

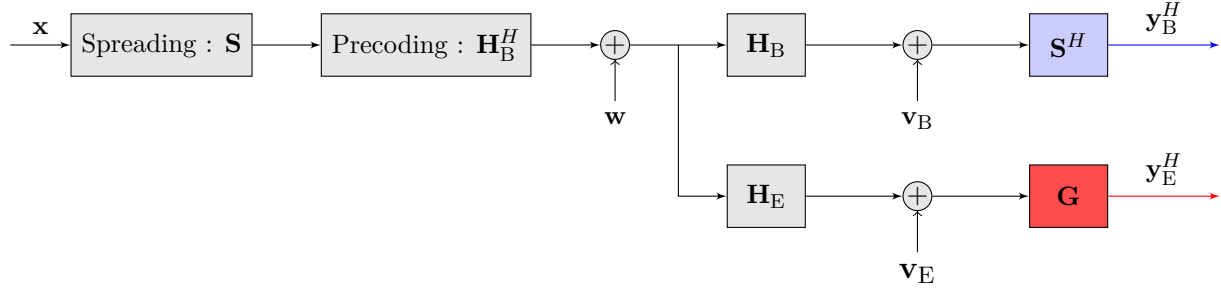
Introducing Correlation among subcarriers

Sidney Golstein

February 25, 2021

1 Short Reminder

We consider that we have Q channel subcarriers, that we spread (via \mathbf{S}) our symbols \mathbf{x} by a factor U called back-off rate (BOR). We send $N = Q/U$ symbols per OFDM block. At the transmitter side, we precode the data by the complex conjugate version of Bob's channel frequency response (\mathbf{H}_B^H). We add an artificial noise signal (\mathbf{w}) which lies in Bob null space, i.e., $\mathbf{S}^H \mathbf{H}_B \mathbf{w} = \mathbf{0}$. At the receiver side, we despread the received sequence at Bob via \mathbf{S}^H , and we decode the received sequence with a particular decoding structure \mathbf{G} at Eve. The communication scheme is presented below:



2 Problem Statement

We want to analyze the effect of the introduction of correlation among channel subcarriers, i.e., the effect of frequency correlation in our system.

In the time-domain, we modelize the channel thanks to L taps:

$$\mathbf{h} = [h_1 \dots h_L]^T \quad (1)$$

The time-domain correlation function is given by:

$$\mathbb{E} [\mathbf{h} \mathbf{h}^H] := C_h(\tau) = \text{diag}(p_1 \dots p_L) \quad (2)$$

where $(p_1 \dots p_L) := P(\tau)$ is the power delay profile (PDP). It gives the power density (per delay unit) incident onto a local area as a function of the propagation delay τ . The Fourier transform of the PDP is the correlation function in the frequency domain:

$$C_H(\Delta f) = \int_{-\infty}^{\infty} P(\tau) e^{-j2\pi \Delta f \tau} d\tau \quad (3)$$

where Δf is the frequency difference between two particular subcarriers. The PDP $P(\tau)$ is often assumed to follow an exponential model:

$$P(\tau) = P_0 e^{-\tau/\sigma_\tau} , \tau \geq 0 \quad (4)$$

where σ_τ is the delay spread which characterizes the PDP duration and depends on the physical environment.

From (3) and (4), we obtain the frequency-domain correlation function as:

$$C_H(\Delta f) = \frac{P_0 \sigma_\tau}{1 + 2\pi j \sigma_\tau \Delta f} \quad (5)$$

Another parameter that characterizes the environment is the coherence bandwidth (Δf_C) which is defined as the bandwidth over which the correlation is above a certain threshold (typically 0.7):

$$\Delta f_C \approx \frac{1}{2\pi \sigma_\tau} \quad (6)$$

In the simulation, we consider a normalized frequency-domain correlation function, i.e. $P_0 \sigma_\tau = 1$. If we have Q channel subcarriers, we create a correlation matrix given by:

$$\mathbf{C}_H = \begin{bmatrix} 1 & c_{1,2} & \dots & c_{1,Q} \\ c_{1,2}^* & 1 & \dots & c_{2,Q} \\ \vdots & & \ddots & \vdots \\ c_{1,Q}^* & c_{2,Q}^* & \dots & 1 \end{bmatrix} \quad (7)$$

