



# Traitement du signal lidar

A. Dabas

# Plan

1. Quelques brefs rappels sur le traitement numérique du signal
  - Numérisation des signaux.
  - Transformation de Fourier des signaux numériques.
  - Démodulation complexe
2. Quelques généralités sur les estimateurs.
  - Estimateur au maximum de vraisemblance.
  - Borne de Cramer Rao
3. Lidar en détection directe
  - Équation lidar (rappel) et propriétés statistiques des signaux.
  - Estimation du coefficient d'atténuation.
  - Estimation d'une épaisseur optique.
4. Lidar en détection hétérodyne
  - Modèle « feuilleté », propriétés statistiques des signaux et modèle de Zrnic.
  - Estimation de la puissance du signal
  - SNR et CNR
  - Estimation Doppler.



# TRAITEMENT DU SIGNAL

# Numérisation des signaux

- Deux opérations

- Échantillonnage: la valeur du signal est prélevée à intervalles de temps réguliers.
- Quantification: la valeur du signal est « codée » informatiquement, généralement en un entier.

- L'échantillonnage

- Caractérisé par la période de temps  $T_s$  entre deux échantillons, ou son inverse  $F_s = 1/T_s$  appelée fréquence d'échantillonnage.

- La quantification

- La plage de numérisation  $[-A/2, A/2]$  est divisée en  $2^b$  intervalles de largeur  $2^{-b}A$  numéroté de 0 à  $2^b-1$ . A chaque valeur du signal, on associe l'indice de l'intervalle dans lequel il se trouve:

$$x_{\text{num}} = \text{floor} \left\{ \left( x_{\text{ana}} + \frac{A}{2} \right) \frac{2^b}{A} \right\}$$

- A est une caractéristique du Convertisseur Analogique/Numérique (fixe ou réglable), on adapte la dynamique du signal en conséquence.
- Ce faisant, on introduit une erreur d'arrondi qui peut prendre la valeur maximale  $2^{-b}A$ :

$$|x_{\text{ana}} - x_{\text{num}}| \leq A/2^b$$

- Le paramètre b est le nombre de bits de quantification. C'est une autre caractéristique du CAN. Plus il est élevé, moins grandes sont les erreurs d'arrondi, mais plus il faut de mémoire pour stocker les signaux.

# Transformation de Fourier analogique

- Définition

$$X(f) = \int_{-\infty}^{+\infty} x(t) \exp(-2j\pi ft) dt \quad \text{où } j^2 = -1$$

C'est une opération linéaire transformant une fonction du temps  $t$  (s) à valeurs réelles ou complexes en fonction de la fréquence  $f$  (en Hz) à valeurs complexes.

- Transformation inverse

$$x(t) = \int_{-\infty}^{+\infty} X(f) \exp(+2j\pi ft) df$$

Cette formule montre qu'il y a autant d'information dans la TF d'une fonction que dans la fonction elle-même. Avec la TF, on voit le même objet mathématique, mais avec un point de vue différent.

- Intérêt de la TF

La TF permet de repérer des périodicités dans un signal. Elles apparaissent par des pics.

Exemple: 
$$x(t) = \begin{cases} \cos(2\pi f_0 t) & \text{si } t \in [-T/2, T/2] \\ 0 & \text{sinon} \end{cases} \xleftrightarrow{\text{TF}} X(f) = \frac{1}{2} \frac{\sin(\pi(f-f_0)T)}{\pi(f-f_0)} + \frac{1}{2} \frac{\sin(\pi(f+f_0)T)}{\pi(f+f_0)}$$

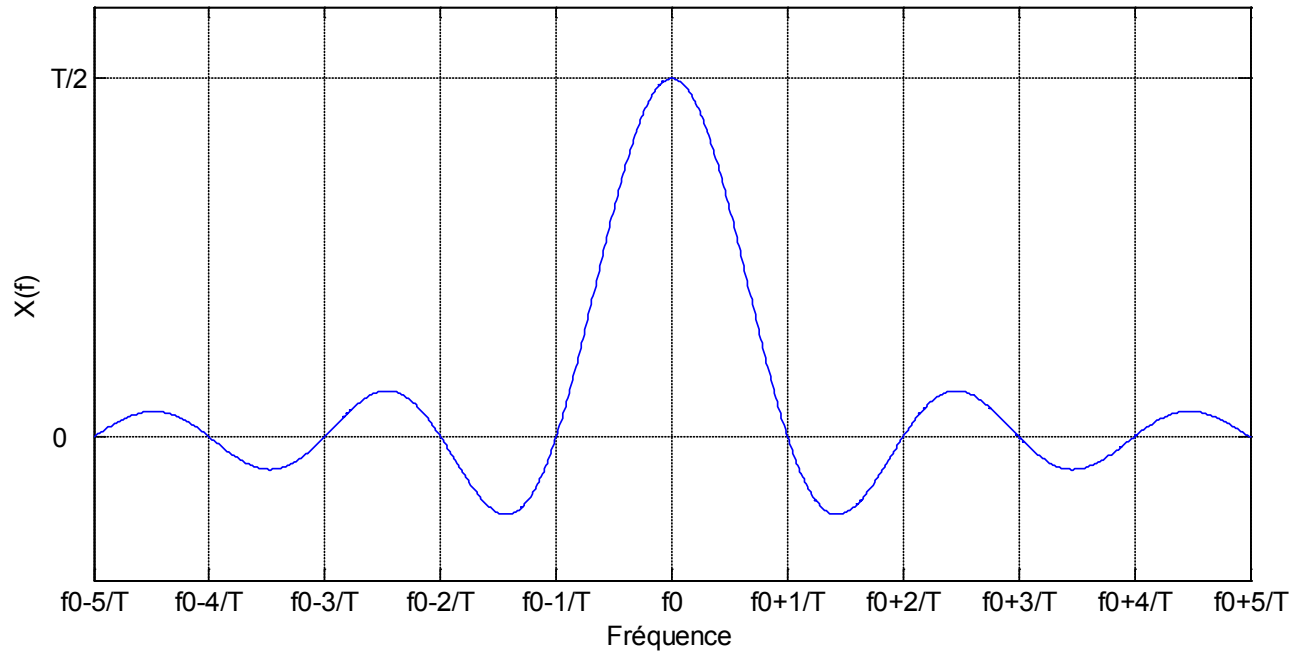
- Incertitude temps-fréquence

Comme le montre l'exemple ci-dessus, la largeur d'une TF est inversement proportionnelle à la durée du signal. On peut montrer de manière très générale

$$\Delta f \Delta t \geq \frac{1}{4\pi} \quad \text{où } \Delta t^2 = \frac{\int t^2 x(t) dt}{\int x(t) dt} \quad \text{et } \Delta f^2 = \frac{\int f^2 X(f) df}{\int X(f) df}$$

Ceci aura des conséquences importantes pour les lidars Doppler : la précision avec laquelle on saura estimer la fréquence du signal est inversement proportionnelle à la durée d'observation...

# Transformation de Fourier analogique



$$\frac{1}{2} \frac{\sin(\pi(f-f_0)T)}{\pi(f-f_0)} = \frac{T}{2} \text{sinc}(\pi(f-f_0)T) \quad \text{où} \quad \text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

# Transformation de Fourier numérique

- Définition

$$\hat{X}(f) = \sum_{k=-\infty}^{+\infty} x(kT_s) \exp(-2j\pi f k T_s)$$

Comme la TF analogique, elle transforme un signal à valeurs réelles ou complexes en un signal à valeurs complexes.

- Transformation inverse

$$x(kT_s) = \frac{1}{F_s} \int_{-F_s/2}^{+F_s/2} \hat{X}(f) \exp(+2j\pi f k T_s) df$$

Conclusion: il y a autant d'information dans la TF du signal numérique que dans le signal numérique lui-même.

