algorithm.

3.3. Problems of conventional system

There are two problems specific to Doppler ultrasound systems employing the conventional gap-filling algorithm. The first problem is that artifacts are newly generated in the predicted output although the low-frequency components already have been removed in the wall filter. For these artifacts, not only D-mode image quality but also audio quality is degraded. The second problem is that discontinuity of D-mode images becomes greater when there are sudden changes in the spectrum (rapid changes in blood-flow velocity). In many cases, the predicted D-mode image has horizontal lines and spikelike noises that are observed near the gaps. Figure 11 shows the D-mode images of a portal vein with moderate changes in velocity. Figure 11(a) is the D-mode image obtained by interleave scanning and Fig. 11(b) is the D-mode image obtained by segment scanning with the conventional gap-filling algorithm. Fig. 11(a) is smooth and free of artifacts, while periodic spikelike noise and low-frequency components are observed in Fig. 11(b).



(a) Interleave scanning image and (b) segment scanning image with gap-filling.

Figure 11. Artifacts in D-mode image caused by the conventional gap-filling algorithm.

The conventional gap-filling algorithm based on AR model (or ARMA model) uses the noise and the predicted output itself as feedback inputs. So it tends to generate the waveforms consisting of multiple changeless frequency components. When noise is $x(n)$, output is $y(n)$, and coefficient series obtained in system identification is a_k, the predicted output is expressed by equation (3). In the gap-filling algorithm based on AR model, the noise with Gaussian distribution or the noise with narrow bandwidth is used for $n(n)$. Assuming that σ^2 is distribution width of noise $n(n)$, the estimated spectrum output $P(f)$ is expressed by equation (4) [4]. T is the sampling time. Equation (4) suggests that AR model can only generate multiple spectral peaks in steady states. This is not suitable for estimating rapid changes in velocity. It has been reported that the time to be considered as steady states in D-mode is approximately 10 ms. Accordingly, conventional systems have been designed to limit the Gap(B) to 10 ms or less. It is also known that if the Gap(B) is longer, number of beams and continuity in B-mode image increase and image quality is improved.

$$y(n) = \sum_{k=1}^{P} a_k \cdot y(n-k) + n(n) \tag{3}$$

$$P(f) = \frac{\sigma^2 \cdot T}{\left|1 - \sum_{k=1}^{P} a_k \cdot \exp\left(-j \cdot 2\pi \cdot f \cdot T(k)\right)\right|^2} \tag{4}$$

To improve image quality of B-mode, a longer gap than D-mode is required. But D-mode image quality is markedly degraded when the gap of B-mode becomes long. It is also important to track blood-flow changes due to pulsation for D-mode image quality. But the gap-filling algorithm based on AR model is insufficient. Diagnostic performance will be substantially improved if long gaps are not filled with changeless spectra but filled with changeful spectra.

3.4. A new gap-filling algorithm based on ARX model

A new algorithm that can reduce spectrum artifacts and stabilize rapid changes in velocity in order to overcome problems shown in Section 3.3 was developed. This algorithm uses not IQ signals but Doppler spectrum parameters as input, and is based on ARX model [5]. The outline of a new D-mode image processing is shown in Fig. 12. After quadrature detection IQ signals are generated. IQ signals are processed by the wall filter and STFT sequentially, and D-mode image is generated. The waveforms with 600 ms time lack (left time-domain IQ signals) show the output of the wall filter, which removes low-frequency clutter. The output of STFT shows the momentary spectra (right frequency-domain periodgram). STFT conducts frequency analysis and carries out the time shift image of spectra [6, 7].

A new gap-filling algorithm in Fig. 12, which is based on ARX model, is shown in Figs. 13 and 14. Figure 13 is a block diagram of system identification for the new gap-filling algorithm. Figure 14 shows ECG and the D-mode image of left ventricular inflow. The spectrum shown in the lower side of Fig. 14 shows mean velocity (V_m) and distribution (σ) and the model spectrum near the time of 1.5 s in D-mode image. In system identification based on ARX model, time sequence data (coefficient series) during the cardiac cycle is calculated. First the power spectrum $SP(f,t)$ is calculated, and then V_m, σ, and spectrum total power (TP) is calculated. The V_m, σ, and TP are the characteristic parameters of a single-peak spectrum model this time.

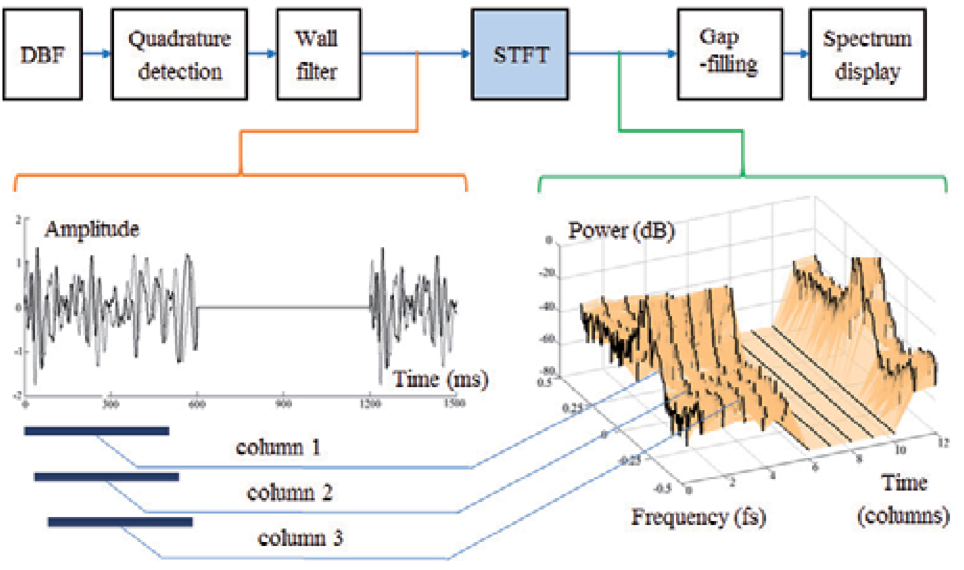


Figure 12. New D-mode image processing.

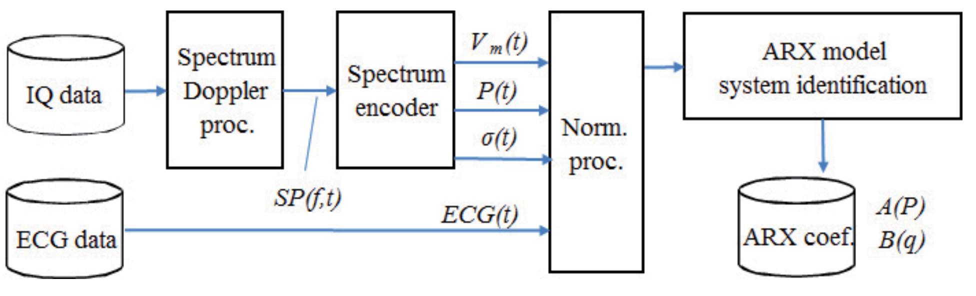


Figure 13. System identification of a new algorithm.

3.5. System design and its confirmation

The formulas of spectrum parameters are shown in equations (5)-(7). For example, the power spectrum is expressed as $P(f)$ and number of FFT points is set to 128. Here, f is frequency, and FFT sampling frequency $f_s(=1/T)$ is 128. The value of f therefore ranges from 0 to 127*f_s/128. Left ventricular inflow shown in Fig. 14 is used as the input. Only positive side components are used in calculation, so value of k ranges from 0 to 95.

$$V_m = \left(\frac{C}{2}\right) \cdot \frac{\sum_{k=0}^{95} f_k \cdot P(f_k)}{\sum_{k=0}^{95} P(f_k)} \tag{5}$$

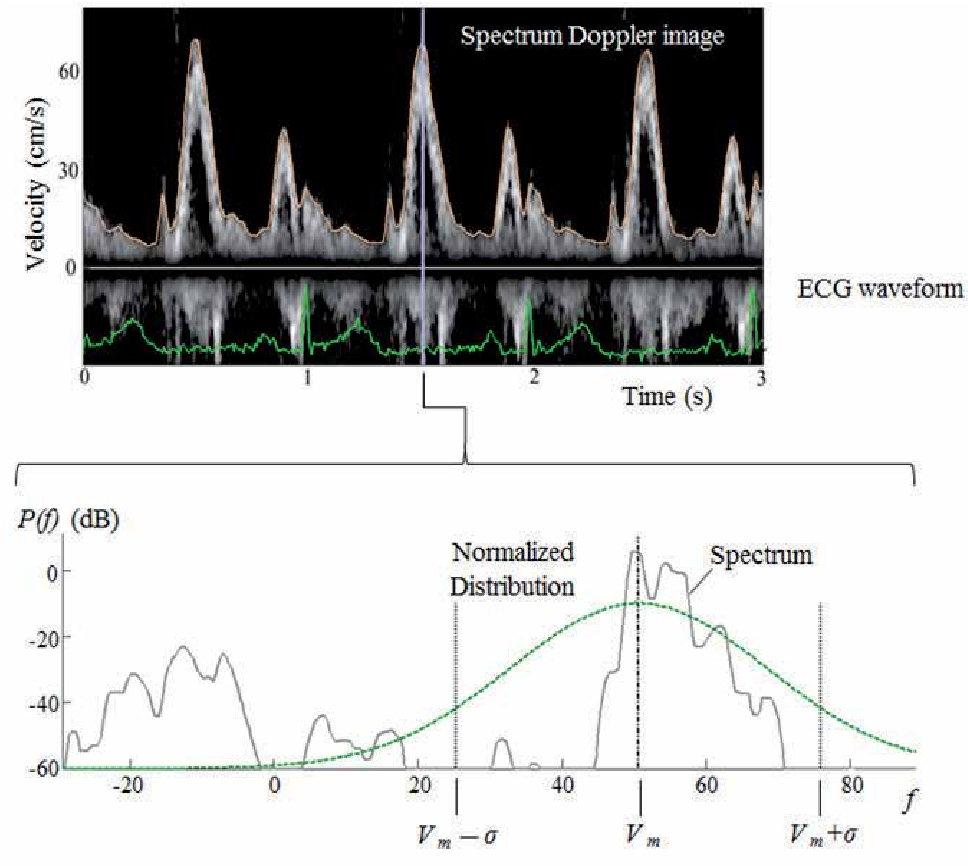


Figure 14. Single-peak spectrum parameter model.

$$\sigma = \sqrt{\frac{\sum_{k=0}^{95}\left(f_k - \frac{2 \cdot V_m}{c}\right)}{96}} \tag{6}$$

$$TP = \sum_{k=0}^{95} P\left(f_k\right) \tag{7}$$

The system that has $V_m(n)$, $\sigma(n)$, and $TP(n)$ waveforms as outputs and ECG waveform as an input was applied to ARX model. These parameters can be obtained by system identification shown in Fig. 13. At the same time these parameters are used for the new gap-filling algorithm. Figure 15 shows the block diagram of the spectra prediction processing based on ARX model. This time *ECG* was newly added as an external input, and only forward prediction was used in the mathematical model.

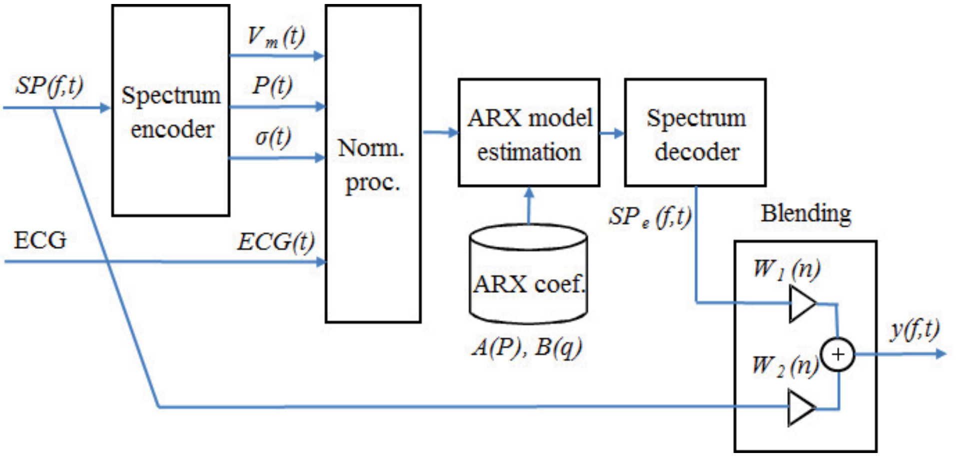


Figure 15. New gap-filling algorithm based on ARX model.

The predicted output of ARX model $\vec{y}(n)$ is shown in equation (8). Here, a_k and b_k are ARX model coefficients. $u(n)$ is the normalized (both time and amplitude) ECG waveform.

$$\vec{y}(n) = \sum_{k=1}^{P} a_k \cdot \vec{y}(n-k) \quad + \sum_{k=1}^{q} b_k \cdot u(n-k) \quad + \quad n(n) \tag{8}$$

$\vec{y}(n)$ is a vector and is expressed by equation (9) as follows:

$$\vec{y}(n) = \begin{bmatrix} V_m(n) \\ \sigma(n) \\ TP(n) \end{bmatrix} \tag{9}$$

The first term and third terms of equation (8) are similar as equation (3). However, the second term of equation (8) means a time-variant system and can generate changeful spectra.

Simulations were applied to confirm ARX model processing and its performance. The data used for simulation was left ventricular inflow, which exhibits rapid changes in velocity. The gap of segment scan was set to 100 ms, which is one order of magnitude larger than the conventional system. The simulation result is shown in Fig. 16. Time 0 to 1 s is a continuous D-mode image (without segment scanning), and time 1 to 2 s is a discontinuous D-mode image (segment scanning). Domains indicated by (a) are actual spectra, and domains indicated by (b) in the figure are spectra estimated by ARX model. Condition of the simulation is as follows: ARX prediction order is 9 to 12, the segment gap is 100 ms, and the blending time is 16.7 ms.

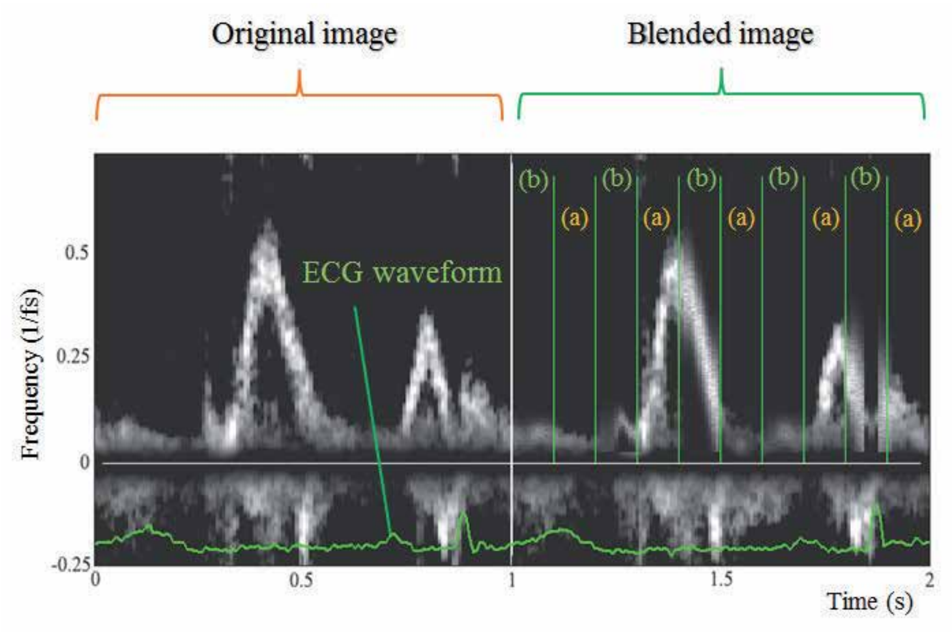


Figure 16. Simulation of predicted spectra based on ARX model.

This result shows that spikelike noises and low-frequency artifacts are reduced compared with the conventional algorithm. Moreover, it is possible to obtain a changeful D-mode image under conditions of rapid changes in velocity and larger segment gaps.

4. Application of system identification to automatic cardiac valve-rejection algorithm

4.1. Problem of blood-flow measurements

The blood-flow diagnoses by Doppler ultrasound system have become popular recently. Peak velocity of blood-flow (STFT envelope waveform) is traced automatically in this system. But valve signals are mixed with the blood-flow signals in the heart. So automatic blood-flow measurements are not correctly recorded. To solve this problem, the mathematical model that has ECG as an input and has Doppler waveform as an output was applied. Using system identification method, a new valve-rejection algorithm was developed [8, 9].

Figure 17 is a Doppler ultrasound diagnostic image for left ventricular outflow. Doppler ultrasound system traces V_p (peak velocity of Doppler waveform) automatically. V_p is superimposed as a bright yellow line. In cardiac blood-flow measurements, cardiac wall noises with strong and low-velocity ingredient generate low-velocity artifacts. The valve noises with strong and high-velocity ingredient generate spikelike trace errors. Although cardiac wall

http://dx.doi.org/10.5772/61164

noises seldom influence V_p waveform, valve noises have unneglective influence on automatic tracing. Valve noises usually mixed in left ventricular outflow in Fig. 17. Thus, users must compensate V_p waveform manually based on the Doppler spectrum image.

In Fig. 17, ECG (green line) is displayed with V_p waveform simultaneously. Also the R-triggers (sharp peaks of ECG) are at 1.48 s and 0.65 s. The systolic phase from the R-trigger of ECG is about 300 ms. Conventional manual trace (cyan line) is displayed between 1.5 s and 1.1 s. Medical doctors estimate left ventricular outflow based on manual traces.

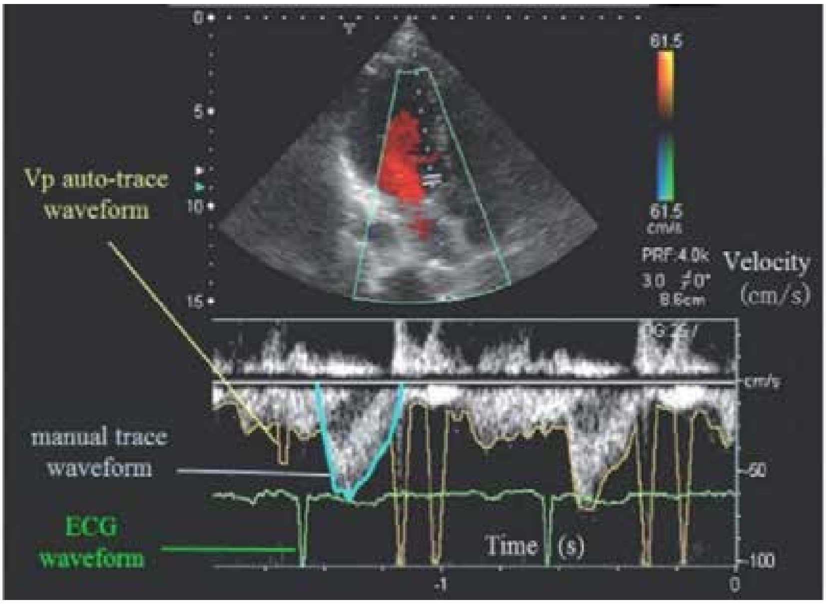


Figure 17. Example of left ventricular outflow.

4.2. Blood-flow measurement

Figure 18 shows a Doppler ultrasound system and automatic blood-flow measurement system. DBF generates echo beams from transmitted and received ultrasound signals. V_p waveform is the spectra envelope of D-mode image, and it is automatically traced and superimposed. B-mode image and D-mode image, and V_p and *ECG* are simultaneously displayed in the same screen. Various blood-flow measurements that combine V_p and *ECG* are known in clinical applications.

4.3. System identification of left ventricle outflow

Many clinical data sets (V_p, *ECG*, diagnostic indices and image) were acquired from numerous volunteers. Using these data sets, system identification shown in Fig. 19 was investigated. V_p has both valve and blood-flow information, so valve components were removed from V_p manually and the ideal blood-flow waveform (V_i) was generated. There were variations in

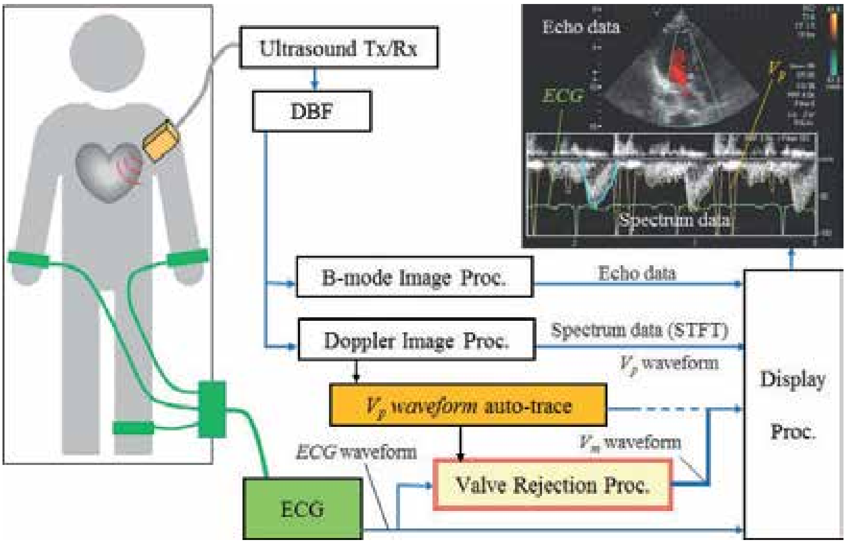


Figure 18. Automatic blood-flow measurement system.

these data sets, such as heartbeat cycles and flow velocity ranges. They were normalized for each data set. Amplitude of V_i was normalized by maximum flow velocity of the systolic phase. Amplitude of *ECG* was normalized by its R-trigger voltage. Heart-beat cycles were normalized and divided by 60. So sampling periods were fixed to 60 (approximately 60 Hz). A low-pass filter with a one-third cutoff (approximately 20 Hz) was applied after normalization, and unnecessary frequency components were rejected. A normalized V_i (NV_i) and a normalized *ECG* (*NECG*) were obtained after filtering process. The system identification block has *NECG* as an input and has NV_i as an output. The coefficient sequences of mathematical model were generated by system identification. A diagnostic image of left ventricular outflow is shown in Fig. 20. V_p and V_i (the aortic valve signal was rejected manually) are shown in Fig. 20(d).

4.4. System design of *Vp* waveform prediction

Figure 21 shows the valve-rejection algorithm using the coefficient sequences of Fig. 19 obtained by system identification. V_p and *ECG* (new data sets) were normalized and filtered. V_e (the prediction waveform of V_p) was generated by system prediction block. Based on differences between *NVP* and V_e, blending times of V_p and V_e were calculated. Based on blending times, the blending weights were changed. Blending weight generator was controlled so that NV_p becomes predominant in systolic phase. After blending process, V_m (the ideal waveform of V_p) was predicted.

Figure 22 shows the waveforms of Fig. 21. V_p and *ECG* are shown in Fig. 22(a) and (b), respectively. A complex heartbeat cycle waveform based on *ECG* is shown in Fig. 22(c). This indicates R-trigger timing, the systolic phase, and the diastolic phase by their amplitude level. V_p and V_e are shown in Fig. 22(d). *Diff* (absolute differences between V_p and V_e) waveform is shown in Fig. 22(e). Weighing functions (V_p-weight and V_e-weight) were generated based on

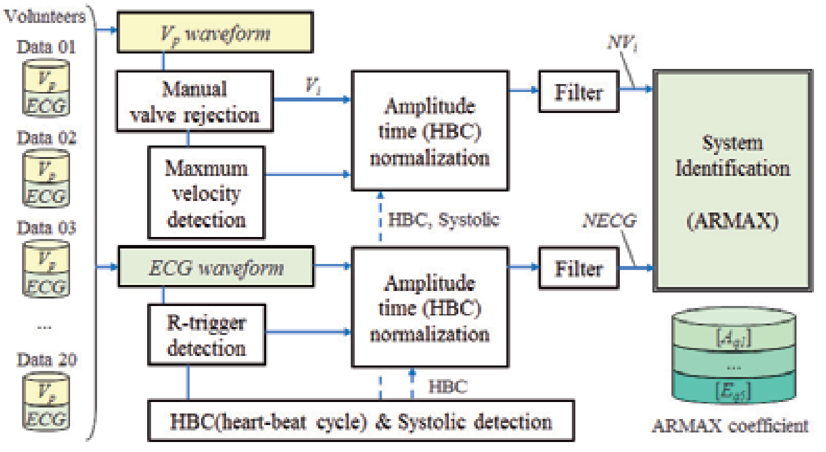
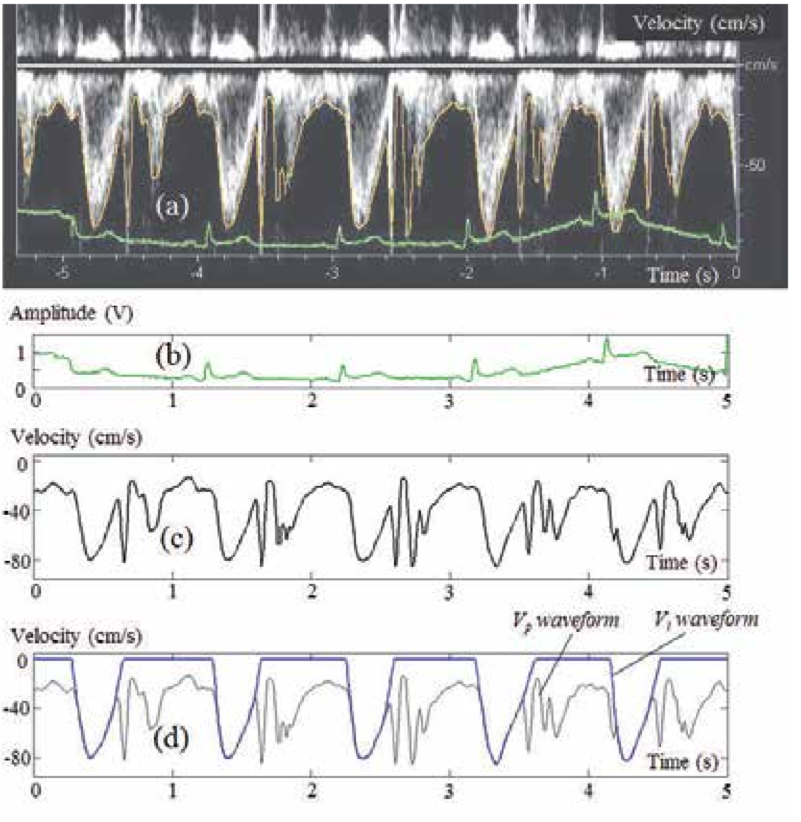


Figure 19. System identification model of Doppler waveform.



