A nearly loss-less compression technology based on CTP and partial sample points calibration for OFDM signal

Yan Li:School of Information Science and Technology,
Sun Yat-sen University,
Guangzhou, 510006, China;
&SYSU.CMU Shunde International Joint Research Institute,
Shunde, 528000, China ;
838627645@qq.com .
Baoxian Yu: School of Information Science and Technology,
Sun Yat-sen University,
Guangzhou, 510006, China;
1531061563@qq.com .
Dongri Su: School of Information Science and Technology,
Sun Yat-sen University,
Guangzhou, 510006, China;
sudongri@gmail.com
Xianhua Dai:School of Information Science and Technology,
Sun Yat-sen University,
Guangzhou, 510006, China;
issdxh@mail.sysu.edu.cn

***Abstract*—Traditional compression algorithms of OFDM signals with a low compression ratio and a low SNR come along with a high error rate, which fail to meet the requirements of OFDM systems. To achieve high compression ratio and low error rate, we propose a new algorithm based on a combination of clipping with tail plug (CTP) and partial sample calibration (PSC). The new algorithm can achieve the nearly loss-less compression and the compression ratio can be as high as 1.86:1 with a error rate less than 10^(-7), while SNR corresponding to quantization error is up to 70dB in 4096-QAM OFDM signal modulation schemes, which meets well for the requirements of the FTTdp+GDSL systems.**

***Keywords*—OFDM systems; clipping with tail plug (CTP); partial sample calibration (PSC); nearly loss-less compression**

I. Introduction

In the OFDM systems[1], the modulated signals (IFFT of transmitted data vector) always exhibits a rather great peak-average ratio[2], which requires to be compressed for effectively transmitting over real channels. The data compression technology[3] officially begins in the late 1930 and early 1940 of the 20th century. In 1952,Huffman invented Huffman code[4], which is a typical data compression technology based on the statistical methods. In 1957 and In 1960, Lloyd and Max separately published the best scalar quantization algorithm in case of signal distribution, that is Lloyd-Max algorithm,which is then extended to the LBG algorithm. However, in most cases, the probability of the data distribution is unknown, in order to compress the data effectively on this condition, Jacob Ziv and Abraham Lempel, two scientists from Israel, firstly offered a data compression technology called LZ77 coding algorithm[5] based on the dictionary. A year later, they improved LZ77,which is referred as LZ78 algorithm[6]. Since then, many other scientists offered new improved data compression algorithms, such as LZW,LZMW,LZAP,LZP and so on. These traditional compression algorithms, however, come along with a high error rate for high-level modulators, e.g.2048 or 4096QAM[7], which fails to meet the OFDM systems in FTTdp+GDSL systems.

In this paper, we propose a new nearly loss-less compression algorithm based on a combination of clipping with tail plug (CTP) and partial sample calibration (PSC). By adjusting the down-sampling factor according to DSL line length, we can obtain the loss-less compression with different compression ratio.

II. Lossy Compression With CTP At The Transmitter

At the transmitter,we use two ways (the high 14 bits and the lowbits) to sample the OFDM signal. As shown in Fig.1: the upper path is composed of CTP, geometric series companding (GSC) and low precision fixed-point(quantization of bits), the down path is the direct 14-bit quantization at down-sampling M. Assume 14-bit quantization is error-free.The implementation process is as follows:

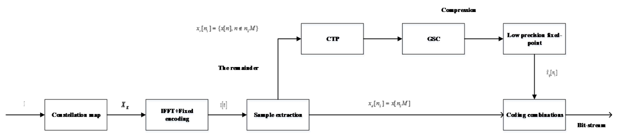

Fig. 1. schematic diagram at the transmitter

(1) IFFT modulated signal length is N .
(2) CTP: clipping with tail plug. It cuts off of the peak over threshold and identifies the cutting-off position, and then inserts sequentially the cutting-off remaining amount at the end of each symbol sequence, the receiver can finally achieve the clipping operation with no deviation.
(3) GSC: geometric series companding. It regulates the size of the first a_1 and common ratio q for the geometric sequence, thus completing switching between linear and nonlinear companding.
(4) Coding combinations: combine $x_d[n_2]$ and $\tilde{x}_q[n_1]$ orderly into a new sequence,which is then sent into the optical fiber.

From Fig.1, compression mainly comes from the process of "CTP+GSC+Low precision fixed-point (quantization of L bits)" in the upper-path. For original OFDM modulated signals with a sampling rate 14bit, the compression ratio in Figure 1 equals to 14:(14*1/M+L*(M-1)/M).

