

# *Fusion de données*

---

## Détection décentralisée Critère de Neyman-Pearson sur des architectures arborescentes

---

### *Decentralized Detection : Neyman-Pearson Criterion on Tree-Structured Architectures*

par **F. BALDIT, J.-P. LE CADRE**

*IRISA/CNRS, Campus de Beaulieu, 35042 Rennes Cedex, France.*

#### Résumé

La multiplicité et la diversité des sources d'information, autant que celle des agents susceptibles de les recevoir et de les traiter, rend de plus en plus illusoire l'usage d'un processeur centralisé pour la gestion des grands systèmes de traitement de l'information. Il est alors nécessaire de prendre en compte les contraintes inhérentes aux systèmes considérés. Celles-ci visent essentiellement à limiter les capacités des communications sous-jacentes au fonctionnement du système. Dans ce contexte, les systèmes de détection décentralisée (D.D.) constituent un exemple typique de l'effort fait pour modéliser de tels systèmes sous des contraintes de communications.

Après un rappel du formalisme adapté à l'étude des architectures arborescentes, trois versions, de généralité croissante, du problème de Neyman-Pearson (N.-P.) sont considérées. Une généralisation du lemme de N.-P. permet de démontrer l'optimalité des règles de décision opérant un seuillage du rapport de vraisemblance de l'observation, ceci pour une architecture arborescente quelconque et des mécanismes de randomisation non-indépendants.

**Mots clés :** Détection décentralisée, architectures arborescentes, hypothèses simples, critère de Neyman-Pearson, randomisation d'équipe.

#### Abstract

*For numerous applications, both the information sources and the processing agents are multiple and of very diverse natures. The interest of a centralized and direct processing of the information sources becomes then quite illusory, especially for "large" systems. It is then necessary to take into account the own system constraints. Frequently, these constraints are due to the communication capacities inside the system. In this context, the decentralized detection (D.D.) systems illustrate the effort made for modelling such systems in the presence of communication constraints.*

*The general formalism for tree-structured architectures is presented in a first time. Then, three versions, of increasing complexity, of the Neyman-Pearson (N.-P.) are presented. Using an extension of the N.-P. lemma it is then possible to demonstrate the optimality of the decision rules based on thresholding the likelihood ratio both for general tree-structured architectures and non-independent randomization schemes.*

**Key words :** *Decentralized detection, tree-structured architectures, simple hypotheses, team randomization.*

## 1. Introduction

Un des principes les plus solidement établis en traitement du signal est celui qui consiste à mettre en œuvre des techniques préservant au mieux l'information disponible dans les données. Ce principe, qui constitue souvent la base (et parfois l'objectif) d'un grand nombre de méthodes, ne peut toujours être respecté. La multiplication et la diversité des sources d'information, autant que celle des agents susceptibles de les recevoir et de les traiter, rend de plus en plus illusoire l'usage d'un processeur centralisé pour la gestion des « grands » systèmes de traitement de l'information.

Pour cette raison, les techniques modernes (de détection, d'estimation et de commande) doivent de plus en plus fréquemment prendre en compte des *contraintes* inhérentes aux systèmes considérés. Celles-ci visent essentiellement à limiter les capacités des *communications* sous-jacentes au fonctionnement du système et à formaliser de cette façon le fait que « tout le monde ne peut être informé de tout ». Dans ce contexte, les systèmes de *détection décentralisée* (D.D.), et plus spécifiquement les *arborescences* de détection décentralisée, sont un exemple typique de l'effort fait pour modéliser de tels systèmes. Ils constituent un domaine pour lequel on a pu reprendre (au moins partiellement) les calculs classiques en théorie de la détection et montrer quelles étaient les

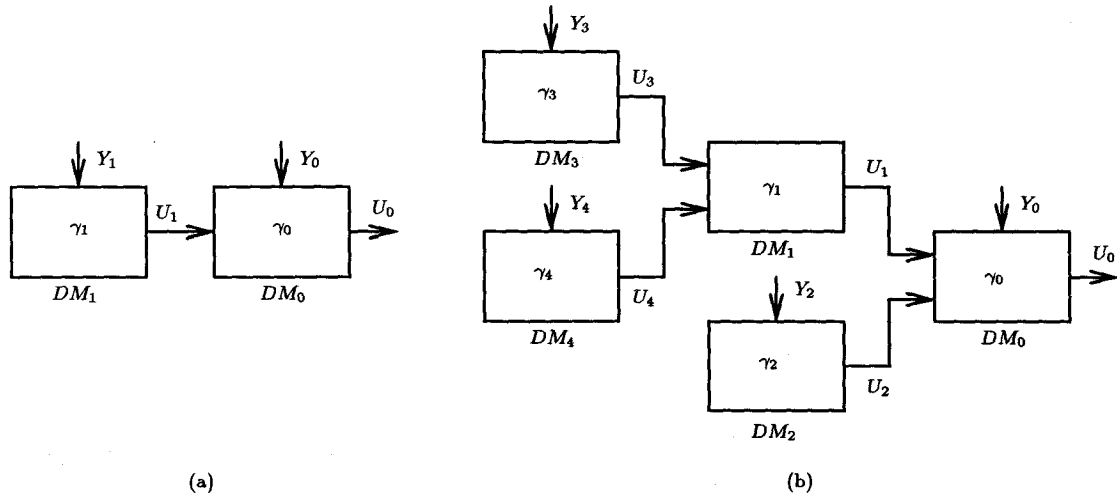


Figure 1. – (a) Architecture de D.D. à deux décideurs en tandem. (b) Architecture de D.D. mixte (fusion et tandem) à cinq décideurs.

règles de décision optimales devant être utilisées par les différents agents, ou « décideurs » du système.

Nous proposons dans cet article une revue des résultats acquis en *détection décentralisée au sens de Neyman-Pearson* (N.-P.) dans le cadre arborescent. Bien que cette étude soit largement motivée par une connaissance théorique des propriétés de ces systèmes, il s'avère que le critère de N.-P. trouve assez naturellement des applications dans un contexte décentralisé. L'exemple le plus typique de ce fait est sans doute le cas des radars multistatistiques, mais, plus généralement, le domaine des traitements « multiplateformes » semble un bon champ d'application de la théorie, [BLC93].

Le plan de l'article est le suivant. Le paragraphe 2 est un rappel du formalisme adapté à l'étude des architectures arborescentes et des différents mécanismes de randomisation qui peuvent être envisagés dans ce cadre. On y donne un énoncé précis du problème de N.-P. en D.D. arborescente. Nous montrons, au paragraphe 3, que trois versions du problème de N.-P., de généralité croissante, peuvent être choisies, associées à trois diagrammes des risques. Le paragraphe 4 établit un lemme important pour la suite, qui est une généralisation du lemme classique de N.-P. A l'aide de ces résultats, on démontre au paragraphe 5 l'optimalité des règles de décision opérant un seuillage du rapport de vraisemblance (r.d.v.) de l'observation. Ce travail est fait d'abord pour une architecture simple à deux décideurs, puis étendu au cas d'une arborescence quelconque. On obtient ce premier résultat dans le contexte du mécanisme de randomisation d'équipe indépendante. Pour le cas général (de la randomisation « non-indépendante ») le paragraphe 6 montre comment on peut décrire toute règle de décision décentralisée optimale au sens de N.-P. comme l'usage coordonné de deux règles de décision globales à randomisation indépendante. Le paragraphe 7 conclut notre étude et donne des références sur quelques pistes possibles pour poursuivre ce travail.

Les lecteurs familiers du sujet de la D.D. noteront que nous nous sommes fortement inspirés de la référence [Tsi93] qui demeure une excellente introduction.

## 2. Le formalisme des architectures arborescentes de D.D.

Le schéma de la fig. (1)-(a) décrit l'architecture de D.D. la plus simple, qu'on appelle architecture en tandem à deux décideurs. Sur cette figure,  $DM = \{DM_0, DM_1\}$  est une équipe de deux décideurs appelés à collaborer dans la réalisation d'un test à deux hypothèses simples  $H_0$  et  $H_1$ .  $DM_1$  peut observer la variable aléatoire (v.a.)  $Y_1$ , de loi  $P_j^1$  sous  $H_j$ . Il ne peut communiquer à  $DM_0$  qu'une version « résumée » (par exemple sur un bit) de  $Y_1$ , sous la forme de la décision  $U_1 = \gamma_1(Y_1)$ , à valeurs dans l'ensemble fini  $D_1$ .  $DM_0$  observe à son tour  $Y_0$ , de loi  $P_j^0$  sous  $H_j$ , reçoit  $U_1$  et doit émettre la décision finale du système,  $U_0 = \gamma_0(U_1, Y_0) \in \{0, 1\}$ , où  $U_0 = j$  signifie que l'hypothèse  $H_j$  est considérée réalisée. Nous supposons que les composantes  $Y_0$  et  $Y_1$  de l'observation complète  $Y = (Y_0, Y_1)$  sont indépendantes sous chaque hypothèse, ceci afin d'éviter les problèmes de complexité algorithmique que pose le cas général, [TA85]. On vérifie que cet exemple regroupe les contraintes évoquées dans l'introduction, puisqu'un canal de capacité illimitée entre  $DM_1$  et  $DM_0$  permettrait à ce dernier de réaliser le test centralisé optimal. D'autre part, cet exemple définit bien un test d'hypothèses décentralisé au sens de N.-P. pourvu qu'on se donne un niveau  $\alpha \in ]0, 1[$ . Les inconnues de ce problème sont les règles de décision  $\gamma_0$  et  $\gamma_1$ , où  $\gamma_1 : Z_1 \rightarrow D_1$ ,  $\gamma_0 : D_1 \times Z_0 \rightarrow D_0 = \{0, 1\}$ . Les ensembles  $Z_i$  et  $D_i$  sont ceux où les v.a.  $Y_i$  et  $U_i$  prennent respectivement leurs

valeurs. On note  $Y = (Y_0, Y_1)$  et  $U = (U_0, U_1)$ , à valeurs dans  $Z = Z_0 \times Z_1$  et  $D = D_0 \times D_1$  respectivement. Les probabilités de fausse alarme et détection sont notées<sup>1</sup>  $J^F(\underline{\gamma}) = \mathbb{P}_0[U_0 = 1]$  et  $J^D(\underline{\gamma}) = \mathbb{P}_1[U_0 = 1]$ , où  $\underline{\gamma} = (\gamma_0, \gamma_1)$  est la stratégie de décision adoptée par l'équipe. La fig. (1)-(b) montre comment généraliser l'exemple précédent au cas d'une arborescence à cinq décideurs, dont le fonctionnement se comprend aisément et respecte la notion de prise de décision hiérarchique. Pour une telle arborescence  $A$  composée de  $N$  décideurs, une stratégie de décision devient un  $N$ -uplet  $\underline{\gamma} = (\gamma_0, \gamma_1, \dots, \gamma_{N-1})$  dont la définition repose entièrement sur les caractéristiques de l'arborescence<sup>2</sup>  $A$ . On note  $\Gamma_A$  l'ensemble de ces stratégies, avec  $\Gamma_A = \Gamma_{A,0} \times \Gamma_{A,1} \times \dots \times \Gamma_{A,N-1}$ . Ceci nous permet de représenter une règle *pure* (déterministe) de décision décentralisée. Pour le cas des règles *mixtes* (incluant un aléa « interne » dans la prise de décision) nous admettons ici la généralité et l'équivalence des deux formalismes suivants. Premièrement, on peut considérer une suite  $(\gamma_i)_{i \geq 0}$  de règles pures *globales* et une v.a.  $I$  entière<sup>3</sup> de loi  $(p_i)_{i \geq 0}$  indépendante des observations. Sous ce formalisme, une règle mixte d'équipe consiste à utiliser la règle pure  $\gamma_i$  avec la probabilité  $p_i$ , ce qu'on note sous forme compacte  $U = \underline{\gamma}_I(Y)$ . Ici, chaque décideur  $DM_j$  reçoit la suite projetée  $(\gamma_{j,i})_{i \geq 0}$ , connaît le programme de génération de la v.a.  $I$  et « choisit au hasard » la règle  $\gamma_{j,i}$  si  $i$  est la réalisation de  $I$  qu'il observe. Un second formalisme *local* équivalent au premier consiste à attribuer à chaque décideur  $DM_j$  une suite  $(\gamma_{j,i})_{i \geq 0}$  de  $\Gamma_{A,j}$  et une v.a.  $I_j$  entière, les v.a.  $(I_0, I_1, \dots, I_{N-1})$  étant globalement indépendantes de l'observation  $Y = (Y_0, Y_1, \dots, Y_{N-1})$  et de loi conjointe fixée.

