

Robustesse minimax et détection

Minimax robustness and detection



Michel BOUVET

Groupe d'Études et de Recherches en Détection Sous-Marine, DCAN Toulon, DCN, 83000 TOULON.

Ancien élève de l'École Polytechnique et Docteur d'État, l'auteur est Ingénieur de l'Armement au Groupe d'Études et de Recherches en Détection Sous-Marine, Centre d'Étude de la Direction des Constructions et Armes Navales de Toulon (DCN). Il a fait un séjour de 4 ans au Laboratoire des Signaux et Systèmes et de 1 an à l'Université de Princeton. Ses domaines d'intérêt sont le traitement du signal, la théorie de la détection, l'analyse spectrale, le traitement d'antenne et les systèmes sonar.

RÉSUMÉ

L'approche robuste en traitement du signal peut être utilisée lorsqu'il existe des incertitudes sur l'observation, bruit non gaussien par exemple. Dans un contexte de détection d'un signal dans une observation bruitée, les notions élémentaires de théorie des jeux nécessaires pour traiter un problème de robustesse minimax sont présentées. Des applications sont étudiées en utilisant différents critères, courbes COR et contraste. On montre en particulier que tous les récepteurs robustes obtenus possèdent dans leur structure un filtre adapté, récepteur optimal dans le cas gaussien.

MOTS CLÉS

Détection, robustesse minimax, récepteur optimal, filtre adapté, bruit non gaussien.

SUMMARY

In signal processing, the robust approach might be used when there are uncertainties on the observation, in a non-Gaussian environment for example. In the case of signal detection problems, the elementary results of game theory that are necessary to deal with a minimax robust approach are presented. Some applications are studied, using different detection criteria, ROC curves or contrast-like criteria. In particular, it is shown that all the robust receivers that are obtained, have in their structure a matched filter, optimal receiver in the Gaussian case.

KEY WORDS

Detection, minimax robustness, optimal receiver, matched filter, non-Gaussian environment.

Notations principales

x , scalaire;
 $p(x)$, densité de probabilité de x scalaire;
 \mathbf{x} , vecteur;
 $P(\mathbf{x})$, densité de probabilité de \mathbf{x} vecteur
 $[P(\mathbf{x}) = \prod_{i=1}^n p(x_i) \text{ dans le cas i. i. d.}]$;

$L(\mathbf{x})$, rapport de vraisemblance;
 ∇ , opérateur gradient;
 \mathbf{s} , signal à détecter;
 Γ , matrice de corrélation du bruit;
 \mathbf{h} , réponse impulsionnelle d'un filtre linéaire;
 ρ , critère de performance;
 $f \in F$, filtre général;
 $m \in M$, point d'opération;
 ε , paramètre de contamination.

I. Notion de robustesse

Parmi tous les développements récents en traitement du signal, la notion de robustesse est peut-être celle qui a mis le plus de temps à s'exprimer. L'idée est relativement simple et naturelle. Comme les propriétés statistiques de l'observation ne sont pas parfaitement connues, les traitements doivent être tels qu'un « petit changement » de ces propriétés n'ait qu'un faible effet sur les performances. Par « petit changement », il faut comprendre soit une faible variation sur un grand nombre d'échantillons de l'observation, soit une forte variation sur un petit nombre. Le premier cas correspondrait à un changement (une incertitude) de 2 % de la variance supposée du bruit, le second à la présence de bruit impulsif, par exemple.

La notion de résistance (résistance) est due à Tukey [1] qui avait précédemment introduit l'idée de robustesse pour un critère particulier de type rapport signal sur bruit (efficiency-robustness) [2]. La résistance correspondait à la version de la robustesse plus orientée vers les données et les estimateurs. La robustesse efficace utilisait l'efficacité relative comme critère de mesure des performances.

La robustesse qualitative (qualitative robustness), introduite par Hampel [3], utilise des propriétés de continuité pour étudier les variations dues à un changement des hypothèses de base. Elle correspond à la contrepartie probabilité de la résistance.

La notion de robustesse a été difficilement traduisible en termes mathématiques et il a fallu attendre le développement de la théorie des jeux pour aboutir à une formation utile en traitement du signal. P. J. Huber introduisit la robustesse minimax (minimax robustness) en 1964 [4] et depuis, beaucoup de travaux ont suivi, aussi bien sous l'aspect traitement du signal qu'en automatique. Une synthèse a récemment été publiée par S. A. Kassam et H. V. Poor [5] en anglais et par V. M. Krasnenker en russe [6]. Le terme « robuste » dans ce contexte remonte à G. E. Box en 1953.

Cet article se fixe comme but de donner rapidement les quelques bases mathématiques de théorie des jeux nécessaires à la bonne compréhension de l'approche minimax robuste et d'étudier quelques cas de récepteurs robustes pour des problèmes de détection en temps discret lorsqu'il existe des incertitudes sur le signal que l'on reçoit et sur la structure statistique du bruit perturbateur. Des résultats originaux sont toutefois présentés au paragraphe V.

II. Robustesse et théorie des jeux

La théorie des jeux appliquée aux tests d'hypothèses remonte aux travaux de Wald [7], repris, toujours dans un cadre statistique, par Ferguson [8] et Huber [9]. Pour un problème de décision binaire ou de détec-

tion d'un signal avec un récepteur à seuil, trois êtres mathématiques sont nécessaires :

- un *critère*, mesure quantitative de performance;
- un *point d'opération*, situation externe donnée par l'environnement, la *nature* (ou l'adversaire);
- un *filtre*, structure du récepteur donnée par l'ingénieur, le *joueur*.

Le joueur cherche à maximiser le critère alors que la nature tend à le minimiser. Il n'y a pas de transfert d'information entre les deux.

Illustrons dès maintenant ce cadre théorique par deux exemples concrets.

La cotation des actions boursières a parfois été considérée comme une situation de « jeu » entre deux joueurs, le vendeur et l'acheteur. Le prix de l'action correspond au critère. Le vendeur, que ce soit un particulier ou l'entreprise elle-même, cherche à augmenter le prix de vente. Le client acheteur (l'investisseur) essaie au contraire d'acheter ses actions au prix le plus bas. Le taux réel de la cotation se fixe ainsi chaque jour à une certaine valeur qui correspond à un compromis entre le prix de vente souhaité par le vendeur et le prix d'achat que peut se permettre l'investisseur.

Le contexte de lutte contre des brouilleurs peut aussi s'étudier à l'aide de la théorie des jeux. Soit un ingénieur (le joueur) qui veut maximiser le rapport signal sur bruit (le critère) en sortie d'un récepteur en présence de brouilleurs (l'adversaire). Ces brouilleurs peuvent être intentionnels (ennemi) ou donnés par l'environnement (la nature). Dans le premier cas, le but de l'adversaire est la minimisation du critère, c'est-à-dire la dégradation du rapport signal sur bruit. Dans le second cas, on peut considérer que la nature est neutre. Si son action est connue, l'ingénieur pourra dimensionner son récepteur de manière optimale pour la situation correspondante (le point d'opération). Si l'effet de la nature est inconnue, on peut utiliser la notion de robustesse.

Le point d'opération n'étant pas connu exactement, l'approche robuste consiste à chercher un système, le filtre, assurant une *limite inférieure optimale des performances*, c'est-à-dire un système robuste minimax. On se place dans le pire des cas.

Soit F l'ensemble des filtres possibles, $f \in F$, et M l'ensemble des points d'opérations possibles, $m \in M$. Étant donné le critère ρ ,

$$\rho: F \times M \rightarrow R$$

le triplet (F, M, ρ) constitue un jeu.