In the above process, M-times down-sampling with a high-bit (14 bits) in down-path, the output signal $x_d[n_2]$, is the same as the original samples at the sampling points, thus the output signal $x_d[n_2]$ is loss-free and approximately error-free, the quantizing errors introduced by quantization exists mainly in the signal $\tilde{x}_q[n_1]$, which is coded by the quantization of L bits.

III. The Nearly Loss-less Decompression Algorithm At The Receiver

A. *Sub-carrier correction*

At the receiver, we firstly separate out the signal $x_d[n_2]$, then decompress the remaining signal $\tilde{x}_q[n_1]$. Recombining $x_d[n_2]$ and $\tilde{x}_q[n_1]$,we'll get the signal $\tilde{x}[n]$,which has the deviation caused by the low precision fixed-point. After FFT demodulation of signal $\tilde{x}[n]$, we'll get the demodulated sequences as $Y_k = Sgn\{FFT[\tilde{x}(\mathrm{n})]\}$, where Sgn{} represents a decision of the constellation diagram according to QAM modulators. In order to correct errors, we need the time-domain signal again as $\tilde{y}[n]$= IFFT(Y_k)). Since the sampling signal $x_d[n_2]$, which is coded directly by 14bits, is approximately error-free, besides, we neglect the additive Gaussian noise effect, the deviated signals between $x_d[n_2]$ and $\tilde{y}[n_2]$ at the down-sampling points n_2 , $err[n_2] = x_d[n_2] - \tilde{y}[n_2]$,will mainly caused by the demodulated errors.

Assume the sub-carrier P has the offset or demodulated error, the size is $A_p e^{j\theta_p}$,then its frequency-domain signal is:

$$Y_p = \boldsymbol{X} + [\underbrace{0,\cdots,0}_{p}, \underbrace{A_p e^{j\theta_p}}_{子载波p}, 0,\cdots,0, \underbrace{A_p e^{-j\theta_p}}_{子载波N-p}, \underbrace{0,\cdots 0}_{p-1}] \qquad (3\text{-}1)$$

N -point IFFT transformation for the signal Y_p in time-domain:

$$\begin{aligned}\tilde{y}_p[n] &= \frac{1}{N}\sum_{k=0}^{N-1} Y_p e^{j2\pi\frac{kn}{N}} \\ &= \frac{1}{N}\sum_{k=0}^{N-1} X e^{j2\pi\frac{kn}{N}} + \frac{1}{N}(A_p e^{j\theta_p} e^{j2\pi\frac{pn}{N}} + A_p e^{-j\theta_p} e^{j2\pi\frac{(N-p)n}{N}}) \\ &= x[n] + \frac{A_p}{N}(e^{j(2\pi\frac{pn}{N}+\theta_p)} + e^{-j(2\pi\frac{pn}{N}+\theta_p)}) \\ &= x[n] + \frac{2A_p}{N}\cos[2\pi\frac{pn}{N}+\theta_p]\end{aligned} \qquad (3\text{-}2)$$

That is, a demodulated error at the sub-carrier P caused a single-frequency sine signal error. M times down-sampling for $\tilde{y}_p[n]$ and subtracting the high bit samples $x_d[n_2]$, we have the error signal in time domain:

$$\begin{aligned}err[n_2] &= x_d[n_2] - \tilde{y}_d[n_2] = x[n_2 M] - \tilde{y}_p[n_2 M] \\ &= -\frac{2A_p}{N}\cos[2\pi\frac{p}{N}n_2 M + \theta_p], (n_2 M \in [0, N])\end{aligned} \qquad (3\text{-}3)$$

Performing N-FFT operation on the signal $err[n_2]$ (fill with zero at $n \neq n_2 M$), we have

$$\begin{aligned}Err[k] &= \sum_{n_2=0}^{N-1} err[n_2] e^{-j2\pi\frac{kn_2}{N}} = -\frac{A_p}{N}\sum_{n_2=0}^{N-1}(e^{j(2\pi\frac{pn_2M}{N}+\theta_p)} + e^{-j(2\pi\frac{pn_2M}{N}+\theta_p)})e^{-j2\pi\frac{kn_2}{N}} \\ &= -\frac{A_p}{N}[e^{j\theta_p}\sum_{n_2=0}^{N-1} e^{j2\pi\frac{pn_2M}{N}} e^{-j2\pi\frac{kn_2}{N}} + e^{-j\theta_p}\sum_{n_2=0}^{N-1} e^{-j2\pi\frac{pn_2M}{N}} e^{-j2\pi\frac{kn_2}{N}}] \\ &= -\frac{2\pi A_p}{N}[e^{j\theta_p}\sum_{l=-\infty}^{+\infty}\delta(2\pi f - 2\pi\frac{p}{N} - 2\pi l) + e^{-j\theta_p}\sum_{l=-\infty}^{+\infty}\delta(2\pi f + 2\pi\frac{p}{N} - 2\pi l)] \\ &(f = \frac{k}{N}, k \in [0, N])\end{aligned} \qquad (3\text{-}4)$$

Therefor, through the time-domain signal $err[n_2]$, we'll be able to detect the constellation mapping point offset for the sub-carrier P , and thus correct the OFDM signals. Similarly, if the demodulated signal offsets occur in two or more sub-carriers, we can calibrate the errors in a similar way above as well.

