

**UNIVERSITE DE LIMOGES**  
**ECOLE DOCTORALE Science – Technologie – Santé**

**Faculté des Sciences et Techniques**

**UMOP, FRE CNRS 2701 – Groupe d'Etude des Systèmes de Télécommunications de l'ENSIL**

**Année : 2004**

**Thèse N° 6-2004**

**Thèse**

**pour obtenir le grade de**

**DOCTEUR DE L'UNIVERSITE DE LIMOGES**

**Discipline : Electronique des Hautes Fréquences et Optoélectronique**

**présentée et soutenue par**

**Leila AZIZI**

**le 23 avril 2004 à l'ENSIL**

**Contribution à l'étude :**

- **Des systèmes de transmission multi-antennes (codage spatio-temporel)**
- **Des turbo-codes associés à des modulations à haute efficacité spectrale**

**Thèse dirigée par : Jean-Michel DUMAS, Jean-Pierre CANCES  
et Vahid MEGHDADI**

**JURY :**

<b>Président :</b>	<b>J.F. DIOURIS</b>	<b>Professeur à l'Ecole Polytechnique de Nantes</b>
<b>Rapporteurs:</b>	<b>P. DUHAMEL</b> <b>J.F. HELARD</b>	<b>Directeur de Recherche CNRS à l'Ecole Supérieure d'Electricité</b> <b>Professeur à l'INSA de Rennes</b>
<b>Examineurs :</b>	<b>J.P. CANCES</b> <b>J.M. DUMAS</b> <b>V. MEGHDADI</b> <b>R. MOHAMMAD KHANI</b> <b>A. REZAIE</b>	<b>Maître de Conférences (HDR) à l'ENSIL, Université de Limoges</b> <b>Professeur à l'ENSIL, Université de Limoges</b> <b>Maître de Conférences à l'ENSIL, Université de Limoges</b> <b>Chercheur à l'Ecole Supérieure d'Electricité</b> <b>Professeur à l'Université Polytechnique de Téhéran</b>



# RESUME

Ce document de thèse s'attache d'abord à étudier les performances de schémas de codage de type turbo associés à des modulations à haute efficacité spectrale (MDP8, MAQ16). Nous distinguons principalement les cas d'un entrelacement au niveau de bits et d'un entrelacement au niveau de symboles. Nous détaillons le principe de fonctionnement de chacun des schémas et comparons leurs performances pour des transmissions sur des canaux de Gauss et des canaux de Rayleigh non sélectifs en fréquence.

Puis, nous abordons le contexte des transmissions multi-antennes. Un premier chapitre est consacré à l'étude de la capacité des systèmes MIMO et met bien en évidence l'accroissement potentiel très important apporté par ce type d'architecture. Cependant pour bien profiter de la diversité spatiale offerte par les systèmes MIMO, il faut pouvoir en réception éliminer les signaux interférents. Nous proposons et étudions plusieurs techniques depuis la plus simple architecture de type VBLAST jusqu'aux codes spatio-temporels de type STBC ou STTC. En particulier, dans le cadre des codes STBC, nous avons développé un calcul original de borne supérieure très précise, et dans le cadre de l'étude des codes STTC, nous avons proposé un nouveau type de code qui combine l'avantage en diversité des codes STBC avec le gain de codage STTC.



**INTRODUCTION ET CONTEXTE**

**GENERAL**



# INTRODUCTION ET CONTEXTE GENERAL

Outre le chapitre 1 qui décrit des architectures d'émetteurs-récepteurs turbo pour les modulations à haute efficacité spectrale, ce document de thèse constitue principalement une contribution à l'étude des systèmes MIMO.

Dans le premier chapitre, nous étudions l'association du codage turbo avec des modulations à haute efficacité spectrale de type MDP-8 au moins, voire MAQ-16, 32... Les modulations codées en treillis, inventées par Ungerboeck en 1982, sont très efficaces pour les communications sur des canaux à bande limitée. Elles sont, entre autres, très utilisées pour les radiocommunications téléphoniques terrestres et satellitaires et permettent d'obtenir des gains de codage de l'ordre de 3-6 dB sans perte de bande passante ou de taux de transmission.

D'autre part, les turbo-codes binaires permettent d'obtenir des taux d'erreur très faibles même à faibles  $E_b/N_0$ . Cependant, leur utilisation n'est pas envisageable pour des canaux à bande passante limitée. Il faut pouvoir améliorer leur efficacité spectrale. L'idée, pour combiner les performances des turbo-codes et l'efficacité spectrale des modulations codées en treillis, est de proposer des schémas de codage combinant les deux approches. On obtient alors des turbo-modulations codées en treillis (Turbo Trellis Coded Modulation schemes). Nous examinons différents schémas de codage/décodage qui combinent le codage turbo et les modulations à haute efficacité spectrale et nous montrons les performances obtenues sur canal de Gauss et canal de Rayleigh non-sélectif en fréquence.

Pour le canal de Gauss, on peut distinguer, pour les turbo-modulations à haute efficacité spectrale, celles qui s'inspirent directement des schémas de turbo-codes en binaire et celles qui utilisent la technique des modulations codées en treillis inspirée des travaux de Ungerboeck. Dans le premier sous-ensemble, cité on distinguera les modulations pragmatiques ("pragmatic" schemes) qui utilisent un seul turbo-code et les modulations multi-niveaux qui utilisent plusieurs codeurs turbo. Dans le cas des schémas qui utilisent les modulations codées en treillis on considérera d'une part les schémas avec poinçonnage alternatif des symboles de parité et d'autre part les schémas avec poinçonnage des bits systématiques. Dans le premier cas l'entrelacement au codeur se fait au niveau des symboles tandis que dans le second cas l'entrelacement a lieu au niveau des bits. Dans le cas des canaux de Rayleigh non-sélectifs en fréquence nous présentons un schéma de modulation turbo appelée "*I-Q* turbo coded modulation" qui s'inspire des modulations pragmatiques mais qui présente une grande souplesse d'utilisation et un gain en performance de l'ordre de 0.5-1 dB par rapport à ces dernières.

A partir du chapitre 2, nous abordons le contexte des systèmes MIMO. Ce travail de thèse s'inscrit dans le cadre de la mise en place de nouveaux services de radiocommunications mobiles à hauts débits. Le principal obstacle à la mise en œuvre de tels systèmes est constitué

par la nature même du canal radiomobile. En effet, le canal radiomobile est un canal sélectif en fréquence et variable dans le temps. Plus on augmente les débits et plus on renforce ce caractère sélectif en fréquence qui se traduit par une forte IES en réception. L'introduction des systèmes multiporteuses de type OFDM pur ou OFDM combiné avec l'étalement de spectre (systèmes MC-CDMA par exemple), permet de réduire efficacement ce problème puisque l'OFDM de taille  $N$  transforme le canal large bande sélectif en fréquence en  $N$  sous-canaux parallèles non-sélectifs en fréquence. Ceci permet, entre autres, d'utiliser un égaliseur beaucoup plus simple en réception que dans le cas des systèmes TDMA. Cependant, la nature variable du canal avec la présence de "fadings" profonds, reste un obstacle majeur à l'augmentation des débits transmis.

Une solution fondamentale au problème de la capacité du canal radiomobile  $H$  a été apportée par les résultats de Télietar et de Foschini. En effet, ces deux chercheurs ont démontré que l'utilisation d'antennes multiples en émission et en réception permettait d'obtenir des accroissements de capacité substantielle à condition de considérer les canaux de transmission entre antennes indépendants et de pouvoir obtenir une estimation précise des coefficients de la réponse impulsionnelle du canal de propagation. Ces résultats ont conduit la communauté scientifique du domaine à s'intéresser aux systèmes MIMO (Multiple Input Multiple Output). La quantification de la capacité de tels canaux MIMO passe par le calcul de la capacité "outage". Dans les canaux à "fadings" la capacité est essentiellement limitée, comme on l'a déjà mentionné, par la présence de phénomènes d'évanouissements profonds sur la liaison. Cependant, la plupart du temps, une capacité bien supérieure est disponible. Ceci est particulièrement important lorsque l'on a affaire à des transmissions par paquets sans codage. Si le système est conçu pour transmettre à un certain débit, un certain pourcentage des paquets sera démodulé avec des erreurs. Ce pourcentage est appelé "outage" et on suppose alors qu'il existe un protocole de reprise des erreurs de transmission (protocole ARQ) qui permet de réordonner la transmission des paquets erronés. Ainsi, dans le cas des canaux à fading, on considère plutôt la capacité "outage" que la capacité de Shannon classique. On définit alors la capacité de transmission "outage"  $C_x$  avec le pourcentage  $x\%$  qui correspond au débit maximum que l'on peut transmettre sans erreur pendant  $(100 - x\%)$  du temps. Ceci revient à dire que l'on vise à obtenir une capacité "outage" de  $x\%$ . En supposant le canal constant sur la durée d'un paquet, la capacité "outage" peut être évaluée par simulations de type Monte Carlo en calculant la capacité pour un grand nombre de réalisations de  $H$ . Le résultat fondamental est que la pente asymptotique de la courbe de la capacité en fonction de  $E_s/N_0$  en dB est déterminée par le nombre minimum  $\min\{n_T, n_R\}$  d'antennes d'émission-réception. Ainsi, un gain de capacité très important, en termes de capacité "outage" peut être réalisé lorsqu'on augmente à la fois le nombre d'antennes d'émission et de réception. En fait, on constate un accroissement quasi-linéaire de la capacité lorsque  $n_T = n_R$ . Par contre, si le nombre d'antennes est augmenté uniquement en émission ou en réception, on n'obtient qu'une translation de la courbe de capacité (par rapport à un système simple antenne) et la pente asymptotique reste constante car la quantité  $\min\{n_T, n_R\}$  reste inchangée.

Ces résultats sont démontrés dans le chapitre 2 du document de thèse présenté. Cependant, s'il est montré que la capacité augmente considérablement avec le nombre d'antennes d'émission et de réception il faut encore pouvoir disposer d'un récepteur performant qui va pouvoir prendre en compte toute cette richesse de diversité créée par l'emploi d'antennes multiples. Le choix de récepteurs à maximum de vraisemblance conduit alors à choisir le signal dont la métrique est la plus faible parmi  $M^{n_T}$  signaux (lorsqu'on a affaire à une modulation à  $M$  états), ce qui conduit à des complexités rapidement prohibitives. C'est ce qui explique pourquoi des systèmes sous-optimaux ont d'abord été introduits comme les systèmes BLAST (Bell-Labs-Layered-Space-Time-Architecture) avec comme principales alternatives les systèmes V-BLAST (Vertical BLAST) et D-BLAST (Diagonal BLAST). Dans ce type de récepteur, on détecte à chaque étape un signal issu d'une antenne alors que les signaux issus des autres antennes sont supprimés à l'aide d'un dispositif ZF (Zero-Forcing) ou par une technique MMSE (Minimum Mean Square Error) et par une suppression d'interférences en utilisant les résultats des étages précédents. Cependant, ce type de dispositif est très sensible aux erreurs de propagation et nécessite un nombre élevé d'antennes d'émission et de réception ce qui peut s'avérer délicat sur un terminal portable de faibles dimensions. Le chapitre 3 est consacré à l'étude du système V-BLAST et nous montrons que les performances obtenues dépendent également beaucoup de l'écart lorsque  $n_R - n_T$  entre le nombre d'antennes de réception et le nombre d'antennes d'émission.

Pour pallier à la nécessité d'utiliser un nombre élevé d'antennes des schémas de transmission à base de codes spatio-temporels ont été mis au point. Dans ce domaine ce sont incontestablement les travaux de Tarokh et d'Alamouti qui ont servi de base aux travaux de recherche actuels. Alamouti a d'abord mis au point un schéma de diversité spatiale simple pour deux antennes d'émission qui s'est imposé comme un standard dans la norme UMTS pour les transmissions hauts-débits (l'utilisation d'antennes multiples n'est qu'optionnelle dans la norme UMTS actuelle). Puis, Tarokh a généralisé le concept d'Alamouti avec le codage spatio-temporel en bloc pour un nombre quelconque d'antennes d'émission et de réception. Il faut bien comprendre que, contrairement à ce que pourrait laisser croire leur nom, ces schémas de transmission n'induisent aucun gain de codage, ils sont simplement des schémas de diversité optimaux. Leur principal intérêt est que le traitement en réception est très simple puisqu'une simple combinaison linéaire (MRC: Maximum Ratio Combiner) des sorties des filtres adaptés permet de récupérer les données transmises avec un maximum de diversité. Le chapitre 4 étudie les schémas d'émission-réception à base de codes STBC, nous introduisons entre-autres les règles de codage-décodage à l'aide de formes matricielles très simples même pour des nombres élevés d'antennes. De plus, nous donnons un calcul de borne supérieure en termes de TEB qui s'avère très précis (plus précis que la borne de Chernoff) sur les canaux de Rayleigh non-sélectifs en fréquence.

Pour améliorer les codes STBC, des schémas de transmission avec diversité utilisant le principe des codes convolutifs ont été mis au point, initialement par Tarokh. Il s'agit des codes STTC (Space Time Trellis Codes). Ces codes introduisent, comme les codes

convolutifs, une corrélation entre les symboles transmis sur les antennes d'émission. Leur décodage fait appel à l'algorithme de Viterbi en supposant connus les paramètres de la réponse impulsionnelle du canal au récepteur. Tarokh a proposé des structures de codes spatio-temporels en treillis pour des modulations de type MDP-4, MDP-8 ou MAQ-16 qui permettent de travailler à 2-3 dB de la capacité de coupure ("Outage Capacity") du canal. Nous avons réalisé et simulé la plupart des codes proposés initialement par Tarokh. De plus, nous proposons dans le chapitre 5, des codes originaux appelés SOSTTC qui combinent le gain de codage des codes STTC avec le gain en diversité optimal des codes STBC. La construction de ces codes est basée sur une règle de partitionnement de la constellation initiale en sous-constellations comme pour la construction des modulations codées en treillis de Ungerboeck. Le critère de partitionnement utilisé est basé sur le calcul du déterminant de la matrice des différences de mots de code. On affecte alors aux différentes sous-constellations ainsi définies des codes STBC différents qui vont permettre de définir les nouveaux états du treillis. La construction de ces codes est assez semblable à celle des M-TCM (Multidimensional-TCM). Les performances obtenues montrent une amélioration sensible par rapport aux codes STTC.

### **Revendications:**

Les contributions originales de cette thèse concernent principalement deux points: l'obtention de bornes supérieures très précises sur le TEB des codes STBC dans le chapitre 4 et l'étude des codes SOSTTC dans le chapitre 5, codes qui combinent le gain de codage des codes STTC avec le gain en diversité optimal des codes STBC.

# **CHAPITRE 1**

**ASSOCIATION TURBO-CODES ET  
MODULATION A HAUTE EFFICACITE  
SPECTRALE**

**(TURBO TRELLIS CODED MODULATION  
SCHEMES)**



# **1- Association turbo-codes et modulation à haute efficacité spectrale (Turbo Trellis Coded Modulation Schemes)**

Dans ce chapitre nous étudions l'association du codage turbo avec des modulations à haute efficacité spectrale de type MDP-8 au moins, voire MAQ-16, 32... Les modulations codées en treillis, inventées par Ungerboeck en 1982 [1], sont très efficaces pour les communications sur des canaux à bande limitée. Elles sont, entre autres, très utilisées pour les radiocommunications téléphoniques terrestres et satellitaires et permettent d'obtenir des gains de codage de l'ordre de 3-6 dB sans perte de bande passante ou de taux de transmission.

D'autre part, les turbo-codes binaires permettent d'obtenir des taux d'erreur très faibles même à faibles  $E_b/N_0$ . Cependant, leur utilisation n'est pas envisageable pour des canaux à bande passante limitée. Il faut pouvoir améliorer leur efficacité spectrale.

L'idée, pour combiner les performances des turbo-codes et l'efficacité spectrale des modulations codées en treillis, est de proposer des schémas de codage combinant les deux approches. On obtient alors des turbo-modulations codées en treillis (Turbo Trellis Coded Modulation schemes). Nous examinons différents schémas de codage/décodage qui combinent le codage turbo et les modulations à haute efficacité spectrale et nous montrons les performances obtenues sur canal de Gauss et canal de Rayleigh non-sélectif en fréquence.

Pour le canal de Gauss, on peut distinguer, pour les turbo-modulations à haute efficacité spectrale, celles qui s'inspirent directement des schémas de turbo-codes en binaire et celles qui utilisent la technique des modulations codées en treillis inspirée des travaux de Ungerboeck. Dans le premier sous-ensemble cité on distinguera les modulations pragmatiques ("pragmatic" schemes) qui utilisent un seul turbo-code et les modulations multi-niveaux qui utilisent plusieurs codeurs turbo. Dans le cas des schémas qui utilisent les modulations codées en treillis on considérera d'une part les schémas avec poinçonnage alternatif des symboles de parité et d'autre part les schémas avec poinçonnage des bits systématiques. Dans le premier cas l'entrelacement au codeur se fait au niveau des symboles tandis que dans le second cas l'entrelacement a lieu au niveau des bits. Dans le cas des canaux de Rayleigh non-sélectifs en fréquence nous présentons un schéma de modulation turbo appelée "*I-Q* turbo coded modulation" qui s'inspire des modulations pragmatiques mais qui présente une grande souplesse d'utilisation et un gain en performance de l'ordre de 0.5-1 dB par rapport à ces dernières.

Le plan de ce chapitre est alors le suivant : nous présentons d'abord les approches s'inspirant directement des schémas de turbo-codes en binaire en examinant les modulations pragmatiques puis les schémas multi-niveaux. Nous abordons ensuite le cas des turbo-modulations codées en treillis en distinguant, d'une part, les schémas avec poinçonnage alternatif des symboles de parité et, d'autre part, les schémas avec poinçonnage des bits systématiques. Enfin, nous terminons par l'examen des schémas "*I-Q* turbo coded modulation" pour le cas du canal de Rayleigh non-sélectif en fréquence.

## 1-1 Les schémas de turbo-modulation binaires

### 1-1-1 Les turbo-modulations pragmatiques

Dans le cas des turbo-modulations pragmatiques [4] on utilise un seul codeur turbo binaire de rendement  $1/n$ . Les sorties du codeur sont multiplexées et convenablement poinçonnées pour obtenir  $\tilde{m}$  symboles de parité et  $m - \tilde{m}$  symboles d'information., comme le montre la figure 1. Ces symboles codés vont ensuite constituer à l'aide d'une correspondance ("mapping") prédéfinie (on utilise classiquement le mapping de Gray) des symboles MDP- $N$  ou MAQ- $N$ . Si on utilise une modulation MDP- $N$ , on a  $m$  symboles codés qui sélectionne un signal MDP- $N$  d'après la correspondance de Gray. Si on utilise une modulation MAQ- $N$ , on a  $m/2$  symboles codés qui sont utilisés pour synthétiser le signal MAQ- $I$  (en phase) et  $m/2$  symboles codés qui sont utilisés pour synthétiser le signal MAQ- $Q$  (en quadrature). Une correspondance de Gray est appliquée sur chacune des voies  $I$  et  $Q$ . L'efficacité spectrale de ce schéma est égal à  $m - \tilde{m}$  bits/s/Hz.

Bien que généralement un entrelaceur soit inséré entre le multiplexeur et le module de poinçonnage dans le but de décorrélérer les signaux sur des canaux non stationnaires à évanouissements, sa présence n'est pas indispensable dans le cas d'un canal de Gauss.

Le récepteur, présenté sur la figure 2, calcule les rapports de vraisemblance logarithmiques au niveau des bits à partir du signal reçu bruité et des sorties pondérées du démodulateur (sorties des filtres adaptés). Le flux des rapports de vraisemblance est désentrelacé puis démultiplexé avant d'entrer dans le décodeur turbo binaire, ce décodeur pouvant être du type SOVA ou MAP.

L'approche pragmatique est simple puisqu'elle ne fait appel qu'à un seul turbo codeur et qu'elle réutilise les principes des turbo-codes en binaire. En modifiant la fonction de poinçonnage et la constellation du signal modulé on peut alors obtenir une large famille de turbo-modulations à haute efficacité spectrale.

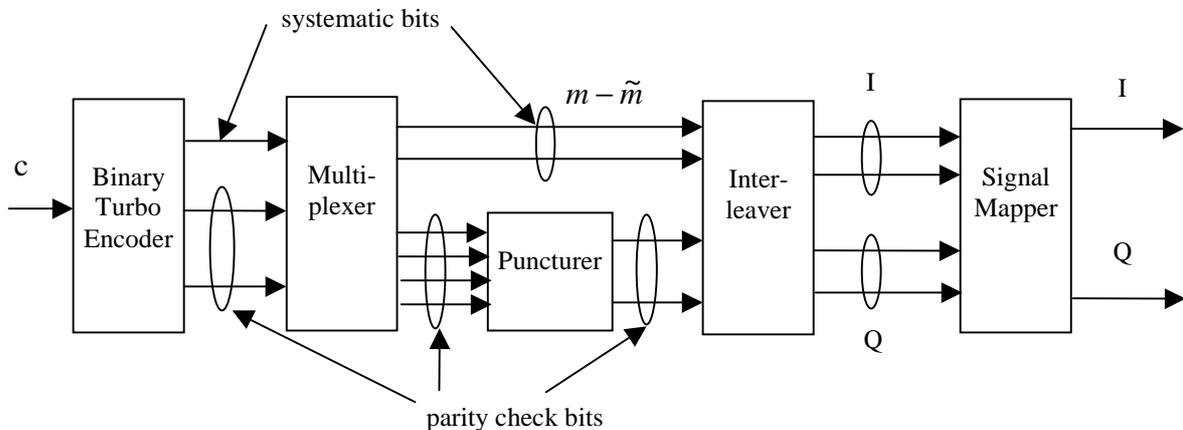
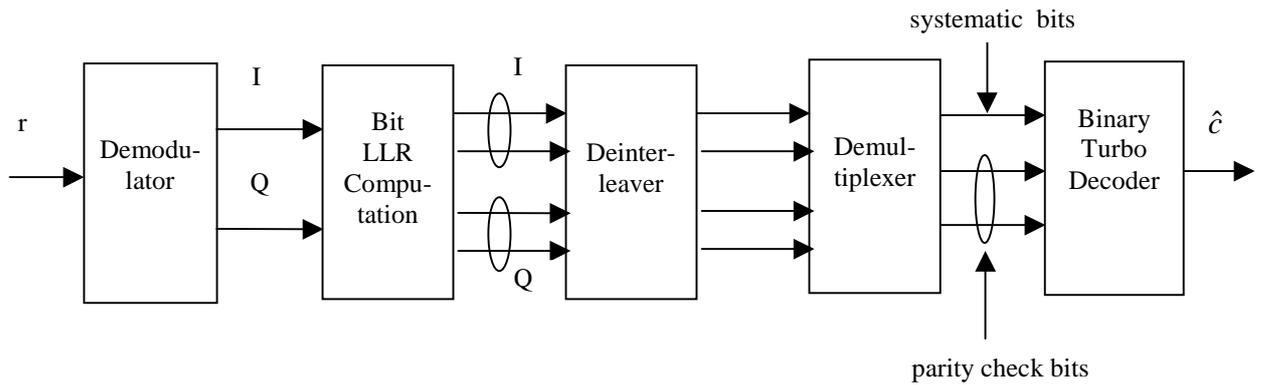


Fig.1 : codeur turbo-modulation pragmatique



**Fig. 2 : décodeur turbo-modulation pragmatique**

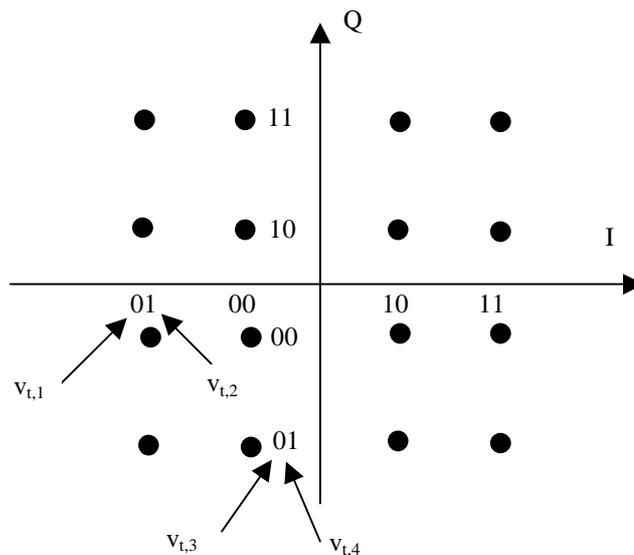
Pour illustrer un exemple concret de turbo-modulation pragmatique nous considérons l'exemple d'une modulation MAQ-16 avec un rendement  $1/2$  avec  $n = 1, m = 3$  et  $\tilde{m} = 4$  (cf figure 1). Un turbo-codeur de rendement  $1/3$  est utilisé. Les bits codés sont multiplexés et poinçonnés d'après la matrice suivante :

$$P = \begin{bmatrix} 11 \\ 10 \\ 01 \end{bmatrix} \quad (1)$$

Ceci conduit à un turbo-code de rendement  $1/2$ . Si on note les quatre bits codés à l'instant  $t$  par  $v_{t,i}, i = 1, 2, 3$  et  $4$ , ils viennent sélectionner un signal MAQ-16  $x_t$ , donné par :

$$x_t = x_{t,I} + j \cdot x_{t,Q} \quad (2)$$

où  $x_{t,I}$  est le signal en-phase ( $I$ ) et  $x_{t,Q}$  est le signal en quadrature ( $Q$ ) du signal transmis  $x_t$ . Une constellation typique MAQ-16 avec mapping de Gray est présentée sur la figure 3.



**Fig. 3 : Modulation MAQ-16 avec mapping de Gray**

En supposant une détection cohérente, le signal reçu à l'instant  $t$  appelé  $r_t$ , peut se mettre sous la forme :

$$r_t = r_{t,I} + j.r_{t,Q} \quad (3)$$

avec les composantes  $I$  et  $Q$  données par

$$\begin{aligned} r_{t,I} &= x_{t,I} + n_{t,I} \\ r_{t,Q} &= x_{t,Q} + n_{t,Q} \end{aligned} \quad (4)$$

$n_{t,I}$  et  $n_{t,Q}$  sont deux échantillons de bruit Gaussiens de moyenne nulle et de variance  $\sigma^2$ . Le rapport de vraisemblance logarithmique (log-likelihood ratio (LLR)) de chaque bit codé  $v_{t,i}$ , noté  $\Lambda(v_{t,i})$ , se calcule par la formule [4], [12].

$$\Lambda(v_{t,i}) = K_c \cdot \log \frac{\Pr\{v_{t,i} = 1 | r_t\}}{\Pr\{v_{t,i} = 0 | r_t\}} \quad (5)$$

où  $K_c$  est une constante. Comme les composantes en phase et en quadrature du signal reçu sont affectées par des composantes de bruit indépendantes, l'application de la règle de Baye montre que les LLR's de  $v_{t,1}$  and  $v_{t,2}$  dépendent uniquement de  $r_{t,I}$ , tandis que les LLR's de  $v_{t,3}$  and  $v_{t,4}$  dépendent uniquement de  $r_{t,Q}$ . On peut alors écrire les LLR's des bits codés avec la formule ci-dessous.

$$\begin{aligned} \Lambda(v_{t,i}) &= K_c \cdot \log \frac{\sum_{x_I: v_{t,i}=1} \exp\left(\frac{-1}{2\sigma^2} \cdot (r_{t,I} - x_I)^2\right)}{\sum_{x_I: v_{t,i}=0} \exp\left(\frac{-1}{2\sigma^2} \cdot (r_{t,I} - x_I)^2\right)}, \quad i = 1,2 \\ \Lambda(v_{t,i}) &= K_c \cdot \log \frac{\sum_{x_Q: v_{t,i}=1} \exp\left(\frac{-1}{2\sigma^2} \cdot (r_{t,Q} - x_Q)^2\right)}{\sum_{x_Q: v_{t,i}=0} \exp\left(\frac{-1}{2\sigma^2} \cdot (r_{t,Q} - x_Q)^2\right)}, \quad i = 3,4 \end{aligned}$$

où  $x_I$  est la composante en-phase et  $x_Q$  la composante en-quadrature du signal modulé  $v_{t,i} = j, j = 0,1$ . Les rapports de vraisemblance sont alors désentrelacés et démultiplexés puis envoyés au turbo-décodeur binaire classique.

Il faut remarquer que les LLR's des bits  $v_{t,1}$  et  $v_{t,2}$  qui sont affectés par le même échantillon de bruit Gaussien  $n_{t,I}$  ne sont pas indépendants. Il en va de même pour  $v_{t,3}$  et  $v_{t,4}$  soumis à l'échantillon de bruit  $n_{t,Q}$ . De plus, pour un signal reçu donné  $r_t$ , les LLR's  $\Lambda(v_{t,i})$  ne sont pas strictement Gaussiens pour tous les bits codés  $i = 1,2,3$  et 4. Ainsi, le turbo-décodeur binaire n'est plus optimal pour le schéma de modulation considéré. Cependant, les résultats de simulation ont montré que, pour des valeurs de SNR's élevés, les LLR's  $\Lambda(v_{t,i})$  sont proches

de variables Gaussiennes de variance  $\sigma^2$  si la constante  $K_c$  est prise égale à  $\sigma^2/2$ . Dans ce cas le turbo-décodeur binaire permet d'obtenir de bonnes performances dans le cadre de l'approche pragmatique [4], [12].

### 1-1-2 Les turbo-modulations multi-niveaux

Le schéma de modulation multi-niveaux, proposé par Imai et Hirakawa [5], est une approche combinant codage et modulation qui utilisent plusieurs turbo-codeurs. L'émetteur, pour un signal dans un alphabet  $M$ -aire, consiste en la juxtaposition de  $l = \log_2 M$  codeurs binaires mis en parallèle comme le montre la figure 4. Pour expliquer le principe de fonctionnement, on note une séquence de messages de longueur  $k$  par

$$c = (c_1, c_2, \dots, c_k) \quad (6)$$

La séquence est partagée en  $l$  blocs. Le bloc  $i$  se met sous la forme

$$c^i = (c_1^i, c_2^i, \dots, c_{k_i}^i) \quad (7)$$

où  $k_i$  est la longueur du bloc numéro  $i$  et

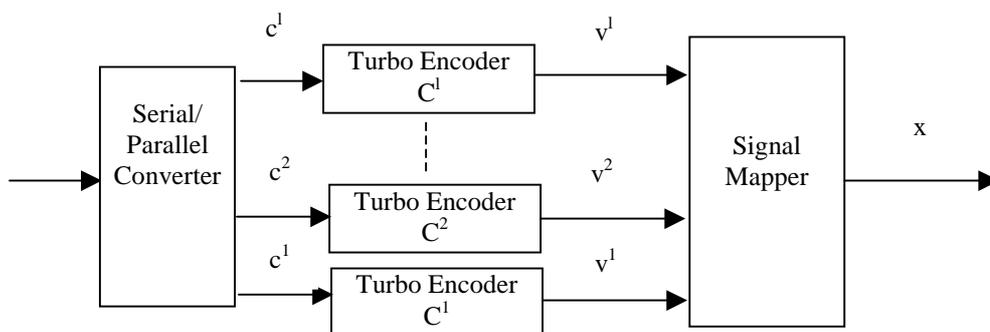
$$k = \sum_{i=1}^l k_i \quad (8)$$

Chaque bloc de message  $c^i$  est codé par un turbo-codeur  $C^i$  qui génère un mot de code

$$v^i = (v_1^i, v_2^i, \dots, v_{n_i}^i) \quad (9)$$

où  $n_i$  est la longueur du code pour le turbo-codeur  $C^i$ . Le rendement du code  $C^i$  est égal à  $R_i = k_i/n_i$ . Dans la suite nous supposons, pour des raisons de simplicité évidentes, que chaque turbo-code a la même longueur  $n$ , ainsi  $n_1 = n_2 = \dots = n_l = n$ . Les sorties des codeurs à l'instant  $t$ , notées  $v_t^i$ ,  $i = 1, 2, \dots, l$ ,  $t = 1, 2, \dots, n$ , forment un mot binaire  $(v_t^1, v_t^2, \dots, v_t^l)$  qui est transformé par un modulateur  $M$ -aire en un signal modulé  $x_t$ . La séquence de signaux transmis s'écrit alors

$$x = (x_1, x_2, \dots, x_t, \dots, x_n) \quad (10)$$

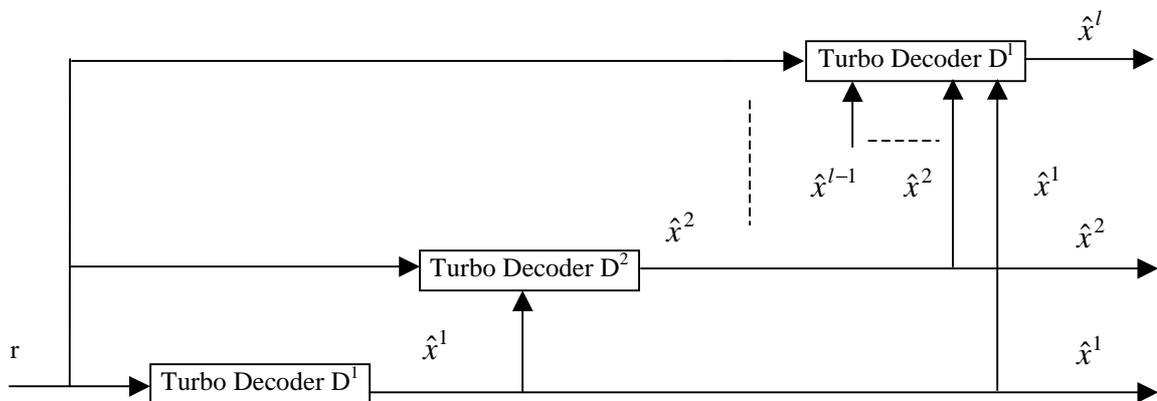


**Fig. 4 : Turbo-codeur multi-niveaux**

Le rendement global du code est  $R = k/nl$ . L'efficacité spectrale du schéma de codage est égale à

$$\eta = R \cdot \log_2 M = \frac{k}{n} \text{ bits/s/Hz} \quad (11)$$

Le décodeur optimal à maximum de vraisemblance opère sur le treillis du code complet égal à la juxtaposition des treillis de chaque codeur élémentaire  $C^i$ . En général la complexité de ce décodeur est telle que sa réalisation pratique est impossible. Heureusement il existe une technique sous-optimale, appelée décodage multi-étages (multistage decoding) dont la performance asymptotique est très proche de celle du décodeur optimal. La structure du décodeur multi-étages est illustrée sur la figure 5.



**Figure 5 : Turbo-décodeur multi-niveaux**

Un problème important dans la conception du code réside dans le choix des turbo-codes élémentaires et de leur rendement. Waschmann et Huber [8] ont proposé une technique particulière pour le choix des turbo-codes élémentaires. Le taux de transmission du code de niveau de modulation  $i$  est choisi égal à la capacité du canal équivalent affecté à ce niveau. Comme la capacité totale du canal est égal à la somme des capacités obtenues pour chaque niveau un tel choix permet, en théorie, dans le cas des longueurs de code infinies, d'éliminer toutes les erreurs sur la transmission. Les turbo-codes opérant près de la capacité limite définie par Shannon sont d'excellents codeurs potentiels pour réaliser le codeur multi-niveaux. De plus, de par l'excellent pouvoir correcteur des turbo-codes, on peut considérer que la propagation d'erreur est négligeable d'un niveau de modulation à l'autre. Cependant, dans le cas de messages de faible longueur, l'utilisation du décodeur multi-niveaux peut conduire à une perte non-négligeable en termes de  $E_b/N_0$  pour obtenir un taux d'erreur donné [8].

## 1-2 Les schémas de turbo-modulation codées en treillis

### 1-2-1 Les schémas avec poinçonnage alternatif des symboles de parité

Une première méthode pour combiner le codage turbo et les modulations codées en treillis consiste à utiliser deux modulations codées en treillis type Ungerboeck réalisées à l'aide de deux codeurs convolutifs récursifs de rendement  $k/(k+1)$  [6-7]. Les principes généraux des modulations codées en treillis sont rappelés en annexe et des exemples de réalisation y sont illustrés. La figure 6 illustre un schéma de turbo-modulation codée en treillis utilisant deux codeurs convolutifs, séparés par des entrelaceurs/désentrelaceurs qui opèrent au niveau des symboles.