On définit le *filtre optimal* pour la situation $m \in M$ comme le filtre $f^*(m) \in F$ tel que

$$(2.1) \quad \rho(f^*(m), m) = \underset{f \in F}{\text{Max}} \rho(f, m).$$

Le critère est alors noté

$$(2.2) \quad \rho^*(m) = \underset{\Delta}{\rho}(f^*(m), m).$$

Pour $f \in F$, la *pire situation* est définie comme $m_p(f) \in M$ vérifiant

$$(2.3) \quad \rho(f, m_p(f)) = \underset{m \in M}{\text{Min}} \rho(f, m).$$

Le but du problème robuste est de trouver le *filtre robuste minimax* $f_r \in F$ tel que

$$(2.4) \quad f_r \in \underset{f \in F}{\text{Arg Max}} \underset{m \in M}{\text{Min}} \rho(f, m),$$

c'est-à-dire réalisant le maximum sur $f \in F$ de $\rho(f, m_p(f))$. En effet un tel filtre assure

$$(2.5) \quad \rho(f_r, m_p(f_r)) \geq \rho(f, m_p(f)), \quad \forall f \in F,$$

c'est-à-dire la meilleure performance pour la pire situation. Ce filtre est surtout intéressant car il permet d'assurer un niveau minimal de performance quel que soit le point d'opération $m \in M$ choisi par la nature. Ceci se lit selon

$$(2.6) \quad \rho(f_r, m_p(f_r)) \leq \rho(f_r, m), \quad \forall m \in M,$$

inégalité obtenue en remplaçant f par f_r dans

$$(2.7) \quad \underset{\mu \in M}{\text{min}} \rho(f, \mu) \leq \rho(f, m), \quad \forall f \in F.$$

Malheureusement, cette approche ne présente que peu d'intérêt, peu d'interprétation existant pour m_p . Cet inconvénient est susceptible d'être levé en supposant l'existence d'un point selle pour le jeu précédent. Un *point selle* est défini comme un couple (f_s, m_s) tel que

$$(2.8) \quad \begin{aligned} f_s &= f^*(m_s); \\ m_s &= m_p(f_s). \end{aligned}$$

On a alors

$$(2.9) \quad \underset{m \in M}{\text{Min}} \underset{f \in F}{\text{Max}} \rho(f, m) = \underset{f \in F}{\text{Max}} \underset{m \in M}{\text{Min}} \rho(f, m).$$

Si (f_s, m_s) est un point selle, alors f_s est le filtre robuste cherché. L'inverse est faux en général. Toutefois, sous certaines conditions de régularité et convexité qui seront précisées plus loin par un théorème, le filtre robuste peut être obtenu non pas comme le «max du min» mais comme le «min du max», c'est-à-dire selon

$$(2.10) \quad \underset{m \in M}{\text{Min}} \underset{f \in F}{\text{Max}} \rho(f, m) = \underset{m \in M}{\text{Min}} \rho(f^*(m), m).$$

L'interprétation est ici plus pratique. On cherche d'abord le filtre optimal, $f^*(m)$, pour un point d'opération donné, avant de chercher le point d'opération le plus défavorable. Le principal avantage réside dans la connaissance préalable pour la recherche d'un récepteur optimal. C'est pourquoi il est souvent plus facile de résoudre le problème de minimax et de prouver l'existence d'un point selle que de résoudre le problème de maximin.

Une représentation simple d'une situation classique de minimax est celle d'une selle ou d'un col de montagne. Le point le plus haut du chemin passant d'une vallée dans l'autre correspond au minimax et au maximin. Une telle situation est représentée sur la figure 1. La ligne de crête correspond au lieu des maximums alors que le talweg est associé au lieu des minimums. Le point selle existe et se trouve à l'intérieur du domaine auquel on s'intéresse. Il est visible sur cette figure que quel que soit le point d'opération m donné par la «nature», la valeur du critère sera supérieure à $\rho(f_r, m_p(f_r))$ selon (2.6). De plus, cette valeur robuste du critère est la valeur maximale qui ait une telle propriété. C'est en cela qu'il est possible de dire que le filtre robuste assure une «limite inférieure optimale des performances».

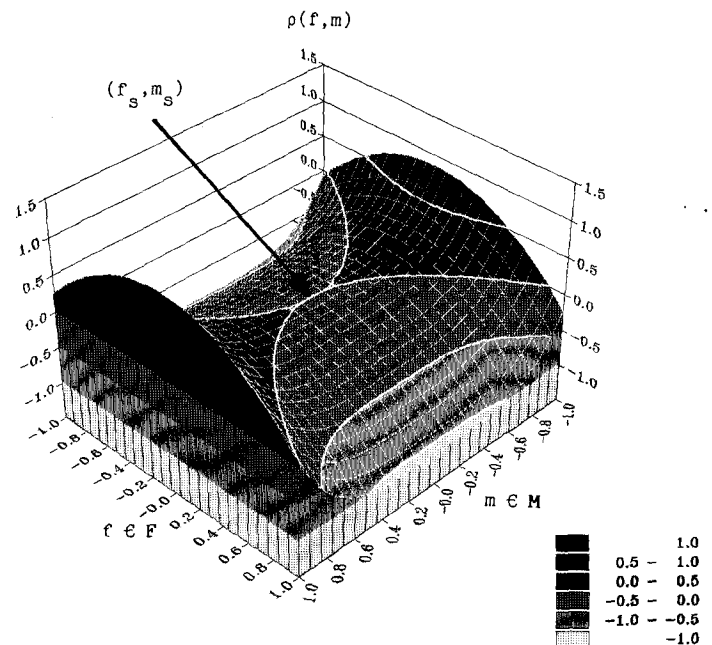


Fig. 1. — Point-selle à l'intérieur.

Il existe des cas moins intéressants où, par exemple, le lieu des maximums est situé sur une frontière du domaine d'intérêt, pour $f=1$ sur la figure 2.

Avant d'étudier quelques cas particuliers en détection, il est bon de rappeler un théorème donnant des conditions suffisantes d'identité entre approches maximin et minimaux [10].

Théorème: Si le jeu (F, M, ρ) est tel que :

- (i) M est un ensemble convexe;
- (ii) $\rho(f, \cdot)$ est convexe sur M pour tout f de F ,
et si (f_p, m_p) est un point régulier pour le jeu, alors les propriétés suivantes sont équivalentes :
 - (a) m_p est la pire situation pour (F, M, ρ) ;

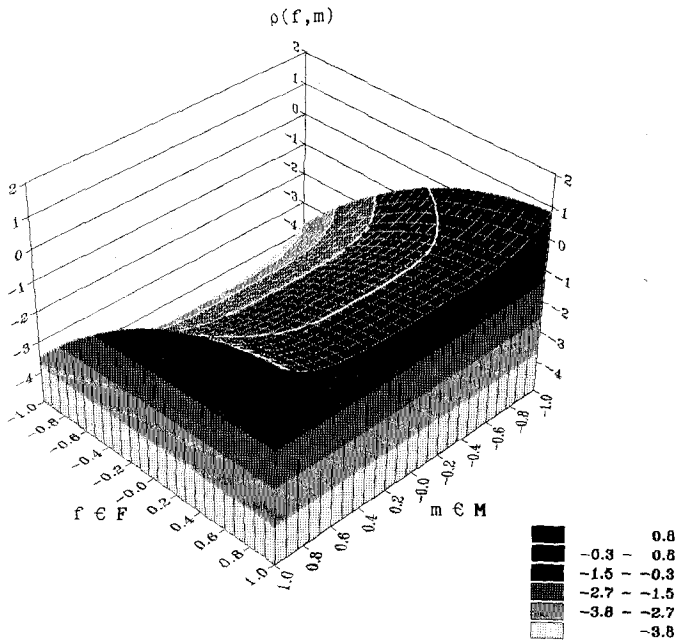


Fig. 2. — Point-selle à l'extérieur.

(b) (f_p, m_p) est un point selle solution pour (F, M, ρ) .

