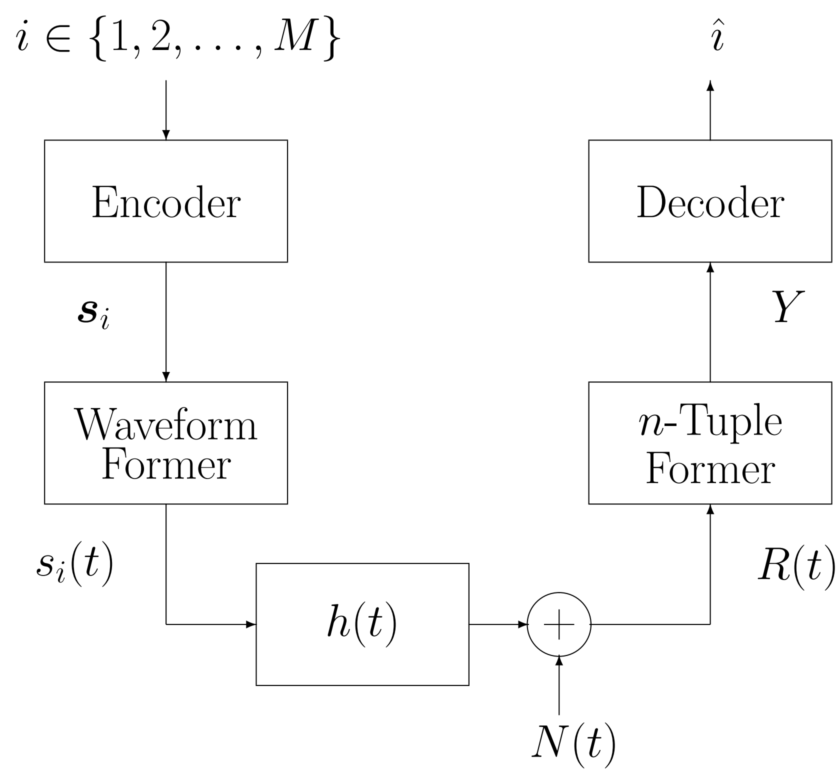
Software-Defined Radio: A Hands-On Course

Basic Digital Communication Link

Last revised: Sept 27, 2016

GOAL

The goal of today's assignment is to implement a software-defined-radio version of a basic point-to-point communication system for bandlimited white Gaussian channels. (See *Principles of Digital Communications*.)



We review the big picture, namely

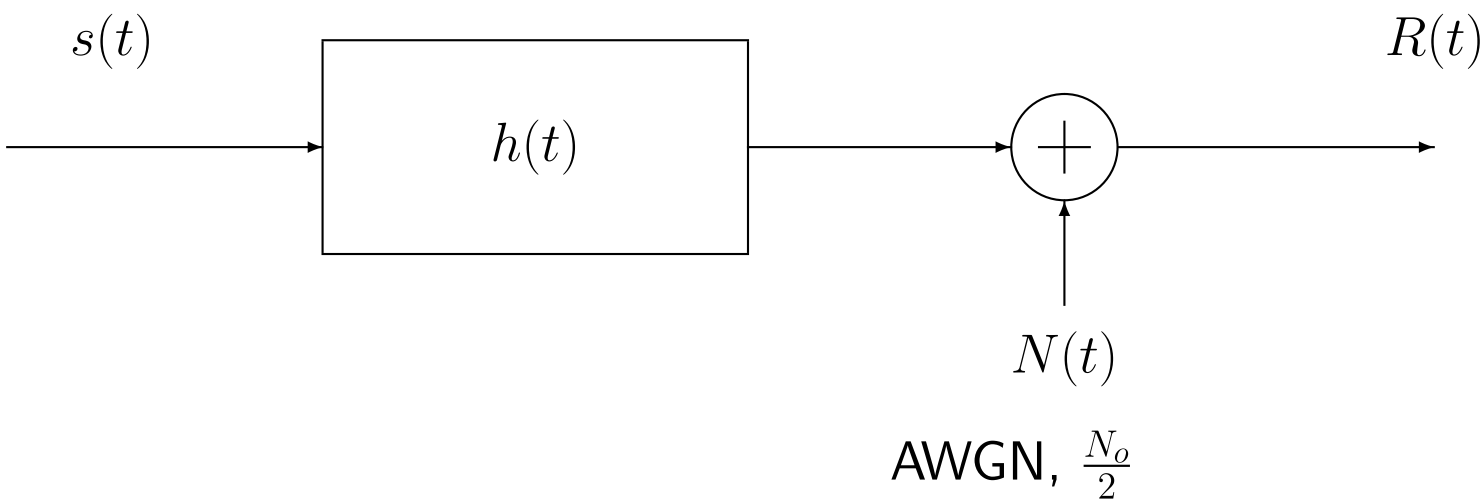
- the notion of orthonormal expansion for the sender's implementation
- the Nyquist criterion, useful for both the sender and the receiver
- the notions of projection and sufficient statistic for the receiver.

