

A novel high spectral efficiency waveform coding-OVTDM

Li Daoben

School of Information and Communication, Beijing Univ. of Posts & Telecomm, Beijing, China

Email address:

Lidaoben2014@163.com

To cite this article:

Li Daoben. A Novel High Spectral Efficiency Waveform Coding-OVTDM. *International Journal of Wireless Communications and Mobile Computing*. Special Issue: 5G Wireless Communication Systems. Vol. 2, No. 4-1, 2014, pp. 11-26. doi: 10.11648/j.wcmc.s.2014020401.12

Abstract: Different from other Coding, based on the Overlapped Multiplexing Principle discovered by author, a novel OVTDM (Overlapped Time Division Multiplexing) Waveform Coding is proposed. Instead of the encoding matrix and mapped signal constellation, any engineering sense band-limited Multiplexing Waveform can be employed. By its data weighted shift overlapped versions, the coding gain and spectral efficiency are both achieved. The heavier the overlap of the data weighted Multiplexing Waveform, the higher the coding gain and spectral efficiency as well as the closer the output to the optimum complex Gaussian distribution. The encoder structures, parameters, optimum and fast decoding algorithms, pre-coding, some implementation problems as well as the bit error performance are estimated and discussed. Simulations show that OVTDM is suitable for high spectral efficiency applications and its spectral efficiency is roughly proportional to SNR.

Keywords: Overlapped Multiplexing Principle, Waveform Encoding, Spectral Efficiency, Shannon Capacity

1. Preface

Shannon Theory is well known the guidance of communications. If let a signaling symbol (pulse) of received signal to carry more information, multi-levels called level division should be employed. Along with the code constraint length increasing, the number of distinguish levels of a received “pulse” approaches to $\sqrt{1+P_s/P_N}$ carrying $\max 0.5 \log_2(1+P_s/P_N)$ bits/symbol information, where P_s/P_N is the signal noise ratio (SNR). Later by level division and assuming channel obey Nyquist criterion with bandwidth strictly limited to B. A continuous channel was easily transformed to a discrete memory-less signaling symbol sequence with rate $2B$ symbols/s, i.e. a discrete memory-less channel. The “Shannon capacity” $C = B \log_2(1+P_s/P_N)$ bps/Hz was easily obtained in this way. Obviously channel capacity may have different even better form if level division and Nyquist criterion are no longer employed.

There are $2K$ combinations of K bits with total duration KT_b , needing $2K$ one to one mapping represented symbols. If the channel is regarded as a memory-less symbol sequence, surely level division, i.e. a signal constellation of $2K$ levels is the only choice. However the received signal is continuous one, why don't employ $2K$ waveforms? It is a general knowledge that in a very noisy environment, people can still distinguish a huge number of weak voices by their waveforms rather than levels.

The fatal weakness of Nyquist Criterion is that it violates the uncertainty principle and the no ISI (Inter-symbol Interference) Nyquist Channel is physically unrealizable. In fact, in any field X (X denotes time T , frequency F , space S , code C as well as their hybrid H) system, the overlapping between adjacent data is unavoidable, the higher the data rate the heavier the ISI. Why don't to utilize ISI adroitly? The Overlapped Multiplexing Principle discovered by [12] reveals that the overlapping between adjacent and neighboring data in any system is never interference but a beneficial coding constraint relation offering benefit coding gain. The destroy fact coming outside the system is only the interference. The channel capacity will be reduced by brutally force equalizing a channel with code constraint into a Nyquist channel of thoroughly losing coding constraint relation.

Leaving from Nyquist criterion and level division, based on the Overlapped Multiplexing Principle and waveform division, a novel OVTDM (Overlapped Time Division Multiplexing) waveform coding scheme is proposed in the paper. By the shift data weighted overlapped version of an engineering sense band-limited Multiplexing Waveform, there appears a OVTDM coding with high spectral efficiency, high coding gain, no coding redundancy, relative low decoding complexity.

It is well known that the transmitted signal should be not only in complex field, but also in complex Gaussian distribution under additive complex Gaussian noise

environment. Unfortunately all nowadays coding need mapping to a signal constellation in complex field. Though most of the sequence to sequence coding is blameless, their final outputs can never be in complex Gaussian distribution, due to the mapped signal constellations are all in uniformly distribution. Even “shaping” scheme may centralize the signal constellation a little. Such modification can never solve their fatal weakness and gives at most 1.53 dB gain.

Although there are non-finite field coding, e.g. [10] and partial response coding, as well as the superimposed coding [9]. However such “superimposition” is not the “shifted overlapping with ISI” and the key point is that they never leave the uniform distributed signal constellation.

FTN (faster than-Nyquist) criticizes Nyquist signaling rate but insists on a strictly band-limited Nyquist channel (Sinc pulse is an exception?). On the other hand, so far FTN is only a little faster than Nyquist rather than much faster than Nyquist like OVTDM, and its drawback is still treats symbol overlapping as an interference rather than a beneficial coding gain.

OVTDM belongs to a novel waveform coding, it is based on waveform division rather than level division and its output automatically approaches to optimum complex Gaussian distribution. OVTDM employs an engineering sense band-limited multiplexing waveform. By its shift data weighed overlapped version, OVTDM will have least output levels and maximum Euclidean branch distance as well as an approaching to optimum complex Gaussian output.

Except Nyquist criterion another obstacle of limiting spectral efficiency η is the signal levels, i.e. the number of points in a signal constellation. People use to put $2K$ levels without coding and at least $2K+1$ levels with coding for $\eta=K\text{bits/symbol}$. Therefore the signal levels will be increased exponentially with η , even shaping may shrink the levels a little. For average power limited channel, the more the signal levels the smaller the distance between them and the lower the noise immunity.

OVTDM essentially is a convolutional waveform coding scheme. Except the near complex Gaussian distribution outputs, OVTDM also has least number of output levels. For binary $(+1,-1)$ data input, the output of a K folds OVTDM only has $K+1$ levels with spectral efficiency $\eta=K\text{bits/symbol}$ and $2K$ distinguished output sequences within the code constraint length KT_b . Since there are only $K+1$ levels for each code node, their Euclidean distance between code nodes can be increased at most. Surely relative higher noise immunity can be achieved.

The output distribution of a K folds OVTDM with binary $(+1,-1)$ input is the K th order binomial distribution approaching to optimum Gaussian distribution with K . It is well known that polynomial and binomial distribution can all approach to Gaussian distribution with K . Two stage concatenate OVTDM structure parallel putting in orthogonal I, Q channels is proposed in the paper, where the 1st stage is a $K1$ th order pure OVTDM (no relative shift) changing binary $(+1,-1)$ input into multilevel real input and the 2nd stage is a $K2$ th order shifted OVTDM making output polynomial

distribution approach to Gaussian. The total spectral efficiency η of such I, Q parallel concatenate OVTDM structure is $\eta=2K1K2\text{bits/symbol}$, and I, Q real distribution outputs together approach to complex Gaussian.

Instead of encoding matrix and mapped signal constellation, any band-limited Multiplexing Waveform can be employed in OVTDM. By its shift data weighed overlapped version, coding gain and spectral efficiency η are both achieved. Then what is the Optimum Multiplexing Waveform? What effect is the channel filtering? Such problems will be discussed in the paper. The channel capacity of OVTDM is roughly linear to SNR rather than logarithm of SNR. The reason is it utilizes physically realizable engineering sense rather than Nyquist sense band-limited waveform. No matter how fast decay with tail spectrum outside filter's bandwidth, physically realizable spectrum tail always extends to infinite. The system capacity should be linearly to SNR (see [15] and appendix C of the paper). The performance of OVTDM can go far beyond the Shannon capacity when spectral efficiency η is high enough, the reasons are as follows:

Leaving obstacle from Nyquist criterion, recovering discrete memory-less sample sequences with no code constraint to waveforms with strong code constraint relation;

Employing waveform division instead of level division;

The code outputs no longer have exponentially increased but algebraically increased levels with spectral efficiency η . The Euclidean distance between nodes of OVTDM increased at most.

No matter how fast decay with tail spectrum outside filter's bandwidth, spectrum tail always extends to infinite rather than strictly no tail.

Even only field time T coding OVTDM is discussed in the paper. It is easily expanded to other field X (OVXDM).

2. OVTDM System Model

A OVTDM Model

The complex envelop (center frequency removed) model of an OVTDM system is shown in Fig.1.

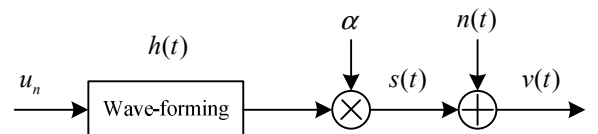


Fig 1. OVTDM model

Assuming that $U \triangleq [u_0, u_1, \dots]^T$ be the transmitted i.i.d. (identical, independent, distribute) data sequence; $h(t)$ be the so called Band-limited Multiplexing Waveform, which is an impulse response function, including all the filters in the system. e.g. wave-forming, transmitter, receiver, channel, equalizer etc. u_n carrying Q bits be the n th transmitted data with duration T ; E_0 be the transmitted data energy; α be the attenuation factor of the channel; $E = \alpha E_0$ be the received data energy, $n(t)$ be the complex envelope of additive white

Gaussian noise with power spectrum density N_0 . Then the received signal's complex envelope is

$$v(t) = \sqrt{2E} \sum_n u_n h(t - nT) + n(t) = s(t) + n(t), \quad (1)$$

Where $h(t) = 0, t \notin (0, \Delta)$ is the employed any physical realizable band-limited multiplexing waveform;

$(K-1)T < \Delta \leq KT$; $K \triangleq \lfloor \Delta/T + 1 \rfloor$ is the number of overlapped folds of $h(t)$; $\lfloor \bullet \rfloor$ is the least integer of \bullet .

When $t \in [nT, (n+1)T]$, $n = 0, 1, 2, \dots$, the received signal can also be represented as

$$v_n(t) = \sqrt{2E} \sum_{i=0}^{K-1} u_{n-i} h_i(t) + n_n(t) = s_n(t) + n_n(t), \quad (2)$$

Where:

$$\begin{cases} s_n(t) \triangleq s(t)[U(t - nT) - U(t - (n+1)T)] \\ n_n(t) \triangleq n(t)[U(t - nT) - U(t - (n+1)T)] \\ v_n(t) \triangleq v(t)[U(t - nT) - U(t - (n+1)T)] \\ h_i(t) \triangleq h(t + iT)[U(t) - U(t - T)] \end{cases} \quad (3)$$

$U(t)$ is the unite step function.

Therefore $v_n(t)$ is just the complex convolution of transmitted data sequence $U \triangleq [u_0, u_1, \dots]^T$ with Multiplexing Waveform sequence $h(t) \triangleq [h_0(t), h_1(t), \dots, h_{K-1}(t)]^T$.