(a) D-mode image, (b) *ECG*, (c) V_p, and (d) V_i.

Figure 20. Example of data sets for system identification.

Diff and the complex heartbeat cycle waveform. V_p-weight and V_e-weight, and V_m are shown in Fig. 22(f).

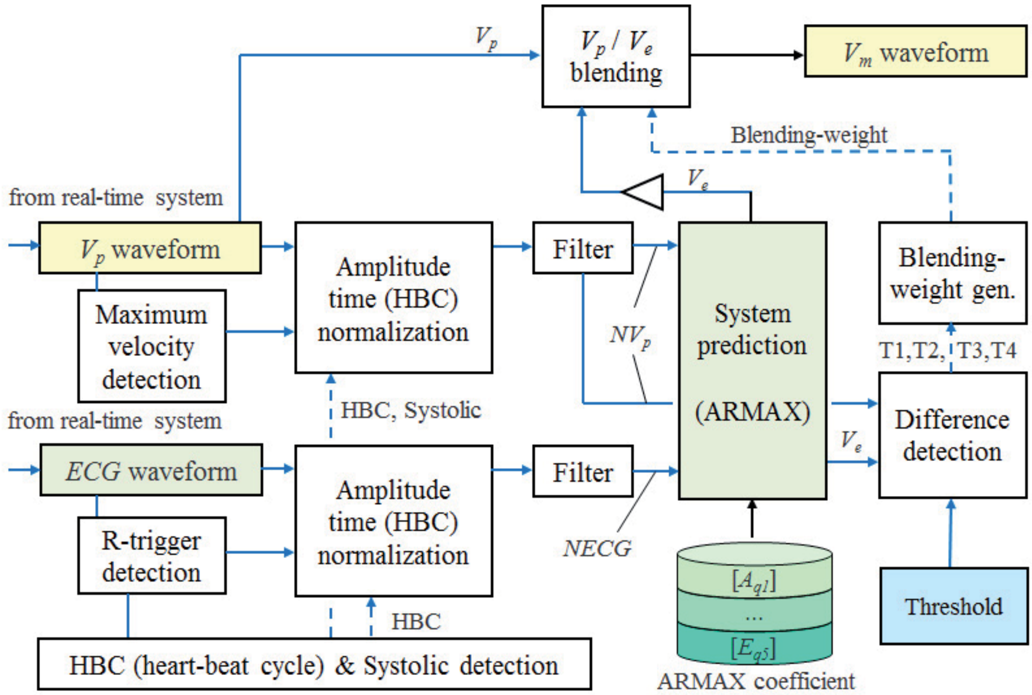


Figure 21. Ideal V_m waveform prediction block diagram.

4.5. Mathematical models

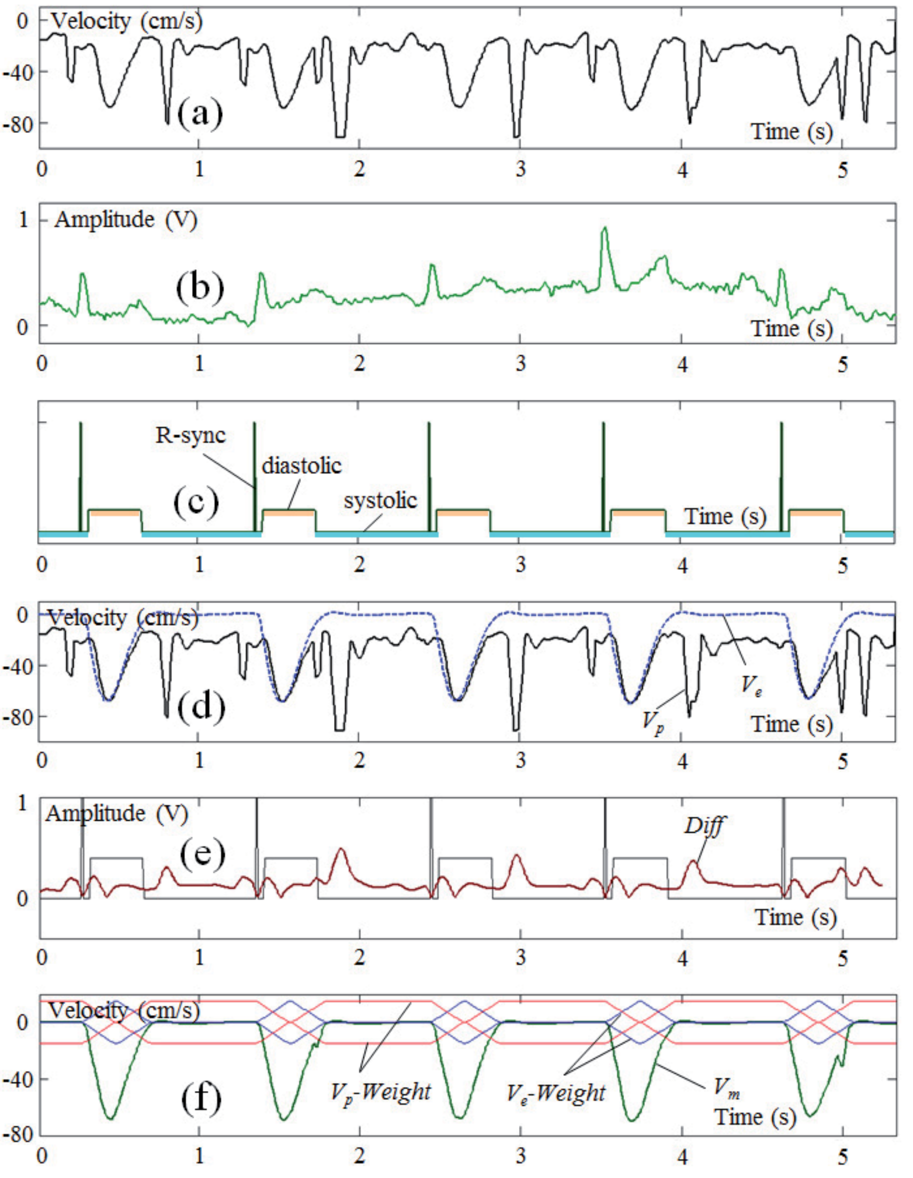
Parametric models such as ARMAX were used as mathematical models [10]. A parametric model is shown in formula (10).

$$P(z) \cdot y(z) = \frac{Q(z)}{R(z)} \cdot u(z) + \frac{S(z)}{T(z)} \cdot w(z) \tag{10}$$

Here, $y(z)$ is an output, and $u(z)$ and $w(z)$ are an input and white noise, respectively. $P(z)$, $Q(z)$, $R(z)$, $S(z)$, and $T(z)$ are the coefficient sequences. In Fig. 19, V_p and *ECG* correspond to $y(z)$ and $u(z)$, respectively. Several models that had different structures and orders were investigated.

4.6. Input data and mathematical model evaluation

Many clinical data sets of left ventricular outflow using a Doppler ultrasound system were acquired. Because outflow varies with individual differences, combined data from numerous



(a) V_p, (b) *ECG*, (c) complex heartbeat cycle, (d) V_p and V_e, (e) *Diff*, and (f) V_p-Weight, V_e-weight, and V_m.

Figure 22. Signal-processing waveforms of system prediction in Fig. 21.

volunteers was used for evaluation. Combined data has different waveforms, heartbeat cycles, and blood-flow sensitivities. Figure 23 shows combined V_p and *ECG*. Both V_p and *ECG* were sampled at 120 Hz sampling rate. The combined 16 heartbeat waveforms were used for simulations. Left ventricular outflows of volunteers A, B, and C are shown in Fig. 24(a)-(c) , respectively.

Several mathematical models were applied and evaluated. Orders of several models such as ARX model, ARMAX model, output error model (OE model), etc., were optimized, respectively. Next, the model fitnesses were evaluated by root-mean-square (RMS) errors. OE model was chosen for the valve-rejection algorithm because of the smallest RMS error.

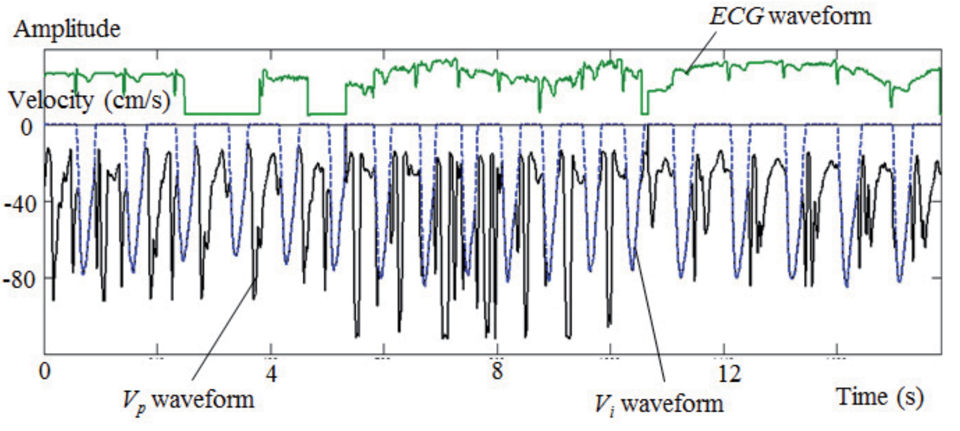
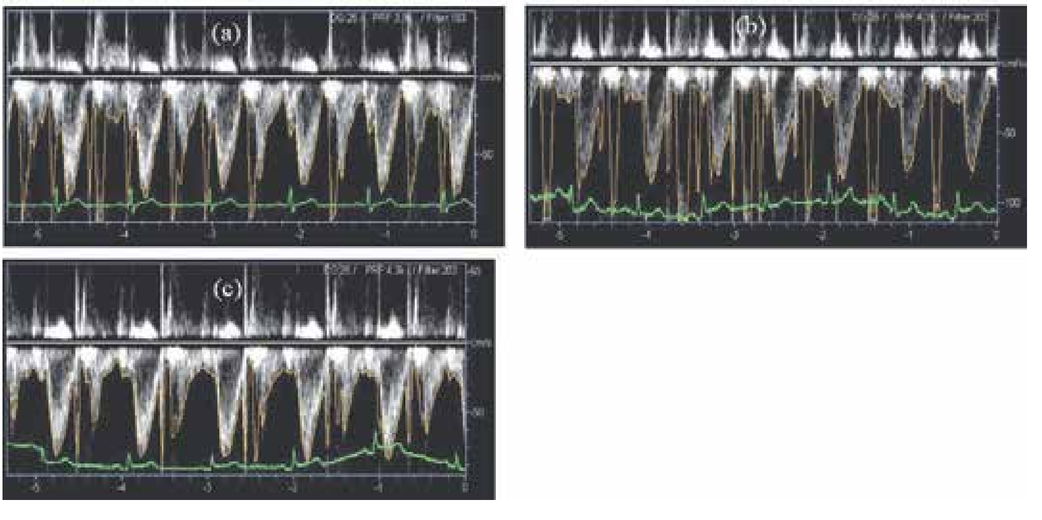


Figure 23. Combined data sets for system identification.



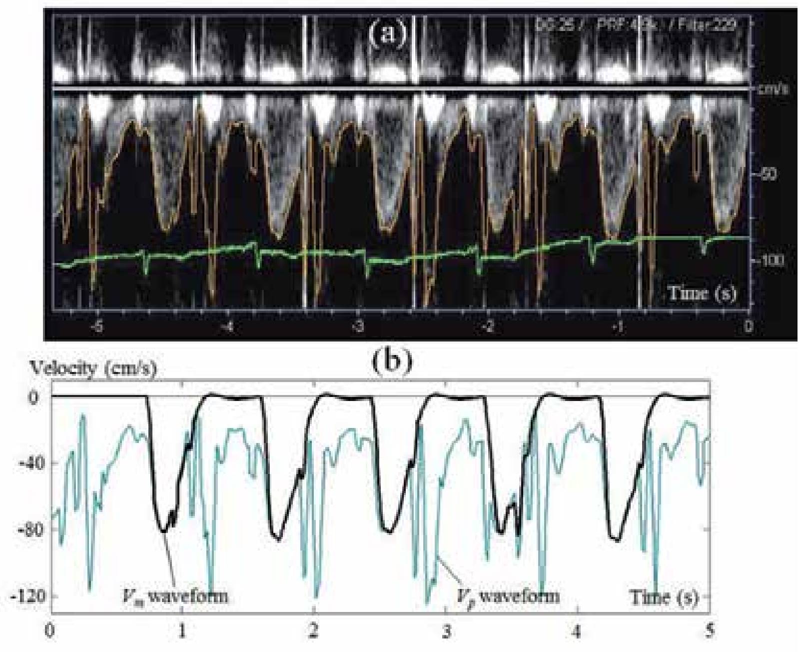
(a) Volunteer A, (b) Volunteer B, and (c) Volunteer C.

Figure 24. Volunteers' blood-flow data.

http://dx.doi.org/10.5772/61164

4.7. Verification

Finally, OE model was chosen as the optimal one for left ventricular outflow. In order to verify its performance of valve-rejection algorithm, the additional data other than volunteer A, B and C was needed. A different volunteer's data (data D) is shown in Fig. 25. Simulation results of the valve-rejection algorithm are shown in Fig. 25(b) using data D. During the first, second, third, and fifth heartbeat cycles, V_m traced ideal outflow. The valve signals were automatically rejected. But the valve signal was not sufficiently rejected during the fourth heartbeat cycle. Although there remains improvement of robustness, high performance of valve rejection was confirmed.



(a) Data sets of volunteer D and (b) V_m waveform.

Figure 25. Verification of valve rejection algorithm.

5. Conclusion

Based on the mathematical model that combined an ECG and biosignals (ultrasound Doppler image parameters, etc.), the system identification method to heart's blood flows was applied. With combination of the image parameter and the ECG, the effectiveness of a new gap-filling algorithm was confirmed. Moreover, with combination of the Doppler blood-flow waveform and the ECG, noises in heart's blood-flow measurement, such as valve regurgitation, were removed, and reliable automatic measurement of left ventricle outflow was realized. System identification using such a statistical method will be an important component for automatic measurement and diagnosis.

Author details

Baba Tatsuro*

Address all correspondence to: EZD03014@nifty.ne.jp

Baba Professional Engineer Office, Japan

References

[1] Satomura, S. Ultrasonic Doppler method for the inspection of cardiac functions. *J. Acoust. Soc. Am.* 1957; 29:1181; DOI: 10.1121/1.1908737.

[2] The terminology and diagnostic criteria committee of the Japan society of ultrasonics in medicine. (2006). Standard measurement of cardiac function indexes. *J. Med. Ultrasonics*, 33, 123-127; DOI: 10.1007/s10396-006-0100-4.

[3] Kristoffersen, K. and Angelsen, B. A. J. (1988). A time-shared ultrasound Doppler measurement and 2-D imaging system. *IEEE Trans. Biomed. Eng.*, 35, 285-295, DOI: 10.1109/10.1383.

[4] Jensen, J. A. (1996). *Estimation of Blood Velocities using Ultrasound: A Signal Processing Approach*. Cambridge University Press. 0-521-46484.

[5] Adachi, S. (1996). *Advanced Level System Identification for Control by MATLAB*. Tokyo Denki University Press. Tokyo. 978-4501323509.

[6] Mukai, F. et al. (2010). An application of system identification theory to the ultrasound Doppler image processing system. *SICE Annual Conference 2010, Proceedings*, 3190-3193. Print ISBN: 978-1-4244-7642-8. INSPEC Accession Number: 11595464.

[7] Baba, T. and Toshiba, Corp. (2013). Ultrasonic imaging apparatus and a method of generating ultrasonic images. US Patent 8597191B2.

[8] Baba, T. et al. (2013). *Automatic Valve-Rejection Algorithm for Cardiac Doppler Ultrasound Systems*. Hindawi Publishing Corporation. ISRN Biomedical Imaging Volume 2013, Article ID 850303, DOI: 10.1155/2013/850303.

[9] Baba, T. et al. and Toshiba, Corp. (2012). Ultrasonic Doppler diagnostic apparatus and measuring method of diagnostic parameter. US Patent 8303507B2.

http://dx.doi.org/10.5772/61164

[10] Baba, T. et al. and Toshiba, Corp. (2014). Ultrasonic diagnostic apparatus, Doppler measurement apparatus, and Doppler measurement method. US Patent 8821401B2.

Chapter 5

Lossy-to-Lossless Compression of Biomedical Images Based on Image Decomposition

Luís M. O. Matos, António J. R. Neves and Armando J. Pinho

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/60650

Abstract

The use of medical imaging has increased in the last years, especially with magnetic resonance imaging (MRI) and computed tomography (CT). Microarray imaging and images that can be extracted from RNA interference (RNAi) experiments also play an important role for large-scale gene sequence and gene expression analysis, allowing the study of gene function, regulation, and interaction across a large number of genes and even across an entire genome. These types of medical image modalities produce huge amounts of data that, for several reasons, need to be stored or transmitted at the highest possible fidelity between various hospitals, medical organizations, or research units.

In this chapter, we study the performance of several compression methods developed by the authors, as well as of image coding standards, when used to compress medical images (computed radiography, computed tomography, magnetic resonance, and ultrasound), RNAi images, and microarray images. The compression algorithms addressed are based on image decomposition, finite-context modeling, and arithmetic coding. In one of the methods, the input image is split into several bitplanes, and each bitplane is encoded using finite-context models and arithmetic coding. In another approach, the intensity levels of a given image are organized in a binary-tree structure, where each leaf node is associated with an image intensity.

The experimental results presented in this chapter are state of the art regarding the compression of some of these types of images. Moreover, several approaches and preprocessing techniques are presented, giving a good hint about new developments that can be studied further. Also, this chapter intends to be used as a reference for comparison with new compression algorithms that may be developed in the future.

Keywords: Image coding, Lossless coding, Progressive decoding, Biomedical images, Image coding standards, Image decomposition, Arithmetic coding, Finite-context modeling

1. Introduction

Image compression is a very important research field. It is fundamental in many different areas, such as biomedical imaging, consumer electronics, and Internet, among others. The goal of an image compression method is to represent an arbitrary image using the smallest possible number of bits.

The use of medical imaging has increased in the last years, especially with magnetic resonance imaging (MRI) and computed tomography (CT) [1]. These types of medical image modalities produce huge amounts of data that, for several reasons, need to be stored or transmitted with the highest possible fidelity between various hospitals and medical organizations.

Also related to biomedical imaging, microarray images play an important role for large-scale gene sequence and gene expression analysis, allowing the study of gene function, regulation, and interaction across a large number of genes and even across an entire genome [2, 3]. The output of a microarray experiment is a pair of 16-bits-per-pixel images, usually with a very high resolution, sometimes exceeding 13,000×4,000 pixels. Consequently, over 200 MB can be required to store a single microarray image. Due to the development of these digital imaging technologies, some concerns appeared regarding efficient ways of storing and transmitting the images. In order to overcome these problems, sophisticated compression methods are required.

Another type of images that are addressed in this chapter are those that can be extracted from RNA interference (RNAi) experiments. Those experiments involve marking cells with various fluorescent dyes to capture three components of interest, namely, DNA, actin, and PH3 channels [4].

Typically, lossless compression algorithms are recommended for dealing with these types of images. In fact, lossless methods are generally required in applications where cost, legal issues, and value play a decisive role, such as in medical imaging or in image archiving [5]. On one hand, the use of lossless algorithms avoids problems of losing diagnostic information vital to identify life-threatening illnesses in the early stages. On the other hand, if a given image is lossless compressed, it is possible to recompress it in the future, using a more efficient algorithm and without losing any information.

Usually, the performance of image coding standards, such as JPEG, JBIG, JPEG-LS, and JPEG2000, falls short when they are applied to certain types of biomedical images. In order to overcome this lower-performance issue, specific compression algorithms are essential. In the class of lossless compression algorithms, there is a specific type that has progressive decoding capabilities. This type of compression is usually designated as "lossy-to-lossless compression." These algorithms are very flexible, because they allow stopping decoding at a certain point, according to the available resources or user requirements. However, the original image can also be obtained, without any loss, if the decoding process goes until the end.

In this chapter, we study the performance of several specific compression methods developed by the authors, as well as of image coding standards, when applied to medical images (computed radiography, computed tomography, magnetic resonance, and ultrasound), RNAi images, and microarray images. The addressed compression algorithms are based on image decomposition, finite-context models, and arithmetic coding. We can divide these algorithms into two categories. In the first one, the input image is split into several bitplanes and each bitplane is encoded individually. This approach is combined with several preprocessing techniques in order to improve the compression efficiency. In the second approach, the intensity levels of a given image are organized in a binary-tree structure, where each leaf node is associated with an image intensity.

We start by presenting the compression algorithms, namely, bitplane decomposition and binary-tree decomposition. We include a brief explanation about finite-context models and arithmetic coding, the entropy coding block used in both approaches. Then, we present a set of experiments that have been performed using the compression algorithms described in Sect. 4, as well as experimental results using the most important image coding standards (e.g., PNG, JBIG, JPEG-LS, and JPEG2000). At the end, we draw some conclusions.

This chapter intends to be a reference for comparison of new compression algorithms that may be developed in the future, for two main reasons. On one hand, the experimental results presented in Sect. 5.2 are state of the art regarding the compression of some of these types of images. On the other hand, several approaches and preprocessing techniques are presented, giving hints about new developments that can be done.

2. Compression methods

The compression algorithms that we address here can be classified into two different categories. One based on bitplane decomposition and the other one based on binary-tree decomposition. For the first category, we use a sophisticated bitplane decomposition approach that was successfully developed for the compression of microarray images [6]. Furthermore, we include some preprocessing techniques, such as segmentation and histogram reduction, in order to improve the compression results. An alternative bitplane decomposition approach based on *bit modeling by pixel value estimates* is also considered.

The other approach, based on binary-tree decomposition, was developed with success for the compression of medical images [7] and microarray images [8]. In this decomposition approach, the intensity levels of a given image are organized in a binary-tree structure, where each leaf node is associated with an image intensity.

In the following subsections, we address these two decomposition approaches and give a brief explanation regarding finite-context models and arithmetic coding. The implementation of the compression algorithms can be found in [9].

2.1. Finite-context models

Markov modeling is widely used in several fields, including image [7, 8, 10, 11, 12] and DNA [13, 14, 15] compression. Finite-context models rely on the Markov property, since an order- k

finite-context model gives a probability distribution for the next symbol, in a sequence of symbols from an alphabet A, taking into account a recent past of depth k. Hence, the finite-context model assigns probability estimates for each symbol, regarding the next outcome, according to a conditioning context, $c_{k,n}$, computed over a finite and fixed number $k>0$ of past outcomes $c_{k,n}=x_{n-k+1\ldots n}=x_{n-k+1}\ldots x_n$ (order- k finite-context model [16, 17, 18] with $|\mathrm{A}|^k$ states). In the example illustrated in Fig. 1, where A0,1 and then $|A|=2$, an order- k model implies having 2^k conditioning states. In this case, $k=5$, so the number of conditioning states is $2^5=32$.