Le codeur opère avec des blocs de message de  $N$  groupes de  $k$  bits d'information où  $N$  est la taille de l'entrelaceur. La séquence de message transmis, notée  $c$ , est donnée par :

$$c = (c_1, c_2, \dots, c_t, \dots, c_N)$$

où  $c_t$  est un groupe de  $k$  bits d'information au temps  $t$

$$c_t = (c_{t,0}, c_{t,1}, \dots, c_{t,k-1})$$

Le codeur transforme la séquence d'entrée en un bloc de  $N$  symboles binaires codés

$$v = (v_1, v_2, \dots, v_t, \dots, v_N)$$

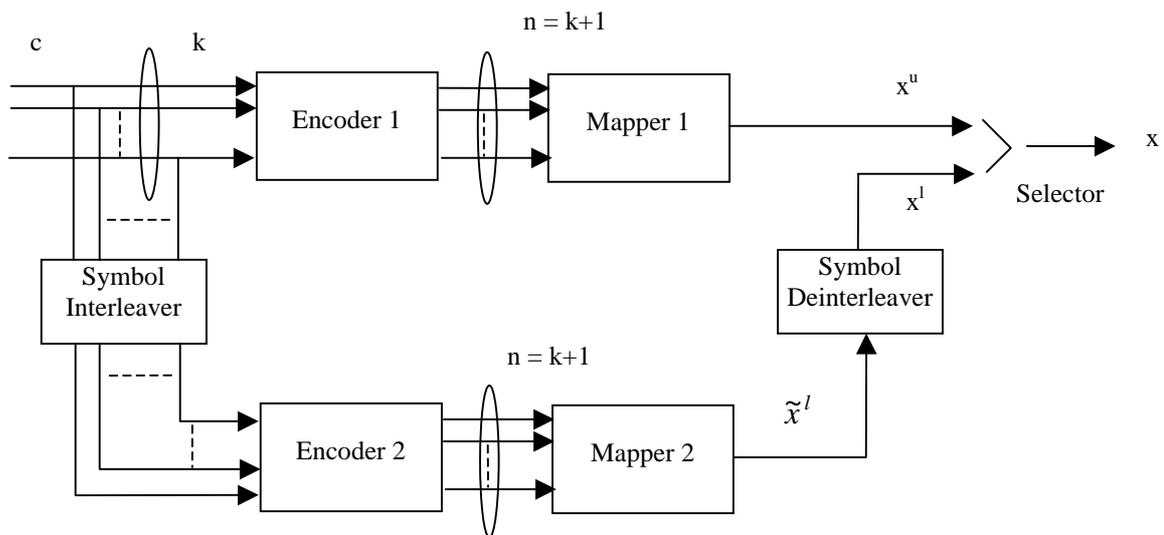


Fig. 6 : Turbo TCM codeur avec poinçonnage des symboles de parité

où  $v_t$  est un symbole codé au temps  $t$ , de longueur  $n = (k+1)$  digits binaires, donné par

$$v_t = (v_{t,0}, v_{t,1}, \dots, v_{t,k})$$

Un entrelaceur permute les bits d'information en travaillant au niveau des symboles. Ainsi, l'ordre d'arrivée de  $k$  bits d'information au niveau de l'entrelaceur reste inchangé.

En ce qui concerne les codes convolutifs utilisés, certains bits d'entrée n'ont pas besoin d'être codés. Dans la pratique ces bits n'ont pas besoin d'être entrelacés mais sont utilisés directement pour sélectionner les signaux dans les sous-constellations. Au récepteur ces bits sont estimés par décodage par sous-constellation [1].

La séquence de symboles codés  $v$  donne naissance à la séquence de signaux modulés  $x''$ , donnée par

$$x'' = (x_1'', x_2'', \dots, x_t'', \dots, x_N'')$$

où  $x_t''$  est un symbole de la constellation de la modulation utilisée, contenant  $2^{k+1}$  points,  $x_t''$  peut s'écrire

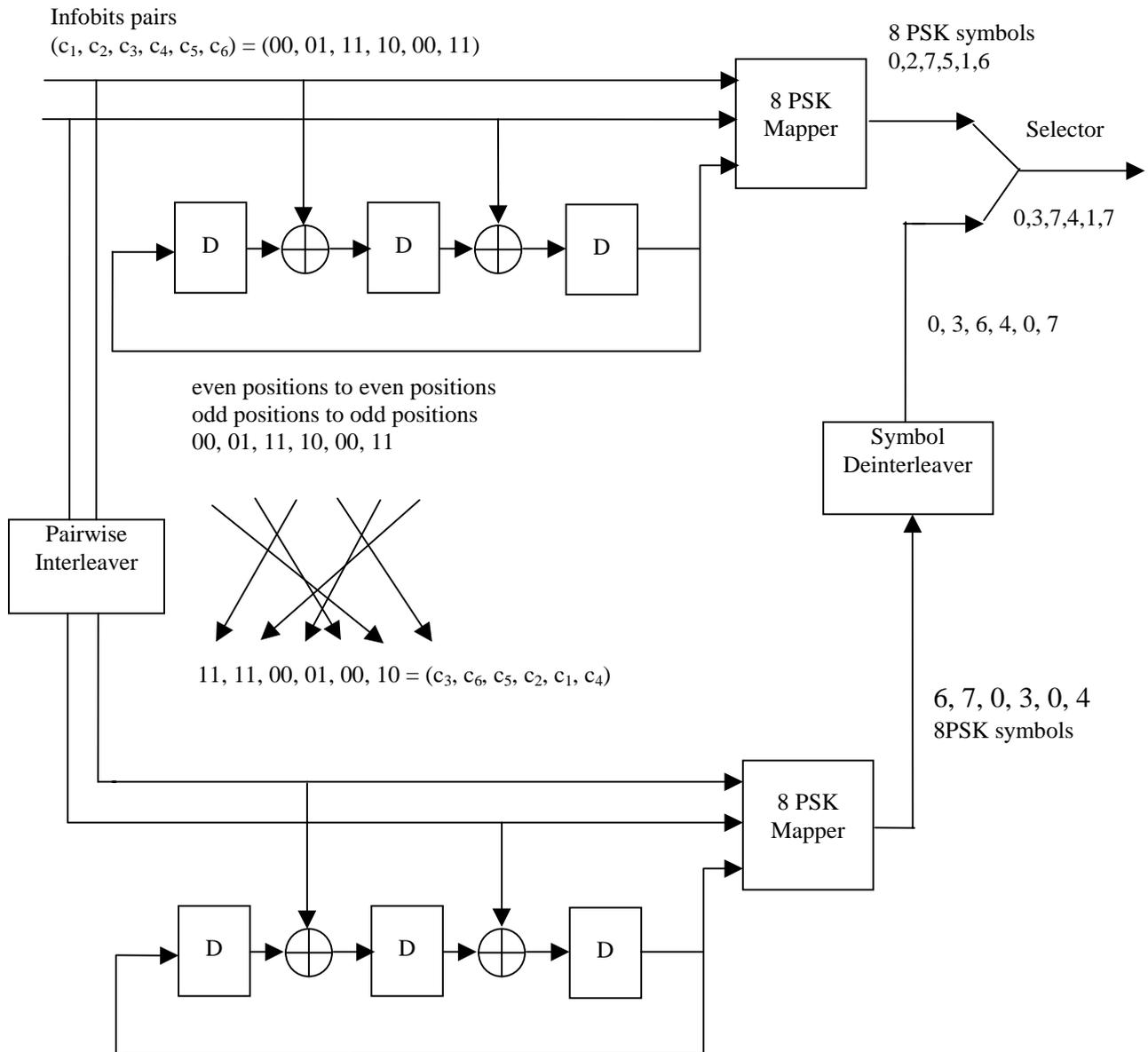
$$x_t'' = x_{t,I}'' + j \cdot x_{t,Q}''$$

$x_{t,I}''$  est le signal en phase et  $x_{t,Q}''$  le signal en quadrature du signal modulé  $x_t''$ . La correspondance ("mapping") entre les bits codés et les symboles modulés est basée sur le partitionnement en sous-constellations proposé par Ungerboeck dans [1] et rappelé en annexe 1.

La sortie du second codeur est désentrelacée. Ceci assure que les  $k$  bits d'information qui déterminent les  $(k+1)$  bits codés de chacun des codeurs supérieur et inférieur sont identiques. Il faut remarquer que les signaux modulés issus de chacun des deux codeurs contiennent chacun  $k$  bits d'information. Ceci constitue une différence importante avec les turbo-codes binaires où

les bits d'information ne sont utilisés qu'une seule fois puisqu'ils peuvent être reconstruits à partir de l'autre séquence reçue en parallèle. Pour empêcher que chaque bit d'information ne soit transmis deux fois, on insère un commutateur (selector, switch) en sortie du schéma de codage qui prend en compte tantôt la sortie de la branche supérieure, tantôt la sortie de la branche inférieure. Dans ce cas, pour une constellation de  $2^{k+1}$  points, on obtient une efficacité spectrale de  $k$  bits/sec/Hz.

Un exemple de codeur 8-états en conjonction avec une modulation MDP-8 est présenté sur la figure 7. Pour le schéma de la figure 7, initialement présenté par Robertson et Woerz, nous avons les paramètres suivants:  $k = 2$  et une taille d'entrelaceur  $N = 6$ .



**Fig. 7 : Exemple de turbo-modulation codée en treillis (poinçonnage des symboles de parité)**

Le processus de codage est alors le suivant. On suppose que nous avons en entrée du codeur une séquence de six groupes de 2 bits ( $k = 2$ ), cette séquence s'écrit:

$$c = (00,01,11,10,00,11)$$

Après codage par le codeur récursif systématique optimisé pour la modulation en treillis correspondante, une séquence MDP-8 est générée par le codeur supérieur, cette séquence est notée  $x^u$

$$x^u = (0,2,7,5,1,6)$$

La séquence d'information est alors entrelacée par un entrelaceur opérant sur les symboles et elle est codée et transformée en séquence modulée par le codeur inférieur, on obtient la séquence  $\tilde{x}^l$

$$\tilde{x}^l = (6,7,0,3,0,4)$$

Après désentrelacement cette séquence devient

$$x^l = (0,3,6,4,0,7)$$

Le commutateur de sortie sélectionne alternativement la sortie supérieure et la sortie inférieure des codeurs. Ainsi, le premier symbole transmis sur le canal provient du codeur supérieur, le second du codeur inférieur, le troisième du codeur supérieur...etc. La séquence transmise pour l'exemple illustrée s'écrit:

$$x = (0,3,7,4,1,7)$$

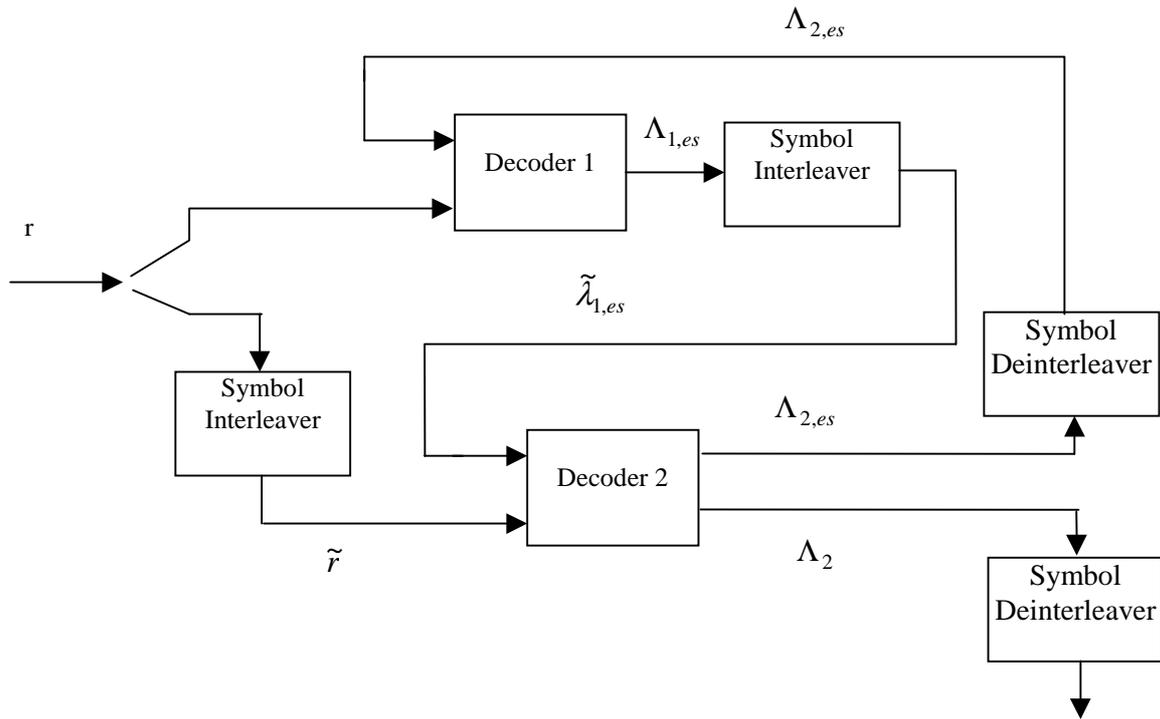
Ainsi, le symbole de parité est alternativement choisi sur la sortie du codeur supérieur et sur la sortie du codeur inférieur. Chaque groupe de bits d'information ( $k = 2$ ) n'apparaît bien qu'une seule fois dans la séquence transmise, on respecte bien la règle de codage des turbo-codes.

Pour ce type de turbo-modulation codée, les codes convolutifs élémentaires doivent être choisis sans branche parallèle pour que chaque bit d'information puisse bénéficier de l'entrelacement. Cependant cette condition peut être relâchée si l'entrelaceur n'a plus à conserver un groupe de  $k$  bits d'information inchangé [15] ou si l'on désire une grande efficacité spectrale [7]. Si l'entrelaceur échange les deux bits d'information dans l'exemple précédent, le code peut alors présenter des branches parallèles. Cette opération assure l'entrelacement des bits d'information à l'intérieur d'un symbole.

Pour trouver de bons codes convolutifs élémentaires on peut utiliser une recherche exhaustive par ordinateur. Le but est alors de maximiser la distance Euclidienne minimale de chaque code tout en sélectionnant de façon aléatoire les symboles de parité (entrelacement) sur le second train codé de symboles [6].

### **1-2-2 Algorithme de décodage Log-MAP pour les Turbo-modulations codées en treillis avec poinçonnage des symboles de parité**

La structure du décodeur est illustrée sur la figure 8. Pour ce type de codage il vaut mieux ne pas utiliser l'algorithme MAP optimal [9] car l'information extrinsèque est souvent très forte ou très faible et peut provoquer des phénomènes de débordement dans les calculs. On lui préfère ici l'algorithme log-MAP beaucoup plus stable numériquement.



**Fig 8 : Décodeur pour Turbo-modulations codées en treillis**

Le procédé de décodage est très semblable à celui des turbo-codes binaires mais on utilise ici des probabilités au niveau des symboles et non plus des bits comme dans le cas binaire. L'algorithme de décodage pour les treillis non-binaires est appelé symbol-by-symbol MAP algorithm.

Le décodeur MAP calcule le LLR de chaque groupe de bits d'information  $c_t = i$ . La sortie pondérée "soft"  $\Lambda(c_t = i)$  est donnée par [2-3]

$$\Lambda(c_t = i) = \log \frac{\Pr\{c_t = i | r_1^\tau\}}{\Pr\{c_t = 0 | r_1^\tau\}} = \log \frac{\sum_{(l', l) \in B_t^i} \alpha_{t-1}(l') \cdot \gamma_t^i(l', l) \cdot \beta_t(l)}{\sum_{(l', l) \in B_t^0} \alpha_{t-1}(l') \cdot \gamma_t^0(l', l) \cdot \beta_t(l)} \quad (12)$$

où  $i$  représente un groupe de bits d'information dans l'ensemble set,  $\{0, 1, 2, \dots, 2^k-1\}$ , et les probabilités  $\alpha_t(l), \beta_t(l)$  et  $\gamma_t(l, l')$  peuvent être calculées récursivement par l'algorithme de Bahl, Cocke, Jelinek and Raviv [9].

Le symbole  $i$  avec le rapport de vraisemblance logarithmique le plus élevé dans l'équation (12),  $i \in \{0, 1, 2, \dots, 2^k-1\}$  est alors choisi comme symbole décidé (hard decision).

Le gros problème de l'algorithme MAP réside dans l'utilisation importante du calcul d'exponentielles et de multiplications pour calculer les LLR's. Cela peut conduire comme on l'a déjà mentionné à des problèmes d'instabilité. Pour éviter ces problèmes on peut travailler au niveau des logarithmes des rapports de vraisemblance. On utilise alors les logarithmes de probabilité pour le calcul de  $\alpha_t(l), \beta_t(l)$  et  $\gamma_t(l', l)$ . Ces logarithmes de probabilité sont alors

notés  $\bar{\alpha}_t(l)$ ,  $\bar{\beta}_t(l)$  et  $\bar{\gamma}_t(l',l)$  respectivement (avec des notations évidentes). L'algorithme de calcul des coefficients  $\bar{\alpha}_t(l)$  procède de la façon suivante (algorithme ascendant)

$$\bar{\alpha}_t(l) = \log \sum_{l'=0}^{M_s-1} e^{\bar{\alpha}_{t-1}(l') + \bar{\gamma}_t(l',l)} \quad (13)$$

avec les conditions initiales

$$\bar{\alpha}_0(0) = 0 \quad \bar{\alpha}_0(l) = -\infty, l \neq 0$$

et les probabilités  $\bar{\beta}_t(l)$  se calculent par l'algorithme descendant

$$\bar{\beta}_t(l) = \log \sum_{l'=0}^{M_s-1} e^{\bar{\beta}_{t+1}(l') + \bar{\gamma}_{t+1}(l,l')} \quad (14)$$

avec les conditions initiales

$$\bar{\beta}_\tau(0) = 0 \quad \bar{\beta}_\tau(l) = -\infty, l \neq 0$$

avec les probabilité de transition sur le canal données par

$$\gamma_t(l',l) = \log \sum_{i=0}^{2^k-1} \gamma_t^i(l',l) \quad (15)$$

et :

$$\gamma_t^i(l',l) = \frac{p_t(i)}{p_t(0)} \cdot \exp\left(-\frac{(r_{t,I} - x_{t,I}^i(l))^2 + (r_{t,Q} - x_{t,Q}^i(l))^2}{2\sigma^2}\right), \text{ for } (l',l) \in B_t^i$$

$$\gamma_t^i(l',l) = 0, \quad \text{otherwise}$$

$B_t^i$  correspond à l'ensemble des transitions  $S_{t-1} = l' \rightarrow S_t = l$  lorsque le symbole d'entrée est  $i$ .  $r_{t,I}$  et  $r_{t,Q}$  sont les composantes en phase et en quadrature du signal reçu.  $p_t(i)$  représente la probabilité à-priori que le symbole  $c_t$  soit égal à  $i$ .

Les coefficients  $\bar{\alpha}_t(l)$  et  $\bar{\beta}_t(l)$  dans les équations (13) et (14) peuvent se calculer en utilisant l'algorithme du Jacobien.

$$\begin{aligned} \log(e^{\delta_1} + e^{\delta_2}) &= \max(\delta_1, \delta_2) + \log(1 + e^{-|\delta_2 - \delta_1|}) \\ &= \max(\delta_1, \delta_2) + f_c(|\delta_2 - \delta_1|) \end{aligned} \quad (16)$$

$f_c(\cdot)$  est une fonction de correction que l'on peut implanter en mémoire ROM (look-up table). L'expression  $\log(e^{\delta_1} + e^{\delta_2} + \dots + e^{\delta_n})$  se calcule de façon récursive en utilisant l'équation (16) par la procédure suivante.

$$\begin{aligned} \log \sum_{i=1}^n e^{\delta_i} &= \log(\Delta + e^{\delta_n}), \Delta = e^{\delta_1} + e^{\delta_2} + \dots + e^{\delta_{n-1}} = e^{\delta} \\ &= \max(\log \Delta, \delta_n) + f_c(|\log \Delta - \delta_n|) \\ &= \max(\delta, \delta_n) + f_c(|\delta - \delta_n|) \end{aligned}$$

Le processus itératif de décodage de l'algorithme symbol-by-symbol MAP est similaire à celui utilisé dans le cadre des turbo-codes binaires mais la nature de l'information échangée entre les deux décodeurs est différente [6]. Dans le cas des turbo-codes binaires on peut décomposer la sortie du décodeur MAP en trois termes: l'information à priori générée par l'autre décodeur, l'information systématique présente à l'entrée du décodeur et l'information extrinsèque générée par le décodeur lors de cette nouvelle itération. L'information extrinsèque dans le cas des turbo-codes binaires est indépendante de l'information à priori et de l'information systématique. Par contre, dans le cas des turbo-modulations codées en treillis, on ne peut plus séparer l'influence de l'information systématique et de l'information extrinsèque, elles sont présentes toutes les deux à l'intérieur d'un symbole codé. Dans ce cas ce ne sera plus l'information extrinsèque seule qui sera échangée entre les deux décodeurs mais l'ensemble de l'information extrinsèque et de l'information systématique. L'ensemble information extrinsèque plus information systématique en sortie du premier décodeur est notée par  $\Lambda_{1,es}(c_t = i)$  et peut s'obtenir par

$$\Lambda_{1,es}(c_t = i) = \Lambda_1(c_t = i) - \log \frac{p_t(i)}{p_t(0)} \quad (17)$$

L'ensemble information extrinsèque plus information systématique  $\Lambda_{1,es}(c_t = i)$  est alors utilisé (après entrelacement) comme probabilité à priori dans le prochain étage de décodage. Après entrelacement cet ensemble est noté  $\tilde{\Lambda}_{1,es}(c_t = i)$ . L'ensemble information extrinsèque plus information systématique du second décodeur est donné par

$$\Lambda_{2,es}(c_t = i) = \Lambda_2(c_t = i) - \tilde{\Lambda}_{1,es}(c_t = i) \quad (18)$$

Lors de la prochaine itération de décodage, le second terme de droite de l'équation (17) est remplacé par l'ensemble information extrinsèque plus information systématique issu du second étage de décodage noté  $\tilde{\Lambda}_{2,es}(c_t = i)$ .

Il faut remarquer que le décodage symbol-by-symbol MAP implique que chaque décodeur ne peut utiliser la même information systématique qu'une seule fois à chaque itération [6]. Pour

les schémas de turbo-modulations codées en treillis chaque décodeur reçoit alternativement la sortie bruitée de son propre codeur ce qui implique que le symbole codé reçu en seconde position appartient à l'autre décodeur et doit être traité comme poinçonné.

Par exemple, si on considère ici le premier décodeur, pour chaque signal reçu impair l'algorithme de décodage procède exactement de la même façon que pour les turbo-codes binaires car le décodeur reçoit bien le symbole en provenance de son propre codeur. La seule différence réside dans la nature de l'information transférée au second décodeur qui comprend l'ensemble information extrinsèque plus information systématique. Cependant dans le cas des signaux reçus pairs, le décodeur reçoit le symbole poinçonné dans lequel le symbole de parité est généré par l'autre codeur. Dans ce cas le décodeur ignore ce symbole en mettant à zéro la métrique de transition correspondante dans le treillis. A ce stage la seule entrée dans le treillis est la composante à priori obtenue à partir de l'autre décodeur. Cette composante contient l'information systématique.

### **1-2-3 Algorithme de décodage SOVA pour les Turbo-modulations codées en treillis avec poinçonnage des symboles de parité**

La structure du décodeur est la même que celle représentée en figure 8. Comme pour le décodage des turbo-codes binaires, on définit la sortie pondérée du décodeur  $\Lambda(c_t = i)$ , où  $i \in \{0, 1, 2, \dots, 2^{k-1}\}$ , comme

$$\Lambda(c_t = i) = \log \frac{\Pr\{c_t = i | r_1^\tau\}}{\Pr\{c_t = 0 | r_1^\tau\}} = \mu_t^0 - \mu_t^i \quad (19)$$

$\mu_t^0$  est la métrique cumulée minimale correspondant à  $c_t = 0$ , et  $\mu_t^i$  est la métrique cumulée minimale correspondant à  $c_t = i$ . On définit la métrique de branche affectée au treillis à l'instant  $t$  par

$$v_t = (r_{t,I} - x_{t,I})^2 + (r_{t,Q} - x_{t,Q})^2 - \log p_t(i) \quad (20)$$

où  $r_{t,I}$  et  $r_{t,Q}$  sont les signaux reçus respectivement sur les voies en phase et en quadrature à l'instant  $t$ .  $x_{t,I}$  et  $x_{t,Q}$  sont les signaux de sortie modulés, en phase et en quadrature, correspondant à  $c_t = i$  et  $p_t(i)$  représente la probabilité à-priori que le symbole  $c_t$  soit égal à  $i$ . La métrique de chemin, pour un chemin  $x$  dans le treillis, peut s'écrire

$$\mu^x = \sum_{t=1}^{\tau} v_t^x \quad (21)$$

Comme dans le cas de l'algorithme symbol-by-symbol MAP, on ne peut pas séparer ici également information extrinsèque et information systématique. Ainsi la sortie d'un décodeur se partage entre la composante de probabilité à priori, d'une part, et l'ensemble information

extrinsèque plus information systématique d'autre part. L'ensemble information extrinsèque plus information systématique en sortie du premier décodeur SOVA est notée par  $\Lambda_{1,es}(c_t = i)$  et peut s'obtenir par

$$\Lambda_{1,es}(c_t = i) = \Lambda_1(c_t = i) - \log \frac{p_t(i)}{p_t(0)} \quad (22)$$

où  $\Lambda_1(c_t = i)$  est la sortie pondérée du premier décodeur et  $\log \frac{p_t(i)}{p_t(0)}$  est l'information à priori pour  $c_t = i$ . L'ensemble information extrinsèque plus information systématique est utilisée comme probabilité à priori dans l'étape suivante de décodage sous la forme

$$\Lambda_{2,es}(c_t = i) = \Lambda_2(c_t = i) - \tilde{\Lambda}_{1,es}(c_t = i) \quad (23)$$

#### **1-2-4 Performance des Turbo-modulation codées en treillis avec poinçonnage des symboles de parité**

On considère ici la performance de schémas de codage de rendement 2/3 en modulation MDP-8 avec une efficacité spectrale de 2 bits/s/Hz et pour des canaux Gaussiens. Les polynômes générateurs des codeurs à 4-8 et 16 états utilisés dans la turbo-modulation codée en treillis sont donnés en table 1 ci-dessous. Le calcul du taux d'erreur binaire se fait avec des entrelaceurs de type S-random [10] de tailles 1024 et 4096. Le nombre d'itérations est égal à 8 pour une taille d'entrelaceur de 1024 et à 18 pour une taille d'entrelaceur égale à 4096. Les résultats de simulation sont illustrés en figures 9-10 et 11. Le décodage est de type SOVA.

Trellis states	generator polynomials g <sub>0</sub>	generator polynomials g <sub>1</sub>	generator polynomials g <sub>2</sub>
4	5	6	7
8	17	15	13
16	35	27	23

**Table 1: codeurs de turbo-modulations codées en treillis rendement 2/3, modulation MDP-8**

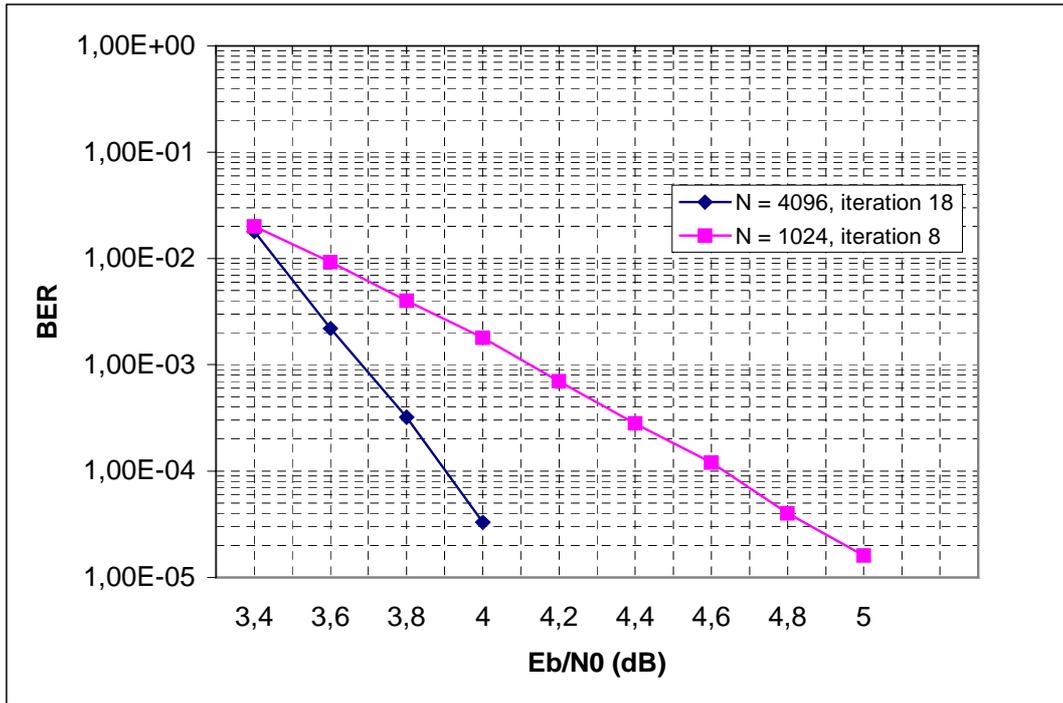


Figure 9 : Performance de turbo-modulations codées en treillis (codeurs à 4 états)

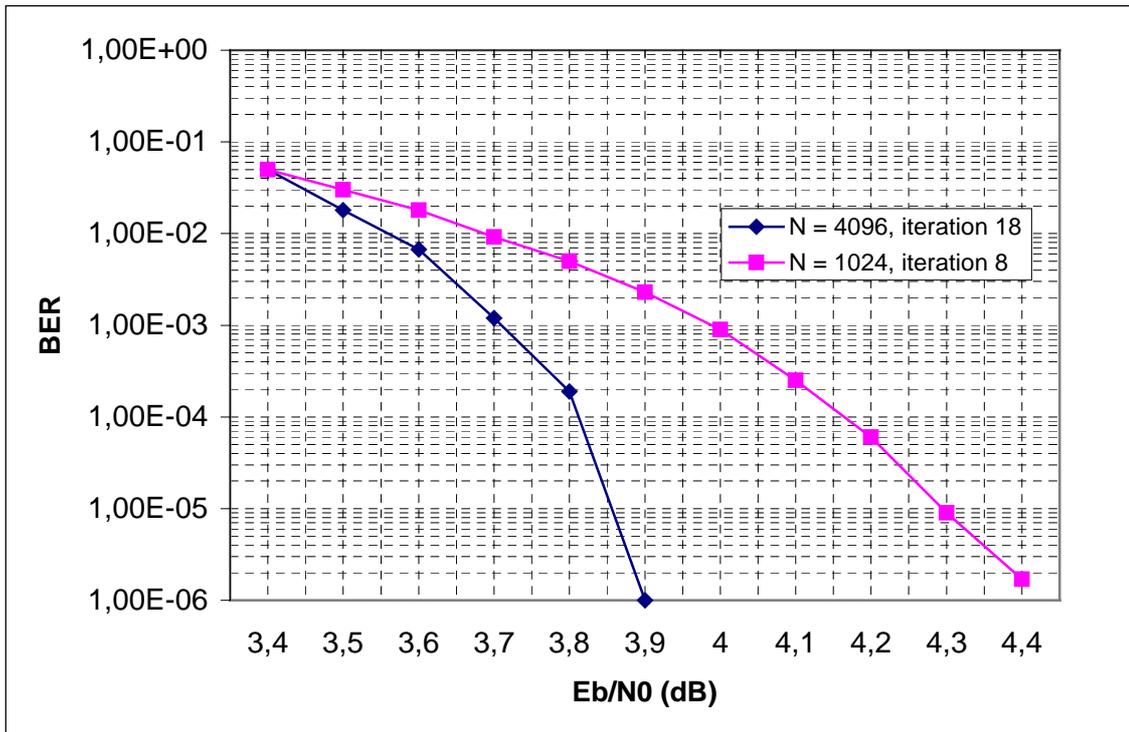


Figure 10 : Performance de turbo-modulations codées en treillis (codeurs à 8 états)

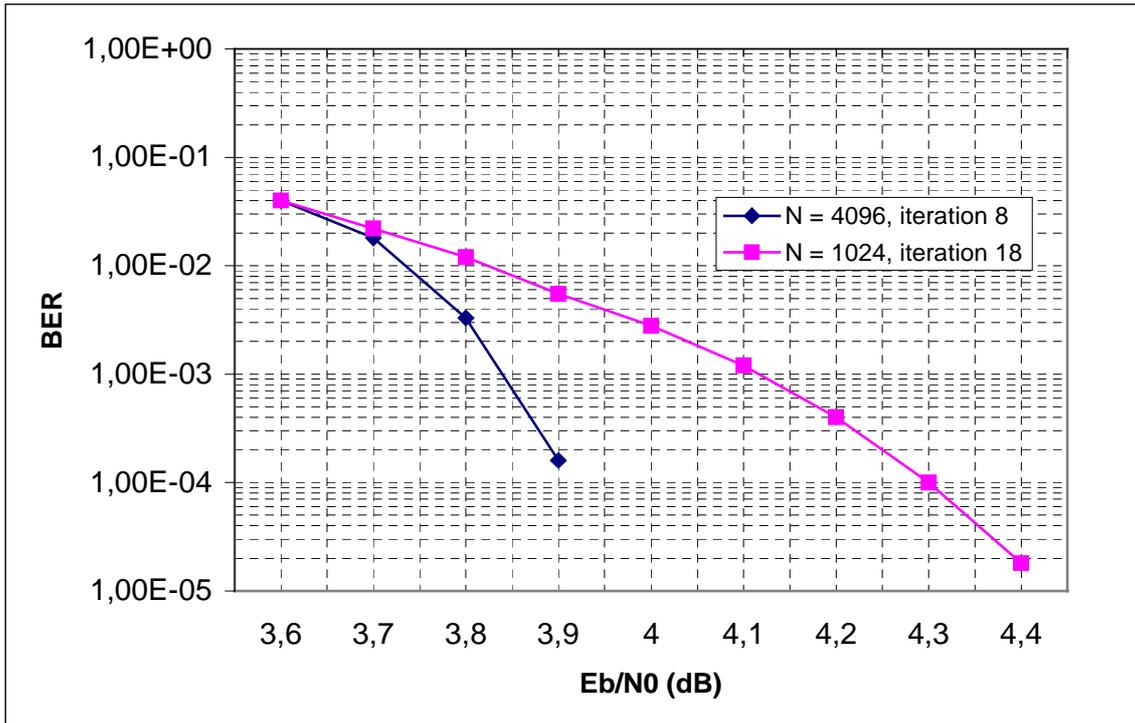


Figure 11 : Performance de turbo-modulations codées en treillis (codeurs à 16 états)

La comparaison entre les algorithmes log-MAP et SOVA est illustrée sur la figure 12 dans le cas du codeur à 8 états avec un entrelaceur de taille 1024. L'algorithme SOVA présente une perte de l'ordre de 0,4 dB pour un BER de  $10^{-3}$  et de 0,2 dB pour un BER of  $10^{-5}$ .