Un point régulier pour le jeu est un point (f_2, m_2) tel que, quelque soit $m \in M$ tel que

$$m_\alpha = (1-\alpha)m_2 + \alpha m \in M \quad \text{pour } 0 \leq \alpha \leq 1,$$

on a $\rho^*(m_\alpha) - \rho(f_2, m_\alpha) = O(\alpha)$, c'est-à-dire d'un ordre supérieur ou égal à α . Tous les points qui seront considérés seront réguliers de par la structure des critères utilisés. Ce point, simple à vérifier, ne sera pas démontré.

Dans la suite nous utiliserons, entre autre, la densité de probabilité (ddp) du bruit comme point d'opération de façon à modéliser l'incertitude du problème par l'incertitude sur la structure du bruit, représentant la « nature » qui est inconnue. Selon les cas, cette ddp, P, sera supposée appartenir à l'un des trois ensembles suivants

$$(2.11) \quad \Pi_\varepsilon = \{ P \in A; P = (1-\varepsilon)Q + \varepsilon C, C \in A \};$$

$$(2.12) \quad \Pi_1 = \{ P \in A; E \{ \mathbf{x} \mathbf{x}^t \} = I \};$$

$$(2.13) \quad \Pi_\nu = \{ P \in A; E \{ \mathbf{x} \mathbf{x}^t \} = \sigma I, \\ 0 < \sigma_1 \leq \sigma \leq \sigma_2 < +\infty \}.$$

A est l'ensemble des ddp symétriques.

Π_ε est un voisinage de ε contamination [9]. En général, ε est supposé fixé, Q est la ddp nominale, connue et C correspond au terme de contamination. Cette notion de ddp de mixture (parfois appelée mélange) sera détaillée plus loin.

Π_1 est l'ensemble des ddp de matrice de covariance identité (ou « de variance unité »).

Π_ν correspond à des bruits blancs de puissance inconnue, comprise entre σ_1^2 et σ_2^2 .

III. Filtres adaptés robustes

La première situation particulière étudiée correspond au cas où le critère ρ est le rapport signal sur bruit en sortie du filtre f et où celui-ci est linéaire et noté \mathbf{h} [10, 11]. Le point d'opération m modélise l'incertitude sur les caractéristiques de l'observation (signal, bruit) et peut être noté

$$(3.1) \quad m = (\mathbf{s}, \Gamma).$$

Cette incertitude porte donc sur le signal \mathbf{s} et la matrice d'autocorrélation du bruit, Γ , le bruit étant supposé additif et de loi centrée.

III.1. POSITION DU PROBLÈME

La valeur du critère en $(\mathbf{h}; \mathbf{s}, \Gamma)$ est

$$(3.2) \quad \rho(\mathbf{h}; \mathbf{s}, \Gamma) = \frac{(\mathbf{h}' \mathbf{s})^2}{\mathbf{h}' \Gamma \mathbf{h}}.$$

Le problème robuste consiste à chercher le filtre réalisant

$$(3.3) \quad \text{Max}_{\mathbf{h} \in \mathbf{H}} \text{Min}_{(\mathbf{s}, \Gamma) \in \mathbf{O}} \rho(\mathbf{h}; \mathbf{s}, \Gamma)$$

où plutôt, sous certaines conditions,

$$(3.4) \quad \text{Min}_{(\mathbf{s}, \Gamma) \in \mathbf{O}} \text{Max}_{\mathbf{h} \in \mathbf{H}} \rho(\mathbf{h}; \mathbf{s}, \Gamma)$$

Ce second problème se ramène donc à chercher le filtre adapté optimal pour un signal et un bruit appartenant à un ensemble de variations \mathbf{O} de l'observation, le bruit n'étant défini que par son second ordre. La seule incertitude sur celui-ci porte sur l'amplitude et la structure de sa matrice d'autocorrélation. L'ensemble \mathbf{O} n'a pas besoin d'être défini précisément tout de suite.

D'après le paragraphe précédent, l'existence d'un point selle, permettant de résoudre le problème robuste (3.3) sous la forme (3.4), sera assurée si le critère est convexe vis-à-vis du couple (\mathbf{s}, Γ) et si \mathbf{H} est un ensemble convexe. Il est aisé de voir que $\rho(\cdot; \cdot)$ est convexe en \mathbf{s} comme $x \rightarrow x^2$ et en Γ comme $x \rightarrow 1/x$ pour $x > 0$. La convexité de \mathbf{O} devra être assurée dans les hypothèses faites pour la modélisation de l'observation; par exemple en supposant que $\|\mathbf{s}\|^2 \leq s^2$ et $\Gamma = \sigma \Gamma_0$, $\sigma_0 \leq \sigma \leq \sigma_1$ avec Γ_0 , $\sigma_0 > 0$ et σ_1 fixés. D'autre part, $\rho(\cdot; \cdot)$ est bornée inférieurement par 0 (Γ est définie positive) et supérieurement par $\rho^*(\mathbf{s}, \Gamma)$, valeur obtenue pour le filtre optimal.

$$(3.5) \quad \mathbf{h}^* = \Gamma^{-1} \mathbf{s},$$

ce qui se montre par l'application sur (3.2) de l'inégalité de Schwarz [11]. Le critère a alors la valeur

$$(3.6) \quad \rho^*(\mathbf{s}, \Gamma) = \mathbf{s}' \Gamma^{-1} \mathbf{s}.$$

Pour minimiser cette expression par rapport à la situation d'opération, deux cas sont à envisager.

III. 2. INCERTITUDES SUR LE BRUIT

Le signal est supposé fixé et connu. Le domaine \mathbf{O} de variations se réduit à un ensemble convexe \mathbf{O}_b de matrices définies positives. Ce cas correspond à une méconnaissance de la puissance et de la structure de coloration du bruit. Le filtre robuste \mathbf{h}_r correspond à la matrice de corrélation du bruit minimisant (3.6) pour $\Gamma \in \mathbf{O}_b$. Ce filtre robuste correspond donc à un cas particulier de bruit, celui associé à Γ minimisant (3.6).

Un exemple est celui d'un bruit blanc de puissance inconnue,

$$(3.7) \quad \mathbf{P} \in \Pi_v$$

Le bruit est blanc mais de variance σ^2 mal connue, que l'on suppose toutefois comprise entre deux bornes, $\sigma_1^2 \leq \sigma^2 \leq \sigma_2^2$.

L'ensemble Π_v étant convexe, le filtre robuste est celui minimisant la valeur maximale du critère, (3.6). Cette valeur est ici $\|\mathbf{s}\|^2/\sigma^2$. Le filtre robuste s'obtient donc pour la valeur maximale de la variance.

$$(3.8) \quad \mathbf{h}_r = \sigma_1 \mathbf{s}$$

obtenue pour $\Gamma = \sigma_1^2 \mathbf{I}$. On se place donc dans le pire des cas, celui où la puissance du bruit est maximale.

Notons que cette modélisation semble raisonnable car il est naturel d'imposer sur le bruit la contrainte

$$(3.9) \quad 0 < \varepsilon \leq \det \Gamma \leq A < +\infty.$$

La borne inférieure permet d'assurer que la détection est non singulière. La borne supérieure assure que le critère ne s'annule pas sur Π_v , ce qui pose des problèmes évoqués dans la suite.

III. 3. INCERTITUDES SUR LE SIGNAL

Le bruit est maintenant de structure connue, supposé blanc. Le domaine de variations \mathbf{O}_s est supposé être la sphère de centre \mathbf{s}_0 et de rayon r , au sens de la norme $L^2(\mathbf{R}^n)$,

$$(3.10) \quad \mathbf{O}_s = \{\mathbf{s} \in \mathbf{R}^n; \|\mathbf{s} - \mathbf{s}_0\| \leq r\}.$$

Cet ensemble est borné, convexe. La valeur optimale du critère étant $\mathbf{s}'\mathbf{s}$, la situation robuste correspondra au signal \mathbf{s} de norme minimale. Deux cas apparaissent.