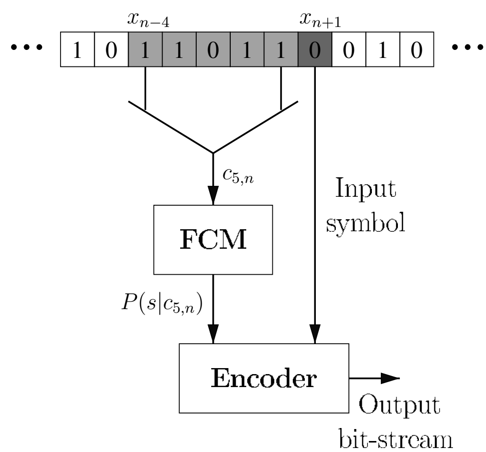


Figure 1. Finite-context model: the probability of the next outcome x_{n+1} is conditioned by the k last outcomes. In this example, $k=5$

The probability estimates $P(x_{n+1}\,|\,x_{n-k+1\ldots n})$ are calculated using symbol counts that are accumulated while processing each pixel of the input image, making them dependent not only on the past k symbols but also on n. The estimator used is

$$P\left(s\,|\,x_{n-k+1\ldots n}\right)=\frac{C\left(s\,|\,x_{n-k+1\ldots n}\right)+\alpha}{C\left(x_{n-k+1\ldots n}\right)+|\mathcal{A}|\alpha}, \tag{1}$$

where $C(s\,|\,x_{n-k+1\ldots n})$ represents the number of times that, in the past, symbol s was found having $c_{k,n}=x_{n-k+1\ldots n}$ as the conditioning context and where

$$C\left(x_{n-k+1\ldots n}\right)=\sum_{a\in\mathcal{A}}C\left(a\,|\,x_{n-k+1\ldots n}\right) \tag{2}$$

is the total number of events that has occurred so far in association with context $c_{k,n}$. Parameter α allows balancing between the maximum likelihood estimator and a uniform distribution

http://dx.doi.org/10.5772/60650

(when the total number of events, n, is large, it behaves as a maximum likelihood estimator). For $\alpha = 1$, (1) is the well-known Laplace estimator.

Initially, all counters are set to zero, i.e., the symbols are assumed to be equally probable. The counters are updated each time a symbol is encoded. However, it is possible to update the counters according to a specific rule. Since the context templates are causal, the decoder is able to reproduce the same probability estimates without needing additional side information.

Table 1 presents a simple example of how a finite context is typically implemented. In this example, we are dealing with an order-5 finite-context model, which means that the context uses the last five encoded symbols to assign the symbol probabilities. Each row of Table 1 represents a probability model that is used to encode the current symbol, using the last five encoded ones. For example, if the last symbols were "11000," i.e., $c_{5,n} = 11000$, then the model sends the following probabilities to the arithmetic encoder (denoted as "Encoder" in Fig. 1): $P0$ and $P1$.

Context, $c_{5,n}$	$C(0 \mid c_{5,n})$	$C(0 \mid c_{5,n})$	$C(c_{5,n})$
00000	23	41	64
⋮	⋮	⋮	⋮
00110	14	34	48
⋮	⋮	⋮	⋮
01100	25	12	37
⋮	⋮	⋮	⋮
11000	28	41	69
⋮	⋮	⋮	⋮
11111	8	2	10

Table 1. A simple example illustrating how finite-context models are implemented. The rows of the table correspond to probability models at a given instant n. In this example, the particular model that is chosen for encoding a symbol depends on the last five encoded symbols (order-5 context)

One important aspect that must be considered is the size of the context. For an order- k model and $\mid$A2 (binary alphabet), the table has 2^k entries and, therefore, its size grows exponentially with k. Using a deeper context, we might achieve higher performance, but this requires also more memory.

2.1.1. Mixtures of finite-context models

In order to attain better compression results, the previous approach can be used in a more robust scheme. The goal here is to use a model mixture with more than one finite-context model. In our case, we decided to use only two different models: the one used by [6] and another one based on *bit modeling by pixel value estimate*. This approach could be extended to use more models, but at a cost of some computation time. In what follows, we will describe how this mixture scheme works.

The per symbol information content average provided by the finite-context model of order-k, after having processed n symbols, is given by

$$H_{k,n} = -\frac{1}{n}\sum_{i=0}^{n-1}\log_2 P\left(x_{i+1} \mid x_{i-k+1\ldots i}\right) \tag{3}$$

bits per symbol (or bits per pixel, bpp, in the case of images). Hence, the $H_{k,n}$ can be viewed as a measure of the performance of the model until that instant. When using several models simultaneously, the probability estimate can be given by a weighted average of the probabilities provided by each model, according to

$$P\left(x_{n+1}\right) = \sum_k P\left(x_{n+1} \mid x_{n-k+1\ldots n}\right) w_{k,n}, \tag{4}$$

where $w_{k,n}$ denotes the weight assigned to model k and

$$\sum_k w_{k,n} = 1. \tag{5}$$

In order to compute the probability estimate for a certain symbol, it is necessary to combine the probability estimates given by (1) using (4). The weight assigned to model k can be computed according to

$$w_{k,n} = P\left(k \mid x_{1\ldots n}\right), \tag{6}$$

i.e., by considering the probability that model k generated the sequence until that point. In that case, we would get

$$w_{k,n} = P\left(k \mid x_{1\ldots n}\right) \propto P\left(x_{1\ldots n} \mid k\right) P\left(k\right), \tag{7}$$

where $P(x_{1\ldots n} \mid k)$ denotes the likelihood of sequence $x_{1\ldots n}$ being generated by model k and $P(k)$ denotes the prior probability of model k. Assuming

$$P\left(k\right) = \frac{1}{K}, \tag{8}$$

where K denotes the number of models, we also obtain

$$w_{k,n} \propto P\left(x_{1\ldots n} \mid k\right). \tag{9}$$

Calculating the logarithm, we get

$$\begin{aligned} \log_2 P\left(x_{1\ldots n} \mid k\right) &= \log_2 \prod_{i=1}^{n} P\left(x_i \mid k, x_{1\ldots i-1}\right) = && \text{(a)} \\ &= \sum_{i=1}^{n} \log_2 P\left(x_i \mid k, x_{1\ldots i-1}\right), && \text{(b)} \end{aligned} \tag{10}$$

which is related to the code length that would be required by model k for representing the sequence $x_{1\ldots n}$. It is, therefore, the accumulated measure of the performance of model k until instant n. In order to improve the global performance, we decided to use a mechanism that progressively forgets past performances of the models. This mechanism allows each model to progressively forget the past and, consequently, to give more importance to the most recent past. Therefore, we rewrite (11) as

$$\begin{aligned} &\sum_{i=1}^{n} \log_2 P\left(x_i \mid k, x_{1\ldots i-1}\right) = && \text{(a)} \\ &= \gamma \sum_{i=1}^{n-1} \log_2 P\left(x_i \mid k, x_{1\ldots i-1}\right) + \log_2 P\left(x_n \mid k, x_{1\ldots n-1}\right), && \text{(b)} \end{aligned} \tag{11}$$

where $\gamma \in 0{,}1$ dictates the forgetting factor to be used. Defining

$$p_{k,n} = \prod_{i=1}^{n} P\left(x_i \mid k, x_{1\ldots i-1}\right) \tag{12}$$

and removing the logarithms, we can rewrite () as

$$p_{k,n} = p_{k,n-1}^{\gamma} P\left(x_n \mid k, x_{1\ldots n-1}\right) \tag{13}$$

and, finally, set the weights to

$$w_{k,n} = \frac{p_{k,n}}{\sum_k p_{k,n}}. \tag{14}$$

Usually, a compression algorithm can be divided into two parts, modeling and coding. The Markov models are responsible for providing a statistical model as reliable as possible to be used later in the coding stage. The coding stage is where the statistical model is used to compress the data. Arithmetic coding is usually the technique that is used. In the following section, we give a brief explanation on how arithmetic coding works.

2.1.2. Arithmetic coding

Arithmetic coding is a compression technique developed by Rissanen [19] in the late 1970s. This method is a good alternative to Huffman coding [20], because it usually generates better compression results. In order to obtain better results, an appropriate probability estimator must be used in the arithmetic encoder (e.g., based on finite-context models, as described above).

This method represents a set of symbols using a single number in the interval [0,1). As the number of symbols of the message grows, the initial interval [0,1) will shrink and the number of bits necessary to represent the interval will increase. When we are processing the pixels of an image in a raster scan order, the probabilities of the intensities of the pixels are conditioned by the context determined by a combination of the already encoded neighboring pixels. The encoder and the decoder estimate this context model dynamically adapting it to the input data, during the encoding/decoding process. According to [16, 17, 18], this arithmetic encoding method generates output bitstreams with average bitrates almost identical to the entropy of the source model.

2.2. Bitplane decomposition

The technique to separate an image into different planes (bitplanes), known as bitplane decomposition, plays an important role in image compression. Usually, each pixel of a grayscale image is represented by 8 bits, or 16 bits as the case of some biomedical images used in the experimental results presented in this chapter. Suppose that the image has $N \times M$ pixels and each one is composed of eight bitplanes, ranging from bitplane 0 for the least significant bitplane (LSBP) to bitplane 7 for the most significant bitplane (MSBP). In fact, plane 0 contains all the lowest-order bits in the bytes comprising the pixels in the image as well as plane 7 holds the most significant bits [21]. Figure 2 illustrates these ideas and Fig. 3 shows the various bitplanes for the image presented on the left. As we can see, the MSBPs (especially 7-4) contain the majority of the visually significant data. On the other hand, the lower planes (namely, planes 0-3) contribute to more subtle details in the image.

The bitplane decomposition technique is very useful on image compression. On one hand, it allows some bi-level compression methods, such as JBIG, to be applied to typical grayscale images. The compression method is applied to each bitplane after the decomposition. On the other hand, it is possible to create sophisticated models that take advantage of this decomposition. For instance, it is possible to use information of the previous bitplanes (usually the MSBPs) to improve the compression performance of the LSBPs.

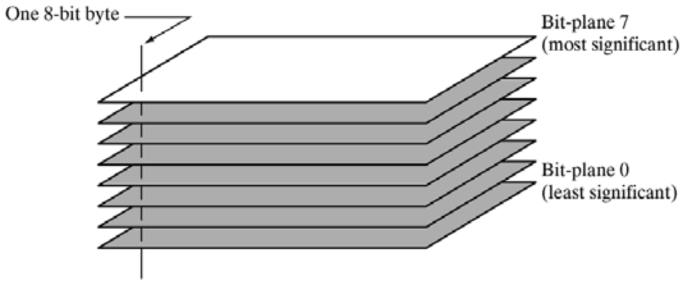


Figure 2. Bitplane representation of an eight-bit image [21]

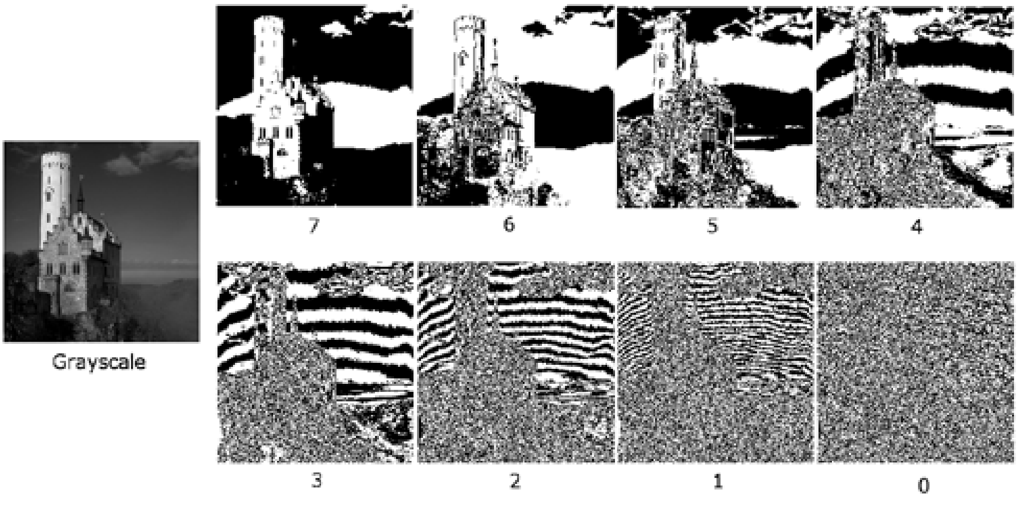


Figure 3. An 8-bit grayscale image and its eight bitplanes. The numbers at the bottom of each image identify the bit-plane, where 0 denotes the less-significant plane and 7 the most significant plane (adapted from [22])

In [23], an embedded image-domain adaptive compression (EIDAC) scheme used with success on simple images (with a reduced number of active intensity values) is presented. EIDAC uses a context-adaptive bitplane coder, where each bitplane is encoded using a binary arithmetic coder. Later, in [24], a compression method inspired from EIDAC for microarrays image is proposed. The same authors improved their method using a more robust 3D context modeling [24]. We used this last approach in our experiments and also other alternative versions using some preprocessing techniques, namely, segmentation and bitplane reduction.

2.2.1. Segmentation

In the literature, we can find several algorithms for image segmentation. In our approach, we decided to use a threshold-based method inspired from the work presented in [25]. The segmentation is attained by means of a dynamic thresholding scheme. By applying a threshold value, the pixels of the image can be split in two sets (background and foreground). For each threshold value, it is possible to obtain the number of pixels and the standard deviation of the intensities of these pixels in each set. The desired threshold is calculated according to (17). Using (17), it is always guaranteed that the weighted sum of the standard deviation of both the background and foreground is minimal,

$$\mathcal{T} = \operatorname{argmin}\left\{ f\left(T\right) \right\},\ T = \left\{ t \in \mathsf{N} \mid 0 \le t \le 2^{16} - 1 \right\}, \tag{15}$$

where

$$f\left(T\right) = \operatorname{stdev}\left(B_T\right) \times \operatorname{size}\left(B_T\right) + \operatorname{stdev}\left(F_T\right) \times \operatorname{size}\left(F_T\right), \tag{16}$$

$B_T\, p \in \mathrm{Image}\, p < T\}$ represents the set of pixels in the background section, $F_T\, p \in \mathrm{Image}\, p \ge T\}$ represents the set of pixels in the foreground section, $\operatorname{stdev}(x)$ is the standard deviation of x, and $\operatorname{size}(y)$ is the number of pixels of set y. Instead of testing all possible threshold values to find the minimum value of $f(T)$, a recursive search algorithm is used to accelerate the search routine. It is possible to use this recursive search algorithm because $f(T)$ plunges down at a certain threshold value, which is chosen as the final threshold value. After the threshold search is completed, a binary map is created where the foreground pixels will be set to "1" and the background pixels to "0."

2.2.2. Bitplane reduction

Bitplane reduction is an interesting method that can further improve compression efficiency by eliminating redundancy in the pixel precision for simple images. Simple images are images where the number of different intensities that occur is very small, compared to the total number of possible intensities. For example, if we have only 24 different intensities out of 256 for an 8-bit image, we only need 5 bits to represent each pixel intensity. This means that, when we are encoding the image, we only need to encode five bitplanes instead of the original eight bitplanes.

Yoo et al. [26] have shown that it is possible to obtain compression gains using the simplest form of bitplane reduction, known as histogram compaction (HC). Later, [23] presented a more robust bitplane reduction method called scalable bitplane reduction (SBR). This approach finds the reduced bitplane codeword by growing a binary tree. The method splits each node of the binary tree into two nodes using a simple MINMAX metric to measure the distortion.

In order to better understand the SBR algorithm, we present a small example in Fig. 4. Initially, we associate all the active pixel values to the root node. In this small example, the image only has four active pixel values. Starting in the root node, the algorithm splits all the children sub-nodes using a MINMAX criterion. The split process in the root node starts by computing the value 32,767 from $(0 + 65{,}535$. The computed value is then used to split the node. All the intensities that are lower or equal to 32,767 are inserted in the left sub-node. On the other hand, the remaining intensities (>32,767) are associated with the right sub-node. After splitting the root node, the SBR algorithm adds a zero to the left node codeword and a one to the right node codeword. This splitting process is repeated until all sub-nodes have only one intensity associated with them. In this specific example, "0" is a complete codeword for the original pixel value zero, due to the fact that the first left node or partition does not have more sub-nodes. As a result of the variable-length codewords, the average codeword length can be less than 3 bits. This is useful because during the encoding process, it is possible to skip some bitplanes of the pixel codewords with a lower length.

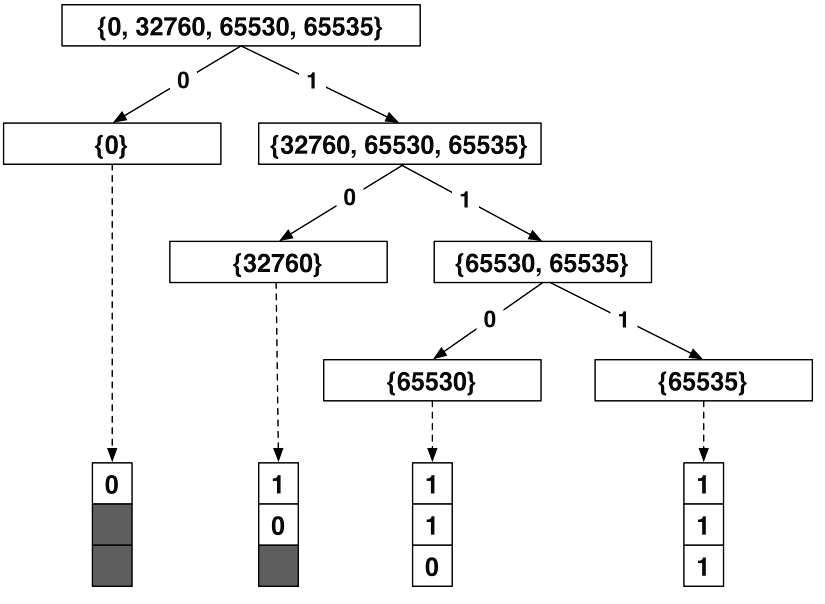


Figure 4. Binary tree to obtain the codewords to represent the active pixel values 0,32,760,65,530,65,535 in the reduced bitplane space after applying the SBR algorithm. Each of the tree nodes is associated with a symbol (an intensity) set to be partitioned, except the end nodes. Each branch defines the bit value of specific symbols at the corresponding bitplane in the reduced bitplane domain. For the intermediate nodes, 32,760,65,530,65,535 and 65,530,65,535, the split process is done using the MINMAX criterion

Taking as example the intensities 0,32,760,65,530,65,535 presented in Fig. 4, we present in Table 2 the codewords for the two histogram reduction methods. As can be seen, the codewords obtained using the HC algorithm are dependent on the number of active intensities. The codeword size can be computed according to

$$S = \left\lceil \log_2(N) \right\rceil, \tag{17}$$

where N denotes the number of active intensities. In the example presented in Fig. 4, we have four different intensities, so the codeword size is $\log_2 4 = 2$. The codeword size is constant for all intensities for the HC method. On the contrary, the resulting codewords obtained using the SBR algorithm have different sizes (see Table 2).

Intensity Value	HC		SBR	
	Codeword	Codeword size	Codeword	Codeword size
0	00	2	0	1
32,760	01	2	10	2
65,530	10	2	110	3
65,535	11	2	111	3

Table 2. A small example showing the differences between the two bitplane reduction methods HC and SBR. The codeword in each column represents the new value that will be assigned to the pixel values of the first column. In HC, the codeword size is constant. In SBR, the codeword size is variable

2.2.3. Simple bitplane coding

In [27], Kikuchi et al. introduced the concept of bit modeling by the pixel value estimates. In their approach, instead of using the true bit values of each bitplane, they used the expectation values of the pixels to build up the contexts. This approach is known as *bit modeling by pixel value estimate*. They extended their work more recently to be applied to various types of images (color, grayscale, color-quantized, bi-level, and halftone) [28] and for HDR (high-dynamic-range) images [29].

Let us consider that a given pixel value at location (i, j) in a given image to be encoded is denoted as $x(i, j)$. Its decoded value is denoted by $y(i, j)$. In order to facilitate the explanation, the location indexes (i, j) are omitted from now on. A typical raster scan order is used to process each pixel of a given image. Similar to Kikuchi's method, as the process of the bitplane coding proceeds to lower bitplanes, the decoded value, y, of the target pixel approaches the true pixel value $x = (x_{16}\, x_{15} \,\dots\, x_1)$ where x_n denotes the n^{th} bit of x in the case of a 16-bit grayscale image.

Contrarily to Kikuchi's method, our approach uses only one type of context, denoted as *neighborhood context* in Kikuchi's work. The contexts are built by the estimates of partially decoded pixels based on the template depicted in Fig. 5. The pixel location of x is labeled by "X" on the 15-pixel template in Fig. 5, and

$$c_k = \begin{cases} 1, & \text{for } y(k) > y \\ 0, & \text{otherwise} \end{cases}, \tag{18}$$

where $k \in \{1, 2, \dots, 15$ denotes the spatial location on the template illustrated in Fig. 5. $y(i)$ and y represent the most recent estimates of the neighboring pixels and the target pixel, respectively.

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

Figure 5. Fifteen-pixel template for building up the context. The target pixel to be encoded is labeled by an "X"

In Kikuchi's method, the inter-bit correlation on a bitplane is not used. Instead, for modeling a target bit, the authors used the pixel value estimates of which more significant bits have been already available at the decoder. Their method is referred to as *bit modeling by pixel values*, where the pixel value estimates are used rather than of the unknown true values at the decoder. Contrarily to Kikuchi's method that uses a nine-pixel template, our approach uses a variable-size 15-pixel template (see Fig. 5). In our case, a greedy search routine is performed in each bitplane in order to obtain the context size that attains the best compression performance (lower bits per pixel as possible). According to Kikuchi's method, the decoded pixel values are spatially correlated to each other as significantly high as the true pixel values which will make sense to use a larger template. Furthermore, since the alphabet size is only two, the probability of having context dilution is low. Similar to Kikuchi's method, we are considering some noncausal locations with respect to the scanning order of the pixels (locations $y3 y4 y10$ and $y11$ in Fig. 5). The usage of noncausal pixels is only possible because the context bits are defined by using the estimates of pixel values, which are available in the decoder.

Let us consider that we are coding an N -bit depth image, where $N \leq 16$. Suppose that the n^{th} bitplane is being encoded at present, where $n \in \{1 \ldots N\}$. For every pixel, the higher bits from the $(n+1)^{\text{th}}$ until N^{th} bitplane are known at the decoder. The other lower n bits are unknown. The value of the unknown part can be distributed over the interval of $[0, 2^n - 1]$. Similar to Kikuchi's method, the values zero and one occur with equal probability in the unknown less significant n bits. Under this assumption, the pixel value estimate of the target pixel is expressed by

$$y^{(n)} = \left\lfloor \frac{y}{2^n} \right\rfloor + 2^{n-1} - 1 \tag{19}$$

at the n^{th} bitplane encoding/decoding, where y is the latest decoded value and . denotes truncation. In Fig. 6, we illustrate an example of a binary representation of the pixel value estimate in the case $n = 8$.

Initially, all the pixel estimate values are set to $y = 2^{N-1} - 1$. The encoding procedure starts at the MSBP and stops at the LSBP. Assuming a target pixel x_n of bitplane n is the one being encoded, under the context of $\{c_k\}$, the pixel estimate y of the target pixel is immediately updated by a simple bit operation as

$$y \leftarrow y + x_n 2^{n-1} - \left\lfloor 2^{n-2} \right\rfloor. \tag{20}$$

The pixel value estimate is used in a coming chance of reference and will be the decoded pixel value, when the decoding is stopped (after all the bitplanes are processed). The most recent estimate of a given pixel is always made up of two parts: its significant bits are those already encoded/decoded true bits and the other less significant bits are 0 (zero) followed by a successive 1s (ones). The value of the less significant bits is equal to the expectation value of the unknown lower bits, if binary symbols of zero and one are assumed to occur in those bits with equal probability.

$$y^{(8)} = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|c|} \hline x_{16} & x_{15} & x_{14} & x_{13} & x_{12} & x_{11} & x_{10} & x_9 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

Figure 6. Binary representation of a pixel value estimate in the case of n=8

2.3. Binary-tree decomposition

Binary trees are also an important data structure that can be used in several algorithms. In the case of image compression, we can associate each leaf node of the binary tree to an image intensity. The binary tree can be viewed as a simple nonlinear generalization of lists; instead of having one way to continue to another element, there are two alternatives that lead to two different elements [30]. Every node (or vertex) in an arbitrary tree has at least two children (see Fig. 7). Each child is designated as left child or right child, according to the position in relation to the tree root.

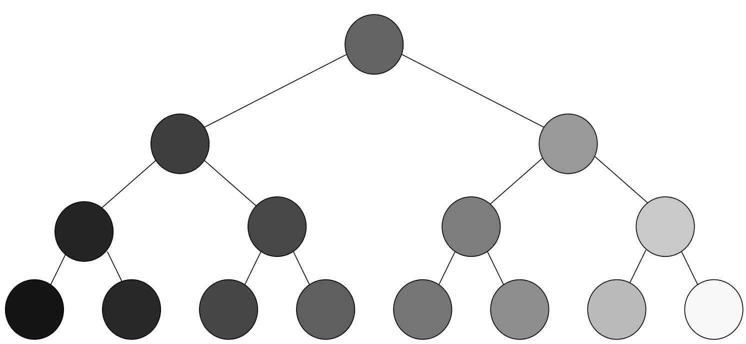


Figure 7. An example of a binary tree with eight intensities or gray levels. Each internal node contains an intensity representative that is computed according to the set of intensities that are associated to the node. The leaf nodes represent the number of different intensities that actually occur; in this case, we have a total of eight intensities

One of the first methods where binary trees were used for image compression was proposed by Chen et al. [31] regarding the compression of color-quantized images. Chen's method uses a binary-tree structure of color indexes instead of a linear list structure. Using this binary-tree structure, Chen's method can progressively recover an image from two colors to all of the colors contained in the original image. Inspired by the work done by [31], a few years later, the authors of [7, 11, 12] developed a lossy-to-lossless method based on binary-tree decomposition and context-based arithmetic coding. In the last approach, the authors studied the performance of their method in several kinds of grayscale images, including medical images. As can been seen, this decomposition approach is very versatile because it can be applied in color-quantized images and also in grayscale images.