Nous noterons respectivement dans la suite  $\Gamma_A$  et  $\Gamma_A^*$  l'ensemble des règles pures et mixtes associées à une arborescence  $A$ , qui satisfont à l'inclusion évidente  $\Gamma_A \subset \Gamma_A^*$ . Nous noterons aussi (avec un léger abus de notation)  $\underline{\gamma}$  une règle de  $\Gamma_A, \bar{\Gamma}_A$  ou  $\Gamma_A^*$ . On remarque à ce stade qu'un ensemble particulier de règles mixtes, dites à *randomisation d'équipe indépendante*, réside entre  $\Gamma_A$  et  $\Gamma_A^*$ , qu'on notera  $\bar{\Gamma}_A$ . C'est l'ensemble des règles mixtes ayant la particularité d'être associées, sous le formalisme mixte local, à des v.a.  $I_j$  mutuellement indépendantes. Ce mécanisme signifie que les décideurs n'ont pas à se coordonner dans leur randomisation. Pour ce formalisme, on remarque qu'une règle globale  $\underline{\gamma} \in \bar{\Gamma}_A$  peut se décrire comme un  $N$ -uplet de règles locales choisies librement, c'est à dire  $\bar{\Gamma}_A = \bar{\Gamma}_{A,0} \times \bar{\Gamma}_{A,1} \times \bar{\Gamma}_{A,N-1}$ . Il revient alors au même de décrire  $\gamma_i \in \bar{\Gamma}_{A,i}$  par la donnée d'une probabilité conditionnelle  $\phi^i = (\phi_1^i, \dots, \phi_{m_i}^i)$ , où  $\phi_k^i : D^i \times Z_i \rightarrow [0, 1]$  est telle que  $\mathbb{P}([U_i = k] | U^i = u^i, Y_i = y_i) = \phi_k^i(u^i, y_i)$ ,  $k = 1, \dots, m_i$ . Dans cette écriture  $D^i$  est l'ensemble des décisions  $u^i$  que  $DM_i$  peut recevoir de ses prédécesseurs immédiats,  $Y_i$  est son obser-

vation propre (éventuellement dégénérée, i.e. constante) et  $U_i$  sa décision, à valeurs dans  $D_i$ , de cardinal  $m_i$ . On désignera cette description d'une règle mixte à randomisation indépendante (i.e. la donnée de  $\phi = (\phi^0, \phi^1, \dots, \phi^{N-1})$ ) sous le nom de *comportementale*. Nous renvoyons le lecteur [Bal94] pour une justification et une compréhension plus complète et détaillée des formalismes qui viennent d'être présentés. Rappelons seulement ici que la randomisation est une technique qui consiste à introduire dans la prise de décision opérée par le statisticien une part de « hasard local », indépendamment de la donnée qui est observée. On sait que ce mécanisme, même s'il peut être d'une portée pratique minime, est important d'un point de vue théorique pour l'étude du problème de détection centralisée au sens de N.-P. La généralisation de ce principe de randomisation au cas des architectures arborescentes de D.D. conduit aux formalismes associés  $\bar{\Gamma}_A$  et  $\Gamma_A^*$ . On peut dire schématiquement que, pour une règle de  $\Gamma_A^*$ , on autorise les différents décideurs de l'arborescence à « coordonner leurs hésitations » (v.a.  $I_j$  non indépendantes), tandis que pour une règle de  $\bar{\Gamma}_A$  ces hésitations personnelles doivent être mutuellement indépendantes. La possibilité offerte par  $\bar{\Gamma}_A$  est une caractéristique propre aux problèmes de D.D. arborescente.

Ayant en main cet ensemble de définitions on peut aborder l'étude du problème de N.-P. sur  $\Gamma_A, \bar{\Gamma}_A$  ou  $\Gamma_A^*$ , qui correspond à

$$\max_{\{\underline{\gamma} \in \Gamma_A, \bar{\Gamma}_A \text{ ou } \Gamma_A^* : J_A^F(\underline{\gamma}) \leq \alpha\}} [J_A^D(\underline{\gamma})]. \quad (1)$$

### 3. Trois versions du problème de D.D. au sens de Neyman-Pearson

Utilisant les notions de règles de décision pures, randomisées et randomisées indépendantes, qui sont respectivement associées aux ensembles  $\Gamma_A, \Gamma_A^*$  et  $\bar{\Gamma}_A$ , on peut définir trois versions du problème (1) associées à une arborescence de D.D.  $A$ . Il résulte de ces trois problèmes trois diagrammes des risques (dont les enveloppes supérieures correspondent à la notion classique de courbe C.O.R) définis comme suit :

**Définition 3.1** Soit  $A$  une arborescence de D.D.. On appelle *diagramme des risques* associés à  $A$  et on note  $Q_A, \bar{Q}_A$  et  $Q_A^*$  les sous-ensembles de  $[0, 1] \times [0, 1]$  définis par :

$$\begin{aligned} Q_A &= \{(J_A^F(\underline{\gamma}), J_A^D(\underline{\gamma})); \underline{\gamma} \in \Gamma_A\} \\ \bar{Q}_A &= \{(J_A^F(\underline{\gamma}), J_A^D(\underline{\gamma})); \underline{\gamma} \in \bar{\Gamma}_A\} \\ Q_A^* &= \{(J_A^F(\underline{\gamma}), J_A^D(\underline{\gamma})); \underline{\gamma} \in \Gamma_A^*\} \end{aligned} \quad (2)$$

Une première étude de la topologie de  $Q_A, \bar{Q}_A$  et  $Q_A^*$  va nous révéler certaines des propriétés de ces diagrammes. Elle nous montrera en quoi la problématique de N.-P. décentralisée diffère du cas centralisé et nous guidera vers l'étude ultérieure en indiquant les questions qui devraient être attaquées.

1. Nous noterons  $\mathbb{P}_j$  et  $\mathbb{E}_j$  la loi de probabilité et l'espérance mathématique lorsque  $H_j$  est l'hypothèse en vigueur.

2. Nous renvoyons à [Bal94] pour toutes les définitions omises dans cet article.

3. i.e. à valeurs dans  $\mathbb{N}$ .

**Proposition 3.1** Soit  $A$  une arborescence de D.D. et  $Q_A, \bar{Q}_A$  et  $Q_A^*$  les diagrammes des risques respectivement associés à  $\Gamma_A, \bar{\Gamma}_A$  et  $\Gamma_A^*$ . On a<sup>4</sup>

- (i)  $Q_A \subset \bar{Q}_A \subset Q_A^*$ ,
- (ii)  $Q_A^* = co(Q_A) = co(\bar{Q}_A)$ ,
- (iii)  $\bar{Q}_A$  et  $Q_A^*$  sont compacts.

**Preuve :** (i) résulte directement des inclusions  $\Gamma_A \subset \bar{\Gamma}_A \subset \Gamma_A^*$ .  
(ii) Soit  $q \in Q_A^*$ ,  $q = (J_A^F(\underline{\gamma}), J_A^D(\underline{\gamma}))$ ,  $\underline{\gamma} \in \Gamma_A^*$ . On a  $U = (U_0, U_1, \dots, U_{N-1}) = \underline{\gamma}_I(Y_0, Y_1, \dots, Y_{N-1}) = \underline{\gamma}_I(Y)$ , où  $(\underline{\gamma}_i)_{i \geq 0}$  est une suite de  $\Gamma_A$  et  $I$  une v.a. entière indépendante de  $Y$ . Il est clair qu'on peut toujours écrire  $U_0 = \gamma_{0,I}(Y)$  (avec un léger abus de notation) de façon à transmettre toute l'influence de  $I$  sur  $\gamma_0$ . Par exemple, dans l'arborescence tandem à deux décideurs,  $U_0 = \gamma_{0,I}(U_1, Y_0) = \gamma_{0,I}(\gamma_{1,I}(Y_1), Y_0)$  (sans abus de notation).  $(\gamma_{0,i})_{i \geq 0}$  est une suite de règles pures, v.a. de  $Z$  dans  $\{0, 1\}$  construite à partir de la suite  $(\underline{\gamma}_i)_{i \geq 0}$  de  $\Gamma_A$ . Donc,

$$\begin{aligned} J_A^F(\underline{\gamma}) &= \mathbb{P}_0[U_0 = 1] \\ &= \mathbb{E}_0 [\mathbb{P}_0([\gamma_{0,I}(Y) = 1] | I)] \\ &= \mathbb{E}_0 \left[ \sum_{i \geq 0} \mathbf{1}_{[I=i]} \cdot \mathbb{P}_0([\gamma_{0,i}(Y) = 1] | I) \right] \\ &= \mathbb{E}_0 \left[ \sum_{i \geq 0} \mathbf{1}_{[I=i]} \cdot \mathbb{P}_0[\gamma_{0,i}(Y) = 1] \right] \\ &= \sum_{i \geq 0} P^I\{i\} \cdot J_A^F(\underline{\gamma}_i) \\ &= \sum_{i \geq 0} P^I\{i\} \cdot q_i^F, \end{aligned}$$

$q_i^F = J_A^F(\underline{\gamma}_i)$ . De la même façon,

$$\begin{aligned} J_A^D(\underline{\gamma}) &= \mathbb{P}_1[U_0 = 1] \\ &= \sum_{i \geq 0} P^I\{i\} \cdot q_i^D, \end{aligned}$$

où  $q_i^D = J_A^D(\underline{\gamma}_i)$ . On voit donc que  $q = (J_A^F(\underline{\gamma}), J_A^D(\underline{\gamma})) = \sum_{i \geq 0} P^I\{i\} \cdot (q_i^F, q_i^D) \in co(Q_A)$ . On vient de montrer que  $Q_A^* \subset co(Q_A)$ . Réciproquement, soit  $q = \sum_{i \geq 0} p_i \cdot q_i$ ,  $p_i \geq 0$  et  $\sum_{i \geq 0} p_i = 1$  une combinaison convexe quelconque de points  $q_i$  de  $Q_A$ . Elle est associée à une suite  $(\underline{\gamma}_i)_{i \geq 0}$  de  $\Gamma_A$ , avec  $q_i = (J_A^F(\underline{\gamma}_i), J_A^D(\underline{\gamma}_i))$ . Considérons la règle  $\underline{\gamma}$  utilisant  $\underline{\gamma}_i$  avec la probabilité  $p_i$  selon le tirage d'une v.a.  $I$  entière, indépendante de  $Y$ , de loi  $P^I\{i\} = p_i$ ,  $i \geq 0$ . Il est clair que  $\underline{\gamma} \in \Gamma_A^*$  et que, par construction,  $\sum_{i \geq 0} p_i \cdot q_i = (\sum_{i \geq 0} P^I\{i\} \cdot J_A^F(\underline{\gamma}_i), \sum_{i \geq 0} P^I\{i\} \cdot J_A^D(\underline{\gamma}_i)) = (J_A^F(\underline{\gamma}), J_A^D(\underline{\gamma})) \in Q_A^*$ . Donc  $Q_A^* \supset co(Q_A)$  et finalement

$Q_A^* = co(Q_A)$ . Montrons que  $Q_A^* = co(Q_A) \cap co(\bar{Q}_A)$  est convexe et contient  $Q_A$ , donc  $Q_A^* = co(Q_A) \subset co(\bar{Q}_A)$ . De même  $\bar{Q}_A \subset Q_A^*$ ,  $co(\bar{Q}_A) \subset co(Q_A^*) = Q_A^*$ , donc  $Q_A^* = co(\bar{Q}_A)$ .