Fig.2 is its model of waveform convolutional encoder with rate 1 and constraint length K. The spectral efficiency reaches $\eta = KQ$ bits/symbol. For the sake of simplicity in Fig.2 tape coefficients h_0, h_1, \dots, h_{K-1} are not number sequence but waveform sequence $h_0(t), h_1(t), \dots, h_{K-1}(t)$.

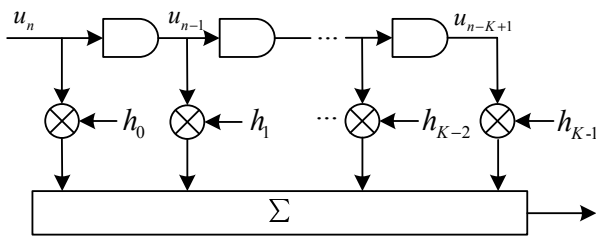


Fig 2. The complex waveform convolutional encoder model of OVTDM with shift unit T and overlapping folds K

For 2Q-nary real data, the output of a K folds OVTDM of real $h(t)$ will have $K(2^Q - 1) + 1$ levels with spectral efficiency KQ bits/symbol. The reason is that since the width of symbol (multiplexing waveform $h(t)$) expanded K times, the original signal bandwidth B should be shrunk to B/K. In order to keep the same data rate there have to be K symbols overlapped together. Let the frame length be L occupying $[K + (L-1)]T$ sec. carrying LQ bits. Then its spectral efficiency η becomes

$$\eta = \frac{LQ}{[K + (L-1)]TB / K} \xrightarrow{L \gg K} \frac{KQ}{BT} \text{ bits/s/Hz} \quad (4)$$

When $L \gg K$, the spectral efficiency η and the capacity of such K folds OVTDM system will increase K times. Its output is binomial distribution of order K, approaching Gaussian distribution when $K \gg 1$.

Similarly, when $h(t)$ is real, for quaternary (+1, -1, +j, -j) independent data steam. Any K folds OVTDM output has $(K+1)^2$ levels with spectral efficiency $\eta = 2K$ bits/symbol. Its output is orthogonal two binomial distributions of order K, approaching complex Gaussian distribution when $K \gg 1$.

OVTDM does destroy the one to one mapping relation between the input and output symbols but keeps the one to one mapping relation between the input sequence and the output sequence [2][12]. For binary data within constraint length K, there are 2K input binary data sequences, but also 2K output waveform sequences. They are absolutely in one to one mapping relation. Similarly, for quaternary data within constraint length K, there are 2K input binary data sequences and 2K output waveform sequences both in orthogonal I and Q channels. One to one mapping relation still kept. Surely MLSD (Maximum Likelihood Sequence Detection) should be employed. From the total 2K possible waveform sequences in each I and Q channel, to select the most possible coded waveform sequence that is nearest to the received signal waveform sequence [2][12].

B Power Spectrum of OVTDM

Let the Multiplexing waveform and its spectrum be $h(t)$ and $H(f)$ respectively. $h(t) \leftrightarrow H(f)$

They are a pair of Fourier transform.

The output waveform of OVTDM and the corresponding spectrum are

$$\sum_n u_n h(t - n/T) \leftrightarrow H(f) \sum_n u_n e^{j2\pi f n/T} \quad (5)$$

Then output power spectrum of OVTDM is

$$E \left\{ \left| H(f) \sum_n u_n e^{j2\pi f n/T} \right|^2 \right\} = E \left\{ \sum_n \sum_{n'} u_n u_{n'}^* e^{j2\pi f (n-n')/T} \right\} |H(f)|^2, \quad (6)$$

Since input data is i.i.d. $E(u_n u_{n'}^*) = E(|u_n|^2) \delta_{n,n'}$, formula (6) becomes

$$|H(f)|^2 \sum_n E(|u_n|^2) \quad (7)$$

For i.i.d. data, since only K adjacent data overlapped together, there are only K terms in summation. Finally the power spectrum of a K folds OVTDM is $K|H(f)|^2$. That is completely same as its band-limited multiplexing waveform $h(t)$.

C The Difference of OVTDM from others

OVTDM employ not level division but waveform division. OVTDM belong to waveform coding. Instead of the encoding

matrix and mapping constellation, any engineering sense band-limited Multiplexing Waveform $h(t)$ can be employed. By its data weighted shift overlapped versions, the coding gain and spectral efficiency η are both achieved. The system performance is only determined by $h(t)$. Then what is the optimum $h(t)$? If the restraint condition is the coding constraint length K , the answer is very simple. Due to the symmetry principle, rectangular $h(t)$ is the optimum, since its coding constraint relation is equal and maximum. However the problem will become complicated for other restraint condition, e.g. spectral efficiency η . Since $h(t)$ should be related to the “time–bandwidth production” of it. First what is required “time–bandwidth production”? 2nd there is so many definitions of “bandwidth” and “time duration” [12]. Which definition is suitable? Or need to re-define? According to the uncertainty principle strictly limited bandwidth and time duration signal is physically unrealizable. The optimum $h(t)$ is no longer the rectangular one. For non-flat $h(t)$, Code constraint length and relation should all be changed. That is a problem of that when moving a little part the whole situation will be complete different. Such question can't be solved by nowadays coding theory. Several $h(t)$ are given for simulation in the paper. They are rectangular, raised cosine, raised cosine spectrum waveform as well as their compounds. Since employing rectangular $h(t)$ has some theoretical meaning, and employing other $h(t)$ can be approached by filtering in engineering. It is lucky that their performance difference is not huge relative to their gain. Which tell us that even the optimum $h(t)$ is found, the performance improvement may not huge. Any way finding optimum $h(t)$ is still an open problem.

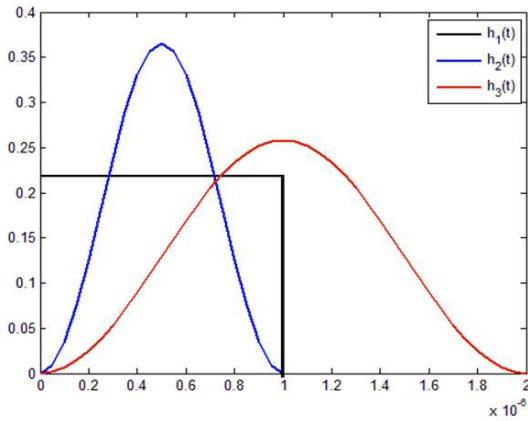


Fig 3. Some basic Multiplexing waveforms

Some tested $h(t)$ in the paper as in [12] are as follows:

$h_1(t) \sim$ Rectangular wave with duration T ,

$$B_e = 1/T, B_0 = 1/T.$$

$h_2(t) \sim$ Raised Cosine wave 1 with duration T ,

$$B_e = 1/T, B_0 = 2/T;$$

$h_3(t) \sim$ Raised Cosine wave 2 with duration $2T$,

$$B_e = 1/2T, B_0 = 1/T.$$

$$h_8(t) = \frac{1}{2} \sin c(\pi t/T) \frac{\cos(\pi t/T)}{(1-t^2/T^2)} \sim \text{Raised Cosine spectrum}$$

wave with most energy within $(-T, T)$, $B_e = 1/T, B_0 = 2/T$.

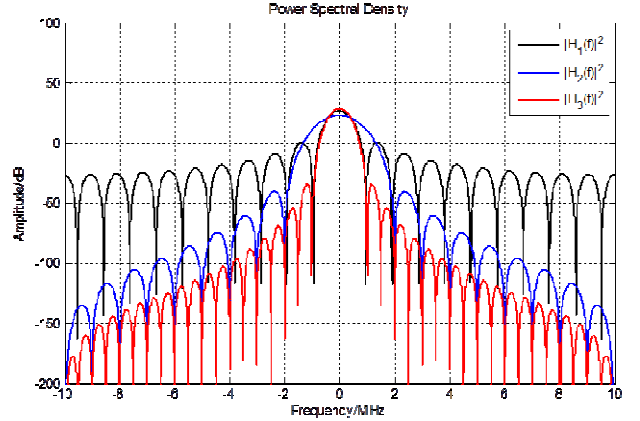


Fig 4. Power spectrum of the basic Multiplexing waveforms

$$h_4(t) = h_1(t) \otimes h_2(t) \sim 2T, B_e = 0.6/T, B_0 = 1/T.$$

$$h_5(t) = h_1(t) \otimes h_3(t) \sim 3T, B_e = 0.4/T, B_0 = 1/T.$$

$$h_6(t) = h_2(t) \otimes h_2(t) \sim 2T, B_e = 0.75/T, B_0 = 1/T.$$

$$h_7(t) = h_2(t) \otimes h_3(t) \sim 3T, B_e = 0.45/T, B_0 = 1/T.$$

The power spectrum of $h_l(t)$ ($l=4,5,6,7$) are the production of their components.

Where: $\sim aT, a=1,2,3$ denotes the time duration of $h_l(t), l=1,2,\dots,7$; B_0, B_e respectively denote the first zero bandwidth and the equivalent noise bandwidth, the last is a fictive rectangular bandwidth of high $H_l(0)$ ($l=1,2,\dots,8$) with the same filtered power in white Gaussian noise.

OVTDM is a waveform coding scheme. The system performance is controlled by $h(t)$. People may worry about $h(t)$ would be destroyed by multipath Rayleigh fading channel. However the random time dispersion of the channel (which causes frequency selective fading) may put additional overlapping to $h(t)$ and has no effect to spectral efficiency η . On the contrary, it is a beneficial factor of improving system performance. Duo to that the additional overlapping increases the coding and implicit diversity gain simultaneously.

3. A Bit Error Probability upper Bound of OVTDM

Just like the convolutional codes, Evaluating on accurate OVTDM's bit error probability is very difficult. That is still an open problem. There were many error probability upper bound

of ISI channel, but all gave a pessimistic and no general result. This chapter will give an optimistic and general result by a “Modified Minimum Euclidean Distance Sphere Bound” [2] and a non-normalized masked distribution.

A Node error event of OVTDM

An error event begins at $t = jT$, ends at

$$t = (j+k+K)T, k = 0, 1, 2, \dots$$

Suppose the correct data sequence be

$\mathbf{u} \triangleq [u_0, u_1, \dots, u_{j-1}, u_j, u_{j+1}, \dots, u_{j+k+K}, u_{j+k+K+1}, \dots, u_{N-1}]$, One error sequence with S errors from \mathbf{u} be

$$\mathbf{u}'_S \triangleq [u_0, u_1, \dots, u_{j-1}, u'_j, u'_{j+1}, \dots, u'_{j+k+K}, u_{j+k+K+1}, \dots, u_{N-1}],$$

$$\mathbf{e}_S \triangleq \frac{1}{2}(\mathbf{u} - \mathbf{u}'_S) = [0, 0, \dots, e_j, e_{j+1}, \dots, e_{j+k}, 0, \dots, 0], \text{ be the}$$

error sequence with S errors, where $e_n \in \{1, 0, -1\}$, $\sum_n e_n^2 = s$, and no consecutive $K-1$ position with no errors, i.e. no $\frac{00\dots 0}{\geq K-1}$ between adjacent $e_n \neq 0$. There are totally $2^S(K-S)^{S-1}$ \mathbf{e}_S event with shortest length $S+K-1$ and longest length $(K-1)(S-1)+K$.