This binary-tree decomposition approach was intended to be used in images with a small number of intensities, usually with eight or less bits per pixel, due to a tight relation between the processing time and the number of different intensities of the image. In this work, we intend to study the performance of this approach to compress biomedical images, such as microarray images, RNAi, etc. Another interesting feature of this approach is its capability of progressive decoding, which means that the decoding process can be stopped at any moment according to a specific distortion metric, obtaining an image with some loss. Moreover, it is possible to obtain the original image without any loss if the full decoding process is performed. In the next two sections, we describe in more detail the compression algorithm that is based on a binary-tree decomposition and context-based arithmetic coding.

2.3.1. Hierarchical organization of the intensity levels

This method is based on a hierarchical organization of the intensity levels of the image. This organization of the intensity levels is attained by means of a binary tree. Each node of the binary tree, n, represents a certain subset, S^n, of the intensities of the image. The root node contains all active pixel values of the image $\mathcal{I} I_1, I_2, \ldots, I_N\}$, where N represents the number of different intensities that occur in the image. Therefore, $\mathsf{S}^n \subset \mathcal{I}$ and $\mathsf{S}^1 \equiv \mathcal{I}$. Each node possesses a representative intensity, I^n, given by

$$I^n = \left\lfloor \frac{I_m^n + I_M^n}{2} \right\rfloor, \tag{21}$$

where I_m^n and I_M^n are, respectively, the smallest and largest pixel value in S^n and where x denotes the largest integer less than or equal to x. Computing the value of I^n according to (23) leads to the smallest possible L_∞ reconstruction error when the intensities associated to node n (those in S^n) are all substituted by I^n. The error is given by

$$\varepsilon_\infty^n = I_M^n - I^n. \tag{22}$$

In order to better understand the construction of the binary tree, we present in Fig. 7 a small example for an image with only five active pixel values 32,50,250,33,768,65,530. The construction of this tree begins with the association to the root node of the set of intensities that occur in the original image. After this association, it is necessary to compute I^1 according to (23). In the example depicted in Fig. 8, $I^1 = \lfloor (32 + 65{,}530)/2 \rfloor = 32{,}781$ and $\varepsilon_\infty^1 = 65{,}530 - 32{,}781$, for the root node. The next step consists of splitting the root node into two sub-nodes and, therefore, splitting S^1 into two subsets. In order to split S^1, we need only to compare the intensity $I \in \mathsf{S}^1$ with I^1. The intensities lower than I^1 are associated with the left node and the other ones with the right one. This procedure is repeated until all nodes are expanded, i.e., until having a tree with N leaves (N is the number of active intensities presented in the original image). The next node to expand is chosen, taking into consideration the smallest possible L_∞ reconstruction error. In case of a tie, one is arbitrarily chosen, although it is necessary that the decoder picks the same one.

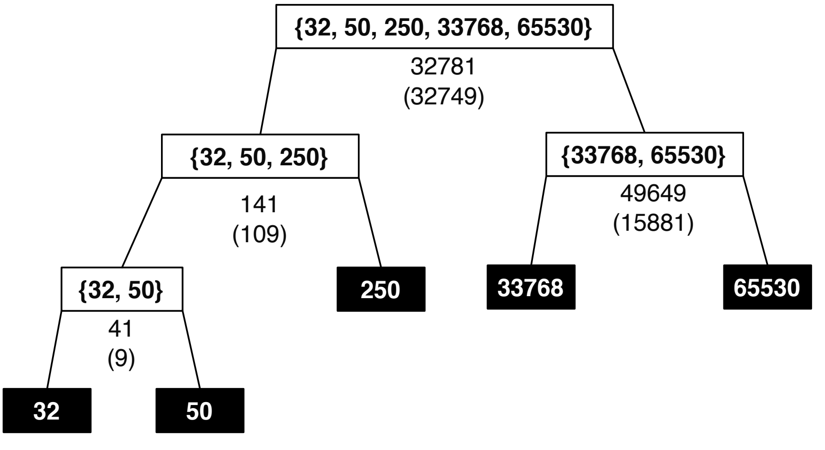


Figure 8. Example of a small binary tree that illustrates the hierarchical organization of the intensity values for an image with five active pixel values 32,50,250,33,768,65,530

In order for the decoder to be able to build the same tree, it is necessary to send the information of the active pixel values to the decoder using a 65,536-bit indicator. In order to encode this indicator, the encoder uses the following strategy. First, the maximum intensity value I_N is sent. After that, a string of I_n bits is transmitted, such that if the n^{th} bit of the string is one, it means that the intensity $n-1$ is present in the image and zero otherwise. The previous 65,536-bit indicator is enough for the decoder to construct exactly the same binary tree.

2.3.2. Encoding pixel locations

After each node is expanded, two new nodes are created, each one with a representative intensity (I_l^n for the left node and I_r^n for the right node). This step can be seen as a region of arbitrary shape, containing zeros (relative to the left node) and ones (regarding the right node), that needs to be communicated to the decoder. The position where the pixels in the image are associated with the parent node that was expanded is known by the decoder. However, it is necessary to communicate to the decoder the zeros and ones that correspond to the pixels that after the expand procedure will be associated to the left and right nodes, respectively. Since the decoder has access to the pixels associated to the parent node that was expanded, it is enough to encode a binary mask, where zero indicates that the pixel needs to change its intensity to I_l^n and one indicates a change to I_r^n. This binary mask is encoded using arithmetic coding based on variable-size finite-context models [16, 17, 18].

The performance of the compression method is directly dependent on the encoding of these binary masks. The encoding efficiency of these binary masks can be controlled by a carefully chosen context modeling that will then drive the binary arithmetic encoder. The context is constructed based on the template depicted in Fig. 9. The number of context pixels can go up to 16 at most, and they are numbered according to their distance to the encoding pixel (represented in gray in Fig. 9). A particular context is represented using a sequence of bits,

$$b_1 b_2 \dots b_k \tag{23}$$

where

$$b_i = \begin{cases} 0, & \text{if} \quad \left| I(i) - I_l^n \right| \leq \left| I(i) - I_r^n \right| \\ 1, & \text{otherwise} \end{cases},$$

and where $I(i)$ denotes the intensity of the pixel in the current reconstructed image corresponding to position i of the context template.

	14	10	15	
13	5	2	6	16
9	1		3	11
	8	4	7	
		12		

Figure 9. Context template used in this work. The use of noncausal pixels is possible, because context information can be obtained from the previous version of the reconstructed image

The value k defines the model order used. In this case, the k value varies as the encoding proceeds. This variation is necessary in order to improve the compression performance. Furthermore, it is expected to have larger mask regions initially in the first nodes that are expanded and smaller regions when $n \approx N$. This variation is also useful to avoid the problem of context dilution. In this research work, we present two modes of context creation. One is denoted as "greedy" where the context size is first chosen using a k value according to [32]. After that, the method tests incrementally several context sizes bigger and smaller than k and stops when it reaches one context that produces worse results than the previous best. In the end, the algorithm has two context sizes. One attained when applying an increment to k and the other one when applying a decrement to k. The best context size is then chosen to encode the binary mask. The other mode is slower because it tests all possible context sizes. This second mode, denoted as "best," always attains the best context size that minimizes the bitrate. For both cases, the context size needs to be sent to the decoder, for each node that is expanded.

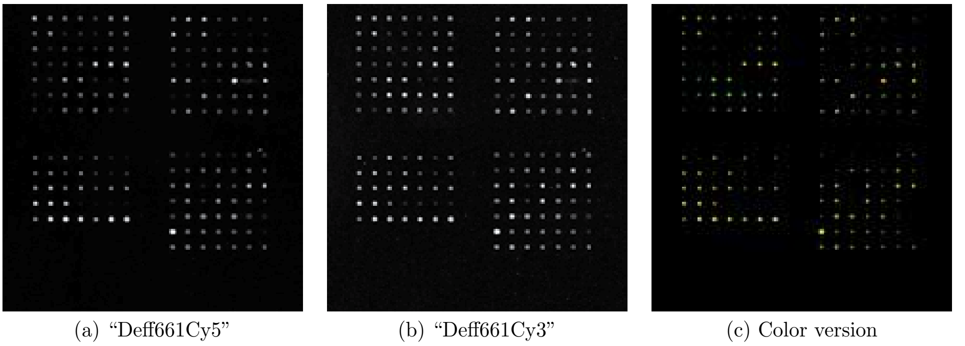
There is also an alternative way to encode the binary mask. If the number of bits required to encode the mask and the context size is bigger than the total number of pixels associated with the node to be expanded, the encoder sent the binary mask as a binary string, without compression. In order for the decoder to differentiate between these two modes, a binary stream is needed to be encoded for each node that is expanded.

3. Evaluation of the compression methods

In this section, we present compression results obtained by the compression methods presented earlier in Sect. 4. We decided to present only the results of the best version of each method in order to avoid extending this section. We start with the description of the data sets used in this work. Then, we provide the obtained compression results. At the end of this section, we present a study of the rate distortion, comparing the two compression approaches described in Sect. 4 and two image coding standards, JBIG and JPEG2000.

3.1. Data sets

To evaluate the compression methods presented in this chapter, we used three types of images: microarray, medical, and RNAi. The output data obtained in a microarray experiment is a pair of 16-bits-per-pixel grayscale images, one from the so-called red channel and the other from the green channel (see Fig. 10). We use a total of 298 microarray images of nine different data sets as described in Table 3.

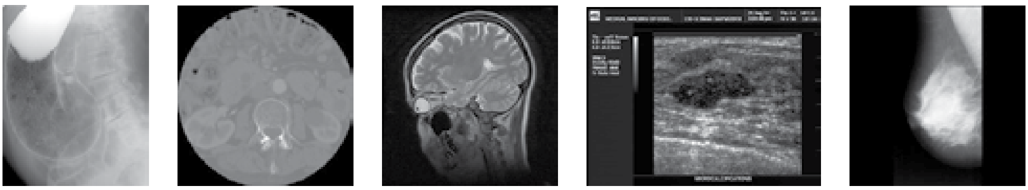


(a) “Deff661Cy5” (b) “Deff661Cy3” (c) Color version

Figure 10. Example of a microarray experiment from the *ISREC* data set. (a) and (b) correspond to the pair of microarray images. (c) a color version created from (a) as the red channel and (b) as the green channel

http://dx.doi.org/10.5772/60650

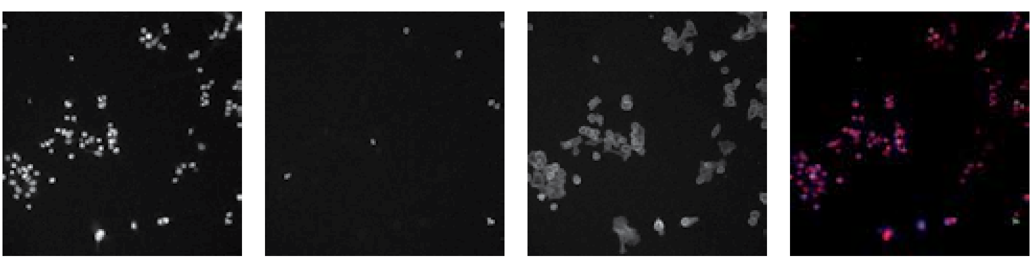
Regarding the medical images, we used images of four modalities: computed radiography (CR), computed tomography (CT), magnetic resonance (MR), and ultrasound (US). In Fig. 11, we can find five examples of medical images used in this work. The data set used in [33] was also considered as a medical reference data set. All the medical image data sets are depicted in Table 4, a total of 370 images.



(a) “cr_4503” (b) “ct_17” (c) “mr_2321” (d) “us_3393” (e) “mdb001”

Figure 11. Example of medical images from different modalities

Finally, we also used a recent and popular image type that can be extracted from RNA interference (RNAi) experiments. Those experiments involve marking of cells with various fluorescent dyes to capture three components of interest, namely, DNA, actin, and PH3 channels [4]. In Table 5, we can find the RNAi image data sets used in this work, a total of 34,560 images. The images were retrieved from the RNAi experiments without any kind of transformation (no normalization was applied). For us, it does not make sense to perform a normalization operation in order to reduce the image depth from 10 to 8 bits per pixel. The normalization is a lossy process, and hence, applying a lossless compression method to images that suffered some loss is not appropriate. In Fig. 12 we present some examples of RNAi images.



(a) DNA (b) PH3 (c) Actin (d) Merged

Figure 12. Examples of RNAi image, (a) DNA channel, (b) action channel, (c) PH3 channel, and (d) all the previous channels merged in an RBG image

In Tables 3-5, we can see several important measures: image size, depth, entropy, intensity usage, and a sparsity measure called Gini index or GI [34]. The GI is a normalized measure that assumes values between zero and one. Values close to zero means that image has a lower sparsity histogram. On the other hand, values close to one represent an image that has a very sparse histogram.

Data sets	Year	Images	Approximate size (cols × rows)	Depth	Average entropy (bpp)	Average intensity usage (percentage)	Gini Index [0-1]
ApoA1 [35]	2001	32	> 1,044×1,041	16	11.038	39.507%	0.494
Arizona [36]	2011	6	=13,800×4,400	16	9.306	82.821%	0.774
IBB [37]	2013	44	= 2,019×6,235	16	8.503	54.072%	0.806
ISREC [38]	2001	14	= 1,000×1,000	16	10.435	33.345%	0.710
Omnibus-LM [39]	2006	25	=12,200×4,320	16	5.713	50.130%	0.726
Omnibus-HM [40]	2006	25	=12,200×4,320	16	7.906	98.076%	0.892
Stanford [41]	2001	40	> 1,900×2,000	16	8.306	27.515%	0.615
Yeast [42]	1998	109	= 1,024×1,024	16	6.614	5.391%	0.518
YuLou [43]	2004	3	> 1,800×1,900	16	9.422	36.906%	0.556
Overall		298		16	7.415	67.003%	0.782

Table 3. Microarray image data sets used in this work. The number of images represents the total number of images that each data set contains (each image corresponds to one channel)

Data sets	Year	Images	Approximate size (cols × rows)	Depth	Average entropy (bpp)	Average intensity usage (percentage)	Gini Index [0-1]
CR	2003	12	> 612× 746	10-16	9.905	11.487%	0.240
CT	2003	12	> 340× 340	12-16	8.032	2.879%	0.389
MR	2003	12	> 256× 256	16	6.837	1.636%	0.624
US	2003	12	> 584× 476	8	4.561	0.358%	0.714
MIAS [33]	1995	322	=1,024×1,024	8	4.544	0.357%	0.681
Overall		370		8-16	5.169	1.596%	0.631

Table 4. Medical image data sets used. The computed radiography (CR), computed tomography (CT), magnetic resonance (MR), and ultrasound (US) data sets can be found in [44]. The last data set, mini-MIAS, contains 322 mammography images

Taking into consideration Table 3, we can see that the image sizes of the microarray data sets are very large (ranging from 1,000×1,000 to 13,800×4,400) and the bit depth is 16 for all the data sets. The average entropy goes from 5.7 bpp to 11 bpp, and the percentage of active intensities varies from 5.4% to 98.1%. The GI measures are all higher than 0.49 for all data sets reaching values close to 0.89 (very sparse) for the *Omnibus-HM* data set.

Data sets		Year	Images	Approximate size (cols × rows)	Depth	Average entropy (bpp)	Average intensity usage (percentage)	Gini Index [0-1]
RNAi D1	| (DNA)	2006	2,304	≈512×512	12	7.408	4.140%	0.274
	| (PH3)		2,304			5.631	1.019%	0.056
	| (Actin)		2,304			7.583	2.287%	0.192
RNAi D2	| (DNA)	2006	2,304	≈512×512	12	6.617	2.925%	0.172
	| (PH3)		2,304			5.572	0.661%	0.050
	| (Actin)		2,304			6.565	1.334%	0.099
RNAi D3	| (DNA)	2006	2,304	≈512×512	12	7.210	4.195%	0.257
	| (PH3)		2,304			5.659	0.981%	0.054
	| (Actin)		2,304			7.653	3.215%	0.242
RNAi D4	| (DNA)	2006	2,304	≈512×512	12	7.530	4.249%	0.286
	| (PH3)		2,304			5.662	1.080%	0.057
	| (Actin)		2,304			7.907	3.070%	0.244
RNAi D5	| (DNA)	2006	2,304	≈512×512	12	7.281	4.046%	0.256
	| (PH3)		2,304			5.622	0.904%	0.052
	| (Actin)		2,304			7.649	2.772%	0.221
RNAi all	| (DNA)	2006	11,520	≈512×512	12	7.209	3.911%	0.249
	| (PH3)		11,520			5.629	0.929%	0.054
	| (Actin)		11,520			7.471	2.536%	0.199
Overall		2006	34,560	≈512×512	12	6.770	2.458%	0.167

Table 5. Five RNAi image data sets extracted from the first five plates that can be found in [45]. Contrary to other authors, the images were extracted from the RNAi files without any loss. The exact image intensity values were used to create the images. No normalization was performed

Regarding the medical image data sets (see Table 4), we can see that the image sizes are lower when compared to the microarray images (except for the *mini-MIAS* data set). The bit depth varies from 8 to 16 bits and the average entropy from 4.5 bpp to 9.9 bpp. The average intensity usage is very low for these medical images. For the *CR* data set, only 11.5% of the total available intensities are actually used. For the other data sets, this measure is even lower (3). For the GI measure, we can see that the *MR*, *US*, and *mini-MIAS* data sets have a GI higher than 0.6. On the other hand, for the *CR* and *CT* data sets, the GI values are lower (0.24 and 0.39, respectively).

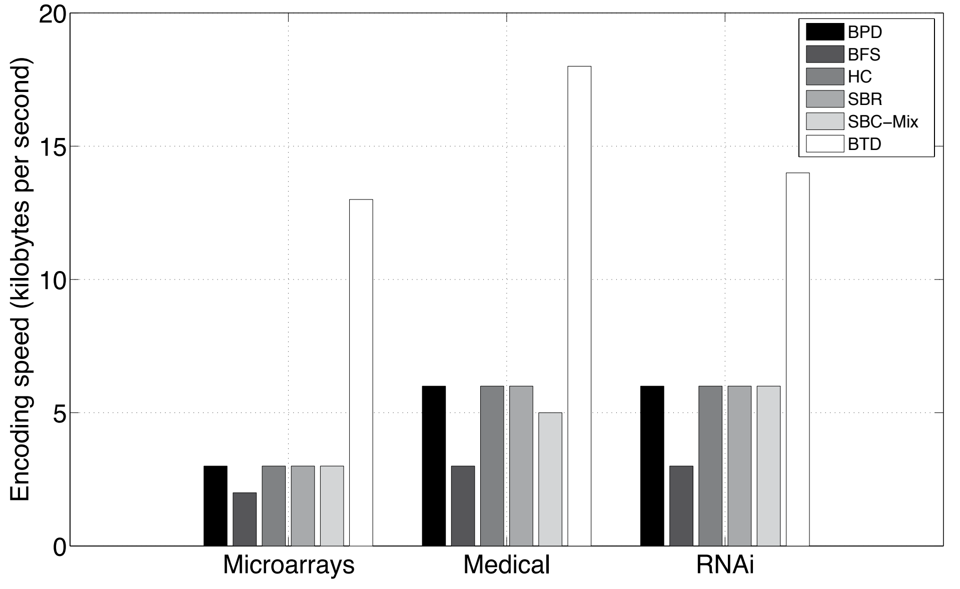


Figure 13. Encoding speed in kilobytes per second for the methods evaluated in this work. BPD (bitplane decomposition), BFS (background-foreground separation), HC (histogram compaction), SBR (scalable bitplane reduction), SBC-Mix (simple bitplane coding-mixture), and BTD (binary-tree decomposition)

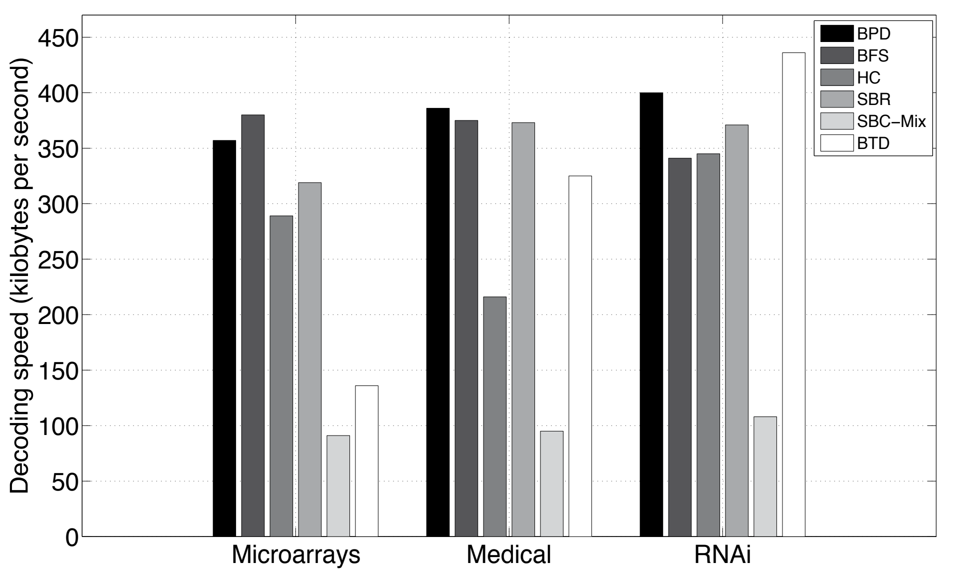


Figure 14. Decoding speed in kilobytes per second for the methods evaluated in this work. BPD (bitplane decomposition), BFS (background-foreground separation), HC (histogram compaction), SBR (scalable bitplane reduction), SBC-Mix (simple bitplane coding-mixture), and BTD (binary-tree decomposition)

Finally, we provide 5 RNAi image data sets in Table 5. We split each data set into three channels (DNA, PH3, and actin). All the images have size 512×512 and depth of 12 bits. Globally, the average entropy is about 6.77 bpp and the average intensity usage is around 2.5%. Regarding the GI, the values are very low (<0.17), which means that these images are not very sparse.

3.2. Experimental results

In this subsection, we present the results obtained using several general purpose compression tools (e.g., gzip, bzip2, PPMd, and LZMA), image coding standards (e.g., PNG, JBIG, JPEG-LS, and JPEG2000), and the methods described in Sect. 4. In Tables 6-8, we can find the compression results for all data set described earlier in Sect. 5.1, using several general compression tools and image coding standards. Default parameters were used in all compression tools. Hence, we did not try to adjust some parameters in order to improve the compression results. Only the lossless mode was imposed.

JBIG results were obtained using version 2.0 of the JBIG-Kit package[1]. The results for the JPEG-LS standard were obtained using version 2.2 of the SPMG JPEG-LS codec[2]. JPEG2000 lossless compression was obtained using version 5.1 of JJ2000 codec with default parameters for lossless compression[3]. For additional reference, we also provided compression results using the gzip, bzip2, ppmd, and lzma general-purpose compression tools.