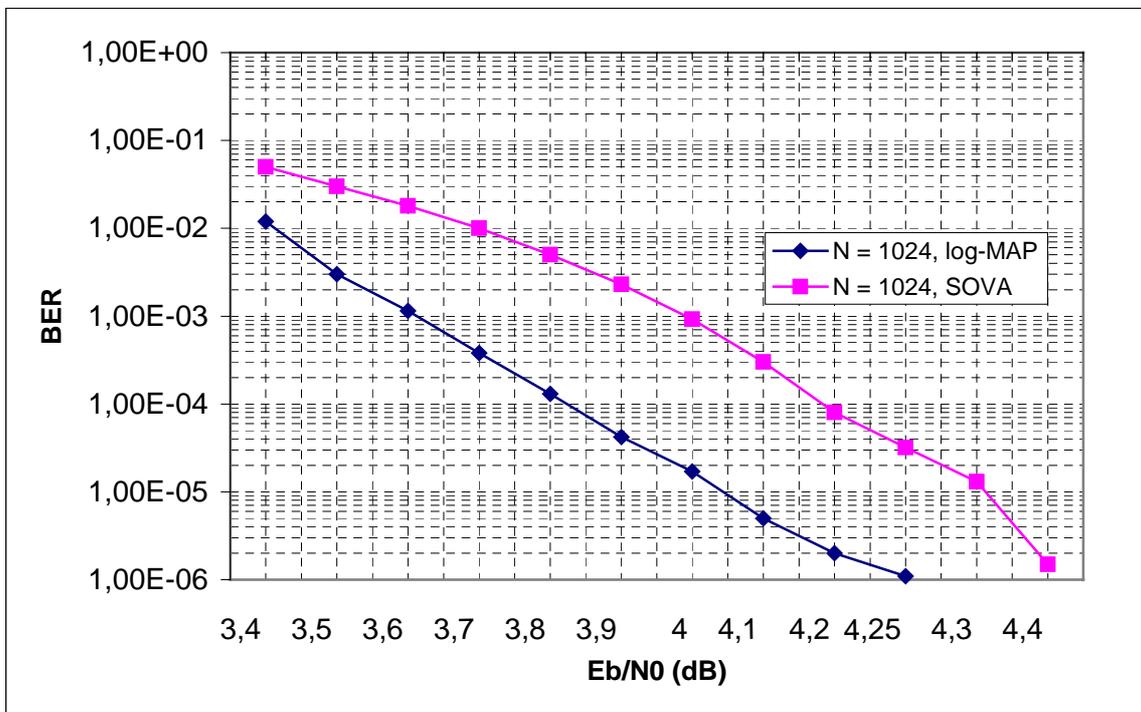


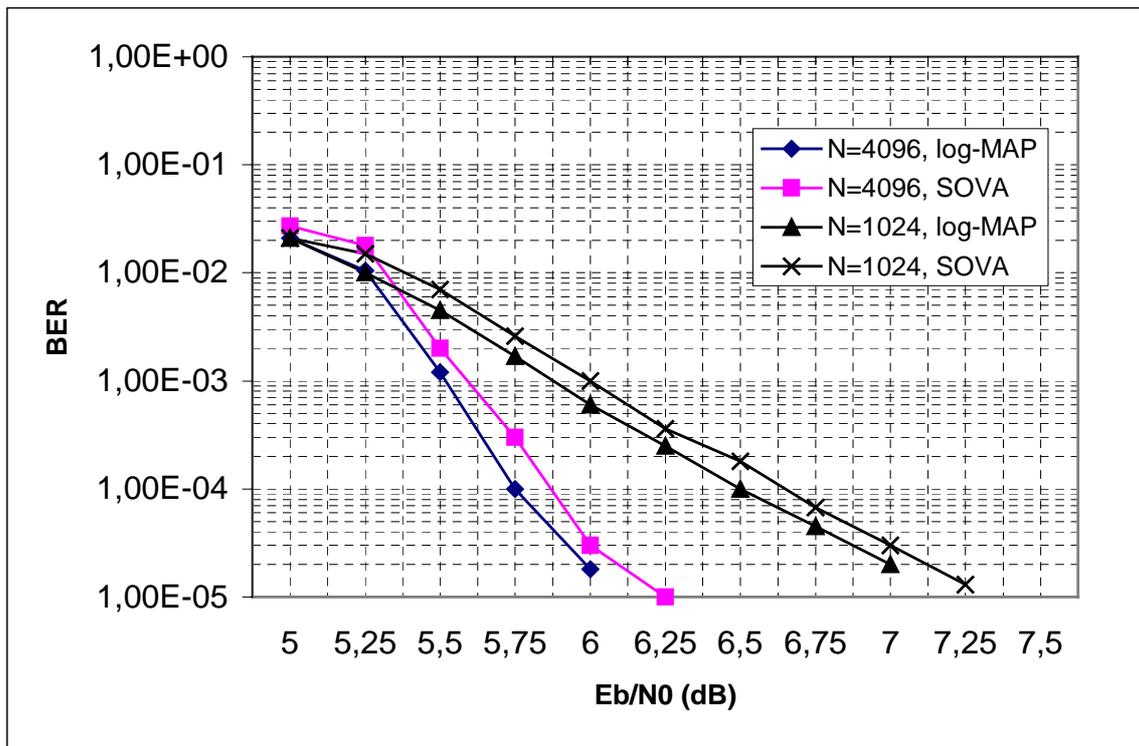
Fig. 12: Comparaison des performances des algorithmes SOVA et log-MAP pour le codeur 8 états, taille d'entrelaceur 1024.

Nous avons étudié également la performance de schémas de turbo-modulations codées en treillis pour la modulation MAQ-16 avec des codeurs de rendement 2/3 et une efficacité spectrale de 3 bits/s/Hz. Les polynômes générateurs des codes convolutifs utilisés sont indiqués sur la table 2. On considère toujours un canal de Gauss.

Trellis States	Generator Polynomials $g_0$	Generator Polynomials $g_1$	Generator Polynomials $g_2$	Generator Polynomials $g_3$
4	5	6	4	7
8	17	15	5	13
16	23	35	33	37

**Table 2: codeurs de turbo-modulations codées en treillis rendement 2/3, modulation MAQ-16**

Les performances sont illustrées sur les figures 13-15, deux tailles d'entrelaceur sont utilisées 1024 et 4096. Il est intéressant de remarquer que la différence entre les deux algorithmes (log-MAP et SOVA) est négligeable lorsque les codeurs ont un nombre d'états faible mais que cette différence augmente avec le nombre des états du codeur atteignant une valeur de 0.3 dB pour un TEB de  $10^{-5}$  dans le cas du codeur à 16 états (cf figure 15)



**Figure 13: Performance de turbo-modulations codées en treillis (codeurs à 4 états), rendement 2/3, modulation MAQ-16, 3 bits/s/Hz, algorithmes SOVA et log-MAP**

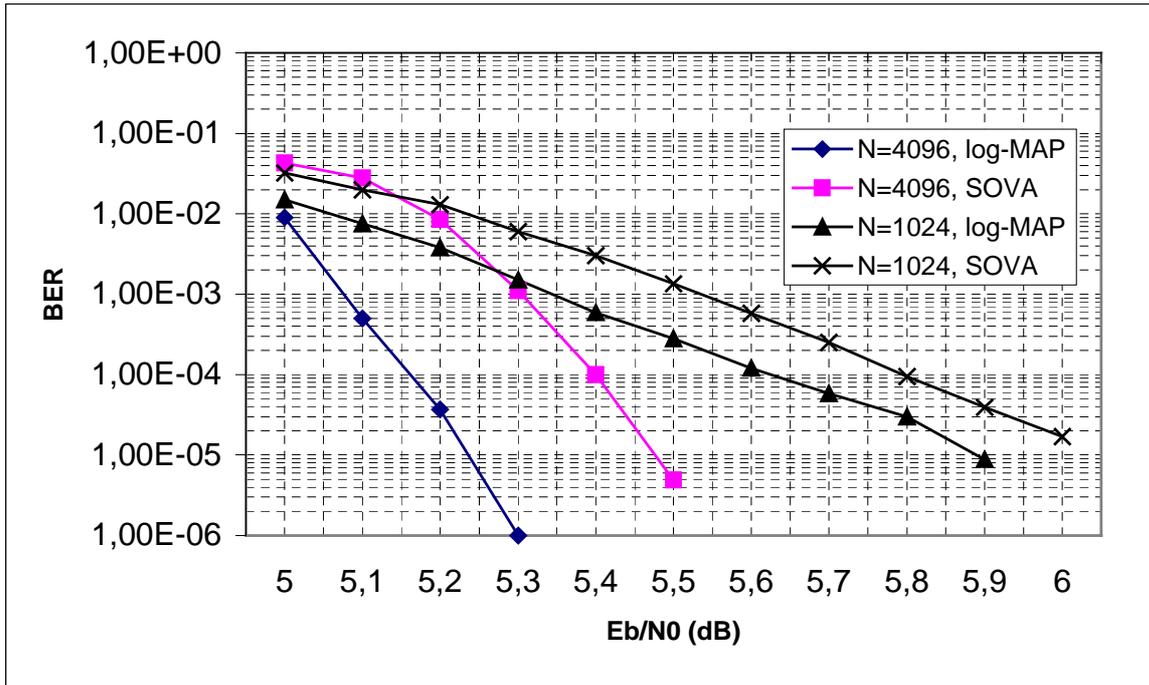


Figure 14: Performance de turbo-modulations codées en treillis (codeurs à 8 états), rendement 2/3, modulation MAQ-16, 3 bits/s/Hz, algorithmes SOVA et log-MAP

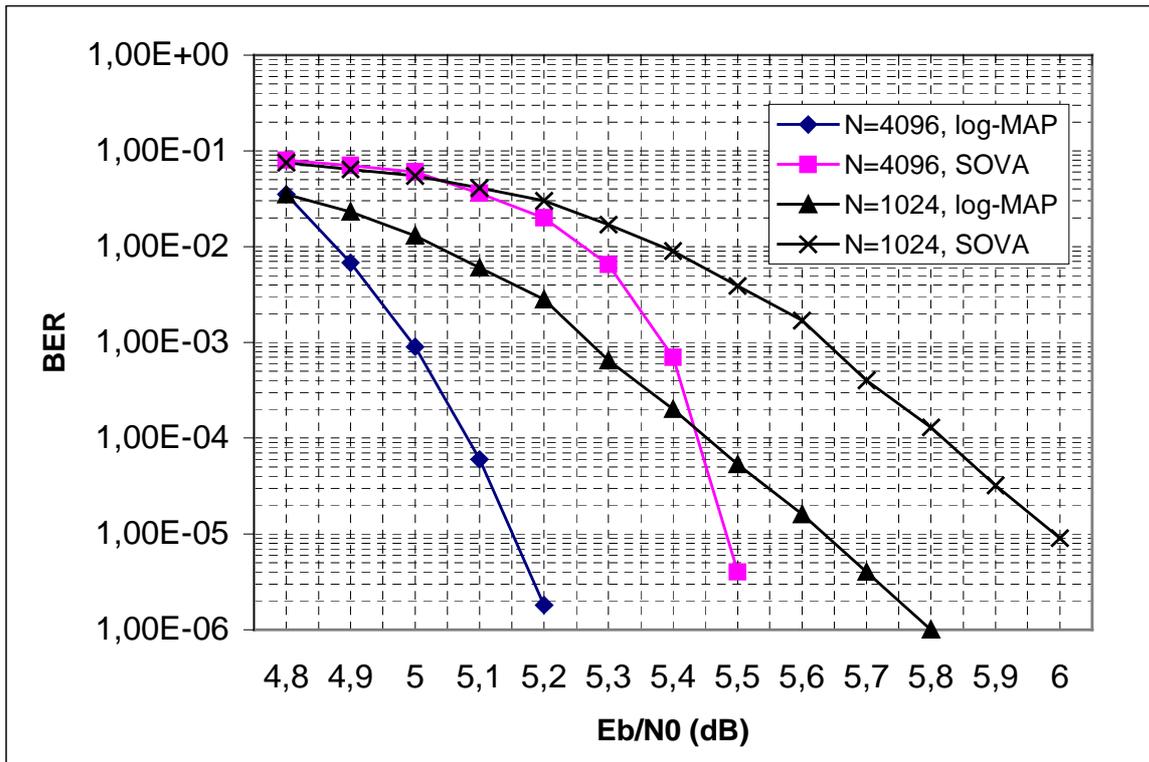


Figure 15: Performance de turbo-modulations codées en treillis (codeurs à 16 états), rendement 2/3, modulation MAQ-16, 3 bits/s/Hz, algorithmes SOVA et log-MAP

### 1-2-5 Les schémas avec poinçonnage des bits systématiques

La concaténation parallèle de deux codes convolutifs récursifs avec poinçonnage des bits systématiques a été proposée par Benedetto, Divsalar, Montorsi et Pollara [11]. L'idée de base est de poinçonner les symboles de sortie de chaque codeur et de sélectionner le motif de poinçonnage de telle sorte que les symboles de sortie du code global avec concaténation parallèle ne contiennent l'information d'entrée qu'une seule fois. Par opposition aux schémas du paragraphe précédent où l'entrelacement avait lieu au niveau des symboles et où le poinçonnage s'opérait au niveau des "digits" de parité, ce nouveau type de codeur utilise un entrelacement au niveau des bits et un poinçonnage au niveau des bits systématiques de chaque codeur convolutif. En effet, il est clair dans ce cas que le gain d'entrelacement sera supérieur permettant un meilleur comportement aux faibles SNR.

La structure globale du code utilise deux codes convolutifs élémentaires de rendement  $k/(k+1)$  en parallèle. En supposant  $k$  pair, chaque encodeur génère un bit codé et  $k$  bits d'information non codés. Pour éviter que l'on transmette deux fois chaque bit d'information, les  $k/2$  premiers bits d'information sont utilisés dans le modulateur supérieur et les  $k/2$  suivants dans le modulateur inférieur. Ceci implique que  $k/2$  bits d'information sont poinçonnés à chaque encodeur. Le nombre total de bits codés est  $k+2$ . Ces bits peuvent être utilisés sur une modulation MDP- $N$  ou MAQ- $N$  avec  $N = 2^{1+k/2}$ , et une efficacité spectrale de  $k/2$  bits/s/Hz est réalisée. Ce schéma est équivalent à une modulation multi-dimensionnelle MDP- $N$  ou MAQ- $N$  puisque deux signaux modulés sont générés dans chaque intervalle. Dans le cas de la MAQ- $N$  on peut aussi faire correspondre les  $1+k/2$  bits codés du premier encodeur au signal modulé sur la voie en phase (I-channel) et les  $1+k/2$  bits codés du second encodeur au signal modulé sur la voie en quadrature (Q-channel) avec  $N = 2^{k+2}$ . Dans ce cas, on obtient une efficacité spectrale de  $k$  bits/s/Hz.

Le décodage se fait selon des principes semblables à ceux évoqués pour le schéma précédent (figure 8), à ceci près que les entrelaceurs travaillant au niveau des bits, **on doit convertir les pondérations des symboles** en sortie du log-MAP **en pondérations sur les bits**. De même, après l'entrelaceur, on doit reconvertir les pondérations sur les bits en pondérations sur les symboles. Pour un symbole  $c$  représentant  $k$  bits d'information  $c = (c_1, c_2, \dots, c_k)$  avec  $c_j = 0, 1$ ,  $j = 1, 2, \dots, k$ , si l'on appelle  $A_e(c)$  l'information extrinsèque du symbole  $c$ , l'information extrinsèque du bit  $j$  peut se calculer par la formule :

$$A_e(c_j) = \log \frac{\sum_{c:c_j=1} e^{A_e(c)}}{\sum_{c:c_j=0} e^{A_e(c)}} \quad (24)$$

Après les opérations d'entrelacement/désentrelacement, les probabilités à priori de chaque symbole sont calculées par la formule :

$$P(c = (c_1, c_2, \dots, c_k)) = \prod_{j=1}^k \frac{e^{c_j \cdot \tilde{\Lambda}_e(c_j)}}{1 + e^{c_j \cdot \tilde{\Lambda}_e(c_j)}} \quad (25)$$

A titre d'exemple, nous montrons le schéma d'une modulation MAQ-16 avec une efficacité de 2 bits/s/Hz en figure 16. Deux codeurs à 16 états de rendement 2/3, sont utilisés. Les polynômes générateurs sont  $g_0 = (23)$ ,  $g_1 = (16)$ ,  $g_2 = (27)$ . La distance libre quadratique de ce code est égale à 7.2.

Les bits codés de chaque encodeur sont poinçonnés et envoyés sur les parties en phase et en quadrature d'une modulation MAQ-16 d'après le "mapping" naturel. Un autre schéma est montré en figure 17 pour une modulation MDP-8. Les polynômes générateurs de chaque encodeur sont  $g_0 = (23)$ ,  $g_1 = (14)$ ,  $g_2 = (16)$ ,  $g_3 = (21)$ ,  $g_4 = (31)$ . La distance libre quadratique de ce code est égale à 6.34.

Les performances du codeur MAQ-16 illustré en figure 16 sont données en figure 18 pour un canal de Gauss avec deux entrelaceurs de taille 16384 ( $128 \times 128$ ). A titre de comparaison nous avons reporté également les performances d'une turbo-modulation pragmatique MAQ-16 (cf paragraphe 1.1) avec la même efficacité spectrale et une taille d'entrelaceur égale à 32768 [16]. On voit clairement que les performances du schéma de concaténation parallèle avec poinçonnage des bits systématiques sont supérieures à celles de la turbo-modulation pragmatique avec un gain de 0.1 dB pour un TEB de  $10^{-5}$ .

Les performances du codeur MDP-8 illustré en figure 17 sont données en figure 19 pour un canal de Gauss avec quatre entrelaceurs de taille 4096 ( $64 \times 64$ ).

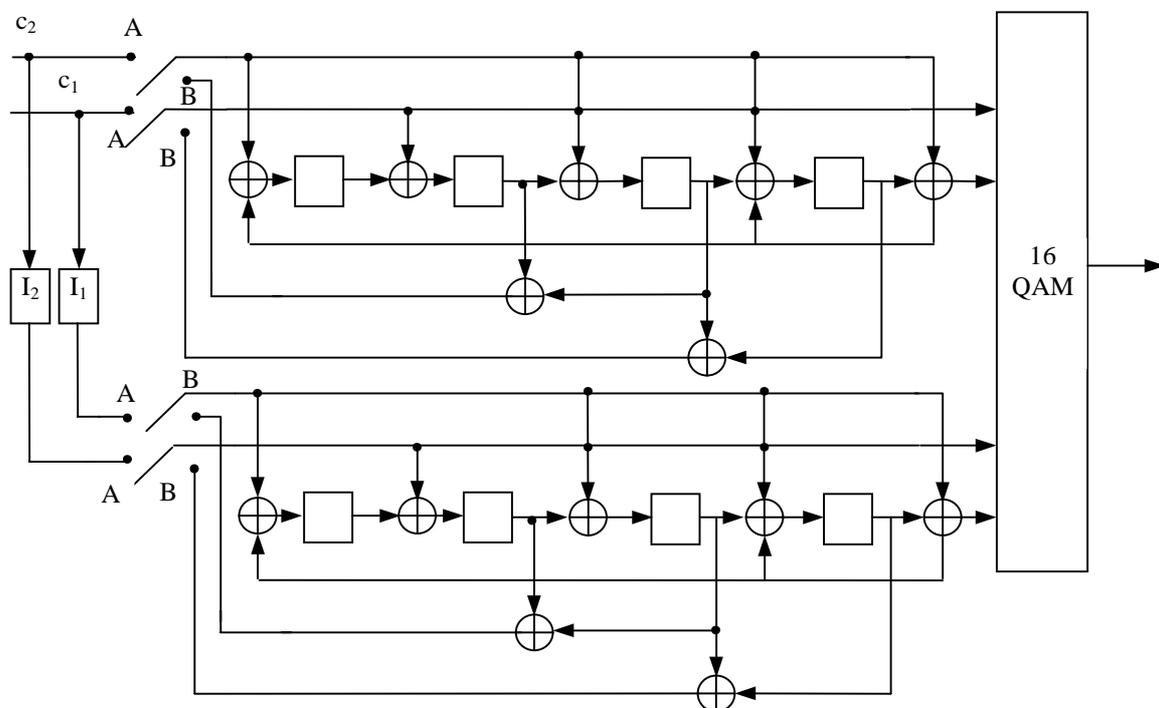
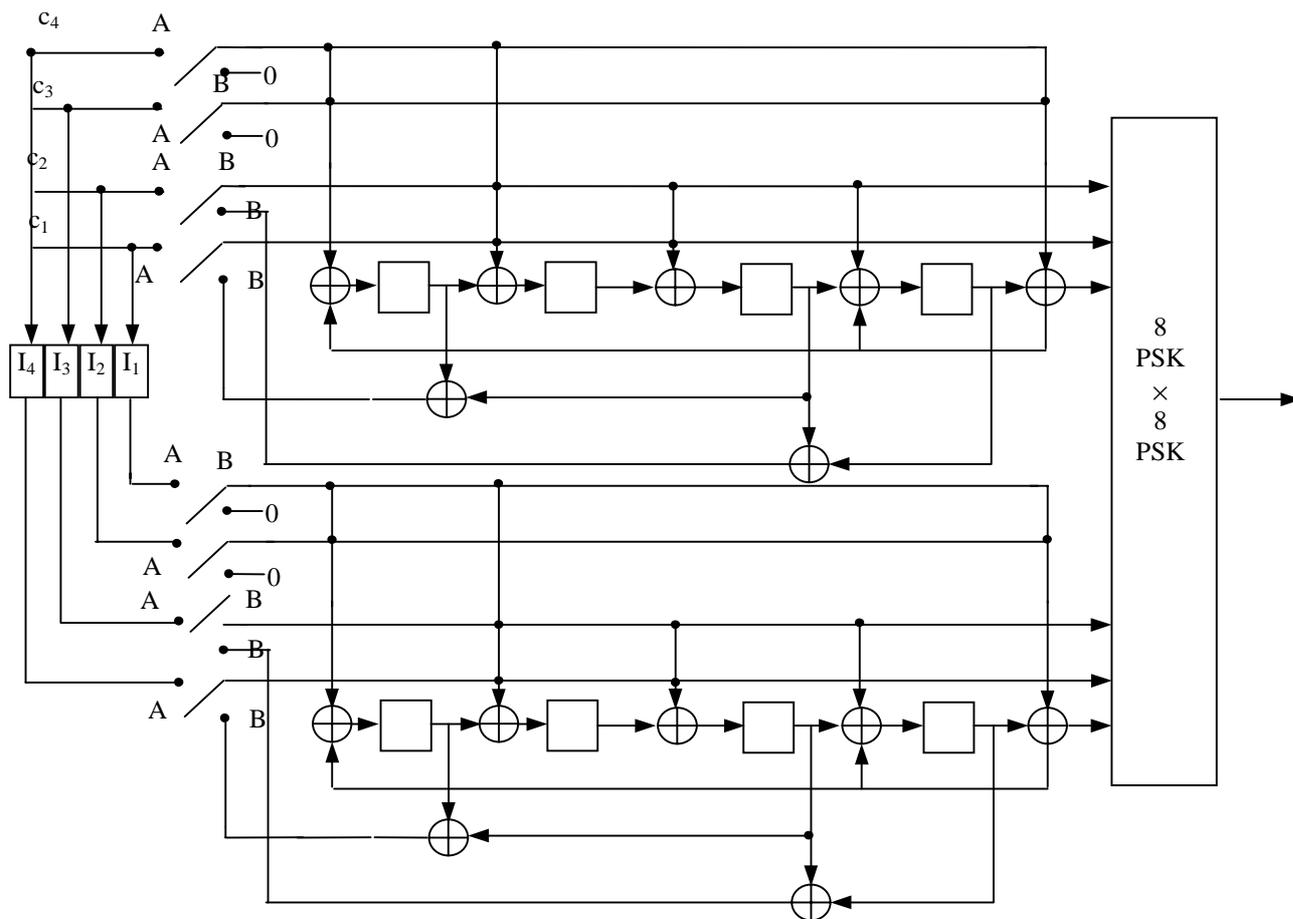


Figure 16: Turbo modulation codée en treillis MAQ-16 avec entrelacement au niveau des bits



**Figure 17: Turbo modulation codée en treillis MDP-8 avec entrelacement au niveau des bits**

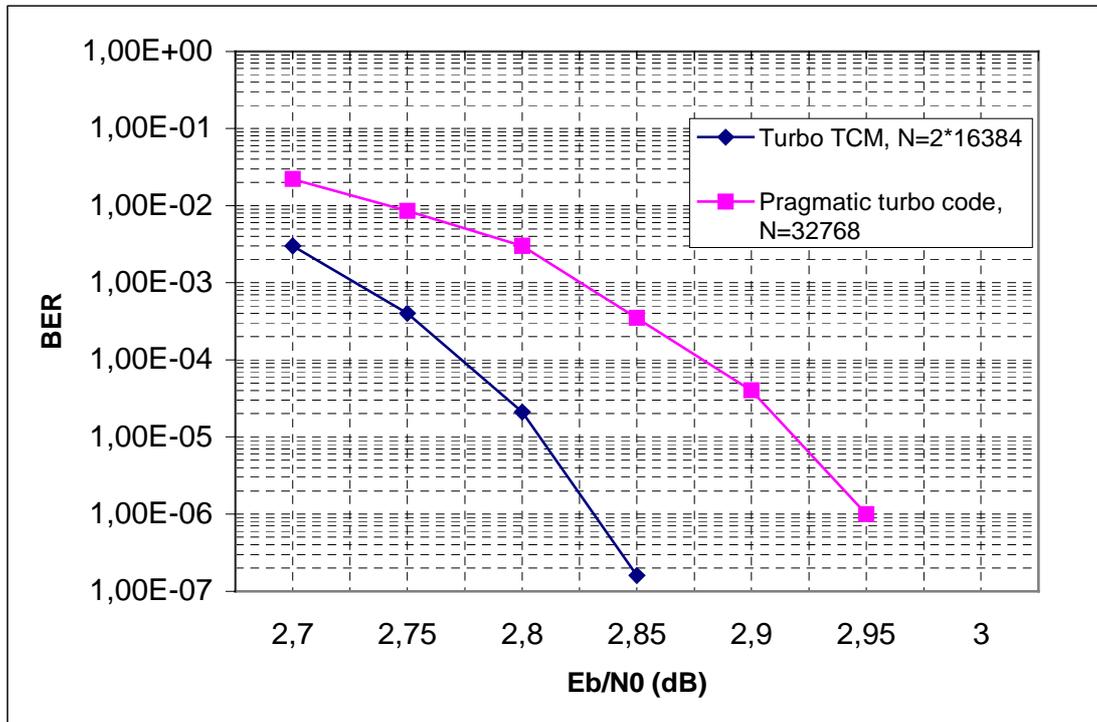


Figure 18 : Performance de la turbo-modulation codée en treillis (codeur figure 16) et comparaison avec une turbo-modulation pragmatique en MAQ-16, efficacité spectrale 2 bits/s/Hz et nombre d'itérations égal à 8.

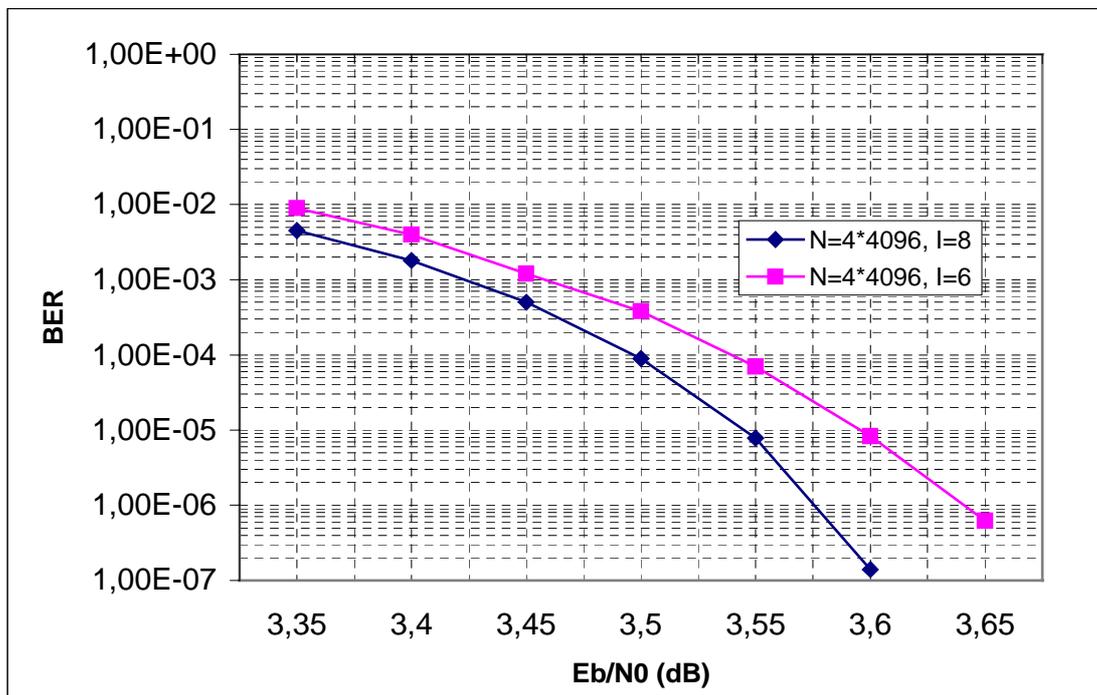


Figure 19 : Performance de la turbo-modulation codée en treillis (codeur figure 17) et comparaison avec une turbo-modulation pragmatique en MDP-8, efficacité spectrale 2 bits/s/Hz, taille d'entrelaceur 16384, nombre d'itérations: 6 et 8.

### **1-3 Les turbo-modulations codées I-Q pour les canaux à évanouissements**

La performance d'un schéma de modulation codée pour un canal de Rayleigh non-sélectif en fréquence dépend fortement de la distance minimale de Hamming entre symboles du code émis. Dans [17], Al-Semari et Fuja ont proposé des schémas I-Q TCM (Modulations Codées en Treillis en Phase et en Quadrature) qui présentent des distances minimale de Hamming entre symboles bien supérieures à celles des schémas conventionnels de modulations codées en treillis. Ils ont montré que des gains importants pouvaient être obtenus en utilisant des codes convolutifs simples optimisés en termes de distance minimale de Hamming en combinaison avec un codage et un mapping sur les voies I et Q. Dans ce paragraphe nous utilisons les turbo-codes optimisés pour le cas binaire avec un mapping sur les voies I et Q pour construire des schémas de turbo-modulations à haute efficacité spectrale [14]. Dans un turbo-codeur I-Q, chaque groupe de bits d'information est associé à ses symboles codés qui sont obtenus alternativement soit à la sortie du premier turbo-codeur, soit à la sortie du deuxième. Les bits codés en sortie des deux codeurs sont modulés séparément sur les voies en phase et en quadrature de la constellation désirée en suivant un mapping de Gray. Ce type de modulation codée s'apparente aux schémas pragmatiques déjà décrits dans la partie 1. Le décodage de ce type de modulation est assuré par un décodeur symbol-by-symbol MAP.

#### **1-3-1 Structure des turbo-modulations codées I-Q**

Nous considérons ici un schéma de décodage turbo en modulation MAQ-16 avec des codeurs de rendement 1/2 et une efficacité spectrale de 2 bits/s/Hz. La structure du codeur est illustrée sur la figure 20. Deux codeurs convolutifs récurrents systématiques identiques de rendement 1/2 et de longueur de contrainte  $v$  sont séparés par un entrelaceur aléatoire de taille  $N$ . Le premier codeur opère directement sur la séquence d'entrée.

$$c = (c_1, c_1, \dots, c_N) \quad (27)$$

La sortie contient la séquence d'information

$$v^{1s} = (v_1^{1s}, v_2^{1s}, \dots, v_N^{1s}) \quad (28)$$

et la séquence codée

$$v^{1p} = (v_1^{1p}, v_2^{1p}, \dots, v_N^{1p}) \quad (29)$$

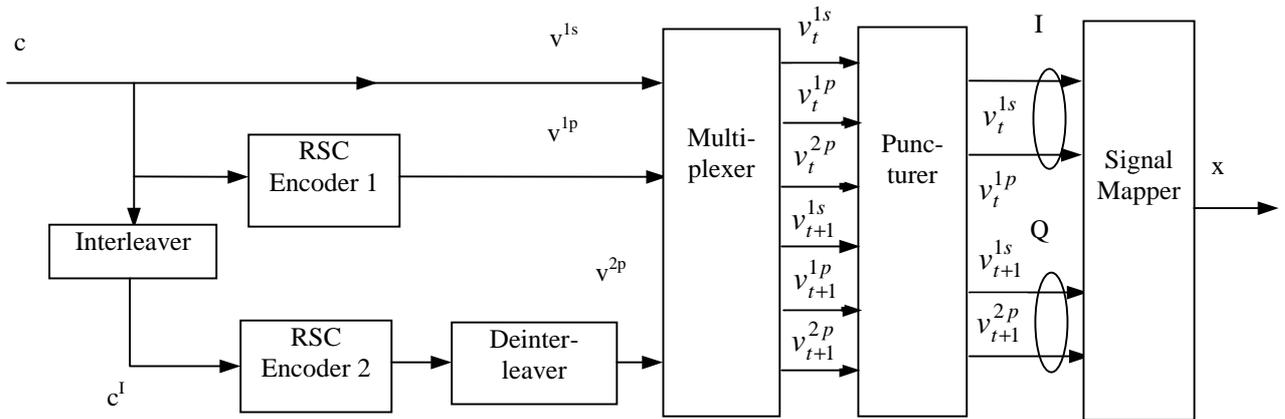
La séquence d'entrée à l'entrée du second codeur est entrelacée à l'aide d'un entrelaceur S-random de taille  $N$ . La sortie codée du second codeur est désentrelacée pour faire en sorte que la séquence de sortie donnée par

$$v^{2p} = (v_1^{2p}, v_2^{2p}, \dots, v_N^{2p}) \quad (30)$$

est ordonnée de la même façon que la séquence d'entrée  $c$ . En utilisant un poinçonnage classique pair/impair, le turbo-code complet obtenu a un rendement de 1/2 et la séquence de sortie codée, appelée  $v$  est constituée de chaque bit d'information suivi par un bit codé en provenance tantôt du premier codeur, tantôt du deuxième codeur. Si l'entrelaceur a une taille  $N$  avec  $N$  paire on peut écrire la séquence de sortie sous la forme

$$v = (v_1^{1s}, v_1^{1p}, v_2^{1s}, v_2^{2p}, \dots, v_{N-1}^{1s}, v_{N-1}^{1p}, v_N^{1s}, v_N^{2p}) \quad (31)$$

La séquence de sortie entre alors dans un modulateur MAQ-16. Dans l'exemple donné, chaque ensemble de deux bits codés vient sélectionner le signal modulé de sortie avec une correspondance (mapping) de Gray sur chacune des voies en phase et en quadrature. La sortie du premier codeur sert à sélectionner le signal modulé sur la voie en phase tandis que la sortie du second codeur sert à sélectionner le signal modulé sur la voie en quadrature.



**Fig. 20: Schéma de turbo-modulations codées I-Q en MAQ-16**

Le principe des turbo-modulations codées I-Q peuvent être facilement appliqués à différents contextes pour obtenir des schémas à grande efficacité spectrale. Par exemple, on peut utiliser a turbo-code de rendement 2/3 en MAQ-64 pour obtenir une efficacité spectrale de 4 bits/s/Hz. Les trois bits de sortie du premier codeur servent à sélectionner la composante en phase du signal MAQ-64 avec un mapping de Gray alors que les trois autres bits de sortie du deuxième encodeur servent à sélectionner la composante en quadrature du signal MAQ-64 avec un mapping de Gray. Avec un poinçonnage pair/impair un signal MAQ-64 est produit tous les deux cycles de codage.