B Node error probability and Bit error probability of OVTDM

Let \mathbf{u} be the correct data sequence, \mathbf{u}'_S be its only alternative, than their pairwise probability is

$$P(\mathbf{u} \rightarrow \mathbf{u}'_S) \triangleq P_e(\mathbf{e}_S) = \text{erfc}\left\{\left[\frac{1}{2N_0} \int_0^{NT+\Delta} |s(t) - s'(t)|^2 dt\right]^{1/2}\right\}, \quad (8)$$

$$\text{Where: } \text{erfc}(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\omega^2/2} d\omega,$$

$$s(t) = \frac{\sqrt{2E_s}}{2} \sum_n u_n h(t-nT) \sim \text{is the correct signal},$$

$$s'(t) = \frac{\sqrt{2E_s}}{2} \sum_n u'_n h(t-nT) \sim \text{is its only alternative}.$$

The Euclidean distance between $s(t)$ and $s'(t)$ is

$$\int_0^{NT+\Delta} |s(t) - s'(t)|^2 dt = 2E_s \sum_n \sum_m e_n e_m h_{n-m}^0, \quad (9)$$

$$(n, m = 0, 1, \dots, N-1)$$

Where:

$$\tilde{h}_{n-m}^0 \triangleq \int_0^{NT+\Delta} h(t-nT) h^*(t-mT) dt = \tilde{h}_{m-n}^{0*} \quad (10)$$

$$(|n-m| = 0, 1, \dots, K-1)$$

$$\text{Since } \sum_n e_n^2 = s, \quad h_0^0 = 2,$$

$$\text{Let: } h_l^0 \triangleq \text{Re } \tilde{h}_l^0 =$$

$$\text{Re} \int_0^{\Delta-lT} h(t) h^*(t+lT) dt, \quad l = 0, 1, \dots, K-1, \quad (11-1)$$

$$\sigma_l \triangleq \frac{1}{S} \sum_n e_n e_{n-l}, \quad (11-2)$$

$$\varepsilon_S \triangleq \sum_{l=1}^{K-1} \sigma_l h_l^0, \quad (11-3)$$

$$d^2 \triangleq E / N_0, \quad (11-4)$$

(11-4) is so called the normalized SNR. Thus (8) becomes

$$P_e(\mathbf{e}_S) = \text{erfc}\left\{[2sd^2(1+\varepsilon_S)]^{1/2}\right\} < \frac{1}{2} \exp\{-sd^2(1+\varepsilon_S)\}, \quad (12)$$

People are interested in when a node error event occurs, the probability of average S bits in error, i.e.

$$P_e(s) = P_e(\cup \mathbf{e}_S), \quad (13)$$

Since in Trellis diagram of OVTDM, each node represents one bit entering into the channel, the bit error probability P_b is

$$P_b = E\{n_b\} = \sum_{s=1}^{\infty} s P_e(s) = P_e(\cup \mathbf{e}_S) \leq \sum_{\mathbf{e}_S} P_e(\mathbf{e}_S), \quad (14)$$

Since there are $2^S(K-1)^{S-1}$ different \mathbf{e}_S with S errors, the union bound on $P_e(s)$ becomes

$$P_e(s) \leq 2^S(K-1)^{S-1} E\{P_e(\mathbf{e}_S)\}, \quad (15)$$

Such a bound only suitable for small S and K . when $S \gg 1$ or $K \gg 1$, (15) may greater than 1 and becomes useless. A best way is to use the “Modified Minimum Euclidean Distance Sphere Bound” [2], that gives

$$P_e(s/\mathbf{u}) \leq (s-1)(K-1)E\{e^{-sd^2(1+\varepsilon_S)} / \mathbf{u}\}, \quad (s \geq 2, K > 2), \quad (16)$$

There is equal probable $(s-1)(K-1)$ \mathbf{u}'_S surrounding \mathbf{u} , then

$$E\{e^{-sd^2(1+\varepsilon_S)} / \mathbf{u}\} \geq \frac{1}{(s-1)(K-1)} e^{-sd^2[1+\text{Min } \varepsilon_S(\mathbf{u})]}, \quad (K > 2) \quad (17)$$

Where: $\varepsilon_{S,i}(\mathbf{u})$ is the conditional ε_S under \mathbf{u} and its only alternative is $\mathbf{u}'_{S,i}$. Thus

$$P_e(s) = \sum_u P(u) P_e(s/u) \leq (s-1)(K-1) \sum_u P(u) E\{ \exp -sd^2(1+\varepsilon_s) \} \\ = (s-1)(K-1) E\{ \exp -sd^2(1+\varepsilon_s) \}, \quad (s \geq 2, K > 2), \quad (18)$$

Note: When $K=2$, u_s with s bits error from u is only consecutive s bits different from u , therefore (17) is hold only for $K > 2$. There should be no $(s-1)(K-1)$ coefficient for $K=2$.

When $S=1, \varepsilon_s=0$, we have

$$P_e(1) = \text{erfc}(2d^2) < e^{-d^2}, \quad (19)$$

An upper bound on P_b is

$$P_b = \sum_{s=1}^{\infty} s P_e(s) \leq e^{-d^2} + \\ + (K-1) \sum_{s=2}^{\infty} (s^2 - s) e^{-sd^2} E\{e^{-sd^2\varepsilon_s}\}, \quad (K > 2) \quad (20)$$

To find an upper bound on P_b , the evaluation of $E\{e^{-sd^2\varepsilon_s}\}$ or its upper bound is of importance.

C One upper bound on Bit error probability of OVTDM
 $K=2$

At such case $\varepsilon_s = \sigma_1 h_1$, and according to appendix B, σ_l only exists even moments as

$$E\{\sigma_l^{2k}\} = \frac{1}{2^{s-1} s^{2k}} \sum_{i=0}^{s-1} \binom{s-1}{i} (s-2i-1)^{2k}, \quad (21) \\ k=0,1,2,\dots, l=1,2,\dots, K-1,$$

$$P_e(s) \leq e^{-sd^2} E\{e^{-sd^2\sigma_1 h_1}\}$$

$$= e^{-sd^2} \sum_{k=0}^{\infty} \frac{(sd^2 h_1)^{2k}}{(2k)!} \cdot \frac{1}{2^{s-1} s^{2k}} \sum_{i=0}^{s-1} \binom{s-1}{i} (s-2i-1)^{2k} \\ = \frac{1}{2^s} e^{-sd^2} \sum_{i=0}^{s-1} \binom{s-1}{i} \left[e^{(s-2i-1)h_1 d^2} + e^{-(s-2i-1)h_1 d^2} \right] \quad (22) \\ = \frac{1}{2^{s-1}} e^{-sd^2} \left[e^{h_1 d^2} + e^{-h_1 d^2} \right]^{s-1},$$

Finally
We have

$$P_b \leq e^{-d^2} \sum_{s=1}^{\infty} s \left[\frac{1}{2} e^{-d^2} (e^{h_1 d^2} + e^{-h_1 d^2}) \right]^{s-1} \\ = e^{-d^2} \left[1 - \frac{1}{2} e^{-d^2} (e^{h_1 d^2} + e^{-h_1 d^2}) \right]^{-2}, \quad (23)$$

(23) is identical to the bound given by Viterbi and Omura [1].

For any physical realizable channel, $h_1^0 < 1$, then for $d^2 \gg 1$

$$P_b < e^{-d^2}, \quad d^2 \gg 1; K=2, \quad (24)$$

The least favorite multiplexing waveform is the rectangular one with $h_1^0 = 1$,

$$h(t) = \sqrt{\frac{2}{KT}} [u(t) - u(t-KT)], \quad K=2, \quad (25)$$

$$P_b \leq 4e^{-d^2} \quad d^2 \gg 1; K=2, \quad (26)$$

$K > 2$

Although the distribution of $\sigma_l (l=1,2,\dots,K-1)$ is known, due to the strong dependence among σ_l , to find the distribution of $P_r(\varepsilon_s)$ is still difficult. So far most of people employ numerate method, but it is unreality when $s \gg 1$ or $K \gg 1$.

Since $E(\varepsilon_s) = 0$, $E\{e^{-sd^2\varepsilon_s}\}$ is controlled by $\varepsilon_s < 0$, especially its terminal of $\varepsilon_s < 0$ when $d^2 \gg 1$. That requires finding the minimum energy of $\sum_n e_n h(t-nT)$, such question had been studied by too many people in the past, there is impossible to list them all. On the other hand $\min \varepsilon_s$ must be related to a least favorite node error sequence, meaning

$$P_r\{\varepsilon_s = \min \varepsilon_s\} = P_r(\sigma_1 = \min \sigma_l). \quad (27)$$

Node: If σ_1 is not the minimum, need reordering.

At the terminal of $\varepsilon_s < 0$, Let $\varepsilon_s \cong a(K,s)\sigma_1$,

Where:

$$a(K,s) \triangleq \frac{\min \varepsilon_s}{\min \sigma_l} = \frac{s}{s-1} |\min \varepsilon_s|, \quad (28)$$

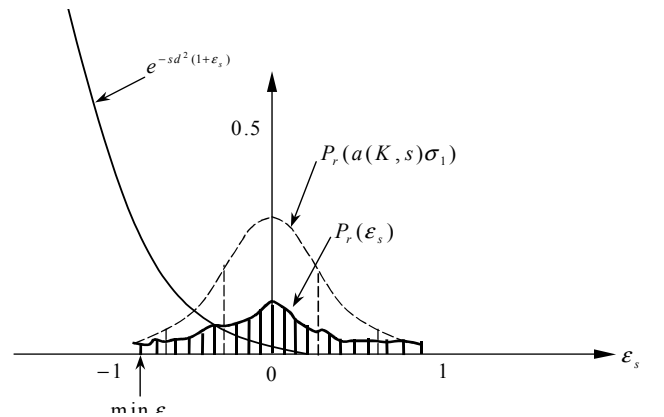


Fig 5. The difference between $P_r(\varepsilon_s)$ and $P_r(a(K,s)\sigma_1)$.