We review the sampling theorem, we recall that it belongs to the family of orthogonal expansions like the Fourier transform and all of its variants.

We discuss questions that have come up in the past.

THE CHANNEL OF INTEREST

We focus on the following waveform channel:



$$h_{\mathcal{F}}(f) = \begin{cases} 1, & |f| \leq B \\ 0, & \text{otherwise.} \end{cases}$$

REVIEW OF THE FUNDAMENTAL CONCEPTS

Digital communication is characterized by a finite (possibly very large) set of messages. Let M be the cardinality of the message set.

Let $\{s_1(t), s_2(t), \ldots, s_M(t)\}$ be the associated finite-energy signals. These signals span an inner product space $\mathcal{W}$ in $\mathcal{L}_2$.

Let $\psi_1(t), \psi_2(t) \ldots, \psi_n(t)$, be an orthonormal basis for $\mathcal{W}$.

For each i, there exists a unique $n$tuple $\boldsymbol{s}_i = (s_{i,1}, s_{i,2}, \ldots, s_{i,n})$ such that

$$s_i(t) = \sum_{k=1}^{n} s_{i,k}\psi_k(t).$$

This explains the structure of the sender.

The approach followed in the previous page underlines the generality of the expansion

$$s_i(t) = \sum_{k=1}^{n} s_{i,k}\psi_k(t).$$

In practice, the signals $\{s_1(t), s_2(t), \ldots, s_M(t)\}$ are not the starting point but the result from choosing the $codewords$ $\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{s}_M$ and an orthonormal basis $\psi_1(t), \psi_2(t) \ldots, \psi_n(t)$ (more on the orthonormal basis later).

The codeword's components are taken from a finite constellation of signal points, e.g. from Quadrature Amplitude Modulation (QAM).

We say that there is no coding when, over the random experiment of selecting a message and sending the corresponding codeword, each codeword component appears as being selected independently from the other components.

On the other hand, coding introduces dependency among components.

Picking the components of $\boldsymbol{s}_i$ from a regular and small-dimensional constellation like QAM simplifies the receiver. From information theory we know that the limitations we incur in doing so are negligible for a well-designed system.

Without loss of optimality, the receiver front-end may project the received signal $r(t) = s(t) + N(t)$ into $\mathcal{W}$. Let $y(t)$ be the resulting signal. We may write

$$y(t) = \sum_{k=1}^{n} y_k \psi_k(t),$$

where

$$y_k = \langle r(t), \psi_k(t) \rangle.$$

Let $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$.

The n-tuple $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ is a sufficient statistic. (We use capitals for random variables.)

A receiver that minimizes the error probability may decide that the transmitted signal is (one of) the signals that minimizes the distance $\| \boldsymbol{y} - \boldsymbol{s}_i \|$.