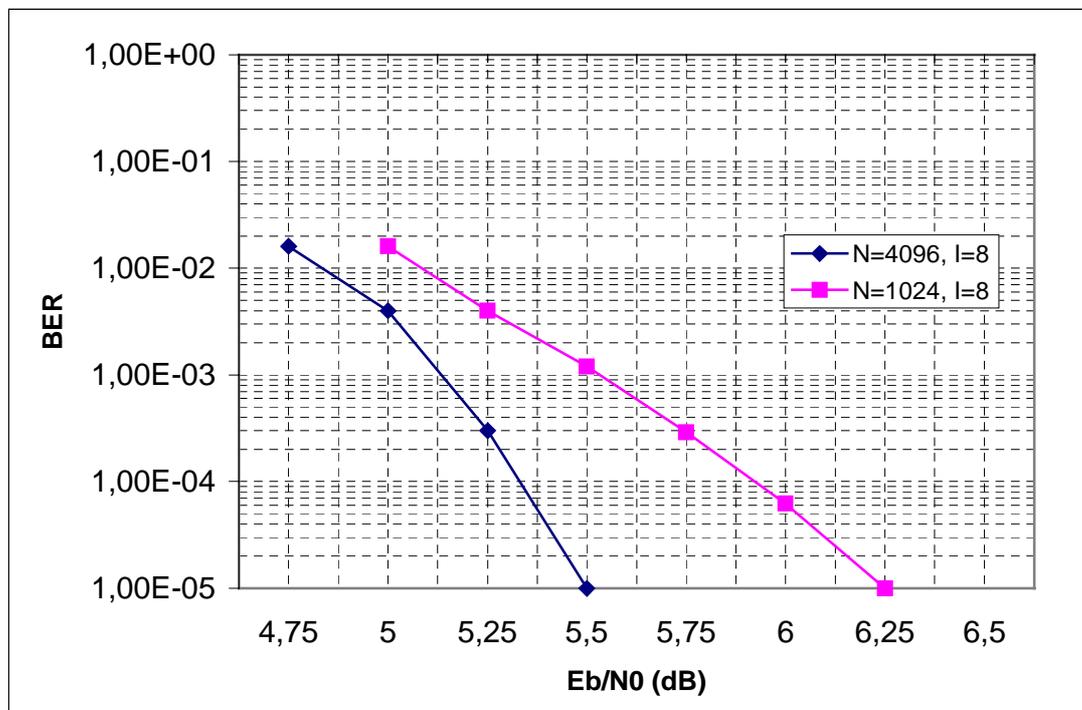
### 1-3-2 Le Decodeur

Dans le système proposé, chaque groupe de bits d'information associé à ses bits de codage est inclus dans le signal modulé de sortie. Un décodeur itératif de type symbol-by-symbol MAP

pour les treillis non-binaires (cf paragraphe 2.2) peut être appliqué. On peut travailler ici directement au niveau des symboles sans avoir à reconvertir les LLR's du niveau symboles au niveau bits. De plus on suppose ici que l'atténuation sur le canal de Rayleigh est parfaitement estimée, elle est alors incluse dans le calcul des métriques de branches dans le treillis.

### 1-3-3 Performance des turbo-modulations I-Q sur canaux de Rayleigh

Le Taux d'Erreurs Binaire (TEB) d'un schéma de codage d'efficacité spectrale 2 bits/s/Hz sur un canal de Rayleigh non-sélectif en fréquence en modulation MAQ-16 est représenté sur la figure 21. Les polynômes générateurs des codes convolutifs utilisés sont  $g_0 = (23)$  et  $g_1 = (35)$ . Dans la simulation on utilise au décodeur un algorithme de type symbol-by-symbol log-MAP. Les tailles d'entrelaceur sont 1024 et 4096.

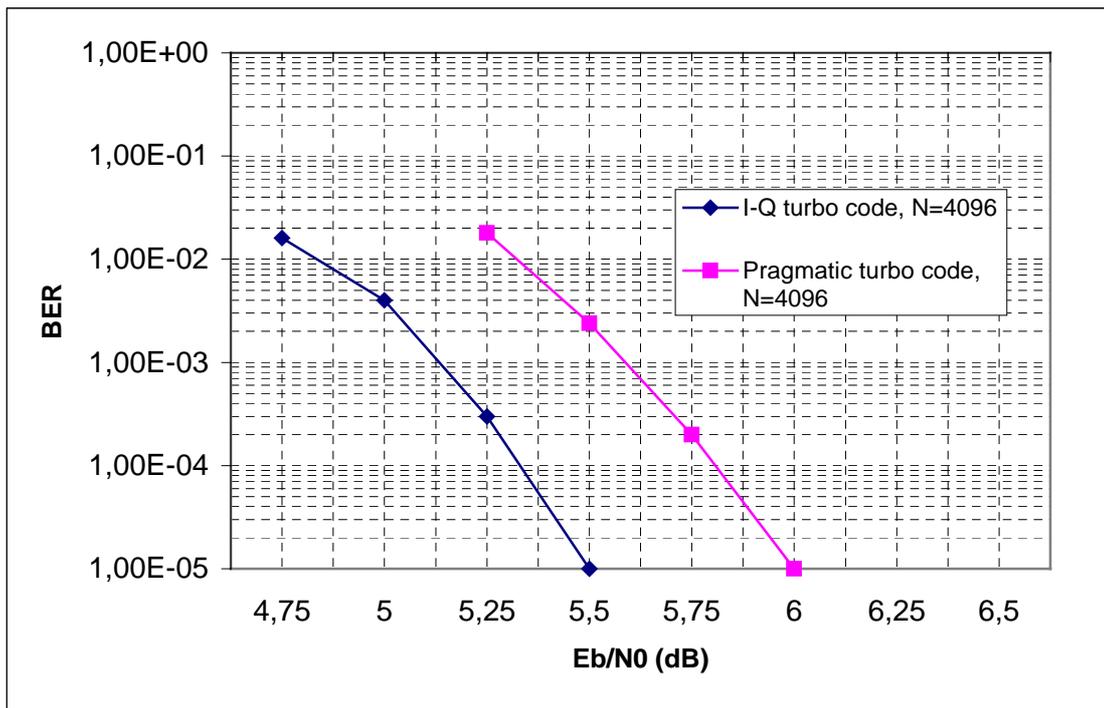


**Fig. 21: Performance d'une turbo-modulation codée I-Q en MAQ-16 avec une efficacité spectrale de 2 bits/s/Hz sur un canal de Rayleigh non-sélectif en fréquence.**

Nous avons comparé les performances de ce schéma de codage avec celles des turbo-modulations pragmatiques (pour une même efficacité spectrale et une même modulation) exposées au paragraphe 1.1.

Les performances de la turbo-modulation pragmatique correspondante peuvent être trouvées dans [4]. Nous utilisons 8 itérations dans l'algorithme symbol-by-symbol log-MAP avec un entrelaceur de taille 4096. Les résultats sont donnés en figure 22. Il est clair d'après la courbe présentée que le schéma de turbo-modulation codée I-Q présente, à taille d'entrelaceur et

efficacité spectrale égales, des performances supérieures au schéma de turbo-modulation codée pragmatique. Le gain est de l'ordre de 0.5 dB à  $TEB=10^{-5}$ .



**Fig. 22: Comparaison des performances des turbo-modulations codées I-Q et des turbo-modulations codées pragmatiques en MAQ-16, efficacité spectrale 2 bits/s/Hz, taille d'entrelaceur 4096, canal de Rayleigh non-sélectif en fréquence**

## 1-4 Conclusion

Nous avons présenté dans ce paragraphe les principaux schémas de codage associant turbo-codes et modulations à haute efficacité spectrale. Pour le canal de Gauss, il semble que les turbo-modulations codées en treillis s'inspirant du principe de partitionnement en sous-constellations de Ungerboeck donnent les meilleurs résultats. Pour le canal de Rayleigh non-sélectif en fréquence, les turbo-modulations codées I-Q basées sur le choix de codeurs convolutifs à distance minimale optimisée constituent le meilleur compromis performances-complexité.

Il est à signaler pour terminer que nous présentons en annexe 2 un système de codage turbo-modulation codée en treillis associé à un code spatio-temporel en bloc de type Alamouti. Les performances remarquables obtenues avec ce dispositif, en particulier avec deux antennes de réception, montrent bien, malgré leur relative complexité, l'intérêt d'utiliser de tels schémas de codage.

## **1-5 Références**

- [1] G. Ungerboeck, "Channel coding with multilevel/phase signaling", *IEEE Trans. Inform. Theory*, vol. 25, n°1, Jan. 1982, pp. 55-67.
- [2] J. Hagenauer, P. Robertson, and L. Papke, "Iterative ('Turbo') decoding of systematic convolutional codes with the MAP and SOVA algorithms", in *Proc. of the 1994 ITG Conference on Source and Channel Coding*, Munich, Oct. 1994, pp. 1-9.
- [3] B. Vucetic, "Iterative decoding algorithm", in *PIMRC'97*, Helsinki, Finland, Sep. 1997, pp.99-120.
- [4] S. LeGoff, A. Glavieux and C. Berrou, "Turbo codes and high efficiency modulation", in *Proc. of IEEE ICC'94*, New-Orleans, LA, May 1994, pp. 645-649.
- [5] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes", *IEEE Trans. Inform. Theory*, vol. 23, n°3, May 1997, pp. 371-377.
- [6] P. Robertson, and T. Woerz, "Novel coded modulation scheme employing turbo codes", *Electronics Letters*, vol. 31, n°18, Aug. 1995, pp. 1546-1547.
- [7] P. Robertson, and T. Woerz, "Bandwidth efficient turbo trellis-coded modulation using punctured component codes", *IEEE Journal on Selected Areas in Communications*, vol. 16, n°2, Feb. 1998, pp. 206-218.
- [8] L.U Wachsmann, and J. Huber, "Power and bandwidth efficient digital communication using turbo codes in multilevel codes", *European Trans. Telecommun.*, vol. 6, n°5, Sept./Oct. 1995, pp. 557-567.
- [9] L. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Trans. Inform. Theory*, vol. IT-20, pp; 284-287, March 1979.
- [10] D. Divsalar and F. Pollara, "Turbo codes for PCS applications", in *Proc. ICC'95*, Seattle, WA, June 1995, pp. 54-59.
- [11] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Parallel concatenated trellis coded modulation", in *Proc. of IEEE ICC'96*, 1996, pp. 974-978.
- [12] T. M. Duman, "Turbo codes and turbo-coded modulation: Analysis and performance bounds", Ph.D. dissertation, *Elect. Comp. Eng. Dep.*, Northeastern Univ., Boston, MA, May 1998.
- [13] D. Divsalar and F. Pollara, "On the design of turbo codes", *TDA Progress Report 42-123*, JPL, Nov. 1995, pp. 99-121.
- [14] J. Yuan, B. Vucetic, and W. Feng, "I-Q turbo coded modulation", in *Proc. of IEEE GLOBECOM'98*, *Communications Theory Mini Conference*, Sydney, Australia, Nov. 1998, pp. 191-195.

- [15] W. Blackert and S. Wilson, "Turbo trellis coded modulation", in Proc. CISS'96, 1996.
- [16] S.A. Barbulescu, W. Farnell, P. Gray, and M. Rice, "Bandwidth efficient turbo coding for high speed mobile satellite communications", in Proc. International Symposium on Turbo Codes, Brest, France, Sept. 1997, pp. 119-126.
- [17] S.A. AL-Semari and T.E. Fuja, "I-Q TCM: Reliable communication over the Rayleigh fading channel close to the cutoff rate", IEEE Trans. Inform. Theory, vol. 43, n°1, Jan. 1997, pp. 250-262.



## ***Annexe 1 : Modulations Codées en Treillis***

Le but de cette annexe est de définir les principes de base de la réalisation des MCT (Modulations Codées en Treillis) qui associent modulation et codage de telle sorte que la distance minimale entre deux signaux codés soit la plus grande possible. Ceci se fait à l'aide d'un mapping particulier qui associe aux bits codés les symboles émis dans la constellation de telle sorte que la distance euclidienne minimale entre eux soit maximisée. Les MCT présentent une bonne efficacité spectrale et permettent avec des codes simples d'obtenir d'excellentes performances. Nous allons examiner les principes de base de réalisation du codage MCT à l'aide des règles données par Ungerboeck.

### **Principes de Partitionnements successifs d'une constellation :**

Les principes cités ci-dessous sont dûs à Ungerboeck (1982). Pour illustrer leur application, nous allons considérer le cas d'une modulation MDP-8 (figure 1). La constellation est partitionnée en sous-constellations avec des distances croissantes entre les points des sous-constellations. Dans le cas de la MDP-8, les points de la constellation sont situés sur un cercle de rayon  $\sqrt{\mathcal{E}}$  et ont une distance minimale de séparation égale à :

$$d_0 = 2 \cdot \sqrt{\mathcal{E}} \cdot \sin\left(\frac{\pi}{8}\right) = \sqrt{(2 - \sqrt{2}) \cdot \mathcal{E}} = 0.765 \cdot \sqrt{\mathcal{E}}$$

Dans la première sous-constellation, les 8 points sont divisés en deux sous-constellations de 4 points chacune de telle sorte que la distance minimale entre les points de ces sous-constellations augmente jusqu'à une valeur  $d_1 = \sqrt{2 \cdot \mathcal{E}}$ . Pour le prochain partitionnement on obtient alors 4 sous-constellations de deux points chacune, dont la distance minimale entre les éléments devient égale à  $d_2 = 2 \cdot \sqrt{\mathcal{E}}$ . Et finalement, on obtient à la fin des partitionnements, 8 sous-constellations contenant chacune un point.

Les principes de partitionnement cités ci-dessus peuvent s'appliquer au cas d'une modulation MAQ 16 comme le montre la figure 2. On commence par diviser la constellation initiale en deux sous-constellations en prenant un point sur deux dans la constellation de départ (points alternés). Ainsi, la distance minimale entre les points dans ces deux sous-constellations devient égale à  $2 \cdot \sqrt{2 \cdot \mathcal{E}}$  alors qu'elle vaut  $2 \cdot \sqrt{\mathcal{E}}$  pour la constellation de départ. On continue alors à partitionner en sous-constellations comme illustré sur la figure 2. Il est intéressant de noter que dans le cas de la modulation MAQ-16, chaque nouveau niveau de partitionnement augmente la distance euclidienne minimale entre les points des constellations d'un facteur  $\sqrt{2}$  soit,  $\frac{d_{i+1}}{d_i} = \sqrt{2}$  pour  $\forall i$ .

Dans ces deux exemples, le partitionnement a été effectué jusqu'à la limite où chaque sous-constellation finale ne contient plus qu'un point. En général, il n'est pas nécessaire d'aller jusque là. Par exemple, dans le cas de la MAQ 16, on peut se contenter d'un seul partitionnement qui aboutit à quatre sous-constellations de quatre points chacune. De la même

façon, dans le cas de la MDP 8 (8 PSK), on peut se limiter à deux partitionnements successifs qui conduisent à quatre sous-constellations de deux points chacune.

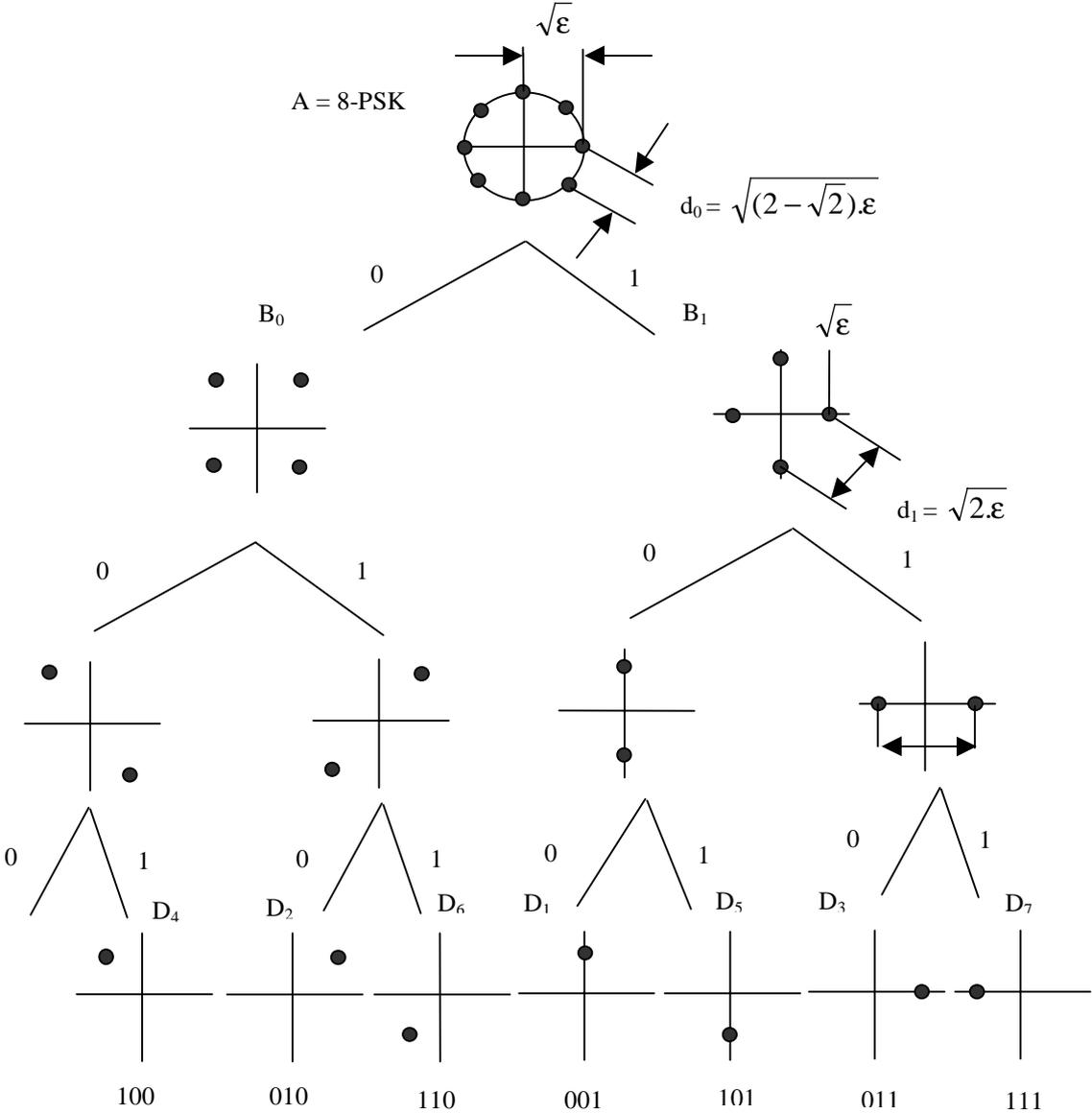
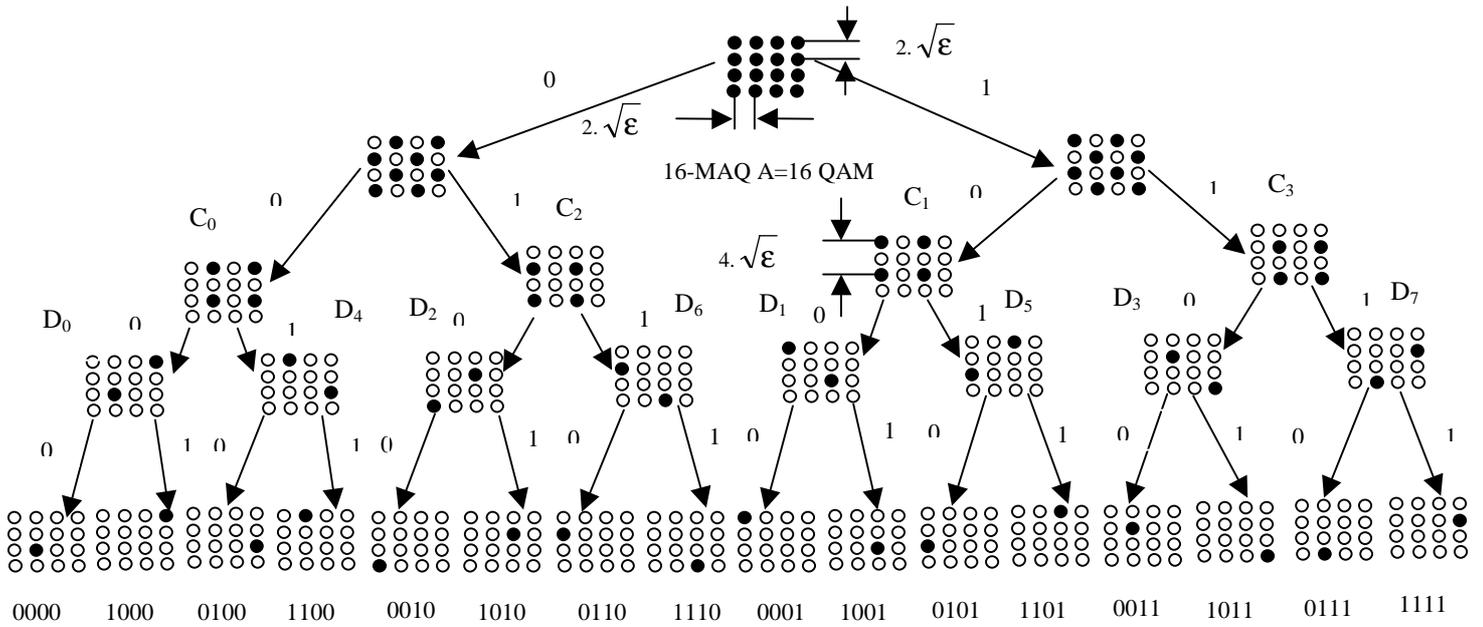
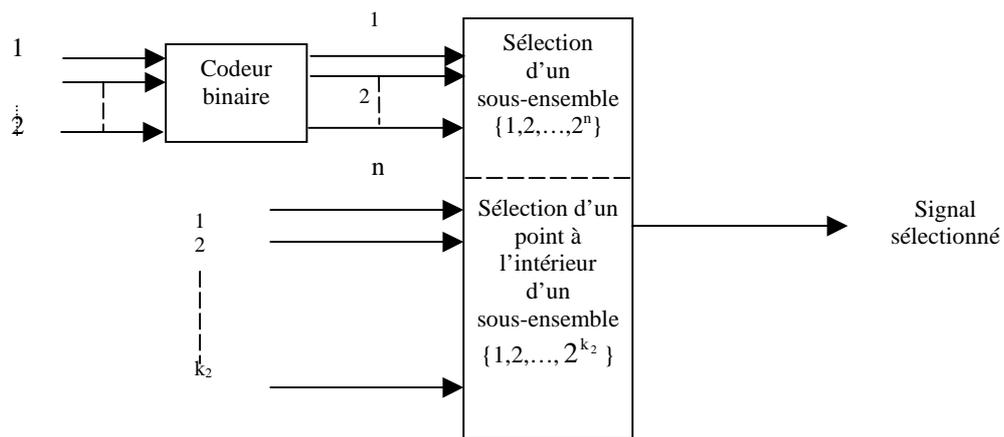


Fig. 1 : partitionnement de constellation MDP-8



**Fig. 2: partitionnement de constellation MAQ-16**

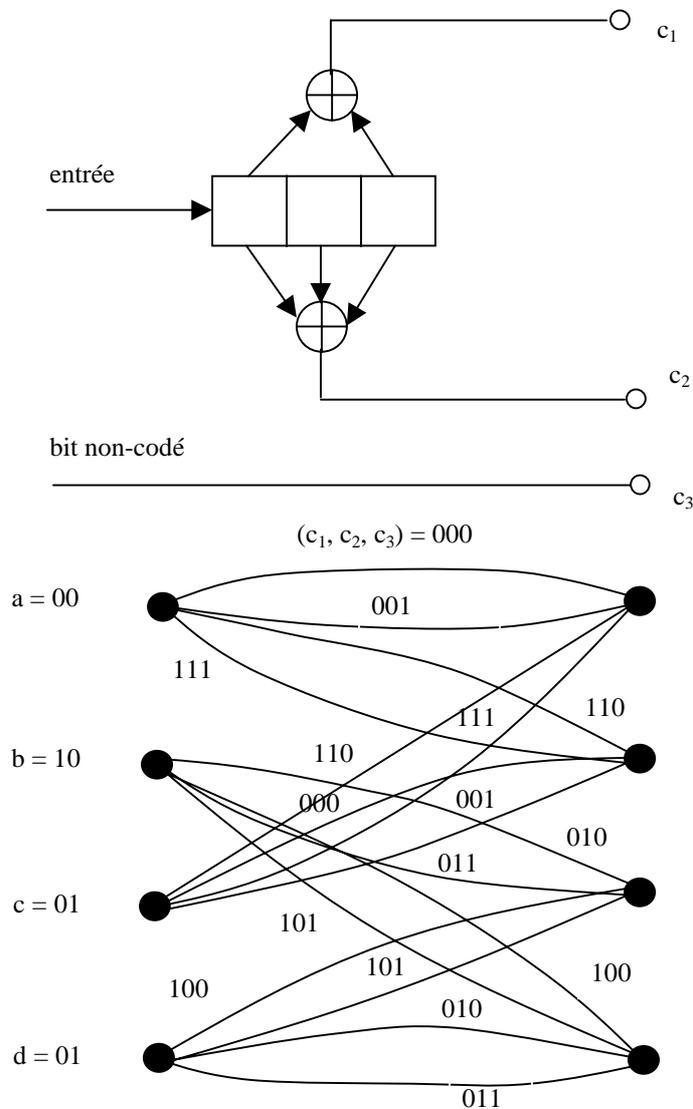
Le degré de partitionnement d'une constellation dépend des caractéristiques du code utilisé. De façon générale, le processus de codage se fait selon le processus illustré sur la figure 3. Un bloc de  $m$  bits d'information est séparé en deux groupes de longueurs  $k_1$  et  $k_2$ . Les  $k_1$  bits sont codés et transformés en  $n$  bits alors que les  $k_2$  bits restent non-codés. Ensuite, les  $n$  bits du codeur sont utilisés pour sélectionner une des  $2^n$  sous-constellations possibles dans le partitionnement de la constellation pendant que les  $k_2$  bits sont utilisés pour sélectionner un des  $2^{k_2}$  points à l'intérieur de chaque sous-constellation. Quand  $k_2 = 0$ , tous les  $m$  bits d'information sont codés.



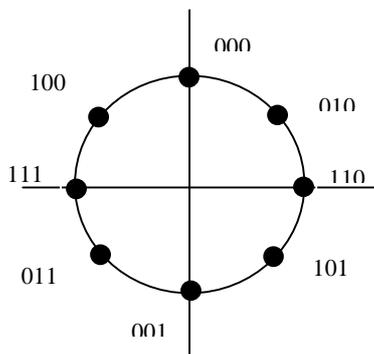
**Fig. 3: Structure générale d'un ensemble codeur/modulateur**

Nous allons maintenant étudier un exemple de codeur utilisant les principes mentionnés ci-dessus. On considère pour cela le codeur de la figure 4 ci-dessous qui utilise un code

convolutif de rendement  $\frac{1}{2}$  pour coder un bit alors que le second bit d'information n'est pas codé. Utilisé conjointement avec une modulation MDP-8, les deux bits codés servent à sélectionner une des quatre sous-constellations dans le diagramme de la figure 1 alors que le bit non codé est utilisé pour sélectionner un des deux points à l'intérieur de chacune des quatre sous-constellations. Dans ce cas, on a  $k_1 = 1$  et  $k_2 = 1$ . Le treillis à quatre états de la figure 4 correspond au treillis classique du codeur convolutif de rendement  $\frac{1}{2}$  avec l'ajout de branches parallèles pour chaque transition dûes à la présence du bit non codé. Ainsi, les bits codés ( $c_1, c_2$ ) sont utilisés pour sélectionner une des quatre sous-constellations contenant deux signaux chacune, tandis que le bit non codé est utilisé pour sélectionner un des deux signaux dans chaque sous-constellation. On peut remarquer que les signaux à l'intérieur d'une sous-constellation sont séparés d'une distance  $d_2 = 2\sqrt{\mathcal{E}}$ . La distance euclidienne entre les branches parallèles est égale à  $d_2$ . La correspondance des bits codés ( $c_1, c_2, c_3$ ) avec le signal émis en MDP-8 pour le codeur de la figure 4 est illustré sur la figure 5.

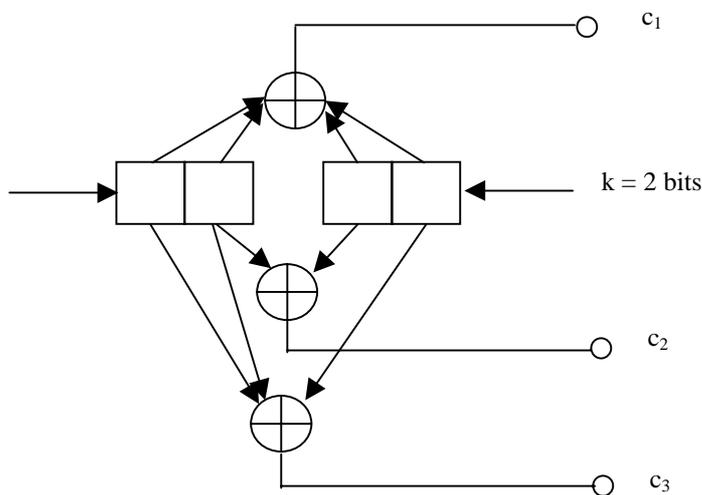


**Fig. 4: codeur pour TCM MDP-8 et diagramme en treillis**



**Fig. 5: Constellation MDP-8 pour le codeur précédent**

Il est bien évidemment possible d'utiliser un code convolutif plus puissant du type 2/3 et ainsi de coder les deux bits d'information. Ce nouveau code donne un treillis à huit états et permet d'obtenir de meilleures performances. Le codeur est représenté sur la figure 6 ci-dessous.



**Fig. 6 : codeur convolutif 2/3**

Nous reprenons ici l'exemple d'une constellation MDP8 utilisé conjointement avec un code convolutif. La modulation MDP4 non-codée est utilisée comme référence pour mesurer le gain de codage. La modulation MDP4 non-codée utilise les indifféremment les points des sous-constellations  $B_0$  et  $B_1$  de la figure 1 pour lesquelles la distance minimale entre les signaux est égale à  $\sqrt{2\mathcal{E}}$ . On peut remarquer que ce signal correspond à un treillis très simple à un seul état avec quatre branches parallèles comme le montre la figure 7(a). Les sous-constellations  $D_0, D_2, D_4$  et  $D_6$  sont utilisées en tant que points dans la constellation.

Dans le cas de la modulation codée MDP8, nous pouvons utiliser le treillis à quatre états de la figure 7(b). Chaque branche dans le treillis correspond à une des sous-constellations  $C_0, C_1, C_2$  ou  $C_3$ . Pour la constellation à huit points, chacune des sous-constellations  $C_0, C_1, C_2$  ou  $C_3$  contient deux signaux. Ainsi, l'état de transition  $C_0$  contient les points de la constellation correspondant aux bits (000, 100) ou encore (0, 4) en notation octale. De la même manière,  $C_2$  contient les points de la constellation correspondant aux bits (010, 110), ou à (2, 6) en

notation octale.  $C_1$  contient les points correspondants à (001, 101) ou (1, 5) en octal et  $C_3$ , les points correspondants à (011, 111) ou (3, 7) en octal. Ainsi, chaque transition dans le treillis à quatre états, contient deux chemins parallèles comme le 7(c). On peut remarquer que chaque paire de chemins qui diverge à partir d'un état donné et qui reconverge vers ce même état après plusieurs transitions possède une distance euclidienne au carré entre eux égale à:  $d_0^2 + 2.d_1^2 = d_0^2 + d_2^2$ . Par exemple, les chemins 0, 0, 0 et 2, 1, 2 sont séparés de la distance:  $d_0^2 + d_2^2 = [(0.765)^2 + 4]\epsilon = 4.585.\epsilon$ . D'un autre côté, la distance euclidienne au carré entre les transitions parallèles est égale à:  $d_2^2 = 4.\epsilon$ . On en conclut que la distance euclidienne minimale entre deux chemins qui divergent à partir d'un état donné et reconvergent vers ce même état dans le cas d'un treillis à quatre états est égale à:  $d_2 = 2.\sqrt{\epsilon}$ . Cette distance minimale dans le treillis est appelée la distance euclidienne libre et est notée  $D_{\text{libre}}$ .

Dans le cas du treillis à quatre états de la figure 7(b),  $D_{\text{libre}} = 2.\sqrt{\epsilon}$ . Lorsque nous comparons à la distance euclidienne  $d_0 = \sqrt{2.\epsilon}$  de la modulation MDP4 non-codée, nous constatons que le treillis à quatre états conduit à un gain de codage égal à 3 dB.

Le treillis à quatre états illustré sur la figure 7(b) est optimum au sens de la maximisation de la distance libre. Il est clair qu'il existe d'autres treillis à quatre états comme celui représenté sur la figure 8 où il existe quatre transitions distinctes depuis chaque état vers tous les autres états, mais aucun de ces codes ne conduit à une valeur de  $D_{\text{libre}}$  plus élevée.

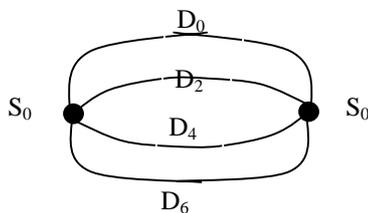


Fig. 7(a): treillis à un état

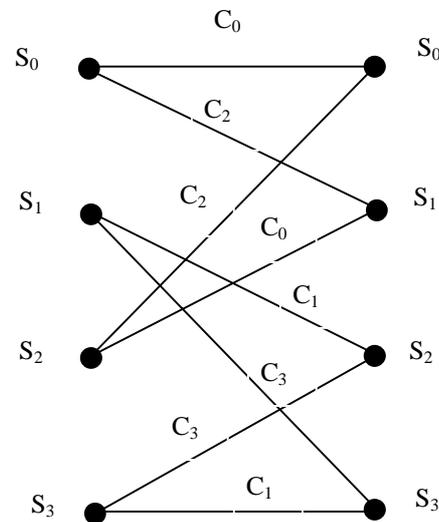


Figure 7(b) : treillis à quatre états

La construction du treillis optimal pour la constellation MDP8 est réalisée à partir des règles de base suivantes :

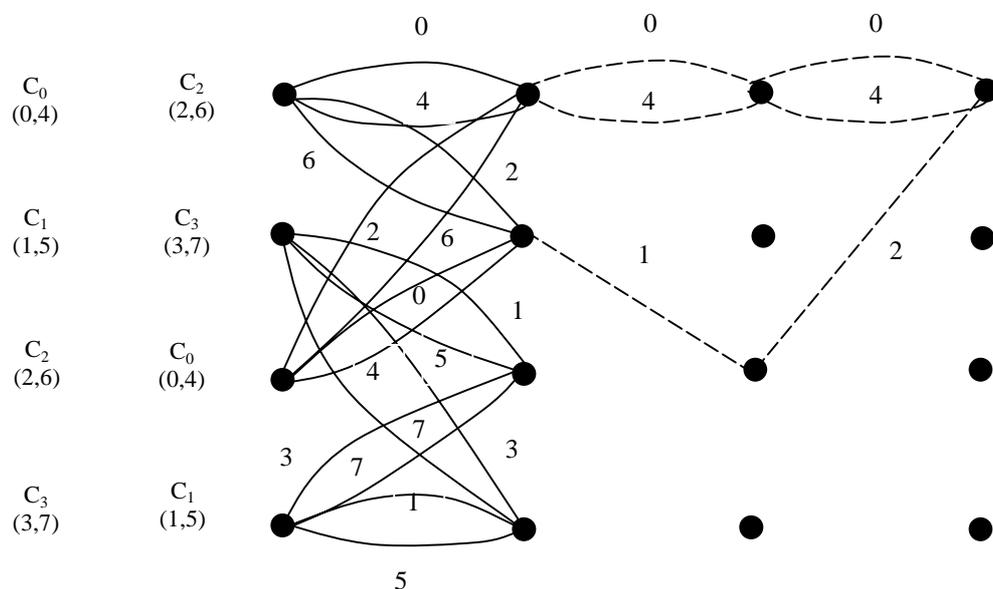
1. Les transitions parallèles lorsqu'elles se produisent sont affectés aux points qui sont les plus éloignés dans la constellation, c'est-à-dire  $d_2 = 2.\sqrt{\epsilon}$  pour la MDP8 dans les quatre sous-constellations  $C_0, C_1, C_2, C_3$ .

2. Les transitions d'un état à l'autre sont affectées aux sous-constellations  $(C_0, C_2)$  ou  $(C_1, C_3)$  qui ont une distance maximale  $d_1 = \sqrt{2}\epsilon$ .
3. Les signaux de la constellation doivent être choisis de façon équiprobable.