where the $c_{i,j}$ coefficients are given by (5). Due to the spreading, there is a spacing of N subcarriers between two symbol components. If we denote by Δf_k the frequency spacing between k subcarriers, we have $\Delta f_k = k\Delta f_N$, where Δf_N is the subcarrier bandwidth. The subcarrier bandwidth is a fraction, depending on N , of the channel coherence bandwidth:

$$\Delta f_N = \frac{\beta}{N} \Delta f_C , \beta \in \mathbb{R}^+ \quad (8)$$

From this, if $\beta/N \ll 1$, the channel is highly correlated, i.e., the frequency spacing between two subsequent symbol components is smaller than the channel coherence bandwidth. If $\beta/N \gg 1$, the channel is strongly uncorrelated, i.e., the frequency spacing between two subsequent symbol components is larger than the channel coherence bandwidth.

In order to generate a correlated channel, we will perform a Choleski decomposition of the correlation matrix. It decomposes a square, symmetric and positive semi-definite matrix into a unique product of a lower triangular matrix and its hermitian, such that:

$$\mathbf{C}_H = \mathbf{T} \mathbf{T}^H \quad (9)$$

where

$$\mathbf{T} = \begin{bmatrix} t_{1,1} & 0 & \dots & 0 \\ t_{2,1} & t_{2,2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ t_{Q,1} & t_{Q,2} & \dots & t_{Q,Q} \end{bmatrix} \quad (10)$$

The Choleski coefficients can be computed in a recursive way thanks to the Cholski algorithm. So, for a given PDP model and delay spread, we can determine the coefficients of the Choleski matrix

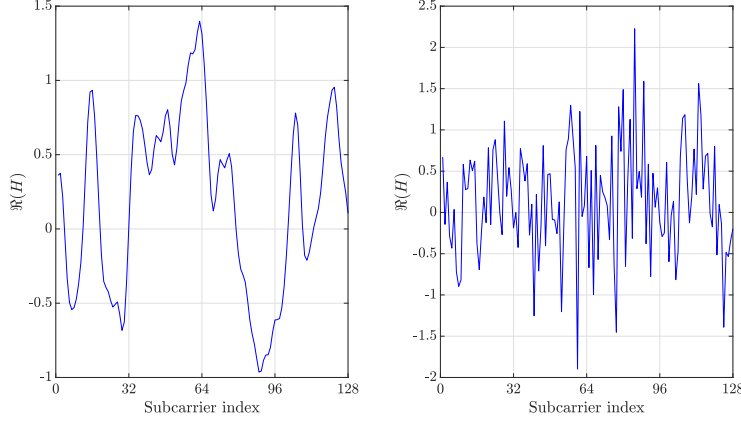


Figure 1: Correlated channel ($\beta/N = 1/6$) ; Uncorrelated channel ($\beta/N = 100$)

The diagonal elements are also positive. Bob correlated channel (\mathbf{H}_B) will then be the product of the Choleski matrix \mathbf{T} with an uncorrelated channel (\mathbf{H}), i.e., $\mathbf{H}_B = \mathbf{T} \mathbf{H}$. As a consequence, a particular channel component will be a weighted sum depending on the channel components:

$$h_{B,j} = \sum_{k=1}^j t_{j,k} h_k \quad (11)$$

Note that for an uncorrelated channel, $\mathbf{T} = \mathbf{I}_Q$ is the identity matrix. In doing so, each subcarrier is independent from the others. We also have:

$$\sum_{k=1}^j |t_{j,k}|^2 = 1 \quad (12)$$

3 SINR Modeling

In this section, we introduce correlation among Bob's subcarriers only. We consider that Eve channel is uncorrelated. *Au début, je pensais intuitivement que ne pas mettre de la corrélation chez Eve était le worst case scénario en terme de sécurité (que ça maximisait la capacité à Eve). Ça s'est avéré ne pas être vrai (cfr plus bas). Mais du coup, je n'ai fait les calculs qu'en considérant de la corrélation à Bob.* As a reminder, the ergodic SINR is modeled in order to find an approximation of the ergodic capacity using the Jensen's inequality:

$$\mathbb{E} [\log_2(1 + X)] \leq \log_2(1 + \mathbb{E} [X]) \quad (13)$$

where X is Bob/Eve SINR. When no correlation was introduced, the inequality bound was tight such that it was used as an approximation.

