Modern Digital Communications: A Hands-On Approach

Rayleigh Flat-Fading

Dr. Nicolae Chiurtu

- Course material of Prof. Bixio Rimoldi -

Last revision: Dec. 19, 2019

Motivation

We have seen that OFDM 'forges' parallel channels of the kind

$$\boldsymbol{Y} = D\boldsymbol{A} + \boldsymbol{Z}$$

where we assume that

- D is the $N \times N$ diagonal matrix of channel coefficients,
- $\boldsymbol{A} \in \mathbb{C}^N$ is the vector of data symbols, and
- $\boldsymbol{Z} \sim \mathcal{CN}(0, N_0 I_N)$.

The diagonal of D is a realization of a Gaussian random vector.

If we assume that the channel coefficients stay constant over several OFDM blocks, then the input/output relationship of the nth parallel channel is

$$Y[n] = \lambda A[n] + Z[n],$$

where λ is a sample from $\sim \mathcal{CN}(0, \sigma^2)$.

What follows can be applied to situations other than OFDM, and the channel strength will be denoted by h instead of λ.

Since we have learned how to estimate the channel coefficients, we assume that h is known to the receiver (but unknown to the transmitter).

Our first objective is to derive the error probability $P_e(h)$, and then the average error probability $E[P_e(H)]$ in the absence of coding.

Without coding (the case assumed here), and since the channel is memoryless, the distribution of $A[n]$ given $Y[1], \ldots, Y[N]$ depends only on $Y[n]$.

We can simplify notation by dropping the index n, i.e., write

$$Y = hA + Z.$$

We are in the PDC setting, where A is the symbol, hA is transmitted across an AWGN channel, and Y is the channel output.

A sufficient statistic is obtained by projecting the channel output onto the signal space, which in this case is the space spanned by $h \in \mathbb{C}$. This gives us

$$W = \left\langle Y, \frac{h}{|h|} \right\rangle = |h|A + \frac{h^*}{|h|}Z.$$

$\frac{h^*}{|h|}Z$ has the same distribution as Z. (Z is circularly symmetric.) Hence, by letting $r = |h|$ and with a slight abuse of notation, we write

$$W = rA + Z.$$

The error probability depends on the signal constellation.

To make our point, we can choose the simplest constellation, namely $\pm\sqrt{\mathcal{E}}$ (BPSK, BPAM, antipodal signaling).

Conditioned on $R = r$, the error probability is

$$P_e(r) = Q\left(\frac{d(r)/2}{\sqrt{N_0/2}}\right)$$

where $d(r)/2$ is half the distance between $\pm r\sqrt{\mathcal{E}}$, i.e., $d(r)/2 = r\sqrt{\mathcal{E}}$, and $N_0/2$ is the variance of the real component of Z.

Thus,

$$P_e(r) = Q\left(\sqrt{\frac{r^2\mathcal{E}}{N_0/2}}\right) = Q\left(\sqrt{\frac{v\mathcal{E}}{N_0/2}}\right),$$

where $v = r^2$.

Notice that $v\mathcal{E}$ is the signal strength at the decoder, so that $\gamma = \frac{v\mathcal{E}}{N_0/2}$ is the signal-energy to noise-power ratio. We are using N_0 rather than $\frac{N_0}{2}$ since the signal constellation in in one dimension, hence what matters is the noise variance along that dimension.

Hence, with a slight abuse of notation, we can rewrite

$$P_e(\gamma) = Q(\sqrt{\gamma}), \tag{1}$$

which is the error probability of BPSK when the signal-energy to noise-power ratio is γ, but remember that here γ varies every time that we use the channel.

Next we compute the average error probability, averaged over all fading states V. As computed in Appendix A,

$$f_V(v) = \begin{cases} \frac{1}{\sigma^2} \exp\left(-\frac{v}{\sigma^2}\right), & \text{if } v \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Hence

$$\bar{P}_e = \int_0^\infty Q\left(\sqrt{\frac{v\mathcal{E}}{N_0/2}}\right) \frac{1}{\sigma^2} \exp\left(-\frac{v}{\sigma^2}\right) dv.$$

Perhaps surprisingly, the above integral can be solved. To do so, we make the change of variable $\alpha = \sqrt{\frac{v\mathcal{E}}{N_0/2}}$ to obtain

$$\bar{P}_e = \int_0^\infty Q(\alpha) \frac{N_0 \alpha}{\mathcal{E}\sigma^2} \exp\left(-\frac{\alpha^2 N_0/2}{\mathcal{E}\sigma^2}\right) d\alpha.$$