For several reasons (including implementation convenience), it is particularly appealing if, for some pulse $\psi(t)$ and some epoch T, we can choose

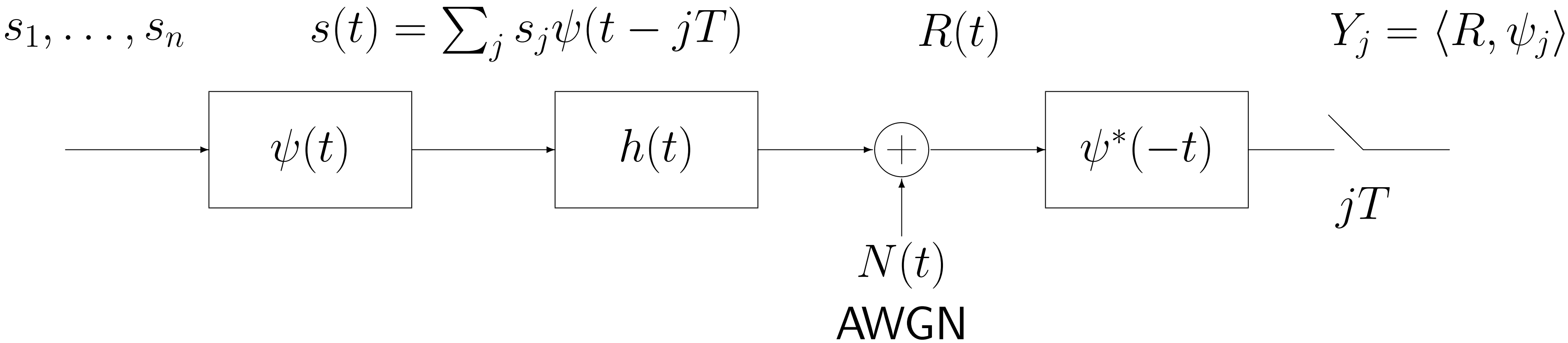
$$\psi_i(t) = \psi(t - iT).$$

Then we may obtain all the $y_k = \langle r(t), \psi_k(t) \rangle$, $k = 1, \ldots, n$ with a single filter, namely the matched filter. If the matched-filter impulse response is chosen to be $\psi^*(-t)$, then $y_k = \langle r(t), \psi_k(t) \rangle$ is obtained by sampling the filter output at $t = kT$.

Nyquist criterion helps us design such a pulse $\psi(t)$ while controlling the power spectral density of the resulting communication signal.

A pulse $\psi(t)$ such that $\{\psi(t - iT) : i \in \mathbb{Z}\}$ is an orthonormal family is referred to as a Nyquist pulse.

HEREAFTER WE ASSUME THE FOLLOWING SETUP

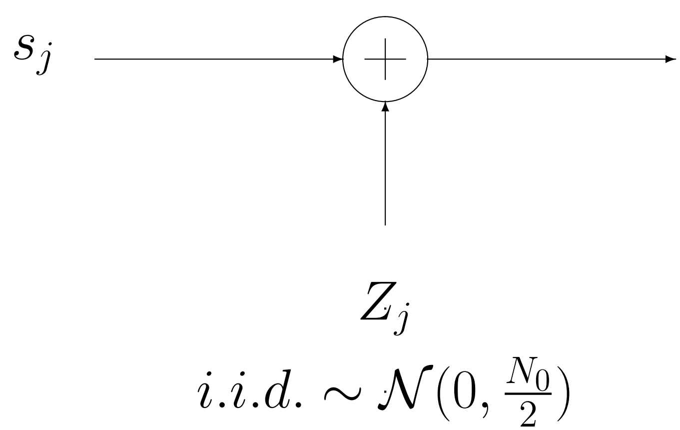


We may assume *baseband communication* since *passband communication* can be ensured by means of an extra "layer" namely the up-conversion before and the down-conversion after the waveform channel.

EQUIVALENT DISCRETE-TIME CHANNEL

We assume that the channel frequency response is 1 over the band occupied by the signal. The time-domain condition is $(\psi \star h)(t) = \psi(t)$, where $\star$ denotes convolution. This means that the channel is "transparent" to the pulse.

Then the input/output behavior of the above block diagram is identical to that of the following *discrete-time AWGN channel* model.



$$Y_j = s_j + Z_j$$

DERIVATION OF THE NYQUIST CRITERION

We are looking for functions $\psi(t)$ with the property

$$\int_{-\infty}^{\infty} \psi(t-nT)\psi^*(t)dt = \delta_n. \tag{1}$$

This means that $\{\psi(t), \psi(t-T), \ldots, \psi(t-nT)\}$ is an orthonormal set. Hence it is an orthonormal basis for the space spanned by $\{s_1(t), s_2(t), \ldots, s_M(t)\}$.