(iii) L'enveloppe convexe d'un compact de  $\mathbb{R}^2$  étant compacte, il suffit de montrer que  $\bar{Q}_A$  est compact. On raisonne sur l'arborescence tandem à deux décideurs afin d'alléger les calculs. Faisons apparaître les règles de décision sous leur forme comportementale et écrivons  $J_A^F(\underline{\gamma})$  et  $J_A^D(\underline{\gamma})$ . En utilisant l'indépendance des observations  $Y_0$  et  $Y_1$  sous  $\mathbb{P}_0$  et  $\mathbb{P}_1$  il vient (on note  $\phi^0(u_1, y_0) = \phi_1^0(u_1, y_0)$ )

$$\begin{aligned} \mathbb{P}_0[U_0 = 1] &= \mathbb{E}_0 [\mathbb{P}_0([U_0 = 1] | U_1, Y_0)] \\ &= \mathbb{E}_0 \left[ \sum_{u_1 \in D_1} \mathbf{1}_{[U_1=u_1]} \cdot \mathbb{P}_0([U_0 = 1] | U_1 = u_1, Y_0) \right] \\ &= \sum_{u_1 \in D_1} \mathbb{E}_0 [\phi^0(u_1, Y_0)] \cdot \mathbb{E}_0 [\mathbb{P}_0([U_1 = u_1] | Y_1)] \\ &= \sum_{u_1 \in D_1} \mathbb{E}_0 [\phi^0(u_1, Y_0)] \cdot \mathbb{E}_0 [\phi_{u_1}^1(Y_1)]. \end{aligned} \quad (3)$$

On montrerait de même que

$$\mathbb{P}_1[U_0 = 1] = \sum_{u_1 \in D_1} \mathbb{E}_1 [\phi^0(u_1, Y_0)] \cdot \mathbb{E}_1 [\phi_{u_1}^1(Y_1)]. \quad (4)$$

On tire de (3)-(4) l'écriture d'un point  $q \in \bar{Q}_A$  à partir d'un  $(m_1 + 1)$ -uplet de probabilités conditionnelles  $\phi = (\phi^0; \phi_1^1, \dots, \phi_{m_1}^1)$ . Soit  $(q^n)_{n \geq 0}$  une suite de points de  $\bar{Q}_A$  et  $(\phi^n)_{n \geq 0}$  la suite de règles de décision associée. On va montrer qu'il est possible d'extraire de  $(q^n)_{n \geq 0}$  une sous-suite  $(q^{\varphi(n)})_{n \geq 0}$  convergente. Le théorème de compacité faible de la boule unité de  $L_\infty(Z_1, \mathcal{Z}_1, \mu^1)$ , [Leh86] Th.3 p.576, permet d'extraire, à l'aide d'un procédé diagonal fini, une sous-suite de  $(\phi^{1,n})_{n \geq 0}$ , soit  $(\phi^{1,\varphi_1(n)})_{n \geq 0}$ , convergeant faiblement vers un élément  $\phi^{1,*} = (\phi_1^{1,*}, \dots, \phi_{m_1}^{1,*})$ .  $\phi^{1,*}$  est un  $m_1$ -uplet de fonctions mesurables à valeurs dans  $[0, 1]$  et de somme unité. De même, on peut extraire de  $(\phi^{0,\varphi_1(n)})_{n \geq 0}$  une sous-suite  $(\phi^{0,\varphi(n)})_{n \geq 0}$ , où  $\varphi(n) = \varphi_0(\varphi_1(n))$ , qui converge faiblement vers  $\phi^{0,*} \in L_\infty(Z_0, \mathcal{Z}_0, \mu^0)$ . La suite  $(\tilde{q}^n)_{n \geq 0}$  associée aux probabilités conditionnelles  $(\tilde{\phi}^n)_{n \geq 0}$ , où  $\tilde{\phi}^n = \phi^{\varphi(n)}$ , converge donc vers  $q^* = (J_A^{F,*}, J_A^{D,*})$ , où  $J_A^{F,*} = \lim_n \sum_{u_1 \in D_1} \mathbb{E}_0 [\tilde{\phi}^{0,n}(u_1, Y_0)] \cdot \mathbb{E}_0 [\tilde{\phi}_{u_1}^{1,n}(Y_1)] = \sum_{u_1 \in D_1} \mathbb{E}_0 [\phi^{0,*}(u_1, Y_0)] \cdot \mathbb{E}_0 [\phi_{u_1}^{1,*}(Y_1)]$  et  $J_A^{D,*}$  est obtenu en remplaçant  $\mathbb{E}_0$  par  $\mathbb{E}_1$ . On a donc la compacité de  $\bar{Q}_A$  pour l'architecture tandem à deux décideurs. Pour  $A$  quelconque, l'indépendance des observations et le mécanisme de randomisation indépendante permettent toujours de développer  $J_A^F(\underline{\gamma})$  et  $J_A^D(\underline{\gamma})$  sous une forme algébrique ne faisant intervenir que des sommes et des produits des espérances, sous  $\mathbb{P}_0$  et  $\mathbb{P}_1$ , des diverses probabilités conditionnelles, [Bal94]. Le raisonnement précédent se généralise donc sans problème et on voit donc que  $\bar{Q}_A$  est compact ainsi que  $Q_A^*$ .  $\square$

4.  $co(Q)$  désigne l'enveloppe convexe de l'ensemble  $Q$ .

On doit principalement retenir de la proposition 3 que le véritable analogue du diagramme des risques centralisé est l'ensemble  $Q_A^*$ , et non  $\bar{Q}_A$ .  $Q_A^*$  est bien convexe, fermé, borné, inclus dans  $[0, 1]^2$ , contenant les points  $(0, 0)$  et  $(1, 1)$  et symétrique par rapport au point  $(1/2, 1/2)$  (il suffit d'échanger  $\phi_0^0$  et  $\phi_1^0$  sur  $DM_0$ ). Par contre  $\bar{Q}_A$  est un « pseudo » diagramme des risques : on sait, par le tracé numérique de certains exemples, que  $\bar{Q}_A$  peut ne pas être convexe. Ce fait n'est pas étonnant puisque le principe de randomisation général est celui appliqué dans  $\Gamma_A^*$ , alors que  $\bar{\Gamma}_A$  est spécifique au problème de D.D.. D'autre part, la preuve du point (iii) de la proposition 3 suggère déjà (c'était prévisible) que la recherche de solutions optimales pour le critère de N.-P. sera plus commode sur  $\bar{\Gamma}_A$  plutôt que sur  $\Gamma_A^*$ . Les remarques suivantes énoncent les questions qui se posent naturellement à la suite de la proposition 3.

**Rq 3.1** Sachant que  $\bar{Q}_A$  n'est pas nécessairement convexe, peut-on caractériser les diagrammes  $\bar{Q}_A$  qui sont convexes et ceux qui ne le sont pas (en fonction des lois  $\mathbb{P}_0$  et  $\mathbb{P}_1$  et de l'arborescence  $A$ ) ? A défaut d'un tel résultat, peut-on exhiber une famille de lois conduisant à un diagramme  $\bar{Q}_A$  convexe (ou non convexe), ou même seulement démontrer, dans un cas précis, qu'il y a (ou non) convexité ? Il est clair que, si  $\bar{Q}_A$  est convexe, on peut assurer que le mécanisme de randomisation indépendante est optimal pour le problème (1) (i.e.  $Q_A^* = \bar{Q}_A$ ).

**Rq 3.2** Existe-t-il, comme dans le problème centralisé, une propriété relative à la sous-optimalité de la randomisation lorsque les lois des r.d.v. des observations sont continues ?

Nous renvoyons le lecteur à [WW92, Bal94] pour la dernière remarque. On montre effectivement que toute randomisation indépendante est inutile dès lors que les lois des r.d.v. des observations propres sont toutes continues (par rapport à la mesure de Lebesgue). En particulier, on démontre pour le cas de l'architecture en « fusion de données » que le « fusionneur »  $DM_0$  n'a pas besoin de mettre en œuvre une règle de décision randomisée, bien que son observation  $(U_1, \dots, U_{N-1})$  soit discrète. Par contre, la première remarque demeure actuellement sans réponse satisfaisante (hormis l'exhibition d'exemples numériques sur lesquels on voit, par exemple, la perte de convexité de  $\bar{Q}_A$ ). Nous démontrons pour l'instant un lemme important, qui va nous servir dans toute la suite.

## 4. Lemme fondamental de Neyman-Pearson en D.D.

Nous traitons d'abord la version du lemme faisant intervenir des contraintes de taille (contraintes d'égalité) sur la règle de décision. Puis on présentera la version généralisée du lemme, avec des contraintes de niveau (contraintes d'inégalité).

### 4.1. LEMME À CONTRAINTES D'ÉGALITÉ

Ce lemme constitue une généralisation possible du lemme classique, [NP33], dans un contexte multidécisionnel. On considère une règle de décision  $\phi$  produisant la décision  $U$ , à valeurs dans  $D = \{1, \dots, m\}$ ,  $m \geq 2$ , à partir d'une observation  $Y$ , v.a. à valeur dans  $Z$ , de loi  $P_\theta$ ,  $\theta \in T = \{0, 1\}$ .  $\phi$  sera donnée sous forme comportementale, c.-à-d. sous la forme d'un  $m$ -uplet  $\phi = (\phi_1, \dots, \phi_m)$  où  $\phi_i : Z \rightarrow [0, 1]$  est mesurable et où  $\sum_{i=1}^m \phi_i = 1$   $\mu$ -p.p..  $\mu$  est ici une mesure dominante  $\sigma$ -finie sur  $Z$ , avec  $P_\theta = f_\theta \cdot \mu$ ,  $\theta = 0, 1$ .  $\phi_i(y)$  représente la probabilité de décider  $U = i$  conditionnellement à l'observation  $Y = y$ . On note  $\Phi$  l'ensemble des règles comportementales  $\phi$ .  $\mathbb{E}_\theta$  est l'espérance mathématique pour la loi  $P_\theta$ ,  $\mathbb{E}_\theta[\phi] = (\mathbb{E}_\theta[\phi_1], \dots, \mathbb{E}_\theta[\phi_m])$ ,  $\theta = 0, 1$ . Soit  $c = (c_1, \dots, c_m) \in \mathbb{R}^m$  et  $\alpha \in S_m$ , où  $S_m = \{x \in \mathbb{R}_+^m; \sum_{i=1}^m x_i = 1\}$  est le simplexe plein standard de  $\mathbb{R}^m$ . On peut définir la fonctionnelle linéaire  $F : \Phi \rightarrow \mathbb{R}$  (qui généralise le critère de puissance du lemme classique) par<sup>5</sup>  $F(\phi) = c' \cdot \mathbb{E}_1[\phi]$  et considérer la contrainte  $\mathbb{E}_0[\phi] = \alpha$  (qui généralise la contrainte de taille). Alors, une première version généralisée du lemme classique s'intéresse à la résolution de

$$\max_{\{\phi \in \Phi; \mathbb{E}_0[\phi] = \alpha\}} \{c' \cdot \mathbb{E}_1[\phi]\}. \quad (5)$$

Pour traiter (5) dans le cas général, on va d'abord supposer que  $c_1 < c_2 < \dots < c_m$ . Dans ce cadre, le lemme central ci-dessous va montrer que les règles de décision optimales pour (5) sont celles possédant la *Structure de Neyman Généralisée monotone*<sup>6</sup> suivante.

