CORRIGÉ.

Problème: Détection d'un pic dans le spectre d'un signal bruité.

1) Moindres carrés.

a) Par définition :

$$\hat{s}_m = \sum_{n=1}^{N} S_n W^{nm}. \tag{1}$$

Il suffit ensuite de développer le module :

$$\epsilon = \sum_{m=1}^{N} (\hat{s}_m - s_m)(\bar{\hat{s}}_m - \bar{s}_m) = \sum_{m=1}^{N} (\sum_{n=1}^{N} S_n W^{nm} - s_m)(\sum_{n'=1}^{N} \bar{S}_{n'} W^{-n'm} - \bar{s}_m)$$

$$= \sum_{m=1}^{N} \left\{ |s_m|^2 - \bar{s}_m \sum_{n=1}^{N} S_n W^{nm} - s_m \sum_{n'=1}^{N} \bar{S}_{n'} W^{-n'm} + \sum_{n=1}^{N} S_n W^{nm} \sum_{n'=1}^{N} \bar{S}_{n'} W^{-n'm} \right\}$$

b) Dérivons :

$$\frac{\partial \epsilon}{\partial S_n} == \sum_{m=1}^{N} \left\{ -\bar{s}_m W^{nm} + W^{nm} \sum_{n'=1}^{N} \bar{S}_{n'} W^{-n'm} \right\}$$

$$\frac{\partial \epsilon}{\partial \bar{S}_n} == \sum_{m=1}^{N} \left\{ -s_m W^{-nm} + W^{-nm} \sum_{n'=1}^{N} S_{n'} W^{n'm} \right\}$$

Ces deux relations sont les conjuguées l'une de l'autre et donnent donc la même condition. Par exemple $\frac{\partial \epsilon}{\partial \bar{S}_n} = 0$ équivaut à :

$$\sum_{m=1}^{N} s_m W^{-nm} = \sum_{m=1}^{N} W^{-nm} \sum_{n'=1}^{N} S_{n'} W^{n'm} = \sum_{n'=1}^{N} S_{n'} \sum_{m=1}^{N} W^{m(n'-n)} = \sum_{n'=1}^{N} S_{n'} \gamma.$$

c) Reste à calculer γ :

$$\gamma = \sum_{m=1}^{N} W^{m(n-n')} = \sum_{m=1}^{N} (W^{(n-n')})^m = \sum_{m=1}^{N} z^m = z \frac{1 - z^N}{1 - z} \tag{2}$$

avec $z = W^{(n-n')}$; alors :

$$z^N = \exp\big(2i\pi(n - n')\big) = 1.$$

Ce qui montre que $\gamma = 0$ pour $n \neq n'$. Quant à $n = n'$, c'est évident puisqu'alors $z = 1$ et qu'ainsi $\gamma = N$.

d) Il s'ensuit directement que :

$$\sum_{n'=1}^{N} S_{n'}\gamma = NS_n$$

ce qui donne le résultat cherché :

$$S_n = \frac{1}{N}\sum_{m=1}^{N} s_m \bar{W}^{nm} = \frac{1}{N}\sum_{m=1}^{N} s_m \exp\left(-2i\pi\frac{nm}{N}\right) \tag{3}$$

2) Loi de l'amplitude spectrale.

a) La transformée de Fourier étant calculée à partir de la sommation d'un (généralement grand) nombre de v.a. (cf. relation (3)) le th. central-limite implique que B_r et B_i sont gaussiennes.

Puisqu'elles sont indépendantes :

$$f_{(B_r,B_i)}(B_r, B_i) = f_{B_r}(B_r)f_{B_i}(B_i) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{B_r^2+B_i^2}{2\sigma^2}\right).$$

b) La relation :

$$Se^{i\varphi} = \mathbb{A} + \mathbb{B}$$

s'écrit aussi sur ses parties réelles et imaginaires :

$$B_r = S\cos\varphi - A_r \quad \text{et} \quad B_i = S\sin\varphi - A_i.$$

Le module du jacobien de la transformation $(S,\varphi) \to (B_r, B_i)$ vaut donc :

$$J = \left|\frac{\partial B_r}{\partial S}\frac{\partial B_i}{\partial \varphi} - \frac{\partial B_r}{\partial \varphi}\frac{\partial B_i}{\partial S}\right| = S$$

d'où, puisque l'on fait ici la transformation inverse :

$$\begin{aligned} f(S,\varphi) &= Jf(B_r(S,\varphi), B_i(S,\varphi)) = Sf(B_r(S,\varphi), B_i(S,\varphi)) \\ &= \frac{S}{2\pi\sigma^2}\exp\left(-\frac{(S\cos\varphi - A_r)^2 + (S\sin\varphi - A_i)^2}{2\sigma^2}\right). \end{aligned}$$

c) La loi marginale se calcule par intégration de la loi conjointe sur l'ensemble des valeurs de φ possibles, c'est-à-dire de 0 à 2π. Commençons par remarquer qu'avec les notations proposées :

$$\begin{aligned} (S\cos\varphi - A_r)^2 + (S\sin\varphi - A_i)^2 &= S^2 + A^2 - 2S(A_r\cos\varphi + A_i\sin\varphi) \\ &= S^2 + A^2 - 2SA\cos(\varphi - \varphi_0). \end{aligned}$$