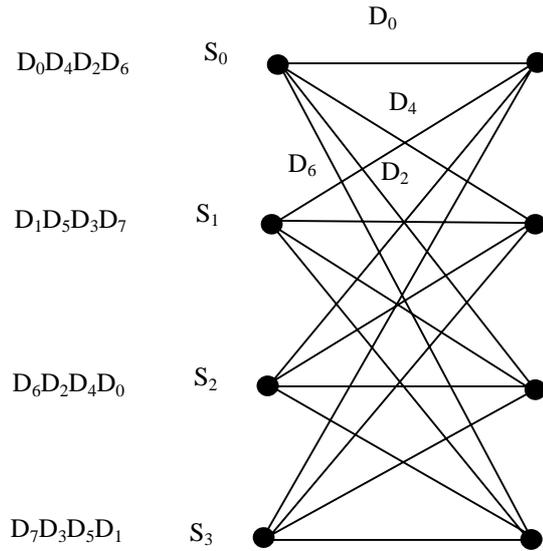
Les règles (1) et (2) garantissent que la distance euclidienne associée aux chemins qui divergent à partir d'un état donné et reconvergent vers ce même état excède la distance euclidienne de la MDP4 non-codée. La règle (3) garantit que le treillis du code a une structure régulière.

La correspondance bits codés et signaux de la constellation n'est pas primordiale et la solution présentée sur la figure 1 n'est pas unique: d'autres correspondances peuvent être appliquées en permutant les sous-constellations de telle sorte que l'accroissement de la distance minimale entre les points des sous-constellations soit respecté d'un partitionnement à l'autre.

Dans le treillis à quatre états, les transitions parallèles sont séparées de la distance euclidienne  $2\sqrt{\epsilon}$ , qui correspond aussi à  $D_{\text{libre}}$ . Ainsi, le gain de codage de 3 dB est limité par la distance des transitions parallèles.



**Fig. 7(c): treillis à quatre états**



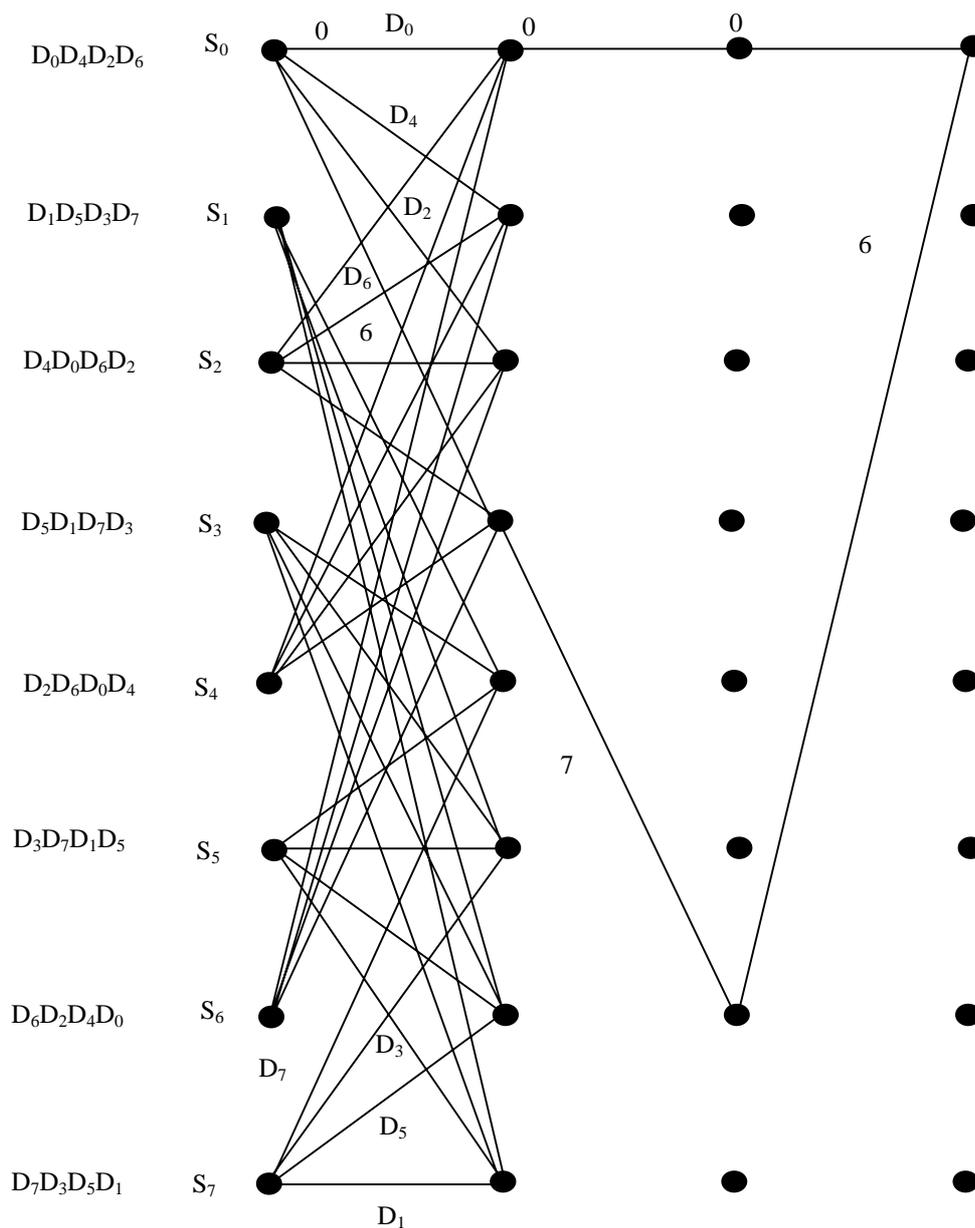
**Fig. 8 : un autre treillis à quatre états possible**

Des gains de performance plus substantiels par rapport à la MDP4 non-codée peuvent être réalisés en utilisant des treillis avec un nombre d'états plus élevés qui permettent d'éliminer les branches parallèles. Des codes avec des treillis de 8 états et plus permettent d'obtenir des transitions distinctes qui conduisent à une augmentation de  $D_{libre}$ .

A titre d'exemple, nous montrons sur la figure 9 le cas d'un treillis à huit états proposé par Ungerboeck en 1982 pour une constellation de type MPD8. Les transitions d'état à état pour maximiser la distance euclidienne libre ont été déterminées à partir des trois règles de base données précédemment. Dans ce cas, le calcul de  $D_{libre}$  conduit au résultat suivant :

$$D_{libre}^2 = d_0^2 + 2.d_1^2 = 4.585.\varepsilon$$

Ceci donne, par rapport à  $d_0^2 = 2.\varepsilon$  dans le cas de la MDP4 non-codée, un gain de 3.6 dB. Ungerboeck (1982, 1987) a également proposé des treillis de rendement 2/3 avec 16, 32, 64, 128 et 256 états qui réalisent des gains de codage allant de 4 à 5.75 dB pour une modulation MDP8.



**Fig. 9: treillis à huit états pour une modulation codée en MDP8**

Les principes de base du partitionnement peuvent s'étendre aux cas de constellations MDPN avec  $N > 8$  ou MAQ- $M$ ,  $M > 8$ , qui possèdent une meilleure efficacité spectrale. Par exemple, 3 (bits/s)/Hz peuvent être obtenus soit par une modulation MDP8 non-codée soit avec une modulation MAQ16. Ungerboeck (1987) a proposé des modulations codées et a évalué les gains de codage obtenus à partir de codes convolutifs de rendements  $1/2$  et  $2/3$  pour la modulation MAQ-16.

Le décodage de Viterbi en soft pour les modulations codées en treillis est réalisé en deux étapes. Puisque chaque branche dans le treillis correspond à un partitionnement de la constellation, la première étape du décodage consiste à déterminer le meilleur point à l'intérieur de chaque sous-constellation, c'est-à-dire le point dans chaque sous-constellation qui est le plus près du signal reçu. Cette étape s'appelle "décodage de la sous-constellation".

Dans la seconde étape, chaque signal sélectionné dans les sous-constellations lors de la première étape et sa métrique euclidienne sont utilisés pour dérouler l'algorithme de Viterbi qui détermine le chemin dans le treillis ayant la métrique cumulée la plus proche de la suite d'échantillons bruités en sortie des filtres adaptés.

## ***Annexe 2 : Association turbo-modulation codée en treillis et code spatio-temporel en bloc (STBC)***

Dans cette annexe nous présentons les performances du schéma de turbo-modulation codée en treillis du paragraphe 2.1, schéma de Robertson et Woerz, associé à un codage spatio-temporel en bloc à deux antennes d'émission de type Alamouti (cf chapitre 3). Le nombre d'antennes de réception est égal à un puis à deux. Un algorithme d'apprentissage de la réponse impulsionnelle du canal basé sur l'insertion de symboles pilotes au sein même du processus turbo est également inclus. Nous montrons comment pondérer les sorties du décodeur STBC avant de les envoyer vers le décodeur log-MAP. Cette pondération des symboles codés s'opère en calculant la distance euclidienne entre les sorties du décodeur STBC (qui est en fait un combineur à maximum de rapport signal à bruit) et tous les points de la constellation MDP-8 utilisée. Nous mettons bien en évidence les gains de capacité considérables obtenus avec un système multi-antennes (en particulier dans le cas de deux antennes de réception). Le papier joint ci-dessous a fait l'objet d'une communication à la conférence ICASSP 2004 à Orlando en Floride.



# **Turbo-TCM and Transmit Antenna Diversity** **In Multipath Fading Channels**

Cances J.P, V. Meghdadi V.N.

GESTE(Groupe d'Etude des Systèmes de Télécommunication)  
ENSIL, University of LIMOGES

**Abstract :** *We consider turbo-trellis-coded transmission over fading MIMO channels with transmit diversity using space-time block codes. We describe the concatenation of Turbo-TCM with a space-time block code and show that, in addition to the transmit diversity, substantial benefits can be obtained by turbo iterations as long as the channel is time-varying during transmission of a coded block or frequency hopping is applied.*

## **I) Introduction**

In future wireless communications systems high data rates need to be reliably transmitted over time-varying bandlimited channels. The wireless channel suffers mainly from time-varying fading due to multipath propagation and destructive superposition of signals received over different paths. Therefore, special techniques have to be applied to enable bandwidth efficient transmission. A very famous one consists in the use of antenna diversity. Usually multiple antennas are used at the receiver using maximum ratio combining (MRC). However, it is hard to efficiently use receive antenna diversity at the mobile receiver since it should remain inexpensive and small. Therefore, receive antenna diversity is only planned at the base station.

However, diversity transmission schemes have been recently introduced to benefit from antenna diversity also in the downlink. A transmit diversity scheme for two transmit antennas and a simple decoding algorithm was introduced by Alamouti in [1] and generalized to an arbitrary number of transmit antennas as space-time block coding by Tarokh & al in [2]. In fact, a space-time block code is not an error event corrector and detector code, it's just a transmit diversity scheme which achieves the maximum possible diversity order of  $n_T \cdot n_R$  for  $n_T$  transmit and  $n_R$  receive antennas under the constraint of allowing simple combining. A space time block code has to be concatenated with an outer code to obtain further coding gain. In recent years "Turbo" codes have been shown to perform close to the Shannon limit in AWGN channels. There have been a number of different proposed schemes combining Turbo codes and higher order modulation [3], [4], [5] and [6]. One of the most outstanding structure is due to Robertson and Worz who proposed a sophisticated scheme using Ungerboeck codes in their recursive systematic form as component codes for Turbo-TCM.

We will first describe the system we propose putting the emphasis on space-time block coding. We show how space-time block coding transforms the fading MIMO channel towards a gaussian SISO channel. We derive a general linear combining rule for space-time block codes. Section 4 explains how we use Turbo-TCM as an outer code and simulation results are eventually provided in section 5.

## **II) Channel model**

Let us consider a discrete-time flat fading MIMO channel model with  $n_T$  transmit antennas and  $n_R$  receive antennas. The tap gain from transmit antenna  $i$  to receive antenna  $j$  is denoted as  $h_k^{(ij)}$ .

These channel taps are considered as independent complex gaussian random variables of equal variance and they satisfy  $E\left\{\left|h_k^{(ij)}\right|^2\right\}=1$ , that means the channel is passive. The observed value at receive antenna  $j$  is given by

$$y_k^{(j)} = \sum_{i=1}^{n_T} h_k^{(ij)} \cdot x_k^{(i)} + n_k^{(j)} \quad (1)$$

In this paper we consider two versions of time variance. The first is the case of quasi-static fading where the tap gains are assumed to be constant over a block and change independently from block to block. This describes a system which applies frequency hopping. The second is the classical Doppler fading, in which the tap gains vary according to a Jakes spectrum. Jakes spectrum is characterized by the product  $f_d \cdot T_s$  of maximal Doppler shift and symbol duration. The additive noise  $n_k^{(j)}$  at each receive antenna  $j$  is assumed to be white and gaussian with spectral power density  $N_0$  in the equivalent lowpass domain.

### III) Space-Time Block Codes

A space-time block code is defined by generalized complex orthogonal design, which is a  $P \times n_T$  matrix  $B_{n_T}$  with orthogonal columns. The entries of  $B_{n_T}$  are elements  $x_k$  of a higher order modulation constellation (M-PSK for example), their conjugates  $x_k^*$  and linear combinations of  $x_k$  and  $x_k^*$ .

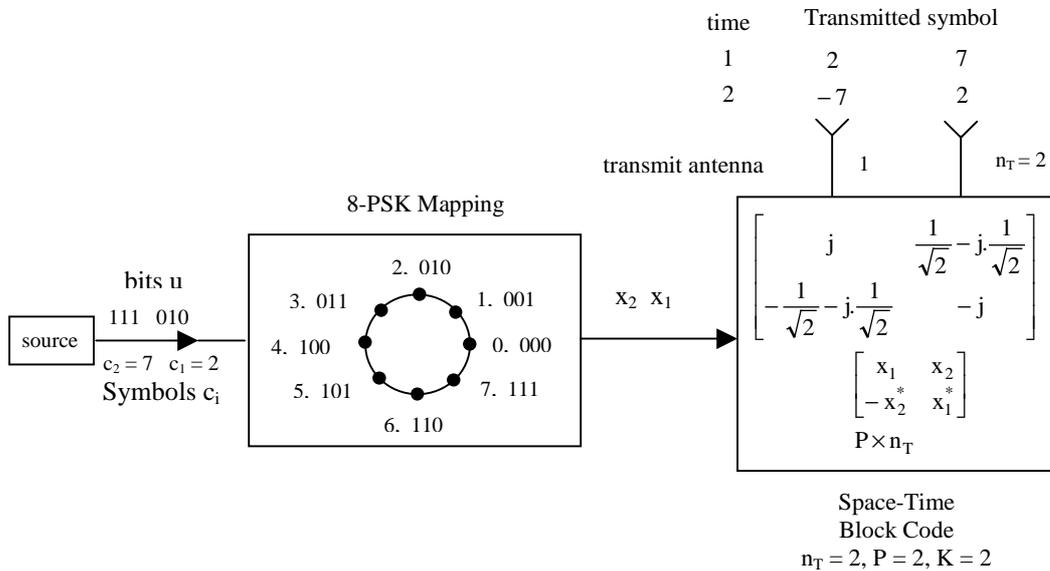


Figure 1 : Space-Time block code transmitter

Figure 1 shows an example for the space-time block code

$$B_2 = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix}$$

which is the scheme proposed by Alamouti in [1]. Inputs to a space-time block encoder is a block of  $K$  complex symbols  $x_k$ ,  $k = 1, \dots, K$ . The space-time block encoder maps the input symbols on entries of  $B_{n_T}$ . All symbols in the same row of  $B_{n_T}$  are transmitted simultaneously from  $n_T$  different antennas. All entries in the same column of  $B_{n_T}$  are transmitted from the same antenna in  $P$  successive time slots. Given the number  $n_T$  of transmit antennas the rate  $R = K/P$  should be maximized. Clearly, each symbol  $x_k$  has to be transmitted with the same energy from all antennas in order to obtain maximum diversity effect. Therefore, the maximum possible rate for space-time codes is  $R = 1$ . For example,  $B_2$  is a rate 1 space-time block code. For  $n_T = 3$  and  $n_T = 4$  the highest rate complex generalized orthogonal designs given in [2] are of rate  $R = \frac{3}{4}$ .

Antenna diversity will be supposed to be used at the receiver premise too using  $n_R$  receive antennas. We assume only the case of flat fading where the tap gains are constant during the transmission of one space-time block code matrix, i.e for  $P.T_S$ . At each receive antenna the observed signal is the superposition of the signals transmitted from the  $n_T$  antennas perturbed by additive white Gaussian noise. However, the orthogonality of the columns of  $B_{n_T}$  allows us to obtain decoupled expressions for the transmitted symbols  $x_k$ , which are independent of the simultaneously transmitted symbols  $x_t$ ,  $t \neq k$ , by simple linear combining. In general, the linear combining for all space-time block codes is done according to

$$\hat{y} = [I_{K \times K} \ 0_{K \times K}] H^H \cdot y + ([0_{K \times K} \ I_{K \times K}] H^H \cdot y)^* = c \cdot \sum_{i=1}^{n_T} \sum_{j=1}^{n_R} |h_k^{(ij)}|^2 \cdot I_{K \times K} x + \hat{n} \quad (1)$$

where  $I_{K \times K}$  is the  $K \times K$  unity matrix,  $y = [y_1^{(1)}, \dots, y_P^{(1)}, \dots, y_1^{(n_R)}, \dots, y_P^{(n_R)}]^T$  and  $x = [x_1, \dots, x_K]^T$  are the received and transmitted symbols, respectively. The received values after combining are denoted  $\hat{y} = [\hat{y}_1, \dots, \hat{y}_K]^T$ . Finally, the  $n_R \cdot P \times 2 \cdot K$  matrix  $H$  is defined such that

$$y = H \begin{bmatrix} x \\ x^* \end{bmatrix} + n \quad (2)$$

The variance of the additive white gaussian noise  $\hat{n} = [\hat{n}_1, \dots, \hat{n}_K]^T$  is  $\sigma_{\hat{n}}^2 = c \cdot \sigma^2 \cdot \sum_{i=1}^{n_T} \sum_{j=1}^{n_R} |h_k^{(ij)}|^2$ ,

where  $\sigma^2$  is the variance of the noise at each receive antenna and  $c$  is a constant, which is  $c = 1$  for the space-time block code  $B_2$ .

Using (1) and the fact that for all space-time block codes the energy gain after combining is equal to the rate loss we can describe the system including the space-time block code, the MIMO fading channel and the linear combiner by a SISO channel as shown in figure 2.

Figure 3 shows for quasi-static fading the probability density function of the per bit SNR normalized to its expected value of the resulting channel of figure 2. It can be seen that the variance of the SNR decreases with the diversity level  $n_T \cdot n_R$  and the fading channel is transformed towards a gaussian channel as it is well known from maximum ratio combining. Due to the chosen normalization, the plots in figure 3 describe any diversity scheme with diversity level  $n_T \cdot n_R$ .

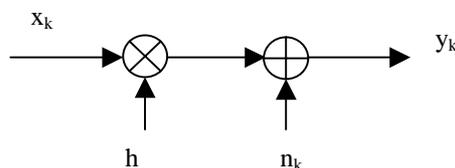


Figure 2 : Equivalent channel model for space-time block code and linear combiner at the receiver

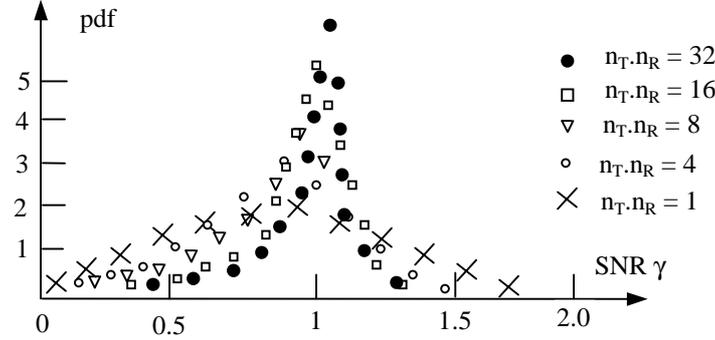


Figure 3 : Probability density function of per bit SNR  $\gamma = \frac{\sum_{i=1}^{n_T} \sum_{j=1}^{n_R} |h_k^{(i,j)}|^2}{n_T \cdot n_R}$  after combining normalized to its expected value

Using SNR distribution we can also easily give analytical results on the bit error probability. Using results in [7] we can derive the bit error probability for BPSK and QPSK with Gray mapping transmitted over quasi-static MIMO fading channels using space-time block codes to be :

$$P_b = \frac{1}{2} \left[ 1 - \mu_b \cdot \sum_{k=0}^{n_T \cdot n_R - 1} C_{2,k}^k \left( \frac{1 - \mu_b^2}{4} \right)^k \right] \quad (3)$$

where  $\mu_b = \frac{1}{\sqrt{1 + \frac{n_T \cdot N_0}{E_b}}}$ . For higher order modulation there exists no closed form solution.

From (1) it is now possible to derive a MAP soft-output detection rule for space-time block codes. This rule is a key point since the output of the space-time block detector is passed to an outer detector. The logarithmic a-posteriori probability can be shown to be

$$\ln P(x_k | \hat{y}_k) = K - \frac{|\hat{y}_k - x_k|^2}{2\sigma^2} - \frac{1}{2\sigma^2} \cdot |x_k|^2 \cdot \left( -1 + c \cdot \sum_{i=1}^{n_T} \sum_{j=1}^{n_R} |h_k^{ij}|^2 \right) + \ln P_a(x_k) \quad (4)$$

The term  $\ln P_a(x_k)$  is an independent a-priori probability obtained from knowledge of the source statistics. If no a-priori information is available, all symbols  $x_k$  are assumed to be equally likely. In the particular case of M-PSK constellations the fourth term in (4) is the same for all possible symbols  $x_k$  and therefore it can be added to the constant term K. In this case the MAP detection rule for equally likely symbols consists in finding the point of the signal constellation with the minimum Euclidean distance to the output of the linear combiner.

#### IV) Turbo-trellis coded modulation and Space-Time Block Codes

From figure 2 it's easy to understand that codes which are optimal for SISO channels are also optimal for MIMO channels with space-time block codes. Trellis-coded modulation schemes appear as an attractive way to interface with Space-Time block codes because the input symbols of a space-time block code are elements of a higher order modulation signal constellation. In this paper we use the Space-Time block code together with a very performant Turbo-TCM scheme as an outer code. This scheme was at first proposed by Robertson and Woz in [6]. It uses Ungerboeck codes in their recursive systematic form as component codes.

Figure 4 shows the encoder, where we used an 8-state 8-PSK Ungerboeck code as component codes. This Ungerboeck code has been shown to be an optimal TCM code for both Gaussian and fading channels. The encoding process is detailed in [6]. The coded 8-PSK symbols are interleaved symbolwise and transmitted over the MIMO channel using a space-time block code.

Infobits pairs

$$(c_1, c_2, c_3, c_4, c_5, c_6) = (00, 01, 11, 10, 00, 11)$$

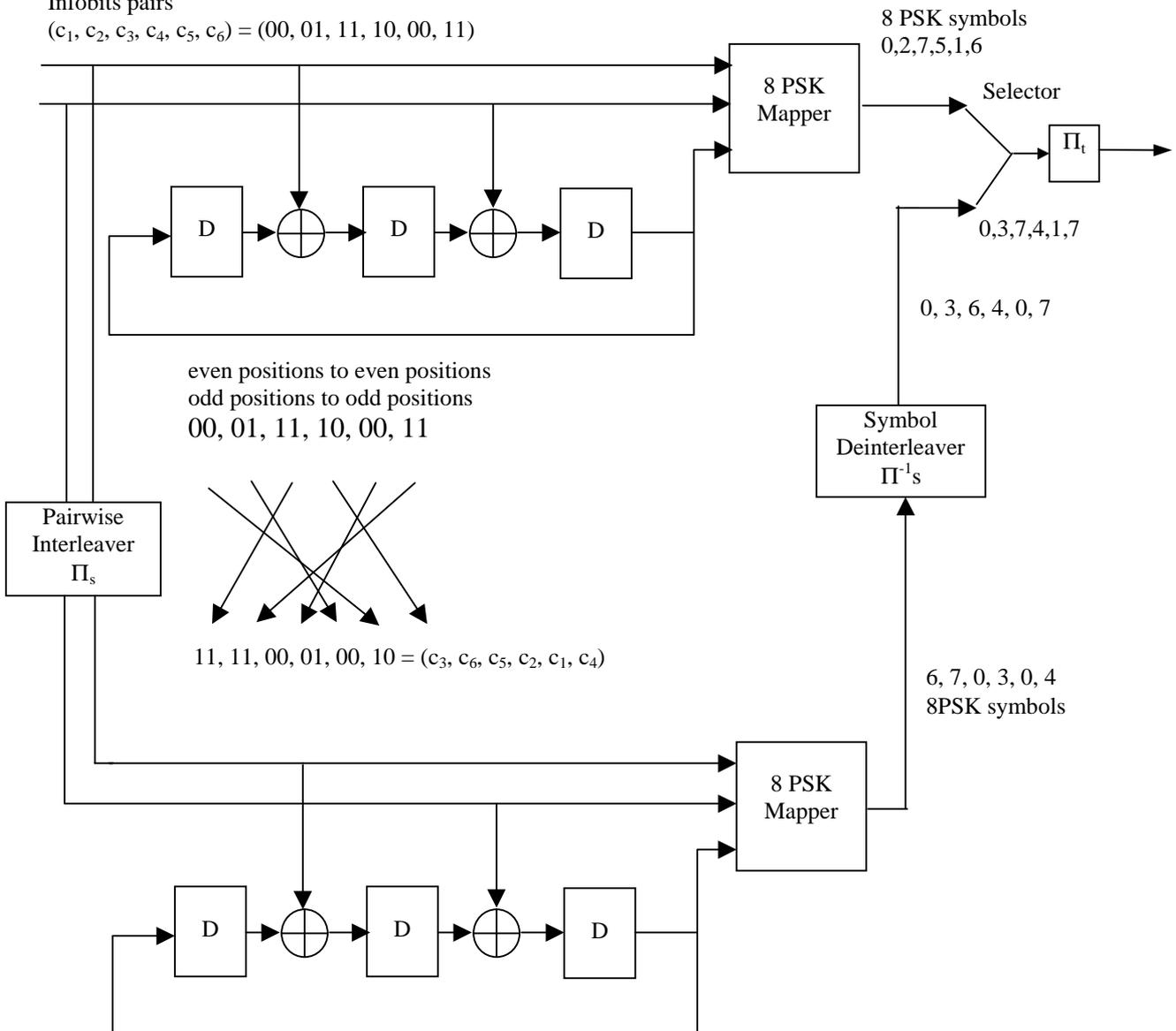


Figure 4 : Turbo Trellis Coded Modulation scheme



are computed according to (1) and (4). These soft values are the channel information for the TTCM decoder. The turbo-TCM decoding is done as described in [6].

It seems possible to think also of a Turbo feedback from the Turbo-TCM decoder to the space-time block code detector. The feedback extrinsic information would then be used as the apriori input  $P_a(x_k)$ . However, this a priori information is only an additive term in the information about symbol  $x_k$  and does not affect the information of other symbols. Therefore, there is no new independent extrinsic information and the feedback will not improve the decoding result. In fact, it's clear that the space-time block code is not a code but a diversity scheme.

The decoding complexity of the concatenated system is mainly determined by the decoding complexity of the outer code. The space-time block code provides diversity with virtually no additional complexity.

Of course, it is also possible to design special codes for turbo space-time coded modulation. The advantage of the method proposed here is besides its good performance the low detection complexity. Furthermore, standard channel coding can be used for the outer code. This makes the system very flexible, because the transmit diversity can just be switched of if only one transmit antenna is available without modifying the channel coding. This is very useful for the application of transmit diversity in existing mobile communications systems.

## **V) Simulation Results**

Simulations were performed for a frame size of 4096 information bits and transmission of 2 bit/s/Hz using 8-PSK. In the turbo-TCM scheme 8-state Ungerboeck codes were used as component codes. All interleavers were pseudo-random [8] of size  $64 \times 64$  (for  $\Pi_c$  and  $\Pi_s$ ). We used only two transmit antennas applying the space-time block code  $B_2$ . The results of BER vs  $E_b/N_0$  is illustrated on figure 6 in the case of the quasi-static fading channel. It's worth noting that there is no significant improvement in BER due to turbo iterations when there is no block space-time code. This can be explained by the fact that only a small fraction of blocks is transmitted in the SNR regions where turbo iterations are very effective. Increasing the size of the interleavers ( $\Pi_s$ ) up to  $128 \times 128$  brings only marginal improvements. However, when the space-time block code is combined with the Turbo TCM modulation the gain compared to the Turbo TCM scheme is obvious. The diversity advantage of this scheme leads to improvements up to 6 dB (TTCM + BSTC : iteration 2) compared to TTCM scheme. We add the case of TCM+BSTC where we replace the TTCM by a classical TCM which uses the same Ungerboeck code as in TTCM as outer code. It's worth noting that TCM + BSTC offers nearly the same performances as TTCM + BSTC. The difference is less than 1 dB.

Figure 7 illustrates the case of fast flat Rayleigh fading (Jakes spectrum with  $f_d.T_s = 0.01$ ). As the channel changes within one frame duration, an additional diversity advantage is obtained by the interleaver between the component codes of the Turbo TCM encoder and by the interleaver between the Turbo TCM encoder and the space time block code. There is a significant additional gain when  $n_R = 2$  receive antennas are used. We add the case of TCM + BSTC too and we found that the curve of this scheme is less step than for Turbo-TCM because of the additional diversity achieved by the interleaver between the component codes of the Turbo TCM.

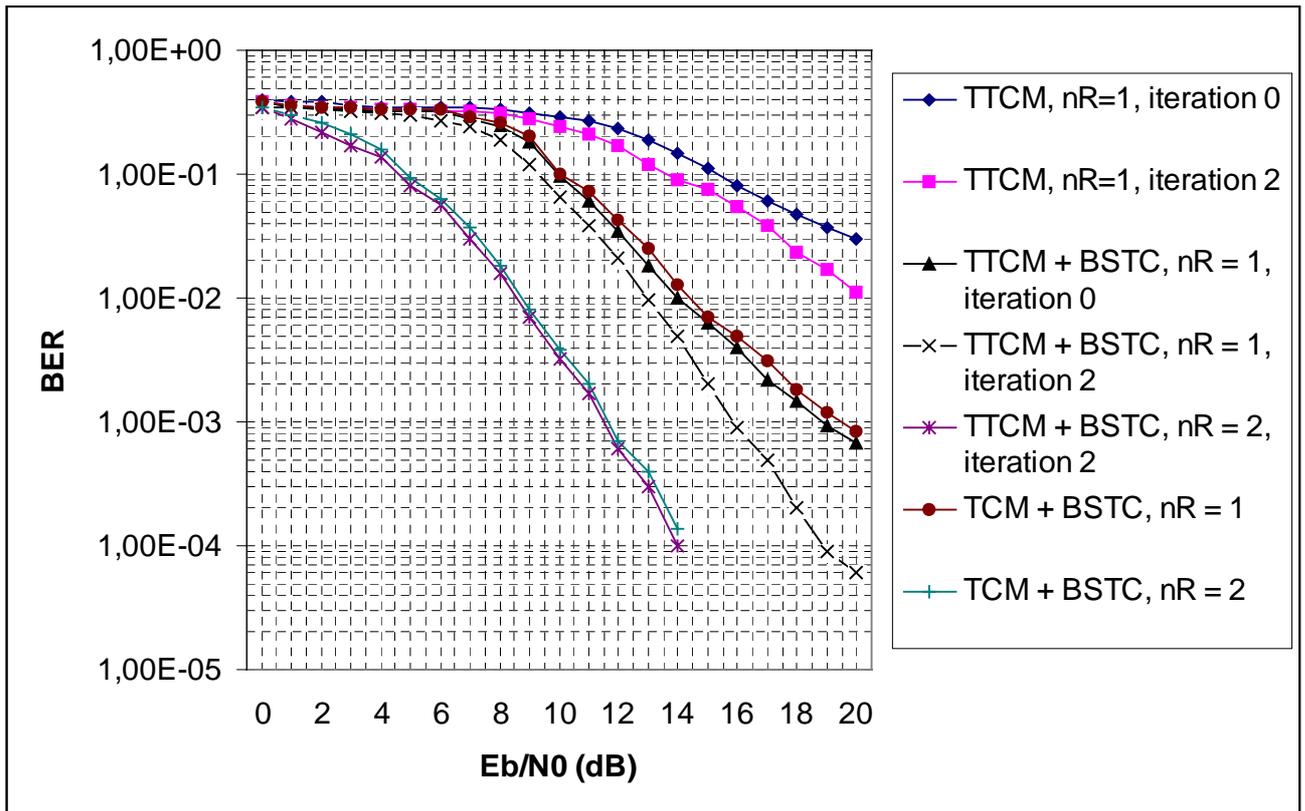


Figure 6 : BER for quasi-static fading channel

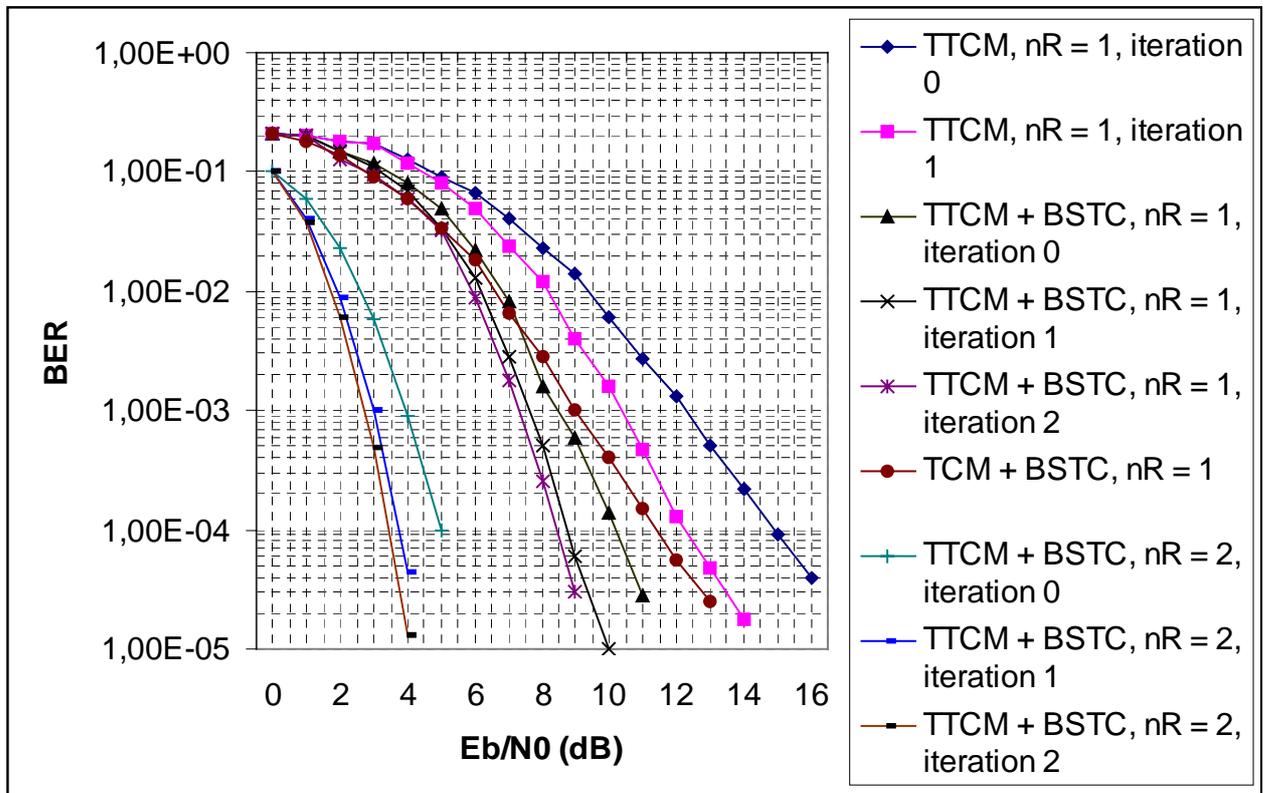


Figure 7 : BER for fast Rayleigh fading channel ( $f_d.T_s = 0.01$ )

## **VI) Conclusion**

We considered in this paper the case of a block space-time code concatenated with a Turbo TCM scheme. We showed that basically the space-time block code together with linear combining at the receiver provides a transformation of the fading MIMO channel to an equivalent SISO channel with a significantly smaller probability of deep fades. Using the equivalent channel model we are able to give analytical results for the BER. Besides, we described the concatenation of turbo-TCM with a space-time block code and showed the outstanding improvement in BER performances due to the additional diversity provided by the block-space time code when we compare this scheme with a classical single sensor receiver.