As a reminder, Bob and Eve channels are spatially uncorrelated.

3.1 Bob SINR

At Bob, the received signal is given by:

$$\mathbf{y}_B^H = \sqrt{\alpha} \mathbf{S}^H |\mathbf{H}_B|^2 \mathbf{S} \mathbf{x} + \mathbf{S}^H \mathbf{v}_B \quad (14)$$

The ergodic SINR is then:

$$\mathbb{E}[\gamma_B] = \frac{\alpha \mathbb{E} \left[\left| \mathbf{S}^H \mathbf{H}_B \mathbf{S} \right|^2 \right]}{\mathbb{E} \left[\left| \mathbf{S}^H \mathbf{v}_B \right|^2 \right]} \quad (15)$$

For a particular symbol n , the expected value of the data component is:

$$\begin{aligned} \mathbb{E} \left[|\text{data}|^2 \right] &= \frac{\alpha}{U^2} \mathbb{E} \left[\left| \sum_{i=0}^{U-1} h_{B,n+iN} \right|^2 \right]^2 \\ &= \frac{\alpha}{U^2} \mathbb{E} \left[\sum_{i=0}^{U-1} |h_{B,n+iN}|^4 + \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} |h_{B,n+iN}|^2 |h_{B,n+jN}|^2 \right] \end{aligned} \quad (16)$$

We can show that:

$$\mathbb{E} \left[\sum_{i=0}^{U-1} |h_{B,n+iN}|^4 \right] = 2U \quad (17)$$

Also:

$$\begin{aligned} \mathbb{E} \left[\sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} |h_{B,n+iN}|^2 |h_{B,n+jN}|^2 \right] &= \sum_{i=0}^{U-1} \sum_{j > i}^{U-1} \left(\sum_{k=1}^{n+iN} 2 |t_{n+iN,k}|^2 |t_{n+jN,k}|^2 + \right. \\ &\quad \left. \sum_{k=1}^{n+iN} \sum_{k' \neq k}^{n+jN} |t_{n+iN,k}|^2 |t_{n+jN,k'}|^2 + \right. \\ &\quad \left. \sum_{k=1}^{n+iN} \sum_{k' > k}^{n+jN} 2R[t_{n+iN,k} t_{n+iN,k'}^* t_{n+jN,k} t_{n+jN,k'}^*] \right) \end{aligned} \quad (18)$$

For the noise component:

$$\mathbb{E} \left[|\text{noise}|^2 \right] = \sigma_B^2 \quad (19)$$

The SINR is simply the ratio between eq.(18) and (19)

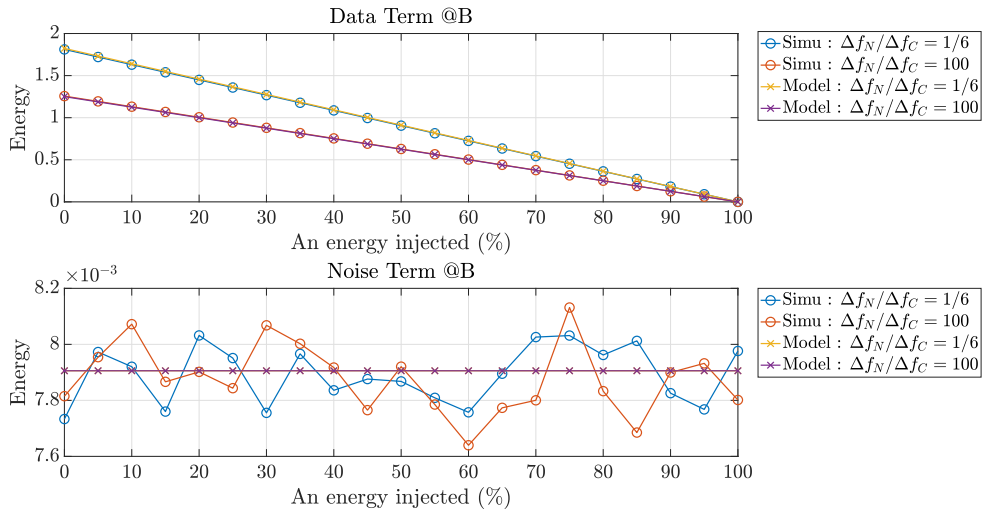


Figure 2: Bob SINR terms fitting, modelization vs simulation

From Fig.2, we observe that the analytic expressions (18) and (19) well fit the simulation results. In particular, we remark that the received data energy increases when the correlation increases.