- Relation entre TF numérique et TF analogique

$$\hat{X}(f) = \sum_{k=-\infty}^{+\infty} X(f - kF_s)$$

- Condition de Shannon et filtre anti-repliement

D'après la formule précédente, si  $X(f)=0$  pour  $f \geq F_s/2$  et  $f \leq -F_s/2$ , alors on a

$$\hat{X}(f) = X(f) \text{ pour } f \in [-F_s/2, F_s/2]$$

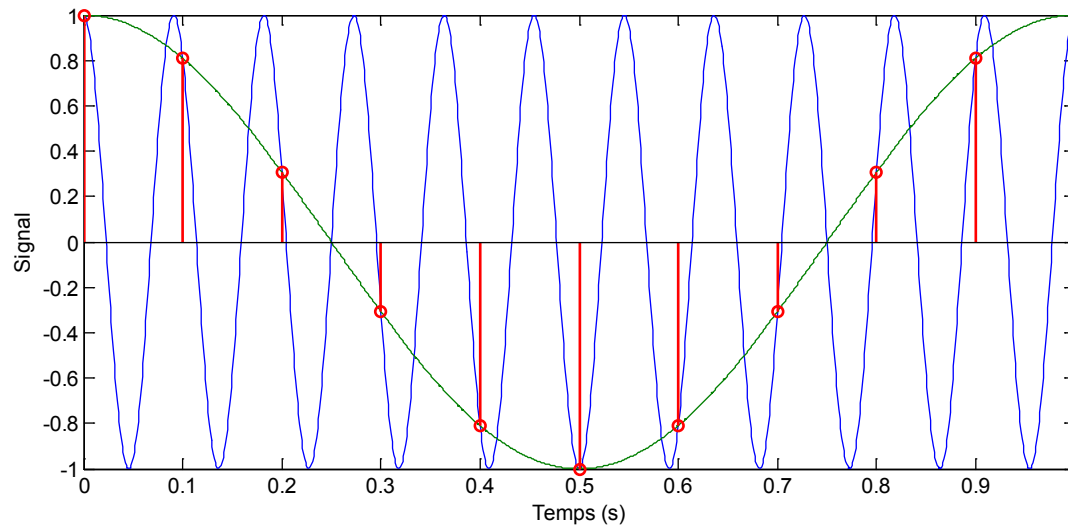
Autrement dit, si la condition précédente est réunie (condition de Shannon), on a autant d'information dans le signal numérisé que dans le signal analogique!

Pour le signal périodique  $\exp(2j\pi f_0 t)$ , la condition de Shannon est vérifiée dès lors qu'il y a au moins 2 points d'échantillonnage par période du signal.

En pratique, on filtre le signal avant échantillonnage pour supprimer les fréquences supérieures à  $F_s/2$ .

# Repliement de spectre

*Illustration du phénomène de repliement quand la condition de Shannon n'est pas vérifiée...*



Bleu :  $y(t) = \cos(2\pi \cdot 11 \cdot t)$  (fréquence 11Hz)

Vert :  $y(t) = \cos(2\pi t)$  (fréquence 1Hz)

Échantillonnés à la fréquence  $F_s=10\text{Hz}$ , les 2 sinusoïdes donnent les mêmes échantillons...



# Transformation de Fourier numérique

- TF et produit de convolution

$$z(kT_s) = (x * y)(kT_s) = \sum_m x(mT_s)y((k-m)T_s) \xleftrightarrow{\text{TF}} \hat{Z}(f) = \hat{X}(f)\hat{Y}(f)$$
$$z(kT_s) = x(kT_s)y(kT_s) \xleftrightarrow{\text{TF}} \hat{Z}(f) = (\hat{X} * \hat{Y})(f) = \int \hat{X}(u)\hat{Y}(f-u)df$$

Autrement dit, la TF transforme un produit en produit de convolution et inversement.

- TF et filtrage numérique

Toute opération de filtrage numérique d'un signal  $x(kT_s)$  peut s'exprimer de la manière suivant:

$$\underbrace{a_0 y(kT_s) + a_1 y((k-1)T_s) + \dots + a_p y((k-p)T_s)}_{\text{Sortie}} = \underbrace{b_0 x(kT_s) + b_1 x((k-1)T_s) + \dots + b_q x((k-q)T_s)}_{\text{Entrée}}$$

Compte tenu des propriétés ci-dessus, on a alors la relation

$$\hat{A}(f)\hat{X}(f) = \hat{B}(f)\hat{Y}(f)$$

- Densité spectrale d'énergie

- Définition

$$\Gamma_x(f) = |\hat{X}(f)|^2$$

Cette fonction de la fréquence s'appelle la densité spectrale d'énergie en vertu de la propriété suivante

$$\int_{-F_s/2}^{F_s/2} \Gamma_x(f) df = F_s \sum_{k=-\infty}^{+\infty} |x(kT_s)|^2$$

# Transformation de Fourier Rapide

En pratique, un signal est numérisé sur une **période d'observation finie**. On a donc un **nombre fini M d'échantillons**. Dans ce cas, on définit la Transformation Rapide de Fourier et son inverse par les équations suivantes:

$$\hat{x}\left(\frac{kF_s}{M}\right) = \sum_{m=0}^{M-1} x(mT_s) \exp\left(-2j\pi \frac{km}{M}\right) \quad x(mT_s) = \sum_{k=0}^{M-1} \hat{x}\left(\frac{kF_s}{M}\right) \exp\left(+2j\pi \frac{km}{M}\right)$$

Les algorithmes de TFR établissent une correspondance entre la valeur de la TF pour les fréquences  $f_k = kF_s/M$  avec  $k=0$  à  $M-1$  et la valeur de  $M$  échantillons de signal  $x(mT_s)$  pour  $m=0$  à  $M-1$ .

$$\begin{bmatrix} x(0) \\ x(T_s) \\ \vdots \\ x((M-1)T_s) \end{bmatrix} \xleftrightarrow{\text{TFR}} \begin{bmatrix} \hat{x}(0) \\ \hat{x}(F_s/M) \\ \vdots \\ \hat{x}((M-1)F_s/M) \end{bmatrix}$$

La résolution spectrale de la TFR est égale à  $F_s/M$  où  $M$  est le nombre d'échantillons temporels.

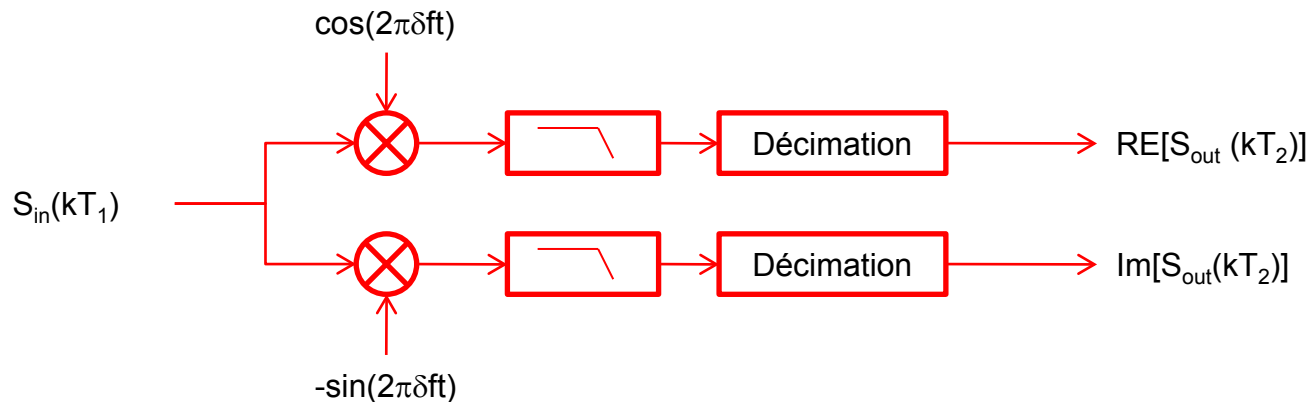
La TFR est optimisée en temps de calcul si  $M$  est une puissance de 2.