## **REFERENCES**

- [1] S. Alamouti, "A simple transmitter diversity technique for wireless communications", IEEE Journal on Selected Areas in Communications, Special Issue on Signal Processing for Wireless Communications, 1998.
- [2] V. Tarokh, H. Jafarkhani, A. Calderbank, "Space-time block codes from orthogonal designs", IEEE Transactions on Information Theory, June 99.
- [3] S. Le Goff, A. Glavieux, C. Berrou, "Turbo-codes and high spectral efficiency modulation", ICC 94, pp. 645-649, vol. 2.
- [4] S. Benedetto, D. Divsalar, G. Montorsi, F. Pollara, "Parallel concatenated trellis coded modulation", ICC 96, Dallas.
- [5] Y. Kofman, E. Zehavi, S. Shamai, "Performance analysis of a multilevel coded modulation system, IEEE Trans. on Commun., vol.42, pp. 299-312, Feb-Mar-Apr. 94.
- [6] P. Robertson, T. Worz, "A novel bandwidth efficient coding scheme employing turbo codes", ICC 96, pp. 962-967.
- [7] J. G. Proakis, "Digital Communications", New York : MCGraw Hill, third edition 95.
- [8] C. Le Dantec, P. Piret, "Algebraic and combinatorial methods producing good interleavers", 2<sup>nd</sup> International Symposium on Turbo Codes & Related Topics, pp. 271-273.



# **CHAPITRE 2**

## **CAPACITE DES CANAUX MIMO**



## 2- Capacité des Canaux MIMO

Dans ce chapitre on étudie l'impact, en termes de capacité, de l'utilisation d'antennes multiples en émission et en réception. On se placera ici dans le contexte d'un système mono-utilisateur. Le but est de trouver des expressions mathématiques qui permettent d'évaluer la capacité de canaux MIMO pour différents modèles de propagation. Nous allons montrer que le gain en capacité apporté par l'utilisation d'antennes multiples est considérable à condition de considérer des fadings et des bruits indépendants sur les antennes de réception.

### 2-1 Introduction

Il est clair que l'utilisation d'antennes multiples est un moyen bien connu pour améliorer les performances d'un système de transmission sur canaux à fading [1]. Ces antennes multiples peuvent s'utiliser à l'émetteur ainsi qu'au récepteur. Cependant, dans un système radio mobile, vu la taille des portables, il apparaît clairement que l'utilisation d'antennes multiples se fera plutôt au niveau de la station de base, sachant qu'au maximum deux antennes pourront être implantées sur le mobile. Dans le cas général on parlera de diversité de transmission lorsqu'on utilise des antennes multiples en émission et de diversité de réception lorsqu'on utilise des antennes multiples en réception. Pour le cas particulier du mobile, la diversité se fait plutôt en liaison montante en réception à la station de base et, en liaison descendante, toujours en émission à la station de base [2].

Récemment les systèmes MIMO (Multi-Input Multi-Output) sont devenus un des sujets les plus étudiés en recherches, car ils sont capables d'augmenter l'efficacité spectrale (capacité) sur une largeur de bande limitée. L'utilisation d'antennes multiples conduit à une dimension supplémentaire dans le degré d'accès multiple au réseau par rapport au cas mono-antenne (Space Division Multiple Access: SDMA) et ainsi offre une solution efficace à l'accroissement des débits pour les générations futures de radiotéléphonie cellulaire. Pour conduire nos calculs nous allons d'abord utiliser le modèle du canal MIMO Gaussien qui est défini comme un canal simple utilisateur gaussien avec des antennes multiples en émission et en réception. De plus, le scénario simple utilisateur nous permet de traiter les informations émises et reçues en commun.

Ces calculs sont essentiellement basés sur les travaux de Telatar [3]. Le plan de ce chapitre est alors le suivant. Nous présentons d'abord le modèle d'un canal MIMO Gaussien. Ensuite nous présentons une étude mathématique préliminaire qui nous permet de calculer la capacité de ce type de canal. Le cas du canal de Rayleigh quasi-statique est ensuite abordé. Ce chapitre se conclue par le cas des canaux non ergodiques et des canaux MIMO multiutilisateurs.

#### **Notations utilisées dans ce chapitre :**

Les matrices et les vecteurs en police gros italique :  $\mathbf{A}_{M \times N}$ ,  $\mathbf{x}_{M \times 1}$

Opérateur Hermitien des matrices ou des vecteurs :  $\mathbf{A}_{N \times M}^\dagger, \mathbf{x}_{1 \times M}^\dagger$   
 Opérateur Transposé des matrices ou des vecteurs :  $\mathbf{A}_{N \times M}^T, \mathbf{x}_{1 \times M}^T$   
 Opérateur Conjugué des matrices ou des vecteurs :  $\mathbf{A}_{M \times N}^*, \mathbf{x}_{M \times 1}^*$

## 2-2 Modélisation du Système

Nous considérons un système mono-utilisateur MIMO avec un canal Gaussien (ref. Figure 1). On dénote le nombre d'antenne d'émission par  $n$ , et le nombre d'antennes de réception par  $m$ . On traite exclusivement un modèle linéaire dans lequel le vecteur reçu  $\mathbf{y}_{m \times 1} = [y_1 \cdots y_j \cdots y_m]^T \in \mathbb{C}^m$  se calcule d'après le vecteur émis  $\mathbf{x}_{n \times 1} = [x_1 \cdots x_i \cdots x_n]^T \in \mathbb{C}^n$  avec l'équation suivante.

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (1)$$

Où  $\mathbf{H}$  est une matrice  $m \times n$  contenant les  $\alpha_{i,j}$  ( gain du trajet entre l'antenne d'émission  $i$  et l'antenne de réception  $j$ ) et  $\mathbf{n} = [n_1 \cdots n_j \cdots n_m]^T$  est le bruit gaussien complexe additif avec la moyenne zéro, et avec les parties réelle et imaginaire indépendantes de la même variance. On suppose également que les bruits des différents antennes de réception sont indépendants et normalisés, donc  $E\{\mathbf{nn}^\dagger\} = \mathbf{I}_{m \times m}$ . Le fait que la puissance de l'émetteur soit égale à  $P$  conduit à la contrainte :

$$\begin{aligned} E\{\mathbf{x}^\dagger \mathbf{x}\} &\leq P \\ \text{tr}(E\{\mathbf{xx}^\dagger\}) &\leq P \end{aligned} \quad (2)$$

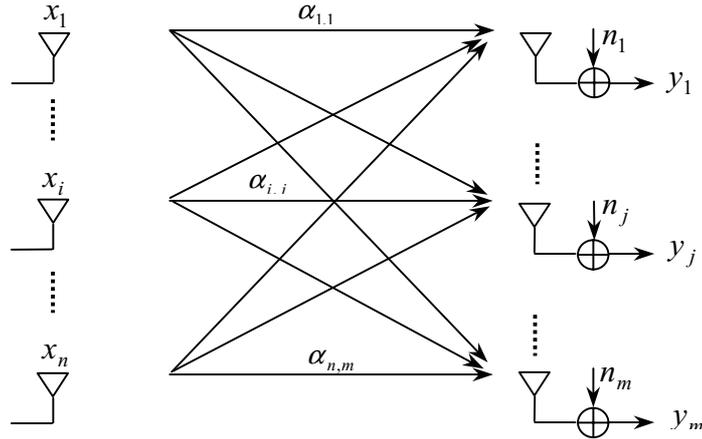
La deuxième inégalité vient du fait que:

$$\text{tr}(\mathbf{A}_{M \times N} \mathbf{B}_{N \times M}) = \text{tr}(\mathbf{B}_{N \times M} \mathbf{A}_{M \times N}) \Rightarrow \mathbf{X}_{1 \times N}^\dagger \mathbf{Y}_{N \times 1} = \text{tr}(\mathbf{Y}_{N \times 1} \mathbf{X}_{1 \times N}^\dagger) \Rightarrow \mathbf{x}^\dagger \mathbf{x} = \text{tr}(\mathbf{xx}^\dagger) \quad (3)$$

La deuxième forme du contraint d'énergie dans (2) sera plus utile dans les discussions à venir.

Jusqu'à maintenant, on n'a pas parlé de la nature de la matrice  $\mathbf{H}$ . On considère trois scénarios différents pour cette matrice.

1.  $\mathbf{H}$  est déterministe.
2.  $\mathbf{H}$  est une matrice aléatoire (pour laquelle on utilisera la notation  $\mathbf{H}$ ), choisie selon une densité de probabilité, et chaque utilisation du canal correspond à une réalisation indépendante de  $\mathbf{H}$ .
3.  $\mathbf{H}$  est une matrice aléatoire, mais fixée une fois choisie.



**Figure 1 : Modèle d'un canal Gaussien multi-antennes**

Dans ce chapitre on étudiera principalement les deux derniers cas. Au cas où la matrice  $\mathbf{H}$  serait aléatoire, on suppose que ses éléments sont des variables aléatoires i.i.d. (Independent Identical Distribution) Gaussien avec la moyenne zéro, et les parties réelle et imaginaire indépendantes, chacune avec la variance  $1/2$ . Par conséquent, chaque élément de la matrice  $\mathbf{H}$  a une phase uniforme ( $[0, 2\pi]$ ) et une loi qui suit une amplitude Rayleigh. Ce choix correspond bien au cas où les antennes d'émission et de réception sont suffisamment espacées pour pouvoir considérer leurs coefficients  $\alpha_{i,j}$  indépendants. Dans tous les cas, on suppose que le récepteur connaît les informations du canal ( $\mathbf{H}$ ), autrement dit, les sorties du canal sont données par la paire  $(\mathbf{y}, \mathbf{H})$  et la distribution de  $\mathbf{H}$  est connue par le récepteur.

### 2-3 Préliminaires mathématiques

Le vecteur aléatoire complexe  $\mathbf{x}_{n \times 1} \in \mathbb{C}^n$  est Gaussien si sa partie réelle et imaginaire sont Gaussiennes. Pour la suite et afin d'éviter les calculs complexes, on définit le vecteur aléatoire réel  $\hat{\mathbf{x}}_{2n \times 1} \in \mathbb{R}^{2n}$  contenant la partie réelle et imaginaire du vecteur  $\mathbf{x}$ ,  $\hat{\mathbf{x}} = \begin{bmatrix} \text{Re}(\mathbf{x}) \\ \text{Im}(\mathbf{x}) \end{bmatrix}$  qui doit être Gaussien aussi. Donc pour désigner la distribution du vecteur aléatoire complexe Gaussien  $\mathbf{x}$ , il suffit juste de trouver les statistiques (l'espérance et la covariance) du vecteur aléatoire réel  $\hat{\mathbf{x}}$ , c'est à dire :

$$E\{\hat{\mathbf{x}}\} \in \mathbb{R}^{2n} \quad , \quad E\{(\hat{\mathbf{x}} - E\{\hat{\mathbf{x}}\})(\hat{\mathbf{x}} - E\{\hat{\mathbf{x}}\})^T\} \in \mathbb{R}^{2n \times 2n}$$

On dit que le vecteur aléatoire complexe gaussien  $\mathbf{x}$  est *circulairement symétrique* (CS) si la matrice d'auto-covariance du vecteur correspondant  $\hat{\mathbf{x}}$ , a la structure suivante (ref. Annexe A).

$$E\{(\hat{\mathbf{x}} - E\{\hat{\mathbf{x}}\})(\hat{\mathbf{x}} - E\{\hat{\mathbf{x}}\})^T\} = \frac{1}{2} \begin{bmatrix} \text{Re}(\mathbf{Q}) & -\text{Im}(\mathbf{Q}) \\ \text{Im}(\mathbf{Q}) & \text{Re}(\mathbf{Q}) \end{bmatrix} = \frac{1}{2} \hat{\mathbf{Q}} \quad (4)$$

où  $\mathbf{Q} \in C^{n \times n}$  est une matrice Hermitienne non-négative définie ;  $\mathbf{Q}$ , plus précisément, désigne la matrice d'auto-covariance du vecteur  $\mathbf{x}$ ,  $\mathbf{Q} = E\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^\dagger\}$  (ref. Annexe A). Donc on peut trouver les statistiques du vecteur  $\mathbf{x}$  à partir de celle du vecteur  $\hat{\mathbf{x}}$ . Il faut aussi remarquer que la partie réelle d'une matrice Hermitienne est symétrique et sa partie imaginaire est antisymétrique et que la matrice  $\hat{\mathbf{Q}}$  est réelle et symétrique.

Comme dans la suite, on utilise plutôt la forme réelle ( $\hat{\mathbf{x}}$ ) d'un vecteur aléatoire complexe ( $\mathbf{x}$ ), il faut expliciter les propriétés de cette projection qui s'applique sur le vecteur complexe ainsi que sur la matrice complexe. Pour chaque  $z_{n \times 1} \in C^n$  et  $A_{n \times n} \in C^{n \times n}$  on définit le vecteur  $\hat{z}_{2n \times 1} \in R^{2n}$  et la matrice  $\hat{A}_{2n \times 2n} \in R^{2n \times 2n}$  comme suivant :

$$\hat{z} = \begin{bmatrix} \text{Re}(z) \\ \text{Im}(z) \end{bmatrix} \quad ; \quad \hat{A} = \begin{bmatrix} \text{Re}(A) & -\text{Im}(A) \\ \text{Im}(A) & \text{Re}(A) \end{bmatrix}$$

**Lemme 1.** La projection T, définie comme  $z \rightarrow \hat{z} = \begin{bmatrix} \text{Re}(z) \\ \text{Im}(z) \end{bmatrix}$  et  $A \rightarrow \hat{A} = \begin{bmatrix} \text{Re}(A) & -\text{Im}(A) \\ \text{Im}(A) & \text{Re}(A) \end{bmatrix}$  a

les propriétés suivantes :

$$\mathbf{C} = \mathbf{AB} \Leftrightarrow \hat{\mathbf{C}} = \hat{\mathbf{A}}\hat{\mathbf{B}} \quad (5-a)$$

$$\mathbf{C} = \mathbf{A} + \mathbf{B} \Leftrightarrow \hat{\mathbf{C}} = \hat{\mathbf{A}} + \hat{\mathbf{B}} \quad (5-b)$$

$$\mathbf{C} = \mathbf{A}^\dagger \Leftrightarrow \hat{\mathbf{C}} = \hat{\mathbf{A}}^T \quad (5-c)$$

$$\mathbf{C} = \mathbf{A}^{-1} \Leftrightarrow \hat{\mathbf{C}} = \hat{\mathbf{A}}^{-1} \quad (5-d)$$

$$\det(\hat{\mathbf{A}}) = \det(\mathbf{AA}^\dagger) = |\det(\mathbf{A})|^2 \quad (5-e)$$

$$\mathbf{z} = \mathbf{x} + \mathbf{y} \Leftrightarrow \hat{\mathbf{z}} = \hat{\mathbf{x}} + \hat{\mathbf{y}} \quad (5-f)$$

$$\mathbf{y} = \mathbf{Ax} \Leftrightarrow \hat{\mathbf{y}} = \hat{\mathbf{A}}\hat{\mathbf{x}} \quad (5-g)$$

$$\text{Re}(\mathbf{x}^\dagger \mathbf{y}) = \hat{\mathbf{x}}^T \hat{\mathbf{y}} \quad (5-h)$$

Les preuves de ces propriétés se trouvent dans l'Annexe B.

**Corollaire 1.**  $U \in C^{n \times n}$  est unitaire si et seulement si  $\hat{U} \in R^{2n \times 2n}$  est orthonormale.

*Preuve :* il suffit juste de suivre (4-a),  $U^\dagger U = I_n \Leftrightarrow \hat{U}^T \hat{U} = \hat{I}_n = I_{2n}$

**Corollaire 2.** Si  $\mathbf{Q} \in C^{n \times n}$  est non-négative définie,  $\hat{\mathbf{Q}} \in R^{2n \times 2n}$  le sera aussi.

*Preuve :* On considère un vecteur complexe quelconque  $\mathbf{x} \in C^n$  et sa projection  $\hat{\mathbf{x}} \in R^{2n}$ . En utilisant (5-g) et (5-h) et la définition d'une matrice non-négative définie, on aura :

$$\hat{\mathbf{x}}^T \underbrace{\hat{\mathbf{Q}} \hat{\mathbf{x}}}_{\hat{\mathbf{y}}} = \text{Re}(\mathbf{x}^\dagger \mathbf{y}) = \text{Re}(\mathbf{x}^\dagger \mathbf{Q} \mathbf{x}) = \mathbf{x}^\dagger \mathbf{Q} \mathbf{x} \geq 0$$

La densité de probabilité (en ce qui concerne la mesure standard de Lebesgue en  $C^n$ ) d'un vecteur *gaussien complexe circulairement symétrique* (GCCS) avec la moyenne  $\boldsymbol{\mu}$  et la matrice de covariance  $\mathbf{Q}$  est donnée par (voir l'annexe C):

$$\begin{aligned}
f_{\mu, \mathbf{Q}}(\mathbf{x}) &= \det(\pi \hat{\mathbf{Q}})^{-1/2} \exp\left(-(\hat{\mathbf{x}} - \hat{\boldsymbol{\mu}})^\dagger \hat{\mathbf{Q}}^{-1}(\hat{\mathbf{x}} - \hat{\boldsymbol{\mu}})\right) \\
&= \det(\pi \mathbf{Q})^{-1/2} \exp\left(-(\mathbf{x} - \boldsymbol{\mu})^\dagger \mathbf{Q}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)
\end{aligned} \tag{6}$$

où la deuxième égalité se déduit de (5-d)-(5-h) et du corollaire 2. L'entropie différentielle d'un complexe Gaussien  $\mathbf{x}$  avec la matrice de covariance  $\mathbf{Q}$  s'écrit alors sous la forme :

$$\begin{aligned}
H(f_{\mathbf{Q}}) &= E_{f_{\mathbf{Q}}} \{-\log(f_{\mathbf{Q}}(\mathbf{x}))\} \\
&= \log(\det(\pi \mathbf{Q})) + \log(e) E\{\mathbf{x}^\dagger \mathbf{Q}^{-1} \mathbf{x}\} = \log(\det(\pi \mathbf{Q})) + \log(e) E\{\text{tr}(\mathbf{x} \mathbf{x}^\dagger \mathbf{Q}^{-1})\} \\
&= \log(\det(\pi \mathbf{Q})) + \log(e) \text{tr}(E\{\mathbf{x} \mathbf{x}^\dagger\} \mathbf{Q}^{-1}) = \log(\det(\pi \mathbf{Q})) + \log(e) \text{tr}(\mathbf{Q} \cdot \mathbf{Q}^{-1}) \\
&= \log(\det(\pi \mathbf{Q})) + \log(e) \text{tr}(\mathbf{I}_n) = \log(\det(\pi \mathbf{Q})) + n \log(e) = \log(e^n \det(\pi \mathbf{Q})) \\
&= \log(\det(\pi e \mathbf{Q}))
\end{aligned} \tag{7}$$

Dans (7) on s'est servi de (3). L'importance des vecteurs aléatoires GCCS est due au lemme suivant : Les variables GCCS donnent l'entropie maximum.

**Lemme 2 :** Supposons que le vecteur aléatoire complexe  $\mathbf{x} \in C^n$  a la moyenne zéro et la matrice de covariance  $\mathbf{Q}$ ,  $E\{\mathbf{x} \mathbf{x}^\dagger\} = \mathbf{Q}$ , autrement dit,  $E\{x_i x_j^*\} = Q_{i,j}$ ,  $1 \leq i, j \leq n$ . Donc l'entropie de  $\mathbf{x}$  satisfait  $H(\mathbf{x}) \leq \log(\det(\pi \mathbf{Q}))$ , avec l'égalité si et seulement si  $\mathbf{x}$  est GCCS avec la même matrice de covariance  $\mathbf{Q}$ .

*Preuve :* Considérons la fonction de densité de  $p(\mathbf{x})$  qui donne exactement la même matrice de covariance  $\mathbf{Q}$ , c-à-d, elle satisfait  $E\{x_i x_j^*\} = \int_{C^n} p(\mathbf{x}) x_i x_j^* d\mathbf{x} = Q_{i,j}$ ,  $1 \leq i, j \leq n$  et on considère également la densité  $f_{\mathbf{Q}}(\mathbf{x})$  qui correspond au vecteur GCCS, donc

$$f_{\mathbf{Q}}(\mathbf{x}) = \det(\pi \mathbf{Q})^{-1} \exp(-\mathbf{x}^\dagger \mathbf{Q}^{-1} \mathbf{x})$$

Comme  $\mathbf{Q}$  est la matrice de covariance de cette densité GCCS, il est évident que  $\int_{C^n} f_{\mathbf{Q}}(\mathbf{x}) x_i x_j^* d\mathbf{x} = Q_{i,j}$ ,  $1 \leq i, j \leq n$ , d'ailleurs d'après cette densité Gaussienne,  $\log(f_{\mathbf{Q}}(\mathbf{x}))$  est une combinaison linéaire des termes  $x_i x_j^*$  et le moyennage de ces termes avec la densité de  $p(\mathbf{x})$  ou avec la densité de  $f_{\mathbf{Q}}(\mathbf{x})$  sont identiques. Ainsi  $E_{f_{\mathbf{Q}}} \{\log(f_{\mathbf{Q}}(\mathbf{x}))\} = E_p \{\log(f_{\mathbf{Q}}(\mathbf{x}))\}$ . Donc :

$$\begin{aligned}
H(p) - H(f_{\mathbf{Q}}) &= -\int_{C^n} p(\mathbf{x}) \log(p(\mathbf{x})) d\mathbf{x} + \int_{C^n} f_{\mathbf{Q}}(\mathbf{x}) \log(f_{\mathbf{Q}}(\mathbf{x})) d\mathbf{x} \\
&= -\int_{C^n} p(\mathbf{x}) \log(p(\mathbf{x})) d\mathbf{x} + \int_{C^n} p(\mathbf{x}) \log(f_{\mathbf{Q}}(\mathbf{x})) d\mathbf{x} \\
&= \int_{C^n} p(\mathbf{x}) \log(f_{\mathbf{Q}}(\mathbf{x})/p(\mathbf{x})) d\mathbf{x} \\
&\leq 0
\end{aligned}$$

Où la dernière inégalité se déduit de l'inégalité d'information ( $D(p\|q) \geq 0$ ), où  $D(p\|q) = \int p \log(p/q)$  est l'entropie relative de deux fonctions de probabilité  $p(x)$  et  $q(x)$ , et l'égalité est atteinte si et seulement si  $p(x) = q(x)$ ,  $\forall x$ ). Cette dernière est elle-même une

conséquence de l'inégalité de Jensen ( $E_p(g(X)) \geq g(E_p(X))$ ) pour la variable aléatoire  $X$  et une fonction  $g(x)$  convexe, si  $g(x)$  est strictement convexe, l'égalité implique que  $X$  soit constant) [11, §2.6, §9.6]. Donc  $H(p) \leq H(f_\rho)$  avec l'égalité si et seulement si  $p(\mathbf{x}) = f_\rho(\mathbf{x}); \forall \mathbf{x} \in C^n$ .

**Lemme 3 :** Si le vecteur  $\mathbf{x} \in C^n$  est GCCS, le vecteur  $\mathbf{y} = A\mathbf{x} \in C^m$ , le sera également pour n'importe quelle matrice  $A \in C^{m \times n}$ .

*Preuve :* Sans perdre la généralité, on peut supposer que le vecteur  $\mathbf{x}$  a la moyenne zéro et la matrice de covariance  $\mathbf{Q} = E\{\mathbf{x}\mathbf{x}^\dagger\}$ . Ainsi la moyenne de  $\mathbf{y}$  est zéro. Comme chaque élément de  $\mathbf{y}$  est une combinaison linéaire des éléments de  $\mathbf{x}$ , il est évident que  $\mathbf{y}$  est un vecteur Gaussien complexe, il reste à prouver qu'il est CS. D'après (5-g)  $\hat{\mathbf{y}} = \hat{A}\hat{\mathbf{x}}$  et

$$E\{\hat{\mathbf{y}}\hat{\mathbf{y}}^\dagger\} = E\{\hat{A}\hat{\mathbf{x}}\hat{\mathbf{x}}^\dagger\hat{A}^\dagger\} = \hat{A}E\{\hat{\mathbf{x}}\hat{\mathbf{x}}^\dagger\}\hat{A}^\dagger = \frac{1}{2}\hat{A}\hat{\mathbf{Q}}\hat{A}^\dagger = \frac{1}{2}\hat{\mathbf{K}}$$

où  $\mathbf{K} = E\{\mathbf{y}\mathbf{y}^\dagger\} = E\{A\mathbf{x}\mathbf{x}^\dagger A^\dagger\} = AE\{\mathbf{x}\mathbf{x}^\dagger\}A^\dagger = A\mathbf{Q}A^\dagger$ .

**Lemme 4 :** Si  $\mathbf{x}$  et  $\mathbf{y}$  sont des vecteurs GCCS indépendants,  $\mathbf{z} = \mathbf{x} + \mathbf{y}$  est également un vecteur GCCS.

*Preuve :* là encore sans perdre de généralité, on suppose que les vecteur  $\mathbf{x}$  et  $\mathbf{y}$  ont la moyenne zéro,  $\mathbf{A} = E\{\mathbf{x}\mathbf{x}^\dagger\}$  et  $\mathbf{B} = E\{\mathbf{y}\mathbf{y}^\dagger\}$ . Donc, en utilisant (5-f), on obtient :

$$E\{\hat{\mathbf{z}}\hat{\mathbf{z}}^\dagger\} = E\{(\hat{\mathbf{x}} + \hat{\mathbf{y}})(\hat{\mathbf{x}} + \hat{\mathbf{y}})^\dagger\} = E\{\hat{\mathbf{x}}\hat{\mathbf{x}}^\dagger\} + E\{\hat{\mathbf{y}}\hat{\mathbf{y}}^\dagger\} = \frac{1}{2}\hat{\mathbf{A}} + \frac{1}{2}\hat{\mathbf{B}} = \frac{1}{2}\hat{\mathbf{C}}$$

où  $\mathbf{C} = E\{\mathbf{z}\mathbf{z}^\dagger\} = E\{(\mathbf{x} + \mathbf{y})(\mathbf{x} + \mathbf{y})^\dagger\} = E\{\mathbf{x}\mathbf{x}^\dagger\} + E\{\mathbf{y}\mathbf{y}^\dagger\} = \mathbf{A} + \mathbf{B}$

## 2-4 Canal Gaussien avec Fonction de Transfert Fixe

Dans le calcul de la capacité de canal, on va commencer par le cas le plus simple où la matrice de transfert  $\mathbf{H}$  est déterministe. Les résultats de ce paragraphe peuvent être déduits de [4, §8]

### **2-4-1 Capacité**

On va d'abord dériver une expression pour la capacité  $C(\mathbf{H}, P)$  de ce type de canal avec les gains constants. Pour ce faire, on maximise l'information mutuelle moyenne  $I(\mathbf{x}; \mathbf{y})$  entre l'entrée et la sortie de notre canal MIMO, sur le choix de la distribution de  $\mathbf{x}$ .

En utilisant le théorème de décomposition en valeur singulière (SVD), chaque matrice  $\mathbf{H} \in C^{m \times n}$  peut s'écrire comme (ref. Annexe D) :

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^\dagger \quad (8)$$

où  $U \in C^{m \times m}$  et  $V \in C^{n \times n}$  sont unitaires ( $U^{-1} = U^\dagger$  et  $V^{-1} = V^\dagger$ ), et  $D \in R^{m \times n}$  est diagonale et non-négative définie. Plus précisément on a les propriétés suivantes :

$D$  : est une matrice diagonale non-négative définie dont les éléments diagonaux sont les racine carrées des valeurs propres de  $HH^\dagger$  (ou  $H^\dagger H$ ).

$U$  : est une matrice unitaire dont les colonnes sont les vecteurs propres de  $HH^\dagger$ .

$V$  : est une unitaire dont les colonnes sont les vecteurs propres de  $H^\dagger H$ .

Donc en utilisant (8), l'équation (1) peut réécrit comme :

$$\begin{aligned} \mathbf{y} &= \mathbf{U} \mathbf{D} \mathbf{V}^\dagger \mathbf{x} + \mathbf{n} \\ \mathbf{U}^\dagger \mathbf{y} &= \mathbf{D} \mathbf{V}^\dagger \mathbf{x} + \mathbf{U}^\dagger \mathbf{n} \end{aligned}$$

Maintenant on utilise un changement d'espace dans lequel le vecteur entré est multiplié par  $V^\dagger$ , et le vecteur sorti par  $U^\dagger$ , avec les définitions suivantes :

$$\tilde{\mathbf{y}} = \mathbf{U}^\dagger \mathbf{y} \quad , \quad \tilde{\mathbf{x}} = \mathbf{V}^\dagger \mathbf{x} \quad , \quad \tilde{\mathbf{n}} = \mathbf{U}^\dagger \mathbf{n}$$

Etant donné que les matrices  $U$  et  $V$  sont unitaires,  $\tilde{\mathbf{n}}$  a exactement la même distribution de  $\mathbf{n}$  (il est toujours Gaussien et avec les mêmes statistiques) et  $E\{\tilde{\mathbf{x}}^\dagger \tilde{\mathbf{x}}\} = E\{\mathbf{x}^\dagger \mathbf{V} \mathbf{V}^\dagger \mathbf{x}\} = E\{\mathbf{x}^\dagger \mathbf{x}\}$  (la puissance totale de  $\tilde{\mathbf{x}}$  et de  $\mathbf{x}$  sont identiques). Donc, le canal original est équivalent au canal suivant :

$$\tilde{\mathbf{y}} = \mathbf{D} \tilde{\mathbf{x}} + \tilde{\mathbf{n}} \quad (9)$$

Dans ce canal équivalent, les statistiques du bruit n'ont pas changé,  $\tilde{\mathbf{n}}$  est Gaussien, avec la moyenne zéro et la partie réelle et imaginaire indépendantes et identiquement distribuées avec la même matrice de covariance,  $E\{\tilde{\mathbf{n}} \tilde{\mathbf{n}}^\dagger\} = \mathbf{I}_m$  et la contrainte d'énergie reste la même aussi  $E\{\tilde{\mathbf{x}}^\dagger \tilde{\mathbf{x}}\} \leq P$ . Comme les matrices  $U$  et  $V$  sont unitaires (et inversibles) la projection  $(x, y, n) \rightarrow (\tilde{x}, \tilde{y}, \tilde{n})$  conserve les statistiques (entropie, informations mutuelles, capacité, etc.) on peut alors considérer le canal équivalent (9) au lieu du canal (1). Une fois ces statistiques calculées pour le canal équivalent, on peut trouver les statistiques correspondantes dans le canal original.

On revient au canal (9), comme  $\text{rank}(H) \leq \min(m, n)$ , au maximum on aura  $\min(m, n)$  valeurs singulière non-nulles (ref. Annexe D). En notant ces dernières par  $\lambda_i^{1/2}$ ,  $i = 1, \dots, \min(m, n)$  (où  $\lambda_i$  sont les valeurs propres non-zéros de  $HH^\dagger$  ou  $H^\dagger H$ ), on peut réécrire (9) sous une forme par composantes scalaires :

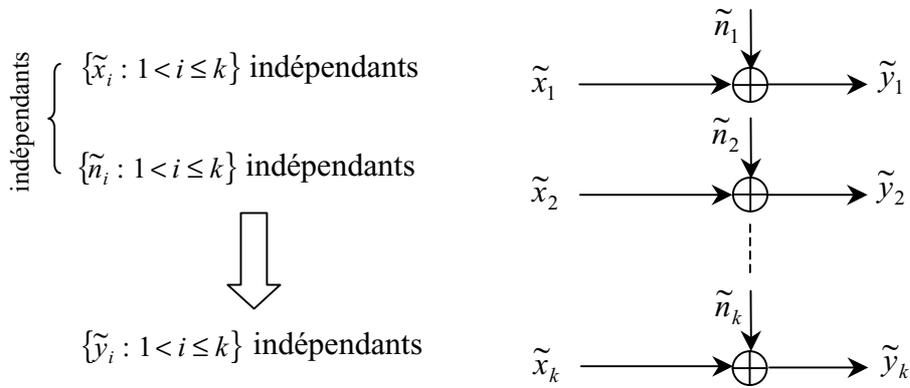
$$\begin{cases} \tilde{y}_i = \lambda_i^{1/2} \tilde{x}_i + \tilde{n}_i & ; \quad 1 \leq i \leq k \\ \tilde{y}_i = \tilde{n}_i & ; \quad k < i \leq m \end{cases} \quad ; \quad k = \min(m, n) \quad (10)$$

Comme on peut remarquer dans (10), si  $m > n$ , il reste des composantes  $\tilde{y}_i$  ( $n < i \leq m$ ) qui sont égaux aux composantes correspondantes du bruit  $\tilde{n}_i$ . Ces composants  $\tilde{y}_i$  pour  $\min(m, n) < i \leq m$  sont indépendants des signaux émis. D'ailleurs si  $n > m$  il y a des

components  $\tilde{x}_i$  ( $m < i \leq n$ ) qui n'interviennent pas dans le canal. Donc pour le calcul de capacité on ne considère que le premier terme de (10). Dans la section suivante on va présenter le calcul détaillé de la capacité de ce canal, mais pour l'instant on accepte que la maximisation de l'information mutuelle nécessite que  $\{\tilde{x}_i : 1 < i \leq \min(m, n)\}$  soient Gaussiens indépendants avec la moyenne zéro et les parties réelles et imaginaires indépendantes. Donc, notre canal équivalent (9) (ou (10)) peut être considéré comme un canal Gaussien parallèle (Figure 2) dont la capacité se trouve dans la littérature [11].

$$C(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \sum_{i=1}^k \log(1 + \lambda_i P_i) \quad (11)$$

où  $P_i = \tilde{Q}_{ii} = E\{|\tilde{x}_i|^2\}$  et  $\sum P_i = P$ .



**Figure 2 : Canaux gaussiens équivalents parallèle**

Donc le problème se réduit à trouver l'ensemble des puissances qui maximisent la capacité de (11), en respectant la contrainte  $\sum P_i = P$ .

$$\text{Max}_{\substack{P_1, \dots, P_k \\ \sum_{i=1}^k P_i = P; P_i \geq 0}} C(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}; P_1, \dots, P_k) = \sum_{i=1}^k \log(1 + P_i \lambda_i)$$

On est confronté à un problème standard d'optimisation qui peut être résolu en utilisant la théorie des 'Multiplieurs de Lagrange' comme suit :

$$J(P_1, \dots, P_k) = \sum_{i=1}^k \log(1 + \lambda_i P_i) + \beta \sum_{i=1}^k P_i$$

$$\frac{\partial J}{\partial P_i} = 0 \Rightarrow \frac{\log(e) \lambda_i}{1 + P_i \lambda_i} + \beta = 0 \Rightarrow P_i = \underbrace{(-\log(e)/\beta)}_{\mu} - \lambda_i^{-1} = \mu - \lambda_i^{-1}$$

Mais comme les puissances  $P_i$  doivent être non-négatives, il n'est pas toujours possible de trouver une solution à cette forme. Dans ce cas on utilise les conditions Kuhn-Tucker pour trouver une solution qui maximise la capacité.

$$P_i = (\mu - \lambda_i^{-1})^+ ; \quad E\{\text{Re}(\tilde{x}_i)^2\} = E\{\text{Im}(\tilde{x}_i)^2\} = \frac{1}{2} (\mu - \lambda_i^{-1})^+ \quad (12)$$

où  $\mu$  est choisi pour satisfaire la contrainte de puissance et  $a^+$  dénote  $\max(0, a)$ . Pour  $P_i$  dans (12) on peut écrire :

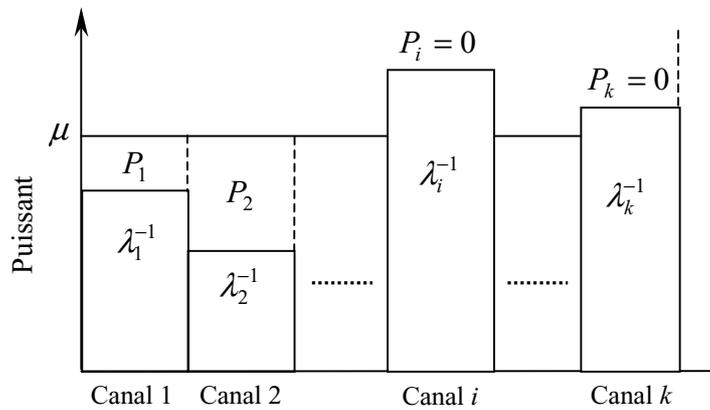
$$\log(1 + \lambda_i P_i) = \log\left(1 + \lambda_i (\mu - \lambda_i^{-1})^+\right) = \begin{cases} \log(1 + \lambda_i \times 0) = \log(1) = 0 & ; \mu < \lambda_i^{-1} \text{ ou } \mu \lambda_i < 1 \\ \log(1 + \mu \lambda_i - 1) = \log(\mu \lambda_i) & ; \mu \geq \lambda_i^{-1} \text{ ou } \mu \lambda_i \geq 1 \end{cases}$$

$$= (\log(\mu \lambda_i))^+$$

Donc la puissance  $P$  et l'information mutuelle maximale  $C$  peuvent être paramétrées comme suivant.

$$P(\mu) = \sum_i (\mu - \lambda_i^{-1})^+, C(\mu) = \sum_i (\ln(\mu \lambda_i))^+ \quad (13)$$

Cette solution connue sous le nom de Water-filling est graphiquement illustrée sur la Figure 3.