If we were completely free to choose $\psi(t)$, it would be easy to design it in the time domain. For instance we could chose a rectangle.

But we want $\psi(t)$ to have a certain characteristic in the frequency domain, e.g. limited bandwidth. (Recall that the power spectral density of the transmitted signal is proportional to $|\psi_{\mathcal{F}}(f)|^2$.) So we are interested in the frequency-domain equivalent of (1).

Define the periodic function

$$g(f) = \sum_{k \in \mathbb{Z}} \psi_{\mathcal{F}}\left(f - \frac{k}{T}\right) \psi^*_{\mathcal{F}}\left(f - \frac{k}{T}\right) = \sum_{k \in \mathbb{Z}} \left| \psi_{\mathcal{F}}\left(f - \frac{k}{T}\right) \right|^2 .$$

We can now rewrite (1) as follows:

$$\delta_n = \int_{-\infty}^{\infty} \psi_{\mathcal{F}}(f) \psi^*_{\mathcal{F}}(f) e^{-j2\pi nTf} df = \int_{-\frac{1}{2T}}^{\frac{1}{2T}} g(f) e^{-j2\pi nTf} df.$$

The last integral is $\frac{1}{T}$ times the Fourier series coefficient A_n of the periodic function $g(f)$. It says that $A_0 = T$ and $A_n = 0$ for $n \neq 0$. Hence $g(f)$ is the constant function that takes the value T everywhere.

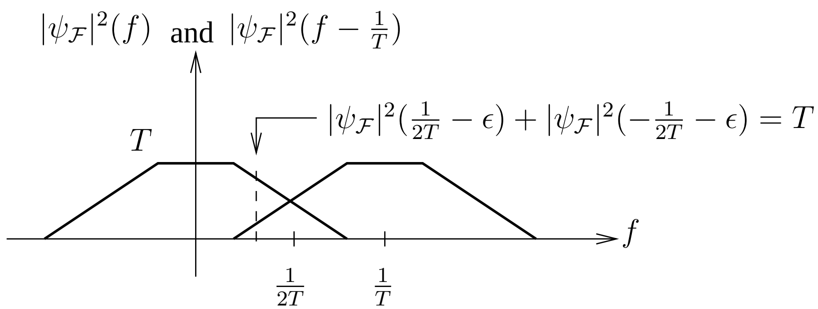
NYQUIST CRITERION

Theorem (Nyquist). Let $\psi(t)$ be an $\mathcal{L}_2$ function. The set $\{\psi(t-nT)\}_{j=-\infty}^{\infty}$ consists of orthonormal functions if and only if

$$\text{l.i.m.} \sum_{k=-\infty}^{\infty} \left|\psi_{\mathcal{F}}(f-\frac{k}{T})\right|^2 = T \quad \text{for } f \in \mathbb{R} \tag{2}$$

□

If $|\psi_{\mathcal{F}}(f)|$ is an even function (always the case if $\psi(t)$ is real-valued) and has support within an interval of width $2/T$, checking Nyquist condition is particularly easy:



A POPULAR PRACTICAL CHOICE

For any *roll-off factor* $\beta \in (0,1)$,

$$|\psi_{\mathcal{F}}|^2(f) = \begin{cases} T, & |f| \leq \frac{1-\beta}{2T} \\ \frac{T}{2}\left(1+\cos\left[\frac{\pi T}{\beta}\left(|f| - \frac{1-\beta}{2T}\right)\right]\right), & \frac{1-\beta}{2T} < |f| < \frac{1+\beta}{2T} \\ 0, & \text{otherwise} \end{cases}$$

fulfills Nyquist criterion.

By taking the inverse Fourier transform of $\psi_{\mathcal{F}}(f)$, we obtain the pulse $\psi(t)$, called *root-raised-cosine pulse*. The result (after some manipulations) is

$$\psi(t) = \frac{4\beta}{\pi\sqrt{T}} \frac{\cos\left((1+\beta)\pi\frac{t}{T}\right) + \frac{(1-\beta)\pi}{4\beta}\operatorname{sinc}\left((1-\beta)\frac{t}{T}\right)}{1-\left(4\beta\frac{t}{T}\right)^2},$$

where

$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}.$$

SAMPLING THEOREM REVISITED

The sampling theorem plays a key role in software-defined radio. Let us review it.