# Démodulation complexe

- Objectif:

La démodulation complexe a pour but de décaler un signal « réel » (tel que le signal brut numérisé en sortie de détecteur) en un signal « imaginaire » (composé de 2 voies en quadrature) dont le contenu spectral est identique mais décalé en fréquence.

Dans un lidar hétérodyne, cette opération permet de supprimer le décalage  $f_e - f_{OL}$  de sorte que la fréquence nulle corresponde à un Doppler nul.



Exemple:  $S_{in}(kT_1) = \cos(2j\pi f kT_1) \dots$



# ESTIMATEURS

# Maximum de vraisemblance

- Vraisemblance

Soit  $x_1, x_2, \dots, x_n$ ,  $n$  observations d'un système physique  $x$  réglé par un paramètre (ou un ensemble de paramètres)  $\theta$ . La fonction de vraisemblance des  $n$  observations est définie par

$$L(\theta) = P(x_1, \dots, x_n | \theta)$$

Autrement dit, la vraisemblance est la fonction qui exprime la probabilité d'observer effectivement  $x_1, x_2, \dots, x_n$  selon la valeur du paramètre  $\theta$ .

- Estimation au maximum de vraisemblance

L'estimation au maximum de vraisemblance du paramètre  $\theta$  est la valeur de ce paramètre pour lequel la probabilité d'observer  $x_2, \dots, x_n$  est maximale. On part du principe que si  $x_2, \dots, x_n$  ont effectivement été observés, c'est qu'ils avaient une bonne chance de l'être...

$$\hat{\theta}_{ML} = \arg[\max(L(\theta))]$$

Assez souvent, on optimise  $\Lambda(\theta) = \log(L(\theta))$  plutôt que  $L(\theta)$ , les calculs sont plus simples...

- Borne de Cramer Rao

On peut montrer que pour n'importe quel estimateur non biaisé de  $\theta$ , on a

$$E\left[\left(\hat{\theta} - \theta\right)^2\right] \geq F^{-1}$$

Où  $F$  est la matrice carrée (matrice de Fisher) définie par  $F_{i,j} = -E\left[\frac{\partial^2 \Lambda(\theta)}{\partial \theta_i \partial \theta_j}\right]$

Cette inégalité définit la Borne de Cramer Rao, limite ultime de la précision (en terme de variance) de tout estimateur  $\rightarrow$  la BCR est une référence. Si un estimateur a une variance proche de la BCR, pas la peine de casser la tête à essayer de faire mieux.

Un estimateur non biaisé dont la variance est égale à la BCR est dit efficace.

S'il existe un estimateur efficace, l'estimateur ML l'est.

# Maximum de vraisemblance

- Exemple

On cherche à trouver l'amplitude (complexe) et la fréquence du signal bruité suivant

$$x(kT_s) = \theta_1 \exp(2j\pi\theta_2 kT_s) + e(kT_s) \quad \text{avec} \quad k = 1, \dots, M$$

où  $e(kT_s)$  est un bruit blanc numérique complexe de statistique gaussienne, c'est-à-dire

$$\langle e(kT_s) \rangle = 0 \quad \langle e^*(kT_s)e(mT_s) \rangle = \sigma_e \delta(m-k) \quad p(e(kT_s)) = \frac{1}{\sqrt{\pi}\sigma_e} \exp\left(-\frac{e(kT_s)^2}{\sigma_e^2}\right)$$

On peut alors facilement montrer que

$$\hat{\theta}_2^{\text{ML}} = \arg\left[\max\left\{\hat{\chi}(f)\right\}^2\right] \quad \text{et} \quad \hat{\theta}_1^{\text{ML}} = \frac{1}{M} \hat{\chi}(\hat{\theta}_2^{\text{ML}})$$

Autrement dit, l'estimation au maximum de vraisemblance de la fréquence du signal est la fréquence qui maximise la densité spectrale du signal, et la valeur de la TF à cette fréquence donne directement (à un facteur multiplicatif près) l'amplitude du signal.



# DÉTECTION DIRECTE

# Équation lidar

- Équation lidar

L'équation peut s'écrire de la manière suivante:

$$S(t) = K \int_0^{+\infty} \text{Imp}\left(t - \frac{2z}{c}\right) \frac{\beta(z)T^2(z)}{z^2} dz + e(t)$$

Où  $\text{Imp}(t)$  décrit l'impulsion laser,  $K$  est une constante multiplicative,  $\beta(z)$  est le coefficient de rétrodiffusion,  $e(t)$  du bruit (Gaussien) et  $T^2(z)$  l'atténuation:

$$T^2(z) = \exp\left(-2 \int_0^z \alpha(x) dx\right)$$

Il est courant de simplifier l'équation lidar ci-dessus en considérant que le rapport  $\beta(z)T^2(z)/z^2$  ne varie pas de manière significative sur la longueur de l'impulsion. On a alors

$$S(z) \approx KE_0 \frac{\beta(z)T^2(z)}{z^2} + e(z) \quad \text{avec} \quad z = ct/2$$



# Estimation de l'atténuation

On se place ici dans le cas où on peut considérer que  $\beta(z)$  et  $\alpha(z)$  est quasi constant sur la porte de mesure, et on cherche à estimer le coefficient  $\alpha$ . Pour cela, on prend

$$U(z) = \log(z^2 S(z))$$

et on cherche à approcher la courbe par une droite dont la pente fournit une estimation de  $\alpha$ . En absence de bruit, on a en effet:

$$U(z) = U_0 - 2\alpha z$$

Soient  $U(z_k)$  les échantillons du signal dont on dispose, on a

$$\hat{\alpha} = \frac{M \sum_{k=0}^{M-1} z(k) U(k) - \sum_{k=0}^{M-1} z(k) \sum_{k=0}^{M-1} U(k)}{M \sum_{k=0}^{M-1} U^2(k) - \sum_{k=0}^{M-1} U(k) \sum_{k=0}^{M-1} U(k)}$$

Problème: Si  $e(t)$  est un bruit gaussien que l'on peut supposer centré (en ayant préalablement corrigé le signal de sa moyenne), il n'est plus de même pour  $U(z)$ . A l'ordre 2, on a en effet

$$\log(z^2 S(z)) = \log(z^2 \bar{S}(z)) + \log\left(1 + \frac{e(z)}{\bar{S}(z)}\right) \approx \bar{U}(z) + \frac{e(z)}{\bar{S}(z)} - \frac{1}{2} \frac{e^2(z)}{\bar{S}^2(z)} \quad \text{d'où} \quad \langle U(z) \rangle = \bar{U}(z) - \frac{1}{2SNR(z)}$$

# Estimation de l'épaisseur optique

Elle consiste à réaliser l'opération précédente avant et après la couche atmosphérique dont on veut connaître l'épaisseur optique, puis à faire la différence entre les droites de régression. En effet, si on note

$$U_1(z) \approx U_1 - 2\alpha_1 z \quad \text{et} \quad U_2(z) \approx U_2 - 2\alpha_2 z$$