Integration by parts yields

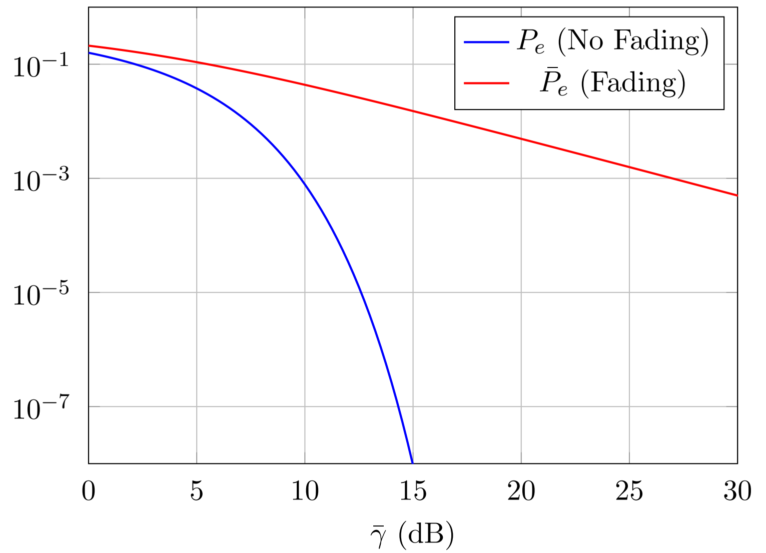
$$\bar{P}_e = \frac{1}{2}\left(1 - \sqrt{\frac{1}{1 + \frac{N_0}{\mathcal{E}\sigma^2}}}\right).$$

If we define the instantaneous signal-energy to noise-power-density $\gamma = \frac{v\mathcal{E}}{N_0/2}$ and compute its average $\bar{\gamma} = \frac{\sigma^2\mathcal{E}}{N_0/2}$, then we can rewrite

$$\bar{P}_e = \frac{1}{2}\left(1 - \sqrt{\frac{\bar{\gamma}}{2+\bar{\gamma}}}\right).$$

The above should be compared to the non-fading situation with average signal-energy to noise-power-density $\bar{\gamma}$. In this case

$$P_e = Q\left(\sqrt{\bar{\gamma}}\right).$$



We see that the error probability in the presence of Rayleigh[1] fading is quite disappointing.

[1] Called this way because the channel strength $r = |h|$ follows a Rayleigh pdf.

Additional insight, for $\bar{\gamma}$ large, is gained by comparing the two approximations

$$P_e \approx \frac{1}{2}\exp\big(-\frac{\bar{\gamma}}{2}\big),$$

obtained via $Q(x) \approx \frac{1}{2}\exp^{-\frac{x^2}{2}}$ and

$$\bar{P}_e \approx \frac{1}{2\bar{\gamma}}.$$

The latter is obtained as follows:

$$\begin{aligned}
\bar{P}_e &= \frac{1}{2}\left(1-\sqrt{\frac{\bar{\gamma}}{2+\bar{\gamma}}}\right) \\
&= \frac{1}{2}\left(1-\sqrt{1-\frac{2}{2+\bar{\gamma}}}\right) \\
&\approx \frac{1}{2}\left(\frac{1}{2+\bar{\gamma}}\right) \\
&\approx \frac{1}{2\bar{\gamma}},
\end{aligned}$$

where we used the approximation $\sqrt{1-\epsilon} \approx 1-\frac{\epsilon}{2}$, which holds for small ϵ, and the fact that $2+\bar{\gamma} \approx \bar{\gamma}$ when $\bar{\gamma}$ is large.

We see that the error probability associated to a non-fading channel is drastically different from the error probability of a fading channel, even when the average signal-to-noise ratio is the same: in the former case P_e decreases exponentially as a function of $\bar{\gamma}$, whereas in the latter case $\bar{P}_e$ is inversely proportional to $\bar{\gamma}$.

In Appendix B we argue that the disappointing behavior of $\bar{P}_e$ is essentially due to the probability that Γ is less than 1.

Diversity

Diversity mitigates the unfavorable effect of Rayleigh fading.