	Compression methods							
Data sets	**Gzip**	**Bzip2**	**PPMd**	**LZMA**	**PNG**	**JBIG**	**JPEG-LS**	**JPEG2000**
ApoA1	12.711	11.068	10.984	11.374	12.568	10.851	10.608	11.063
Arizona	11.263	9.040	8.980	9.402	11.017	8.896	8.676	9.107
IBB	10.453	9.081	8.495	8.985	10.090	9.344	9.904	10.516
ISREC	12.464	10.922	10.730	11.126	12.476	10.925	11.145	11.366
Omnibus (LM)	7.124	5.346	4.977	5.527	6.781	5.130	4.936	5.340
Omnibus (HM)	9.558	7.523	7.219	7.787	9.160	7.198	6.952	7.587
Stanford	9.972	7.961	7.809	8.273	9.776	7.906	7.684	8.060
Yeast	7.672	6.075	5.794	6.389	8.303	6.888	8.580	9.079
YuLou	11.434	9.394	9.285	9.708	11.428	9.298	8.974	9.515
Average	9.044	7.189	6.859	7.388	8.729	7.051	6.996	7.511

Table 6. Lossless compression results, in bits per pixel (bpp), using gzip, bzip2, PPMd, LZMA, PNG, JBIG, JPEG-LS, and JPEG2000 for the microarray image data sets. Default compression parameters have been used for all algorithms. The best results are highlighted in bold

1 http://www.cl.cam.ac.uk/ mgk25/jbigkit

2 The original website of this codec, http://spmg.ece.ubc.ca, is currently unavailable. However, it can be obtained from http://sweet.ua.pt/luismatos/codecs/jpeg_ls_v2.2.tar.gz

3 The original website of this codec, http://jj2000.epfl.ch, is currently unavailable. Nevertheless, this codec can be obtained from http://sweet.ua.pt/luismatos/codecs/jj2000_5.1-src.zip

	Compression methods							
Data sets	**Gzip**	**Bzip2**	**PPMd**	**LZMA**	**PNG**	**JBIG**	**JPEG-LS**	**JPEG2000**
CR	9.284	6.140	6.230	7.197	7.991	6.152	5.784	5.845
CT	8.790	5.604	5.847	6.996	9.411	8.822	8.089	8.334
MR	8.342	5.313	5.514	6.663	9.795	9.940	9.321	9.389
US	3.496	2.742	2.642	2.820	3.061	2.908	2.750	3.138
mini-MIAS	2.637	1.696	1.590	2.022	1.672	1.562	1.416	1.443
Average	3.452	2.242	2.163	2.660	2.484	2.183	2.005	2.040

Table 7. Lossless compression results, in bits per pixel (bpp), using gzip, bzip2, PPMd, LZMA, PNG, JBIG, JPEG-LS, and JPEG2000 for the medical image data sets. Default compression parameters have been used for all algorithms. The best results are highlighted in bold

		Compression methods							
Data sets		**Gzip**	**Bzip2**	**PPMd**	**LZMA**	**PNG**	**JBIG**	**JPEG-LS**	**JPEG2000**
RNAi D1	| (DNA)	9.120	6.846	6.832	7.314	9.857	6.802	6.311	6.533
	| (PH3)	7.461	5.933	5.804	6.008	11.552	5.994	5.806	5.830
	| (Actin)	9.274	6.807	6.839	7.353	11.801	6.790	6.304	6.452
RNAi D2	| (DNA)	8.358	6.469	6.389	6.714	9.421	6.469	6.094	6.233
	| (PH3)	7.396	5.884	5.760	5.963	12.420	5.964	5.793	5.804
	| (Actin)	8.309	6.378	6.315	6.638	12.449	6.422	6.032	6.122
RNAi D3	| (DNA)	8.948	6.735	6.637	7.153	9.665	6.704	6.238	6.436
	| (PH3)	7.490	5.952	5.819	6.030	11.756	6.042	5.812	5.842
	| (Actin)	9.362	6.874	7.387	7.431	10.582	6.838	6.353	6.509
RNAi D4	| (DNA)	6.243	6.925	6.915	7.415	10.014	6.878	6.372	6.612
	| (PH3)	7.491	5.950	5.824	6.028	11.548	6.028	5.813	5.845
	| (Actin)	9.586	7.014	7.070	7.629	10.955	6.977	6.460	6.630
RNAi D5	| (DNA)	9.007	6.765	6.745	7.198	9.775	6.729	6.259	6.459
	| (PH3)	7.450	5.928	5.798	6.001	11.745	5.994	5.804	5.825
	| (Actin)	9.346	6.857	6.893	7.414	11.185	6.828	6.342	6.493
RNAi all	| (DNA)	8.935	6.748	6.703	7.159	9.746	6.716	6.255	6.455
	| (PH3)	7.458	5.929	5.801	6.006	11.804	6.004	5.806	5.829
	| (Actin)	9.176	6.786	6.900	7.293	11.394	6.771	6.298	6.441
Average		8.523	6.488	6.468	6.819	10.982	6.497	6.120	6.242

Table 8. Lossless compression results, in bits per pixel (bpp), using gzip, bzip2, PPMd, LZMA, PNG, JBIG, JPEG-LS, and JPEG2000 for the RNAi image data sets. Default compression parameters have been used for all algorithms. The best results are highlighted in bold

According to the results depicted in Tables 6-8, it seems that, globally, JPEG-LS is the best image coding standard. Among the general compression methods used, PPMd is the one that attained the best compression results.

Tables 9-11 provide lossless compression results using the methods described in Sect. 4. BPD (bitplane decomposition) corresponds to the results using method [6]; BFS (background-foreground separation) corresponds to using the previous approach with segmentation. The columns HC (histogram compaction) and SBR (scalable bitplane reduction) correspond to the approaches that use bitplane reduction in method [6]. The SBC-Mix (simple bitplane coding-mixture) corresponds to the approach described in Sect. 4.2.3 using a mixture of finite-context models. The last column, denoted as BTD (binary-tree decomposition), corresponds to the results obtained using the method described in Sect. 4.3.

If we look closely to Table 9, we can see that the method based on binary-tree decomposition is the one that attained the best compression results among all the others, for the case of microarray images. The best method is about 11% better than JPEG-LS. However, the method SBC-Mix is the best for three data sets: *ApoA1*, *ISREC*, and *Yulou*.

Regarding the experiments performed on medical images and taking into consideration the results presented in Table 10, we can conclude once again that the method based on binary-tree decomposition outperforms all the others and is about 9% better than JPEG-LS.

	Compression methods						
Data sets	**BPD**	**BFS**	**HC**	**SBR**	**SBC**	**SBC-Mix**	**BTD**
ApoA1	10.194	10.234	10.231	10.232	10.205	10.142	10.194
Arizona	8.242	8.245	8.244	8.243	10.308	8.219	8.186
IBB	7.974	7.982	7.978	7.978	8.537	7.966	7.943
ISREC	10.159	10.193	10.195	10.199	10.260	10.148	10.198
Omnibus (LM)	4.567	4.570	4.561	4.565	4.645	4.545	4.539
Omnibus (HM)	6.471	6.479	6.473	6.472	6.581	6.443	6.400
Stanford	7.379	7.349	7.350	7.349	7.403	7.305	7.303
Yeast	5.453	5.395	5.527	5.466	5.492	5.326	5.318
YuLou	8.619	8.641	8.626	8.626	8.669	8.591	8.592
Average	6.284	6.288	6.286	6.285	6.437	6.257	6.235

Table 9. Average compression results, in bits per pixel using the methods described in Sect. 4, BPD (bitplane decomposition), BFS (background-foreground separation), HC (histogram compaction), SBR (scalable bitplane reduction), SBC-Mix (simple bitplane coding-mixture), and BTD (binary-tree decomposition), for the microarray image data sets

	Compression methods						
Data sets	**BPD**	**BFS**	**HC**	**SBR**	**SBC**	**SBC-Mix**	**BTD**
CR	5.338	5.207	5.291	5.214	5.764	5.137	5.136
CT	4.836	4.840	4.857	4.875	8.302	4.762	4.809
MR	4.654	4.702	4.835	4.833	9.382	4.594	4.783
US	2.472	2.564	2.478	2.471	2.722	2.439	2.462
mini-MIAS	1.390	1.378	1.391	1.391	1.430	1.358	1.352
Average	1.877	1.853	1.874	1.865	2.016	1.826	1.822

Table 10. Average compression results, in bits per pixel using the methods described in Sect. 4, BPD (bitplane decomposition), BFS (background-foreground separation), HC (histogram compaction), SBR (scalable bitplane reduction), SBC-Mix (simple bitplane coding-mixture), and BTD (binary-tree decomposition), for the medical image data sets

		Compression methods						
	Data sets	**BPD**	**BFS**	**HC**	**SBR**	**SBC**	**SBC-Mix**	**BTD**
RNAi D1	| (DNA)	6.250	6.293	6.250	6.255	6.225	**6.179**	6.186
	| (PH3)	5.559	5.570	5.570	5.569	5.571	**5.545**	5.555
	| (Actin)	6.255	6.311	6.244	6.258	6.220	**6.185**	6.205
RNAi D2	| (DNA)	5.959	5.986	5.968	5.968	5.960	**5.918**	5.930
	| (PH3)	5.537	5.547	5.539	5.540	5.545	**5.518**	5.525
	| (Actin)	5.919	5.948	5.904	5.928	5.904	**5.876**	5.892
RNAi D3	| (DNA)	6.175	6.213	6.169	6.164	6.145	**6.106**	6.111
	| (PH3)	5.580	5.598	5.584	5.585	5.590	**5.564**	5.572
	| (Actin)	6.313	6.366	6.299	6.306	6.277	**6.236**	6.247
RNAi D4	| (DNA)	6.309	6.355	6.312	6.317	6.289	**6.239**	6.244
	| (PH3)	5.577	5.593	5.582	5.584	5.587	**5.561**	5.570
	| (Actin)	6.423	6.484	6.412	6.419	6.395	**6.348**	6.363
RNAi D5	| (DNA)	6.195	6.233	6.194	6.198	6.166	**6.126**	6.133
	| (PH3)	5.558	5.568	5.564	5.567	5.569	**5.543**	5.552
	| (Actin)	6.299	6.355	6.286	6.298	6.263	**6.224**	6.240
RNAi all	| (DNA)	6.178	6.216	6.179	6.182	6.157	**6.113**	6.121
	| (PH3)	5.562	5.575	5.568	5.569	5.573	**5.546**	5.555
	| (Actin)	6.242	6.293	6.229	6.242	6.212	**6.174**	6.189
Average		5.994	6.028	5.992	5.998	5.980	**5.945**	5.955

Table 11. Average compression results, in bits per pixel using the methods described in Sect. 4, BPD (bitplane decomposition), BFS (background-foreground separation), HC (histogram compaction), SBR (scalable bitplane reduction), SBC-Mix (simple bitplane coding-mixture), and BTD (binary-tree decomposition), for the RNAi image data sets

Finally, in Table 11, we present the results for the RNAi images, and in this case, method SBC-Mix is the one that attained the best compression results, almost 3% better than JPEG-LS. The difference between the SBC-Mix and BTD is very low (close to 0.01 bpp). In our opinion, it seems that the methods described in Sect. 4 attain better compression results when compared to JPEG-LS for images with higher GI values (sparse images). This is the reason why the binary-tree decomposition approach attains between 9% to 11% better results than JPEG-LS for the microarray and medical images. On the other hand, the RNAi images have a lower GI (around 0.17), therefore the lower performance of our methods when compared to JPEG-LS.

In Figs. 15 and 16, we can see the encoding and decoding speed in kilobytes per second for the compression methods described in Sect. 4. In the encoding phase, it seems that the method based on binary-tree decomposition is the fastest one among all the others for all data sets. In the decoding phase, the fastest method is different from data set to data set. For the microarray and medical image data sets, the methods based on bitplane decomposition are faster than the method based on binary-tree decomposition. On the other hand, for the RNAi images, the method based on binary-tree decomposition is faster than all the others. The SBC-Mix method is the slowest one in all data sets. The lower decoding speed for the SBC-Mix is due to the mixture procedure. For each symbol that is processed, the model weights need to be updated, which is a very time-consuming task.

3.3. Rate distortion study

The two decomposition approaches described in Sect. 4 have progressive decoding capabilities (also known as lossy to lossless). Due to that, it is important to understand what is the error induced in the image during the decoding phase when only part of the data is decoded. This characteristic allows also to obtain the original image without any loss.

In Figs. 15 and 16, we present the rate distortion curves of two images, "array1" from the *YouLou* data set and "cr_17218" from the medical *CR* data set, according to two metrics: L_2-norm (root-mean-square error or RMSE) and L_∞-norm (maximum absolute error or MAE). We used JBIG, JPEG2000, the approach based on bitplane decomposition [6], and the other based on binary-tree decomposition [8].

Regarding Fig. 15, we can notice that the method based on binary-tree decomposition outperforms all the other methods in terms of L_2-norm and L_∞-norm. JBIG and method [6] have a similar performance in terms of rate distortion. The JPEG2000 standard attains globally worse results in terms of L_2. Furthermore, we can notice a sudden degradation of the rate distortion, for higher bitrates L_∞. We believe that this phenomenon is probably related to the default parameters used by the encoder, which might not be well suited for microarray images.

In Fig. 16, we can see in the first plot, related to the L_2-norm, that method [6] attains the worse distortion results. JBIG is slightly better but worse than JPEG2000 and the method [8]. We can also see that JPEG2000 outperforms all the other methods in terms of L_2-norm for lower bitrates. For bitrates higher than 2 bpp, the best method is the one based on binary-tree decomposition. For the L_∞-norm, JPEG2000 and method [6] are the ones that attained the

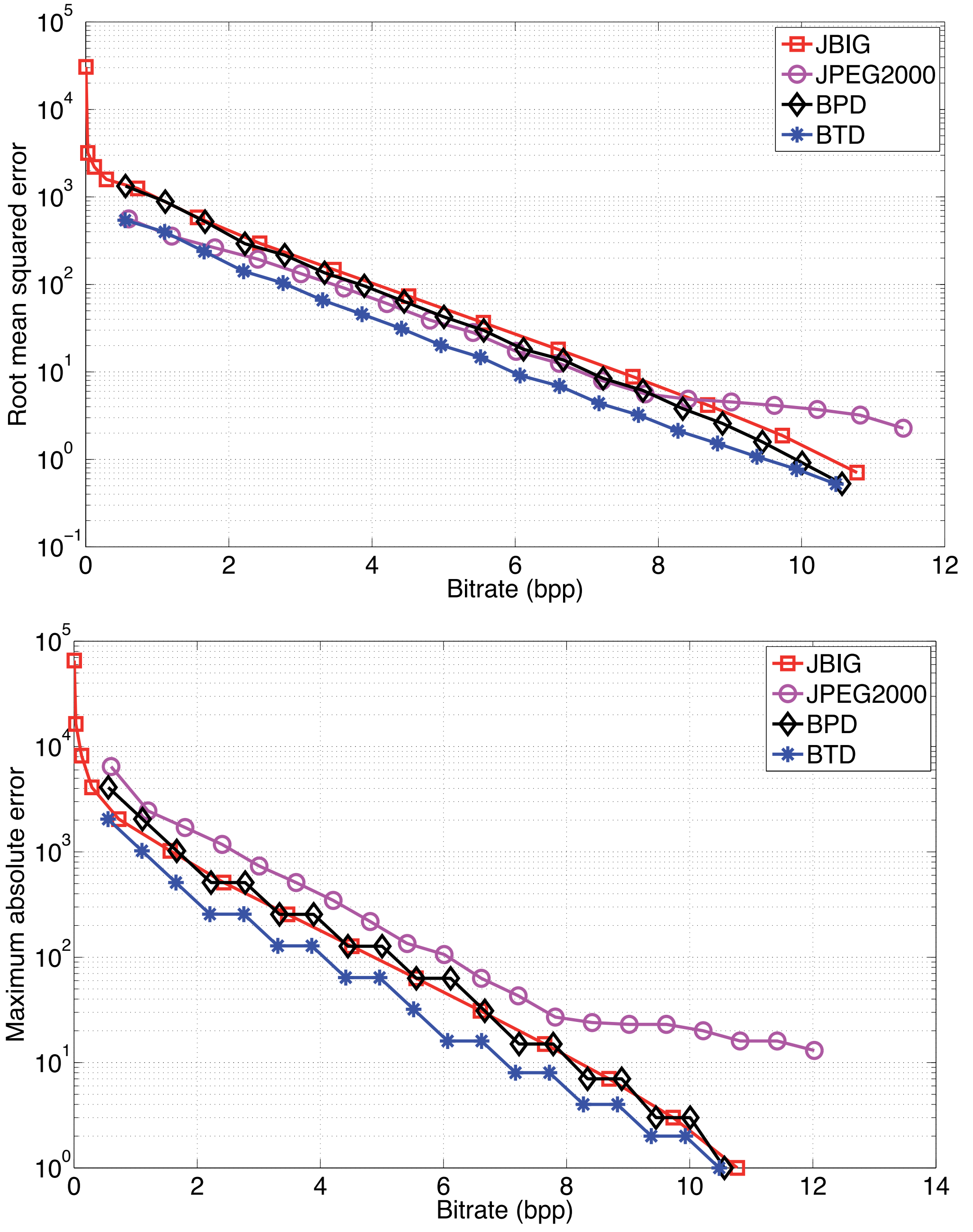


Figure 15. Rate distortion curves showing the performance of the bitplane decomposition (BPD), the binary-tree decomposition (BTD), and the JPEG2000, and JBIG standards, in lossy-to-lossless mode for the microarray image "array1" from the *YuLou* data set. Results are given both for the L_2 (root-mean-square error or RMSE) and L_∞ (maximum absolute error or MAE) norms

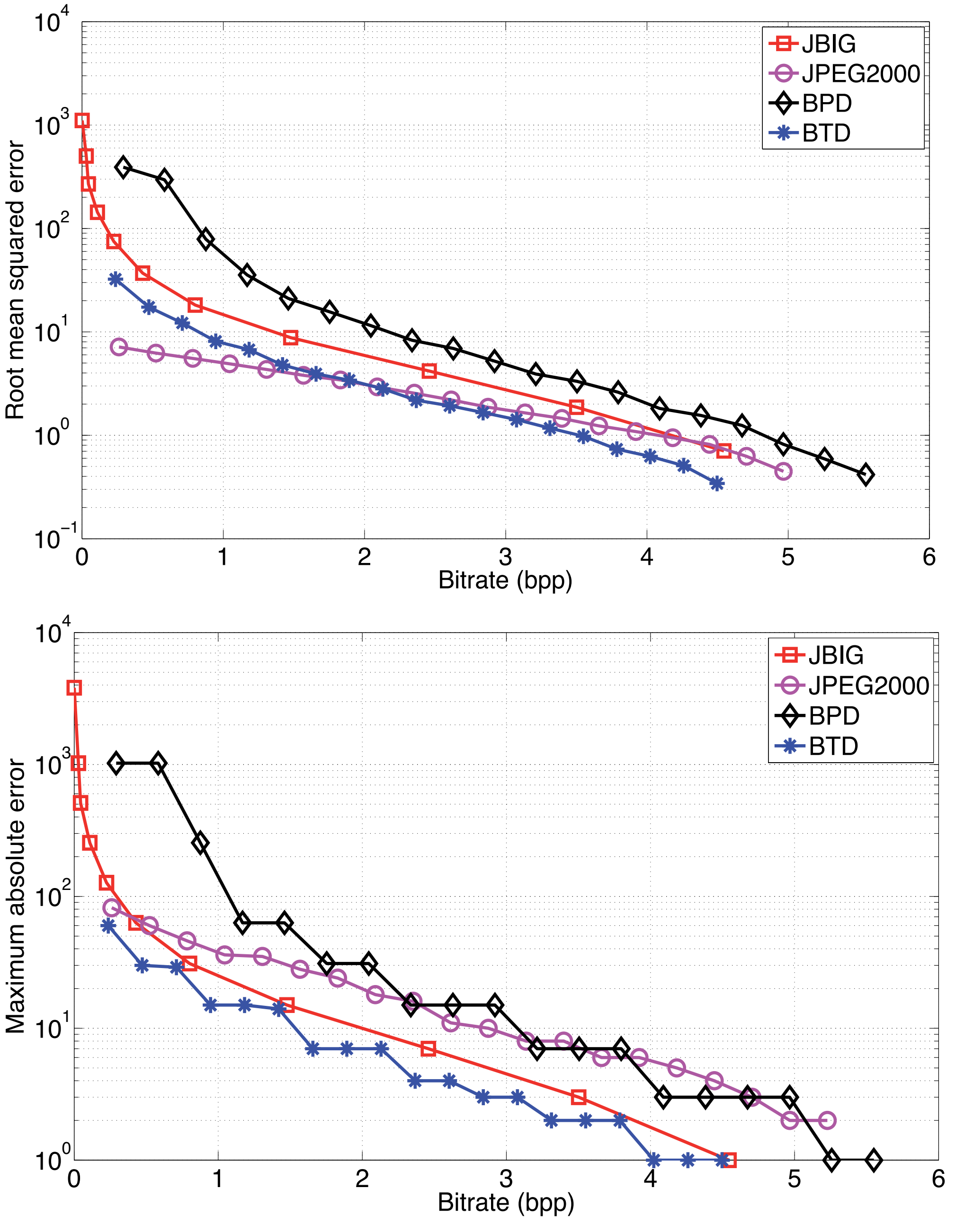


Figure 16. Rate distortion curves showing the performance of the bitplane decomposition (BPD), the binary-tree decomposition (BTD), and the JPEG2000 and JBIG standards, in lossy-to-lossless mode for the medical image "cr_17218" from the *CR* data set. Results are given both for the L_{∞} (root-mean-square error or RMSE) and $(>8\,\mathrm{bpp})$ (maximum absolute error or MAE) norms

worse results. On the other hand, method [8] outperforms all the others, regardless of the bitrate. We did not include results for other images to avoid extending this chapter, but the results obtained with the medical images are very similar with the results obtained using RNAi images. On the other hand, we can conclude that the rate distortion results for the microarray images are not similar to the other data set types (medical and RNAi).

4. Conclusions

The use of biomedical imaging has increased in the last years. These biomedical images include microarray, medical, and RNAi images. In this chapter, we have studied two image decomposition approaches to be applied to this type of images in order to compress them efficiently.

We presented a set of comprehensive results regarding the lossless compression of biomedical images using general coding methods (e.g., gzip), image coding standards (e.g., JPEG2000), and the two image decomposition approaches that rely on finite-context models and arithmetic coding.

From the obtained experimental results, we conclude that JPEG-LS gives the best compression results, among the image coding standards, but lacks lossy-to-lossless capability, which may be a decisive functionality if remote transmission over possibly slow links is a requirement.

We developed compression algorithms based on two decomposition approaches: bitplane decomposition and binary-tree decomposition. In the bitplane decomposition, we also used segmentation, bitplane Reduction, and an approach based on bit modeling by the pixel value estimates. All the developed methods allow lossy-to-lossless compression and are based on finite-context models and arithmetic coding. According to the obtained results, the approach based on binary-tree decomposition seems to be the one that attains the best compression results. The only exceptions are the RNAi images, for which the best method is the one based on a mixture of finite-context models.

In terms of coding speed, the compression algorithm based on a binary-tree decomposition seems to be the fastest one among all the others in the encoding phase. In the decoding stage, the approaches based on bitplane decomposition seem to be faster for the microarray and medical image data sets. For the RNAi images, the method based on binary-tree decomposition seems to be the fastest one.

A rate distortion study was also performed and, according to the obtained results, it seems that the method based on binary-tree decomposition attains the best results for the majority of the cases. The results obtained by the developed compression algorithms have been compared with general-purpose compression methods and also with image coding standards. According to the results, we can conclude that these compression methods have better compression performance in all the image test sets used.

http://dx.doi.org/10.5772/60650

Acknowledgements

This work was developed at IEETA (Institute of Electronics and Telematics Engineering of Aveiro).It was partially supported by FEDER, through the Operational Program Competitiveness Factors (COMPETE), and by National Funds through FCT (Foundation for Science and Technology) in the context of a PhD Grant (FCT reference SFRH/BD/85855/2012) and projects PEst-OE/EEI/UI0127/2014 and Incentivo/EEI/UI0127/2014.

Author details

Luís M. O. Matos*, António J. R. Neves and Armando J. Pinho

*Address all correspondence to: luismatos@ua.pt

IEETA/DETI, University of Aveiro, Aveiro, IEETA/DETI, Portugal

References

[1] A. Nait-Ali and C. Cavaro-Menard. *Compression of Biomedical Images and Signals*. Wiley-ISTE, 1st edition, 2008.

[2] P. Hegde, R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J. Earle-Hughes, E. Snesrud, N. Lee, and John Q. A concise guide to cDNA microarray analysis. *Biotechniques*, 29(3):548–562, September 2000.

[3] S. K. Moore. Making chips to probe genes. *IEEE Spectrum*, 38(3):54–60, March 2001.

[4] C. Falschlehner, S. Steinbrink, G. Erdmann, and M. Boutros. High-throughput RNAi screening to dissect cellular pathways: a how-to guide. *Biotechnology Journal*, 5(4): 368–376, 2010.

[5] D. S. Taubman and M. W. Marcellin. *JPEG2000: Image Compression Fundamentals, Standards and Practice*. Kluwer Academic Publishers, 2002.

[6] A. J. R. Neves and A. J. Pinho. Lossless compression of microarray images using image-dependent finite-context models. *IEEE Transactions on Medical Imaging*, 28(2):194–201, February 2009.

[7] A. J. Pinho and A. J. R. Neves. Progressive lossless compression of medical images. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-2009*, Taipei, Taiwan, April 2009.

[8] L. M. O. Matos, A. J. R. Neves, and A. J. Pinho. Compression of microarrays images using a binary-tree decomposition. In *Proceedings of the 22th European Signal Processing Conference, EUSIPCO-2014*, pages 531–535, Lisbon, Portugal, September 2014.

[9] MENT, 2014. Last accessed on February, 2015. http://bioinformatics.ua.pt/software/ment.

[10] A. J. R. Neves and A. J. Pinho. Compression of microarray images. In S. Miron, editor, *Signal Processing*, pages 429–448. INTECH, March 2010.

[11] A. J. Pinho and A. J. R. Neves. L-infinity progressive image compression. In *Proceedings of the Picture Coding Symposium, PCS-07*, Lisbon, Portugal, November 2007.

[12] A. J. Pinho and A. J. R. Neves. Lossy-to-lossless compression of images based on binary tree decomposition. In *Proceedings of the IEEE International Conference on Image Processing, ICIP-2006*, pages 2257–2260, Atlanta, Georgia, USA, October 2006.

[13] A. J. Pinho, D. Pratas, and P. J. S. G. Ferreira. Bacteria DNA sequence compression using a mixture of finite-context models. In *Proceedings of the IEEE Workshop on Statistical Signal Processing*, Nice, France, June 2011.