— $\mathbf{0} \in \mathbf{O}_s$.

Ceci se produit si $\|\mathbf{s}_0\| \leq r$. Ce cas met en évidence une limitation de l'approche robuste car la valeur minimale du critère est 0, valeur nulle atteinte pour $\mathbf{s} = \mathbf{0}$. Le «filtre» correspondant est le filtre nul, qui n'est pas d'un grand intérêt pratique!

— $\mathbf{0} \notin \mathbf{O}_s$.

C'est le cas où $\|\mathbf{s}_0\| > r$. Le signal \mathbf{s}_r réalisant le minimum du critère est celui de norme minimale,

$$(3.11) \quad \mathbf{s}_r = \left(1 - \frac{r}{\|\mathbf{s}_0\|}\right) \mathbf{s}_0.$$

Le filtre robuste est associé à ce signal,

$$(3.12) \quad \mathbf{h}_r = \left(1 - \frac{r}{\|\mathbf{s}_0\|}\right) \mathbf{s}_0.$$

La valeur minimale du critère est alors

$$(3.13) \quad \rho_r(\mathbf{h}_r; \mathbf{s}_r) = \left(1 - \frac{r}{\|\mathbf{s}_0\|}\right)^2 \mathbf{s}_0' \mathbf{s}_0.$$

III. 4. DISCUSSION. LIMITATIONS

Tout d'abord, il faut remarquer que si le signal et le bruit peuvent présenter des incertitudes en même temps, il est logique de supposer que leurs fluctuations respectives sont indépendantes. Ceci correspond au cas où il n'y a pas de transfert d'information entre les deux joueurs (ou entre le joueur et la nature). L'ensemble \mathbf{O} peut alors se mettre sous la forme

$$\mathbf{O} = \mathbf{O}_b \times \mathbf{O}_s,$$

produit cartésien des deux ensembles, où l'on suppose $\mathbf{0} \in \mathbf{O}$. Le problème de minimisation est ainsi séparé en deux. Avec les notations précédentes et en utilisant \mathbf{O}_b défini par (3.7) et \mathbf{O}_s défini par (3.10), le filtre robuste est, de manière évidente,

$$(3.14) \quad \mathbf{h}_r = \sigma_1 \left(1 - \frac{r}{\|\mathbf{s}_0\|}\right) \mathbf{s}_0.$$

Ensuite, des hypothèses de convexité ont été faites (celles du théorème précédent). La convexité de l'ensemble \mathbf{O} est suffisante, mais n'est évidemment pas nécessaire comme précisé dans [8]. Par exemple, la figure 3 représente un cas où, l'incertitude portant sur le signal, l'ensemble \mathbf{O}_s n'est pas convexe mais où le signal \mathbf{s}_1 conduit à l'existence d'un point-selle, \mathbf{O}_s étant un sous-ensemble de \mathbf{R}^2 .

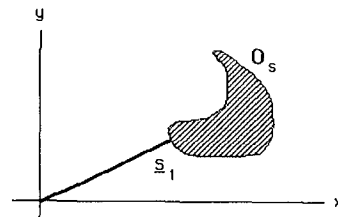


Fig. 3. — Exemple d'ensemble \mathbf{O}_s non convexe.

La convexité du critère n'est pas non plus indispensable, bien que pratique pour être sûr que le minimum trouvé est global et pas seulement local.

Enfin, une limitation de l'approche robuste vient du domaine de variations des situations possibles.

(i) Critère nul

S'il existe une situation annulant le critère optimal, le filtre robuste annulera la valeur du critère, ce qui n'est pas très intéressant.

Un tel cas a été présenté dans le cas où le signal pouvait être nul, ou en tout cas très faible. Un autre cas est le suivant : \mathbf{O} est un ensemble de couples de ddp, (P_0, P_1) correspondant aux statistiques de l'observation sous les deux hypothèses; \mathbf{O} se décomposant en un produit cartésien $\mathbf{O}_0 \times \mathbf{O}_1$, si il y a recouvrement dans un certain domaine, c'est à dire une situation où $P_0 = P_1$, celle-ci correspondra à un cas complètement dégradé : on est incapable de discriminer deux situations identiques. Alors, le critère optimal sera aussi nul.

(ii) *Difficulté de modéliser les incertitudes*

L'approche robuste permet de travailler avec des incertitudes sur la nature. Celles-ci ne doivent pas être trop importantes, comme dans l'approche non paramétrique, sous peine d'annuler le critère comme nous venons de le voir. Il est parfois difficile pour un problème réel de modéliser l'incertitude, le domaine de variations de la situation devant être ni trop petit, sous peine de ne pas représenter l'observation, ni trop grand, sous peine d'avoir les problèmes évoqués précédemment. Les deux prochains paragraphes présentent des exemples de cette modélisation de l'incertitude. De plus l'hypothèse de séparation de \mathbf{O} en un produit cartésien de deux ensembles indépendants n'est pas toujours vérifiée. Par exemple, le bruit (Γ) peut être corrélé avec le signal (\mathbf{s}), par exemple si il y a des interférences (communications), de la réverbération (sonar) ou du fouillis (radar).

IV. Récepteurs robustes au sens des courbes COR

La notion de filtre adapté n'est qu'un cas particulier, linéaire, de récepteur. Dans ce paragraphe, l'hypothèse de linéarité du récepteur ne sera pas faite *a priori*.

L'incertitude portera sur les statistiques de l'observation, $P_0(\cdot)$ sous H_0 et $P_1(\cdot)$ sous H_1 , le bruit étant supposé blanc, en général de variance connue, unité pour simplifier. Ces statistiques peuvent souvent se mettre sous la forme d'une loi de mixture comportant deux termes,

$$(4.1) \quad \begin{cases} P_i(\mathbf{x}) = (1 - \varepsilon_i) Q_i(\mathbf{x}) + \varepsilon_i C_i(\mathbf{x}), \\ 0 \leq \varepsilon_i \leq 1, \quad i = 0, 1. \end{cases}$$

Dans cette expression qui modélise les incertitudes par le terme $\varepsilon_i C_i(\mathbf{x})$, $Q_i(\mathbf{x})$ est une loi nominale ou « de référence ». Les ddp réelles sont ainsi supposées d'écart peu de $Q_i(\cdot)$. Les paramètres de contamination, ε_i , sont supposés connus et, en général, faibles devant 1. Le terme de contamination, $C_i(\cdot)$, est une ddp perturbatrice, supposée continûment dérivable. Cette modélisation conduit à Π_ε (2.11) et correspond à ce qu'Huber appelle un voisinage de ε -contamination [9]. Le lecteur intéressé par les modèles de mixture peut se reporter à [12].

Il faut noter que les paramètres de ce modèle, $\varepsilon_0, \varepsilon_1, Q_0(\cdot), Q_1(\cdot)$, ainsi que C_0 et C_1 , ensembles de variations de $C_0(\cdot)$ et $C_1(\cdot)$, doivent être choisis de telle façon qu'un recouvrement des ensembles de définition de $P_0(\cdot)$ et $P_1(\cdot)$ soit évité. Conformément à l'esprit de cet article, l'hypothèse H_0 sera celle « bruit seul » et H_1 , « signal plus bruit ». Par la suite, le signal sera supposé déterministe, ce qui revient à remplacer $P_1(\mathbf{x})$ par $P_0(\mathbf{x} - \mathbf{s})$.

De telles ddp sont d'un intérêt pratique évident. Elles correspondent à un bruit proche d'un bruit nominal, souvent gaussien où un terme de perturbation modélise l'écart à cette référence [12]. Par exemple, un bruit acoustique sous-marin peut correspondre à une loi de référence gaussienne et à une perturbation asymétrique (exponentielle) pour prendre en compte un moment empirique d'ordre 3 non nul (skewness) ou une perturbation à grande variance (laplacienne ou gaussienne) pour modéliser un bruit impulsif. Ce dernier cas est courant, produit par des phénomènes biologiques (crevettes claqueuses, dauphins) et naturels (craquements de la banquise) qui font que les queues de la ddp empirique du bruit sont plus grandes que celles associées à la loi gaussienne.