**Définition 4.1** Soit  $\phi = (\phi_1, \dots, \phi_m) \in \Phi$  (considérer  $m \geq 2$  quelconque dans cette définition). On dit que  $\phi$  possède la structure de Neyman généralisée monotone d'ordre  $m$  si il existe une famille  $t = (t_1, \dots, t_{m-1})$  de seuils,  $0 \leq t_1 \leq \dots \leq t_{m-1} \leq \infty$  tels que, pour  $\mu$ -presque tout  $y \in Z$  on ait

$$\phi_1(y) = \begin{cases} 1 & \text{si } f_1(y) < t_1 \cdot f_0(y), \\ 0 & \text{si } f_1(y) > t_1 \cdot f_0(y), \end{cases}$$

$$\phi_i(y) = \begin{cases} 0 & \text{si } f_1(y) < t_{i-1} \cdot f_0(y), \\ 1 & \text{si } t_{i-1} \cdot f_0(y) < f_1(y) < t_i \cdot f_0(y), \\ 0 & \text{si } f_1(y) > t_i \cdot f_0(y), \end{cases} \quad i = 2, \dots, m-1,$$

$$\phi_m(y) = \begin{cases} 0 & \text{si } f_1(y) < t_{m-1} \cdot f_0(y), \\ 1 & \text{si } f_1(y) > t_{m-1} \cdot f_0(y). \end{cases}$$

On note  $\Phi^*$  l'ensemble des règles  $\phi \in \Phi$  ayant la S.N.G. monotone et  $\Phi_\alpha^* = \{\phi \in \Phi^*; \mathbb{E}_0[\phi] = \alpha\}$ , pour un

5. On note  $x' \cdot y$  le produit scalaire usuel de  $\mathbb{R}^m$ .

6. S.N.G. monotone en abrégé.

$\alpha \in S_m$ . Une règle  $\phi$  telle que  $\mathbb{E}_0[\phi] = \alpha$  (ensemble noté  $\Phi_\alpha$ ) est dite de *taille*  $\alpha$ . La *puissance* de  $\phi$  est la quantité  $F(\phi)$ .

**Lemme 4.1** Soit à résoudre, dans le cadre énoncé ci-dessus, le problème (5), où on suppose que  $c_1 < c_2 < \dots < c_m$ . Alors,

(i) Existence.  $\Phi_\alpha^* \neq \emptyset$  : il existe une règle  $\phi$  de taille  $\alpha$  ayant la S.N.G. monotone,

(ii) C.N.S..  $\phi$  est solution optimale de (5) ssi  $\phi \in \Phi_\alpha^*$ .

**Preuve :** La preuve repose essentiellement sur une correspondance biunivoque entre  $\Phi_\alpha$  et  $\Psi_\alpha$  et entre  $\Phi_\alpha^*$  et  $\Psi_\alpha^*$ , où  $\Psi_\alpha$  et  $\Psi_\alpha^*$  renferment des fonctions  $\psi$  sur lesquelles on va utiliser les techniques du lemme classique. Plus précisément,  $\Psi$  est l'ensemble des  $(m-1)$ -uplets  $\psi = (\psi_1, \dots, \psi_{m-1})$ , où  $\psi_j : Z \rightarrow [0, 1]$  est mesurable,  $j = 1, \dots, m-1$ .  $\Psi^*$  est l'analogie de  $\Phi^*$ , c.-à-d. l'ensemble des  $\psi \in \Psi$  pour lesquelles il existe une famille  $t = (t_1, \dots, t_{m-1})$  de seuils  $0 \leq t_1 \leq \dots \leq t_{m-1} \leq \infty$  tels que, pour  $\mu$ -presque tout  $y \in Z$  et pour  $j = 1, \dots, m-1$  on ait

$$\psi_j(y) = \begin{cases} 0 & \text{si } f_1(y) < t_j \cdot f_0(y), \\ 1 & \text{si } f_1(y) > t_j \cdot f_0(y). \end{cases}$$

Enfin, on pose  $a = (a_1, \dots, a_{m-1})$ , où  $a_j = \sum_{i=j+1}^m \alpha_i$ .  $\Psi_\alpha$  et  $\Psi_\alpha^*$  sont définis de façon analogue à  $\Phi_\alpha$  et  $\Phi_\alpha^*$ . Définissons une correspondance  $f$  de  $\Phi$  sur  $\Psi$  par  $\psi = f(\phi) = (\psi_1, \dots, \psi_{m-1})$  où  $\psi_j = \sum_{i=j+1}^m \phi_i$ ,  $j = 1, \dots, m-1$ . Il n'est pas difficile (mais long et fastidieux) de vérifier que  $f$  réalise une bijection de  $\Phi_\alpha^*$  sur  $\Psi_\alpha^*$ .  $f^{-1}$  est donnée par  $\phi = f^{-1}(\psi) = (\phi_1, \dots, \phi_m)$  où  $\phi_m = \psi_{m-1}$ ,  $\phi_{i+1} = \psi_i - \psi_{i+1}$ ,  $i = m-2, \dots, 1$  et  $\phi_1 = 1 - \psi_1$ .

(i) Existence. Pour montrer que  $\Phi_\alpha^* \neq \emptyset$  il suffit de montrer que  $\Psi_\alpha^* \neq \emptyset$ . Soit  $\alpha : \mathbb{R} \rightarrow [0, 1]$ ,  $\alpha(t) = P_0[f_1 > t \cdot f_0]$ .  $1 - \alpha(t)$  est la fonction de répartition, sous  $P_0$ , du r.d.v.  $f_1/f_0$ , définie  $P_0$ -p.s. Donc  $\alpha$  est continue à droite, décroissante,  $\alpha(x) = 1$  si  $x < 0$ ,  $\alpha(\infty) = 0$  et  $\alpha(t-) - \alpha(t) = P_0[f_1 = t \cdot f_0]$ . Construisons la famille  $t = (t_1, \dots, t_{m-1})$  de seuils associée à  $\Psi$ , en partant de  $t_{m-1}$  jusqu'à  $t_1$  et de sorte que  $\alpha(t_j) \leq a_j \leq \alpha(t_j-)$ . Ce seuil existe toujours et il est unique pourvu qu'on prenne le plus petit en posant  $t_j = \inf\{t : \alpha(t) \leq a_j \leq \alpha(t-)\}$ . Posons alors, pour tout  $y \in Z$

$$\psi_j(y) = \begin{cases} 0 & \text{si } f_1(y) < t_j \cdot f_0(y), \\ g_j = \frac{a_j - \alpha(t_j)}{\alpha(t_j-) - \alpha(t_j)} & \text{si } f_1(y) = t_j \cdot f_0(y), \\ 1 & \text{si } f_1(y) > t_j \cdot f_0(y), \end{cases} \quad j = m-1, \dots, 1. \quad (6)$$

Dans (6), l'expression médiane a toujours un sens, car si  $\alpha(t_j-) = \alpha(t_j)$  alors  $P_0[f_1 = t_j \cdot f_0] = 0$ . Or  $P_0[f_1 = t_j \cdot f_0] = 0 \Rightarrow \mu(\{f_1/f_0 = t_j\} \cap \{f_0 > 0\}) = 0 \Rightarrow P_1[f_1 = t_j \cdot f_0] = 0$  et donc l'événement  $\{f_1 = t_j \cdot f_0\}$  est dans ce cas  $\mu$ -p.s. impossible. On a par construction  $0 \leq t_1 \leq \dots \leq t_{m-1} \leq \infty$  car

7. À ne pas confondre avec la taille  $\alpha$ .

8. On note  $\alpha(t-) = \lim_{x \rightarrow t} \alpha(x)$ .

$0 \leq a_{m-1} \leq \dots \leq a_1$  et  $\psi$  ainsi construite est un élément de  $\Psi^*$ . De plus  $\psi \in \Psi_\alpha^*$  car

$$\begin{aligned} \mathbb{E}_0[\psi_j] &= P_0[f_1 > t_j \cdot f_0] + \frac{a_j - \alpha(t_j)}{P_0[f_1 = t_j \cdot f_0]} \cdot P_0[f_1 = t_j \cdot f_0] \\ &= \alpha(t_j) + a_j - \alpha(t_j) \\ &= a_j, \quad j = 1, \dots, m-1. \end{aligned}$$

Enfin,  $\Psi_\alpha^* \neq \emptyset$ , donc  $\Phi_\alpha^* \neq \emptyset$ , ce qui prouve l'existence.

(ii) Condition suffisante. Soit  $\phi \in \Phi_\alpha^*$  et  $\phi^* \in \Phi_\alpha$ . On va montrer que  $F(\phi) \geq F(\phi^*)$ . Exprimons  $F(\phi)$  à partir de  $\psi = f(\phi)$ . Il vient

$$\begin{aligned} F(\phi) &= \sum_{i=1}^m c_i \cdot \mathbb{E}_1[\phi_i] \\ &= c_1 \cdot (1 - \mathbb{E}_1[\psi_1]) + \sum_{i=2}^{m-1} c_i \cdot (\mathbb{E}_1[\psi_{i-1}] - \mathbb{E}_1[\psi_i]) + c_m \cdot \psi_{m-1} \\ &= c_1 + (c_2 - c_1) \cdot \mathbb{E}_1[\psi_1] + \dots + (c_{j+1} - c_j) \cdot \mathbb{E}_1[\psi_j] \\ &\quad + \dots + (c_m - c_{m-1}) \cdot \mathbb{E}_1[\psi_{m-1}] \\ &= c_1 + \sum_{j=1}^{m-1} (c_{j+1} - c_j) \cdot \mathbb{E}_1[\psi_j]. \end{aligned} \quad (7)$$

Donc, si  $\psi^* = f(\phi^*)$ ,

$$F(\phi) - F(\phi^*) = \sum_{j=1}^{m-1} (c_{j+1} - c_j) \cdot (\mathbb{E}_1[\psi_j] - \mathbb{E}_1[\psi_j^*]). \quad (8)$$

Comme, par hypothèse,  $c_{j+1} - c_j > 0$  il suffit de montrer, avec la technique du lemme ordinaire que  $\mathbb{E}_1[\psi_j] - \mathbb{E}_1[\psi_j^*] \geq 0$ ,  $j = 1, \dots, m-1$ . Posons  $S_j^+ = [\psi_j - \psi_j^* > 0]$  et  $S_j^- = [\psi_j - \psi_j^* < 0]$ . Sur  $S_j^+$  on a  $\psi_j > 0$ , donc  $f_1(y) \geq t_j \cdot f_0(y)$  et sur  $S_j^-$  on a  $\psi_j(y) < 1$ , donc  $f_1(y) \leq t_j \cdot f_0(y)$ . Par conséquent

$$\begin{aligned} &\int_Z (\psi_j - \psi_j^*) \cdot (f_1 - t_j \cdot f_0) d\mu \\ &= \int_{S_j^+ \cup S_j^-} (\psi_j - \psi_j^*) \cdot (f_1 - t_j \cdot f_0) d\mu \geq 0, \quad j = 1, \dots, m-1. \end{aligned} \quad (9)$$

En développant (4.1.) et en utilisant le fait que  $\mathbb{E}_0[\psi_j] = \mathbb{E}_0[\psi_j^*] = a_j$  il vient  $\mathbb{E}_1[\psi_j] - \mathbb{E}_1[\psi_j^*] \geq 0$ ,  $j = 1, \dots, m-1$ . Revenant à (8), on voit que  $F(\phi) \geq F(\phi^*)$ , ce qu'il fallait montrer.