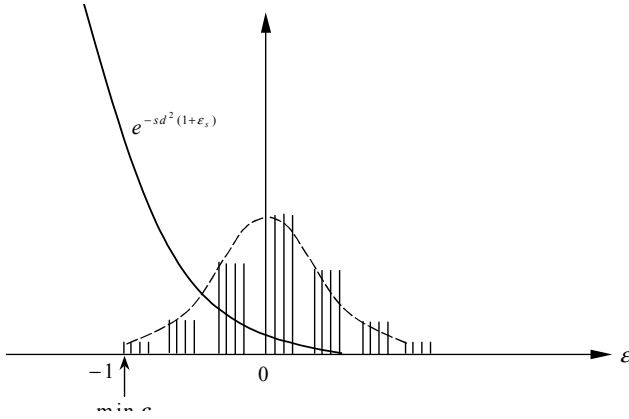


Fig 6. Non-normalized $P_r[a(K,s)\sigma_1]$ with $K-1$ times spectrum line

Fig.5 shows the difference between the numerical evaluated real $P_r(\epsilon_s)$ and $P_r(a(K,s)\sigma_1)$. ϵ_s , $a(K,s)\sigma_1$ are all discrete random variables, $P_r(a(K,s)\sigma_1)$ is always above $P_r(\epsilon_s)$, due to $P_r(\epsilon_s)$ has $K-1$ times more spectrum lines than $P_r(a(K,s)\sigma_1)$, and at the terminal of $\epsilon_s < 0$, this two probabilities are equal.

Since ϵ_s is consisted of the summation of $K-1$ weighted identical distributed dependent $\sigma_l (l=1,2,\dots,K-1)$. Between adjacent two spectrum lines of $a(K,s)\sigma_1$, there are no more than $K-1$ spectrum lines of ϵ_s .

Since $P_r(\epsilon_s)$ is unknown, $P_r(a(K,s)\sigma_1)$ is known. If adding $K-1$ lines between two adjacent lines of $a(K,s)\sigma_1$, a new non-normalized distribution $P_r[a(K,s)\sigma_1]$ will be obtained, that has total probability greater than 1, and $P_r(\epsilon_s)$ will be masked by it (Fig.6), except in very few case, at the terminal of $\epsilon_s > 0$. However it is no effect to the final integration, due to at $\epsilon_s > 0$, $e^{-sd^2(1+\epsilon_s)} \ll 1$, the final integration is controlled only by $\epsilon_s < 0$ (Fig.5).

If replacing $P_r(\epsilon_s)$ by $P_r[a(K,s)\sigma_1]$, the expectation operation for $e^{-sd^2\epsilon_s}$ will be increased $A(K,d,s)$ times.

Where:

$$A(K,d,s) = \sum_{i=0}^{K-1} e^{-i \frac{a_s d^2}{K-1}} = \frac{1 - e^{-a_s d^2}}{1 - e^{-\frac{a_s d^2}{K-1}}} \quad (29)$$

$$0 < a_s < |\text{Min } \epsilon_s| / s.$$

When $d^2 \gg K$, $A(K,d,s) \cong 1$;

$$K \gg d^2, A(K,d,s) \cong K-1;$$

Otherwise $1 \leq A(K,d,s) \leq K-1$.

After such modification, we have

$$P_e(s) \leq A(K,d,s)(K-1)(s-1)e^{-sd^2} E\{e^{-sd^2 a(K,s)\sigma_1}\}, \quad (30)$$

Since

$$E\{e^{-sd^2 a(K,s)\sigma_1}\} = \sum_{k=0}^{\infty} \frac{(sd^2 a)^{2k}}{(2k)!} E\{\sigma_1^{2k}\}, \quad (31)$$

Substitute $E\{\sigma_1^{2k}\}$ into (31), we have

$$P_e(s) < e^{-d^2} \frac{A(K,d,s)(s-1)}{(s-1)^{s-2}} B^{s-1}(K,d,s), \quad (32)$$

Where

$$B(K,d,s) \triangleq (K-2)e^{-d^2} + \frac{1}{2}e^{-d^2} [e^{-a(K,s)d^2} + e^{a(K,s)d^2}], \quad (33)$$

$$\text{Thus } P_b < e^{-d^2} \{1 + (K-1) \sum_{s=2}^{\infty} (s^2 - s) [\frac{1}{K-1} B(K,d,s)]^{s-1}\},$$

$$K > 2, \quad (34)$$

$$\text{Especially } P_b < e^{-d^2} \{1 + (K-1) \sum_{s=2}^{\infty} (s^2 - s) [\frac{1}{K-1} B(K,d,s)]^{s-1}\},$$

$$d^2 \gg K, K > 2, \quad (35)$$

$$P_b < e^{-d^2} \{1 + (K-1) \sum_{s=2}^{\infty} (s^2 - s) A(K,d,s) [\frac{1}{K-1} B(K,d,s)]^{s-1}\},$$

$$K \gg d^2, K > 2, \quad (36)$$

Now the key point becomes to evaluate $a(K,s)$ or equivalently the minimum Euclidean distance between any two received signals. Such task seems very difficult, due to the fact when $K \gg 1$ or $s \gg 1$, the total number of the node error sequences is very huge. However, there is no need to evaluate all $a(K,s)$, only evaluate several small s 's $a(K,s)$ is enough. Then, in general such evaluations is not difficult but easy. Even most physical $a(K,s)$ can be found by intuition way (appendix A). The reason is that, for the physical realizable $h(t)$, series $|h_1^0|, |h_2^0|, \dots, |h_{K-1}^0|$ is strictly monotonic decreasing with attenuation coefficient less than any arithmetic one. If the evaluation from $s=1, 2, \dots$, until s_M , $a(K, s_M) \leq 1$, then just let $a(K,s) = a(K, s_M), \forall s > s_M$, and s_M usually is not large.

Thus for $|h_1^0| \leq 1$ channel, a loser upper bound on P_b can be obtained by let all $a(K,s) = 1$ as

$$P_b < e^{-d^2} [1 + (K-1) \sum_{s=2}^{\infty} \frac{(s^2 - s)}{(K-1)^{s-1}} A(K,d,s) B^{s-1}(K,d)],$$

$$(K > 2, |h_1^0| \geq 1, \quad (37)$$

Where

$$B(K,d) \triangleq (K-2)e^{-d^2} + \frac{1}{2}(1 + e^{-2d^2}), \quad (38)$$

When $d^2 \gg 1$, we have

$$P_b < e^{-d^2} [1 + (K-1) \sum_{s=2}^{\infty} \frac{(s^2-s)}{2^{s-1}(K-1)^{s-1}}] \\ = e^{-d^2} \{1 + [1 - \frac{1}{2(K-1)}]^3\}, \quad (d^2 \gg K, K > 2, |h_1^0| \leq 1), \quad (39)$$

It is a pity that $|h_1^0| \leq 1$ exists only for small K . However, no matter in what situation, in $\sum_n e_n \tilde{h}(t-nT)$ at least one whole $h(t-n'T)$ should be survived for odd s . Therefore for odd s , $|h_1^0| > 1$ is impossible, $|h_1^0| > 1$ only exists for even s . Since $a(K, s)$ is monotonic decreasing function of s , there must exist a s_M , for $s \geq s_M$, $a(K, s) \leq 1$, and let $a(K, s) = 1$ ($s \geq s_M$). Thus only the term with $a(K, s) > 1$ should be remained. Therefore a loser upper bound on P_b can be obtained as

$$P_b < e^{-d^2} \{1 + (K-1) \sum_{s=2}^{\infty} \frac{(s^2-s)}{(K-1)^{s-1}} A(K, d, s) B^{s-1}(K, d)\} \\ + (K-1) \sum_{s=2,4,\dots,s_M} \frac{s^2-s}{(K-1)^{s-1}} A(K, d, s) \\ \times [B^{s-1}(K, d, s) - B^{s-1}(K, d)], \quad (K > 2). \quad (40)$$

When $d^2 \gg K$,

$$P_b < e^{-d^2} \{1 + [1 - \frac{1}{2(K-1)}]^3 \\ + (K-1) \sum_{s=2,4,\dots,s_M} \frac{s^2-s}{2^{s-1}(K-1)^{s-1}} [e^{(s-1)[a(K,s)-1]d^2} - 1]\} \\ (K > 2, d^2 \gg K), \quad (41)$$

When $K \gg d^2$,

$$P_b < e^{-d^2} \{[1 + 2(K-1)B(K, d)(1 - \frac{1}{K-1} B(K, d))^{-3}] \\ + (K-1)^2 \sum_{s=2,4,\dots,s_M} \frac{s^2-s}{(K-1)^{s-1}} [B^{s-1}(K, d, s) - B^{s-1}(K, d)]\} \\ (K > 2, K \gg d^2). \quad (42)$$

and (42) are the final results, In fact (42) is suitable for any SNR d^2 , due to it uses the largest upper bound $(K-1)$ on $A(K, d, s)$.

D Bit error probability upper bound of OVTDM for special multiplexing waveform $h(t)$ ($K \gg 1, d^2 \gg K$)

Two special multiplexing waveforms $h(t)$ are considered in this section, they are:

1) Rectangular waveform

$$h(t) \triangleq \sqrt{\frac{2}{KT}} [u(t) - u(t-KT)];$$

2) Truncated exponential waveform

$$h(t) \triangleq 2\sqrt{\alpha} e^{-\alpha} [u(t) - u(t-KT)], \quad \alpha = b/KT, \quad b \geq 7$$

Note: Larger truncate number b will cause smaller cut off power that becomes interference, b has no effect to the conclusion!

Due to the following reasons to study above two waveforms have both theoretical and engineer importance:

1. According to the symmetry principle, rectangular $h(t)$ uniformly distributes signal's energy to all time delay, it belongs to the least favorite waveform, but has the simplest decoding complexity at condition of all $h(t)$ have same η .
2. Physical realizable $h(t)$ can be looked upon as the different exponential waveforms' linear combination.
3. By employing a special "Perfect Complete Complementary Orthogonal Code Pairs Mate" [12], rectangular $h(t)$ can be working well in OFDM system without losing performance.