– *Critère de Neyman-Pearson*

Dans cette première partie, le critère de performance est le critère classique de Neyman-Pearson. On cherche la situation la moins favorable telle que le récepteur optimal associé (le rapport de vraisemblance) minimise la probabilité de détection β sous la contrainte que la probabilité de fausse alarme α ne dépasse pas un certain seuil α^* .

(i) *Faible contamination*

En bruit iid, et sous l'hypothèse de paramètres ε_0 et ε_1 faibles et connus, Martin et Schwartz [13] ont montré que la situation la moins favorable est donnée par les ddp univariées suivantes,

$$(4.2) \quad p_{0r}(x) = \begin{cases} (1 - \varepsilon_0) q_0(x) \\ \text{si } q_1(x)/q_0(x) < k_0, \\ (1 - \varepsilon_0) \frac{q_1(x)}{k_0} \\ \text{si } q_1(x)/q_0(x) \geq k_0; \end{cases}$$

$$(4.3) \quad p_{1r}(x) = \begin{cases} (1 - \varepsilon_1) q_1(x) \\ \text{si } q_1(x)/q_0(x) > k_1, \\ (1 - \varepsilon_1) \frac{q_1(x)}{k_1} \\ \text{si } q_1(x)/q_0(x) \leq k_1. \end{cases}$$

Les paramètres k_0 et k_1 sont des paramètres de normalisation assurant que $p_{0,r}(\cdot)$ et $p_{1,r}(\cdot)$ sont des ddp [13, 14]. Le test correspondant est

$$(4.4) \quad T(\mathbf{x}) = \prod_{i=1}^n t(x_i),$$

$$t(x) = \text{Max} \left\{ k_1, \min \left\{ k_0, \frac{q_1(x)}{q_0(x)} \right\} \right\}.$$

Ces résultats relativement généraux ont une application intéressante. Pour la détection d'un signal déterministe dans un bruit proche du Gaussien, $Q_0(\cdot)$ est la loi de Gauss centrée, de variance unité et $Q_1(\mathbf{x})$ vaut $Q_0(\mathbf{x}-s)$. La solution robuste au sens des courbes COR consiste en un corrélateur-écréteur. Cette structure se représente selon la figure 4. Par rapport au récepteur classique réalisant la simple corrélation, la présence de l'écréteur peut se justifier par le fait que les échantillons à grandes valeurs sont susceptibles d'avoir été générés par un bruit de type impulsif (provenant de la contamination) et que leur limitation s'impose.

Curieusement, il est possible de montrer que la recherche du récepteur robuste *localement optimal* conduit à un écréteur-corrélateur, c'est-à-dire à effectuer une permutation des deux principales opérations présentes sur la figure 4, corrélateur et écréteur [6]. Dans la mesure où cette notion de récepteur localement optimal sera utile par la suite, précisons-la.

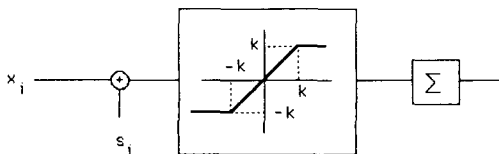


Fig. 4. — Corrélateur-écréteur.

(ii) Récepteur localement optimal

Le rapport de vraisemblance,

$$(4.5) \quad L(\mathbf{x}) = \frac{P_1(\mathbf{x})}{P_0(\mathbf{x})} = \frac{P_0(\mathbf{x}-s)}{P_0(\mathbf{x})},$$

est le récepteur optimal au sens du critère de Neyman-Pearson (courbes COR) ainsi que pour le critère de déflexion. Ceci sera précisé plus loin. Les problèmes les plus « intéressants » en détection se posent en particulier lorsque le signal est faible par rapport au bruit. On peut alors développer le rapport de vraisemblance (4.5) en puissances du signal (ou du rapport signal sur bruit), en utilisant un développement de Taylor, et ne conserver que le terme linéaire en fonction du signal [15].

Ce terme,

$$(4.6) \quad T_{l_0}(\mathbf{x}) = -\frac{s^t \nabla P_0(\mathbf{x})}{P_0(\mathbf{x})}$$

dépend de la dérivée logarithmique de la ddp sous H_0 . L'opérateur ∇ est le gradient par rapport à \mathbf{x} . Celui-ci, au signe près, est une non-linéarité précédant la corrélation avec le signal, comme indiqué sur la figure 5.

Le récepteur (4.6) est appelé « localement optimal », c'est-à-dire « optimal pour la déflexion locale » [15]. Le terme « localement » correspond à la région des faibles signaux.

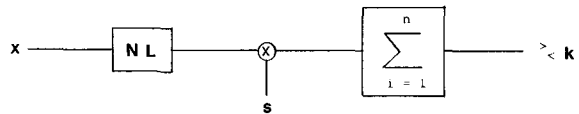


Fig. 5. — Récepteur localement optimal.

(iii) Cas général

Le développement ayant conduit à (4.4) a été effectué sous l'hypothèse d'un faible ε . On peut obtenir le récepteur robuste, toujours pour le critère de Neyman-Pearson et dans Π_c , pour la détection d'un signal déterministe dans un bruit de mixture sans cette hypothèse. La formulation du problème est plus restrictive mais le résultat est plus général.

Théorème [16] : Si :

- (i) le bruit est stationnaire et blanc au sens fort;
 - (ii) $p_0(x) = (1-\varepsilon)q(x) + \varepsilon c(x)$, $\varepsilon \in [0,1[$ connu;
 - (iii) $p_1(x_i) = p_0(x_i - s_i)$;
 - (iv) $q(\cdot)$ est symétrique, centrée et continûment dérivable;
 - (v) $q(\cdot)$ est fortement unimodale [c'est-à-dire $-\log q(x)$ est convexe];
 - (vi) $c(\cdot)$ est bornée et symétrique;
- alors le récepteur localement optimal robuste au sens des courbes COR existe et est donné par

$$(4.7a) \quad T(\mathbf{x}) = \sum_{i=1}^n l_r(x_i) s_i,$$

$$(4.7b) \quad l_r(x) = -\frac{\dot{f}_r(x)}{f_r(x)},$$

$$(4.7c) \quad f_r(x) = \begin{cases} (1-\varepsilon)q(x) & \text{si } |x| < a; \\ (1-\varepsilon)q(a)e^{-b(|x|-a)} & \text{si } |x| \geq a. \end{cases}$$

Dans (4.7b), $\dot{f}_r(x)$ désigne la dérivée de f_r par rapport à x , et dans (4.7c), les valeurs a et b sont telles que $f_r(\cdot)$ soit une ddp,

$$(4.8) \quad \int_{-a}^a q(x) dx + 2\frac{q(a)}{b} = (1-\varepsilon)^{-1},$$

et continûment dérivable,

$$(4.9) \quad b = -\dot{q}(a)/q(a).$$

La non-linéarité l_r , (4.7b), correspond à la dérivée logarithmique de f_r et provient de la même idée que

celle ayant conduit à (4.6), c'est à dire de la limitation du développement en série de Taylor du rapport de vraisemblance associé à f_r .