[14] A. J. Pinho, A. J. R. Neves, D. A. Martins, C. A. C. Bastos, and P. J. S. G. Ferreira. Finite-context models for DNA coding. In S. Miron, editor, *Signal Processing*, pages 117–130. INTECH, March 2010.

[15] A. J. Pinho, A. J. R. Neves, V. Afreixo, Carlos A. C. Bastos, and P. J. S. G. Ferreira. A three-state model for DNA protein-coding regions. *IEEE Transactions on Biomedical Engineering*, 53(11):2148–2155, November 2006.

[16] T. C. Bell, J. G. Cleary, and I. H. Witten. *Text Compression*. Prentice Hall, 1990.

[17] K. Sayood. *Introduction to Data Compression*. Morgan Kaufmann, 3rd edition, 2006.

[18] D. Salomon. *Data Compression - The Complete Reference*. Springer, 4th edition, 2007.

[19] J. Rissanen and G. G. Langdon, Jr. Arithmetic coding. *IBM Journal of Research and Development*, 23(2):149–162, March 1979.

[20] David Huffman. A method for the construction of minimum-redundancy codes. *Proceedings of the IRE*, 40(9):1098–1101, September 1952.

[21] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice Hall, 2nd edition, 2002.

[22] P. Buonora and F. Liberati. A format for digital preservation of images: a study on JPEG 2000 file robustness. *D-Lib Magazine*, 14(7/8), August 2008.

[23] Y. Yoo, Y. G. Kwon, and A. Ortega. Embedded image-domain compression using context models. In *Proceedings of the IEEE International Conference on Image Processing, ICIP-99*, volume I, pages 477–481, Kobe, Japan, October 1999.

[24] A. J. R. Neves and A. J. Pinho. Lossless compression of microarray images. In *Proceedings of the IEEE International Conference on Image Processing, ICIP-2006*, pages 2505–2508, Atlanta, Georgia, USA, October 2006.

[25] A. Neekabadi, S. Samavi, S. A. Razavi, N. Karimi, and S. Shirani. Lossless microarray image compression using region based predictors. In *Proceedings of the IEEE International Conference on Image Processing, ICIP-2007*, volume 2, pages 349–352, San Antonio, Texas, USA, September 2007.

[26] Y. Yoo, Y. G. Kwon, and A. Ortega. Embedded image-domain adaptive compression of simple images. In *Proceedings of the 32nd Asilomar Conference on Signals, Systems, and Computers*, volume 2, pages 1256–1260, Pacific Grove, California, USA, November 1998.

[27] H. Kikuchi, K. Funahashi, and S. Muramatsu. Simple bit-plane coding for lossless image compression and extended functionalities. In *Proceedings of the Picture Coding Symposium, PCS-09*, pages 1–4, Chicago, Illinois, USA, May 2009.

[28] H. Kikuchi, R. Abe, and S Muramatsu. Simple bitplane coding and its application to multi-functional image compression. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, E95.A(5):938–951, 2012.

[29] H. Kikuchi, T. Deguchi, and M. Okuda. Lossless compression of LogLuv32 HDR images by simple bitplane coding. In *Picture Coding Symposium, PCS-2013*, pages 265–268, San Jose, California, USA, December 2013.

[30] S. Baase and A. V. Gelder. *Computer Algorithms: Introduction to Design and Analysis*. Addison Wesley, 3rd edition, November 1999.

[31] X. Chen, S. Kwong, and J.-F. Feng. A new compression scheme for color-quantized images. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(10):904–908, October 2002.

[32] A. J. Pinho and A. J. R. Neves. A context adaptation model for the compression of images with a reduced number of colors. In *Proceedings of the IEEE International Conference on Image Processing, ICIP-2005*, volume 2, pages 738–741, Genova, Italy, September 2005.

[33] mini-MIAS, 1995. Last accessed on February, 2015. http://peipa.essex.ac.uk/pix/mias.

[34] D. Zonoobi, A. A. Kassim, and Y. V. Venkatesh. Gini index as sparsity measure for signal reconstruction from compressive samples. *IEEE Journal of Selected Topics in Signal Processing*, 5(5):927–932, September 2011.

[35] ApoA1, 2001. Last accessed on February, 2015. http://www.stat.berkeley.edu/ terry/zarray/Html/apodata.html or http://sweet.ua.pt/luismatos/microarrays/apoa1.html.

[36] Arizona, 2011. Last accessed on February, 2015. http://deic.uab.es/ mhernandez/media/imagesets/arizona.tar.bz2 or http://sweet.ua.pt/luismatos/microarrays/arizona.html.

[37] IBB, 2013. Last accessed on February, 2015. http://deic.uab.es/ mhernandez/media/imagesets/ibb.tar.bz2 or http://sweet.ua.pt/luismatos/microarrays/ibb.html.

[38] ISREC, 2001. Last accessed on February, 2015. http://ccg.vital-it.ch/DEA/module8/P5_chip_image/images or http://sweet.ua.pt/luismatos/microarrays/isrec.html.

[39] Omnibus-LM, 2006. Last accessed on February, 2015. http://deic.uab.es/ mhernandez/media/imagesets/omnibus.txt or http://sweet.ua.pt/luismatos/microarrays/omnibus-lm.html.

[40] Omnibus-HM, 2006. Last accessed on February, 2015. http://deic.uab.es/ mhernandez/media/imagesets/omnibus.txt or http://sweet.ua.pt/luismatos/microarrays/omnibus-hm.html.

[41] Stanford, 2001. Last accessed on February, 2015. ftp://smd-ftp.stanford.edu/pub/smd/transfers/Jenny or http://sweet.ua.pt/luismatos/microarrays/stanford.html.

[42] Yeast, 1998. Last accessed on February, 2015. http://genome-www.stanford.edu/cellcycle/data/rawdata/individual.html or http://sweet.ua.pt/luismatos/microarrays/yeast.html.

[43] YuLou, 2004. Last accessed on February, 2015. http://www.cs.ucr.edu/yuluo/MicroZip or http://sweet.ua.pt/luismatos/microarrays/yulou.html.

[44] MEDNAT, 2004. Last accessed on February, 2015. http://sun.aei.polsl.pl/ rstaros/mednat.

[45] BBBC017v1, 2006. Last accessed on February, 2015. http://www.broadinstitute.org/bbbc/BBBC017.

Section 3

Role of DSP in Power Conversion Systems

Chapter 6

Application of DSP in Power Conversion Systems — A Practical Approach for Multiphase Drives

Hugo Guzman, Mario Bermúdez, Cristina Martín,
Federico Barrero and Mario Durán

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/60450

Abstract

Digital Signal Processing is not a recent research field, but has become a powerful technology to solve engineering problems in the last few decades due to the introduction by Texas Instruments in 1982 of the Digital Signal Processor. Fast digital signal processors have quickly become a cornerstone of high-performance electrical drives, where power electronic conversion systems have heavy online computation burdens and must be controlled using complex control algorithms. In this sense, multiphase drives represent a particularly interesting case of study, where the computational cost highly increases with each extra phase. This technology has been recognized in recent times as an attractive electrical drive due to its usefulness in traction, more-electric aircraft applications and wind power generation systems. However, the complexity of the required control algorithms and signal processing techniques notably increases in relation with conventional three-phase drives. This chapter makes a revision of the necessities of a high-performance multiphase drive from the digital signal processing perspective. One of the most powerful Texas Instruments' digital signal processor (TMS320F28335) is used, and specific control algorithms, electronic circuits and acquisition processing methods are designed, implemented and analyzed to show its interest in the development of a high-performance multiphase drive.

Keywords: Digital Signal Processing, Digital Signal Processors, Multiphase Electrical Drives, Field-Oriented Control, Predictive Current Control Techniques

1. Introduction

Digital Signal Processing in Power Electronics is an important research area which mainly covers problems concerning the design and realization of algorithms using Digital Signal Processors (DSPs). A wide variety of solutions can be found at industry and in research fields involving DSP applications, mostly in accordance to conventional three-phase systems. This chapter will analyze the interest and utility of Digital Signal Processing in the specific field of multiphase electric drives. This area has significantly accelerated in the last decade due to their intrinsic features like higher fault tolerance, reduced torque ripple or better power splitting. With the number of conventional electrical drives continuously growing, the interest in multiphase ones is also rising, although they are not yet so common due to their complexity. The development of modern power electronic switches and the ability of modern DSPs to implement complex algorithms are allowing the control of multiphase drives in applications such as aerospace actuators, wind energy conversion systems, oil pumping or ultrahigh-speed elevators.

Power conversion systems are composed of different electrical and electronic components that need to be managed following specific constraints depending on the final application. Moreover, such systems are designed to work under different conditions and states; thus, several control algorithms can be found in a single DSP, and the designed control unit (based on the DSP) must be able to change between these different states, considering measured components and actuating on the overall equipment to work as desired. For this purpose, Digital Signal Processing based on DSPs must be done considering not only the inner control algorithm implementation, but also the proper signal adaptation going in or out the DSP.

The major benefit to readers of this chapter is the acquisition of specific knowledge concerning the processing of signals from current and/or speed sensors using a digital signal processor and the subsequent generation of the signals needed to control the power switches in a multiphase power converter, following the implemented controller commands. Hardware and software design for the specific application are included as illustrative examples for the considered systems.

The chapter is organized as follows:

First of all, the overall system features and constraints are presented, demonstrating the necessity for digital signal processing and the implementation of DSPs for such application. The specific characteristics needed for the multiphase drive application are also presented and the selected DSP (TMS320F28335) is described. Subsequently, the electrical and electronic hardware is explained and the acquisition/adaptation PCBs, developed in order to communicate the DSP with the overall system, are presented. The next section explains the software designed to control the power converters and electrical equipment, the current and speed sensors measurement and treatment, the implemented controller algorithm and the semiconductor switching signals. Finally, some results are depicted and conclusions are presented.

http://dx.doi.org/10.5772/60450

2. Power conversion systems using multiphase induction machines

Although the deadline for a complete replacement of petroleum-based fuels is uncertain, it is clear that the human being is in a race against time to replace conventional energy sources. In the last decade, the technology evolution in areas such as renewable energies or electric propulsion has been intense to achieve this goal [1]. Government policies are promoting the use of renewable energy sources (wind, photovoltaic) and new transportation systems (hybrid and electrical vehicles) aiming to fulfill environmental, economical and social needs [2]. From the technological point of view, the present time is a moment of challenge and the success of a technological approach will depend on the degree in which the society demands are satisfied. As far as the electrical vehicles (EVs) are concerned, where digital signal processing and power electronic fields are finding an important niche of application, some of the desirable features of an automotive drive include:

- High overall efficiency over wide speed and torque range.
- Fast torque response for vehicle acceleration.
- High reliability in different operation conditions.
- Safety and fault tolerance.
- Low maintenance and improved robustness.
- Reasonable cost.

These general characteristics are advantageous regardless of the wide range of EVs available in the market, which includes battery-powered electric vehicles (BEV) and plug-in electric vehicles (PHEV), hybrid electric vehicles (HEV) or fuel-cell electric vehicles (FCEV) [3]. Hybrid topologies combining internal combustion engine (ICE) and electrical motors (EM) can be classified into series, parallel and series-parallel types. Each vehicle topology has its own specific features and the general control of the vehicle can follow different approaches [4]. Nevertheless, the structure of the electric propulsion system is similar in all topologies and essentially consists of a battery, an electronic converter, an electric motor, and a speed and/or torque sensor [1]. Figure 1 presents a general schematic of an electric propulsion drive.

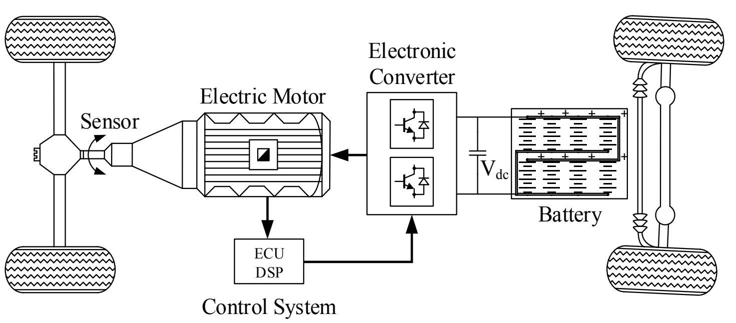


Figure 1. Example of an electric vehicle propulsion system.

The propulsion system is completed with the mechanical part, including transmission and wheels. The energy conversion process is controlled by the electric drive, which is the core of the electric propulsion system. In motor mode, the energy comes from the batteries to the wheels in a controlled manner by means of drive control, while in regenerative braking the energy is reversed back to the batteries. Failure of one of the drive phases may result in a complete breakdown of the propulsion system unless a proper post-fault control is applied to improve the vehicle reliability.

The electric propulsion has been traditionally obtained from conventional DC machines but nowadays three-phase AC machines are more appropriate solutions due to their lower cost and higher reliability [5]. Consequently, the standard option for the drive in the propulsion system is the use of three-phase induction or permanent magnet synchronous motors (PMSM), two-level voltage source inverters (VSI's), field-oriented control (FOC) and pulse width modulation (PWM). However, many other options can be explored for optimal drive design:

- Motor type: PMSM provide higher torque density, small size and weight, and good efficiency [6], but squirrel cage induction motors possess interesting features for electric propulsion of EVs including low cost, low maintenance and ruggedness [5].
- Number of phases: Three-phase motors are available off-the-shelf and can benefit from the economy of scales. However, multiphase motors can provide higher robustness due to its inherent fault-tolerant characteristic, lower noise and vibrations [7], higher torque density by current harmonic injection [8], lower per phase current rating [9] or ripple-free post-fault drive operation [10].
- Converter type: Although matrix [11] or multilevel [12] converters have been proposed for EV applications, the standard option is the use of a two-level VSI. This option becomes natural for EVs with multiphase motors due to the Amps per phase current reduction. DC buck/boost converters can also be included into the drive topology [1], but the motor and inverter control are not affected.
- Control strategy: The electric drive control can be implemented using any of the extensive options available in literature, from the simple scalar control to artificial-intelligence-based controls [5]. Control schemes based on model reference adaptive control (MRAC) or sliding mode control (SMC) can be found in [13], and popular control strategies like the field-oriented control (FOC) or the direct torque control (DTC) have been addressed in [12]. Nevertheless, the finite-control set model based predictive control (FCS-MPC) has been introduced in the last decade and it is a promising candidate to replace standard schemes (or to become a part of other existing methods like FOC) due to its simplicity and good performance [14-16]. The increment in the available computing power of modern microprocessors makes this strategy now plausible for controlling conventional and multiphase drives, and a wide variety of predictive current controllers have been recently presented [17-21].

The final decision on a particular topology is always a trade-off between different desirable features. The topology explored in this work uses a five-phase induction motor supplied from a two-level inverter and controlled with a predictive control strategy, to promote the reliability, efficiency and performance characteristics of the propulsion drive (Fig. 2).

http://dx.doi.org/10.5772/60450

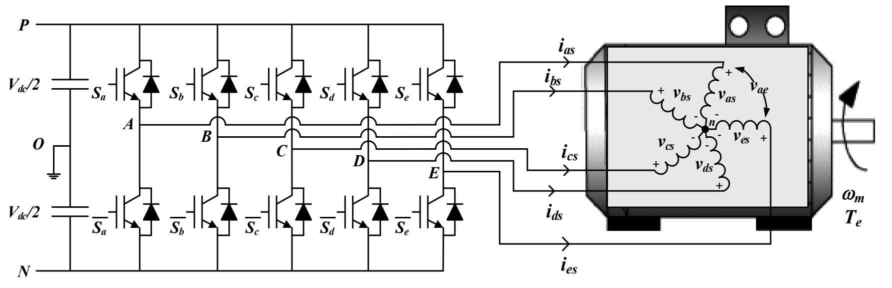


Figure 2. Schematic diagram of a five-phase two-level induction motor drive.

The key features of the proposed propulsion system and drive control include:

- Improved efficiency with lower inverter losses and motor losses (predictive controller and multiphase motor drive).
- High robustness and low maintenance (induction motor).
- Fast torque response (predictive control).
- Fault-tolerant operation (multiphase motor).

3. Hardware description: Power converter and electronic control unit

An experimental test rig has been implemented to emulate the aforementioned multiphase propulsion drive. This test-bench is shown in Fig. 3. The system is composed of a symmetrical five-phase induction motor (IM) with distributed windings that has been constructed rewinding a conventional three-phase IM. The electrical drive is mechanically coupled to a DC motor, which can provide a programmable mechanical load torque to the system. Notice that each motor (DC motor and 5-phase IM) is controlled independently using two different power converters. An incremental encoder is also coupled to the shaft to measure the rotational speed.

The five-phase motor is driven by two conventional three-phase two-level voltage source inverters (VSIs), with an independent external DC power supply as the DC-Link. A DSP-based Electronic Control Unit (ECU) is used in order to control the power converters switching sequence, depending on the demanded control action. Switching signals are sent from the ECU to the power converter drivers, being 0 V for OFF state and 15 V for ON state. For control purposes, four phase currents are measured with hall-effect sensors included on the power converters and the remaining fifth current is estimated considering that the IM is arranged with only one isolated neutral point. The hall-effect current sensors allow measurements in a range of ±25 A. Each semiconductor driver provides an error signal that triggers ON (0 V) under low- or high-voltage condition and overcurrent operation. The aforementioned signals are classified as input or output signals of the ECU. Switching signals are regarded as output signals while error signals and current/speed measurements are considered as input signals.

The ECU is connected to a host PC using a standard RS232 cable. Then, the user of the system can manage the entire system using a developed host PC software that governs the ECU. The software that runs in the DSP, configures its internal peripherals, the communication protocol, the data acquisition system, and the control algorithm is programmed using the DSP's manufacturer proprietary software.

The measured speed and current signals as well as the digital errors obtained from the power converter (i.e., errors generated in the semiconductors' drivers) are received and processed in the DSP. If no error is detected, the processor executes the control algorithm and provides the switching signals to the power converters semiconductors (the implemented algorithm will be described in the next section).

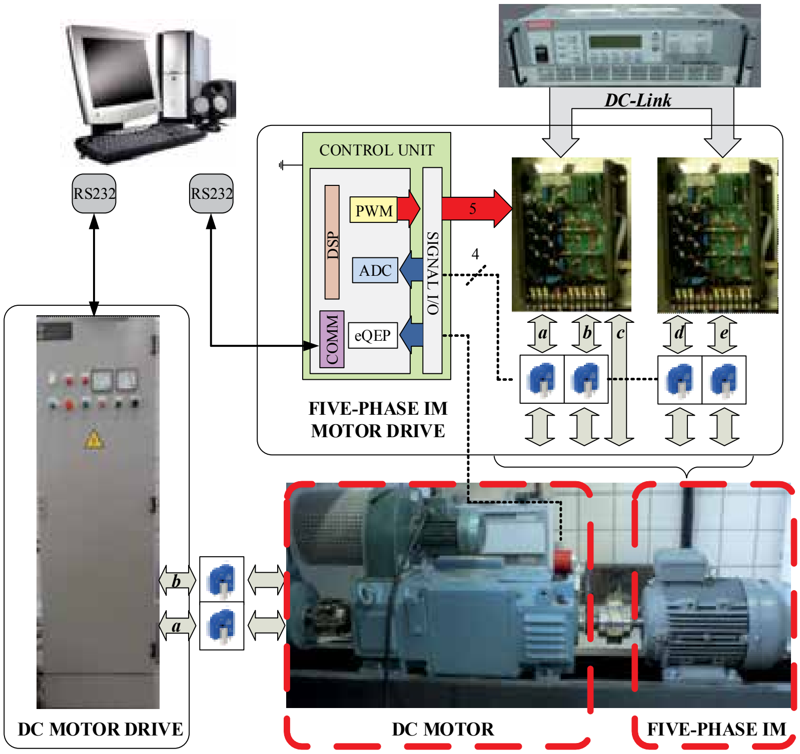


Figure 3. Multiphase drive experimental test-bench.

A summary of the described input and output (I/O) signals is provided in Table 1, where their electrical characteristics are also detailed.

In order to properly select a DSP among those commercially available, it is necessary to identify the number and type of input/output signals that the system will handle. A more detailed scheme of the DSP, including the managed I/O signals, is presented in Fig. 4. As schematically

Name	Quantity	Characteristics	Type
Phase Current	4	±25 A	Input
DC-Link Voltage	1	300 V	–
Rotational Speed	1	±1000 rpm	Input
Driver Error	5	0/15 V	Input
General Purpose Output	5-	–	Output
Semiconductor switching signals	5	0/15 V	Output

Table 1. Electronic Control Unit I/O signals and their electrical characteristics.

shown, error and measured signals (currents and speed) correspond to DSP's inputs, while communication, semiconductor switching and auxiliary general purpose signals are considered as outputs. In what follows, a short summary of I/O signals of the ECU is presented.

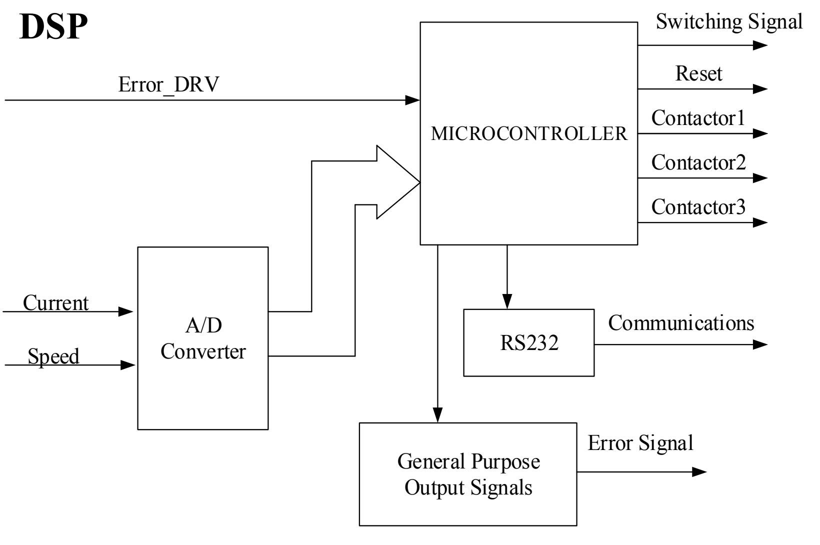


Figure 4. Signal requirements for the ECU.

The semiconductors' drivers possess an error signal that is automatically triggered in case of an excessive temperature, overcurrent or overvoltage condition on the power converter. In case that an abnormal working condition is detected, the ECU should interrupt the power converter switching process, avoiding further damage in the electric and electronic system.

Due to the analog nature of the measured signals (remember that currents and speed sensors are required), an Analog-to-Digital Converter (ADC) stage is necessary before providing measured values to the processor. Such stage is commonly found on commercially available DSPs, being only necessary to select its working conditions and enabling signals.

The semiconductor switching signals are provided from the DSP to the drivers by means of the enhanced Pulse Width Modulator (ePWM) modules. Therefore, for the designed experimental test-rig, the DSP must include at least five ePWM modules in order to control the power converter connected to the five-phase electrical drive.

Different auxiliary signals, found on the General Purpose Input Output (GPIO) module, are considered for reset, system configuration and initialization purposes. For instance, the reset signal is used to restart the system in the presence of a system error or if it is demanded by the test-rig operator. Regarding the system configuration, three general purpose (GPIO) signals are used to manage and control the power converter through electric contactors. Remaining GPIOs are used to turn ON/OFF warnings and error light indicators and to trigger external measurement equipment such as an oscilloscope.

The chosen DSP must also manage the communication protocol RS232 in order to allow an external operator to control the system using an implemented Human Machine Interface (HMI) software that runs in a personal computer. Consequently, a Serial Communication (SCI) module must be integrated in the ECU.

Taking into account the previous requirements, the TMS320F28335 Texas Instruments DSP has been selected among other devices. The characteristics of the TMS320F28335 are summarized in Table 2.

Name	Characteristics
Power supply	5 V
Output voltage range	-0.3 to 4.6 V
Input voltage range	-0.3 to 4.6 V
High-level input/output voltage	3.3 V
Low-level input/output voltage	0 V
Floating-point Unit	Yes
PWM outputs	6
32-Bit Capture inputs of auxiliary PWM outputs	6
12-Bit Analog-to-Digital Converter (ADC)	16
Serial Peripheral Interface (SPI)	1
Serial Communications Interface (SCI)	3
General purpose I/O pins	88

Table 2. Texas Instruments TMS320F28335 DSP characteristics.

Notice that the DSP I/O signal voltage threshold is between 0/3.3 V, while the rest of the signals of the system work at higher voltage levels. Consequently, it is necessary to implement some voltage adaptation stages between the different system signals and the DSP. In what follows, an explanation of the different adaptation stages for current, speed, error and switching signals is presented.

a. Measured current adaptation stage

As mentioned above, the measuring range of hall-effect sensors is ±25 A, being the maximum/ minimum possible current to be measured on a power converter phase. Moreover, the hall-effect sensor has a transformation ratio of 1/1000, which means that the output current of the sensor is ±25 mA. The aim of the current adaptation circuit is to transform the current measurement into an equivalent voltage signal within the allowed DSP I/O threshold. The electrical circuit schematic of the adaptation stage is presented in Fig. 5, where the hall-effect sensor is represented by an ideal current source. In general, the designed circuit contains the following stages:

- Transformation of the measured current into an equivalent voltage by means of the measurement resistance (R_M).
- Voltage divider and low-pass active filter, reducing measurement noise.
- Signal voltage adaptation stage (0/3.3 V).
- Safety diodes that prevent the output signal from exceeding the voltage limits.