Assume that $s(t)$ is a continuous $\mathcal{L}_2$ function, and $s_{\mathcal{F}}(f) = 0,\ f \notin [-B, B]$. Then we can write

$$s(t) = \sum_i s(iT) \operatorname{sinc}\left(\frac{t - iT}{T}\right)$$

for any $T < \frac{1}{2B}$.

Recall that the Fourier transform of $\operatorname{sinc}(\frac{t}{T})$ is the rectangular function

$$\operatorname{sinc}_{\mathcal{F}}(f) = \begin{cases} T, & |f| \le B \\ 0, & \text{otherwise.} \end{cases}$$

It is straightforward to see that the above rectangular function fulfills the Nyquist criterion. Hence

$$\psi(t) = \frac{1}{\sqrt{T}} \operatorname{sinc}\left(\frac{t}{T}\right)$$

has the property

$$\langle \psi(t), \psi(t - kT) \rangle = \delta_k.$$

We call such a pulse $\psi(t)$ a Nyquist pulse.

We have discovered that a simple rescaling turns the sampling theorem into an orthonormal expansion, namely

$$\begin{aligned} s(t) &= \sum_i s(iT) \operatorname{sinc}\left(\frac{t - iT}{T}\right) && \textit{(sampling thm)} \\ &= \sum_i s_i \psi(t - iT) && \textit{(orthonormal expansion)} \end{aligned}$$

where

$$s_i = s(iT)\sqrt{T}.$$

To remember the relationship $s_i = s(iT)\sqrt{T}$ we may compare terms in

$$\sum |s_i|^2 = \int |s(t)|^2 dt \approx \sum |s(iT)|^2 T$$

where the equality holds since an orthonormal expansions relates a signal to the corresponding n-tuple of coefficients via a unitary transformation (i.e. a norm-preserving transformation). The expression on the right is Riemann's approximation to the integral. As a byproduct we observe that Riemann's "approximation" is actually exact when we use samples that fulfill the sampling theorem.

An even easier way to remember that $s_i = s(iT)\sqrt{T}$ is to notice that for a fixed signal $s(t)$, the result of the sum $\sum |s(iT)|^2$ grows essentially linearly with the number of sample points, i.e., with the sampling rate $1/T$. To have a chance of being a coefficient of the orthonormal expansion, the term $s(iT)$ has to be multiplied by $\sqrt{T}$.

To summarize: If we sample a signal, we have to multiply the samples by $\sqrt{T}$ if we want the norm of the resulting discrete-time signal to be equal that of the original continuous-time signal.

We recommend that you normalize the discrete-time version of $\psi(t)$ that you use to implement the sender and the receiver. (See below why). With MATLAB we can normalize a signal just by dividing it by its norm.

Something Useful for the Assignment

You may find the following to be useful (not done in PDC).

How to efficiently create the samples of

$$s(t) = \sum_k s_k \psi(t - kT)$$

from the sequence of coefficients s_k and from the samples of $\psi(t)$?

Assumption and Notation:

$$\begin{aligned} s[k] &:= s(kT_s) \\ \psi[k] &:= \psi(kT_s) \\ T &= NT_s. \end{aligned}$$

Now

$$
\begin{aligned}
s[n] := s(nT_s) &= \sum_k s_k \psi(nT_s - kT) \\
&= \sum_k s_k \psi((n - kN)T_s) \\
&= \sum_k s_k \psi[n - kN] && (3) \\
&= \sum_k \hat{s}_k \psi[n - k] && (4)
\end{aligned}
$$

where we have defined

$$
\hat{s}_k = \begin{cases} s_{\frac{k}{N}} & \text{if } k \text{ is an integer multiple of } N \\ 0 & \text{otherwise.} \end{cases}
$$

In words, $\hat{s}_k$ is obtained from s_k by inserting $N-1$ zeros between consecutive samples.