**Condition nécessaire.** Soit  $\phi^*$  une solution optimale de (5). On peut choisir, grâce à (i), une règle  $\phi \in \Phi_\alpha^*$ . On va montrer, par l'absurde, que si  $\phi^*$  ne possède pas la S.N.G. monotone de  $\phi$ , alors  $F(\phi) > F(\phi^*)$ . Soit  $\psi^* = f(\phi^*)$  et  $\psi = f(\phi) \in \Psi_\alpha^*$ . Soit  $S_j^+$  et  $S_j^-$  les événements définis au point (ii) à l'aide des seuils  $t_j$  associés à  $\phi$ . On pose  $S_j = (S_j^+ \cup S_j^-) \cap \{f_1 \neq t_j \cdot f_0\}$ ,  $j =$

$$\phi_1(y) = \begin{cases} 1 & \text{si } f_1(y) < t_1 \cdot f_0(y), \\ \gamma_1 & \text{si } f_1(y) = t_1 \cdot f_0(y), \\ 0 & \text{si } f_1(y) > t_1 \cdot f_0(y), \end{cases}$$

$$\phi_i(y) = \begin{cases} 0 & \text{si } f_1(y) < t_{i-1} \cdot f_0(y), \\ 1 - \gamma_{i-1} & \text{si } f_1(y) = t_{i-1} \cdot f_0(y), \\ 1 & \text{si } t_{i-1} \cdot f_0(y) < f_1(y) < t_i \cdot f_0(y), \\ \gamma_i & \text{si } f_1(y) = t_i \cdot f_0(y), \\ 0 & \text{si } f_1(y) > t_i \cdot f_0(y), \end{cases} \quad i = 2, \dots, m-1,$$

$$\phi_m(y) = \begin{cases} 0 & \text{si } f_1(y) < t_{m-1} \cdot f_0(y), \\ 1 - \gamma_{m-1} & \text{si } f_1(y) = t_{m-1} \cdot f_0(y), \\ 1 & \text{si } f_1(y) > t_{m-1} \cdot f_0(y). \end{cases}$$

$1, \dots, m-1$ . Supposons qu'il existe un  $k \in \{1, \dots, m-1\}$  tel que  $\mu(S_k) > 0$ . Alors

$$\begin{aligned} & \int_Z (\psi_k - \psi_k^*) \cdot (f_1 - t_k \cdot f_0) d\mu \\ &= \int_{S_k^+ \cup S_k^-} (\psi_k - \psi_k^*) \cdot (f_1 - t_k \cdot f_0) d\mu \\ &= \int_{S_k} (\psi_k - \psi_k^*) \cdot (f_1 - t_k \cdot f_0) d\mu > 0. \end{aligned}$$

Donc

$$\mathbb{E}_1[\psi_k] - \mathbb{E}_1[\psi_k^*] > t_k \cdot (\mathbb{E}_0[\psi_k] - \mathbb{E}_0[\psi_k^*]) = 0. \quad (10)$$

En réutilisant (8), il vient, grâce à  $c_{j+1} - c_j > 0$

$$F(\phi) - F(\phi^*) = \sum_{j=1}^{m-1} (c_{j+1} - c_j) \cdot (\mathbb{E}_1[\psi_j] - \mathbb{E}_1[\psi_j^*]) > 0. \quad (11)$$

L'équation (11) est une contradiction, qui est levée si  $\mu(S_j) = 0$  pour  $j = 1, \dots, m-1$ , ce qui signifie que  $\psi_j^*$  ne peut être différente de  $\psi_j$  que sur les événements  $[f_1 = t_j \cdot f_0]$ . Ainsi  $\psi^*$  possède la structure de Neyman de  $\psi$ , ce qui équivaut à dire que  $\phi^* = f^{-1}(\psi^*)$  possède la S.N.G. monotone de  $\phi$ , et donc  $\phi^* \in \Phi_\alpha$ .  $\square$

**Rq 4.1** La preuve du lemme 4.1. montre qu'il est inutile, pour une règle de décision solution optimale du problème (5), de randomiser sur les événements  $[f_1 = t_j \cdot f_0]$  dès lors que ces derniers sont  $P_0$ -négligeables. C'est en particulier le cas si le r.d.v. de  $Y$  possède une loi continue par rapport à la mesure de Lebesgue de  $\mathbb{R}_+$ . D'autre part, la même preuve montre que l'optimalité est atteinte en choisissant une règle à S.N.G. monotone constante sur ces mêmes événements. La structure d'une telle règle découle directement de celle des fonctions  $\psi_j$  de la preuve, cf. (6) et s'écrit, si  $(\gamma_1, \dots, \gamma_m) = f^{-1}(g_1, \dots, g_{m-1})$  (voir formule centrée en haut de page).

Cette écriture signifie que toute « hésitation » inhérente au fonctionnement de la règle  $\phi$  ne peut se produire qu'entre deux décisions consécutives  $U = i$  et  $U = i+1$ . Pour toute observation

$y$  telle que  $f_1(y) = t_i \cdot f_0(y)$ , on tire aléatoirement la décision  $U$ , avec une probabilité  $\gamma_i$  pour  $U = i$  et  $1 - \gamma_i$  pour  $U = i+1$ . Il résulte de ce fait que la règle optimale peut être vue comme un mécanisme de randomisation sur un nombre fini de règles pures. Ce point sera utilisé plus loin, lorsque nous aurons abordé l'optimisation d'arborescences de D.D. au sens de N.-P.

**Rq 4.2** De la preuve du lemme 4.1. on tire un procédé de construction d'une règle  $\phi$  optimale pour (5). Il suffit pour cela de placer successivement sur la courbe C.O.R. associée à  $P_0$  et  $P_1$  les niveaux successifs  $a_{m-1}, \dots, a_2, a_1$ . On en déduit immédiatement la valeur optimale de  $F$  (à partir de l'équation (7)) et des seuils  $t_{m-1}, \dots, t_2, t_1$  (qui sont les pentes, ou les sous-gradients, aux points considérés). La fig. (2) illustre cette construction pour  $m = 3$ . Il est bon aussi de noter que ces seuils sont essentiellement uniques, comme c'est le cas pour le lemme classique.

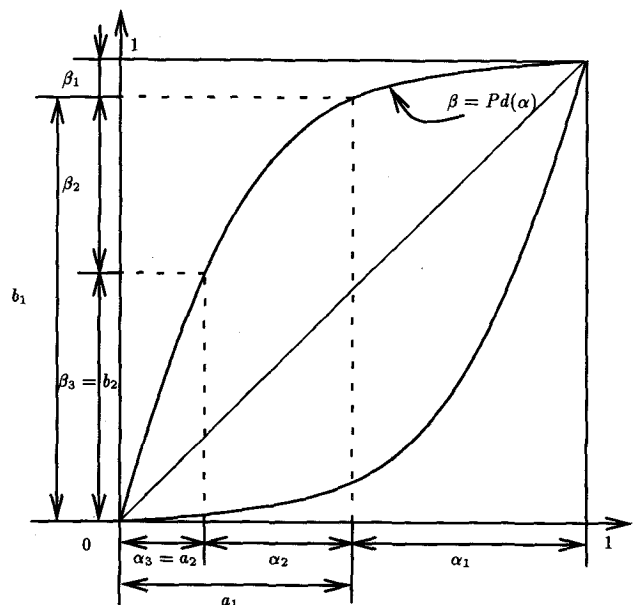


Figure 2. - Règle à S.N.G. monotone,  $m = 3$ .

**Rq 4.3** On vérifie sans peine que la S.N.G. monotone de la définition 4.1. est celle d'un *seuillage monotone du r.d.v. des observations*. Nous allons voir plus loin que cette structure caractérise les règles de D.D. optimales au sens de N.-P. Celles-ci possèdent donc une structure identique à celles qu'on obtient dans un cadre Bayésien à coût terminal, ce qui vient précisément du caractère terminal du critère de N.-P.

Il existe plusieurs généralisations du lemme 4.1.. On peut, dans un premier temps, s'intéresser au relachement de l'hypothèse  $c_1 < c_2 < \dots < c_m$ , c.-à-d. supposer que les  $c_i$  ne sont ni ordonnés ni deux à deux distincts. Quand les  $c_i$  ne sont plus ordonnés, on devine facilement que cela revient à abandonner le caractère monotone de la définition 4.1., en faisant intervenir la permutation<sup>9</sup>  $\sigma$  de  $D$  qui ordonne les  $c_i$ , c.-à-d. telle que  $c_\sigma = (c_{\sigma(1)}, \dots, c_{\sigma(m)})$  vérifie  $c_{\sigma(1)} \leq \dots \leq c_{\sigma(m)}$ . Ensuite, on peut rassembler les termes égaux de  $c_\sigma$  à l'aide d'une partition  $\pi$  de  $\{1, \dots, m\}$  en intervalles disjoints consécutifs, soit  $\pi = \{I_1, \dots, I_{m'}\}$ , où  $m' \leq m$ . Noter ici que  $m' = 1$  est interdit (tous les  $c_i$  sont égaux), car alors (5) devient trivial. Si  $X = X_1 \times \dots \times X_m$  est un ensemble produit (où les  $X_i$  sont supposés être isomorphes à un ensemble  $E$  muni d'une addition interne), on peut choisir des représentants  $i_1, i_2, \dots, i_{m'} \in \{1, \dots, m\}$  des classes  $I_1, \dots, I_{m'}$  et poser  $X'_1 = X_{i_1}, \dots, X'_{m'} = X_{i_{m'}}$ . Alors, on pourra faire intervenir une opération  $\Pi$  associée à  $\pi$ ,  $\Pi : X = X_1 \times \dots \times X_m \rightarrow X'_1 \times \dots \times X'_{m'} = X'$  qui associe à tout  $m$ -uplet  $x = (x_1, \dots, x_m)$  un  $m'$ -uplet  $x' \in X'$  tel que  $x' = \Pi(x) = (x'_1, \dots, x'_{m'})$  où  $x'_j = \sum_{i \in I_j} x_i$ ,  $j = 1, \dots, m'$ . On peut résumer l'usage de  $\sigma$  et  $\pi$  sur l'exemple suivant. Si  $m = 3, c = (c_1, c_2, c_3) = (3, 1, 3)$  et  $\alpha = (3/8, 1/4, 3/8)$ , alors  $\sigma(1) = 2, \sigma(2) = 3, \sigma(3) = 1; c_\sigma = (1, 3, 3); \pi = \{I_1, I_2\}$  où  $I_1 = \{1\}$  et  $I_2 = \{2, 3\}, m' = 2$  et  $\Pi(\alpha_\sigma) = \Pi(1/4, 3/8, 3/8) = (1/4, 3/4)$ . Noter ici que le résultat final est indépendant de l'ordre dans lequel on fait agir les opérations d'ordonnement et de partitionnement. Nous pouvons donc nous ramener au cas du lemme 4.1., ce qui permet d'étendre la définition de la structure de Neyman généralisée de la façon suivante.

**Définition 4.2** Soit  $\phi = (\phi_1, \dots, \phi_m) \in \Phi$  une règle de décision donnée sous forme comportementale. On dit que  $\phi$  possède la structure de Neyman généralisée si il existe une permutation  $\sigma$  de  $D = \{1, \dots, m\}$  et une partition  $\pi$  de  $D$  tels que  $\phi' = \Pi(\phi_\sigma)$  soit une règle à structure de Neyman généralisée monotone d'ordre<sup>10</sup>  $m' = |\pi|$ .

On peut maintenant donner la généralisation du lemme 4.1..

**Proposition 4.1** Soit à résoudre le problème (5) où  $c = (c_1, \dots, c_m) \in \mathbb{R}^m$  possède au moins deux termes distincts. Alors, si  $\sigma$  est une permutation de  $D = \{1, \dots, m\}$  qui ordonne les  $c_i$  et  $\pi$  la partition de  $D$  en  $m'$  intervalles disjoints consécutifs

associée à la relation d'égalité sur les  $c_{\sigma(i)}$ ,

(i) existence : il existe une règle  $\phi$  à structure de Neyman généralisée construite à partir de  $\sigma, \pi$  et de taille  $\alpha$ ,

(ii) C.N.S. :  $\phi$  est solution optimale de (5) ssi elle possède la structure de Neyman généralisée associée à  $\sigma, \pi$  et est de taille  $\alpha$ .