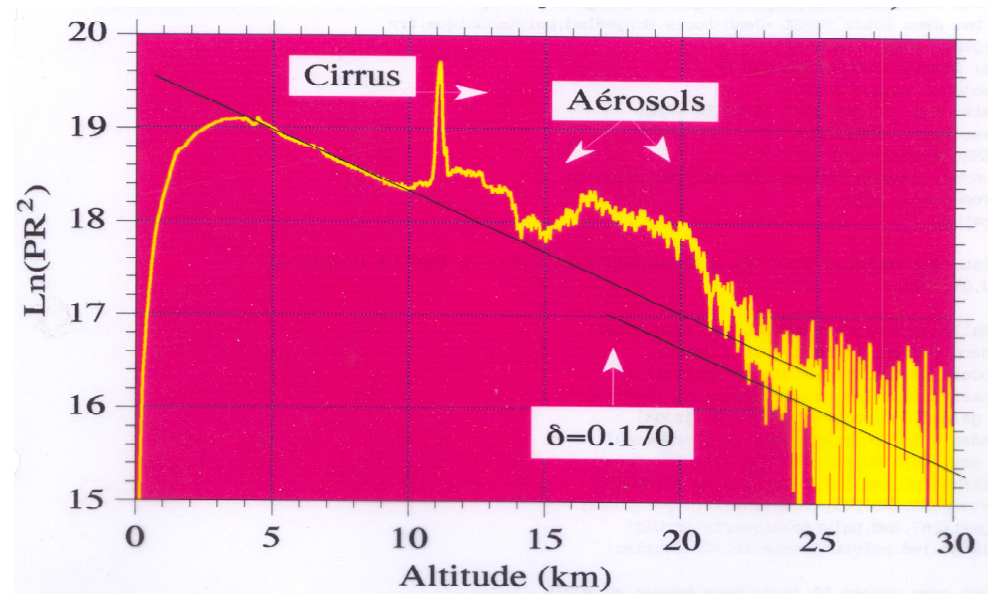
Les droites avant et après la couche, on a

$$\hat{\delta} = \frac{U_1 - U_2}{2}$$

Mais là encore, attention, on a un problème de biais dû à la non-linéarité du log appliqué au bruit. On a en effet d'après la page précédente:

$$\langle \hat{\delta} \rangle = \delta + \frac{1}{4} \left[ \frac{1}{\text{SNR}_2} - \frac{1}{\text{SNR}_1} \right]$$

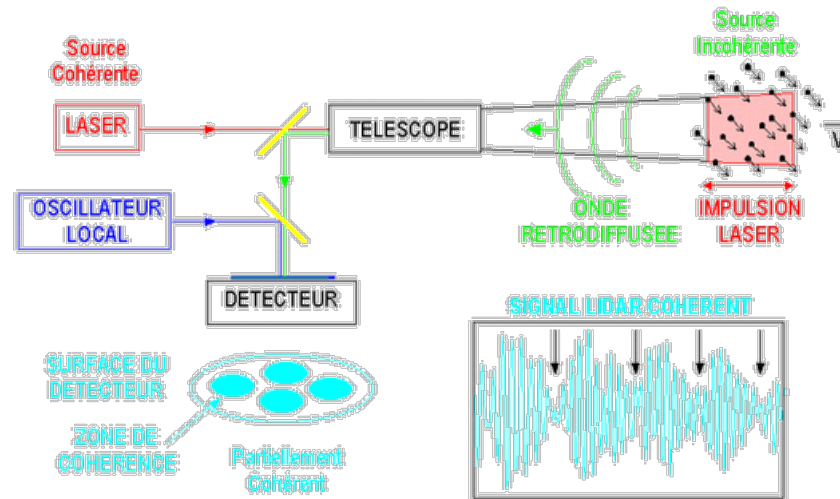
Ceci conduit à une surestimation de l'épaisseur optique...





# DÉTECTION HÉTÉRODYNE

# Rappel du principe de fonctionnement



L'onde em rétrodiffusée est superposée physiquement au faisceau d'un laser continu (Oscillateur Local). Le détecteur produit un courant électrique mesurant l'interférence entre les deux:

$$\begin{aligned}
 i(t) &= K \int_D \|E_r(\vec{x}, t) \exp(2j\pi\nu_r t) + E_{oi}(\vec{x}) \exp(2j\pi\nu_{oi} t)\|^2 d\vec{x} \\
 &= K \int_D \|E_r(\vec{x}, t)\|^2 d\vec{x} + K \int_D \|E_{oi}(\vec{x})\|^2 d\vec{x} + 2K \operatorname{Re} \left\{ \exp(2j\pi(\nu_r - \nu_{oi})t) \int_D E_r(\vec{x}, t) E_{oi}^*(\vec{x}) d\vec{x} \right\} \\
 &= i_r(t) + i_{oi} + 2\gamma \sqrt{i_r(t) i_{oi}} \cos(2\pi(\nu_r - \nu_{oi})t + j\phi)
 \end{aligned}$$

Signal direct  
(basse fréquence)

Signal hétérodyne  
(fréquence intermédiaire)

$\gamma$  = rendement hétérodyne

# Modèle feuilleté (1/3)

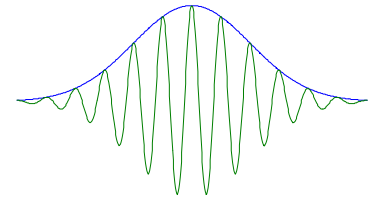
Salamitou et al., *Applied Optics*, 34, 499-506

## • Onde émise

$$E_e(\vec{x}, t, 0) = g_e(t) e_e(\vec{x}, 0) \exp(2j\pi\nu_e t)$$

Enveloppe temporelle

Profil spatial transverse (normalisé)



## • Onde émise propagée à la distance z

$$E_e(\vec{p}, t, z) = g_e(t - z/c) \exp(2j\pi\nu_e(t - z/c)) \int_S e_e(\vec{x}, 0) G(\vec{x}, \vec{p}, t) d\vec{x}$$

Enveloppe retardée du temps de propagation.

Porteuse déphasée à cause du temps de propagation.

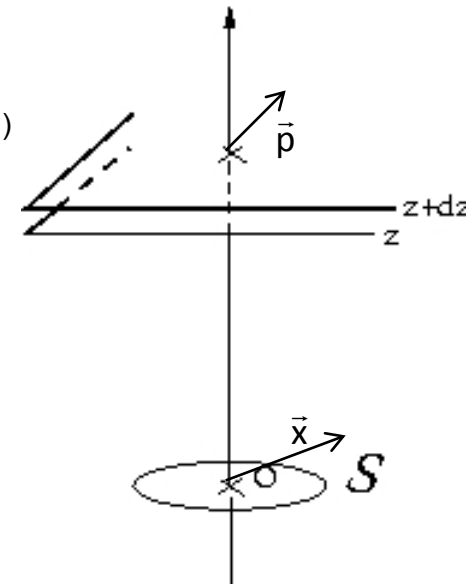
Propagateur (fonction de Green)

## • Onde diffusée par les diffuseurs du « feuillet » [z, z+dz]

Chaque particule du feuillet « renvoie » l'onde qu'elle reçoit...

$$E_r(\vec{p}, t, z) = \sum_{\text{diffuseurs } i} \alpha_i g_e(t - z_i/c) \exp(2j\pi\nu_e(t - z_i/c)) \int_S e_e(\vec{x}, 0) G(\vec{x}, \vec{p}_i, t) d\vec{x}$$

$\alpha_i$  caractérise la capacité de la cible  $i$  à diffuser l'onde em.