3.2 Eve SINR

SDS Decoding Structure

When Eve simply despread the received signal, we have

$$\mathbf{y}_E^{SDS} = \sqrt{\alpha} \mathbf{S}^H \mathbf{H}_B^H \mathbf{H}_E \mathbf{S} \mathbf{x} + \mathbf{S}^H \mathbf{v}_E + \sqrt{1-\alpha} \mathbf{S}^H \mathbf{H}_E \mathbf{w} \quad (20)$$

The ergodic SINR is then:

$$\mathbb{E} [\gamma_E^{SDS}] = \frac{\alpha \mathbb{E} \left[\left| \mathbf{S}^H \mathbf{H}_E \mathbf{H}_B^H \mathbf{S} \right|^2 \right]}{\mathbb{E} \left[\left| \mathbf{S}^H \mathbf{v}_E \right|^2 + (1-\alpha) \left| \mathbf{S}^H \mathbf{H}_E \mathbf{w} \right|^2 \right]} \quad (21)$$

For a particular symbol n , the ergodic data component energy is:

$$\begin{aligned} \mathbb{E} [|\text{data}|^2] &= \frac{\alpha}{U^2} \mathbb{E} \left[\left| \sum_{i=0}^{U-1} h_{B,n+iN}^* h_{E,n+iN} \right|^2 \right] \\ &= \frac{\alpha}{U^2} \mathbb{E} \left[\sum_{i=0}^{U-1} |h_{B,n+iN}|^2 |h_{E,n+iN}|^2 + \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} h_{B,n+iN}^* h_{B,n+jN} h_{E,n+iN} h_{E,n+jN}^* \right] \\ &= \frac{\alpha}{U^2} \mathbb{E} \left[\sum_{i=0}^{U-1} |h_{B,n+iN}|^2 \cdot 1 \right] = \frac{\alpha}{U} \end{aligned} \quad (22)$$

Eq.(22) gives the same result as if no correlation was introduced at Bob. Obviously the AN and noise terms will not depend on Bob correlation. For the noise component:

$$\mathbb{E} [|\text{noise}|^2] = \sigma_E^2 \quad (23)$$

For the AN component:

$$\mathbb{E} [|\text{AN}|^2] = \frac{(1-\alpha)}{U} \sigma_{AN}^2 = (1-\alpha) \quad (24)$$

The SINR is therefore identical as in the case where Bob channel is uncorrelated.

Je ne montre pas les calculs pour le matched filter ni le own channel decoder parce que le problème qui suit apparaît déjà pour la capacité à Bob. Mais, les mêmes soucis apparaissent aussi pour les structures de décodage à Eve.

4 Jensen's innequality issue

In the past works, the Jensen innequality was used as a good approximation when no correlation was introduced among Bob's subcarriers. When the correlation is introduced, the approximation does not hold anymore, as seen in Fig(3). In particular, we observe that, when $\Delta f_N / \Delta f_C = 100$, i.e., uncorrelated channel, the fit between analytic and simulation curves is almost perfect. However, when $\Delta f_N / \Delta f_C = 1/6$, i.e., highly correlated channel, the ergodic capacity decreases but the approximation suggests that it should increase. In order to understand the behaviour of the analytic model of the capacity derived using the Jensen's innequality as an approximation, we will analyze