**Preuve :** (i) Il suffit de remarquer que  $F(\phi) = \sum_{i=1}^m c_{\sigma(i)} \cdot \mathbb{E}_1[\phi_{\sigma(i)}] = \sum_{j=1}^{m'} c'_j \cdot \mathbb{E}_1[\phi'_j] = F'(\phi')$  où  $\phi' = \Pi(\phi_\sigma)$  et où  $c' = (c_{i_1}, \dots, c_{i_{m'}})$  est construit à partir des représentants  $i_1, \dots, i_{m'}$  des classes  $I_1, \dots, I_{m'}$  de  $\pi$ . On est ainsi ramené aux conditions du lemme 4.1., mais avec une taille  $\alpha' = \Pi(\alpha_\sigma)$ . Le point (i) de ce lemme donne l'existence de  $\phi' \in \Phi_{\alpha'}^*$ , qui est à structure de Neyman généralisée monotone d'ordre  $m'$ . Pour montrer le point (i) de notre proposition il suffit de montrer qu'on peut toujours construire  $\phi$  à partir de  $\phi'$ , telle que  $\mathbb{E}_0[\phi] = \alpha$ . Il suffit pour cela d'utiliser un mécanisme de randomisation local. Si  $U'$  est la décision sur  $\{1, \dots, m'\}$  produite par  $\phi'$ , on produit  $U$  sur  $\{1, \dots, m\}$  avec  $\phi$  selon le tirage aléatoire suivant. Quand  $U' = j$ ,  $\phi$  produit  $U = i$ , où  $i \in \sigma(I_j)$  (les indices de  $I_j$  sont ceux associés à des termes de  $c_\sigma$  qui sont identiques), avec une probabilité  $\alpha_i/\alpha'_j$ , où  $\alpha'_j = \sum_{i \in \sigma(I_j)} \alpha_i$ . Ainsi  $\mathbb{E}_0[\phi_i] = \mathbb{P}_0[U = i] = \mathbb{P}_0([U = i] | [U' = j]) \cdot \mathbb{P}_0[U' = j] = (\alpha_i/\alpha'_j) \cdot \alpha'_j = \alpha_i$ ,  $i = 1, \dots, m$ . Avec l'exemple cité précédemment, si  $U' = 2$ , on tire  $U$  aléatoirement dans  $\sigma(I_2) = \{1, 3\}$ , soit  $U = 1$  avec une probabilité de  $\alpha_1/(\alpha_1 + \alpha_3)$  et  $U = 3$  avec une probabilité de  $\alpha_3/(\alpha_1 + \alpha_3)$ .

(ii) C.N. Soit  $\phi$  la solution optimale de (5). Elle est, par hypothèse, de taille  $\alpha$ . Soit  $\sigma$  la permutation qui ordonne les termes de  $c$  et  $\pi = \{I_1, \dots, I_{m'}\}$  la partition de  $\{1, \dots, m\}$  associée à la relation d'égalité sur les  $c_{\sigma(i)}$ . On a, par hypothèse,  $m' \geq 2$ . Posons  $\phi' = \Pi(\phi_\sigma)$  et  $\alpha' = \Pi(\alpha_\sigma)$ . Posons également  $c' = (c_{i_1}, \dots, c_{i_{m'}})$ , où  $i_1 \in I_1, \dots, i_{m'} \in I_{m'}$ . Alors on voit, avec l'écriture de  $F(\phi)$  du point (i) que  $\phi'$  est solution optimale du problème (5) associé à  $c'$  et  $\alpha'$  (raisonner par l'absurde et utiliser la technique de construction d'une règle de  $\Phi_\alpha$  à partir d'une règle de  $\Phi_{\alpha'}$ , du point (i)). Donc, par la C.N. du lemme 4.1,  $\phi'$  est à S.N.G. d'ordre  $m'$ .

C.S. Réciproquement, si  $\phi$  est de taille  $\alpha$  et si elle possède la S.N.G. c'est qu'il existe une permutation  $\sigma$  et une partition  $\pi$  telles que  $\phi' = \Pi(\phi_\sigma)$  soit à S.N.G. d'ordre  $m'$ . Donc, avec la C.S. du lemme 4.1,  $\phi'$  est solution optimale du problème (5) associé à  $c'$  et  $\alpha'$  (qui sont définis comme pour la C.N.). Or  $F'(\phi') = \sum_{j=1}^{m'} c'_j \cdot \mathbb{E}_1[\phi'_j] = \sum_{i=1}^m c_i \cdot \mathbb{E}_1[\phi_i] = F(\phi)$ . Par l'absurde, si  $\phi$  n'était pas solution optimale de (5), on trouverait  $\phi^* \in \Phi_\alpha$  telle que  $F(\phi^*) > F(\phi)$ , i.e. telle que  $F'(\Pi(\phi^*_\sigma)) > F'(\phi')$  et donc  $\phi'$  ne serait pas solution optimale de (5) (associée à  $c', \alpha'$ ).  $\square$

**Rq 4.4** La proposition 4.1. montre clairement que l'égalité de certains termes  $c_i$  revient simplement à diminuer le cardinal de l'alphabet sur lequel on code la décision. On verra plus loin que ce type de comportement est intuitivement sous-optimal dès qu'il correspond à la règle de décision d'un décideur qui transmet le

9. Ou une de ces permutations si certains termes sont égaux.

10.  $|A|$  désigne le cardinal de l'ensemble  $A$ .



meilleur résumé de son observation au décideur qui lui succède dans une arborescence de D.D..

La dernière généralisation du lemme 4.1. conduit à modifier la contrainte de taille en une contrainte de niveau.

## 4.2. LEMME À CONTRAINTE DE NIVEAU

Comme pour le lemme classique, on peut tenter de modifier la contrainte de taille du problème (5) en une contrainte de niveau. Il est alors naturel, pour  $h = (h_1, \dots, h_m)$ ,  $c = (c_1, \dots, c_m) \in \mathbb{R}^m$  et  $\alpha \in ]0, 1[$  de s'intéresser à

$$\max_{\{\phi \in \Phi; h' \cdot \mathbb{E}_0[\phi] \leq \alpha\}} \{c' \cdot \mathbb{E}_1[\phi]\}. \quad (12)$$

Pour étudier (12) il faut d'abord s'apercevoir que l'ensemble réalisable de ce problème peut être vide. Si on pose  $S_\alpha = \{\alpha_1 \in S_m; h' \cdot \alpha_1 \leq \alpha\}$ , on voit que (12) est sans solution dès que  $S_\alpha = \emptyset$ . Ensuite, il est clair qu'une solution optimale de (12) va nécessairement mettre en œuvre une solution du type de celles rencontrées dans l'étude du problème (5). Nous allons démontrer le résultat suivant, qui confirme et précise ces idées.

**Proposition 4.2** *Soit à résoudre le problème (12), où  $c, h \in \mathbb{R}^m$  et  $\alpha \in ]0, 1[$ . Alors (12) admet une solution optimale ssi  $S_\alpha \neq \emptyset$ , auquel cas toute solution optimale est une règle de décision  $\phi$  à structure de Neyman généralisée (au sens de la définition 4.2.).*

**Preuve :** Pour démontrer l'existence d'une solution optimale à (12) ssi  $S_\alpha \neq \emptyset$ , on va associer à (12) un second problème qui lui est équivalent. Soit

$$F_\alpha : S_m \rightarrow \mathbb{R} \\ \alpha' \mapsto F_{\alpha'} = \max_{\{\phi \in \Phi; \mathbb{E}_0[\phi] = \alpha'\}} \{c' \cdot \mathbb{E}_1[\phi]\}. \quad (13)$$

Il est facile de montrer que (12) et (14)

$$\max_{\{\alpha' \in S_\alpha\}} \{F_{\alpha'}\} \quad (14)$$

est bien définie et on sait, grâce au lemme 4.1., qu'elle est même continue (grâce à la continuité de la courbe C.O.R.). De plus,  $S_\alpha$  est un compact de  $\mathbb{R}^m$ , donc (14) (et donc (12)) admet une solution ssi  $S_\alpha \neq \emptyset$ . La deuxième partie est une conséquence du lemme 4.1.. Toute solution  $\phi^*$  de (12) est associée à une solution  $\alpha^*$  de (14), laquelle correspond, par (13), à une règle  $\phi^*$ , qui est à S.N.G. (lemme 4.1., (iii)).  $\square$

La proposition 4.2. va nous servir au paragraphe suivant, pour montrer que les règles optimales au sens de N.-P. possèdent la S.N.G.. On doit remarquer ici qu'aucun argument évident ne permet d'affirmer que l'optimum de (12) est atteint au bord de la contrainte  $h' \cdot \mathbb{E}_0[\phi] \leq \alpha$ . Notons aussi, en revenant à la preuve du lemme 4.1., que le problème (12) peut se formuler comme un problème d'optimisation non linéaire sous contraintes linéaires. Il suffit pour cela de prendre la forme équivalente (14) dans laquelle

$F_{\alpha'}$  est exprimée à l'aide de l'équation (7). En reprenant les notations déjà utilisées au § 4.1., le lecteur vérifiera qu'on peut mettre (12) sous la forme

$$\max\{\sum_{j=1}^{m-1} [(c_{j+1} - c_j) \cdot Pd(a_j)]\} \\ \text{sous } \begin{cases} 0 \leq a_{m-1} \leq \dots \leq a_1 \leq 1, \\ \sum_{j=1}^{m-1} (h_{j+1} - h_j) \cdot a_j \leq \alpha. \end{cases} \quad (15)$$

On a supposé, dans (15), pour simplifier l'écriture, que  $c_1 < c_2 < \dots < c_m$ . Ce problème n'admet pas de solution explicite évidente, mais on voit que la courbe C.O.R.  $Pd : [0, 1] \rightarrow [0, 1]$  associée à  $(P_0, P_1)$  intervient directement. Il apparaît que les courbes C.O.R. « locales » (i.e. associées à chaque décideur d'une arborescence) sont des éléments essentiels pour la détermination des règles optimales, [TPK91, TPK93, Bal94]. Nous pouvons maintenant passer à l'optimisation concrète d'arborescences de D.D. au sens de N.-P. La nécessité de l'argumentation « personne par personne » dont on va se servir oblige de restreindre l'étude du problème général (1) au plus grand sous-ensemble séparable de  $\Gamma_A^*$ , à savoir  $\bar{\Gamma}_A$ .

## 5. Optimisation sur $\bar{\Gamma}_A$

Nous procédons par ordre de généralité croissante, d'abord sur une arborescence tandem à deux décideurs, puis sur une arborescence quelconque.

### 5.1. DEUX DÉCIDEURS EN TANDEM

On rappelle que l'hypothèse de randomisation indépendante permet de représenter la règle  $\gamma \in \bar{\Gamma}_A$  soit par les couples  $((\gamma_{i,j})_{i \geq 0}, I_j)$ ,  $j = 0, 1$ , soit sous la forme comportementale  $\phi = (\phi^0, \phi^1)$ . Fixons donc  $DM_0$  et recherchons la règle optimale utilisée par  $DM_1$ .

#### 5.1.1. Optimisation de $DM_1, DM_0$ fixé

Il est immédiat de voir que l'écriture de l'influence de  $DM_1$  sur la fonction objective et sur la contrainte du problème (1) correspond aux équations (3)-(4) qui peuvent s'écrire

$$J_A^D(\gamma) = \sum_{u_1 \in D_1} \mathbb{P}_1([U_0 = 1] | U_1 = u_1) \cdot \mathbb{E}_1[\phi_{u_1}^1(Y_1)] \quad (16)$$

$$J_A^F(\gamma) = \sum_{u_1 \in D_1} \mathbb{P}_0([U_0 = 1] | U_1 = u_1) \cdot \mathbb{E}_0[\phi_{u_1}^1(Y_1)] \quad (17)$$

En posant  $c_{u_1} = \mathbb{P}_1([U_0 = 1] | U_1 = u_1)$  et  $h_{u_1} = \mathbb{P}_0([U_0 = 1] | U_1 = u_1)$ ,  $u_1 \in D_1$ , on fait apparaître un problème qui est celui que nous avons traité au § 4.2., puisque les termes  $c_{u_1}$  et

$h_{u_1}$  sont indépendants de  $\phi^1$ . Plus précisément, toute règle  $\phi^1$  optimale pour  $DM_1$  à  $DM_0$  fixé est solution de

$$\max_{\{\phi^1; c_1^1. \mathbb{E}_0[\phi^1(Y_1)] \leq \alpha\}} [h_1^1. \mathbb{E}_1[\phi^1(Y_1)]] . \quad (18)$$

Mais (18) n'est autre que (12), qui est traité dans la proposition 4.2., adaptée à la situation de  $DM_1$ . Donc, on est assuré, si une solution  $\phi^1$  existe, qu'elle possède la S.N.G. (définition 4.2.). Passons à l'optimisation de  $DM_0$  quand  $DM_1$  est fixé.