B. *The nearly loss-less decompression*

Based on the above mentioned in A,loss-less decompression can be summarized as follows:

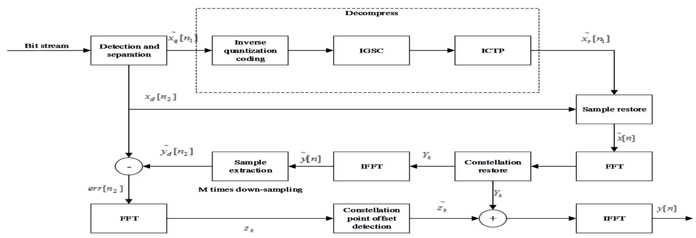

Fig. 2. the nearly loss-less data compression schematic diagram.

The main function modules：

(1)Detection and separation:divide the receiving sequence into two sub-sequences, that is the 14-bit-fixed-point sub-sequence $x_d[n_2]$ and the L -bit- fixed-point sub-sequence $\tilde{x}_q[n_1]$.

(2)Decompression:the process of "Inverse quantization+IGSC+ICTP)",the operation corresponds to the compression procedure at the sender.

(3)Sample restore:combine $x_d[n_2]$ and $\tilde{x}_q[n_1]$ to form a complete OFDM symbols in order of original sample time, the length of the sequence $\tilde{x}[n]$ is N .

(4)FFT:The discrete Fourier transform.

(5)Constellation restore: After FFT and the anti-normalization transformation, the data points on each sub-carrier are mapped to the corresponding constellation diagram so as to correct minor deviations on the constellation diagram.

(6)Constellation point offset detection: N -point FFT transformation for the signal $err[n_2]$,we'll get the frequency domain signal Z_k . Determine whether there is the constellation mapping point offset by detecting whether the real part and the imaginary part of Z_k exceeds a preset threshold value.

(7)IFFT:Inverse discrete Fourier transform.

Note that the calibration method above, the M times down-sampling frequency must meet the formula $f_s^{'} = 1/M \geq 2\,p_{\max}/N$ where $p_{\max}$ is the maximum sub-carrier number of demodulated errors occurring, for avoiding aliasing. That is, the sub-carrier number of demodulated errors occurring should meet $p_{\max} \leq N/2M$. Therefore, this new algorithm mainly applies for low frequency carrier correction, e.g. DSL errors currently focused mainly on low frequencies, so the algorithm is of great value.

IV. Modeling And Simulation

In simulations, the simulation parameters are listed in Table 1.

In order to secure greater compression ratios (greater than 1.45:1) and bit error rate is less than 10^{-7} magnitude,proper values should be given to parameters L and M . Table 2 shows some typical parameters and values.

Therefore, as long as $M > 4$,we can ensure that the bit error rate is less than 10^{-7} magnitude and the compression ratio is greater than 1.4:1 .if we want to ensure the maximum compression ratio on different length DSL,the optimal values should be given to parameters L and M ,which are shown in table 3.

TABLE I. SIMULATION PARAMETER SETTINGS

Parameters	Values
DSLline lengthLoopLen [m][a]	50,100,150,200,250,300
number of effective sub-carrier	4096
Length of FFT/IFFTI	8192
Carrier top bits	12
IFFTOutput fixed bit-width	14
Second quantization in bit width	9,8,7
Constant β	1.6σ
Fixed-length DSL standard deviation σ [b]	1292, 1355, 1400, 1100,880,730
OFDM simulation symbol number	10000
Down sampling factors M	2,4,8,16,32,64,128,256,512

Note*1:DSL line is within 300 meters, each 50 meters for the test site.
Note*2:DSL standard deviation is set to constants so as to reduce computational complexity, simulations show that the instantaneous standard deviation for each symbol is almost equal to the setting value,which respectively corresponds to the DSL line length(50~300)from left to right.

TABLE II. TYPICAL PARAMETERS L AND M

Down-sampling factor M		Quantization bits wide L		
		7	8	9
DSL length	50	-	-	4 ~ 512
	100	-	4 ~ 8	4 ~ 512
	150	-	4 ~ 16	4 ~ 512
	200	4	4 ~ 32	4 ~ 512
	250	4 ~ 16	4 ~128	4 ~ 512
	300	4 ~ 32	4 ~128	4 ~ 512

Note:"-"indicates non-existent.