# Modèle feuilleté (2/3)

- Onde diffusée propagée jusqu'au plan de détection

$$E_r(\vec{y}, t, 0) = \sum_{\text{diffuseurs } i} \alpha_i g_e(t - 2z_i/c) \exp(2j\pi v_e(t - 2z_i/c)) \int_S e_e(\vec{x}, 0) G(\vec{x}, \vec{p}_i, z_i) G^*(\vec{p}_i, \vec{y}, -z_i) d\vec{x}$$

$$\approx g_e(t - 2z/c) \exp(2j\pi v_e t) \sum_{\text{diffuseurs } i} \alpha_i \exp(-2jkz_i) \int_S e_e(\vec{x}, 0) G(\vec{x}, \vec{p}_i, z_i) G^*(\vec{p}_i, \vec{y}, -z_i) d\vec{x}$$

$$k = 2\pi/\lambda$$

- Mélange optique  $E_{ol}(\vec{x}, t) = g_{ol} \exp(2j\pi f_{ol} t) e_{ol}(\vec{x}, 0)$

$$x(t) = \int_S E_r(\vec{y}, t, 0) E_{ol}^*(\vec{y}, t) d\vec{y} \approx \exp(2j\pi(v_e - v_{ol})t) g_e \left( t - \frac{2z}{c} \right) g_{ol} \underbrace{\sum_{\text{diffuseurs } i} \alpha_i \exp(-2jkz_i) \iint_{S^2} e_e(\vec{x}, 0) e_{ol}^*(\vec{y}, 0) G(\vec{x}, \vec{p}_i, z_i) G(\vec{p}_i, \vec{y}, -z_i) d\vec{x} d\vec{y}}_{\kappa(z, t)}$$

Réponse impulsionnelle de la couche élémentaire  $[z, z+dz]$

- Intégration dans la profondeur

$$x(t) = \exp(2j\pi(v_e - v_{ol})t) g_{ol}^* \underbrace{\int_0^{+\infty} g_e(t - 2z/c) \kappa(z, t) dz}_{\text{Produit de convolution}}$$

Produit de convolution

# Modèle feuilleté (3/3)

- Décalage Doppler

La fonction  $\kappa(z,t)$  dépend de  $t$  à travers la position  $z_i$  des particules. Pendant le temps que la particule voit l'impulsion laser (de l'ordre de la microseconde), la trajectoire de la particule peut être approchée au premier ordre:

$$z_i(t) = z_i^0 + v_r^i t$$

On a alors

$$\kappa(z,t) = \sum_i \alpha_i \exp(-2jkz_i^0) \exp(-2jkv_r^i t) \iint_{S^2} e_e(\vec{x}) e_{ol}^*(\vec{y}) G(\vec{x}, \vec{p}_i, z_i) G(\vec{p}_i, \vec{y}, -z_i) d\vec{x} d\vec{y}$$

↓

$$\text{Décalage Doppler } \Delta v_i = -2v_r^i / \lambda$$

- Back-Propagated Local Oscillator

Dans l'équation ci-dessus, la double intégrale représente un mélange optique entre l'onde émise propagée jusqu'à la position de la cible  $i$

$$e_e(\vec{p}_i, z_i) = \int_S e_e(\vec{x}, 0) G(\vec{x}, \vec{p}_i, z_i) d\vec{x}$$

et l'onde issue de l'oscillateur local « rétro-propagée » jusqu'à la même position

$$e_{bplo}(\vec{p}_i, z_i) = \int_S e_{ol}(\vec{y}, 0) G^*(\vec{y}, \vec{p}_i, z_i) d\vec{y}$$

On peut en effet écrire

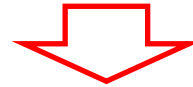
$$\kappa(z,t) = \sum_i \alpha_i \exp(-2jkz_i^0) \exp(-2jkv_r^i t) e_e(\vec{p}_i, z_i) e_{bplo}^*(\vec{p}_i, z_i)$$

Le BPLO permet d'exprimer le mélange optique dans le plan de la cible plutôt que celui du récepteur (et accessoirement de limiter le nombre d'intégrales dans les équations...).

# Statistique de la réponse impulsionnelle

$$\kappa(z, t) = \sum_i \alpha_i \exp(-2jkz_i^0) \exp(-2jkv_r^i t) e_e(\vec{p}_i, z_i) e_{bplo}^*(\vec{p}_i, z_i)$$

- $\kappa(z, t)$  est la somme sur un grand nombre de diffuseurs d'une quantité aléatoire: les diffuseurs étant en effet répartis aléatoirement dans le feuillet  $[z, z+dz]$  avec  $dz \gg \lambda/2$ , le terme de phase  $\exp(-2jkz_i)$  est lui-même aléatoire.
- $\kappa(z, t)$  représente une marche aléatoire dans le plan complexe. Chaque pas va dans n'importe quelle direction. La marche est constituée d'un grand nombre de pas.



$\kappa(z, t)$  suit une loi normale dans le plan complexe.

Sa moyenne est nulle:  $\langle \kappa(z, t) \rangle = 0$

Les statistiques associées à deux feuillets sont indépendantes  $\langle \kappa^*(z, t) \kappa(z', t) \rangle = 0$

- On peut par ailleurs montrer la relation suivante

$$\langle \kappa^*(z, t) \kappa(z, t') \rangle = \left\langle \sum_i |\alpha_i|^2 |e_e(\vec{p}_i, z_i)|^2 |e_{bplo}(\vec{p}_i, z_i)|^2 \right\rangle \approx \underbrace{\gamma(z) \frac{A\beta(z)\Gamma^2(z)}{z^2}}_{\text{}} \langle \exp(-2jkv_r^i(t-t')) \rangle$$

Obtenu par identification avec équation lidar

Où  $\gamma(z)$  est le rendement hétérodyne. Ci-dessus, l'approximation fait l'hypothèse qu'il n'y a pas de corrélation entre les propriétés optiques et dynamiques de l'air. Ce n'est pas toujours le cas (combinaison d'un gradient optique et d'un gradient dynamique).



# Chatoisement

- Valeur instantannée  $x(t) = \exp(2j\pi(v_e - v_{oi})t) g_{oi}^* \int_0^{+\infty} g_e(t - 2z/c) \kappa(z, t) dz$ 
  - Le signal  $x(t)$  étant proportionnel à une somme<sup>0</sup> de variables aléatoires Gaussiennes complexes, elle est elle-même une **variable aléatoire Gaussienne complexe**.
  - Si on pose

$$x(t) = \rho(t) \exp(j\phi(t))$$

on a

$$p(\phi(t)) = \frac{1}{2\pi} \quad p(\rho(t)) = \frac{2\rho(t)}{\langle \rho^2(t) \rangle} \exp\left(-\frac{\rho^2(t)}{\langle \rho^2(t) \rangle}\right) \quad p(\rho^2(t)) = \frac{1}{\langle \rho^2(t) \rangle} \exp\left(-\frac{\rho^2(t)}{\langle \rho^2(t) \rangle}\right)$$

Autrement dit, la valeur instantannée du signal possède une phase aléatoire uniformément répartie entre 0 et  $2\pi$ , et une puissance qui obéit à une loi exponentielle négative.

- Transformée de Fourier

Comme la TF est une combinaison linéaire de variables aléatoires complexes Gaussiennes, c'est aussi une variable aléatoire complexe Gaussienne. On a donc

$$\hat{X}(f) = \sum_{k=0}^{M-1} x(kT_s) \exp(-2j\pi f k T_s)$$

On retiendra ici que les composantes de la densité spectrales suivent toutes une loi exponentielle négative.

$$p(\arg(\hat{X}(f))) = \frac{1}{2\pi} \quad p(|\hat{X}(f)|^2) = \frac{1}{\langle |\hat{X}(f)|^2 \rangle} \exp\left(-\frac{|\hat{X}(f)|^2}{\langle |\hat{X}(f)|^2 \rangle}\right)$$

# Chatoiement

A noter que tous ces résultats demeurent si on rajoute un bruit Gaussien au signal lidar (ce qui est le cas en pratique), car ce bruit n'en modifie pas la nature statistique: on a toujours à faire avec une variable aléatoire Gaussienne complexe. Seule les auto-corrélation et les puissances changent (il faut rajouter celles du bruit).

# Auto-corrélation et spectre

- Auto-corrélation

D'après le modèle « feuilleté »:

$$\langle x^*(t)x(t') \rangle = A \int_0^{+\infty} \gamma(z) \frac{\beta(z)T^2(z)}{z^2} g_e^* \left( t - \frac{2z}{c} \right) g_e \left( t' - \frac{2z}{c} \right) \langle \exp(-2jkv_r^i(t'-t)) \rangle dz$$

Si l'on fait l'hypothèse 1./ que les propriétés optiques de l'air varient peu sur une longueur d'impulsion et 2./ que la dispersion de vitesse est également négligeable, on a

$$\langle x^*(t)x(t') \rangle \propto \underbrace{\int_0^{+\infty} g_e^* \left( t - \frac{2z}{c} \right) g_e \left( t' - \frac{2z}{c} \right) dz}_{\text{autocorrélation de l'impulsion}}$$

Autrement dit, l'autocorrélation du signal est fixée par celle de l'impulsion laser.