**Figure 3 : Water-filling pour les canaux parallèles**

**Remarque 1 (Réciprocité) :** Comme les valeurs propres non-zéros de  $\mathbf{H}^\dagger \mathbf{H}$  sont les mêmes que celles de  $\mathbf{H} \mathbf{H}^\dagger$ , les capacités des canaux correspondant aux matrices  $\mathbf{H}$  et  $\mathbf{H}^\dagger$  sont les mêmes.

**Remarque 2 :** Pour  $n > m$ , il y a des composants  $\tilde{x}_i$  ( $m < i \leq n$ ) qui ne jouent aucun rôle dans le canal et n'ont aucun effet sur l'information mutuelle. Il est donc logique que leurs puissances soient considérées nulle ( $P_i = 0$  ;  $i = k+1, \dots, n$ ).

**Remarque 2 :** Dans le canal équivalent (9), pour avoir l'information mutuelle maximum, il faut que  $\{\tilde{x}_i : 1 < i \leq k\}$  soient gaussien et avec la matrice d'auto covariance diagonale  $\tilde{\mathbf{Q}}$ . En plus pour atteindre la capacité maximum, les puissances doivent satisfaire la contrainte de puissance dans (13) et elles sont calculées par la méthode 'water-filling'. D'après cette dernière les puissances correspondantes aux valeurs propres zéros sont nulles et celles qui correspondent aux valeurs propres identiques sont identiques.

$$\begin{aligned} \lambda_i = 0 &\Rightarrow P_i = 0 \\ \lambda_i = \lambda_j &\Rightarrow P_i = P_j \quad ; \quad i, j = 1, \dots, k \end{aligned}$$

Cela nous rappelle une règle générale de la théorie d'information qui dit que pour arriver à la capacité de canal, les statistiques des signaux d'entrée doivent ressembler au canal. En plus d'après les définitions des  $\tilde{\mathbf{x}}$  et  $\mathbf{V}$ , nous avons  $\mathbf{Q} = \mathbf{V} \tilde{\mathbf{Q}} \mathbf{V}^\dagger$  et  $\mathbf{H}^\dagger \mathbf{H} = \mathbf{V} \mathbf{D}^v \mathbf{V}^\dagger$  ( $\mathbf{D}^v = \mathbf{D}^T \mathbf{D}$ ), et en considérant la ressemblance entre  $\tilde{\mathbf{Q}}$  et  $\mathbf{D}^v$ , on peut apercevoir la ressemblance entre  $\mathbf{Q}$  et  $\mathbf{H}^\dagger \mathbf{H}$  pour le canal original (1).

Pour mieux comprendre les relations (12) et (13) et la remarque 2, nous allons considérer quelques exemples.

**Exemple 1 :** Supposons  $H_{i,j} = 1$  ;  $\forall i, j$ . Le canal est complètement corrélé. La décomposition SVD minimisée de  $\mathbf{H}$  (ref. Annexe D) peut s'écrire comme :

$$\mathbf{H} = \begin{bmatrix} \sqrt{1/m} \\ \vdots \\ \sqrt{1/m} \end{bmatrix}_{m \times 1} (\sqrt{mn}) \begin{bmatrix} \sqrt{1/n} & \dots & \sqrt{1/n} \end{bmatrix}_{1 \times n}$$

Pour ce canal la seule valeur propre non nulle de  $\mathbf{H}^\dagger \mathbf{H}$  (ou  $\mathbf{H} \mathbf{H}^\dagger$ ) est  $mn$  et donc la seule valeur singulière non nulle de  $\mathbf{H}$  est  $\sqrt{mn}$  et les premières colonnes des matrices  $\mathbf{U}$  et  $\mathbf{V}$  (correspondant à cette valeur singulière) sont respectivement  $\sqrt{1/m} [1, \dots, 1]^\dagger$  et  $\sqrt{1/n} [1, \dots, 1]^\dagger$ . Pour trouver la capacité de ce canal, on revient aux formules (12) et (13).

$$\lambda_1 = mn, \lambda_2 = \dots = \lambda_k = 0 \Rightarrow P_2 = \dots = P_n = 0, P_1 = P = \mu - (mn)^{-1} \Rightarrow \mu = \frac{1 + Pmn}{mn}$$

$$C = \log(\mu \lambda_1) = \log\left(\frac{1 + Pmn}{mn} mn\right) \Rightarrow C = \log(1 + Pmn)$$

Comme on l'a remarqué, pour arriver à cette capacité la matrice d'auto covariance de  $\tilde{\mathbf{x}}$  ( $\tilde{\mathbf{Q}}$ ) est une matrice diagonale dont tous les éléments sont zéros sauf le premier. Donc la matrice d'auto covariance de  $\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$  qui atteint cette capacité a la structure suivante.

$$\left. \begin{aligned} \tilde{\mathbf{Q}} &= E\{\tilde{\mathbf{x}}\tilde{\mathbf{x}}^\dagger\} = \text{diag}(P, 0, \dots, 0) \\ \mathbf{Q} &= E\{\mathbf{x}\mathbf{x}^\dagger\} = \mathbf{V}\tilde{\mathbf{Q}}\mathbf{V}^\dagger \end{aligned} \right\} \Rightarrow \mathbf{Q} = \sqrt{1/n} [1, \dots, 1]^\dagger \cdot P \cdot \sqrt{1/n} [1, \dots, 1] = \frac{P}{n} \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix}$$

Autrement dit  $E\{x_i x_j^*\} = P/n$  ;  $\forall i, j$ , c-à-d les émetteurs transmettent tous à la même puissance. On peut remarquer encore une fois la ressemblance des matrices  $\mathbf{Q}$  et  $\mathbf{H}^\dagger \mathbf{H}$ . On note que même si la puissance d'émission de chaque émetteur est  $P/n$ , comme leurs signaux au récepteur sont cohérents, la puissance reçue à chaque récepteur est  $Pn$  ( $n^2 P/n$ ) avec un bruit de puissance 1. Et comme chaque récepteur voit le même signal et les bruits aux récepteurs sont non-corrélés, le rapport total signal à bruit est  $Pmn$  (En fait la puissance du signal est multipliée par  $m^2$  mais la variance du bruit est multipliée par  $m$ ).

**Exemple 2 :** On suppose que  $n = m$  et  $\mathbf{H} = \mathbf{I}_n$ . Pour ce canal non-corrélé, nous avons :

$$U = V = \mathbf{D} = \mathbf{I}_n \Rightarrow \lambda_1 = \lambda_2 = \dots = \lambda_n = 1 \Rightarrow P = \sum_{i=1}^n (\mu - \lambda_i) = n(\mu - 1) \Rightarrow \mu = 1 + \frac{P}{n}$$

$$C = \sum_{i=1}^n \log(1 + P/n) \Rightarrow C = n \log(1 + P/n)$$

Donc pour atteindre la capacité tous les signaux ont la même puissance ( $P/n$ ), c'est à dire :

$$\tilde{\mathbf{Q}} = \frac{P}{n} \mathbf{I}_n \Rightarrow \mathbf{Q} = \mathbf{V}\tilde{\mathbf{Q}}\mathbf{V}^\dagger = \mathbf{V} \cdot \frac{P}{n} \mathbf{I}_n \cdot \mathbf{V}^\dagger = \frac{P}{n} \mathbf{V}\mathbf{V}^\dagger = \frac{P}{n} \mathbf{I}_n$$

Ainsi afin d'atteindre la capacité d'un canal non-corrélé, il faut avoir une source non-corrélée (les composants de  $\mathbf{x}$  sont i.i.d. avec  $E\{x_i x_j^*\} = P/n \delta_{i,j}$ ). On ne peut pas pourtant déduire de cette conclusion que pour atteindre la capacité, il faut un codage indépendant pour chaque émetteur. Il est vrai que la capacité de ce canal peut être atteinte en divisant la séquence de données arrivée en  $t$  séquences qui seront codées et modulée séparément et envoyées sur les différents émetteurs. Supposons que  $Nt$  bits vont être transmis, on a deux possibilités pour atteindre la capacité de ce canal :

Soit nous allons les séparer en  $t$  groupes chacun de  $N$  bits et utiliser chaque groupe afin de sélectionner un des  $2^N$  signaux pour chaque émetteur,

Soit nous allons utiliser tous les  $Nt$  bits pour sélectionner un des  $2^{Nt}$  signaux.

La deuxième alternative donne une probabilité d'erreur nettement meilleure que la première, au prix d'une complexité beaucoup plus importante. En effet le log de la probabilité d'erreur pour ces deux cas-là diffère par un facteur de  $t$ . (voir 'error exponents' des canaux parallèles dans [4, pp. 149-150].)

De toute manière, vu la source qui permet d'atteindre la capacité, on peut considérer ce canal MIMO, comme  $n$  canaux parallèles indépendant avec le  $SNR = P/n$  pour chaque canal, ce

qui confirme la capacité obtenue. Dans le cas où le nombre d'antennes de réception et d'émission ( $m = n$ ) augmenterait pour un canal non-corrélé, la capacité se calcule selon :

$$\lim_{n \rightarrow \infty} C = n \log(1 + P/n) = \lim_{1/n \rightarrow 0} \frac{\log(1 + P/n)}{1/n} = \log(e)$$

### 2-4-2 Autre expression de la capacité

Dans la section précédente on a supposé sans le démontrer que la source  $\tilde{\mathbf{x}}$  qui permettait d'atteindre la capacité de canal (9) devait être gaussienne avec la matrice d'auto covariance  $\tilde{\mathbf{Q}}$  diagonale et on a également supposé vérifiée la formule (11). Maintenant nous allons présenter un calcul plus détaillé de la capacité ; pour cela on revient au canal original (1) ( $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$ ). Pour ce canal l'information mutuelle  $I(\mathbf{x}; \mathbf{y})$  peut s'écrire comme :

$$I(\mathbf{x}; \mathbf{y}) = H(\mathbf{y}) - H(\mathbf{y}|\mathbf{x}) = H(\mathbf{y}) - H(\mathbf{n}) \quad (14)$$

Comme  $H(\mathbf{n})$  ne dépend pas de la source, la maximisation de  $I(\mathbf{x}; \mathbf{y})$  est équivalente à maximiser  $H(\mathbf{y})$ . On note que si  $\mathbf{x}$  satisfait la contrainte de puissance  $E\{\mathbf{x}^\dagger \mathbf{x}\} \leq P$ , alors  $\mathbf{x} - E\{\mathbf{x}\}$  le fera aussi. On peut donc se concentrer sur  $\mathbf{x}$  en le supposant de moyenne zéro. En outre, si  $\mathbf{x}$  est de moyenne nulle avec la matrice de covariance  $\mathbf{Q} = E\{\mathbf{x}\mathbf{x}^\dagger\}$ , les statistiques de  $\mathbf{y}$  seront les suivantes :

$$E\{\mathbf{y}\} = \mathbf{0} \quad ; \quad E\{\mathbf{y}\mathbf{y}^\dagger\} = \mathbf{H}E\{\mathbf{x}\mathbf{x}^\dagger\}\mathbf{H}^\dagger + E\{\mathbf{n}\mathbf{n}^\dagger\} = \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger + \mathbf{I}_m$$

D'après le lemme 2 parmi les  $\mathbf{y}$  avec de telles statistiques,  $H(\mathbf{y})$  sera maximum, quand  $\mathbf{y}$  est GCCS, ce qui sera le cas quand  $\mathbf{x}$  est GCCS (Lemme 3 et 4).

$$\text{Max } I(\mathbf{x}; \mathbf{y}) \equiv \text{Max } H(\mathbf{y}) \Rightarrow \mathbf{y} : \text{GCCS} \Rightarrow \mathbf{x} : \text{GCCS}$$

Donc désormais, on peut se concentrer sur  $\mathbf{x}$  GCCS. Dans ce cas-là, l'information mutuelle peut se calculer à partir de (7).

$$\begin{aligned} I(\mathbf{x}; \mathbf{y}) &= \log(\det(\pi e(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger))) - \log(\det(\pi e\mathbf{I}_m)) \\ &= \log(\det(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger)) = \log(\det(\mathbf{I}_n + \mathbf{Q}\mathbf{H}^\dagger \mathbf{H})) \end{aligned} \quad (15)$$

où la dernière égalité se déduit de l'identité des déterminants :

$$\det(\mathbf{I}_M + \mathbf{A}_{M \times N} \mathbf{B}_{N \times M}) = \det(\mathbf{I}_N + \mathbf{B}_{N \times M} \mathbf{A}_{M \times N}) \quad (16)$$

Donc il reste juste à maximiser la quantité  $I(\mathbf{x}; \mathbf{y})$  dans (15) par rapport aux matrices non-négatives définies  $\mathbf{Q}$ , en respectant la contrainte de puissance  $\text{tr}(\mathbf{Q}) \leq P$ . Et comme la quantité  $\log(\det(\mathbf{I} + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger))$  sera utilisée très fréquemment dans ce chapitre, on propose la notation suivante :

$$\text{Max } I(\mathbf{x}; \mathbf{y}) \equiv \text{Max}_{\text{tr}(\mathbf{Q}) \leq P} \Psi(\mathbf{Q}; \mathbf{H}) = \log(\det(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger)) \quad (17)$$

Comme  $\mathbf{H}^\dagger \mathbf{H}$  est Hermitienne et non-négative définie, elle peut être diagonalisée (ref. Annexe D),  $\mathbf{H}^\dagger \mathbf{H} = \mathbf{V} \mathbf{A} \mathbf{V}^\dagger$ , avec la matrice unitaire  $\mathbf{V}$  et non-négative diagonale  $\mathbf{A} = \text{diag}(\lambda_1, \dots, \lambda_n)$ . En remplaçant  $\mathbf{H}^\dagger \mathbf{H}$  par sa décomposition en valeurs propres et en appliquant l'identité de déterminant, on aura :

$$\begin{aligned} \det(\mathbf{I}_m + \mathbf{H} \mathbf{Q} \mathbf{H}^\dagger) &= \det(\mathbf{I}_n + \mathbf{Q} \mathbf{H}^\dagger \mathbf{H}) = \det(\mathbf{I}_n + \mathbf{Q} \mathbf{V} \mathbf{A} \mathbf{V}^\dagger) \\ &= \det(\mathbf{I}_n + (\mathbf{Q} \mathbf{V} \mathbf{A}^{1/2})(\mathbf{A}^{1/2} \mathbf{V}^\dagger)) = \det(\mathbf{I}_n + \mathbf{A}^{1/2} \mathbf{V}^\dagger \mathbf{Q} \mathbf{V} \mathbf{A}^{1/2}) = \det(\mathbf{I}_n + \mathbf{A}^{1/2} \tilde{\mathbf{Q}} \mathbf{A}^{1/2}) \end{aligned} \quad (18)$$

où  $\tilde{\mathbf{Q}} = \mathbf{V}^\dagger \mathbf{Q} \mathbf{V}$  est également une matrice non-négative définie si et seulement si  $\mathbf{Q}$  l'est, de plus, selon (3) on peut vérifier que :

$$\text{tr}(\tilde{\mathbf{Q}}) = \text{tr}(\mathbf{V}^\dagger \mathbf{Q} \mathbf{V}) = \text{tr}(\mathbf{Q} \mathbf{V} \mathbf{V}^\dagger) = \text{tr}(\mathbf{Q})$$

Donc on aura la même contrainte de puissance pour  $\tilde{\mathbf{Q}}$  ( $\text{tr}(\tilde{\mathbf{Q}}) \leq P$ ) et la maximisation sur  $\mathbf{Q}$  peut être transposée sur  $\tilde{\mathbf{Q}}$ .

$$\text{Max}_{\substack{\mathbf{Q} \\ \text{tr}(\mathbf{Q}) \leq P}} I(\mathbf{x}; \mathbf{y}) \equiv \text{Max}_{\substack{\tilde{\mathbf{Q}} \\ \text{tr}(\tilde{\mathbf{Q}}) \leq P}} \Psi(\mathbf{Q}, \mathbf{H}) \equiv \text{Max}_{\substack{\tilde{\mathbf{Q}} \\ \text{tr}(\tilde{\mathbf{Q}}) \leq P}} \Psi(\tilde{\mathbf{Q}}, \mathbf{A}) = \log(\det(\mathbf{I}_n + \mathbf{A}^{1/2} \tilde{\mathbf{Q}} \mathbf{A}^{1/2})) \quad (19)$$

Pour la suite, on se concentre sur la matrice  $\mathbf{I}_n + \mathbf{A}^{1/2} \tilde{\mathbf{Q}} \mathbf{A}^{1/2}$ . Cette dernière est également une matrice hermitienne non-négative définie et pour une matrice non-négative définie quelconque  $\mathbf{A}$ , on a l'inégalité :  $\det(\mathbf{A}) \leq \prod_i A_{ii}$ , donc :

$$\det(\mathbf{I}_n + \mathbf{A}^{1/2} \tilde{\mathbf{Q}} \mathbf{A}^{1/2}) \leq \prod_i (1 + \lambda_i \tilde{Q}_{ii}) \quad (20)$$

avec l'égalité quand  $\tilde{\mathbf{Q}}$  est diagonale. Donc en résumé, pour atteindre la capacité de canal, la source  $\mathbf{x}$  doit être premièrement GCCS avec la moyenne zéro, et deuxièmement avoir une matrice de covariance  $\mathbf{Q}$  telle que  $\tilde{\mathbf{Q}} = \mathbf{V}^\dagger \mathbf{Q} \mathbf{V}$  soit diagonale. Pour une telle source (19) s'écrit sous la forme suivante :

$$\text{Max}_{Q_{11}, \dots, Q_{nn}} C(\mathbf{x}; \mathbf{y}) = \log(\det(\mathbf{I}_n + \mathbf{A}^{1/2} \tilde{\mathbf{Q}} \mathbf{A}^{1/2})) = \log\left(\prod_i (1 + \lambda_i \tilde{Q}_{ii})\right) = \sum_i (1 + \lambda_i \tilde{Q}_{ii}) \quad (21)$$

Il reste seulement à obtenir les éléments diagonaux optimaux de la matrice  $\tilde{\mathbf{Q}}$  (les mêmes puissances du canal équivalent (9)) qui peuvent être trouvés par la méthode 'water-filling' (voir Figure 3).

$$\tilde{Q}_{ii} = (\mu - \lambda_i^{-1})^+ ; \quad i = 1, \dots, n \quad (22)$$

où le paramètre  $\mu$  est choisi, pour satisfaire la contrainte de puissance et la capacité correspondant à ces éléments sera définie par :

$$P(\mu) = \sum_i (\mu - \lambda_i^{-1})^+, C(\mu) = \sum_i (\ln(\mu \lambda_i))^+ \quad (23)$$

Donc on arrive finalement exactement aux même résultats présentés que ceux de la section précédente (équations (12) et (13)).

### 2-4-3 « Error Exponents »

Connaître la capacité d'un canal n'est pas toujours suffisant. Par contre, il peut être intéressant de savoir à quel point peut-on s'approcher de cette capacité. Les « erreurs exponents » donnent une réponse partielle à cette question en calculant une borne supérieure de la probabilité d'erreur qui peut être atteinte avec des codes en bloc de largeur  $N$  et de rendement  $R$ . Cette borne supérieure connue comme la borne des codes aléatoires s'écrit :

$$P(\text{erreur}) \leq \exp(-N.E_r(R))$$

où  $E_r(R)$ , l'exposant de codage aléatoire, est donné par :

$$E_r(R) = \max_{0 \leq \rho \leq 1} (E_0(\rho) - \rho R)$$

et où  $E_0(\rho)$  se calcule en maximisant l'expression suivante sur toutes les distributions  $q_x$  satisfaisant la contrainte d'énergie.

$$E_0(\rho, q_x) = -\log \left( \int \left[ \int q_x(\mathbf{x}) p(\mathbf{y}|\mathbf{x})^{1/(1+\rho)} d\mathbf{x} \right]^{1+\rho} d\mathbf{y} \right) \quad (24)$$

Dans notre cas  $p(\mathbf{y}|\mathbf{x}) = \det(\pi \mathbf{I}_r)^{-1} \exp(-(\mathbf{y} - \mathbf{H}\mathbf{x})^\dagger (\mathbf{y} - \mathbf{H}\mathbf{x}))$  et en choisissant  $q_x$  comme une distribution gaussienne  $f_Q$  et après quelques manipulations (preuve dans l'annexe E), nous arrivons à :

$$E_0(\rho, \mathbf{Q}) = \rho \log \left( \det(\mathbf{I}_r + (1 + \rho)^{-1} \mathbf{H} \mathbf{Q} \mathbf{H}^\dagger) \right) = \rho \Psi \left( (1 + \rho)^{-1} \mathbf{Q}, \mathbf{H} \right) \quad (25)$$

où la fonction  $\Psi$  est déjà définie dans (17). Avec notre hypothèse sur la distribution de  $q_x$ , la maximisation de  $E_0(\rho, q_x)$  sur  $q_x$  est équivalente à la maximisation de  $E_0(\rho, \mathbf{Q})$  sur  $\mathbf{Q}$ . Cette maximisation de  $E_0$  sur  $\mathbf{Q}$  constitue exactement le même problème que de maximiser l'information mutuelle de l'expression (17) sauf que  $\text{tr}((1 + \rho)^{-1} \mathbf{Q}) \leq P/(1 + \rho)$ . Donc on peut dire que :

$$E_0(\rho) = \rho C(P/(1 + \rho), \mathbf{H}) \quad (26)$$

Le choix de  $q_x$  Gaussien n'est pas optimal, et certaines autres distributions donnent de meilleurs résultats [4, §7.3]. Néanmoins, l'expression ci-dessus est une borne inférieure très proche de  $E_0$  et donc elle donne une bonne approximation de la borne supérieure de la probabilité d'erreur.

## 2-5 Canal Gaussien avec Fading de Rayleigh

Cette fois-ci, on suppose que la matrice  $\mathbf{H}$  n'est pas déterministe, elle est une matrice aléatoire et indépendante de  $\mathbf{x}$  et  $\mathbf{n}$  pour laquelle on utilise la notation  $\mathbf{H}$ . La matrice  $\mathbf{H}$  est censée d'être connue au récepteur (Estimation parfaite du canal) mais pas à l'émetteur. Le canal peut donc être considéré avec une entrée  $\mathbf{x}$  et deux sorties  $\mathbf{y}$  et  $\mathbf{H}$  (Figure 4)

$$(\mathbf{y}, \mathbf{H}) = (\mathbf{H}\mathbf{x} + \mathbf{n}, \mathbf{H}) \quad (27)$$

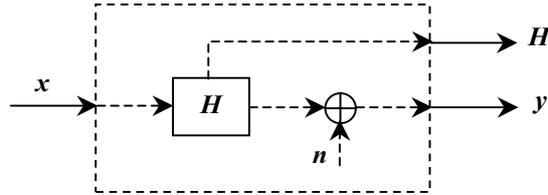


Figure 4 : Canal de Gauss avec fading de Rayleigh et estimation parfaite

On suppose que les éléments de  $\mathbf{H}$  sont indépendants, gaussiens complexes de moyenne zéro et de parties réelle et imaginaire indépendantes avec une variance  $1/2$  pour chacune d'entre elles. Par conséquent pour chaque élément, la phase a une distribution uniforme  $[0, 2\pi]$  et l'amplitude a une distribution de Rayleigh avec l'espérance de l'amplitude au carré égale à 1. Ceci correspond à la modélisation d'un canal à fading de Rayleigh avec une distance de séparation suffisante entre les antennes d'émission et de réception afin d'avoir l'indépendance entre les éléments de  $\mathbf{H}$ . Pour calculer la capacité, nous avons d'abord besoin de montrer qu'une telle matrice  $\mathbf{H}$  est invariante avec les transformations unitaires.

**Lemme 5 :** Supposons  $\mathbf{H} \in \mathbb{C}^{m \times n}$  est une matrice complexe gaussienne avec les éléments distribués identiquement, et les parties réelle et imaginaire de chaque élément sont indépendantes avec la moyenne zéro et la variance identique. Pour n'importe quelle matrice unitaire  $\mathbf{U} \in \mathbb{C}^{m \times m}$  et  $\mathbf{V} \in \mathbb{C}^{n \times n}$ , la distribution de  $\mathbf{U}\mathbf{H}\mathbf{V}^\dagger$  est identique à celle de  $\mathbf{H}$ .

*Preuve :* Il suffit juste de démontrer que  $\mathbf{G} = \mathbf{U}\mathbf{H}$  a la même distribution que  $\mathbf{H}$ . La preuve suit ensuite d'une application de ce dernier pour  $\mathbf{G}^\dagger$ . Comme les colonnes de  $\mathbf{H}$  sont indépendantes, les colonnes de  $\mathbf{G}$  le sont également. Donc il reste à vérifier si chaque colonne de  $\mathbf{G}$  a la même distribution que  $\mathbf{H}$ . Comme les colonnes de  $\mathbf{H}$  sont GCCS, les colonnes de  $\mathbf{G}$  le seront aussi. Si  $\mathbf{G}_j$  et  $\mathbf{H}_j$  représentent respectivement les colonnes  $j$  de  $\mathbf{G}$  et de  $\mathbf{H}$ , on aura donc :

$$E\{\mathbf{G}_j \mathbf{G}_j^\dagger\} = \mathbf{U} E\{\mathbf{H}_j \mathbf{H}_j^\dagger\} \mathbf{U}^\dagger = \mathbf{U} (\mathbf{K} \mathbf{I}_m) \mathbf{U}^\dagger = \mathbf{K} \mathbf{U} \mathbf{U}^\dagger = \mathbf{K} \mathbf{I}_m = E\{\mathbf{H}_j \mathbf{H}_j^\dagger\}$$

Dans cette section on suppose que le canal est sans mémoire et pour chaque utilisation de ce canal, une réalisation indépendante de  $\mathbf{H}$  est appliquée. Dans ce cas la capacité peut se calculer comme l'information mutuelle maximale. De toute manière les résultats qui suivent sont valides pour les canaux pour lesquels  $\mathbf{H}$  est générée avec un processus ergodique et

seulement les statistiques de  $\mathbf{H}$  du premier ordre sont nécessaires pour déterminer la capacité de canal.

### 2-5-1 Capacité

Comme le récepteur connaît la matrice  $\mathbf{H}$ , on considère le canal de l'expression (27). Pour ce canal l'information mutuelle entre l'entrée et les sorties est :

$$\begin{aligned} I(\mathbf{x}; (\mathbf{y}, \mathbf{H})) &= I(\mathbf{x}; \mathbf{H}) + I(\mathbf{x}; \mathbf{y} | \mathbf{H}) \\ &= I(\mathbf{x}; \mathbf{y} | \mathbf{H}) \\ &= E_{\mathbf{H}} \{ I(\mathbf{x}; \mathbf{y} | \mathbf{H} = \mathbf{H}) \} \end{aligned} \quad (28)$$

On sait d'après la section précédente que le choix de  $\mathbf{x}$  (avec la contrainte d'avoir la matrice de covariance  $\mathbf{Q}$ ) qui maximise  $I(\mathbf{x}; \mathbf{y} | \mathbf{H} = \mathbf{H})$  est un vecteur GCCS de même covariance  $\mathbf{Q}$  et  $\Psi(\mathbf{Q}; \mathbf{H}) = \log(\det(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger))$  est l'information mutuelle maximale correspondante. Nous devons donc maximiser l'expression suivante

$$\text{Max}_{\substack{\mathbf{Q} \\ \text{tr}(\mathbf{Q}) \leq P}} \Psi(\mathbf{Q}) = E \{ \Psi(\mathbf{Q}; \mathbf{H}) \} = E \{ \log(\det(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger)) \} \quad (29)$$

où  $\mathbf{Q}$  est non-négative définie et satisfait  $\text{tr}(\mathbf{Q}) \leq P$ .

Comme  $\mathbf{Q}$  est non-négative définie, nous pouvons l'écrire comme  $\mathbf{Q} = \mathbf{V}\mathbf{D}\mathbf{V}^\dagger$  où  $\mathbf{V}$  est unitaire et  $\mathbf{D}$  est diagonale et non-négative définie. Ces deux matrices sont également les matrices des vecteurs propres et des valeurs propres de  $\mathbf{Q}$ . Avec cette expression de  $\mathbf{Q}$ , (29) se réécrit sous la forme suivante :

$$\Psi(\mathbf{Q}) = E \{ \log(\det(\mathbf{I}_m + (\mathbf{H}\mathbf{V})\mathbf{D}(\mathbf{H}\mathbf{V})^\dagger)) \} = \Psi(\mathbf{D}) \Rightarrow \text{Max}_{\substack{\mathbf{Q} \\ \text{tr}(\mathbf{Q}) \leq P}} \Psi(\mathbf{Q}) \equiv \text{Max}_{\substack{\mathbf{D} \\ \text{tr}(\mathbf{D}) \leq P}} \Psi(\mathbf{D}) \quad (30)$$

Selon le lemme (5), la distribution de  $\mathbf{H}\mathbf{V}$  est la même distribution que  $\mathbf{H}$  et par conséquent  $\Psi(\mathbf{Q}) = \Psi(\mathbf{D})$ . Nous pouvons donc juste se concentrer sur  $\mathbf{Q}$  diagonale et non-négative définie. Pour une matrice diagonale  $\mathbf{Q}$  donnée et une matrice de permutation  $\mathbf{\Pi}$ , on considère  $\mathbf{Q}^\Pi = \mathbf{\Pi}\mathbf{Q}\mathbf{\Pi}^\dagger$  (elle sera aussi diagonale, mais avec les éléments permutés par rapport à  $\mathbf{Q}$ ). Comme  $\mathbf{H}\mathbf{\Pi}$  a la même distribution que  $\mathbf{H}$  (les matrices de permutation sont unitaires),  $\Psi(\mathbf{Q}^\Pi) = \Psi(\mathbf{Q})$ . En plus, on note que pour n'importe quelle matrice  $\mathbf{H}$ , la projection  $\mathbf{Q} \mapsto \mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger$  est linéaire et préserve le caractère positif. Comme  $\log(\det(\cdot))$  est concave sur l'ensemble des matrices positive définies, la projection  $\mathbf{Q} \mapsto \Psi(\mathbf{Q}; \mathbf{H}) = \log(\det(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger))$  sera également concave. Donc pour la matrice  $\tilde{\mathbf{Q}}$  définie comme suit (la moyenne de toutes les permutations possibles pour une matrice diagonale non-négative définie  $\mathbf{Q}$ ), on aura :

$$\tilde{\mathbf{Q}} = \frac{1}{n!} \sum_{\Pi} \mathbf{Q}^\Pi \Rightarrow \Psi(\tilde{\mathbf{Q}}) \geq \Psi(\mathbf{Q}) \quad ; \quad \text{tr}(\tilde{\mathbf{Q}}) = \text{tr}(\mathbf{Q}) \quad (31)$$

*Preuve :* Le premier terme se déduit directement de la concavité de  $\Psi(\mathbf{Q}; \mathbf{H})$  (et par conséquent de  $\Psi(\mathbf{Q})$ ) par rapport à  $\mathbf{Q}$  et du fait que  $\Psi(\mathbf{Q}^\Pi) = \Psi(\mathbf{Q})$ .

$$\Psi(\tilde{\mathbf{Q}}) = \Psi\left(\frac{1}{n!} \sum_{\Pi} \mathbf{Q}^\Pi\right) \geq \sum_{\Pi} \frac{1}{n!} \Psi(\mathbf{Q}^\Pi) = \frac{1}{n!} \sum_{\Pi} \Psi(\mathbf{Q}) = \Psi(\mathbf{Q}) \Rightarrow \Psi(\tilde{\mathbf{Q}}) \geq \Psi(\mathbf{Q})$$

Pour le deuxième terme, on se rappelle que  $\mathbf{Q}$  est diagonale ( $\mathbf{Q} = \text{diag}(q_1, q_2, \dots, q_n)$ ) et chaque  $\mathbf{Q}^\Pi$  est une permutation de  $\mathbf{Q}$ . Parmi  $n!$  permutations possibles, sur chaque élément diagonal, nous avons  $(n-1)!$  fois les termes  $q_1, q_2, \dots, q_n$ . Donc :

$$\begin{aligned} \mathbf{Q} = \text{diag}(q_1, \dots, q_n) & \Rightarrow \tilde{\mathbf{Q}} = \frac{1}{n!} (\sum q_i) (n-1)! \mathbf{I}_n = \frac{\sum q_i}{n} \mathbf{I}_n = \alpha \mathbf{I}_n \\ \tilde{\mathbf{Q}} = \frac{1}{n!} \sum_{\Pi} \mathbf{Q}^\Pi & \Rightarrow \text{tr}(\tilde{\mathbf{Q}}) = n \left( \frac{\sum q_i}{n} \right) = \sum q_i = \text{tr}(\mathbf{Q}) \end{aligned}$$

On s'aperçoit que  $\tilde{\mathbf{Q}}$  est une multiple de la matrice d'identité et on conclut d'après (31) que  $\mathbf{Q}$  optimale doit être de la forme  $\alpha \mathbf{I}_n$ . Il est clair que le maximum est atteint, quand  $\alpha$  est le plus grand possible, c'est à dire  $\alpha = P/n$ . On peut récapituler les résultats de cette section sous la forme suivante.

Pour atteindre la capacité 1-  $x$  doit être GCCS. 2- D'après (30) on peut s'occuper uniquement des matrices  $\mathbf{Q}$  diagonales. 3- Parmi ces matrices diagonale non-négative définie, celle qui est optimale, est de la forme  $\alpha \mathbf{I}_n$ . 4- En respectant la contrainte d'énergie,  $\alpha$  optimal est  $P/n$ .

Tout cela peut être résumé dans le théorème suivant.

**THEOREME 1.** Le vecteur aléatoire  $\mathbf{x}$  qui nous permet d'atteindre la capacité, est GCCS avec la moyenne zéro et la matrice de covariance  $(P/n) \mathbf{I}_n$ . La capacité correspondante sera :

$$C = E\left\{\log\left(\det\left(\mathbf{I}_m + (P/n) \mathbf{H} \mathbf{H}^\dagger\right)\right)\right\} \quad (32)$$

On va présenter le calcul précis de (32) dans la section suivante. Mais déjà lorsque  $m$  est fixe, on remarque selon la loi des grands nombres que  $\frac{1}{n} \mathbf{H} \mathbf{H}^\dagger \rightarrow \mathbf{I}_m$  quand  $n$  devient très grand.

Donc la capacité pour les grandes valeurs de  $n$  prend la forme

$$C \rightarrow \log(\det(\mathbf{I}_m + P \mathbf{I}_m)) = \log(\det((1+P) \mathbf{I}_m)) = \log((1+P)^m) = m \log(1+P) \quad (33)$$

On s'aperçoit dans ce cas que la capacité est proportionnelle au nombre d'antennes de réception.