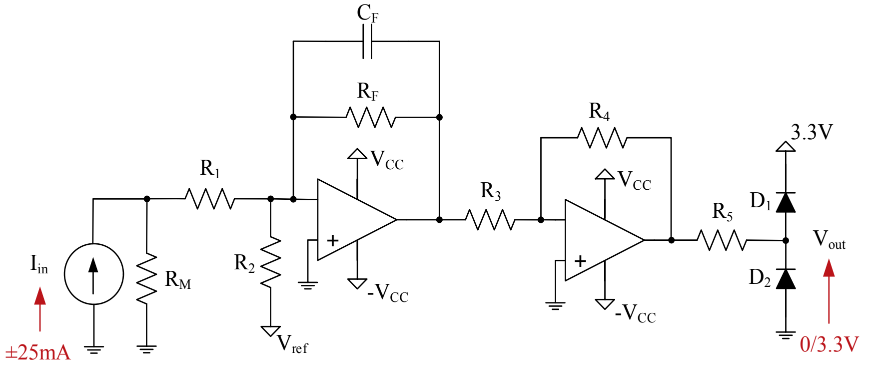


Figure 5. Schematic of the measured current adaptation circuit.

b. Measured speed adaptation stage

The test-rig speed is measured by means of an incremental encoder from the manufacturer Hohner with reference 10-11657-2500. This device measures speeds between ±1000 rpm and gives two types of squared waves that are 90 electrical degrees out of phase, which are usually called channel A and channel B. The reading of one channel provides information about the speed of rotation, while the acquisition of the second channel allows obtaining the rotational direction. Another signal called channel I (or index) is also available, which gives the position of absolute zero on the encoder shaft. This signal is a squared impulse with the phase and width centered on channel A. The set of three signals (Fig. 6) has an equivalent output voltage of ±5 V and must be adapted to the DSP's voltage threshold (0/3.3 V).

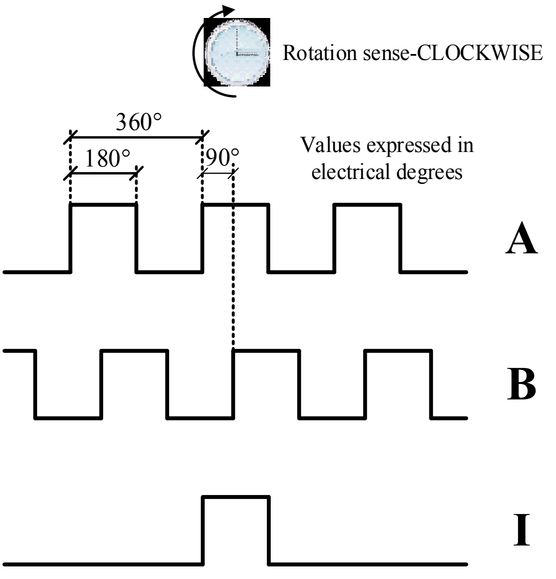


Figure 6. Graphic representation of the A, B and I incremental signals obtained from the Hohner's encoder with reference 10-11657-2500.

The speed measurement adaptation circuit shown in Fig. 7 consists of the following stages:

- Optocoupler circuit that allows the electrical isolation of the speed sensor and the electronic control unit.
- Voltage divider.
- Unity gain buffer amplifier for circuit stage isolation.
- Safety diodes that prevent the output signal from exceeding the voltage limits.

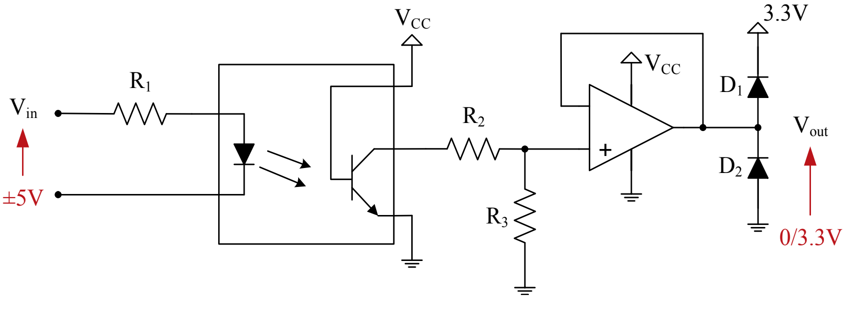


Figure 7. Schematic of the measured speed adaptation circuit.

The three adapted signals are then received by the eQEP peripheral module in the DSP, which is in charge of calculating the rotational speed. This module offers two different first-order

http://dx.doi.org/10.5772/60450

approximations for speed estimation, which should be implemented for either high- (1) or low-speed operation (2) [22]:

$$v(k) \cong \frac{x(k) - x(k-1)}{T} = \frac{\Delta X}{T} \tag{1}$$

$$v(k) \cong \frac{X}{t(k) - t(k-1)} = \frac{X}{\Delta T} \tag{2}$$

where

$v(k)$: Velocity at the time instant k

$x(k)$: Position at the time instant k

$x(k-1)$: Position at the time instant k-1

T: Fixed unit time or inverse of velocity calculation rate

Δx: Incremental position movement in unit time

$t(k)$: Time instant k

$t(k-1)$: Time instant k-1

X: Fixed unit position

ΔT: Incremental time elapsed for unit position movement

(1) is the conventional approach for speed estimation in electrical drives working at high speed. This method is based on counting the number of encoder pulses $[x(k) - x(k-1)]$ in an established period of time (*T*) which can be defined as the inverse of the DSP calculation rate (150 MHz in our case) [22]. The encoder count (position) is read at the beginning of each speed control loop. Then, the speed estimate is computed by multiplying the number of pulses by the known constant $1/T$ (Fig. 8).

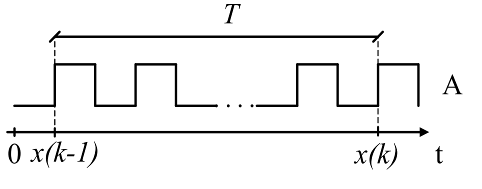


Figure 8. Electrical drive speed estimation under high-speed operation.

However, the speed estimation method based on (1) has an inherent accuracy limit directly related to the resolution of the electrical drive encoder and the sampling period *T*, making difficult to obtain an accurate speed estimation at low-speed working conditions [22]. Consequently, at low-speed operation the calculation method is changed to (2), in order to obtain a more precise measurement. In this case, the motor speed is calculated by measuring the elapsed time between successive pulses (Fig. 9). The width of each pulse is defined by the motor speed for a given encoder resolution (2500 pulses per revolution in our case).

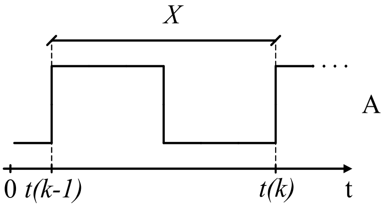


Figure 9. Electrical drive speed estimation under low-speed operation.

The DSP eQEP peripheral module is mainly formed by the units shown in the right side of Fig. 10 [23] and its functioning can be summarized as follows [22]. The encoder adapted signals A, B, I in Fig. 6, now labeled as EQEPA, EQEPB and EQEPI, are received by the Quadrature Decoder Unit (QDU), which calculates the speed direction and generates the clock for the pulses counter. The working characteristics of this pulses counter are set in the Position Counter and Control Unit (PCCU) [22]. The eQEP peripheral includes an integrated Quadrature Capture Unit (QCAP) to measure the elapsed time between consecutive pulses as shown in Fig. 9. This feature is typically implemented for low-speed measurement using equation (2). In addition, the eQEP peripheral contains a 16-bit watchdog timer (QWDOG) used to ensure the proper module operation. Finally, a 32-bit Unit Time Base (UTIME) is included to generate periodic interrupts, based on the internal clock (SYSCLKOUT), for the encoder signals measurement and speed/direction estimation.

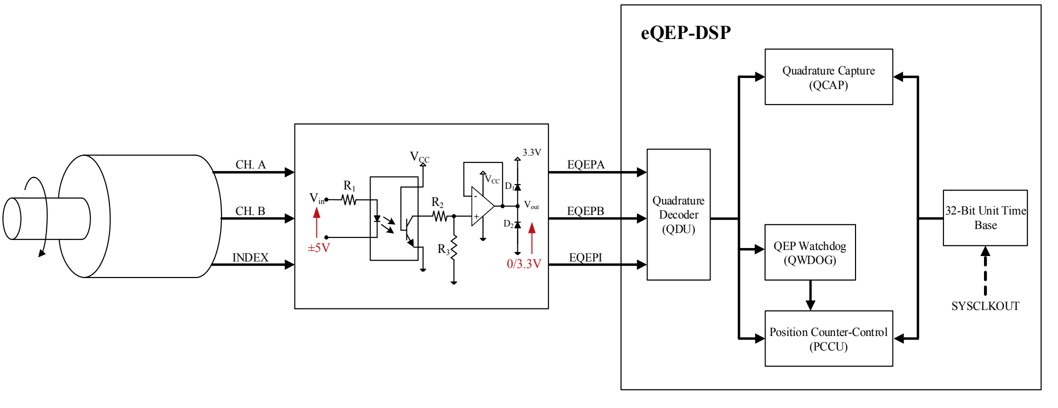


Figure 10. Electrical drive speed estimation process.

c. Power converter signal adaptation stage

In the power converter it is necessary to carry out the signal adaptation of both the semiconductor's switching and driver error signals.

http://dx.doi.org/10.5772/60450

The semiconductors drivers operate with a voltage threshold between 0/15 V, switching ON when triggered with 15 V and OFF with 0 V. The semiconductors' switching signals come from the DSP of the ECU. Then their voltage level is given between 0/3.3 V and must be adapted to meet the 0/15 V driver threshold. Each phase of the power converter consists of two semiconductors, which, for safety reasons, cannot be active at the same time to avoid short circuits. This implies that the switching signal of one semiconductor will be the complementary of the other one. The switching signal adaptation circuit (Fig. 11) contains the following stages:

- Optocoupler circuit that allows the electrical isolation between the power converter and the adaptation circuit.
- Inverter circuit to obtain the complementary switching signal, setting in this way the signals of the two semiconductors of each converter leg.
- Darlington transistor in order to raise the voltage from 0/3.3 V (DSP) to 0/15 V (semiconductor driver).

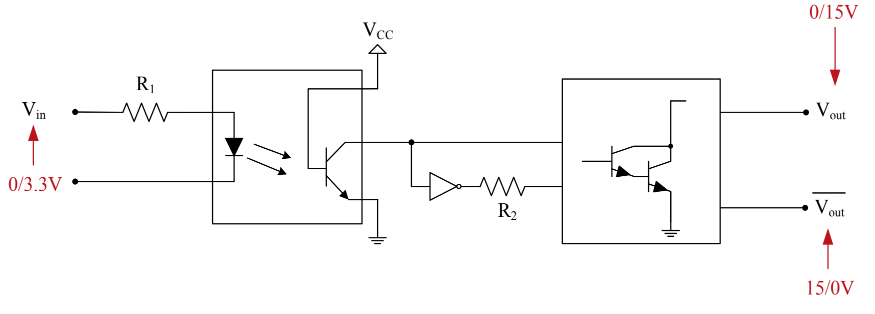


Figure 11. Schematic of the semiconductor switching signal adaptation circuit.

Correspondingly, the semiconductors driver error signals are given in a voltage level of 0/15 V. As in previous cases, such voltage level exceeds the DSP's allowed voltage threshold (0/3.3 V), and consequently must be adapted. The adaptation circuit (Fig. 12) consists of an optocoupler, which provides an electrical isolation between the power converter and the electronic control unit, and the required voltage level adaptation. Moreover, a LED is placed in series with the optocoupler resistance in order to generate a light signal in case of an error.

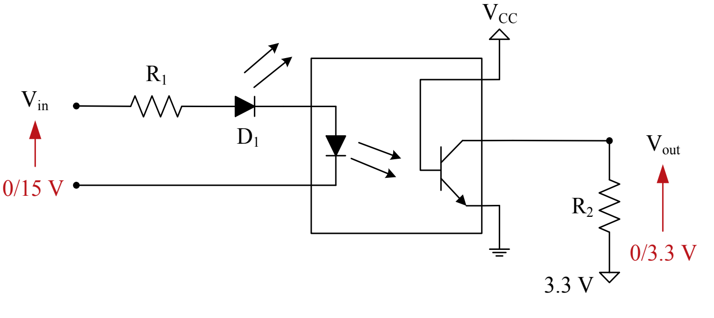


Figure 12. Error signal adaptation circuit schematic.

The aforementioned designed circuits along with the selected DSP are finally mounted on a set of Printed Circuit Boards (PCB), constituting the Modular Electronic Control Unit (MECU, patent pending) shown in Fig. 13 and placed within the power converter module, allowing to measure four phase currents, the electrical machine speed, handle the power converter error signals and control the semiconductors' switching sequence.

Figure 13. Modular electronic control unit printed circuit boards.

4. Software design: Data acquisition, signal processing and control algorithm implementation

In this section the control algorithm is detailed. Before describing the algorithm, it is necessary to understand how the hardware and the software meet. The electrical drive power converters include some current sensors to measure the stator currents in the multiphase machine. Voltage levels are obtained, corresponding to the measured stator currents, and these values are adapted to the available DSP's voltage levels and then converted into digital values using the analog-to-digital converter module of the DSP. The implemented algorithm uses these digital values to control the multiphase machine. The complete system operation flowchart is presented in Fig. 14.

http://dx.doi.org/10.5772/60450

When the operation of the system starts, the control algorithm guarantees that error signals are not activated before switching the power converter. Notice that if an error in a driver is detected while the system is on, the DSP stops the switching process, shutting down the system to avoid any damage. In any case, the control algorithm includes a stop signal to allow the test-rig operator to halt the entire system anytime. The control algorithm operates during normal operation of the multiphase drive, reading stator currents and forcing these values to follow the reference ones.

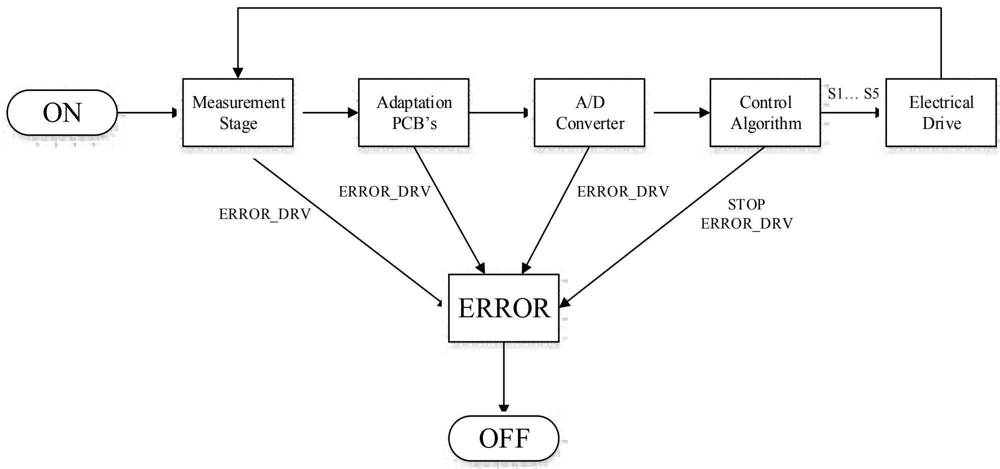


Figure 14. System operation flowchart.

The most common control strategy in the multiphase drive case is the well-known Rotor-Flux Oriented Control (RFOC) method, based on multiple inner current control loops (linear stator current PI controllers and PWM modulators) commanded by outer controllers (speed/torque and flux PI controllers). A modified RFOC controller is implemented in our case. While the external control loops of a conventional high-performance RFOC drives is maintained, an alternative to the inner current controller based on the model predictive control (MPC) is applied. The inner control method is also known as finite-control set model-based predictive control method (FCS-MPC) because the number of switching states in a power converter is finite and only one switching state is applied during the whole sampling period. Notice that the FCS-MPC method is used as a case example in this work due to its recent interest in the development of high-performance three-phase and multiphase drives [15, 16, 24, 25] and power converters [26, 27].

The basic scheme of FCS-MPC is presented in Fig. 15, and runs in the processor unit as follows. The control action is obtained solving an optimization problem at each sampling period, where a model of the real system is used to predict its output. This prediction is carried out for each possible input, or switching vector, of the power converter to determine which one minimizes a defined cost function (the basis of the optimization problem). Notice that the model of the real system, also called predictive model, must be used considering all possible voltage vectors of the power converter, which suggests a high-computational cost for the digital signal processing technique.

Different cost functions can be used to express different control criteria. In our case, an absolute current error is used to define the cost function. The reference stator currents, $\underline{i}_s^*(k)$, and the stator phase currents of the machine, $\underline{i}_s(k)$, are measured and processed by the DSP every sampling period. The machine state-space model is then used to predict the current evolution, $\hat{\underline{i}}_s(k+1)$, depending on the different possible switching states $S_i^j(k+1)$ and considering the VSI DC-Link voltage and measured phase currents $\underline{i}_s(k)$. Subsequently, the cost function J is evaluated considering predicted and reference stator phase currents, and the switching vector that provides the lowest value of the cost function, $S_i^{optimum}(k+1)$, is applied to the power converter during the next sampling period. This process is depicted in the flow diagram shown in Fig. 16.

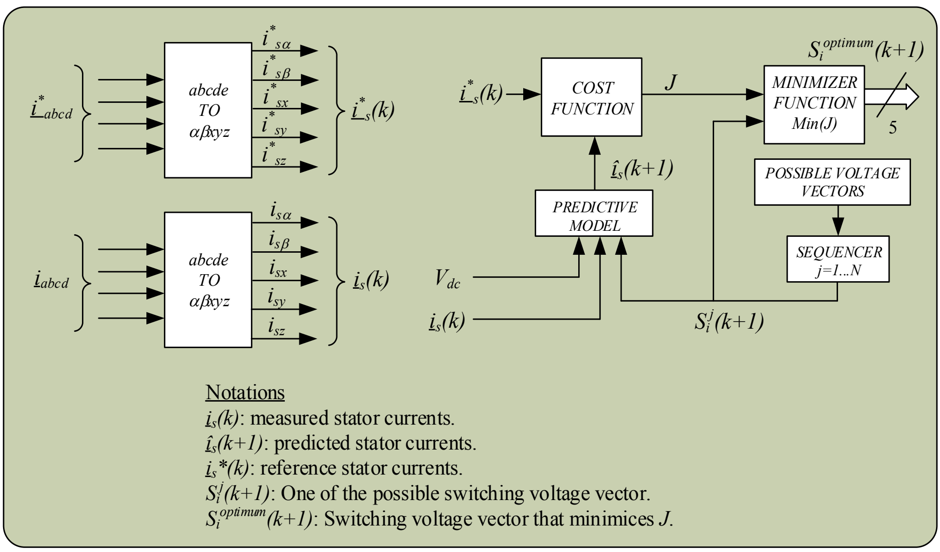


Figure 15. Finite-Control Set Model-Based Predictive Control scheme.

It is important to note that different cost functions (3) can be defined depending on the specific application in order to include different control constraints (C_i). In our case, the cost function only depends on the reference and predicted currents in the stationary reference frame. However, different control criteria aimed to optimize the multiphase drive performance such as DC-Link voltage balancing, switching stress minimization, common-mode voltage reduction or stator current harmonic minimization can be included, increasing also the complexity of the control strategy and the digital signal processing capacity of the entire system. Every considered constraint can have a different degree of importance in the cost function using weighting factors (W_i). This is one of the advantages of the FCS-MPC technique, a flexible control method where different constraints can be easily introduced without increasing the complexity of the algorithm.

$$J = W_1C_1 + W_2C_2 + \ldots + W_iC_i \tag{3}$$

http://dx.doi.org/10.5772/60450

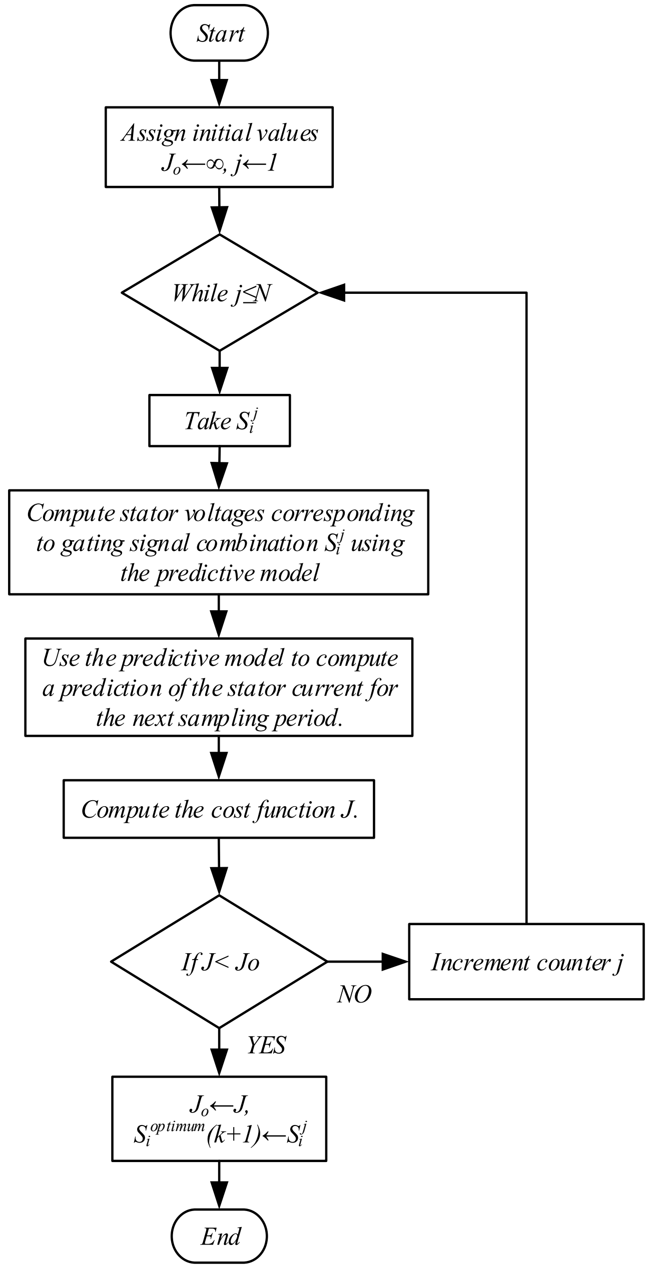


Figure 16. FCS-MPC flow diagram.

FCS-MPC can be classified in two categories depending on the prediction horizon N_p they implement. Prediction horizon can be defined as the number of future states in time that the controller predicts in order to select the most suitable control action [25]. The shortest prediction horizon can be defined as N_p=1, where measured variables are determined in the instant

k, then the optimum switching state is calculated for $k+1$ and applied at $k+1$. Larger prediction horizons ($N_p \geq 2$) predict the behavior of the electrical drive for future instants $k+2$, $k+3$, ... and select the optimum VSI switching state to be applied on $k+1$ (Fig. 17). It has been demonstrated that larger prediction horizon results in better performance [27]. However, the increase in the prediction horizon results in higher computational cost and is not suitable for real-time implementation in low and medium power drive applications [28]. One disadvantage of the FCS-MPC technique is that, unlike other aforementioned techniques (such as PI-PWM methods), it provides a variable switching frequency in the power converter.

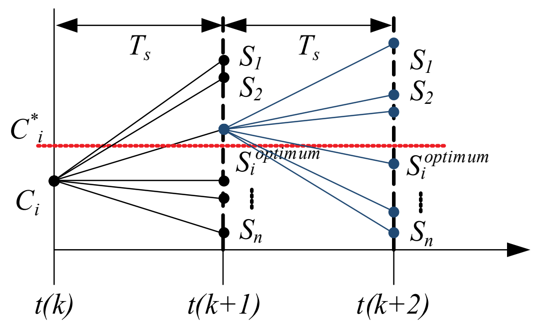


Figure 17. FCS-MPC prediction horizon principle.