TABLE III. OPTIMAL PARAMETERS L AND M

DSL length	L	M	avgRatio	avgBER
50	9	512	1.4876	7.54E-08
100	8	8	1.5451	4.55E-07
150	8	16	1.6088	3.91E-07
200	8	32	1.6416	7.64E-08
250	7	16	1.8128	7.30E-08
300	7	32	1.8632	3.60E-08

When parameters L and M reach the optimum solution,draw the input/output SNR, as shown in Fig.3.

V. Conclusion

In this paper,we have proposed a new compression algorithm for OFDM signal that offers the nearly loss-less decompression based on partial sample calibration. The compressed complexity of the algorithm is extremely low, whereas the decompressed complexity is relatively high. By adjusting down-sampling factor M for different length of DSL line, we can achieve different compression ratios of nearly loss-less compression, simulation results show that the compression ratio can be as high as 1.86:1.

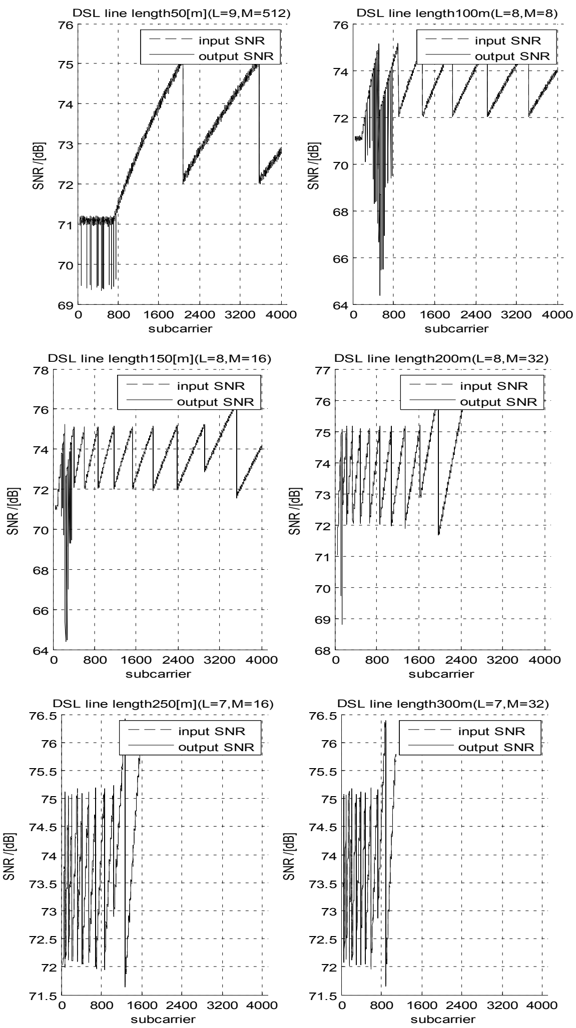

Fig. 3. the input/output SNR when parameters L and M reach the optimum solution

ACKNOWLEDGMENT

The work of this paper is supported by the Guangdong Natural Science Foundation under the grant number 408161212044, and the Scientific Research Project of Sun Yat-sen University-Carnegie Mellon University Shunde International Joint Research Institute under grant number YB2012110175.

REFERENCES

[1] Morant M, Llorente R, Herrera J, et al. Integrated FTTH and In-building Fiber-Coax OFDM Field Trial[J]. Photonics Technology Letters, IEEE, 2014, PP(99): 1.

[2] More A P, Somani S B. The reduction of PAPR in OFDM systems using clipping and SLM method[C]. Information Communication and Embedded Systems (ICICES), 2013 International Conference on. IEEE, 2013: 593-597.

[3] Sayood K. Introduction to data compression 4th[M]. Newnes, 2012: 13.

[4] Huffman D A. A method for the construction of minimum redundancy codes[J]. Proceedings of the IRE, 1952, 40(9): 1098-1101.

[5] Ziv J, Lempel A. A universal algorithm for sequential data compression[J]. IEEE Transactions on information theory, 1977, 23(3): 337-343.

[6] Lempel A. Compression of individual sequences via variable-rate coding[J]. IEEE Transactions on Information Theory, 1978, 24(5): 530-536.

[7] Deepa T, Kumar R. Performance analysis of μ -law companding & SQRT techniques for M-QAM OFDM systems[C]. Emerging Trends in Computing, Communication and Nanotechnology (ICE-CCN), 2013 International Conference on. IEEE, 2013: 303-307.