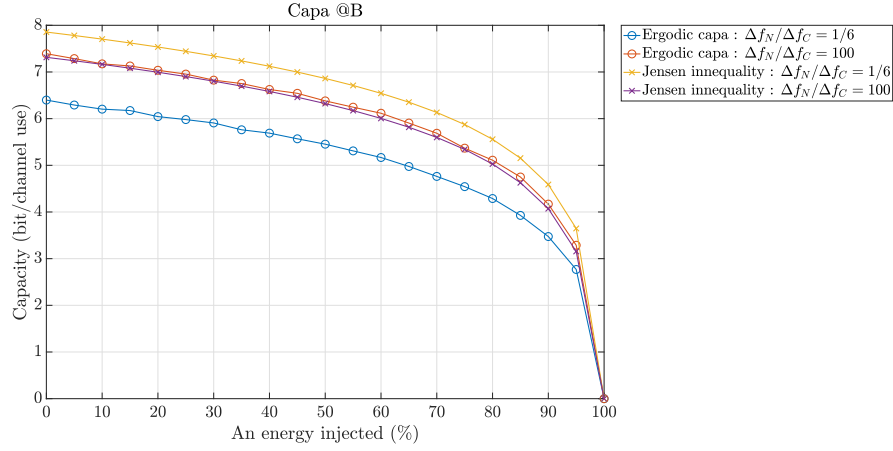


Figure 3: Bob capacity, variable correlation, ergodic vs Jensen's innequality

eq.(16). When Bob's channel is fully correlated, i.e., all subcarriers are an identical complex scalar, $h_{B,n+iN} = h_{B,n+jN}$, $\forall i, j$. Therefore, eq.(16) becomes:

$$\begin{aligned} \mathbb{E} [|\text{data}|^2] &= \frac{\alpha}{U^2} \mathbb{E} \left[\left| \sum_{i=0}^{U-1} |h_{B,n+iN}|^2 \right|^2 \right] \\ &= \frac{\alpha}{U^2} \mathbb{E} \left[\sum_{i=0}^{U-1} |h_{B,n+iN}|^4 + \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} |h_{B,n+iN}|^2 |h_{B,n+jN}|^2 \right] \\ &= \frac{\alpha}{U^2} (2U + 2U(U-1)) = 2\alpha \end{aligned} \quad (25)$$

When Bob's channel is totally uncorrelated, eq.(16) becomes:

$$\begin{aligned} \mathbb{E} [|\text{data}|^2] &= \frac{\alpha}{U^2} \mathbb{E} \left[\left| \sum_{i=0}^{U-1} |h_{B,n+iN}|^2 \right|^2 \right] \\ &= \frac{\alpha}{U^2} \mathbb{E} \left[\sum_{i=0}^{U-1} |h_{B,n+iN}|^4 + \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} |h_{B,n+iN}|^2 |h_{B,n+jN}|^2 \right] \\ &= \frac{\alpha}{U^2} (2U + U(U-1)) = \frac{\alpha(U+1)}{U} \end{aligned} \quad (26)$$

Eq.(25) and (26) suggest that Bob ergodic SINR increases when the correlation increases:

$$\left. \frac{\alpha(U+1)}{U\sigma_B^2} \right|_{\text{no correl}} \leq \mathbb{E} [\gamma_B] \leq \left. \frac{2\alpha}{\sigma_B^2} \right|_{\text{full correl}} \quad (27)$$

This behaviour can be seen in Fig.(4). This implies that the approximated capacity $\log_2 (1 + \mathbb{E} [\gamma_B])$ should increase with an increase of the correlation. However, it is observed in Fig.(4) that the ergodic capacity $\mathbb{E} [\log_2 (1 + \gamma_B)]$ decreases with an increase of the correlation.

Je ne sais pas expliquer pourquoi l'ergodic capacit  se comporte diff remment que l'ergodic SINR... Any idea? J'ai lu qu'on pouvait peut- tre estimer l'erreur provoqu e par Jensen en calculant le Jensen's gap. Mais  a ne m'a pas l'air trivial du tout...

Another approximation of the capacity is therefore needed.