Lorsque les hypothèses précédentes ne sont pas vérifiées – notamment au regard de la dispersion des vitesses radiales, l'autocorrélation du signal diminue.

- Spectre

Dans le cas où le signal est stationnaire, on peut facilement montrer que

$$\langle |\hat{X}(f)|^2 \rangle = \sum_{k=1-M}^{M-1} (M-|k|) \gamma_x(kT_s) \exp(-2j\pi f k T_s) \quad \text{où} \quad \gamma_x(kT_s) = \langle x^*(t)x(t+kT_s) \rangle$$

D'où, si on applique l'équation ci-dessus

$$\langle |\hat{X}(f)|^2 \rangle \propto \sum_{k=1-M}^{M-1} (M-|k|) \gamma_g(kT_s)$$

Et si enfin on suppose que la durée d'observation  $MT_s$  est grande devant la durée d'impulsion

$$\langle |\hat{X}(f)|^2 \rangle \propto |\hat{G}_e(f)|^2$$

A la base, le spectre du signal est fixé par le spectre de l'impulsion. Le spectre est ensuite élargi par le fenêtrage (durée de la porte de mesure) et les hétérogénéités à l'intérieur du volume de mesure.

# Modèle de Zrnic

Dans le cas où l'atmosphère est homogène à l'échelle de l'impulsion et où la dispersion des vitesses radiales à cette échelle est faible, nous avons démontré que

$$\langle |\hat{X}(f)|^2 \rangle \propto |\hat{G}_e(f)|^2$$

Par ailleurs, nous avons vu que

$$p(\arg(\hat{X}(f))) = \frac{1}{2\pi} \quad p(|\hat{X}(f)|^2) = \frac{1}{\langle |\hat{X}(f)|^2 \rangle} \exp\left(-\frac{|\hat{X}(f)|^2}{\langle |\hat{X}(f)|^2 \rangle}\right)$$

Enfin, nous n'avons pas démontré, mais nous admettons que si la porte de mesure est grande devant l'impulsion, et si l'homogénéité des propriétés atmosphériques est toujours vérifiée à l'échelle de la porte, alors les fluctuations aléatoires de deux composantes spectrales sont indépendantes.

Le modèle de Zrnic combine ces considérations et propose une manière très simple de simuler un signal lidar ayant les bonnes propriétés statistiques

$$\begin{bmatrix} \sqrt{\xi(0) |\hat{G}_e(0)|^2} \exp(j\phi_0) \\ \sqrt{\xi(1) \left| \hat{G}_e\left(\frac{1}{M} F_s\right) \right|^2} \exp(j\phi_1) \\ \vdots \\ \sqrt{\xi(M-1) \left| \hat{G}_e\left(\frac{M-M}{M} F_s\right) \right|^2} \exp(j\phi_{M-1}) \end{bmatrix} \xleftrightarrow{\text{TFR}^{-1}} \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(M-1) \end{bmatrix}$$

Où les  $\zeta(k)$  sont des variables aléatoires de loi exponentielle négative (de moyenne 1) et les  $\phi(k)$  une phase aléatoire uniformément répartie entre 0 et  $2\pi$ .

# Estimation de la puissance

- Estimateur

Le signal détecté est numérisé est noté  $s(t) = x(t) + n(t)$  où  $x(t)$  est le signal utile et  $n(t)$  est le bruit de détection.

En détection hétérodyne, la source dominante de bruit est le bruit de grenaille (*shot noise*) qui a la particularité d'être blanc, c'est-à-dire que

$$\langle n^*(kT_s)n(lT_s) \rangle = P_n \delta(k-l) \quad \text{d'où} \quad \langle |\hat{N}(f)|^2 \rangle = \left\langle \left| \sum_{k=0}^{M-1} n(kT_s) \exp(-2j\pi f k T_s) \right|^2 \right\rangle = NP_n$$

Ce dernier résultat justifiant le terme de bruit blanc.

L'estimateur usuel de la puissance du signal sur  $M$  échantillons successifs du signal est défini par:

$$\hat{P}_s = \frac{1}{M} \sum_{k=0}^{M-1} |s(kT_s)|^2$$

- Biais et variance

On peut facilement montrer:

$$\langle \hat{P}_s \rangle = P_s = P_x + P_n$$

C'est-à-dire que l'estimateur est non biaisé. Pour le moment d'ordre 2 on a:

$$\langle \hat{P}_s^2 \rangle = \frac{1}{M^2} \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \langle s^*(k)s(k)s^*(l)s(l) \rangle = \frac{1}{M^2} \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \left[ \langle s^*(k)s(k) \rangle \langle s^*(l)s(l) \rangle + \underbrace{\langle s^*(k)s^*(l) \rangle}_{=0} \underbrace{\langle s(k)s(l) \rangle}_{=0} + \langle s^*(k)s(l) \rangle \langle s^*(l)s(k) \rangle \right]$$

d'où

$$\text{Var}[\hat{P}_s] = \frac{1}{M^2} \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \langle |s^*(k)s(l)|^2 \rangle = \frac{P_x^2}{M^2} \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} |\gamma_x(k,l)|^2 + \frac{P_n^2}{M}$$

avec  $\gamma_x(k,l) = \frac{1}{P_x} \langle x^*(k)x(l) \rangle$

# Estimation de la puissance

- Nombre de tavelures

Le **nombre  $M$  de tavelures temporelles** du signal lidar à l'intérieur de la porte de mesure est défini par la relation

$$M = \frac{M^2}{\sum_{k=0}^{M-1} \sum_{l=0}^{M-1} |\gamma_x(k, l)|^2}$$

On a

$$\text{Var}[\hat{p}_s] = \frac{P_x^2}{M} + \frac{P_n^2}{M}$$

Donc  $M$  est le nombre d'échantillons indépendants à l'intérieur de la porte de mesure, où bien encore le nombre de réalisations indépendantes du phénomène de chatoiement à l'intérieur de cette porte.

- Cas où l'estimation est faite sur un nombre  $N$  de signaux

Dans ce cas l'estimation devient 
$$\hat{p}_s = \frac{1}{N} \frac{1}{M} \sum_{n=1}^N \sum_{k=0}^{M-1} |s(k)|^2$$

Et sa variance vaut

$$\text{Var}[\hat{p}_s] = \frac{P_x^2}{NM} + \frac{P_n^2}{NM}$$

# Estimation de puissance

- Loi de probabilité

En fait, si l'on revient à l'expression de l'estimateur et à la statistique de la valeur instantanée du signal, on s'aperçoit que l'estimation de la puissance est une somme de variables aléatoires Gaussiennes complexes. On sait qu'une telle somme suit une loi du khi2 à m degrés de liberté

$$p(\hat{p}_s) = \frac{m^m}{\Gamma(m)} \frac{\hat{p}_s^{m-1}}{P_s^m} \exp\left(-m \frac{\hat{p}_s}{P_s}\right)$$

Où

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt$$

Est la fonction Gamma (pour n entier,  $\Gamma(n+1)=n!$ ) et m est défini par le rapport

$$m = \frac{P_s^2}{\text{Var}[P_s]}$$

D'après les résultats précédents, on a donc:

$$\frac{1}{m} = \frac{1}{NM} \left( \frac{\text{CNR}}{1+\text{CNR}} \right)^2 + \frac{1}{NM} \left( \frac{1}{1+\text{CNR}} \right)^2$$

Où le CNR (Carrier to Noise Ratio) est défini par le rapport

$$\text{CNR} = \frac{P_x}{P_n}$$

# CNR vs. SNR

- Définition du SNR et relation avec le CNR

En lidar hétérodyne, on distingue le CNR et le SNR

Le CNR (défini dans le transparent précédent) est rapport de la puissance de la porteuse du signal Doppler (le signal hétérodyne) sur la puissance moyenne du bruit de détection. C'est le paramètre de bruit pertinent pour estimer la précision des mesures de vitesse.