The main idea is to create the conditions to make many, say L, independent observations of the transmitted symbols. Even if each observation is the output of a Rayleigh faded channel, the probability that all L channels are in a deep fade at the same instant becomes small when L is sufficiently large.

There are many ways to create diversity:

- time diversity exploits the time-varying nature of the fading channel.
- frequency diversity exploits the frequency-selective nature of multipath channels.
- spatial diversity exploits the position-dependent selectivity of multipath channels.

All forms of diversity have the same effect on the probability of error, and can be studied in a similar manner.

For example, with OFDM we could do frequency diversity by transmitting each symbol over L of the N parallel channels. If the L channels are sufficiently separated in frequency, they experience independent fading. The price for diversity in this case is a reduction of the data rate to $\frac{N}{L}$ symbols per OFDM block.

If $\boldsymbol{Y}$ is the vector that consists of the L channel outputs that carry the same symbol a, then we have

$$\boldsymbol{Y} = \boldsymbol{h}a + \boldsymbol{Z},$$

where $\boldsymbol{h}$, the vector of coefficients, is assumed to be a realization of $H \sim \mathcal{CN}(0, \sigma^2 I_L)$, a is the transmitted symbol which, for comparison with the previously studied case, we assume to be in $\{\pm\sqrt{\mathcal{E}}\}$, and $\boldsymbol{Z} \sim \mathcal{CN}(0, N_0 I_L)$.

As before, we assume that $\boldsymbol{h}$ is known to the receiver.

We are in a typical PDC situation: To minimize the error probability, we project the observed $\boldsymbol{Y}$ onto $\boldsymbol{h}$ to form the sufficient statistic

$$\begin{aligned} Y &= \left\langle \boldsymbol{Y}, \frac{\boldsymbol{h}}{\|\boldsymbol{h}\|} \right\rangle \\ &= \left\langle \boldsymbol{h}a + \boldsymbol{Z}, \frac{\boldsymbol{h}}{\|\boldsymbol{h}\|} \right\rangle = a\|\boldsymbol{h}\| + Z, \end{aligned} \tag{2}$$

where $Z \sim \mathcal{CN}(0, N_0)$.

Let $Y = y$. A maximum likelihood receiver decides

$$\hat{a} = \begin{cases} +\sqrt{\mathcal{E}} & \text{if } y > 0 \\ -\sqrt{\mathcal{E}} & \text{otherwise.} \end{cases}$$

When $a = -\sqrt{\mathcal{E}}$, a decision error occurs if $\Re\{Z\} > \sqrt{\mathcal{E}}\|\boldsymbol{h}\|$. The probability that this occurs is $Q\left(\frac{\sqrt{\mathcal{E}}\|\boldsymbol{h}\|}{\sqrt{N_0/2}}\right) = Q\left(\sqrt{\frac{\mathcal{E}\|\boldsymbol{h}\|^2}{N_0/2}}\right) = Q\big(\sqrt{\gamma}\big)$, with $\gamma = \frac{\mathcal{E}\|\boldsymbol{h}\|^2}{N_0/2}$.

By symmetry, we obtain the same result when $a = +\sqrt{\mathcal{E}}$.

Hence

$$P_e(\gamma) = Q\big(\sqrt{\gamma}\big) \tag{3}$$

is the error probability, conditioned on the event $\Gamma = \gamma$, where $\Gamma = \frac{\mathcal{E}\|\boldsymbol{H}\|^2}{N_0/2}$.

Notice that (1) and (3) are the same function of γ; what changes is the distribution of Γ.

$\Gamma = \frac{\mathcal{E}\|\boldsymbol{H}\|^2}{N_0/2}$ has a chi-square distribution with $2L$ degrees of freedom, namely

$$f_\Gamma(\gamma) = \frac{1}{(L-1)!\left(\frac{\mathcal{E}\sigma^2}{N_0/2}\right)^L}\,\gamma^{L-1}\,e^{-\gamma\left(\frac{N_0/2}{\mathcal{E}\sigma^2}\right)},\ \text{for}\ \gamma \geq 0$$

(see Appendix C).

To compute the average error probability $\bar{P}_e$, we integrate $Q\big(\sqrt{\gamma}\big)$ against $f_\Gamma(\gamma)$ as we did earlier for the special case of $L = 1$.