In order to apply the FCS-MPC to a real high-performance multiphase drive, a Digital Signal Processor (DSP) is required due to the heavy computational cost involved. All on-line digital signal processing algorithms are implemented in the DSP. This is the case of the predictive model based on the machine modeling equations, which must be discretized to be programmed in the DSP [29, 30]. In our case, the stator phase currents and rotor flux in the stationary reference frame are assumed as state variables [29], and the obtained equations can be summarized as follows:

$$\frac{d}{dt}\overrightarrow{i_{s\alpha\beta}} = -\left(\frac{1}{\sigma\tau_s} - \frac{1-\sigma}{\sigma\tau_r}\right)\overrightarrow{i_{s\alpha\beta}} + \frac{1-\sigma}{\sigma\tau_r L_m}\overrightarrow{\lambda_{r\alpha\beta}} - j\frac{1-\sigma}{\sigma L_m}\omega_r\overrightarrow{\lambda_{r\alpha\beta}} + \frac{1}{\sigma L_s}\overrightarrow{v_{s\alpha\beta}} \tag{4}$$

$$\frac{d}{dt}\overrightarrow{i_{sxy}} = -\frac{1}{\tau_{ls}}\overrightarrow{i_{sxy}} + \frac{1}{L_{ls}}\overrightarrow{v_{sxy}} \tag{5}$$

$$\frac{d}{dt}\overrightarrow{i_{sz}} = -\frac{1}{\tau_{ls}}\overrightarrow{i_{sz}} + \frac{1}{L_{ls}}\overrightarrow{v_{sz}} \tag{6}$$

$$\frac{d}{dt}\overrightarrow{\lambda_{r\alpha\beta}} = \frac{L_m}{\tau_r}\overrightarrow{i_{s\alpha\beta}} - \frac{1}{\tau_r}\overrightarrow{\lambda_{r\alpha\beta}} - j\omega_r\overrightarrow{\lambda_{r\alpha\beta}} \tag{7}$$

http://dx.doi.org/10.5772/60450

$$\overrightarrow{\lambda_{s\alpha\beta}} = \sigma L_s \overrightarrow{i_{s\alpha\beta}} + k_r \overrightarrow{\lambda_{r\alpha\beta}} \tag{8}$$

where k_r, τ_r, σ, τ_s and τ_{ls} are obtained from the electrical parameters (stator, rotor and mutual inductances, L_s, L_r and L_m, and stator and rotor resistances, R_s and R_r) of the machine, as it is defined in (9), (10), (11), (12) and (13), respectively.

$$k_r = \frac{L_m}{L_r} \tag{9}$$

$$\tau_r = \frac{L_r}{R_r} \tag{10}$$

$$\sigma = 1 - \frac{L_m^2}{L_s L_r} \tag{11}$$

$$\tau_s = \frac{L_s}{R_s} \tag{12}$$

$$\tau_{ls} = \frac{L_{ls}}{R_s} \tag{13}$$

Rewriting equations (4)-(8) in terms of their real and imaginary components in matrix form:

$$\frac{d}{dt}\left[x\right] = \left[A\right]\left[x\right] + \left[B\right]\left[u\right] \tag{14}$$

$$\left[y\right] = \left[C\right]\left[x\right] \tag{15}$$

where

$$\left[x\right] = \left[i_{s\alpha}\ i_{s\beta}\ i_{sx}\ i_{sy}\ i_{sz}\ \lambda_{r\alpha}\ \lambda_{r\beta}\right]^T \tag{16}$$

$$\left[u\right] = \left[v_{s\alpha}\ v_{s\beta}\ v_{sx}\ v_{sy}\ v_{sz}\right]^T \tag{17}$$

$$\left[y\right]=\left[\lambda_{s\alpha}\ \lambda_{s\beta}\right]^{T} \tag{18}$$

$$\left[A\right]=\begin{bmatrix} -\left(\frac{1}{\sigma\tau_s}-\frac{1-\sigma}{\sigma\tau_r}\right) & 0 & 0 & 0 & 0 & \frac{1-\sigma}{\sigma\tau_r L_m} & \frac{1-\sigma}{\sigma L_m}\omega_r \\ 0 & -\left(\frac{1}{\sigma\tau_s}-\frac{1-\sigma}{\sigma\tau_r}\right) & 0 & 0 & 0 & -\frac{1-\sigma}{\sigma L_m}\omega_r & \frac{1-\sigma}{\sigma\tau_r L_m} \\ 0 & 0 & -\frac{1}{\tau_{ls}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{\tau_{ls}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{\tau_{ls}} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{\tau_r} & -\omega_r \\ 0 & 0 & 0 & 0 & 0 & \omega_r & -\frac{1}{\tau_r} \end{bmatrix} \tag{19}$$

$$\left[B\right]=\begin{bmatrix} \frac{1}{\sigma L_s} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sigma L_s} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{L_{ls}} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{L_{ls}} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{L_{ls}} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{20}$$

$$\left[C\right]=\begin{bmatrix} \sigma L_s & 0 & 0 & 0 & 0 & k_r & 0 \\ 0 & \sigma L_s & 0 & 0 & 0 & 0 & k_r \end{bmatrix} \tag{21}$$

Next, the machine state-space equations (14)-(15) are discretized using a sample period T_s and assuming constant inputs and constant electrical parameters during the whole sampling period [30]. Notice that the matrix $[A]$ includes constant and variable components that depend on the instantaneous value of the electrical speed (ω_r). The matrix $[A]$ has been divided into a constant matrix $[A_c]$ and a speed-dependent one $[A_\omega]$, to simplify the discretization process as follows:

$$\left[A\right]=\left[A_c\right]+\left[A_\omega\right] \tag{22}$$

$$x\left[k+1\right]=\left[\Phi\right]x\left[k\right]+\left[\Gamma\right]u\left[k\right] \tag{23}$$

$$y\left[k+1\right]=\left[C\right]x\left[k+1\right] \tag{24}$$

$$\left[\Phi\right]=e^{\left[A\right]T_s}=e^{\left(\left[A_c\right]+\left[A_\omega\right]\right)T_s}=e^{\left[A_c\right]T_s}\cdot e^{\left[A_\omega\right]T_s} \tag{25}$$

$$\left[\Gamma\right]=\int_0^{T_s}e^{\left[A\right]\tau}\left[B\right]d\tau=e^{\left[A_c\right]T_s}\left[B\right]T_s \tag{26}$$

The obtained [Φ] matrix is based on a constant term $e^{\left[A_c\right]T_s}$, which is calculated off-line, and a time-varying term $e^{\left[A_\omega\right]T_s}$, which can be defined as stated in [30] applying the Cayley-Hamilton theorem:

$$e^{\left[A_\omega\right]T_s}=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & L_m\dfrac{1-\cos\left(\omega_r T_s\right)}{\sigma L_s L_r} & L_m\dfrac{\sin\left(\omega_r T_s\right)}{\sigma L_s L_r} \\ 0 & 1 & 0 & 0 & 0 & -L_m\dfrac{\sin\left(\omega_r T_s\right)}{\sigma L_s L_r} & L_m\dfrac{1-\cos\left(\omega_r T_s\right)}{\sigma L_s L_r} \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cos\left(\omega_r T_s\right) & -\sin\left(\omega_r T_s\right) \\ 0 & 0 & 0 & 0 & 0 & \sin\left(\omega_r T_s\right) & \cos\left(\omega_r T_s\right) \end{bmatrix} \tag{27}$$

To summarize, the implemented predictive current control (PCC) technique is based on a fast inner current FCS-MPC controller with an outer speed PI RFOC-based regulator, offering better system performance (higher control bandwidth) than using conventional RFOC methods with cascaded PI controllers [25]. The inner PCC controller has been described before, where the predicted stator currents in the stationary reference frame are used in order to control the multiphase drive. The outer speed controller generates current references in a rotating reference frame from an outer PI-based speed control loop. Constant flux operation (reference speed under its nominal value) is supposed, and a constant d -current reference is imposed. These stator current references are then mapped in the α - β - x - y stationary reference frame in order to be used in the cost function (28) as it is shown in Fig. 18. The same methodology has also been applied in recent scientific studies [16, 20, 24].

The reference stator current vector in the x - y plane can be null or non-null depending on the type of multiphase machine. The α - β stator current components contribute to torque production, while x - y stator current components do not in distributed windings multiphase machines. This is our case with the five-phase induction machine with distributed windings used for experimentation. Then, zero reference is set in our controller for the x - y current components to avoid harmonic generation and reduce the copper losses in the drive. The weighting factors in the cost function (28) are adjusted in order to favor those switching states (32 for the five-phase two-level VSI used), which maximize α - β currents and at the same time provide minimum x - y currents. The overall control aim is to generate the desired electric torque to the drive, which implies the generation of sinusoidal stator current references in a - b - c - d - e phase coordinates in steady state. In the stationary α - β - x - y reference frame and steady state, the control aim is equivalent to generate a reference stator current vector in the α - β plane, which is constant in magnitude but changing in its electrical angle following a circular trajectory.

$$J = A\left|\overline{\overline{i_{s\alpha}}}\right| + B\left|\overline{\overline{i_{s\beta}}}\right| + C\left|\overline{\overline{i_{sx}}}\right| + D\left|\overline{\overline{i_{sy}}}\right| \tag{28}$$

Each α - β - x - y current term in the cost function is defined as:

$$\overline{\overline{i_{s\alpha}}} = i^*_{s\alpha}\left(k+1\right) - \hat{i}_{s\alpha}\left(k+1\right), \overline{\overline{i_{s\beta}}} = i^*_{s\beta}\left(k+1\right) - \hat{i}_{s\beta}\left(k+1\right) \tag{29}$$

$$\overline{\overline{i_{sx}}} = i^*_{sx}\left(k+1\right) - \hat{i}_{sx}\left(k+1\right), \overline{\overline{i_{sy}}} = i^*_{sy}\left(k+1\right) - \hat{i}_{sy}\left(k+1\right) \tag{30}$$

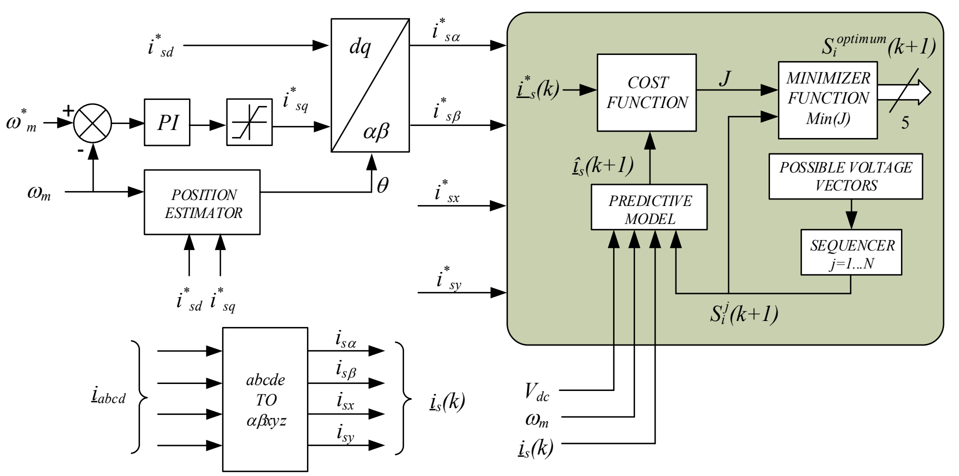


Figure 18. General scheme of the implemented PCC controller.

5. Application results

In this section, experimental results are provided to validate the interest in the design of high-performance and complex power drives of real-time digital signal processing implemented using DSPs.

A DC-Link voltage of 300 V is used and the switching frequency is fixed to 10 kHz. A constant *d*-current component of 0.52 A is also imposed, and a maximum phase current of 2.1 A is selected for safety reasons. The five-phase induction machine specifications are detailed in Table 3, while the estimated electrical parameters of the machine, used in the discretized predictive model implemented in the DSP, are summarized in Table 4.

	Parameter	**Value**
Original Machine	Rated current	7.13 A
	Rated power	4 kW
	Rated speed	2880 rpm
Five-Phase Machine	Conductor	Copper
	Diameter	0.7 mm
	Number of pole pairs	3
	Number of slots	30
	Number of turns	165
	Slots per-phase per-pole	1
	Type of winding	Single layer
	Winding pitch	5
	Rated power	1 kW
	Rated current	2.5 A
	Rated voltage (peak value)	127 V

Table 3. Five-phase induction machine specifications.

Parameter	**Value**	**Parameter**	**Value**
R_s (Ω)	12.85	L_r (mH)	768.80
σL_s (mH)	151.65	M (mH)	688.92
L_s (mH)	768.80	τ_r (ms)	179.49

Table 4. Electrical parameters of the five-phase induction machine.

The developed PCC is tested under various working conditions, including steady state operation and different torque and speed references to show that PCC provides fast reference tracking and accurate steady-state behavior at different operating points. A speed and torque response test is conducted first. Obtained results are shown in Fig. 19, where the speed response is depicted in the upper plot and the q-current in the lower plot. Under no load condition (T_L =0), a reference speed step change from 0 to 500 rpm (Fig. 19, upper plot, red line) is demanded on (t=0.1 s). Steady state operation and correct speed reference tracking is obtained at approximately (t=1.5 s). Notice that a small q-current component is observed at steady state with no load condition, due to the fact that the IM is mechanically coupled to the DC-Machine in the experimental test-rig (Fig. 3), requiring a small current in order to maintain the reference speed (Fig. 19, lower plot). Next, a load torque of approximately 28% of the nominal torque (T_n) is demanded at (t=1.8 s). Due to the load torque change, the speed reference tracking is instantaneously affected, being necessary for the system to provide an extra q-current until the reference speed is achieved. As observed in the q-current component plot, the implemented predictive current controller provides proper current tracking under the different working conditions, providing almost instantaneous response to reference current changes.

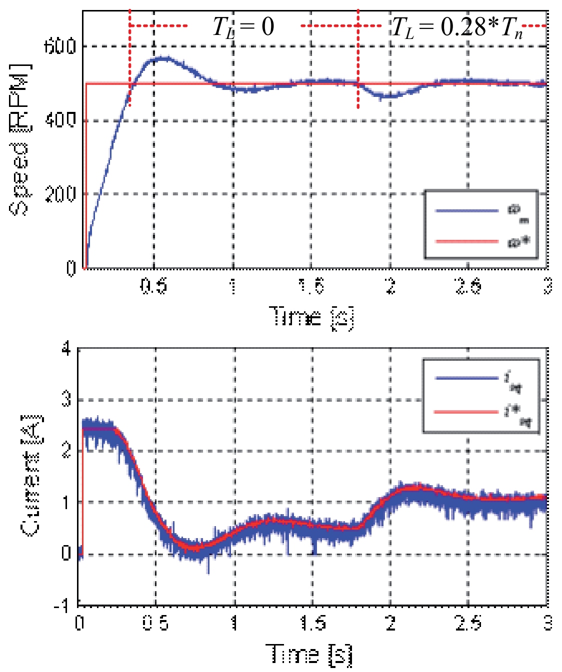


Figure 19. Speed and torque response test. The speed response is shown in the upper plot while the q-current under the different working conditions is shown in the lower plot. First, the electrical drive speed performance is evaluated with a reference speed step change from 0 to 500 rpm, under no load condition, at t =0.1 s. Then, at t =1.8 s a load torque of 28% of the nominal torque (T_L) is demanded.

http://dx.doi.org/10.5772/60450

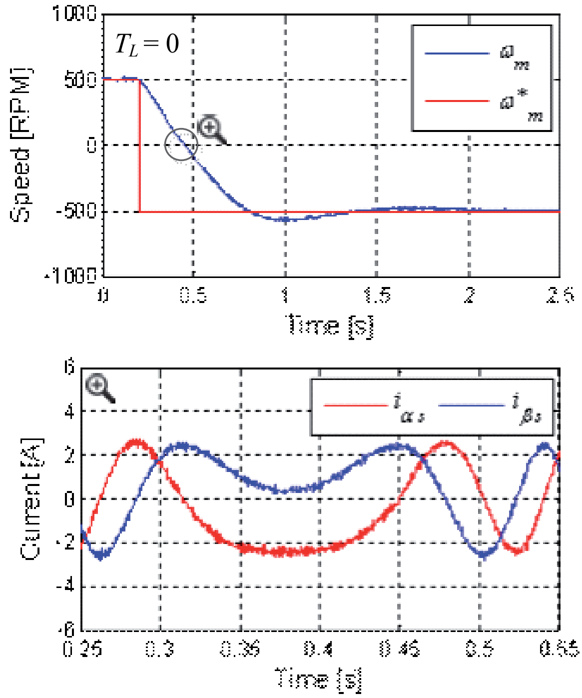


Figure 20. Reversal test. The speed reference is changed from 500 rpm to -500 rpm, under no load condition. The speed response is shown in the upper figure while a zoom-in of the obtained current waveforms in the α - β plane during the zero-crossing speed operating point are shown in the lower plot.

A speed reversal test changing from 500 rpm to -500 rpm is also applied under no load condition (although the IM and DC machines are mechanically coupled, no extra load torque is demanded). The speed is depicted in the upper plot of Fig. 20, while a zoom-in of the obtained α - β currents during the zero-speed crossing point are shown in the lower plot. Notice that once again the implemented controller provides an adequate performance, effectively achieving the speed reference. Finally, the machine currents under steady state operation are shown in Fig. 21. Phase currents are shown in the upper plot, while the α - β and x -y plots are shown in the lower plot. Notice that phase currents are symmetrical, equally displaced and do not exceed the maximum current limit of 2.1 A.

One of the main concerns in applications where digital signal processing is involved is the computational cost of the implemented algorithm. This can even be considered as the most critical issue in real-time applications, based on DSPs or microprocessors. Then, determining if the control unit is able to implement demanded operations in certain programmed sampling periods is fundamental. This computational burden is associated with the required mathematical task and DSP's modules (mainly eQEP and A/D for the data acquisition, SCI for external communication operations, and ePWM for the power converter control actions). The real-time implemented algorithm is analyzed in detail in Fig. 22. This figure shows the relative time-consuming load of every implemented task in a sampling period. It can be appreciated that the most time-demanding task corresponds to the PCC control, followed by the analog-to-digital conversion process, while other implemented processes (such as data logging and PC-DSP communication) and the speed PI control loop are not heavy from a computational load perspective.

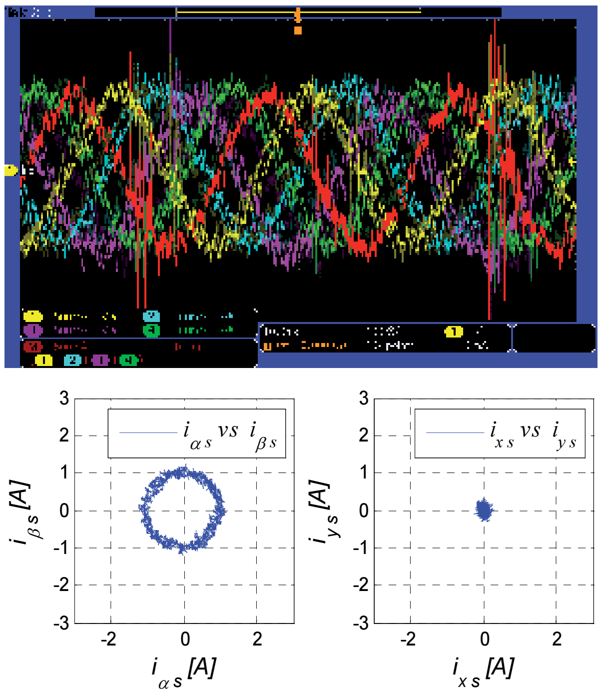


Figure 21. Steady state phase currents. Phase currents are shown in the upper plot, while currents in the α - β and x - y plane are shown in the lower plot.

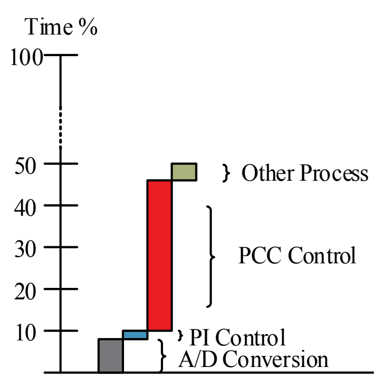


Figure 22. Algorithm real-time implementation details and task distribution in a sampling period.

http://dx.doi.org/10.5772/60450

6. Conclusion

Digital signal processing constitutes the cornerstone of numerous applications. Power electronics and the development of high-performance drives is one of them. In general, the development of a high-performance drive is a complex field where data acquisition, data adaptation and discretization, data processing, control systems and finally physical interaction using actuators with the real system are required. This complexity exponentially increases if multiphase drives are involved, as it is the case analyzed in this chapter, where a high-performance multiphase drive based on a symmetrical five-phase induction machine with sinusoidally distributed windings is analyzed from the perspective of the digital signal processing field. The specific requirements for data acquisition, processing and control in the studied electrical drive are detailed, starting with the required signals for its safety operation. Insights of the electrical adaptation stages, needed to process the data in the DSP, are then provided. The electronic adaptation circuits are also shown like a fundamental part in the processing task. Finally, the control system is summarized (the PCC method is used as a case example due to its recent interest in the field) and experimental results are provided to show the system operation.

Author details

Hugo Guzman[1*], Mario Bermúdez[2], Cristina Martín[2], Federico Barrero[2] and Mario Durán[1]

*Address all correspondence to: hugguzjim@uma.es

1 Universidad de Málaga, España

2 Universidad de Sevilla, España

References

[1] C.C. Chan, "The State of the Art of Electric, Hybrid and Fuel Cell Vehicles," *Proc IEEE*, vol. 95, no. 4, pp. 704–718, 2007.

[2] Global EV Outlook: Understanding the Electric Vehicle Landscape to 2020, Apr. 2013. [Online]. Available: http://www.iea.org/publications/globalevoutlook_2013.pdf

[3] C.C. Chan, A. Bouscayrol and K. Chen, "Electric, Hybrid, and Fuel-Cell Vehicles: Architectures and Modeling," *IEEE Transact Vehicular Technol*, vol. 59, no. 2, pp. 589–598, 2010.

[4] S.G. Wirasingha and A. Emadi, "Classification and Review of Control Strategies for Plug-In Hybrid Electric Vehicles," *IEEE Transact Vehicular Technol*, vol. 60, no. 1, pp. 111–122, 2011.

[5] M.E.H. Benbouzid, D. Diallo and M. Zeraoulia, "Advanced Fault-Tolerant Control of Induction-Motor Drives for EV/HEV Traction Applications: From Conventional to Modern and Intelligent Control Techniques," *IEEE Transact Vehicular Technol*, vol. 56, no. 2, pp. 519–528, 2007.

[6] K.T. Chau, C.C. Chan and C. Liu, "Overview of Permanent-Magnet Brushless Drives for Electric and Hybrid Electric Vehicles," *IEEE Transact Indust Electron*, vol. 55, no. 6, pp. 2246–2257, 2008.

[7] S. Williamson and S. Smith, "Pulsating Torque and Losses in Multiphase Induction Machines," *IEEE Transact Indust Applic*, vol. 39, no. 4, pp. 986–993, 2003.

[8] S. Dwari and L. Parsa, "Fault–Tolerant Control of Five–Phase Permanent–Magnet Motors With Trapezoidal Back EMF," *IEEE Transact Indust Electron*, vol. 58, no. 2, pp. 476–485, 2011.

[9] E. Levi, R. Bojoi, F. Profumo, H.A. Toliyat and S. Williamson, "Multiphase Induction Motor Drives – A Technology Status Review," *IET Electric Power Applic*, vol. 1, no. 4, pp. 489–516, 2007.

[10] J.R. Fu and T.A. Lipo, "Disturbance-Free Operation of a Multiphase Current-Regulated Motor Drive with an Opened Phase," *IEEE Transact Indust Applic*, vol. 30, no. 5, pp. 1267–1274, 1994.

[11] T.F. Podlesak, D.C. Katsis, P.W. Wheeler, J.C. Clare, L. Empringham and M. Bland, "A 150-kVA Vector-Controlled Matrix Converter Induction Motor Drive," *IEEE Transact Indust Applic*, vol. 41, no. 3, pp. 841–847, 2005.

[12] F. Khoucha, S.M. Lagoun, K. Marouani, A. Kheloui and M.E.H. Benbouzid, "Hybrid Cascaded H-Bridge Multilevel-Inverter Induction-Motor-Drive Direct Torque Control for Automotive Applications," *IEEE Transact Indust Electron*, vol. 57, no. 3, pp. 892–899, 2010.

[13] Y. Wang, X. Zhang, X. Yuan and G. Liu, "Position-Sensorless Hybrid Sliding-Mode Control of Electric Vehicles With Brushless DC Motor," *IEEE Transact Vehicular Technol*, vol. 60, no. 2, pp. 421–432, 2011.

[14] S. Kouro, P. Cortés, R. Vargas, U. Ammann and J. Rodríguez, "Model Predictive Control—A Simple and Powerful Method to Control Power Converters," *IEEE Transact Indust Electron*, vol. 56, no. 6, pp. 1826–1838, 2009.

[15] M.R. Arahal, F. Barrero, S. Toral, M.J. Durán, R. Gregor, "Multi–phase current control using finite–state model–predictive control," *Control Engin Pract*, vol. 17, no. 5, pp. 579–587, 2009.

[16] F. Barrero, M.R. Arahal, R. Gregor, S. Toral and M.J. Durán, "A Proof of Concept Study of Predictive Current Control for VSI-Driven Asymmetrical Dual Three-Phase AC Machines," *IEEE Transact Indust Electron*, vol. 56, no. 6, pp. 1937–1954, 2009.

Edited by Sudhakar Radhakrishnan

This book is recommended to readers who can ponder on the collection of chapters authored/co-authored by various researchers as well as to researchers around the world covering the field of digital signal processing. This book highlights current research in the digital signal processing area such as communication engineering, image processing and power conversion system. The entire work available in the book mainly focusses on researchers who can do quality research in the area of digital signal processing and related fields. Each chapter is an independent research, which will definitely motivate young researchers to further study the subject. These six chapters divided into three sections will be an eye-opener for all those engaged in systematic research in these fields.

IntechOpen

ISBN 978-953-51-5764-9

9 789535 157649