Le SNR est le rapport de la puissance moyenne du bruit (au carré) sur la variance de celle-ci. C'est le paramètre pertinent quand on s'intéresse aux estimations de puissance de signal.

On a la relation suivante entre les 2:

$$\frac{1}{\text{SNR}} = \frac{1}{\text{NM}} \left( \frac{\text{CNR}}{1 + \text{CNR}} \right)^2 + \frac{1}{\text{NM}} \left( \frac{1}{1 + \text{CNR}} \right)^2$$

- Estimation du CNR

Elle consiste à estimer  $P_s$  sur une porte de mesure, puis à estimer la puissance du bruit  $P_n$  sur une portion de signal sur laquelle on est sûr qu'il n'y a que du bruit (distance très éloigné du lidar par exemple). On fait alors le rapport entre les deux:

$$\hat{\text{CNR}} = \frac{\hat{P}_s}{\hat{P}_n} - 1$$

On peut alors montrer que la variable

$$\hat{z} = \frac{1 + \hat{\text{CNR}}}{1 + \text{CNR}}$$

Suit une loi de Fisher à  $m = \text{NSNR}^{-1}$  et  $\text{NM}_n$  degrés de libertés ( $\text{NM}_n$  est le nombre d'échantillons de bruit sur lesquels on a estimé la puissance du bruit).



# Estimation Doppler

- Maximum de vraisemblance

Dans le cas où le modèle de Zrnic s'applique, on sait que le spectre  $\hat{\Gamma}_s(f) = |\hat{S}(f)|^2$  du signal a la forme moyenne

$$\Gamma_s(f|P_x, P_n, f_0) = P_n T_s + P_x \Gamma_e(f - f_0)$$

Où  $\hat{\Gamma}_s(f) = |\hat{G}_e(f)|^2$ .

Comme on connaît la statistique des composantes spectrales du signal (composantes indépendantes entre elles et chacune de loi exponentielle négative), on peut écrire un estimateur au maximum de vraisemblance. La log-vraisemblance prend en effet la forme suivante

$$\Lambda(P_x, P_n, f_0) = \Lambda_0 - \sum_{k=0}^{M-1} \frac{\hat{\Gamma}_s(kF_s/N)}{\Gamma_s(f|P_x, P_n, f_0)}$$

En pratique, cet estimateur est peu pratique car il faut maximiser une fonction dépendant de 3 paramètres indépendants, ce qui est toujours coûteux en temps de calcul.

En revanche, il apparaît que l'estimateur au maximum de vraisemblance consiste à essayer de trouver les caractéristiques d'un filtre dont la réponse spectrale est  $\Gamma_s^{-1}$  qui produit en sortie un minimum de puissance (pour maximiser la log-vraisemblance). Bref, il s'agit de trouver le filtre adapté au signal...

- BCR

Toujours dans le cas des conditions de Zrnic, on peut développer une formule pour la BCR de l'estimation de fréquence

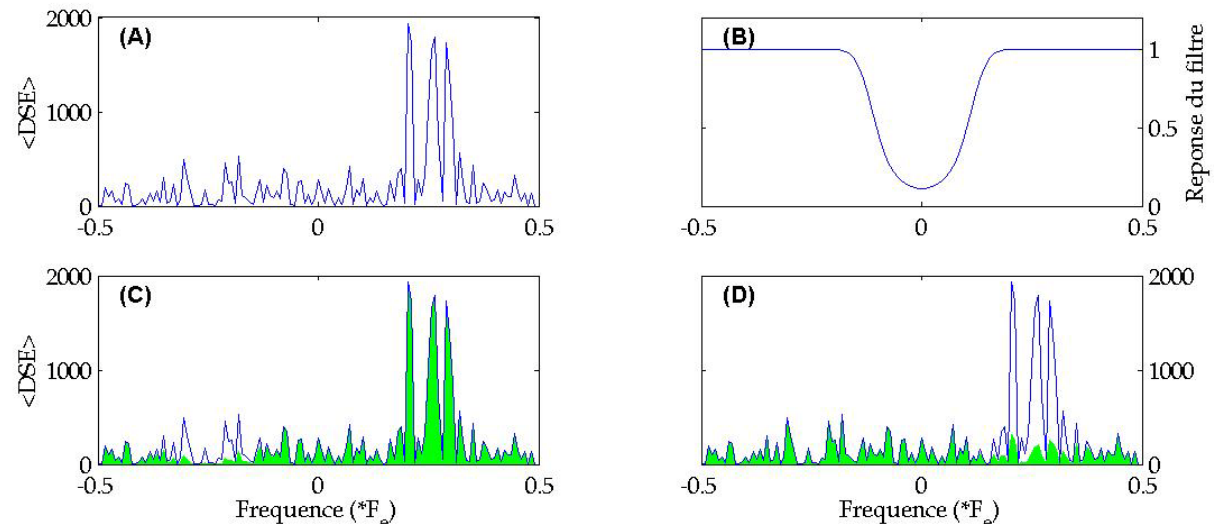
$$\text{BCR} = \left[ -2\pi^2 T_s^2 N \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} (k-l)^2 \text{Cov}_{ss}(k,l) \text{Cov}_s^{-1}(k,l) \right]^{-1/2}$$

Où  $\text{Cov}_{ss}$  désigne la matrice de variance covariance du signal ( $\text{Cov}_{ss}(k,l) = \gamma_s(k,l)$ ).

# Estimation Doppler

- Filtrage adapté

Tous les estimateurs de fréquence utilisés en lidar Doppler hétérodyne sont de type « filtrage adapté », c'est-à-dire qu'on fait passer le signal par un filtre dont on optimise les paramètres jusqu'à obtenir une puissance en sortie de signal maximale ou minimale. La fréquence du signal est ensuite reliée aux paramètres optimisés.



**Illustration:** (A): exemple de spectre de signal lidar réel. (B): filtre que l'on cherche à adapter au signal (de type encoche, il supprime le contenu du signal dans une bande finie, ici centrée autour de 0). L'adaptation porte ici uniquement sur la position centrale du filtre. (C): le filtre est positionnée hors de l'écho Doppler. Le spectre après filtrage est représenté par l'aire verte qui se différencie peu du spectre initial car seule une fraction de bruit est supprimée. (D): cette fois-ci, le filtre est positionné sur l'écho Doppler, le spectre du signal filtré n'est quasiment plus que du bruit, une partie importante de la puissance a été filtrée.

# Filtrage adapté

- Définition du filtre

En général, on part d'un filtre que l'on définit par sa fonction de réponse spectrale. Exemple:

$$\psi(f) = \frac{1}{1 + \frac{S_0}{\sqrt{2\pi}\sigma} \exp\left(-\frac{f^2}{2\sigma^2}\right)}$$

Puis on « adapte » ce fil en jouant le plus souvent uniquement sur sa fréquence centrale (on considère que les éventuels autres paramètres sont de moindre importance). L'adaptation consiste alors à minimiser la fonction

$$J(\hat{f}) = \frac{1}{M} \sum_{k=0}^{M-1} \hat{\Phi}(f_k + \hat{f}) \Psi(f_k) \quad \text{où} \quad \hat{\Phi}(f) = \frac{1}{N} \sum_{n=0}^{N-1} \left| \sum_{m=0}^{M-1} s_n(mT_s) \exp(-2j\pi f m T_s) \right|^2 \quad \text{et} \quad f_k = \frac{kF_s}{M}$$

- Adaptation

L'optimisation de la fonction  $J(f)$  est plus efficace si on calcule sa valeur de la manière suivante

Avec 
$$J(\hat{f}) = -d_0 + 2\text{Re} \left\{ \sum_{m=0}^{M-1} d_m \exp(-2j\pi \hat{f} m T_s) \right\}$$
 ← Se calcule pour de nombreuses valeurs de  $f$  avec une TF