In going from (3) to (4) we are using the fact that the discrete-time signal $\cdots + s_1\psi[n] + s_2\psi[n-N] + \cdots$ is the same as $\cdots + s_1\psi[n] + 0\psi[n-1] + \cdots + 0\psi[n-(N-1)] + s_2\psi[n-N] + 0\psi[n-N+1] + \cdots$

What we have gained is that (4) is the convolution of the upsampled symbols sequence $\hat{s}_k$ and the sampled pulse $\psi[k]$. `MATLAB` provides functions to upsample and to convolve.

About generating the discrete-time root raised-cosine pulse $\psi[n]$ in MATLAB

The continuous-time root raised-cosine pulse $\psi(t)$ is completely specified by two parameters: the symbol time T and the roll-off factor β.

For the continuous-time pulse $\psi[n]$, instead of T we need the number-of-samples per symbol-interval T. (Once sampled, the notion of time no longer makes sense. All you have, is a sequence of numbers.) Let us call this `SPS` (for samples per symbol).

You can generate a truncated version of $\psi[n]$ with the MATLAB function `rccosdeisgn`. The length of the truncated pulse is specified in terms of the number of samples per symbols, denoted by `SPAN`.

To summarize, `rcosdesign` requires the mandatory parameters `BETA`, `SPAN`, and `SPS`. See `help rccosdeisgn` for additional parameters.

Frequently Asked Questions

Q1: If the sampling theorem is an orthogonal expansion, then I should be able to obtain the samples from a projection. I don't see the projection.

A1: Consider the sampling theorem written as an orthonormal expansion. The coefficients are then computed according to $s_i = \langle s(t), \psi_i(t) \rangle$, where $\psi_i(t) = \psi(t - iT)$ and $\psi(t)$ is the normalized sinc. A matched filter implementation of this consists of a lowpass filter with frequency response

$$\psi_{\mathcal{F}}(f) = \begin{cases} \sqrt{T}, & |f| \le \frac{1}{2T} \\ 0, & \text{otherwise} \end{cases}$$

and output sampled at time $t = iT$. But the above filter does nothing to the signal (which vanishes outside the frequency interval $[-\frac{1}{2T}, \frac{1}{2T}]$) except for scaling it by $\sqrt{T}$.

Hence the sampled output $s_i = \langle s(t), \psi_i(t) \rangle$ equals $\sqrt{T} s(iT)$, as expected.

Q2: Why do we care whether or not we normalize the sampled version of $\psi(t)$?

A2: Several reasons:

(1): If you normalize, in absence of noise the ith matched filter output will be exactly s_i. Verifying that it is indeed the case is a "sanity check" you should do. It also allows you to use the slicer (also called demodulator) defined for your constellation without having to figure out how to rescale the matched filter output.

(2): To implement the AWGN channel you need to figure out the correct variance of the Gaussian noise that the channel adds to each sample. When we specify the SNR, it is the SNR at the output of the matched filter the one that matters. If $\boldsymbol{N} = (N_1, N_2, \ldots, N_n)$ is a random vector with i.i.d. components and $\boldsymbol{\psi} = (\psi_1, \psi_2, \ldots, \psi_n)$ has unit norm, then $\langle \boldsymbol{N}, \boldsymbol{\psi} \rangle$ has the same variance as the components of $\boldsymbol{N}$.

In summary, normalization can save you a headache in designing and debugging your implementation.

Q3: How can we test the end-to-end system?

A3: One test is to check that the error probability is what it should be. For 4-QAM, the symbol (as opposed to bit) error probability is $2Q - Q^2$ where $Q = Q(\frac{d}{2\sigma})$. Here $\frac{d}{2}$ is half the minimum distance between 4-QAM points at the matched filter output and σ^2 is the noise variance at the same point.

It is useful to know that the average energy of m-PAM with symbol at $\{\pm 1, \pm 3, \ldots, \pm(m-1)\}$ is $\frac{m^2-1}{3}$.

If $S = X + iY$ where X and Y are independent m-PAM points and $i = \sqrt{-1}$, then S belongs to M-QAM with $M = m^2$. Due to the independence of X and$Y$, the average energy of S is twice that of X, i.e., it is twice that of m-PAM, namely $\frac{2(M-1)}{3}$.