### 5.1.2. Optimisation de $DM_0$ , $DM_1$ fixé

$DM_1$  étant fixé,  $DM_0$  reçoit l'observation  $(U_1, Y_0)$  de statistique fixée et est donc confronté à un test classique de N.-P.. Sa règle de décision est à structure de Neyman, qu'on peut voir comme un cas particulier de la S.N.G. de la définition 4.2.. Sans discuter en détail la structure des règles employées par  $DM_0$  et  $DM_1$  (en particulier la question de savoir si  $DM_1$  utilise tout son alphabet de codage  $D_1$ ), on peut en tout cas montrer que celles-ci peuvent toujours être choisies monotones. Le raisonnement est le même que celui déjà utilisé dans le cas Bayésien, [Bal94]. Si  $DM_1$  n'emploie pas une règle monotone, on peut toujours rendre celle-ci monotone par une permutation adéquate de  $D_1$ ,  $DM_0$  tenant compte de ce changement dans son « décodage » de  $U_1$ . L'optimalité de la nouvelle règle est bien conservée, puisque  $U_0$  n'est pas modifiée. On est donc assuré de garder des règles à S.N.G., cette fois-ci monotones. En convenant d'appeler  $\underline{\gamma} \in \bar{\Gamma}_A$  règle à S.N.G. monotone toute règle globale dont les composantes  $\gamma_0, \dots, \gamma_{N-1}$  sont à S.N.G. monotone (i.e. associés à des fonctions de décision  $\phi^i$  possédant la S.N.G. monotone), nous pouvons résumer l'ensemble de ces résultats de la façon suivante.

**Proposition 5.1** Soit  $A$  une arborescence de D.D. à deux décideurs en tandem. L'ensemble des solutions optimales du problème de N.-P. (1) associé à  $A$  et restreint à  $\bar{\Gamma}_A$  rencontre nécessairement l'ensemble des règles  $\underline{\gamma} = (\gamma_0, \gamma_1) \in \bar{\Gamma}_A$  qui sont à structure de Neyman généralisée monotone.

Nous voyons que les résultats qualitatifs de la proposition 5.1. révèlent, comme pour le cas Bayésien à coût terminal, l'importance des règles de décision à base de seuillage du r.d.v. de l'observation complète. Avant de généraliser ce résultat au cas d'une arborescence quelconque, quelques remarques sont utiles.

**Rq 5.1** La deuxième partie de notre étude (optimisation de  $DM_0$  à  $DM_1$  fixé) montre que toute solution  $\underline{\gamma}$  de (1) vérifie  $J_A^F(\underline{\gamma}) = \alpha$ , à moins qu'il existe un test sur  $DM_0$  de taille inférieure à  $\alpha$  et de puissance unité. En particulier si les lois du couple  $(U_1, Y_0)$  sont équivalentes on est sûr d'avoir  $J_A^F(\underline{\gamma}) = \alpha$ . Cette remarque montre que la courbe C.O.R. associée à  $\bar{Q}_A$  est une fonction qui, à défaut d'être concave, est du moins (continue et) croissante.

**Rq 5.2** Il apparaît que les règles utilisées par  $DM_0$  et  $DM_1$  utilisent des seuils qui sont reliés par des équations de couplage inter-décideurs. L'écriture de ces relations est compliquée dans le

cas de N.-P., à cause de la contrainte de fausse alarme et des problèmes de randomisation associés. Rappelons aussi que la résolution de ces équations correspond aux conditions nécessaires du principe d'optimalité « personne par personne » et ne conduit pas forcément à des optima locaux.

## 5.2. ARCHITECTURE ARBORESCENTE À N DÉCIDEURS

**Proposition 5.2** Soit  $A$  une arborescence de D.D. que l'on optimise au sens de N.-P. pour le problème (1) restreint à  $\bar{\Gamma}_A$ .

- (i) Il existe une règle  $\underline{\gamma} \in \bar{\Gamma}_A$  optimale pour ce problème,
- (ii) toute règle  $\underline{\gamma} \in \bar{\Gamma}_A$  optimale pour ce problème est à structure de Neyman généralisée,
- (iii) toute règle optimale peut être choisie à S.N.G. monotone.

**Preuve :** (i) L'existence découle de la compacité de  $\bar{Q}_A$  (proposition 3.1.).

(ii) Soit  $\underline{\gamma} \in \bar{\Gamma}_A$  une règle optimale pour (1). On applique le principe d'optimalité « personne par personne » à un décideur  $DM_i \in DM$ . Pour  $i = 0$  le résultat est évident,  $DM_0$  effectuant un test classique de N.-P.. Si  $i \neq 0$ , on écrit

$$\begin{aligned} J_A^F(\underline{\gamma}) &= \mathbb{P}_0[U_0 = 1] \\ &= \mathbb{E}_0[\mathbb{P}_0(\{U_0 = 1\} | U_i)] \\ &= \mathbb{E}_0 \left[ \sum_{u_i \in D_i} \mathbb{1}_{[U_i = u_i]} \cdot \mathbb{P}_0(\{U_0 = 1\} | U_i = u_i) \right] \\ &= \sum_{u_i \in D_i} \mathbb{P}_0(\{U_0 = 1\} | U_i = u_i) \cdot \mathbb{E}_0[\mathbb{P}_0(\{U_i = u_i\} | Y_i, U^i)] \\ &= \sum_{u_i \in D_i} \mathbb{P}_0(\{U_0 = 1\} | U_i = u_i) \cdot \mathbb{E}_0[\phi_{u_i}^i(Y_i, U^i)] . \end{aligned} \quad (19)$$

De la même façon

$$J_A^D(\underline{\gamma}) = \sum_{u_i \in D_i} \mathbb{P}_1(\{U_0 = 1\} | U_i = u_i) \cdot \mathbb{E}_1[\phi_{u_i}^i(Y_i, U^i)] . \quad (20)$$

Posons  $h_i = (h_1, \dots, h_{m_i})$  et  $c_i = (c_1, \dots, c_{m_i})$  où  $h_{u_i} = \mathbb{P}_0(\{U_0 = 1\} | U_i = u_i)$  et  $c_{u_i} = \mathbb{P}_1(\{U_0 = 1\} | U_i = u_i)$ .  $h_i$  et  $c_i$  sont indépendants de  $\gamma_i$ . On retrouve donc, avec la maximisation de (20) sous une contrainte de niveau  $\alpha$  pour (19) le problème (12) qui a été précédemment étudié (§ 4.2.). En appliquant la proposition 4.2., on voit que  $\gamma_i$  (ou  $\phi^i$ ) est nécessairement à S.N.G..

(iii) Pour obtenir le caractère monotone, on effectue une « récurrence spatiale finie », qui part des « feuilles » de l'arborescence (décideurs initiaux) et se dirige vers la racine  $DM_0$ . Au rang 0 de la récurrence, tous les décideurs initiaux transforment leur règle de décision de façon à la rendre monotone, à l'aide des permutations adéquates. Leurs successeurs immédiats, informés de cette opération, décodent la décision reçue en appliquant la permutation associée. De cette façon, l'information convoyée à  $DM_0$  n'est pas changée. Au rang  $n$  de la récurrence,  $DM_i$  ( $i \neq 0$ ) rend monotone sa règle de décision et informe son successeur de ce changement,

qui en tient compte dans sa propre règle. De proche en proche, on atteint ainsi toute l'arborescence  $A$ , toutes les règles (exceptée  $\gamma_0$ ) sont rendues monotones et l'optimalité est conservée.  $\gamma_0$  demeure donc optimale pour cette règle, ce qui implique sa monotonie, à cause de la monotonie des règles appliquées par les décideurs qui la précèdent.  $\square$

**Rq 5.3** Notons qu'une conséquence de la proposition 5.2. est que tout mécanisme de randomisation (indépendante, mais aussi d'équipe, comme on va le voir au § 6) peut se réaliser par utilisation d'un nombre fini de règles pures. Ceci découle de la S.N.G., du nombre fini de seuils qui les paramétrisent et de la remarque 4.1.

**Rq 5.4** Il est bon aussi de noter que la caractérisation des solutions optimales du problème de N.-P. sur l'ensemble  $\Gamma_A$  des règles pures n'a pas été abordée. La raison de ce fait est qu'on est tout simplement incapable de caractériser la structure de ces règles. On peut, il est vrai, rechercher les règles pures à S.N.G. qui sont optimales au sens de N.-P. Mais il faut bien voir qu'il ne s'agit là que d'un sous-ensemble de  $\Gamma_A$ , dont on ne sait pas démontrer l'optimalité. Pour se convaincre de cette difficulté, il faut repenser au problème centralisé. Dans le lemme classique de N.-P., il est faux de dire que les règles pures optimales sont à structure de Neyman. Par exemple, lorsque l'observation est discrète, l'ordre obtenu en classant les observations selon leur rapport de vraisemblance n'est pas forcément celui qu'il faut utiliser pour « seuiller » les observations. Par conséquent, les articles dans lesquels on observe des courbes C.O.R. décentralisées obtenues en choisissant des règles à S.N.G. pures ne sont associées au diagramme des risques  $Q_A$  que sous certaines hypothèses, par exemple celle assurant que  $Q_A = \bar{Q}_A$ .

On voit avec la proposition 5.2. que la question de la monotonie est indifférente au critère de N.-P., ce qui est normal compte tenu du caractère terminal de ce dernier. Par contre, il n'a pas été possible de préciser ici les questions concernant la dégénérescence possible des seuils, *i.e.* le fait que certains d'entre eux peuvent être égaux. Cependant, on pressent assez bien que les seuils utilisés par les décideurs doivent être tous distincts, sans quoi on perd la possibilité de transmettre une information plus « fine » au décideur suivant. L'égalité de deux seuils consécutifs dans une règle à S.N.G. revient en effet à diminuer d'une unité le cardinal de l'alphabet de codage. Nous passons maintenant à l'optimisation d'une arborescence de D.D. sur l'ensemble  $\Gamma_A^*$ .

## 6. Optimisation sur $\Gamma_A^*$

Comme on peut le prévoir, l'étude de (1) sur  $\Gamma_A^*$  ne va pas nous permettre d'améliorer la connaissance de la structure locale des règles de décision optimales au sens de N.-P. Cependant, on va voir que toute règle globale utilisant un mécanisme de

randomisation d'équipe peut être vue comme une randomisation globale sur les règles de  $\bar{\Gamma}_A$ .

**Proposition 6.1** Soit  $A$  une arborescence de D.D. qu'on optimise sur  $\Gamma_A^*$  au sens de N.-P. et soit  $\underline{\gamma} \in \Gamma_A^*$  une règle de décision optimale pour ce problème. Alors,

(i) Il existe une loi a priori  $\pi = \{\pi_0, \pi_1\}$  sur  $H = \{H_0, H_1\}$  telle que  $\underline{\gamma}$  soit optimale pour le problème Bayésien minimisant la probabilité d'erreur de  $DM_0$  quand  $\pi$  est la loi a priori,

(ii) il existe deux règles  $\underline{\gamma}^1, \underline{\gamma}^2 \in \bar{\Gamma}_A$  à S.N.G. monotone telles que  $\underline{\gamma} \in [\underline{\gamma}^1, \underline{\gamma}^2]$ . De plus  $\underline{\gamma}^1$  et  $\underline{\gamma}^2$  sont solutions optimales du problème Bayésien associé à  $\pi$ ,

(iii) si de plus le r.d.v. des observations possède des lois continues on peut toujours choisir  $\underline{\gamma}^1$  et  $\underline{\gamma}^2$  dans  $\Gamma_A$ .