Et 
$$d_m = \frac{D_m}{N} \sum_{n=0}^{N-1} \sum_{m'=0}^{M-m-1} s_n^*(m' T_s) s((m+m') T_s)$$
 ← Auto-corrélation fenêtrée

$$D_m = \frac{1}{M} \sum_{k=0}^{M-1} \Psi(f_k) \exp\left(-2j\pi \frac{km}{M}\right)$$
 ← Ne dépend que du filtre

# Quelques exemples de filtre adaptable

- Pulse-Pair

$$\Psi_{pp}(f) = \frac{1}{2}(1 - \cos(2\pi f T_s))$$

Filtre « large », donc pas forcément bien adapté au signal lidar, généralement « étroit », mais d'une très grande simplicité de mise en œuvre et très rapide, car  $D_m = 0.5\delta(0) - 0.5\delta(1) - 0.5\delta(-1)$ , et donc

$$J_{pp}(\hat{f}) = -d_0 + 2\text{Re}\{d_1 \exp(-2j\pi\hat{f}T_s)\}$$

D'où

$$\hat{f}_{pp} = \frac{1}{2\pi T_s} \arg(d_1) = \frac{1}{2\pi T_s} \arg\left[\frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=0}^{M-2} s_n^*(kT_s) s((k+1)T_s)\right]$$

Estimation de l'auto-corrélation

- Levin

$$\Psi_{LEV}(f) = \frac{1}{1 + \frac{S_0}{\sqrt{2\pi}\sigma} \exp\left(-\frac{f^2}{2\sigma^2}\right)}$$

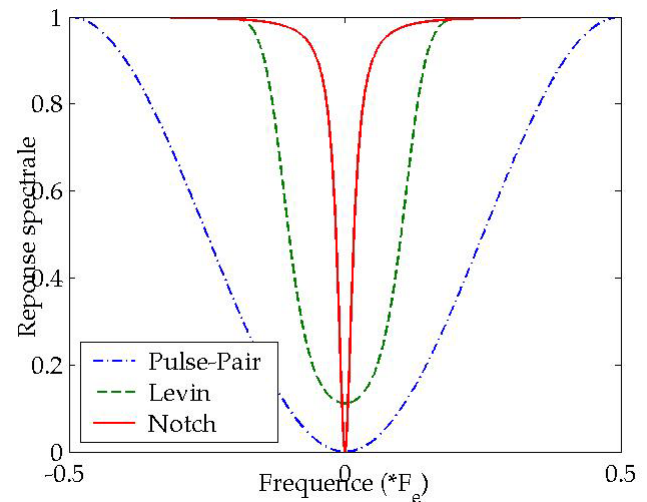
- Notch

$$\Psi_{notch}(f) = \frac{2 - 2\cos(2\pi f T_s)}{1 + r^2 - 2r \cos(2\pi f T_s)}$$

Filtre autorégressif

$$y(k) - ry(k-1) = x(k) - x(k-1)$$

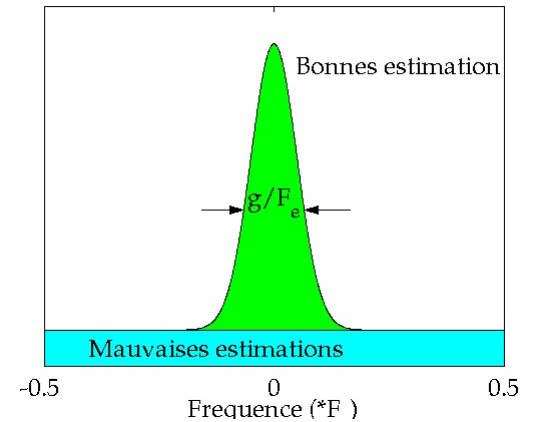
Peut être optimisé par des techniques de poursuite de fréquence utilisées en automatisme.



# Précision de la mesure

- Distribution des erreurs

$$\text{prob}(\hat{\nu}) = \frac{b-1}{F_e} + \frac{b}{\sqrt{2\pi}g} \exp\left(-\frac{(\hat{\nu}-\nu)^2}{2g^2}\right)$$



La distribution des erreurs combine un « pic » de bonnes estimations de largeur  $g$  sur une distribution uniforme de mesures « aberrantes » (l'estimateur a pris un pic spectral de bruit pour le signal).

A faible CNR, la fiabilité chute

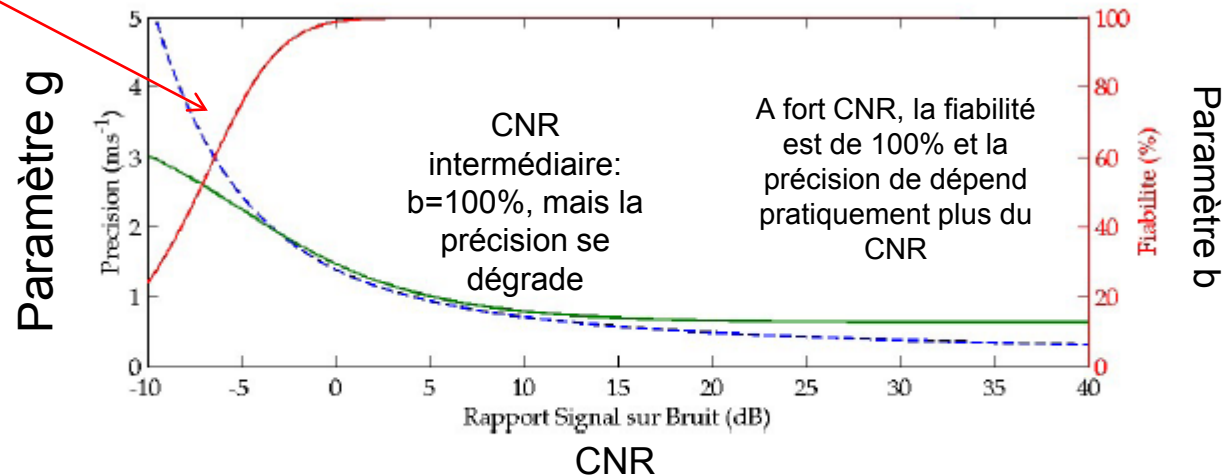


FIG. 3.5 – Précision (écart-type des erreurs des bonnes estimations) et fiabilité des estimations de fréquence en fonction du rapport signal sur bruit (en dB) pour l'algorithme de Levin appliqué à des signaux de type Zrnic de largeur spectrale  $w = 0.625$  MHz, et de longueur  $M = 32$ . Les tirets représentent la borne de Cramer-Rao.

# Précision de la mesure

- Ecart-type  $\sigma$

Ce paramètre varie quasiment plus à fort CNR. Aussi, quand on veut augmenter la précision d'une mesure à fort CNR, il ne sert à rien d'augmenter l'énergie des impulsions émises (le CNR augmentera bien mais ça n'aura pratiquement pas d'impact sur la précision), il vaut mieux augmenter le nombre de signaux lidars « accumulés » (donc privilégier la cadence du laser pour avoir un plus grand nombre de tirs dans une durée de temps fixée).

- Fiabilité

La fiabilité se dégrade rapidement dès qu'on passe en-dessous d'un seuil. Celui-ci peut-être très bas (exemple: -22dB pour le lidar Léosphere...).

Il existe un modèle permettant d'anticiper le taux de fiabilité en fonction des caractéristiques de la mesure:

A. Dabas. Semi-empirical model for the reliability of a matched filter frequency estimator for Doppler lidar. *Journal of Atmospheric and Oceanic Technology*, **16** :19–28, 1999.

Voir aussi les études publiées dans

G. Frehich and M. J. Yadlowsky: Performance of Mean-Frequency Estimators for Doppler Radar and Lidar. *Journal of Atmospheric and Oceanic Technology*, **11** :1217, 1994.