Par intégration :

$$f(S) = \int_0^{2\pi} f(S,\varphi)\ \mathrm{d}\varphi = \frac{S}{2\pi\sigma^2}\exp\left(-\frac{S^2+A^2}{2\sigma^2}\right)\int_0^{2\pi}\exp\left(-\frac{SA}{\sigma^2}\cos(\varphi-\varphi_0)\right)\ \mathrm{d}\varphi$$

Puisque l'on intègre une fonction 2π-périodique il revient au même d'intégrer sur n'importe quel intervalle de longueur 2π, par exemple $[\varphi_0, \varphi_0+2\pi]$. En faisant un changement de variable on élimine alors φ_0 :

$$f(S) = \int_0^{2\pi} f(S,\varphi)\ \mathrm{d}\varphi = \frac{S}{2\pi\sigma^2}\exp\left(-\frac{S^2+A^2}{2\sigma^2}\right)\int_0^{2\pi}\exp\left(-\frac{SA}{\sigma^2}\cos(\varphi)\right)\ \mathrm{d}\varphi$$

$$= \frac{S}{\sigma^2}\exp\left(-\frac{S^2+A^2}{2\sigma^2}\right) I_0\left(\frac{AS}{\sigma^2}\right),$$

c.q.f.d. avec :

$$f_0 = \frac{1}{\sigma^2}\exp\left(-\frac{A^2}{2\sigma^2}\right).$$

f est bien sûr nulle pour $S \leq 0$ puisque $S > 0$.

d) $(A_r+B_r)^2 = A_r^2 + B_r^2 + 2A_rB_r$: le premier terme est une valeur certaine (non aléatoire). Le troisième terme a autant de chance d'être positif que négatif puisque B_r est centrée. Par contre le deuxième terme est toujours positif ce qui crée un biais lorsque l'on prend le module d'un signal bruité : il est plus probable que S soit supérieur à A qu'inférieur. C'est la raison pour laquelle l'abscisse du maximum des lois est toujours supérieur à 1. Lorsque dans un signal bruité on estime l'amplitude d'une onde par une méthode de moindres carrés généralisés on trouvera donc un résultat surestimé, c'est-à-dire plus grand que la valeur réelle.

3) Signification de la détection d'un pic.

a) Pour $A = 0$ on a évidemment $I_0(AS/s^2) = I_0(0) = 1$ et $f_0 = 1/\sigma^2$ d'où le résultat. C'est une d.d.p. car $f \geq 0$ et car

$$\int_{-\infty}^{\infty} f(S)\ \mathrm{d}S = \int_0^{\infty}\frac{S}{\sigma^2}\exp\left(-\frac{S^2}{2\sigma^2}\right)\ \mathrm{d}S = \left[-\exp\left(-\frac{S^2}{2\sigma^2}\right)\right]_0^{\infty} = 1$$

b) De la même façon :

$$P_k = \int_{k\sigma}^{\infty} f(S)\ \mathrm{d}S = \left[-\exp\left(-\frac{S^2}{2\sigma^2}\right)\right]_{k\sigma}^{\infty} = \exp\left(-\frac{k^2}{2}\right)$$

La probabilité qu'il ne le dépasse pas vaut évidemment $1 - P_k$.

c) Pour N réalisations, il faut multiplier entre elles les probabilités : $p_k = (1-P_k)^N$. La probabilité qu'au moins un parmi les N dépasse ce seuil est donc $1 - p_k = 1 - (1-P_k)^N$.

d) $1-(1-P_k)^N$ tend vers 1 quand N tend vers l'infini donc en augmentant N on est presque sûr de voir des pics. C'était évident. Il faut néanmoins le quantifier. A partir des relations précédentes :

k	P_k	p_k	$1-p_k$
2	0,135	$2\ 10^{-65}$	1
3	0,011	$1\ 10^{-5}$	0,999989
4	0,0003	0,71	0,29
5	0,000004	0,9962	0,0038

La probabilité que le pic soit du au hasard est faible si $k=4$, ou 5 pour être plus « confortable » (c'est-à-dire 5σ).

4) Test numérique.

a) Soit u(x) la d.d.p. d'une v.a. uniforme. Il faut chercher le changement de variable S(x) telle que $f(S)\ \mathrm{d}S = u(x)\ \mathrm{d}x$. En intégrant on arrive directement à :

$$\int_{-\infty}^{S(x)} f(S')\ \mathrm{d}S' = \int_{-\infty}^{x} u(x')\ \mathrm{d}x'$$

c'est-à-dire $F(S(x)) = U(x)$. Vues les formes de ces fonctions :

$$F(S(x)) = \int_0^{S(x)} \frac{S'}{\sigma^2} \exp\left(-\frac{S'^2}{2\sigma^2}\right)\ \mathrm{d}S' = 1-\exp\left(-\frac{S(x)^2}{2\sigma^2}\right)$$

et

$$U(x) = \int_0^x \mathrm{d}x' = x$$

Il faut donc appliquer la transformation :

$$S(x) = \sigma\sqrt{-2\ln(1-x)}$$

aux réalisations $x \in [0,1[$.

b) Le nombre moyen de pics dépassant $k\sigma$ est NP_k c'est-à-dire 135, 11, 0,3 et 0 pour $k = 2,3,4,5$.

c) Cela correspond à la réalisation donnée (cf. figure), pour laquelle on voit que ces nombres valent 11 pour $k=3$ et 1 pour $k=4$.