La probabilité de fausse alarme, α , est donnée par

$$(4.10) \quad \alpha \leq \alpha^* = 1 - \operatorname{erf} \left(\frac{\ln r_m^2}{r_m^2 - 1} \right)$$

où

$$(4.11) \quad r_m^2 = 1 + \frac{1}{2} \frac{\varepsilon}{1 - \varepsilon} \times \frac{b^2}{\int_0^a q(x) l_r^2(x) dx + b^2 \int_a^\infty q(x) dx}$$

Le récepteur robuste localement optimal au sens des courbes COR dans un bruit de mixture correspond au récepteur localement optimal adapté au bruit de référence, suivi d'un limiteur. La philosophie de cette limitation correspond au cas d'un bruit perturbateur impulsif. Ce bruit, à grande variance, entraîne beaucoup de fausses alarmes à cause des valeurs importantes de ses échantillons. Le « meilleur » moyen de s'affranchir de ces « bouffées » de fausse alarme lorsque le bruit devient très impulsif est de ne pas prendre en compte les échantillons trop grands. En effet, ceux-ci sont suspects de n'être que des valeurs du bruit. Avec une proportion de ε , le bruit peut être considéré comme généré par $C(x)$ et on vérifie que plus ε est petit, plus le seuil de limitation a est grand.

— Application au bruit de Gauss généralisé

Si $q(x)$ est la loi de Gauss généralisée,

$$(4.12) \quad q(x) = \frac{c}{2 A(c) \Gamma(1/c)} \exp \left(- \left(\frac{|x|}{A(c)} \right)^c \right),$$

avec

$$A(c) = \left(\Gamma \left(\frac{1}{c} \right) / \Gamma \left(\frac{3}{c} \right) \right)^{1/2},$$

$\Gamma(\cdot)$ étant la fonction Gamma, la non-linéarité robuste, devient

$$(4.13) \quad l_r(x) = \begin{cases} \frac{c}{A(c)} \left(\frac{|x|}{A(c)} \right)^{c-1} \operatorname{sgn}(x) & \text{si } |x| \leq a \\ \frac{ca^{c-1}}{A(c)} \operatorname{sgn}(x) & \text{si } |x| > a \end{cases}$$

où

$$(4.14) \quad \int_{-a}^a q(x) dx + \left[\frac{A(c)}{a} \right]^{c-1} \times \left(\Gamma \left(\frac{1}{c} \right) \right)^{c-1} e^{-(a/A(c))^c} = \frac{1}{1 - \varepsilon}.$$

On vérifie que pour $c=2$, la non-linéarité a bien la même forme que celle obtenue pour la loi de Gauss (voir fig. 4).

— Robustesse du détecteur du signe

Si on suppose que la ddp de mixture est telle que la partie la plus importante, $q(x)$, soit une loi de Laplace,

$$(4.15) \quad q(x) = \frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}|x|/\sigma},$$

et que la perturbation, $c(x)$, soit continue à l'origine et de médiane 0 (ou symétrique, ce qui est plus restrictif), il est alors possible de montrer que le détecteur du signe est le récepteur robuste au sens des courbes COR pour la détection d'un signal constant [17].

V. Récepteurs robustes au sens du contraste

Dans cette partie où des résultats originaux sont présentés, seront utilisés deux critères de comparaison dérivés du contraste. Le critère de contraste est défini de manière très générale pour un récepteur $T(\cdot)$ comme [18]

$$(5.1) \quad C = \frac{[E\{T(\mathbf{x})/H_1\} - E\{T(\mathbf{x})/H_0\}]^2}{\operatorname{Var}_\pi(T(\mathbf{x}))}$$

où le dénominateur représente la variance de $T(\mathbf{x})$ calculée à partir de la ddp suivante

$$(5.2) \quad P_\pi = (1 - \pi) P_0 + \pi P_1,$$

π étant un paramètre compris entre 0 et 1 et $P_i, i=0,1$, étant la ddp de l'observation sous l'hypothèse i . Le contraste est un critère de type rapport signal sur bruit. Le numérateur correspond à une mesure de l'augmentation, de la valeur de sortie du récepteur, due au signal. Le dénominateur est associé à une mesure de la puissance du bruit lorsque les deux hypothèses sont prises en compte.

Lorsque π est égal à 0, on obtient le critère de déflexion [18],

$$(5.3) \quad D = \frac{[E\{T(\mathbf{x})/H_1\} - E\{T(\mathbf{x})/H_0\}]^2}{\operatorname{Var}\{T(\mathbf{x})/H_0\}}.$$

La déflexion est donc obtenue en n'utilisant que la variance du bruit au dénominateur. C'est un critère classique mesurant le rapport signal sur bruit en sortie du récepteur.

La même approche locale que précédemment pour le rapport de vraisemblance est envisageable pour la déflexion. Si on développe D en fonction du signal, supposé faible par rapport à la variance du bruit, dans un problème de détection, le terme linéaire de ce développement est proportionnel à la déflexion locale, ou efficacité, définie comme [15]

$$(5.4) \quad d = \frac{1}{\text{Var}(T(\mathbf{x})/H_0)} \times \left[\int_{\mathbf{R}^n} \mathbf{s}' \nabla(P(\mathbf{x})) T(\mathbf{x}) d\mathbf{x} \right]^2.$$

Dans cette expression, \mathbf{s} est le signal déterministe qu'on cherche à détecter et $P(\cdot)$ représente la loi du bruit.

Il faut remarquer que pour les deux premiers critères, déflexion et déflexion locale, une structure de récepteur est invariante par homothétie et translation. Une fonctionnelle affine d'un récepteur possède les mêmes critères de contraste.

V. 1. DÉFLEXION

En explicitant la dépendance de la déflexion en $T(\cdot)$ et en $P(\cdot)$, elle s'exprime selon

$$(5.5) \quad D(T, P) = \frac{\left\{ \int_{\mathbf{R}^n} T(\mathbf{x}) (P(\mathbf{x}-\mathbf{s}) - P(\mathbf{x})) d\mathbf{x} \right\}^2}{\int_{\mathbf{R}^n} [T(\mathbf{x}) - E_0\{T(\mathbf{x})\}]^2 P(\mathbf{x}) d\mathbf{x}}.$$

La maximisation de la déflexion par rapport à la structure du récepteur conduit au rapport de vraisemblance [18]. Ce résultat s'obtient en utilisant l'inégalité de Schwarz, qui entraîne

$$(5.6) \quad \forall T, D(T, P) \leq D^*(P) = D(L_p, P) \\ = \int_{\mathbf{R}^n} \frac{P^2(\mathbf{x}-\mathbf{s})}{P(\mathbf{x})} d\mathbf{x}.$$

où $L_p(\mathbf{x})$ est le rapport de vraisemblance associé à $P(\cdot)$ et à \mathbf{s} , c'est-à-dire défini comme en (4.5) en remplaçant P_0 par P . Malheureusement, dans ce cas général, la minimisation de $D^*(P)$ donnée par (5.6) est un problème complexe dont un élément de solution est donné dans [19].

Par contre, si le récepteur $T(\mathbf{x})$ présent dans (5.5) est contraint à avoir une structure linéaire,

$$(5.7) \quad T(\mathbf{x}) = \mathbf{h}' \mathbf{x},$$

l'expression de la déflexion (5.5) se réduit à

$$(5.8) \quad D(T, P) = \frac{E\{\mathbf{h}' \mathbf{x}/H_1\}^2}{E\{(\mathbf{h}' \mathbf{x})^2/H_0\}}.$$

Identique au critère (3.2). Ce fait est, bien entendu, naturel car la contrainte de linéarité (5.7) conduit à utiliser une structure de filtre adapté. Les résultats du paragraphe III s'appliquent donc : on obtient le filtre adapté, $\mathbf{h} = \Gamma^{-1} \mathbf{s}$.