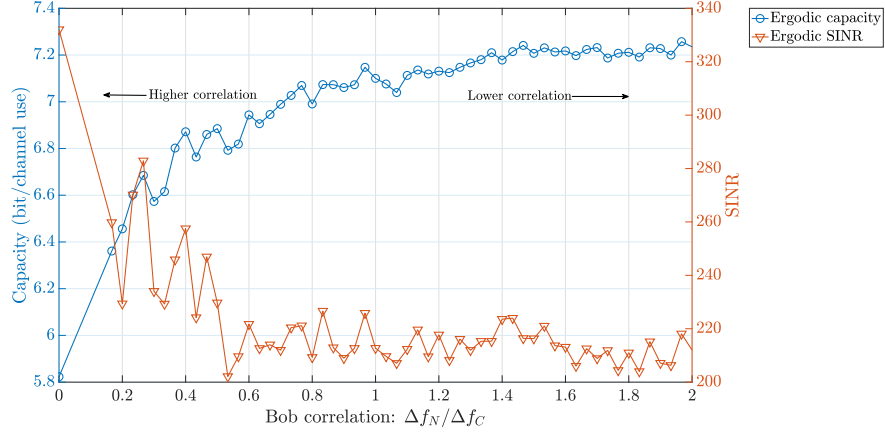


Figure 4: Bob SINR/capacity as a function of the frequency correlation

5 Second order Taylor approximation of the capacity

In order to better approximate the ergodic capacity, it was decided to perform a second order Taylor expansion of $\mathbb{E} [\log_2 (1 + \gamma)]$ around $x_0 = \mathbb{E} [\gamma]$, where γ represents Bob or Eve SINR, depending on the investigated scenario:

$$\mathbb{E} [\log_2 (1 + \gamma)] \approx \log_2 (1 + \mathbb{E} [\gamma]) - \frac{\text{var}(\gamma)}{2(1 + \mathbb{E} [\gamma])^2} \quad (28)$$

where $\text{var}(\gamma) = \mathbb{E} [\gamma^2] - (\mathbb{E} [\gamma])^2$ is the SINR variance, which is quite complicated to derive.

Auriez-vous une idée pour une meilleure approximation?

Bob variance derivation

The terme $\mathbb{E} [\gamma_B^2]$ is difficult to derive. In fact:

$$\mathbb{E} [\gamma_B^2] = \left(\frac{\alpha \mathbb{E} [|\mathbf{S}^H \mathbf{H}_B|^2 \mathbf{S}|^2]}{\mathbb{E} [|\mathbf{S}^H \mathbf{v}_B|^2]} \right)^2 = \frac{\alpha^2 \mathbb{E} [|\mathbf{S}^H \mathbf{H}_B|^2 \mathbf{S}|^4]}{\mathbb{E} [|\mathbf{S}^H \mathbf{v}_B|^4]} \quad (29)$$

Since $v_B \sim \mathbf{CN}(0, \sigma_B^2) \sim N(0, \sigma_B^2/2) + 1jN(0, \sigma_B^2/2)$, it is easy to show that ¹:

$$\mathbb{E} [|\mathbf{S}^H \mathbf{v}_B|^4] = 2(\sigma_B^2)^2 \quad (30)$$

¹Personnal note: cfr FC17 verso for the derivation

In what concerns the data term, for a particular symbol n , the derivation is quite long:

$$\begin{aligned}
\mathbb{E} [|\text{data}|^4] &= \frac{\alpha^2}{U^4} \mathbb{E} \left[\left| \sum_{i=0}^{U-1} |h_{B,n+iN}|^2 \right|^4 \right] \\
&= \frac{\alpha^2}{U^4} \mathbb{E} \left[\sum_{i=0}^{U-1} |h_{B,n+iN}|^8 + 4 \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} |h_{B,n+iN}|^2 |h_{B,n+jN}|^6 \right. \\
&\quad + 3 \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} |h_{B,n+iN}|^4 |h_{B,n+jN}|^4 + 6 \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} \sum_{k \neq i,j}^{U-1} |h_{B,n+iN}|^4 |h_{B,n+jN}|^2 |h_{B,n+kN}|^2 \\
&\quad \left. + \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} \sum_{k \neq i,j}^{U-1} \sum_{l \neq i,j,k}^{U-1} |h_{B,n+iN}|^2 |h_{B,n+jN}|^2 |h_{B,n+kN}|^2 |h_{B,n+lN}|^2 \right]
\end{aligned} \tag{31}$$

Knowing that:

$$\begin{aligned}
\mathbb{E} [h_{B,i} h_{B,j}] &\neq 0 \\
h_{B,n+iN} &= \sum_{k=1}^{n+iN} t_{n+iN,k} h_k
\end{aligned} \tag{32}$$

the resolution of eq(31) is tedious. Par exemple le second terme en $h^2 h^6$ donne lieu à 10 termes à calculer. Chacun de ces 10 termes correspond à plusieurs lignes de calcul. Le dernier terme de (31) en $h^2 h^2 h^2 h^2$ donne lieu au calcul de 6 termes, dont un fait 13 lignes de calculs... Néanmoins je les ai tous calculé et vérifié un par un. Si tout est décorrélé, ce calcul est très simple car à ce moment l'expected value d'un produit est égal au produit des expected values. Il faut juste savoir que: $\mathbb{E} [|h|^2] = 1$, $\mathbb{E} [|h|^4] = 2$, $\mathbb{E} [|h|^6] = 6$ et $\mathbb{E} [|h|^8] = 24$. Après, je me rends compte que le calcul est ultra simple aussi dans le cas *full correlated*. Dans le sens où on aurait plus que du $\mathbb{E} [|h|^8]$ pour tous les termes car $h_{B,n+iN} = h_{B,n+jN}, \forall i, j$. Peut-être qu'alors l'approximation serait meilleure??? (voir la suite pour comprendre que l'approximation n'est pas top). Est-ce que vous désirez les calculs complets??