After repeated integrations by part, we obtain

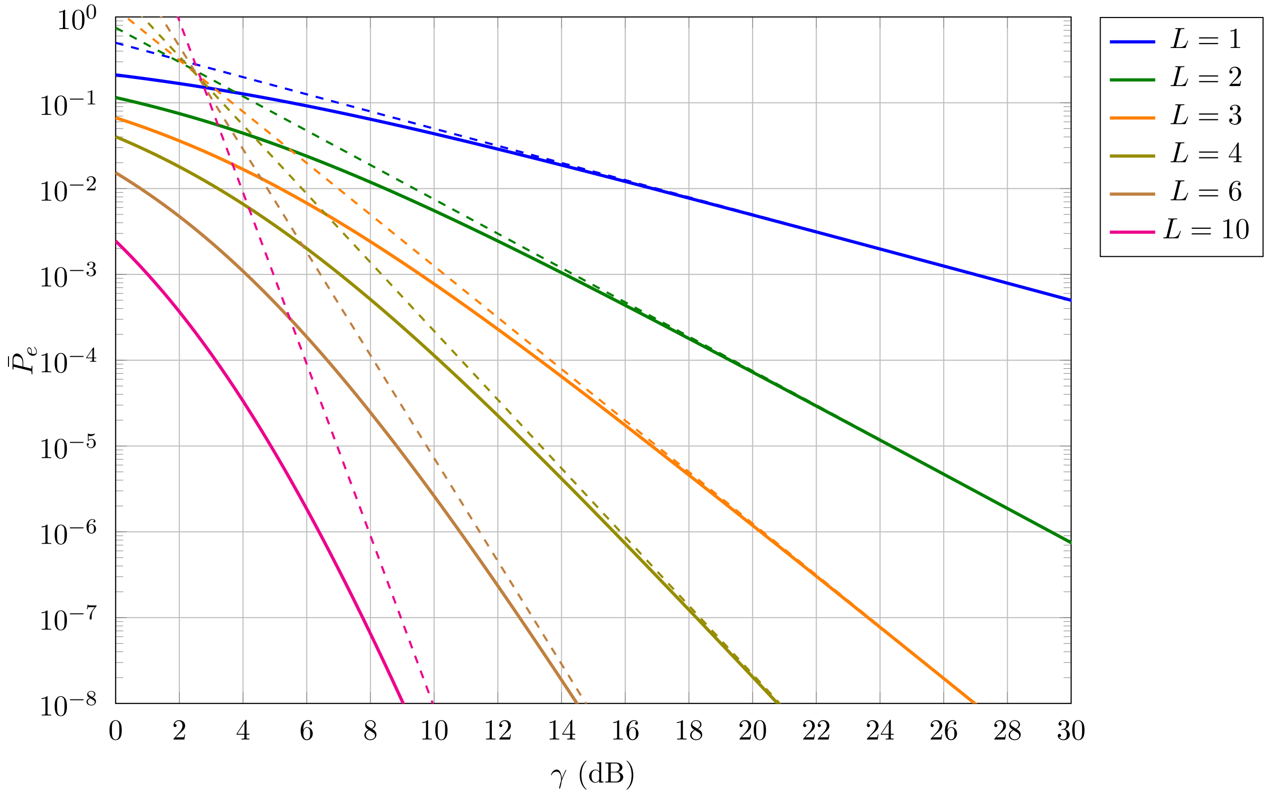
$$\bar{P}_e(\bar{\gamma}) = \left(\frac{1}{2}\left(1 - \sqrt{\frac{\bar{\gamma}}{2+\bar{\gamma}}}\right)\right)^L \sum_{l=0}^{L-1} \binom{L-1+l}{l} \left(\frac{1}{2}\left(1 + \sqrt{\frac{\bar{\gamma}}{2+\bar{\gamma}}}\right)\right)^l.$$

For large $\bar{\gamma}$, the above approximates to

$$\bar{P}_e(\bar{\gamma}) \approx \left(\frac{1}{2\bar{\gamma}}\right)^L \binom{2L-1}{L}. \qquad (*)$$

Observe that for $L = 1$ we get back our earlier approximation.

Observe also that, in essence, the effect of order-L diversity is to raise to the power L the original error probability.