V. 2. DÉFLEXION LOCALE

Toujours pour le problème de détection d'un signal déterministe dans un bruit additif de ddp $P(\cdot)$, le critère de déflexion locale s'exprime selon (5.4). L'utilisation de l'inégalité de Schwarz conduit au récepteur maximisant ce critère [15, 18],

$$(5.9) \quad T_{l_0}(\mathbf{x}) = \frac{\mathbf{s}' \nabla P(\mathbf{x})}{P(\mathbf{x})}.$$

On obtient le récepteur localement optimal. La valeur maximale de la déflexion locale est alors

$$(5.10) \quad d^*(P) = \int_{\mathbf{R}^n} \frac{(\mathbf{s}' \nabla P(\mathbf{x}))^2}{P(\mathbf{x})} d\mathbf{x}.$$

Il est facile d'exprimer $d^*(P)$ sous la forme matricielle,

$$(5.11) \quad d^*(P) = \mathbf{s}' F_p \mathbf{s}$$

avec

$$(5.12) \quad F_p \stackrel{\Delta}{=} E\{\nabla \log P(\mathbf{x}) (\nabla \log P(\mathbf{x}))'\}.$$

Cette matrice carrée $n \times n$ est la matrice d'information de Fisher de $P(\cdot)$ [9]. Plus précisément c'est la matrice d'information du problème d'estimation d'un paramètre de position, disons $\boldsymbol{\theta}$, dans la ddp vectorielle $P(\mathbf{x}-\boldsymbol{\theta})$ [4].

La recherche de la ddp minimisant la fonctionnelle $d^*(P)$ s'effectue, muni de la remarque précédente, en utilisant l'inégalité de Cramer-Rao vectorielle [20]. De manière générale, celle-ci exprime que, pour une variable aléatoire \mathbf{Y} de \mathbf{R}^n possédant une densité $f(\mathbf{Y}, \boldsymbol{\theta})$, $\boldsymbol{\theta}$ appartenant à \mathbf{R}^p ,

$$(5.13) \quad E_{\boldsymbol{\theta}}\{(\mathbf{Y} - E_{\boldsymbol{\theta}}(\mathbf{Y}))(\mathbf{Y} - E_{\boldsymbol{\theta}}(\mathbf{Y}))'\} \\ - (\text{grad } E_{\boldsymbol{\theta}}(\mathbf{Y}))' (M(\boldsymbol{\theta}))^{-1} (\text{grad } E_{\boldsymbol{\theta}}(\mathbf{Y}))$$

est définie positive, où $M(\boldsymbol{\theta})$ est la matrice d'information de Fisher du problème d'estimation d'éléments ij ,

$$(5.14) \quad m_{i,j} \stackrel{\Delta}{=} E_{\boldsymbol{\theta}}\left\{ \frac{\partial}{\partial \theta_i} \log L(\boldsymbol{\theta}) \cdot \frac{\partial}{\partial \theta_j} \log L(\boldsymbol{\theta}) \right\},$$

$L(\boldsymbol{\theta})$ étant la vraisemblance, et $\text{grad}(\mathbf{V})$ étant la matrice d'élément $i, j : \partial V_i / \partial \theta_j$. Il faut évidemment que $L(\boldsymbol{\theta})$ soit différentiable, que $\nabla \log L(\boldsymbol{\theta})$ soit centrée et de carré intégrable et que $M(\boldsymbol{\theta})$ soit inversible. Pour le problème d'estimation d'un paramètre de position, on a, de manière générale,

$$(5.15) \quad f(\mathbf{Y}, \boldsymbol{\theta}) = P(\mathbf{Y} - \boldsymbol{\theta}).$$

Les différents termes intervenant dans la matrice (5.13) sont alors

$$(5.16) \quad E_{\boldsymbol{\theta}}\{(\mathbf{Y} - E_{\boldsymbol{\theta}}(\mathbf{Y}))(\mathbf{Y} - E_{\boldsymbol{\theta}}(\mathbf{Y}))'\} = \Gamma,$$

matrice de corrélation associée à $P(\mathbf{x})$,

$$(5.17) \quad \text{grad}(E_{\boldsymbol{\theta}}(\mathbf{Y})) = \text{Id},$$

la matrice identité $n \times n$,

$$(5.18) \quad M(\theta) = F_p,$$

donnée par (5.12).

L'inégalité de Cramer-Rao prouve alors que la matrice $\Gamma - F_p^{-1}$ est définie positive, sous la réserve que $P(x)$ soit différentiable et que $E\{\nabla \log P(x)\}$ soit nulle, ce qui est vérifiée si $P(x)$ est symétrique. L'utilisation du théorème de l'annexe A, F_p^{-1} et Γ étant deux matrices carrées symétriques définies positives, prouve que $F_p - \Gamma^{-1}$ est définie positive, ce qui s'exprime comme

$$(5.19) \quad s^t \Gamma^{-1} s \leq s^t F_p s = d^*(P).$$

Il ne reste plus qu'à remarquer que le terme de gauche de cette inégalité, $s^t \Gamma^{-1} s$, est la valeur de la déflexion locale pour la loi de Gauss centrée de matrice de covariance Γ pour prouver que $d^*(P)$ possède un minimum pour cette loi de Gauss. Ce minimum est global dans l'ensemble des ddp différentiables et symétriques (ou telles que $E\{\nabla \log P(x)\} = 0$), car $d^*(P)$ est convexe, ce qui est démontrée dans l'annexe B. Une démonstration équivalente prouve que la déflexion locale (5.4) est convexe en P à T fixé.

Le filtre robuste est alors, au signe près, le filtre adapté. Grâce à l'invariance par homothétie de la déflexion locale, on peut utiliser celui-ci,

$$(5.20) \quad T(x) = s^t \Gamma^{-1} x.$$

Lorsque le récepteur est contraint à être un filtre linéaire, et en utilisant le critère de déflexion locale, celui-ci s'exprime selon

$$(5.21) \quad d(h, P) = \frac{\left(\int_{\mathbf{R}^n} h^t x P(x) dx \right)^2}{\int_{\mathbf{R}^n} (h^t x)^2 P(x) dx}.$$

De la même manière que précédemment, on obtient le filtre adapté, $s^t \Gamma^{-1} x$.

V. 3. DISCUSSION

Dans les trois derniers cas étudiés dans ce paragraphe, le filtre robuste existe et correspond à la loi gaussienne : c'est le filtre adapté. L'obtention du filtre adapté est relativement naturelle. Si il y a une incertitude sur la loi du bruit, il semble logique que le récepteur « robuste » réalise la corrélation de l'observation avec le signal attendu, pour obtenir le plus grand « contraste » visuel à la sortie.

Une autre propriété intéressante est que le filtre adapté peut être vu comme la projection du rapport de vraisemblance sur l'espace des filtres linéaires. En d'autres termes, le filtre adapté minimise l'erreur quadratique moyenne entre le rapport de vraisemblance et les filtres linéaires

$$(5.22) \quad E\{\|L(x) - h^t x\|^2\} \geq E\{\|L(x) - s^t \Gamma^{-1} x\|^2\}.$$

Enfin, l'obtention de la loi de Gauss comme la loi « la pire » est évidemment à relier à l'entropie : à puissance constante, celle-ci est maximisée pour la loi de Gauss.

VI. Conclusion

Les bases théoriques élémentaires de l'étude d'un problème de détection d'un signal dans un bruit en utilisant l'approche robuste ont été détaillées. En utilisant différentes contraintes sur la structure du récepteur et plusieurs critères de mesure des performances, des récepteurs robustes ont été obtenus. Leur structures font appel au filtre adapté au signal attendu, suivi (et parfois précédé) d'une non-linéarité. Lorsque le bruit a été modélisé selon une mixture impulsive, cette non-linéarité est un écréteur, limitant l'influence des impulsions dues au bruit. Lorsque le bruit a été supposé de second ordre fixé cette non-linéarité disparaît. Toutefois l'approche robuste utilise trop peu d'éventuelles informations disponibles sur la structure de l'observation. Elle semble ainsi inférieure en qualité à une « bonne » approche adaptative, ce qui a été montré sur des exemples de bruits réels en [21] et [22].

Annexe A

Théorème : Si A et B sont deux matrices carrées symétriques définies positives, $A - B$ est définie positive entraîne que $B^{-1} - A^{-1}$ est définie positive.