Eve SDS decoder variance

Also derived but less complicated since:

$$\begin{aligned}
\mathbb{E} [|\text{data}|^4] &= \frac{\alpha^2}{U^4} \mathbb{E} \left[\left| \sum_{i=0}^{U-1} h_{B,n+iN}^* h_{E,n+iN} \right|^4 \right] \\
&= \frac{\alpha^2}{U^4} \mathbb{E} \left[\sum_{i=0}^{U-1} |h_{B,n+iN}|^4 |h_{E,n+iN}|^4 + 2 \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} |h_{B,n+iN}|^2 |h_{B,n+jN}|^2 |h_{E,n+iN}|^2 |h_{E,n+jN}|^2 \right] \\
&= \frac{2\alpha^2}{U^4} \mathbb{E} \left[\sum_{i=0}^{U-1} |h_{B,n+iN}|^4 + \sum_{i=0}^{U-1} \sum_{j \neq i}^{U-1} |h_{B,n+iN}|^2 |h_{B,n+jN}|^2 \right]
\end{aligned} \tag{33}$$

6 Results

Approximation accuracy comparison

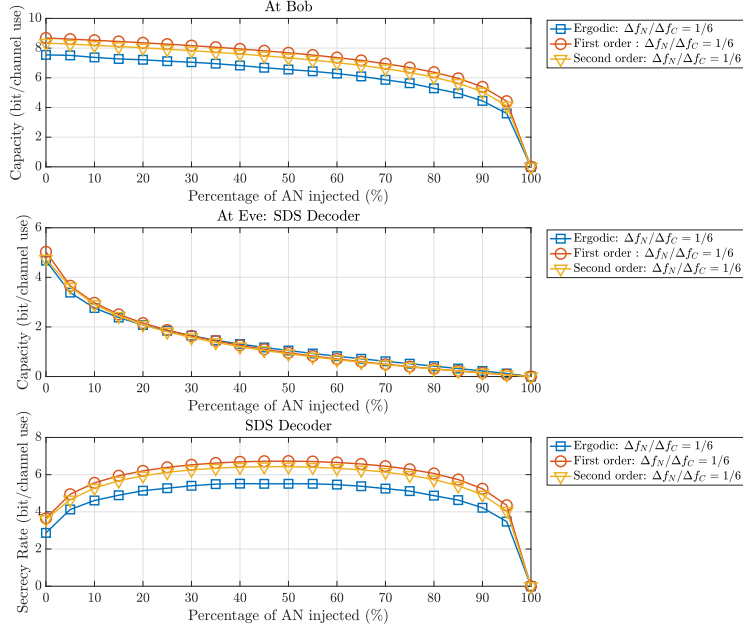


Figure 5: High correlation, $Q = 16$, $U = 8$, approximation comparison

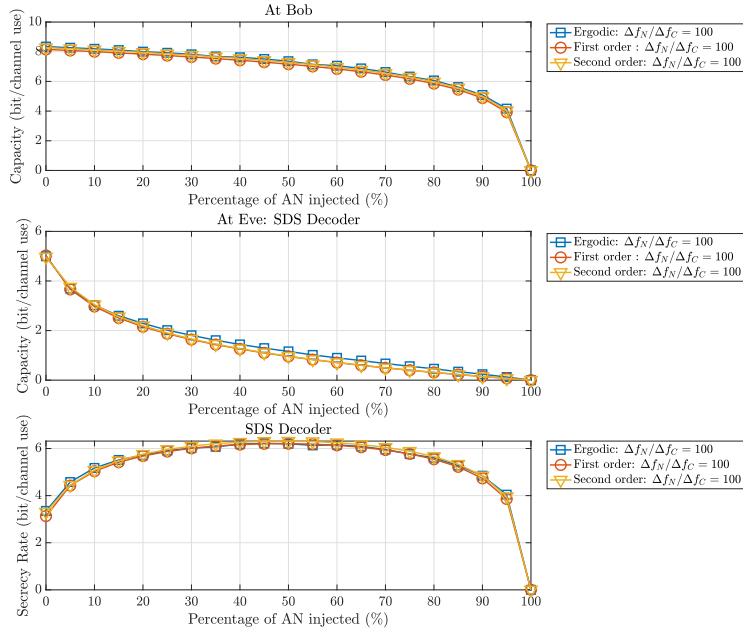


Figure 6: No correlation, $Q = 16$, $U = 8$, approximation comparisons

Fig5 and 6 compare the accuracy of the two approximations (first order using Jensen, second

order using Jensen plus the variance term) with the ergodic results, for 2 different environments, i.e., high correlation (fig5) and low correlation (fig6).

As expected, when no correlation is introduced, the first order approximation approximates well the ergodic curves. There is no need to use another approximation. We also observe that the second order approximation is better than the first order approximation when the correlaton is introduced. However, the SR bound is not tight. Furthermore, it is an upper bound when we would be happier with a lower one in order to a priori know at least the rate at which to securely communicate.