1. Rectangular $h(t)$

$$\text{Its } h_l^0 = 2(1 - l/K), \quad (l = 1, 2, \dots, K-1), \quad (43)$$

Its least favorite node error sequence is

$$xx \underbrace{00 \dots 0}_{K-2} xx \underbrace{00 \dots 0}_{K-2} \dots xx \underbrace{00 \dots 0}_{K-2}$$

Where: xx denotes polar alternative $+$ $-$ errors. The least favorite waveform of $\sum_n e_n h(t-nT)$ is shown in Fig.7 as

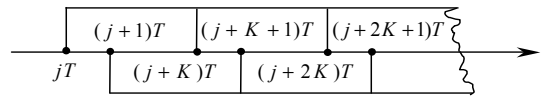


Fig. 7. Least favorite $\sum_n e_n h(t-nT)$ for rectangular $h(t)$

It can be found:

$$\text{Min } \varepsilon_s = \begin{cases} -1 + 2/sK, & (s = 2, 4, 6, \dots) \\ -(s-1)/s, & (s = 3, 5, 7, \dots) \end{cases} \quad (44)$$

$$a(K, s) = \begin{cases} \frac{s}{s-1} [1 - \frac{2}{sK}], & (s = 2, 4, 6, \dots) \\ 1, & (s = 3, 5, 7, \dots) \end{cases} \quad (45)$$

Thus an upper bound for rectangular $h(t)$ on P_b is

$$\begin{aligned}
P_b &= P_e(1) + \sum_{s=3,5,7,\dots} sP_e(s) + \sum_{s=2,4,6,\dots} sP_e(s) \\
&\leq e^{-d^2} \{1 + (K-1) \sum_{s=3,5,7,\dots} \frac{s^2-s}{2^{s-1}(K-1)^{s-1}}\} \\
&\quad + (K-1)e^{-2d^2/K} \sum_{s=2,4,6,\dots} \frac{s^2-s}{2^{s-1}(K-1)^{s-1}} \\
&= \frac{1}{2}e^{-d^2} \left\{ 2 + \left[1 - \frac{1}{2(K-1)}\right]^{-3} + \left[1 + \frac{1}{2(K-1)}\right]^{-3} \right\} \\
&\quad + \frac{1}{2}e^{-2d^2/K} \left\{ 1 - \left[1 - \frac{1}{2(K-1)}\right]^{-3} - \left[1 + \frac{1}{2(K-1)}\right]^{-3} \right\}, (d^2 \gg K),
\end{aligned} \quad (46)$$

$$P_b < 2e^{-d^2} + \frac{3}{2(K-1)}e^{-2d^2/K}, (d^2 \gg K, K \gg 1) \quad (47)$$

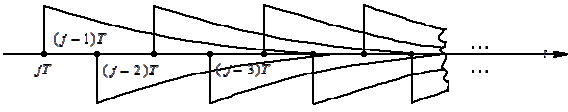


Fig 8. Least favorite $\sum e_n h(t-nT)$ for truncated exponential $h(t)$

2. Truncated exponential $h(t)$

When $K \gg 1$,

$$h_l^0 = 2e^{-2\alpha T} \cong 2(1-2\alpha T), (l=1,2,\dots,K-1), \quad (48)$$

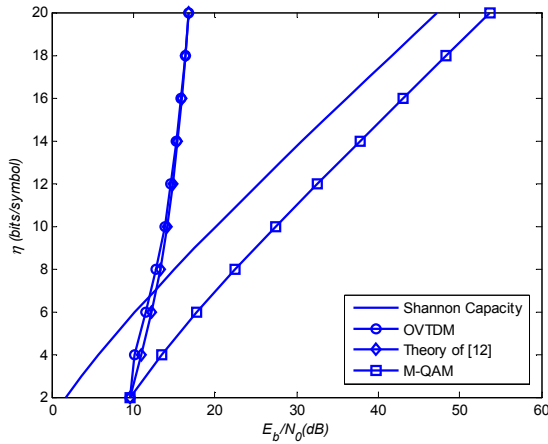


Fig 9. Simulation and theoretical Comparison for rectangular $h(t)$ $\{\eta \sim \text{spectral efficiency (bits/symbol)}, d^2 = E_b / N_0, P_b = 10^{-5}\}$

The least favorite error sequence is the continues polar alternative $+-+-+ \dots$ one (Fig.8)

It can be found:

$$a(K,s) = \begin{cases} \frac{s}{s-1} [1 - \alpha T], & (s=1,3,5,\dots) \\ 1, & (s=2,4,6,\dots) \end{cases}, \quad (49)$$

Thus when $d^2 \gg K$, an upper bound on P_b for truncated exponential $h(t)$ is

$$\begin{aligned}
P_b &\leq e^{-d^2} \{1 + (K-1) \sum_{s=3,5,7,\dots} \frac{s^2-s}{2^{s-1}(K-1)^{s-1}}\} \\
&\quad + (K-1)e^{-\alpha T d^2} \sum_{s=2,4,6,\dots} \frac{s^2-s}{2^{s-1}(K-1)^{s-1}} e^{-(s-1)\alpha T d^2} \\
&= \frac{1}{2}e^{-d^2} \{2 + [1 - 1/2(K-1)]^{-3} + [1 + 1/2(K-1)]^{-3}\} \\
&\quad + \frac{1}{2}e^{-2\alpha T d^2} \{[1 - 1/2(K-1)]^{-3} + [1 + 1/2(K-1)]^{-3}\} \\
&\quad d^2 \gg K
\end{aligned} \quad (50)$$

Since $\alpha T = b/K$, then when $d^2 \gg K$,

$$\begin{aligned}
P_b &\leq \frac{1}{2}e^{-d^2} \{2 + [1 - 1/2(K-1)]^{-3} + [1 + 1/2(K-1)]^{-3}\} \\
&\quad + \frac{1}{2}e^{-2bd^2/K} \{[1 - 1/2(K-1)]^{-3} + [1 + 1/2(K-1)]^{-3}\}, \quad (51) \\
&\quad d^2 \gg K
\end{aligned}$$

(5 When $d^2 \gg K, K \gg 1$,

$$P_b < 2e^{-d^2} + \frac{3}{2(K-1)}e^{-2bd^2/K}, (d^2 \gg K, K \gg 1), \quad (52)$$

After deeply study (46), (47), (51), (52), It may be find when d^2/K is kept a constant, P_b are roughly the same. That means that spectral efficiency η (or channel capacity) of OVTDM is roughly proportional to the normalized SNR d^2 . Such conclusion is identical to the conjecture of [12] and also similar to [15], since no matter rectangular or exponential $h(t)$, their bandwidth is defined only by engineering sense rather than Nyquist strictly sense. Regardless the decaying speed, the spectrum tail always extends to infinite. When the strict bandwidth is infinite, for Nyquist $H(f)$ [15] has proved that the channel capacity is linearly to SNR. For any $H(f)$ appendix C get the same conclusion when SNR is high. Simulations shown next in the paper also verify such conclusion. Since among different $h(t)$ with identical η , Rectangular $h(t)$ is the worst due to the Symmetric Principle and shortest constraint length, And among lots of simulated $h(t)$, when $\eta > 6$, their bit error probability performances all close to such bound (Fig.13). Therefore (46), (47) can be looked upon as an upper bound on bit error probability for any $h(t)$.

4. Two Stage Concatenated OVTDM

A Two stage Concatenated OVTDM structure and Implementation

Two stage concatenate OVTDM structure parallel putting in orthogonal I, Q channels is proposed in Fig. 10, where the 1st stage is a K_1 th order pure OVTDM (no relative shift) changing binary (+1,-1) input into multilevel real input and the 2nd stage is a K_2 th order shifted OVTDM making its polynomial distribution output approach to Gaussian

distribution. The total spectral efficiency of such I, Q parallel concatenate OVTDM structure is $\eta = 2K_1K_2$ bits/symbol, and I, Q real distribution outputs together approach to complex Gaussian distribution.

The multiplexing waveform of the 1st stage OVTDM is a rectangular one with duration K_1T_b . The multiplexing waveform of the 2nd stage OVTDM, denoted within blue line block from transmitter to receiver, is a real waveform $h(t)$ with duration $T_\lambda = \lambda K_1K_2T_b$ ($\lambda \geq 1$). Where $\lambda \geq 1$ is the waveform duration expanding coefficient corresponding to a rectangular $h(t)$ after filtering. Time duration T_λ of $h(t)$ is the time width of occupying at least 99.9% of the total energy

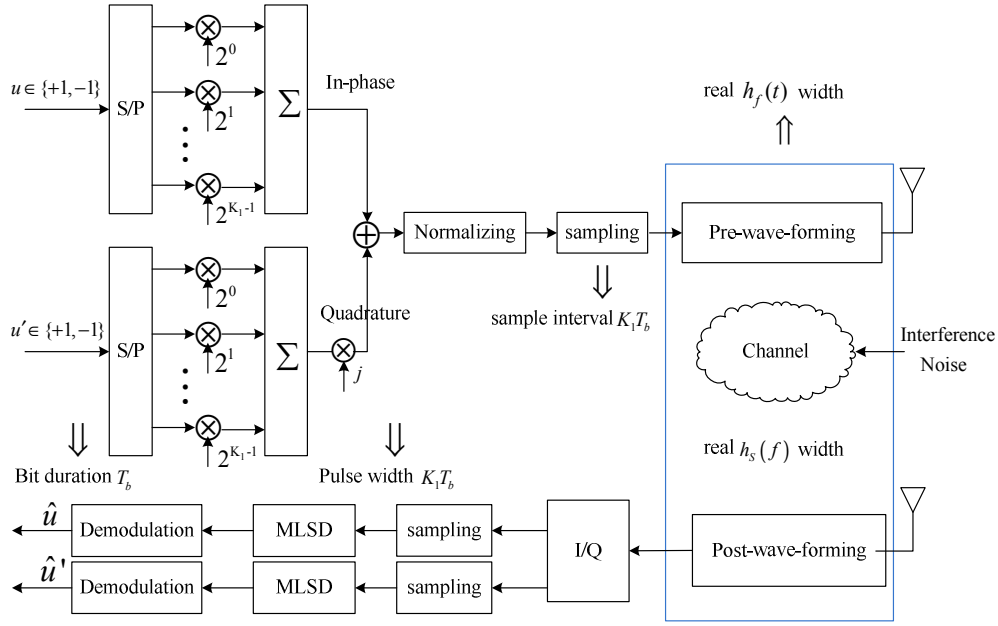


Fig 10. Concatenate OVTDM structure. The Multiplexing waveform of the 1st stage is a rectangular one of width K_1T_b , The Multiplexing waveform of the 2nd stage is $h(t)$ of width $T_\lambda = \lambda K_1K_2T_b$, $\eta = 2K_1K_2$.

The 2nd stage of shifted OVTDM with multiplexing waveform $h(t)$ can be also denoted by Fig.10, with shifted interval K_1T_b and overlapping folds K_2^λ . The impulse response of $h_f(t)$ consists of all filters in transmitter, e.g. root transmitter filter, pre-wave-forming, pre-equalizer etc. The impulse response of $h_s(t)$ consists of all filters in receiver, e.g. root receiver filter, post-wave-forming, post-equalizer etc. The total multiplexing waveform of the 2nd stage OVTDM is $h(t) = h_f(t) \otimes h_s(t)$.

MLSD should be employed in such concatenated OVTDM system. From the total $2^{K_1K_2^\lambda}$ possible waveform sequences in each I and Q channel, to select the most possible concatenated OVTDM coded waveform sequence that is nearest to the received signal waveform sequence [2].

Coding steps:

Let the I, Q orthogonal channels' binary (+1,-1) data input sequences be

$$I: \dots, u_{c,n-1}^T, u_{c,n}^T, u_{c,n+1}^T, \dots$$

of $h(t)$, depending on the required η . Because the cut off energy of $h(t)$ will become interference, which should be less than the threshold SNR's noise level at least 10-15 dB, larger η should choose larger percentage of the total energy of $h(t)$ (larger λ).

The shift interval of the 2nd stage OVTDM is K_1T_b . In order to keep the spectral efficiency η unchanged in a filtered system. Overlapped folds of the 2nd stage OVTDM would be automatically increased from K_2 to $K_2^\lambda = \lceil \lambda K_2 \rceil$, where $\lceil \cdot \rceil$ is the least integer that is greater or equal to \cdot .

$$Q: \dots, u_{s,n-1}^T, u_{s,n}^T, u_{s,n+1}^T, \dots$$

$$\text{Where } u_{\bullet,n}^T = \underbrace{[u_{\bullet,nK_1}, u_{\bullet,nK_1+1}, u_{\bullet,nK_1+2}, \dots, u_{\bullet,(n+1)K_1-1}]}_{K_1},$$

$\bullet \sim$ denote either c or s.