**Preuve :** (i)  $Q_A^*$  est un ensemble compact de  $[0, 1]^2$  qui contient les points  $\{(q^F, q^D); q^F \in [0, 1]\}$ . Donc il existe un point  $(\alpha, q^{D,*})$  qui maximise la puissance  $q^D$  et qui correspond à une règle de décision de taille  $\alpha$ . Tout autre point  $(q^F, q^D)$  de niveau  $\alpha$  ( $q^F \leq \alpha$ ) est de puissance inférieure,  $q^D \leq q^{D,*}$ , sinon  $Q_A^*$  ne serait pas convexe. Donc  $(\alpha, q^{D,*})$  est une solution optimale sur  $\Gamma_A^*$  de (1) et  $(\alpha, q^{D,*}) = (J_A^F(\underline{\gamma}), J_A^D(\underline{\gamma}))$ ,  $\underline{\gamma} \in \Gamma_A^*$ . Comme  $Q_A^*$  est un convexe fermé et que<sup>11</sup>  $q^* = (\alpha, q^{D,*}) \in Fr(Q_A^*)$  on peut appliquer le théorème de l'hyperplan d'appui en  $q^*$  : il existe un vecteur  $\lambda = (\lambda^F, \lambda^D)$  tel que,  $\forall q = (q^F, q^D) \in Q_A^*$ ,  $q \cdot \lambda \geq q^* \cdot \lambda$ . Autrement dit,  $q^*$  minimise  $q \cdot \lambda$  sur  $Q_A^*$ .  $\lambda$  est le vecteur normal de l'hyperplan d'appui en  $q^*$ . Il pointe vers l'intérieur de  $Q_A^*$ . Comme  $q^* \neq (0, 0)$ ,  $q^* \neq (1, 1)$  et  $\{(q^F, q^D); q^F \in [0, 1]\} \subset Q_A^*$ , on a  $\lambda^F \geq 0$  et  $\lambda^D < 0$  (cf. fig. (3)). Posant  $\pi_0 = \lambda^F / (\lambda^F - \lambda^D)$  et  $\pi_1 = -\lambda^D / (\lambda^F - \lambda^D)$ , on voit que  $q^*$ , c.-à-d.  $\underline{\gamma}$ , minimise  $\pi_0 \cdot J_A^F + \pi_1 \cdot (-J_A^D)$  qui est le

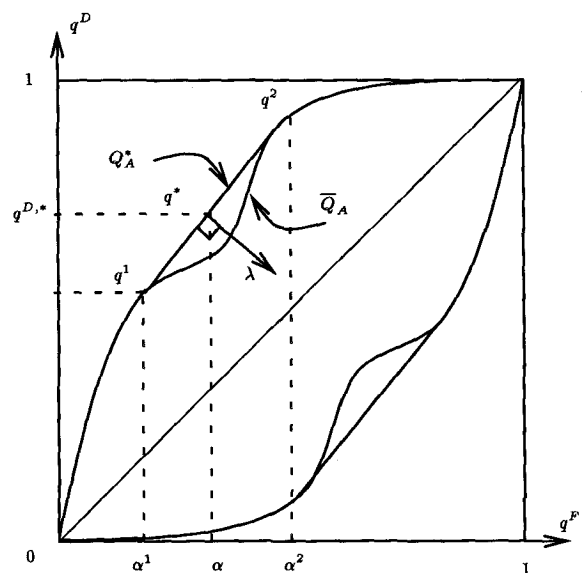


Figure 3. – Construction d'une règle  $\underline{\gamma} \in \Gamma_A^*$  à partir de deux règles de  $\bar{\Gamma}_A$ .

11.  $Fr(A)$  désigne la frontière de  $A$ .

risque Bayésien moyen associé au critère de la probabilité d'erreur de  $DM_0$ .

(ii)  $Q_A^* = \text{co}(\overline{Q}_A)$  et  $q^* \in \text{Fr}(Q_A^*)$ . Par le théorème de Carathéodory, [Ber90], dans  $\mathbb{R}^2$ ,  $q^*$  est une combinaison convexe de deux points de  $\overline{Q}_A$ , soit  $q^1$  et  $q^2$ , associés à deux règles  $\underline{\gamma}^1$  et  $\underline{\gamma}^2$  de  $\overline{\Gamma}_A$ , soit  $q^* = p.q^1 + (1-p).q^2$ ,  $0 \leq p \leq 1$  et  $\underline{\gamma} \in [\underline{\gamma}^1, \underline{\gamma}^2]$ . Il reste à montrer que  $\underline{\gamma}^1$  et  $\underline{\gamma}^2$  sont sur l'hyperplan d'appui associé à  $q^*$ . Si  $p = 0$  ou  $1$  c'est trivial. Sinon, supposons par l'absurde que l'un des points  $q^1$  et  $q^2$  ne soit pas sur l'hyperplan. Alors, l'hyperplan et le segment  $[q^1, q^2]$  n'ont que  $q^*$  en commun, et  $q^1$  et  $q^2$  sont de part et d'autre de l'hyperplan. C'est une contradiction, puisqu'un des deux points ( $q^1$  ou  $q^2$ ) ne peut plus vérifier l'inégalité  $q' \cdot \lambda \geq q^* \cdot \lambda$ . Ainsi  $\underline{\gamma}^1, \underline{\gamma}^2, \underline{\gamma}$  sont sur le même hyperplan d'appui, ce qui signifie, revenant au (i), qu'elles sont solutions du même problème Bayésien.

(iii) Ce point est clair, car on a vu que  $\overline{Q}_A = Q_A$  si le modèle est à r.d.v. continu.  $\square$

Pour conclure ce paragraphe, insistons sur la caractéristique particulière des règles de  $\overline{\Gamma}_A$  et  $\Gamma_A^*$ . Pour cela, il est important de comprendre la mise en œuvre pratique de ces deux types de règles et ce qui les différencie. Si  $\underline{\gamma} \in \overline{\Gamma}_A$ , nous avons montré l'optimalité des règles de décision locales à S.N.G. monotones. Pour de telles règles, chaque décideur dispose d'une famille (unique) de seuils à l'aide de laquelle il quantifie le r.d.v. de son observation totale. Eventuellement, on peut avoir à mettre en œuvre un mécanisme de randomisation locale (i.e. indépendante) si les lois des r.d.v. chargent des points. Si  $\underline{\gamma} \in \Gamma_A^*$ , on vient de voir la nécessité d'une randomisation globale. Celle-ci consiste, pour chaque décideur, à utiliser deux familles de seuils, chaque famille étant choisie aléatoirement selon une loi de Bernoulli globale connue et utilisée par tous les décideurs. On peut voir de telles règles comme étant à S.N.G., mais pour des seuils aléatoires. Il y a dans ce cas nécessité de synchroniser les décideurs, en distribuant à chacun non seulement le paramètre  $p$  de la loi de Bernoulli, mais aussi la « graine » (conditions initiales) du programme de simulation de la loi. Ce mécanisme peut s'interpréter comme une « coordination silencieuse » des divers décideurs. Le problème de l'identification effective des règles  $\underline{\gamma} \in \Gamma_A^*$  optimales se fera sans difficulté du moment qu'on saura identifier  $\overline{Q}_A$ . Alors, les zones où  $\overline{Q}_A$  perd sa convexité sont identifiables et  $Q_A^*$  est l'enveloppe convexe de  $\overline{Q}_A$  sur ces zones (par exemple sur la zone  $[\alpha_1, \alpha_2]$  de la fig. (3)). Si cette construction géométrique est possible « à la main » (ou plutôt numériquement), on doit reconnaître que la question théorique consistant à démontrer qu'un diagramme  $\overline{Q}_A$  donné perd sa convexité en certains points demeure ouverte. Par conséquent, on ne sait pas répondre actuellement à la question : « Quand a-t-on  $\overline{Q}_A = Q_A^*$  ? » Mais, à défaut de preuves, le tracé de  $\overline{Q}_A$  montre que pour certaines lois (et certaines architectures) la réponse est affirmative (i.e.  $\overline{Q}_A = Q_A^*$ ) tandis qu'elle est négative pour d'autres.

## 7. Conclusion

Comme nous l'avions annoncé dans l'introduction, les problèmes de D.D. arborescente permettent, moyennant l'hypothèse de l'indépendance conditionnelle des observations, de reprendre la démarche classique (centralisée) recherchant la structure des règles de décision optimales. Le résultat central que nous avons ainsi obtenu indique l'optimalité des règles globales qui sont construites à partir de règles locales opérant un seuillage du r.d.v. de l'observation. Cependant, le critère de N.-P. engendre en D.D. de nouveaux mécanismes de randomisation qui intègrent explicitement la possibilité que peuvent avoir les différents décideurs de « coordonner leur randomisation ». Par ailleurs, on a vu que (i) on sait seulement caractériser la structure des règles optimales et (ii) certaines questions théoriques (par exemple la question : « quand a-t-on  $\overline{Q}_A = Q_A^*$  ? ») demeurent actuellement ouvertes. Seul le point (i) a fait l'objet de travaux réellement satisfaisants, qui ont réussi à développer des méthodes numériques pour déterminer les valeurs des seuils employés par les décideurs. Ces méthodes sont performantes et ont une portée à la fois pratique et théorique, [TPK91, Bal94]. En guise de conclusion et pour revenir à un point de vue de traiteur de signal, rappelons que l'hypothèse de l'indépendance conditionnelle des observations est très restrictive, puisqu'elle oblige à se limiter aux situations du type « signal connu dans du bruit indépendant de capteur à capteur ». Une piste de recherche qui devrait sérieusement motiver la communauté de traitement de signal est celle visant, comme le suggère [Tsi93], à développer des algorithmes capables de gérer la complexité algorithmique intrinsèque de ces problèmes en s'inspirant des méthodes approximatives qu'on connaît pour « résoudre » les problèmes NP-complets.

## BIBLIOGRAPHIE

- [Bal94] F. Baldit. *Détection décentralisée. Théorie et pratique des architectures arborescentes*. PhD thesis, Université de Rennes I, July 1994.
- [Ber90] M. Berger. *Géométrie*. Volume II, Nathan, 90.
- [BLC93] F. Baldit and J.-P. Le Cadre. Comparative study of centralized and decentralized localization methods. In *ICASSP93*, 1993.
- [Leh86] E. L. Lehmann. *Testing Statistical Hypotheses*. John Wiley & Sons, second edition, 1986.
- [NP33] J. Neyman and E.S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans. Roy. Soc. Ser. A*, 231 :289-337, 1933.
- [TA85] J. N. Tsitsiklis and M. Athans. On the complexity of decentralized decision making and detection problems. *IEEE Transactions on Automatic Control*, 30(5) :440-446, 1985.
- [TPK91] Z.B. Tang, K.R. Pattipati, and D.L. Kleinman. Optimization of detection networks : part i-tandem networks. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(5) :1044-1059, Sept/Oct 1991.

[TPK93] Z.B. Tang, K.R. Pattipati, and D.L. Kleinman. Optimization of detection networks : part ii-tree structures. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(1) :211–221, jan./feb. 1993.

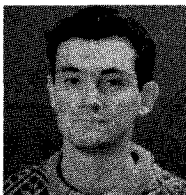
[Tsi93] J. Tsitsiklis. *Advances in Statistical Signal Processing*, chapter Decentralized Detection, pages 297–344. Volume 2 : Signal detection, Greenwich, CT : Jai Press, 1993.

[WW92] P. Willett and D. Warren. The suboptimality of randomized tests in distributed and quantized detection systems. *IEEE Transactions on Information Theory*, Vol.38(No.2) :355–361, March 1992.

*Manuscrit reçu le 20 juin 1994.*

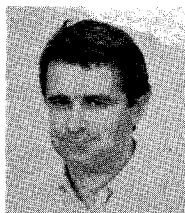
## LES AUTEURS

### F. BALDIT



Frédéric Baldit a obtenu le diplôme d'ingénieur de l'ICPI en 88 (Electronique et Traitement de l'Information). Il a ensuite poursuivi une formation en traitement du signal obtenant le DEA «Signal, Images et Paroles» de l'INPG en 89. Son activité de recherche a été essentiellement consacrée à l'étude théorique des problèmes de détection répartie, faisant l'objet d'une thèse de doctorat de l'université de Rennes I soutenue le 11 juillet 94.

### J.-P. LE CADRE



Jean-Pierre Le Cadre a soutenu une thèse de 3<sup>e</sup> cycle en juin 1982 puis une thèse de doctorat d'Etat en octobre 1987, toutes deux en traitement du Signal et à l'INPG. De 1980 à 1988, ses travaux portent essentiellement sur le traitement d'antenne appliqué à l'acoustique sous-marine et sont effectués au GERDSM (DCAN Toulon). Depuis le 1<sup>er</sup> octobre 1988, J.P. Le Cadre est affecté à l'IRISA en qualité de Chargé de recherche au CNRS. Il est responsable du groupe de travail n<sup>o</sup> 4 (GT4, «antennes») du GdR TdSI.