What is interesting though, is that the SR curves present the same concavity. In particular, the analytic determination of the optimal amount of data to inject (optimal α) can be accurately found using the first order approximation (I did it).

Fig.7 shows the maximal SR value as a function of the amount of correlation at Bob. We observe that the first order approximation increases with an increase of the correlation, as already explained. The second order remains quasi constant.

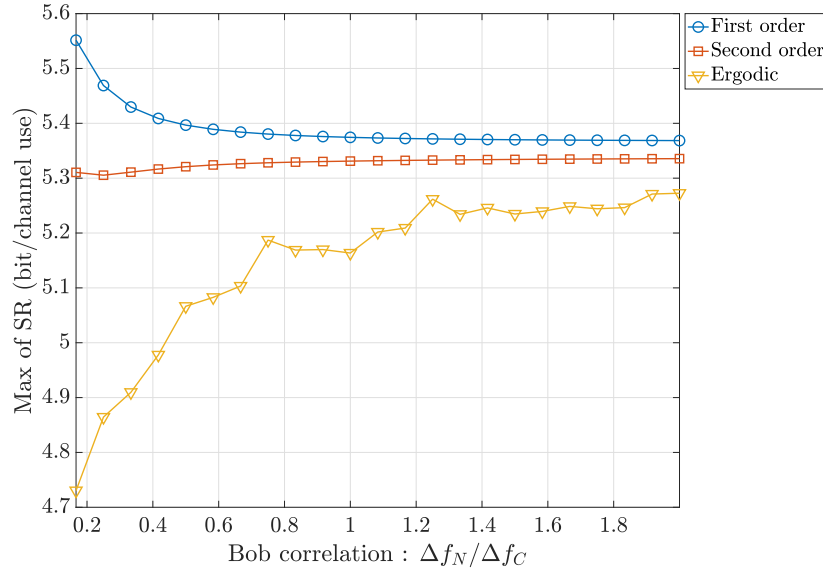


Figure 7: Maximum of SR, SDS Decoder at Eve, variable correlation at Bob, $Q = 32$, $U = 4$, approximation comparaison

Ici, les questions suivantes se posent:

- L'approximation du second est-elle utilisable?
- Y'a t'il une autre approximation qui pourrait être utilisée?

Other obtained results

- Determination of the optimal α to maximize the ergodic SR
- First and second order approximations for Eve matched filter and own channel decoders

- Required SNR at Bob in order to guarantee a given SR, i.e. $\text{SNR} = \infty$ at Eve. No analytic expression, only via simulations since the analytic approximations are not tight enough.
- Simulations (no analytic model yet) with variable correlation among Eve's subcarriers, i.e., influence of Eve correlation on the performances. It seem that introducing correlatin at Eve decreases the SR performances

POUR CONCLURE:

Je suis un peu embêté car je ne sais pas trop comment valoriser ces résultats. Est-ce que tout étudier via des simulations seulement est intéressant? Dans le sens où les formules dérivées ne sont pas terribles. Ca sera un peu bête de tout jeter...