Performing the 1st stage OVTDM operation. That is the K_1 folds Pure OVTDM operation for I, Q channel with rectangular multiplexing waveform of width K_1T_b respectively.

Performing the 2nd stage OVTDM operation. That is the K_2^λ ($\lambda \geq 1$) folds shifted OVTDM operation for I, Q channel with $h(t)$ of width $T_\lambda = \lambda K_1K_2T_b$ ($\lambda \geq 1$) respectively. Where the shifted interval is K_1T_b , $K_2^\lambda = \lceil \lambda K_2 \rceil$.

Finally the complex envelop of the received signal of such concatenate OVTDM is

$$S(t) = \sqrt{2E} \left[\sum_n u_{c,n}^T H(t - nK_1 T_b) + j \sum_n u_{s,n}^T H(t - nK_1 T_b) \right] + \tilde{n}(t), \quad (53)$$

Where: E is bit energy of the received signal; $n(t)$ is the complex envelop of the noise.

$$H(t) = \left(2^{-K_1+1} \sum_{i=1}^{2^{K_1-1}} (2i-1)^2 \right)^{-\frac{1}{2}} \begin{bmatrix} h(t) \\ 2h(t) \\ \vdots \\ 2^{K_1-1} h(t) \end{bmatrix}, \quad (54)$$

$\left(2^{-K_1+1} \sum_{i=1}^{2^{K_1-1}} (2i-1)^2 \right)^{-\frac{1}{2}}$ is the normalized coefficient;

The multiplexing waveform vector of the 2nd stage OVTDM is $H(t)$;

The multiplexing waveform of the 2nd stage OVTDM is $h(t)$.

Summary: The spectral efficiency of the proposed concatenate OVTDM is $\eta = 2K_1K_2$ bits/symbol. The system bandwidth is determined by $T_\lambda = \lambda K_1 K_2 T_b$ ($\lambda \geq 1$). Power spectrum is determined by $h(t)$. The larger the $K_1 K_2$ or λ , the higher the decoding complexity as well as the narrower the system bandwidth. The number of its output levels is $2^{2K_1} (K_2^\lambda + 1)^2$, However QAM signal's level with the same spectral efficiency is $2^{2K_1 K_2}$, which is much more than the concatenate OVTDM.

B Decoding Algorithms and Complexity of Concatenate OVTDM

Based on the complex convolutional coding model of Fig.2 and the orthogonal property between I, Q channels, Decoding procedure can be done independently in I and Q channel. The optimum decoding algorithm is well known the MLSD, i.e. the Viterbi algorithm with state (node in trellis) number $2^{K_1(K_2^\lambda-1)}$ and input level number 2^{K_1} [2][3][12]. The decoding complexity will be increased exponentially with $K_1 K_2^\lambda$. Fast algorithms like Fano, Stack or Sphere decoding algorithm can also be employed when $K_1 K_2^\lambda$ is large.

C Overlapping Parameter Selection of Concatenate OVTDM

The spectral efficiency η depends on the production of $K_1 K_2^\lambda$. How to put allocation on K_1, K_2^λ ? In fact, K_1 only offer a transform from binary data to multilevel data, it offers no coding gain. All the coding gain is offered by K_2 . For rectangular $h(t)$, According to the symmetric principle, indeed $K_1 = 1$ is the best allocation. However for uneven $h(t)$, the situation is different, it is still an open problem and can only be determined by simulation.

It is well known that under the uncertainty principle time duration limited signal has infinite bandwidth. Frequency

bandwidth limited signal has infinite duration. Both time duration and bandwidth limited signal is physically unrealizable. If a smaller λ (or K_2^λ) is chosen, the decoding complexity is reduced. However the signal power outside its processing "bandwidth" (time duration or frequency bandwidth) is larger, that will become interference. On the contrary, when a larger λ (or K_2^λ) is chosen, the signal power outside its processing "bandwidth" is smaller, but the decoding complexity would be increased. There should be some "compromise" chosen. Surely "compromise" will reduce the system performance and an "error floor" will be appeared, due to that the cut off signal's power (interference) is fixed. The smaller power that outside its processing "bandwidth", the lower the "error floor". "Compromise" should be considered among system "bandwidth", error floor, λ (or K_2^λ). Error floor is determined by the cut off power and the spectral efficiency η . In engineering the following considerations should be noticed:

1. Employing FIR filter with finite time duration $h(t)$;
2. Choosing suitable λ (or K_2^λ);
3. Employing a multiplexing waveform $h(t)$ that is robust from filtering;
4. Employing "Minimum error probability equalizer" [2] to equalize the real multiplexing waveform $h'(t)$ approach to the required $h(t)$.

5. Simulation Results

Some summary simulations of OVTDM employing different multiplexing waveforms are given from Fig.11 to Fig.12. Where the system parameters respectively are $K_1=1$, $K_2=1, 2, 3, 4, 5, 6, 7$, the corresponding spectral efficiency η respectively are 2, 4, 6, 8, 10, 12, 14bits/symbol etc. The spectral efficiency η measured by equivalent noise bandwidth [12] for different multiplexing waveform $h(t)$ is given by Fig.13.

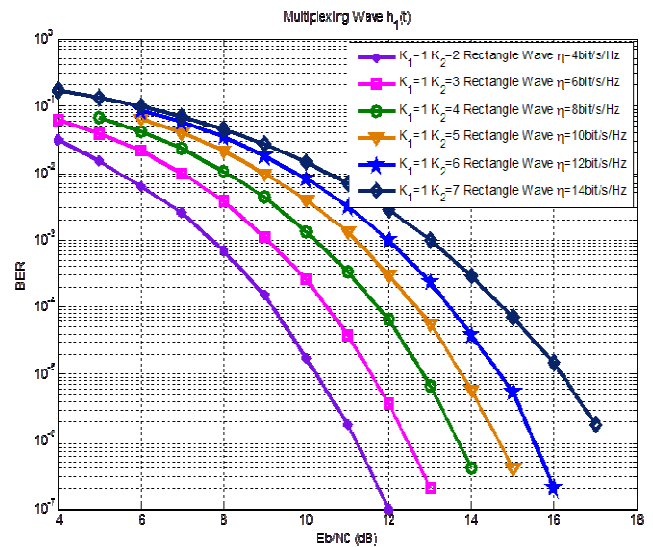


Fig 11. Bit error probability performances of OVTDM with different η , ($h_1(t)$)

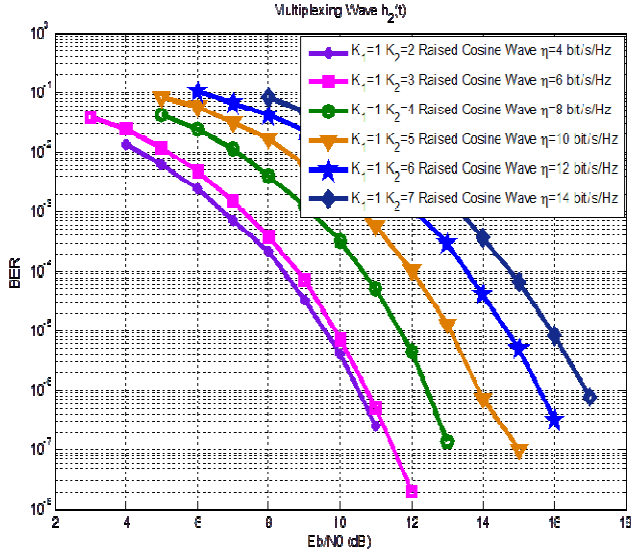


Fig 12. Bit error probability performances of OVTDM with different $\eta, h_2(t)$

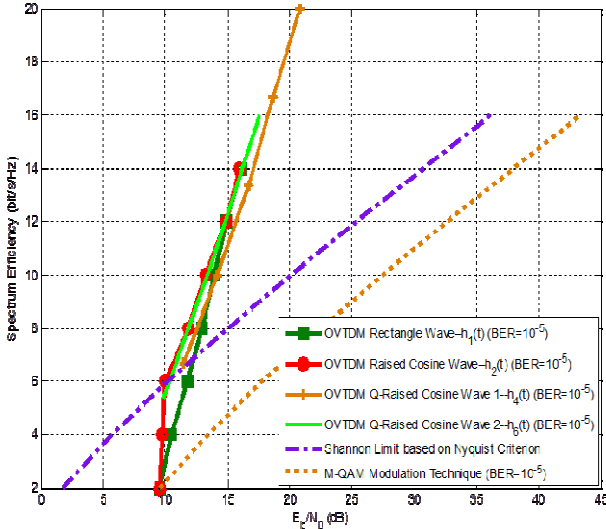


Fig 13. Simulation comparisons of (η, d^2) relations for OVTDM employing different $h(t)$ [12] ($P_b = 1 \times 10^{-5}$, equivalent noise bandwidth)

It can be seen that when $\eta \geq 6$, Relation between η and the threshold $d^2 \triangleq E_b / N_0$ (at BER $P_b = 1 \times 10^{-5}$) is roughly a linear relation, which is identical to the conjecture of [12].

Conclusion: The performance of OVTDM is much better than the M-QAM, Especially when spectral efficiency $\eta \geq 6$, Performance of OVTDM begin to go beyond Shannon limit. The larger the η , the higher the gain over Shannon limit.

6. OVTDM with Pre-Coding

In a pre-coded OVTDM system, the binary (+1,-1) input of Fig.14 is not directly from the source but from a pre-coded output. Fig.14 are some simulations employing TPC (64, 51) and TPC (64, 57) respectively. TPC is the abbreviation of Turbo Product Code or Turbo Array Code. The row and column codes of them are BCH (64, 57) and BCH (64, 51)

respectively. The code rate respectively are $(57/64)2=0.7932$, and $(51/64)2=0.6350$.

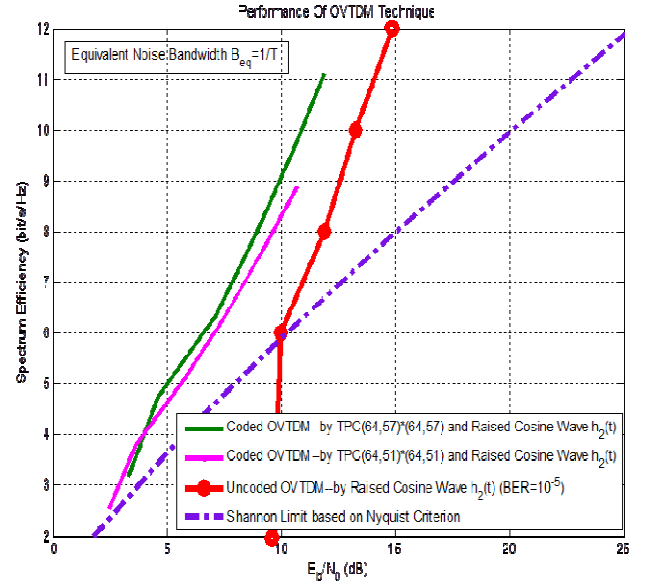


Fig 14. The spectral efficiency of TPC pre-coded OVTDM (equivalent noise bandwidth, $h_2(t)$ [12], $P_b = 1 \times 10^{-5}$)

It can be seen that the pre-coding gain is larger when the spectral efficiency of OVTDM is low. However when the spectral efficiency of OVTDM is getting higher pre-coding gain will become lower. For other kind pre-coding we believe that the same conclusion will be obtained.