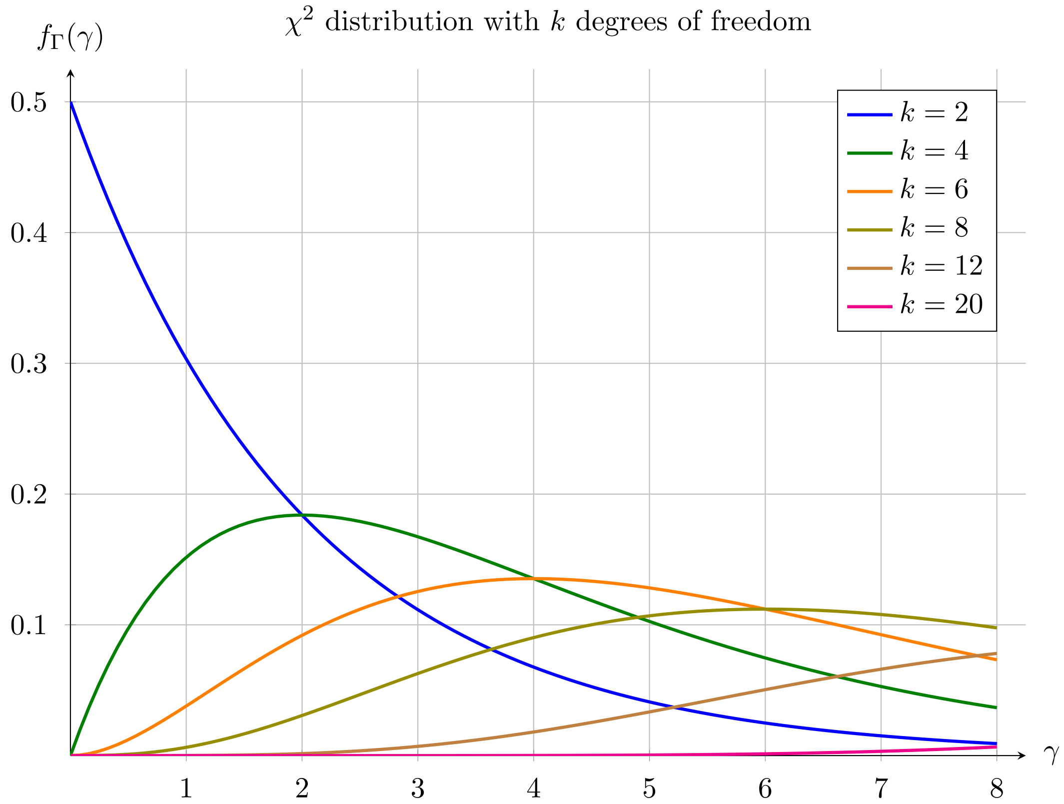


(Dashed lines show the straight-line asymptotics of the curves, as approximated in $(*)$)

The action of combining two or more observables like Y_i and Y_j into $Y = Y_i h_i + Y_j h_j$ as we have done above (see (2)), is commonly referred to as $\mathit{maximal\text{-}ratio\ combining}$.

The name stems from the fact that this way of combining does indeed maximize the signal-energy to noise-variance ratio.

The following plots of $f_\Gamma(\gamma)$ (the χ^2 distribution with k degrees of freedom) help to understand why diversity makes a difference: it drastically changes the probability that Γ is below 1.



Appendix A: the PDF of R and that of V

$H = X + jY$ is zero-mean, circularly symmetric, of variance σ^2.

Equivalently, X and Y are zero-mean, independent Gaussian random variables of variance $\sigma^2/2$ each.

Hence

$$f_H(h) = f_{X,Y}(x,y) = \frac{1}{2\pi\frac{\sigma^2}{2}} \exp\left(-\frac{x^2+y^2}{2\frac{\sigma^2}{2}}\right) = \frac{1}{\pi\sigma^2} \exp\left(-\frac{|h|^2}{\sigma^2}\right).$$

Next, let r, ϕ be the polar coordinates of the Cartesian point x, y.

Recall that

$$f_{R,\Phi}(r,\phi) = |\det J| f_{X,Y}\big(r\cos(\phi), r\sin(\phi)\big),$$

where J is the Jacobian of the coordinate transformation

$$(r,\phi) \to (x,y) = \big(r\cos(\phi), r\sin(\phi)\big).$$

i.e.,

$$\det J = \begin{vmatrix} \cos(\phi) & -r\sin(\phi) \\ \sin(\phi) & r\cos(\phi) \end{vmatrix} = r.$$

Plugging in,

$$f_{R,\Phi}(r,\phi) = \begin{cases} \frac{r}{\pi\sigma^2}\exp\big(-\frac{r^2}{\sigma^2}\big), & \text{if } r \geq 0 \text{ and } 0 \leq \phi < 2\pi, \\ 0, & \text{otherwise.} \end{cases}$$

Integrating over ϕ yields

$$f_R(r) = \begin{cases} \frac{2r}{\sigma^2} \exp\left(-\frac{r^2}{\sigma^2}\right), & \text{if } r \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

which is a Rayleigh distribution.