Démonstration : A étant symétrique définie positive, il existe une matrice M définie positive symétrique telle que

$$(A.1) \quad A = M^t M.$$

On exprime le caractère défini positif de $A - B$ par

$$(A.2) \quad \forall x \in \mathbf{R}^n, \quad x^t A x \geq x^t B x.$$

En posant

$$(A.3) \quad y = M x, \quad K = M^{-t} B M^{-1},$$

il vient

$$(A.4) \quad \forall y \in \mathbf{R}^n, \quad y^t y \geq y^t K y.$$

B et M étant symétriques définies-positives, K l'est aussi. D'après (A.4) ses valeurs propres sont inférieu-

res ou égales à 1. Donc les valeurs propres de K^{-1} sont supérieures ou égales à 1, ce qui s'exprime de manière similaire à (A. 4) par

$$(A. 5) \quad \forall y \in \mathbb{R}^n, \quad y^t y \leq y^t K^{-1} y.$$

En posant

$$(A. 6) \quad z = My$$

et en exprimant K^{-1} en fonction de M et B,

$$(A. 7) \quad K^{-1} = MB^{-1}M^t,$$

il vient, compte tenu de (A. 1)

$$(A. 8) \quad \forall z \in \mathbb{R}^n, \quad z^t A^{-1} z \leq z^t B^{-1} z,$$

ce qui prouve que $B^{-1} - A^{-1}$ est définie positive.

Annexe B

CONVEXITÉ DE $d^*(P)$

La convexité de $d^*(P)$ se démontre en utilisant le résultat suivant (lemme 6 de [4]) :

Si $y_1 > 0, y_2 > 0, 0 \leq \alpha \leq 1$, alors

$$(B. 1) \quad \frac{(\alpha x_1 + (1-\alpha)x_2)^2}{\alpha y_1 + (1-\alpha)y_2} \leq \alpha \frac{x_1^2}{y_1} + (1-\alpha) \frac{x_2^2}{y_2}.$$

Ceci se montre en posant

$$(B. 2) \quad \begin{cases} Y = \alpha y_1 + (1-\alpha)y_2 \\ \gamma = \alpha \frac{y_1}{Y} \end{cases}$$

et en remarquant

$$(B. 3) \quad \frac{(\alpha x_1 + (1-\alpha)x_2)^2}{\alpha y_1 + (1-\alpha)y_2} = Y \left(\gamma \frac{x_1}{y_1} + (1-\gamma) \frac{x_2}{y_2} \right)^2$$

qui, par convexité de $x \rightarrow x^2$, est inférieur ou égal à

$$(B. 4) \quad Y \left(\gamma \left(\frac{x_1}{y_1} \right)^2 + (1-\gamma) \left(\frac{x_2}{y_2} \right)^2 \right) = \alpha \frac{x_1^2}{y_1} + (1-\alpha) \frac{x_2^2}{y_2}.$$

La convexité de $d^*(P)$ se démontre alors en remplaçant dans (B. 1) x_1 par $s^t \nabla P(x)$, x_2 par $s^t \nabla Q(x)$, y_1 par $P(x)$ et y_2 par $Q(x)$ [en remarquant que si $P(x)$ ou $Q(x)$ s'annule, le membre de droite de l'inégalité i. e., (B. 4), devient infini] et en intégrant :

$$(B. 5) \quad \int_{\mathbb{R}^n} \frac{(\alpha s^t \nabla P(x) + (1-\alpha)s^t \nabla Q(x))^2}{\alpha P(x) + (1-\alpha)Q(x)} dx \leq \alpha \int_{\mathbb{R}^n} \frac{(s^t \nabla P(x))^2}{P(x)} dx + \beta \int_{\mathbb{R}^n} \frac{(s^t \nabla Q(x))^2}{Q(x)} dx$$

Remerciements

Cette recherche s'est principalement effectuée au Laboratoire des Signaux et Systèmes (Gif-sur-Yvette) et a été supportée en partie par la Direction des Constructions Navales sous le contrat n° C. 86. 48. 826. 220. 000.

L'auteur tient à exprimer sa gratitude au professeur B. Picinbono pour son accueil et d'intéressantes discussions sur la robustesse et à remercier les rapporteurs pour leurs nombreuses remarques constructives qui ont grandement amélioré la clarté de cet article.

Manuscrit reçu le 23 juin 1986.

BIBLIOGRAPHIE

- [1] J. W. TUKEY, Useable resistant/robust techniques of analysis, *Proceedings of the 1st ERDA Statistical Symposium*, Los Alamos, 3-5 novembre 1976.
- [2] J. W. TUKEY, A survey of sampling from contaminated distributions, dans *contributions to Probability and Statistics*, I. OLKIN éd. Stanford University Press, 1960.
- [3] F. R. HAMPELL, A general qualitative definition of robustness, *Annal. Math. Stat.*, 42, 1971, p. 1887-1895.
- [4] P. J. HUBER, Robust Estimation of a location parameter, *Annal. Math. Stat.*, 35, 1964, p. 73-101.
- [5] S. A. KASSAM et H. V. POOR, Robust techniques for signal processing: a survey, *Proceedings of IEEE*, 73, n° 3, 1985, p. 433-481.
- [6] V. M. KRASNENKER, Stable (robust) detection methods for signal against a noise background (survey), traduit de *Automatika i Telemekhanika*, n° 5, 1980, p. 65-88.
- [7] A. WALD, *Statistical decision functions*, John Wiley, 1950.
- [8] T. S. FERGUSON, *Mathematical statistics: a decision theoretic approach*, Academic Press, 1967.
- [9] P. J. HUBER, *Robust statistics*, John Wiley, 1981.
- [10] S. VERDU et H. V. POOR, Minimax robust discrete-time matched filters, *IEEE Transactions on Com.*, 31, n° 2, 1983, p. 208-215.
- [11] B. PICINBONO, Detection with uncertainties: non parametric, robust on adaptive approaches, dans *NATO ASI on Adaptive methods in underwater acoustics*, H. URBAN éd., Reidel, p. 497-514.
- [12] M. BOUVET, Détection en environnement non gaussien : différentes approches et utilisation des modèles de mixture, *Traitement du Signal*, vol. 4, n° 2, juin 1987.
- [13] R. D. MARTIN et S. C. SCHWARTZ, Robust detection of a known signal in nearly Gaussian noise, *IEEE Transactions on IT*, 17, n° 1, 1971, p. 50-56.
- [14] P. J. HUBER, A robust version of the probability ratio test, *Ann. Maths. Stat.*, 36, 1965, p. 1753-1758.
- [15] M. BOUVET, Expansions of the likelihood ratio and applications, *IEEE Transactions on ASSP*, 34, n° 4, août 1986.

- [16] S. A. KASSAM, Non parametric and robust detection: dependant data and conditional tests, *Ph. D. Dissertation*, EECS Dept, Princeton University, 1975.
- [17] J. B. THOMAS, Non parametric detection, *Proceedings of IEEE*, 58, n° 5, 1970, p. 623-631.
- [18] B. PICINBONO et P. DUVAUT, Detection and contrast, chap. 7 de *Stochastic processes in underwater acoustics*, C. BAKER éd., Springer Verlag, 1986.
- [19] M. BOUVET, Étude de certains problèmes de détection et normalisation adaptatives. Application à l'acoustique sous-marine, *Thèse de Doctorat d'État*, Université de Paris-XI, 5 février 1987.
- [20] C. FOURGEAUD et A. FUCHS, *Statistique*, Dunod, 1967.
- [21] M. BOUVET et S. C. SCHWARTZ, Comparison between adaptive and robust receivers for signal detection in ambient underwater noise, soumis aux *IEEE Transactions on ASSP*, juin 1985.
- [22] M. BOUVET et S. C. SCHWARTZ, Signal detection and normalization in underwater noises modeled as a Gaussian-Gaussian mixture, *Information Sciences and Systems Laboratory*, Report n° 18, Princeton University, janvier 1986.