7. OVTDM in Random Time Varying Channel

A OVTDM performance in Flat Rayleigh Fading Channel

The performance of OVTDM in flat Rayleigh fading (i.e. no time dispersion) channel is studied in this section. At the receiver, Optimum detection is still the Matched filtering + MLSD. As comparison a high order QAM of the same spectral efficiency η with different explicit multiple-diversity is employed. Because of that OVTDM itself already have some implicit diversity gain, that is similar to the so called "multipath diversity", the larger the K_2 , the higher the implicit diversity gain of the OVTDM. There is never need adding extra diversity for the system employing OVTDM. However the high order QAM can't be worked well in flat Rayleigh fading channel.

Fig.15 shows the bit error probability performance of OVTDM with parameter ($K_1=1, K_2=3, \eta=6$) employing raised cosine multiplexing waveform [12], working in flat fading channel. As comparison a 64QAM with different explicit multiple diversity are also given. The slope of the error probability curve of such OVTDM is about the same with the 7th order explicit diversity of 64QAM, and has about 4dB gain (at BER $P_b = 1 \times 10^{-5}$).

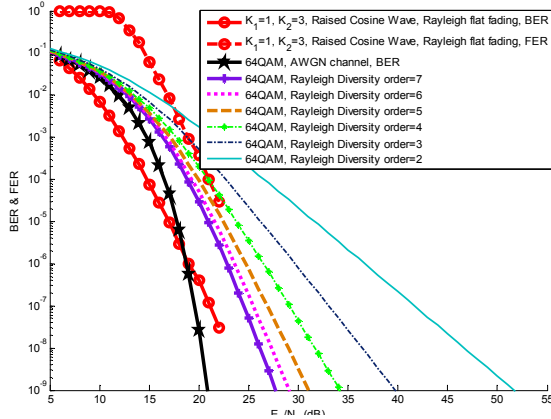
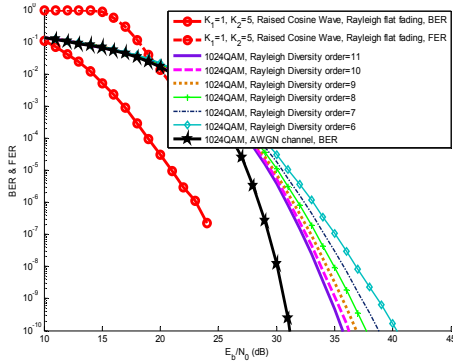
(K1=1, K2=3, $\eta=6$, $h_2(t)$, Frame length 1000)**Fig 15.** Simulation result of OVTDM in flat Rayleigh fading channel

Fig.16 shows the bit error probability performance of OVTDM with parameter ($K_1=1$, $K_2=5$, $\eta=10$) employing raised cosine multiplexing waveform [12], working in flat fading channel. As comparison a 1024QAM with different explicit multiple diversity are also given. The slope of the error probability curve of such OVTDM is about the same with the 11th order explicit diversity of 1024QAM, and has about 8dB gain (at BER $P_b = 1 \times 10^{-5}$).

(K1=1, K2=5, $\eta=10$, $h_2(t)$, Frame length 1000)**Fig 16.** Simulation result of OVTDM in flat Rayleigh fading channel

Conclusion: The performance of OVTDM working in flat Rayleigh fading channels is far beyond the high order QAM with the same η . When $\eta > 6$, Performance of OVTDM working in flat Rayleigh fading channel is even better than the high order QAM with the same η working in AWGN channel.

B OVTDM Multiplexing Waveforms in Multipath Fading Channel

The designed multiplexing waveform and the corresponding spectrum respectively are $h(t)$ and $H(f)$. That is a pair of Fourier transform.

$$h(t) \leftrightarrow H(f) = H^F(f)H^C(f)H^S(f) \quad (55)$$

Where $H^F(f)$, $H^S(f)$, $H^C(f)$ respectively are the transfer functions of transmitter, receiver as well as the channel

For AWGN channel, $H^C(f) \equiv 1$.

However for multipath Rayleigh fading channel, $H^C(f)$ is a random time varying function with impulse response function of its Fourier transform.

$$H^C(f) \leftrightarrow \sum_i \tilde{a}_i \delta(t - \tau_i) \quad (56)$$

Where $\tau_i (i = 0, 1, \dots)$ represent different multipath delay; $\tilde{a}_i = (a_{c,i} + ja_{s,i})$, ($i = 0, 1, \dots$) are all i.i.d. (independent identical distributed) complex zero mean Gaussian random variables with the following properties (Rayleigh fading)

$$\frac{1}{2} E(\tilde{a}_i \tilde{a}_{i'}^*) = \bar{a}_i^2 \delta_{i,i'} (\forall i, i' = 0, 1, \dots);$$

$E(a_{c,i} a_{c,i'}) = E(a_{s,i} a_{s,i'}) = \bar{a}_i^2 \delta_{i,i'} \forall i, i'$, \bar{a}_i^2 denotes the i th path's normalized average power (variance);

$$E(a_{c,i} a_{s,i'}) = E(a_{c,i}) = E(a_{s,i'}) = 0, \forall i, i';$$

Multipath spread $\Delta \triangleq \max_{i \neq j} (\tau_i - \tau_j)$.

Obviously in multipath Rayleigh fading channel, multiplexing waveforms of I and Q orthogonal channels respectively are

$$h^I(t) = \sum_i a_{c,i} h(t - \tau_i), h^Q(t) = \sum_i a_{s,i} h(t - \tau_i) \quad (57)$$

$h^I(t)$, $h^Q(t)$ are all random time varying multiplexing waveforms with waveform duration expanding coefficient λ_c related to Δ . Spectral efficiency η only depends on the original designed $h(t)$ and is unrelated to Δ . Larger Δ can cause larger $K_2^{\lambda_c}$ making decoding complexity increased. However larger additional implicit diversity and coding gain are simultaneously achieved.

Where

$$K_2^{\lambda_c} = \lceil (T_\lambda + \Delta) / K_1 T_b \rceil, \lambda_c = (T_\lambda + \Delta) / K_1 K_2 T_b. \quad (58)$$

C OVTDM performance in Multipath Fading Channel

The bit error probability performance of OVTDM employing Raised Cosine multiplexing waveform in multipath Rayleigh fading channel is studied in this section. At the receiver, Optimum detection is still the Matched filtering + MLSD. However M-QAM can't be worked well in such channel.

In multipath Rayleigh fading channel, according to (57) multiplexing waveform of I and Q channel are no longer the same $h(t)$ but respectively become $h^I(t)$, $h^Q(t)$ which are all random time-varying waveforms with larger waveform duration expanding coefficient λ_c related to the multipath spread Δ . Although spectral efficiency η of OVTDM only depends on the original designed $h(t)$ and is unrelated to Δ . Larger Δ can cause larger λ_c and larger $K_2^{\lambda_c}$ making

decoding complexity increased. However larger additional implicit diversity and coding gain are simultaneously achieved. The larger the λ_c / λ , the larger the additional coding and implicit diversity gain.

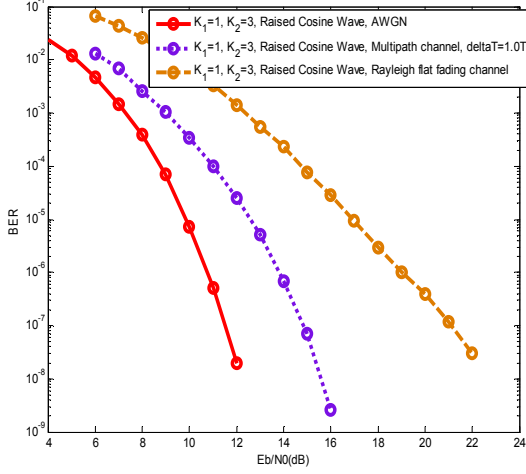


Fig 17. Simulation result of OVTDM in Multipath Rayleigh fading channel ($K_1=1, K_2=3, \eta=6, h_2(t), \Delta=T$)

Fig.17 is one of such examples, which shows the bit error probability performance of OVTDM in multipath Rayleigh fading channel with parameter ($K_1=1, K_2=3, \eta=6$) employing raised cosine multiplexing waveform [12]. Under conditions of the following:

- 1) Ideal channel estimation achieved;
- 2) The time dispersion power spectrum of the channel is uniformly distributed with multipath spread Δ equal to T ;

It can be seen that the performance of OVTDM working in Multipath Rayleigh fading channel is much better than that of in flat Rayleigh fading channel. According to [2], comparing with the uniform distributed time dispersion power spectrum of the Multipath Rayleigh fading channel, for uneven time dispersion power spectrum channel, the bit error probability curve will shift to right with the same slope at the high SNR region. The shifted value is depended on its uneven degree. However no matter which shape of the time dispersion power spectrum of the channel, when $\Delta \gg T$ the performance of OVTDM working in such Multipath Rayleigh fading channel will converge to that of in AWGN channel [2].

8. Conclusion

Nyquist criterion and level division are the two obstacles of communications. After leaving them a novel waveform coding OVTDM by waveform division is revealed in the paper. The required $d^2 \triangleq E_b / N_0$ is roughly proportional to its spectral efficiency (system capacity). That is an amazing result! So far we never find any coded modulation scheme with better performance than OVTDM. However OVTDM is only at its initial stage. OVTDM still has some unsolved open problems. We do hope there appears better waveform coding latter by other people.

Acknowledgment

The author would like to thank The NNSF (National Nature Science Foundation) of China for their continuously long term support by key projects. Thanks also to my colleagues in Daoben Lab. of Research Institute of Tsinghua Univ. China for their testifying simulations by hardware platform. Special thanks also to Dr. J.K. Omura, and Prof. Chen Junbi, Prof. Tong Zhipeng for their helpful suggestions. Thanks also to my hundreds graduate students they had been working hard successively for over 10 years under the guidance of Overlapped Multiplexing Principle, No their simulations, feedback discussions, a simple principle could not become a systematic OVTDM theory. Specially thank to my Ph.D. Students Wang Ying and Liu Bingchao who offered the final simulations of the paper, and Dong Xuan, Gong Peizhu, Zhao Yujie etc. for their final independent testifying simulations.