Finally, we determine the distribution of $V = R^2$, namely the exponential distribution

$$f_V(v) = \begin{cases} \frac{1}{\sigma^2} \exp\left(-\frac{v}{\sigma^2}\right), & \text{if } v \geq 0, \\ 0, & \text{otherwise.} \end{cases} .$$

Appendix B: The Disappointing Behavior of $\bar{P}_e$ Explained

To gain insight about the disappointing behavior of $\bar{P}_e$, let us break down the computation of $\bar{P}_e$ as follows:

$$\begin{aligned}
\bar{P}_e &= \int_0^\infty Q(\sqrt{\gamma}) f_\Gamma(\gamma) d\gamma \\
&= \int_0^1 Q(\sqrt{\gamma}) f_\Gamma(\gamma) d\gamma + \int_1^\infty Q(\sqrt{\gamma}) f_\Gamma(\gamma) d\gamma \\
&\geq Q(1) \int_0^1 f_\Gamma(\gamma) d\gamma \\
&= Q(1) Pr\{\Gamma \leq 1\} \\
&= Q(1)\left(1 - e^{\frac{1}{\bar{\gamma}}}\right) \\
&\approx Q(1)\frac{1}{\bar{\gamma}}
\end{aligned}$$

This shows that the probability of the event $\Gamma \leq 1$ is essentially $\frac{1}{\bar{\gamma}}$ and this alone explains why $\bar{P}_e$ cannot decay faster than $\frac{1}{\bar{\gamma}}$.

Appendix C: The Chi-Square Distribution

The chi-square distribution with $2L$ degrees of freedom, also denoted χ^2_{2L} distribution, is

$$\chi^2_{2L}(\gamma) = \frac{1}{(L-1)!\,2^L}\,\gamma^{L-1}\,e^{-\frac{\gamma}{2}},\ \text{for}\ \gamma \geq 0.$$

It is the distribution of $\sum_{i=1}^{2L} X_i^2$ where $X_1, \ldots, X_{2L}$ are zero-mean iid Gaussian random variables, each of unit variance.

If $Y_i = \alpha X_i$, so that Y_i has variance α^2, then $\sum_{i=1}^{2L} Y_i^2 = \alpha^2 \sum_{i=1}^{2L} X_i^2$, and its density is

$$\frac{1}{\alpha^2}\chi^2_{2L}\big(\frac{\gamma}{\alpha^2}\big) = \frac{1}{(L-1)!\,(2\alpha^2)^L}\,\gamma^{L-1}\,e^{-\frac{\gamma}{2\alpha^2}},\ \text{for}\ \gamma \geq 0.$$

Now suppose that

$$\tilde{\Gamma} = \sum_{i=1}^{2L} Y_i^2,$$

where

$$Y_i^2 = \frac{\mathcal{E}}{N_0/2}\frac{\sigma^2}{2}X_i^2 = \frac{\mathcal{E}\sigma^2}{N_0}X_i^2.$$

If $X_i \sim \mathcal{N}(0,1)$, then $\tilde{\Gamma}$ has the same statistic as $\Gamma = \frac{\mathcal{E}\|H\|^2}{N_0/2}$.

Hence

$$f_\gamma(\gamma) = f_{\tilde{\Gamma}}(\gamma) = \frac{1}{\alpha^2}\chi^2_{2L}(\frac{\gamma}{\alpha^2})$$

with

$$\alpha^2 = \frac{\mathcal{E}\sigma^2}{N_0}.$$

Therefore

$$f_\Gamma(\gamma) = \frac{1}{(L-1)!\left(\frac{\mathcal{E}\sigma^2}{N_0/2}\right)^L}\,\gamma^{L-1}\,e^{-\gamma\left(\frac{N_0/2}{\mathcal{E}\sigma^2}\right)},\ \text{for } \gamma \geq 0.$$