Appendix

A Some simple waves' $a(K, s)$

To evaluate $a(K, s)$ or equivalently the minimum energy of the least favorite waveform $\sum_n e_n h(t-nT)$ is of importance.

If $h(t) \geq 0, \forall t \in [0, \Delta]$, Then in general, the least favorite error sequence is continues polar alternative $+-+-+ \dots$ one, except rectangular $h(t)$. Any no error 0's between the adjacent errors, or no polar alternative will make $\sum_n e_n h(t-nT)$ with larger energy. Thus

$$\text{Min } \varepsilon_s = -\frac{1}{s} \sum_{l=1}^{\text{Min}(K,s)-1} (-1)^{l-1} (s-l) h_l^0, \quad (s \geq 2, K \geq 2), \quad (\text{A-1})$$

$$\text{Max } \varepsilon_s = \frac{1}{s} \sum_{l=1}^{\text{Min}(K,s)-1} (s-l) h_l^0, \quad (s \geq 2, K \geq 2), \quad (\text{A-2})$$

$$a(K, s) = \frac{s}{s-1} \text{Min } \varepsilon_s = -\frac{1}{s-1} \sum_{l=1}^{\text{Min}(K,s)-1} (-1)^{l-1} (s-l) h_l^0, \quad (\text{A-3})$$

B B) The distribution of σ_l

The definition of σ_l is

$$\sigma_l \triangleq \frac{1}{s} \sum_n e_n e_{n-l}, \quad (l=1, 2, \dots, K-1), \quad \sum_n e_n^2 = s, \quad (\text{B-1})$$

σ_l can also be represented as

$$\sigma_l \triangleq \frac{1}{s} \sum_{i=0}^{n_l} x_i, \quad (l=1, 2, \dots, K-1); n_l = 1, 2, \dots, s-1; s \geq 2; K \geq 2), \quad (\text{B-2})$$

Where: binary $x_i \in \{-1, +1\}$, $P_r(x_i = 1) = P_r(x_i = -1) = 0.5$;

n_l is the number of gap length equal to $l-1$ between adjacent errors ;

n_l, x_i are independent random variables;

σ_l is compound binomial distributed with conditional characteristic function as

$$\theta_{\sigma_l}(\omega / n_l) \triangleq E\{e^{j\omega\sigma_l} / n_l\} = \text{Cos}^n(\omega / s), \quad (\text{B-3})$$

Since in a node error event, following the first error the second error can occur equal likely at the past $K-1$ positions. That means

$$P_r(n_l) = (K-1)^{-n_l} \left(\frac{s-1}{n_l} \right) \left(1 - \frac{1}{K-1} \right)^{s-n_l-1}, \quad (\text{B-4})$$

Then the characteristic function of σ_l is

$$\begin{aligned} \theta_{\sigma_l}(\omega) &= \sum_{n_l=0}^{s-1} \theta_{\sigma_l}(\omega / n_l) P_r(n_l) \\ &= \frac{1}{(K-1)^{s-1}} [\text{Cos}(\omega / s) + K-2]^{s-1}, \quad (s \geq 2, K \geq 2), \end{aligned} \quad (\text{B-5})$$

(B-5) is independent on l , therefore σ_l are identical distributed, and so the subscript σ_l of l can be ignored.

Since $\theta_{\sigma}(\omega)$ is an even function of ω , σ only exists even number of moments as

$$\begin{aligned} E\{\sigma^{2k}\} &= (-1)^k \frac{d^{2k}}{d\omega^{2k}} \theta_{\sigma}(\omega) \Big|_{\omega=0} \\ &= \frac{1}{(K-1)^{s-1}} \sum_{l=0}^{s-1} \sum_{i=0}^l \binom{s-1}{l} \binom{l}{i} \frac{(K-2)^{s-i-1} (l-2i)^{2k}}{2^l s^{2k}}, \quad (\text{B-6}) \\ &\quad (k=1, 2, \dots; s \geq 2; K \geq 2). \end{aligned}$$

Especially

$$E\{\sigma\} = 0, \quad E\{\sigma^2\} = \frac{s-1}{s^2(K-1)}, \quad (\text{B-7})$$

$$E\{\sigma^4\} = \frac{s-1}{s^2(K-1)^2} [(K-1) + 3(s-2)], \quad (\text{B-8})$$

When $s \gg 1$,

$$\lim_{s \rightarrow \infty} \text{Log} \theta_{\sigma}(\omega) = \frac{s-1}{2s^2(K-1)} \omega^2, \quad (\text{B-9})$$

It is just the logarithm characteristic function of a Gaussian random variable with mean 0 and variance $(s-1)/s^2(K-1)$.

Although n_l ($l=1, 2, \dots$) are not independent, x_i ($i=0, 1, 2, \dots$) are independent, σ_l ($l=1, 2, \dots$) are uncorrelated.

$$E\{\sigma_l \sigma_m\} = \frac{s-1}{s^2(K-1)} \delta_{m,l}, \quad (m, l=1, 2, \dots, K-1), \quad (\text{B-10})$$

C The capacity of channel of non-strictly limited

bandwidth with high SNR

OVTDM employs physical realizable rather than strictly band-limited $h(t) \leftrightarrow H(f)$. In engineer people are interested in $H(f)$ with a long tail. What is the channel capacity C for such $H(f)$ in AWGN channel with high SNR?

When received signal power $P_s \gg 1$, we cut off $H(f)$ by a fictive rectangular filter $G_{B_f}(f)$ with bandwidth $B_f > B_e$, where B_e is the equivalent noise bandwidth or other engineer sense bandwidth, and divide $G_{B_f}(f)$ into many rectangular sub-channels $G_{B_i}(f)$ with band-width $B_i \rightarrow 0$, they are all Nyquist ones such that

$$\sum_i G_{B_i}(f) = G_{B_f}(f) = \begin{cases} 0, & |f| > B_f \\ 1, & |f| \leq B_f \end{cases}$$

Basic assumptions:

$$P_s^i \triangleq P_s \int_{-\infty}^{\infty} |H(f) G_{B_i}(f)|^2 df \gg N_0 B_i, \quad \forall i; \quad \sum_i B_i = B_f > B_e.$$

Where: P_s^i is the signal power of the i th sub-channel; $\sum_i P_s^i \cong P_s$, $\sum_i C_i = C$.

If both B_f, P_s are sufficiently large, $H(f)G_{B_f}(f)$ close to $H(f)$ well under MMSE criterion. The i th sub-channel's capacity C_i is

$$C_i = B_i \text{Log}_2(1 + P_s^i / N_0 B_i), \quad (\text{B-11})$$

$$\text{Thus } 2^{C_i} = (1 + P_s^i / N_0 B_i)^{B_i} \xrightarrow{B_i \rightarrow 0} e^{P_s^i / N_0}.$$

$$2^C = \prod_i 2^{C_i} = 2^{\sum_i C_i} \xrightarrow{B_i \rightarrow 0} e^{\sum_i P_s^i / N_0} = e^{P_s / N_0}. \quad (\text{B-12})$$

Finally we have

$$C = \frac{P_s}{N_0} \text{Log}_2 e, \quad (\text{B-13})$$

When SNR is high, (B-13) is identical to [15], but [15] is only for Nyquist $H(f)$ with infinite bandwidth. Any way when $H(f)$ strictly bandwidth extends to very wide, no matter $H(f)$ is engineer sense band-limited or not, Channel capacity will be always linearly to SNR.

References

- [1] Viterbi A J, Omura J K. Principle of Digital Communication & Coding. McGraw-Hill, 1979.
- [2] Li Daoben. The statistical Theory of Signal Detection & Estimation. China Academic Press, 2-nd edition, Sept. 2005.
- [3] Li Daoben. An Overlapped Time Division Multiplexing Transmission Scheme. PCT application number: PCT/CN2006/001585.

- [4] Li Daoben. An Overlapped Frequency Division Multiplexing Transmission Scheme. PCT application number: PCT/CN2006/002012.
- [5] Li Daoben. An Overlapped Blocked Time, Frequency and Space Division Multiple Accesses Transmission Scheme. PCT application number: PCT/CN2006/000947.
- [6] Li Daoben. A Coded Division Multiple Accesses Transmission Scheme. PCT application number: PCT/CN2007/000536.
- [7] Forney G.D. and Jr. Ungerboeck G. Modulation and coding for linear Gaussian channels. IEEE Trans. on Information Theory. Vol. 44(6). 1998: 2348.
- [8] Simon, M. K., and Alouini, M. S. Digital Communication over Fading Channels – A Unified Approach to Performance Analysis, 1st ed. Wiley, 2000.
- [9] X. Ma and Ping Li, Coded modulation using superimposed binary codes. IEEE Trans. Information Theory, Vol.50, No.10, Oct. 2004.
- [10] U Erez, R Zamir. Achieving $1/2 \log(1 + \text{SNR})$ on the AWGN channel with lattice. IEEE Trans. on Information Theory, Vol.50, No.12, Dec. 2004.
- [11] John G. Proakis. Digital Communications 4th edition. New York: McGraw Hill, 2001.
- [12] Li Daoben. A High Spectral Efficient Waveform Encoding OVTDM Theory and Applications. China Academic Press, Nov. 2013.
- [13] Cho, K., and Yoon, D., "On the general BER expression of one- and two-dimensional amplitude modulations", IEEE Trans. Commun. Vol. 50, Number 7, pp. 1074-1080, 2002.
- [14] Robert G. Gallager, Principles of Digital Communication. 2010.
- [15] Thomas M. Cover, Joy A. Thomas "Elements of Information Theory" John Wiley & Sons, Inc. 2006.

Biography



Li Daoben is a professor of School of Information and Communication of Beijing University of Posts And telecom. Beijing China. He received his M.S. and Ph.D. degrees from the Beijing University of Posts & Telecom, China, in 1961 and 1964 respectively. He is now a full professor of

BUPT. His research interests are in information and coding theory as well as mobile communications. He is the inventor of LAS-CDMA, OVCDM and OVCDMA. In 1999, he founded LinkAir Communications, Inc., located in Santa Clara, USA, from 1999-2005 had been the Chief Scientist and Vice President of LinkAir Communications, Inc.

From 1981 to 1982, he was a Visiting Scholar at the University of Hawaii. From 1982 to 1984, he was a Visiting Scholar at the University of California at Los Angeles. From 1989 to 1991, he was a Visiting Professor and Ph.D. students' supervisor at Northeastern University, Boston, Massachusetts, USA.

Prof. Li has published over 200 papers and five books. He holds 119 PCT international patents on LAS-CDMA, OVXDM, OVCDM, OVCDMA and others. He was ever in charge the following National Key projects: 3 NNSF, 2 863; 2 National major on mobile Comm. etc. in China.

E-mail: lidaoben2014@163.com;

Tel. 86-13901256221, 86-10-82295359.

Address: 6-2-3A02, # 15, Deshengmen Xidajie Street, Beijing, China.100082.