### 2-5-2 Evaluation de la capacité

Même si l'expression (32) est facile à évaluer pour  $m = n = 1$ , cette évaluation devient compliquée pour les valeurs plus importantes de  $m$  et  $n$ . Dans cette section on va montrer comment évaluer l'expression (32). En utilisant l'identité du déterminant on peut dire que :

$$\det(\mathbf{I}_m + (P/n)\mathbf{H}\mathbf{H}^\dagger) = \det(\mathbf{I}_n + (P/n)\mathbf{H}^\dagger\mathbf{H})$$

Et on définit la matrice  $\mathbf{W}$  selon

$$\mathbf{W}_{p \times p} = \begin{cases} \mathbf{H}\mathbf{H}^\dagger & m < n \\ \mathbf{H}^\dagger\mathbf{H} & m \geq n \end{cases}, \quad p = \min\{m, n\}, \quad q = \max\{m, n\} \quad (34)$$

$\mathbf{W}_{p \times p}$  est une matrice non-négative définie et par conséquent elle possède les valeurs propres  $(\lambda_1, \lambda_2, \dots, \lambda_p)$  réelles et non-négatives. En utilisant (34) l'expression (31) peut s'écrire sous la forme suivante.

$$C = E\left\{\log\left(\det(\mathbf{I}_p + (P/n)\mathbf{W}^\dagger)\right)\right\} \quad (35)$$

Cette dernière peut se simplifier en se servant de la décomposition en valeurs propres de  $\mathbf{W}$  ( $\mathbf{W} = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger$  où  $\mathbf{D}$  est une matrice diagonale contenant les valeurs propres de  $\mathbf{W}$  (les mêmes valeurs singulières de  $\mathbf{H}$ ), et  $\mathbf{U}$  la matrice unitaire des vecteurs propres) et de l'identité du déterminant (16).

$$\begin{aligned} \det(\mathbf{I}_p + (P/n)\mathbf{W}^\dagger) &= \det(\mathbf{I}_p + (P/n)\mathbf{U}\mathbf{D}\mathbf{U}^\dagger) = \det(\mathbf{I}_p + (P/n)\mathbf{U}^\dagger\mathbf{U}\mathbf{D}) \\ &= \det(\mathbf{I}_p + (P/n)\mathbf{U}^\dagger\mathbf{U}\mathbf{D}) = \det(\mathbf{I}_p + (P/n)\mathbf{D}) = \prod_{i=1}^p (1 + (P/n)\lambda_i) \end{aligned} \quad (36)$$

Donc on peut finalement réécrire la capacité en fonction des valeurs propres  $\lambda_1, \dots, \lambda_p$  de  $\mathbf{W}$ .

$$C = E\left\{\log\left(\prod_{i=1}^p (1 + (P/n)\lambda_i)\right)\right\} = E\left\{\sum_{i=1}^p \log(1 + (P/n)\lambda_i)\right\} \quad (37)$$

Et le moyennage sur  $\mathbf{H}$  est remplacé par le moyennage sur les valeurs propres de  $\mathbf{W}$ . Pour ce faire, on a besoin de connaître la distribution de ces valeurs propres. La loi de distribution de  $\mathbf{W}$  s'appelle la 'distribution Wishart avec les paramètres  $p$  et  $q$ ' et la loi de probabilité conjointe de ses valeurs propres dans l'ordre décroissant est : (voir [5] ou [6], une simple présentation se trouve également dans l'annexe F)

$$P_{\lambda, \text{ordered}}(\lambda_1, \dots, \lambda_p) = K_{p,q}^{-1} e^{-\sum_i \lambda_i} \prod_i \lambda_i^{p-q} \prod_{j>i} (\lambda_i - \lambda_j)^2, \quad \lambda_1 \geq \dots \geq \lambda_p$$

où  $K_{p,q}$  est la facture de normalisation. Les valeurs propres non-ordonnées auront donc la densité suivante.

$$p_\lambda(\lambda_1, \dots, \lambda_p) = (p! K_{p,q})^{-1} e^{-\sum_i \lambda_i} \prod_i \lambda_i^{p-q} \prod_{j>i} (\lambda_i - \lambda_j)^2 \quad (38)$$

Vu la symétrie de  $p_\lambda$  par rapport aux  $\lambda_1, \dots, \lambda_p$ , l'espérance de l'expression (37) se réécrit sous la forme suivante.

$$\begin{aligned} C &= E \left\{ \sum_{i=1}^p \log(1 + (P/n)\lambda_i) \right\} = \sum_{i=1}^p E \{ \log(1 + (P/n)\lambda_i) \} \\ &= p E \{ \log(1 + (P/n)\lambda_1) \} \end{aligned} \quad (39)$$

Donc l'espérance que l'on veut calculer ne dépend que de la distribution d'une des valeurs propres non-ordonnées. Pour calculer la densité de  $\lambda_1$ , nous aurons juste besoin de faire l'intégration sur  $\lambda_2, \dots, \lambda_p$ .

$$p_{\lambda_1}(\lambda_1) = \int \dots \int p_\lambda(\lambda_1, \lambda_2, \dots, \lambda_p) d\lambda_2 \dots d\lambda_p \quad (40)$$

Ce qui complique cette intégration, c'est le dernier terme de (38) (à cause de ce terme, l'intégrale (40) n'est pas séparable). On peut pourtant noter que  $\prod_{i<j} (\lambda_i - \lambda_j)$  est le déterminant d'une matrice *Vandermonde*.

$$\mathbf{D}(\lambda_1, \dots, \lambda_p) = \begin{bmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_p \\ \vdots & & \vdots \\ \lambda_1^{p-1} & \dots & \lambda_p^{p-1} \end{bmatrix} \Rightarrow \det(\mathbf{D}(\lambda_1, \dots, \lambda_p)) = \prod_{i<j} (\lambda_i - \lambda_j) \quad (41)$$

Et on peut réécrire  $p_\lambda$  de (38) comme suivant.

$$p_\lambda(\lambda_1, \dots, \lambda_p) = (p! K_{p,q})^{-1} \det(\mathbf{D}(\lambda_1, \dots, \lambda_p))^2 \prod_i \lambda_i^{p-q} e^{-\lambda_i} \quad (42)$$

Toute la complexité du calcul réside dans le terme  $\det(\mathbf{D}(\lambda_1, \dots, \lambda_p))$ . Avec les opérations linéaires sur les lignes, on peut transformer  $\mathbf{D}(\lambda_1, \dots, \lambda_p)$  en  $\tilde{\mathbf{D}}(\lambda_1, \dots, \lambda_p)$  dont le déterminant est proportionnel à  $\det(\mathbf{D}(\lambda_1, \dots, \lambda_p))$ .

$$\tilde{\mathbf{D}}(\lambda_1, \dots, \lambda_p) = \begin{bmatrix} \varphi_1(\lambda_1) & \dots & \varphi_1(\lambda_p) \\ \vdots & & \vdots \\ \varphi_p(\lambda_1) & \dots & \varphi_p(\lambda_p) \end{bmatrix} \Rightarrow \det(\tilde{\mathbf{D}}(\lambda_1, \dots, \lambda_p)) = K \det(\mathbf{D}(\lambda_1, \dots, \lambda_p)) \quad (43)$$

où  $\varphi_1, \dots, \varphi_p$  sont le résultat de l'application de la procédure d'orthogonalisation de Gram-Schmidt à la séquence  $1, \lambda, \lambda^2, \dots, \lambda^{p-1}$  dans l'espace des fonctions réelles avec le produit scalaire suivant.

$$\langle f, g \rangle = \int_0^\infty f(\lambda) g(\lambda) \lambda^{q-p} e^{-\lambda} d\lambda \quad (44)$$

Suivant la procédure de Gram-Schmidt les fonctions  $\varphi_1, \dots, \varphi_p$  seront orthonormales selon le produit scalaire (44) et on verra que chaque  $\varphi_k(\lambda)$  est une polynôme de l'ordre  $(k-1)$  par rapport à  $\lambda$ .

$$\langle \varphi_i, \varphi_j \rangle = \int_0^\infty \varphi_i(\lambda) \varphi_j(\lambda) \lambda^{q-p} e^{-\lambda} d\lambda = \sigma_{i,j} \quad (45)$$

Donc, en utilisant la définition du déterminant, le déterminant de  $\tilde{\mathbf{D}}(\lambda_1, \dots, \lambda_p)$  se développe suivant :

$$\det(\tilde{\mathbf{D}}(\lambda_1, \dots, \lambda_p)) = \sum_{\alpha} (-1)^{\text{per}(\alpha)} \prod_i \tilde{D}_{\alpha_i, i} = \sum_{\alpha} (-1)^{\text{per}(\alpha)} \prod_i \varphi_{\alpha_i}(\lambda_i) \quad (46)$$

où la sommation s'effectue sur toutes les permutations possibles de  $\{1, 2, \dots, p\}$  (il en existe  $p!$ ), et  $\text{per}(\alpha)$  est 0 ou 1, cela dépend du signe de la permutation (paire ou impaire). L'avantage de cette démonstration du déterminant de  $\mathbf{D}$  (46), par rapport à celle de (41) est le fait que  $p_{\lambda}$  dans (38) devient séparable sur  $\lambda_1, \dots, \lambda_p$  pour chaque permutation. Donc l'expression (42) peut encore se mettre sous la forme suivante.

$$p_{\lambda}(\lambda_1, \dots, \lambda_p) = C_{p,q} \sum_{\alpha, \beta} (-1)^{\text{per}(\alpha) + \text{per}(\beta)} \prod_i \varphi_{\alpha_i}(\lambda_i) \varphi_{\beta_i}(\lambda_i) \lambda_i^{p-q} e^{-\lambda_i} \quad (47)$$

où  $C_{p,q}$  est une constante, plus précisément  $C_{p,q} = (p! K_{p,q} K)^{-1}$ . En faisant l'intégration sur  $\lambda_2, \dots, \lambda_p$  et selon l'orthonormalité des fonctions  $\varphi_1, \dots, \varphi_p$  ((45)), on arrive à :

$$\begin{aligned} p_{\lambda_1}(\lambda_1) &= C_{p,q} \sum_{\alpha, \beta} (-1)^{\text{per}(\alpha) + \text{per}(\beta)} \varphi_{\alpha_1}(\lambda_1) \varphi_{\beta_1}(\lambda_1) \lambda_1^{q-p} e^{-\lambda_1} \prod_{i \geq 2} \delta_{\alpha_i, \beta_i} \\ &= C_{p,q} \sum_{\alpha} (-1)^{2\text{per}(\alpha)} \varphi_{\alpha_1}(\lambda_1) \varphi_{\alpha_1}(\lambda_1) \lambda_1^{q-p} e^{-\lambda_1} \\ &= C_{p,q} \sum_{\alpha} \varphi_{\alpha_1}(\lambda_1)^2 \lambda_1^{q-p} e^{-\lambda_1} \end{aligned} \quad (48)$$

$$\begin{aligned} &= C_{p,q} \cdot (p-1)! \sum_{i=1}^p \varphi_i(\lambda_1)^2 \lambda_1^{q-p} e^{-\lambda_1} \\ &= \frac{1}{p} \sum_{i=1}^p \varphi_i(\lambda_1)^2 \lambda_1^{q-p} e^{-\lambda_1} \end{aligned}$$

La deuxième égalité vient du fait que le terme  $\prod_{i \geq 2} \delta_{\alpha_i, \beta_i}$  est non nul, si et seulement si  $\alpha_i = \beta_i$ ,  $\forall i \geq 2$  et donc  $\alpha_i = \beta_i$ ,  $\forall i$ , c'est à dire à condition d'avoir deux permutations identiques  $\alpha = \beta$ . La quatrième égalité vient du fait que pour  $\alpha$  il existe  $p!$  combinaisons possibles mais pour  $\alpha_1$  il n'y en a que  $p$ , donc pour chaque de  $p$  combinaisons de  $\alpha_1$  nous avons  $(p-1)!$  combinaison possibles. La dernière égalité se déduit de l'orthonormalité des

fonctions  $\varphi_1, \dots, \varphi_p$ ,  $\int_0^\infty \varphi_i(\lambda_1)^2 \lambda_1^{q-p} e^{-\lambda_1} d\lambda_1 = 1$  et  $C_{p,q}$  qui satisfait à  $\int_0^\infty p_{\lambda_1}(\lambda_1) d\lambda_1 = 1$  doit donc être égale à  $1/p!$ . Donc pour le calcul de  $p_{\lambda_1}(\lambda_1)$  (48) et la capacité (39), Il nous reste seulement à définir les fonctions  $\varphi_1, \dots, \varphi_p$ . Vu la procédure l'orthonormalisation Gram-Schmidt sur la séquence  $1, \lambda, \lambda^2, \dots, \lambda^{p-1}$  (45), on peut exprimer ces fonctions sous la forme :

$$\begin{cases} \varphi_{k+1}(\lambda) = \left[ \frac{k!}{(k+q-p)!} \right]^{1/2} L_k^{q-p}(\lambda) \\ L_k^{q-p}(x) = \frac{1}{k!} e^x x^{p-q} \frac{d^k}{dx^k} (e^{-x} x^{q-p+k}) \end{cases}, \quad k = 0, 1, \dots, p-1 \quad (49)$$

où  $L_k^{q-p}(x)$  est le polynôme Laguerre associé d'ordre  $k$  (voir [7., §8.90,8.97]). Tout cette procédure de calcul de la capacité peut être récapitulée dans le théorème suivant.

**THEOREME 2.** La capacité du canal avec  $n$  émetteurs et  $m$  récepteurs sous la contrainte de puissance  $P$  est égale à :

$$C = \int_0^\infty \log(1 + P\lambda/n) \sum_{k=0}^{m-1} \frac{k!}{(k+q-p)!} [L_k^{q-p}(\lambda)]^2 \lambda^{q-p} e^{-\lambda} d\lambda \quad (50)$$

où  $p = \min\{m, n\}$  et  $q = \max\{m, n\}$ , et  $L_k^{q-p}$  sont les polynômes Laguerre associés.

**Exemple 3.** Considérons  $n = 1$ , dans ce cas-là  $p = 1$  et  $q = m$ . Etant donné que  $L_0^{q-p}(\lambda) = 1$ , l'application de (50) donne l'expression suivante pour la capacité.

$$C = \frac{1}{\Gamma(m)} \int_0^\infty \log(1 + P\lambda) \lambda^{m-1} e^{-\lambda} d\lambda \quad (51)$$

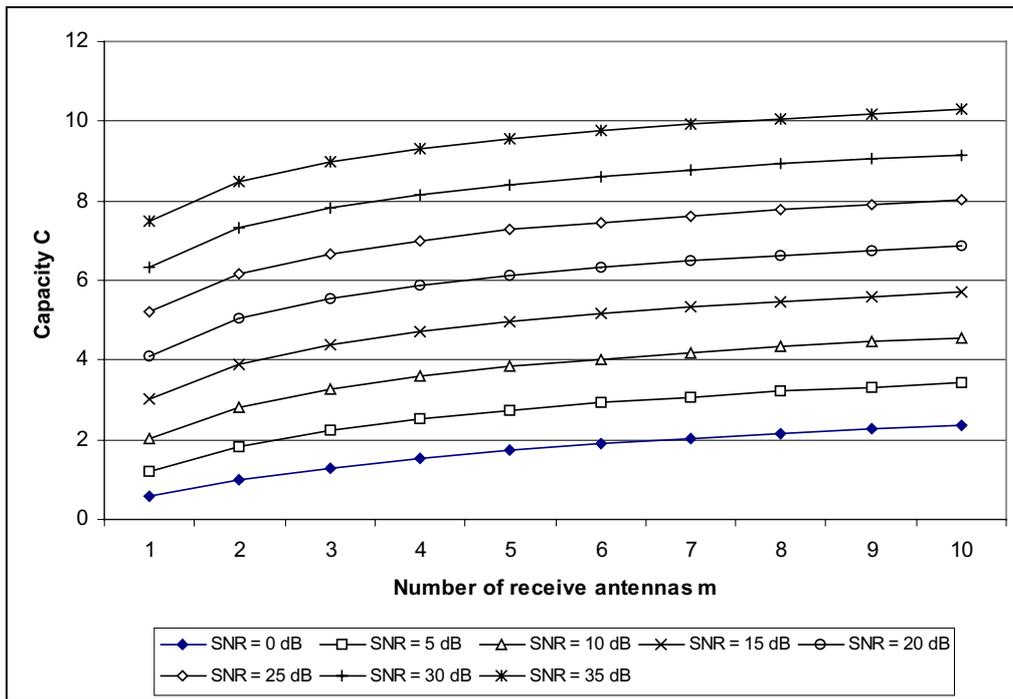
où  $\Gamma(m)$  est la fonction Gamma ( $\Gamma(m) = (m-1)!$ ). Les valeurs de cette intégrale sont représentées dans le tableau 1 ainsi que dur la courbe de la figure 5, pour  $1 \leq m \leq 10$  et  $P$  variant de 0dB à 35dB en incrément de 5dB. Il faut noter que pour les valeurs importantes de  $m$ , la capacité asymptotique est en  $\log(1 + Pm)$ .

$\frac{P}{m}$	0dB	5dB	10dB	15dB	20dB	25dB	30dB	35dB
1	0.5963	1.1894	2.0146	3.0015	4.0785	5.1988	6.3379	7.4845
2	1.0000	1.8133	2.8132	3.9066	5.0377	6.1824	7.3315	8.4822
3	1.2982	2.2146	3.2732	4.3922	5.5329	6.6808	7.8310	8.9820
4	1.5321	2.5057	3.5913	4.7204	5.8646	7.0136	8.1642	9.3153
5	1.7236	2.7327	3.8333	4.9679	6.1138	7.2634	8.4141	9.5652
6	1.8853	2.9183	4.0285	5.1663	6.3133	7.4632	8.6141	9.7652
7	2.0250	3.0752	4.1919	5.3319	6.4796	7.6298	8.7807	9.9319

8	2.1479	3.2110	4.3324	5.4740	6.6222	7.7726	8.9235	10.075
9	2.2576	3.3306	4.4556	5.5985	6.7471	7.8975	9.0485	10.200
10	2.3565	3.4375	4.5654	5.7091	6.8580	8.0086	9.1596	10.311

**Tableau 1 : Valeurs de l'intégrale (51)**

La capacité d'un canal fading simple émetteur, multi récepteur en nats. Le gain du trajet entre l'émetteur et chaque récepteur a une phase uniforme et une amplitude de Rayleigh dont l'espérance du carré est égale à 1. Ces gains sont indépendants. Le nombre de récepteurs est égale à  $m$  et  $P$  est le SNR.



**Figure 5 : Capacité en nats en fonction de  $m$  pour  $n = 1$  et différentes valeurs de  $P$ .**

**Exemple 4.** Considérons cette fois-ci  $m = 1$ . Comme exemple précédente ( $p = 1$  et  $q = n$ ) en appliquant (50), on trouve l'expression suivante pour la capacité.

$$C = \frac{1}{\Gamma(n)} \int_0^\infty \log(1 + P\lambda/n) \lambda^{n-1} e^{-\lambda} d\lambda \quad (52)$$

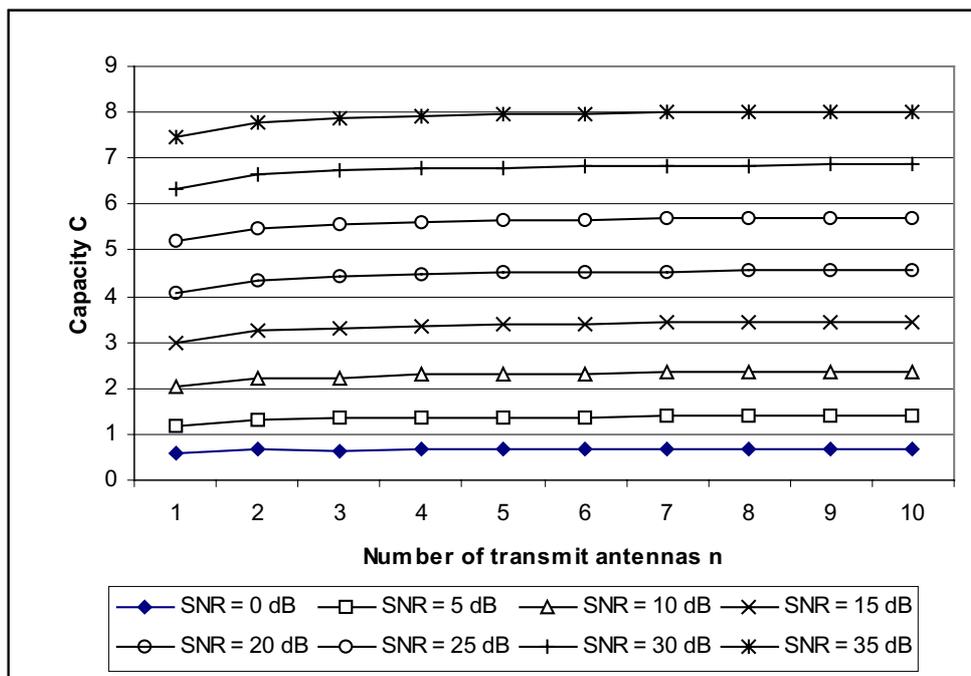
Comme on avait déjà noté dans (33), la capacité s'approche de  $\log(1 + P)$  quand  $n$  devient très grand. Les valeurs de la capacité de (52) sont représentées dans le tableau 2 ainsi que sur la courbe de la figure 6.

$P$	0dB	5dB	10dB	15dB	20dB	25dB	30dB	35dB
-----	-----	-----	------	------	------	------	------	------

$n$								
1	0.5963	1.1894	2.0146	3.0015	4.0785	5.1988	6.3379	7.4845
2	0.6387	1.2947	2.1947	3.2411	4.3540	5.4923	6.6394	7.7893
3	0.6552	1.3354	2.2608	3.3236	4.4441	5.5854	6.7334	7.8837
4	0.6640	1.3570	2.2947	3.3646	4.4882	5.6305	6.7789	7.9293
5	0.6695	1.3702	2.3152	3.3891	4.5142	5.6571	6.8057	7.9561
6	0.6733	1.3793	2.3289	3.4053	4.5314	5.6746	6.8233	7.9738
7	0.6760	1.3858	2.3388	3.4169	4.5436	5.6870	6.8358	7.9863
8	0.6781	1.3907	2.3462	3.4255	4.5527	5.6963	6.8451	7.9956
9	0.6797	1.3946	2.3519	3.4322	4.5598	5.7034	6.8523	8.0028
10	0.6810	1.3977	2.3565	3.4375	4.5654	5.7091	6.8580	8.0086

**Tableau 2 : Valeurs de l'intégrale (52)**

La capacité d'un canal fading muti émetteur, simple récepteur en nats. Le gain du trajet entre chaque émetteur et le récepteur a une phase uniforme et une amplitude de Rayleigh dont l'espérance du carré est égale à 1. Ces gains sont indépendants. Le nombre d'émetteurs est égale à  $n$  et  $P$  est le rapport signal à bruit.



**Figure 6 : Capacité en nat en fonction de  $n$  pour  $m = 1$  et différentes valeurs de  $P$**

**Exemple 5.** Considérons  $m = n$ , donc  $p = q = n$ , en appliquant (50) la capacité s'écrit comme suit.

$$C = \int_0^{\infty} \log(1 + P\lambda/n) \sum_{k=0}^{m-1} L_k(\lambda)^2 e^{-\lambda} d\lambda \quad (53)$$

où  $L_k(\lambda) = L_k^0(\lambda)$  est polynôme la guerre de l'ordre  $k$ . La Figure 7 illustre la capacité pour différentes valeurs de  $n$  et  $P$ . D'après la figure, on aperçoit bien que la fonction linéaire de  $n$  peut être une bonne approximation pour la capacité. Pour prouver cela mathématiquement on revient à l'expression (39). En utilisant la distribution empirique comme une approche de la

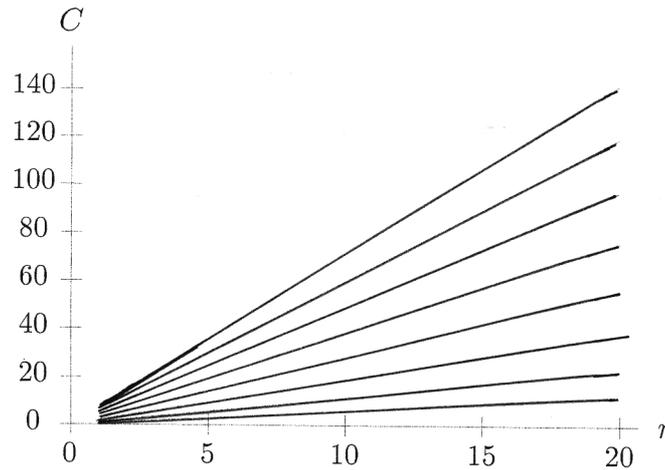
distribution cumulative (CDF : Cumulative Distribution Function), on peut réécrire (39) comme suivant.

$$\begin{aligned}
 C &= pE\{\log(1+(P/n)\lambda)\} = p \int_0^\infty \log\left(1+\frac{P}{n}\lambda\right) p_\lambda(\lambda) d\lambda \\
 &= p \int_0^\infty \log\left(1+\frac{P \cdot p}{n} \nu\right) p_\nu(\nu) d\nu = p \int_0^\infty \log\left(1+\frac{P \cdot p}{n} \nu\right) dF^{\frac{1}{p}W}(\nu)
 \end{aligned} \tag{54}$$

où  $F^A(x)$  est la distribution empirique des valeurs propres d'une matrice Hermitien  $A_{p \times p}$ , définie comme suivant.

$$F^A(x) = \frac{\text{le nombre des valeurs propres de } A \text{ inférieur à } x}{p}$$

On note que si  $\lambda$  est la valeur propre de  $W$ ,  $\nu = \frac{\lambda}{p}$  sera la valeur propre de la matrice  $\frac{W}{p}$ .



**Figure 7 : Capacité en fonction de  $n$  pour  $m = n$  et différentes valeurs de  $P$**

Un résultat très général de la théorie des matrices aléatoire (voir [8]) dit que pour  $W$  définie comme ci-dessus, quand  $q = \max\{m, n\}$  et  $p = \min\{m, n\}$  s'augmentent et le rapport  $q/p$  s'approche d'une limite  $\tau \geq 1$ , on aura :

$$\frac{dF^{\frac{1}{p}W}(\nu)}{d\nu} \rightarrow \begin{cases} \frac{1}{2\pi} \sqrt{\left(\frac{\nu_+}{\nu} - 1\right)\left(1 - \frac{\nu_-}{\nu}\right)} & \text{pour } \nu \in [\nu_-, \nu_+] \\ 0 & \text{sinon,} \end{cases} \tag{55}$$

où  $\nu_\pm = (\sqrt{\tau} \pm 1)^2$ . Donc, dans la limite des valeurs importantes de  $m$  et  $n$ , l'expression (54) se met sous la forme suivante.

$$C = \frac{p}{2\pi} \int_{\nu_-}^{\nu_+} \log\left(1+\frac{Pp}{n}\nu\right) \sqrt{\left(\frac{\nu_+}{\nu} - 1\right)\left(1 - \frac{\nu_-}{\nu}\right)} d\nu \tag{56}$$

Pour le cas considéré dans l'exemple 5,  $p = q = m = n$ , donc  $v_- = 0$ ,  $v_+ = 4$ , et en remplaçant ces valeurs, (56) s'écrit comme suivant.

$$C \cong p \int_0^4 \log(1 + Pv) \frac{1}{\pi} \sqrt{\frac{1}{v} - 4} dv \quad (57)$$

D'après cette dernière  $C$  sera linéaire par rapport à  $p$ , comme on avait déjà remarqué dans la Figure 7.

Maintenant on présente quelques remarques observés dans les exemples précédentes.

*Remarque 2.* Le résultat de la théorie des matrices aléatoires utilisé dans l'exemple 5 s'applique en général pour les matrices aléatoires qui ne sont pas nécessairement gaussiennes. Donc l'expression (57) est toujours valable, si et seulement si les éléments de  $\mathbf{H}$  sont i.i.d. avec la variance 1.

*Remarque 3.* La propriété de réciprocity que nous avons observé pour  $\mathbf{H}$  déterministe (on a la même capacité pour le canal  $\mathbf{H}^h$  que pour  $\mathbf{H}$ ) n'est pas valable pour  $\mathbf{H}$  aléatoire. Il suffit juste de comparer l'exemple 3 et 4 où les matrices  $\mathbf{H}$  correspondantes sont transposées l'une de l'autre. Dans l'exemple 3, la capacité augmente sans limite quand  $m$  augmente, tandis que dans l'exemple 4 la capacité est limitée.

Cependant, la permutation de  $m$  et  $n$ , n'a aucun effet sur la matrice  $\mathbf{W}$ , et la capacité dépend seulement de  $P/n$  et des valeurs propres de  $\mathbf{W}$ . Donc si  $C(m, n, P)$  dénote la capacité d'un canal avec  $m$  récepteurs,  $n$  émetteurs et une puissance totale  $P$ , on peut dire :

$$C(a, b, Pb) = C(b, a, Pa)$$

*Remarque 4.* Dans la procédure de calcul de la capacité du théorème, nous avons obtenu la densité d'une des valeurs propres non-ordonnée de la matrice complexe Wishart  $\mathbf{W}$ . On est également capable de trouver la densité jointe de n'importe quel nombre ( $k$ ) des valeurs propres non-ordonnées de  $\mathbf{W}$  comme suivant.

$$p_{\lambda_1, \dots, \lambda_k}(\lambda_1, \dots, \lambda_k) = \frac{(p-k)!}{p!} \det(D_k(\lambda_1, \dots, \lambda_k)^h D_k(\lambda_1, \dots, \lambda_k)) \prod_{i=1}^k \lambda_i^{q-p} e^{-\lambda_i}$$

où

$$D_k(\lambda_1, \dots, \lambda_k) = \begin{bmatrix} \varphi_1(\lambda_1) & \cdots & \varphi_1(\lambda_k) \\ \vdots & & \vdots \\ \varphi_p(\lambda_1) & \cdots & \varphi_p(\lambda_k) \end{bmatrix}$$

### 2-5-3 « Error Exponents » :

Comme pour le cas  $\mathbf{H}$  déterministe, nous pouvons calculer l'« error exponent » dans le cas du canal à fading. Dans ce cas-là, il y a un calcul d'espérance sur  $\mathbf{H}$  qui s'ajoute à l'expression (24).

$$E_0(\rho, q_x) = -\log \iint \left[ \int q_x(\mathbf{x}) p(\mathbf{y}, \mathbf{H}|\mathbf{x})^{V(1+\rho)} d\mathbf{x} \right]^{1+\rho} d\mathbf{y} d\mathbf{H} \quad (58)$$

Comme  $\mathbf{H}$  est indépendante de  $\mathbf{x}$ ,  $p(\mathbf{y}, \mathbf{H}|\mathbf{x}) = p_{\mathbf{H}}(\mathbf{H})p(\mathbf{y}|\mathbf{x}, \mathbf{H})$  et (58) peut se mettre sous la forme suivante.

$$E_0(\rho, q_x) = -\log \left( E_{\mathbf{H}} \left\{ \int \left[ \int q_x(\mathbf{x}) p(\mathbf{y}|\mathbf{x}, \mathbf{H})^{y(1+\rho)} d\mathbf{x} \right]^{1+\rho} d\mathbf{y} \right\} \right) \quad (59)$$

On note que :

$$p(\mathbf{y}|\mathbf{x}, \mathbf{H}) = \det(\pi \mathbf{I}_m)^{-1} \exp \left( -(\mathbf{y} - \mathbf{H}\mathbf{x})^\dagger (\mathbf{y} - \mathbf{H}\mathbf{x}) \right)$$

Et pour  $q_x = f_Q$ , la distribution gaussienne avec la matrice de covariance  $\mathbf{Q}$ , on peut utiliser les résultats du cas  $\mathbf{H}$  déterministe pour conclure :

$$E_0(\rho, q_x) = -\log \left( E_{\mathbf{H}} \left\{ \det \left( \mathbf{I}_m + (1+\rho)^{-1} \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger \right)^{-\rho} \right\} \right) \quad (60)$$

Tenant compte que  $A \rightarrow \det(A)^{-\rho}$  est une fonction convexe, le raisonnement précédemment utilisé pour montrer que  $\mathbf{Q} = (P/n)\mathbf{I}_n$  maximise l'information mutuelle, montre que  $\mathbf{Q}$  maximise également  $E_0$ , et on obtient finalement.

$$E_0(\rho, q_x) = -\log \left( E_{\mathbf{H}} \left\{ \det \left( \mathbf{I}_m + \frac{P}{n(1+\rho)} \mathbf{H}\mathbf{H}^\dagger \right)^{-\rho} \right\} \right) \quad (61)$$

Pour calculer efficacement  $E_0$ , on devrait représenter la densité des valeurs propres Wishart comme un déterminant Vandermonde (exactement comme dans la section précédente) et orthonormaliser les monômes  $1, \lambda, \lambda^2, \dots, \lambda^{p-1}$ , avec le produit scalaire ci-dessous (toujours avec la technique d'orthonormalisation Gram-Schmidt).

$$\langle f, g \rangle = \int_0^\infty f(\lambda)g(\lambda)\lambda^{q-p} \left( 1 + \frac{P}{n(1+\rho)}\lambda \right)^{-\rho} e^{-\lambda} d\lambda$$

où le facteur multiplicatif dans l'orthonormalisation est la valeur de l'espérance dans (61). Comme pour le cas  $\mathbf{H}$  déterministe, le choix de la distribution gaussienne est sous-optimale, mais ce choix aboutit aux expressions les plus simples.

## **2-6 Canaux non Ergodiques**

Nous avons remarqué au début de la section précédente que l'information mutuelle peut permettre de calculer la capacité quand le canal est sans mémoire (memoryless) et dans le cas où pour chaque utilisation du canal nous avons une réalisation indépendante de  $\mathbf{H}$ . Cela n'est pas le seul cas où l'information mutuelle représente la capacité de canal. En particulier si le processus qui génère  $\mathbf{H}$  est ergodique, on peut obtenir les relations arbitrairement proches de l'information mutuelle.

Au contraire pour le cas où  $\mathbf{H}$  serait choisi aléatoirement au début de chaque temps et reste fixe pour tous les utilisateurs du canal, l'information mutuelle n'est pas en générale égale à la

capacité de canal. Dans cette section nous allons nous concentrer sur un tel cas quand les entrées de  $\mathbf{H}$  sont i.i.d. GCCS avec la moyenne zéro et  $E\{|h_{ij}|^2\}=1$ , c'est à dire la même distribution que nous avons analysée dans la section précédente.

### 2-6-1 Capacité

Dans le cas décrit ci-dessus la capacité au sens de Shannon du canal est nulle. De toute façon il existe une probabilité non-nulle que  $\mathbf{H}$  réalisée soit incapable de la supporter, peu importe la longueur de notre code. D'autre part, on peut parler d'un compromis entre *la probabilité outage* et *le taux de transmission supportable*. Avec un taux donné  $R$  et une puissance  $P$ , on peut trouver  $P_{out}(R, P)$  telle que pour tous les taux inférieurs à  $R$  et tous les  $\delta$ , il existe un code satisfaisant la contrainte de puissance  $P$  pour lequel la probabilité d'erreur est inférieure à  $\delta$ :

$$P_{out}(R, P) = \inf_{\substack{\mathbf{Q}: \mathbf{Q} \succeq 0 \\ \text{tr}(\mathbf{Q}) \leq P}} \mathbb{P}(\Psi(\mathbf{Q}, \mathbf{H}) < R) \quad (62)$$

où

$$\Psi(\mathbf{Q}, \mathbf{H}) = \log(\det(I_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger))$$

Cette approche est utilisée dans [9] dans un problème similaire.

Dans cette section comme dans la précédente, nous supposons que les entrées de  $\mathbf{H}$  sont gaussiennes indépendantes avec la moyenne zéro et les parties réelle et imaginaire indépendantes avec la variance  $1/2$ .

*Exemple 6.* Considérons  $n = 1$ . Dans ce cas il est clair que  $\mathbf{Q} = P$  est optimale. La probabilité outage sera donc.

$$P_{out}(R, P) = \mathbb{P}(\log(\det(I_m + \mathbf{H}P\mathbf{H}^\dagger)) < R) = \mathbb{P}(\log(\det(1 + P\mathbf{H}^\dagger\mathbf{H})) < R) \quad (63)$$

Comme  $\mathbf{H}^\dagger\mathbf{H}$  est une variable aléatoire  $X^2$  de degré de liberté  $2m$  et la moyenne  $m$ , on peut calculer la probabilité d'outage comme suivante.

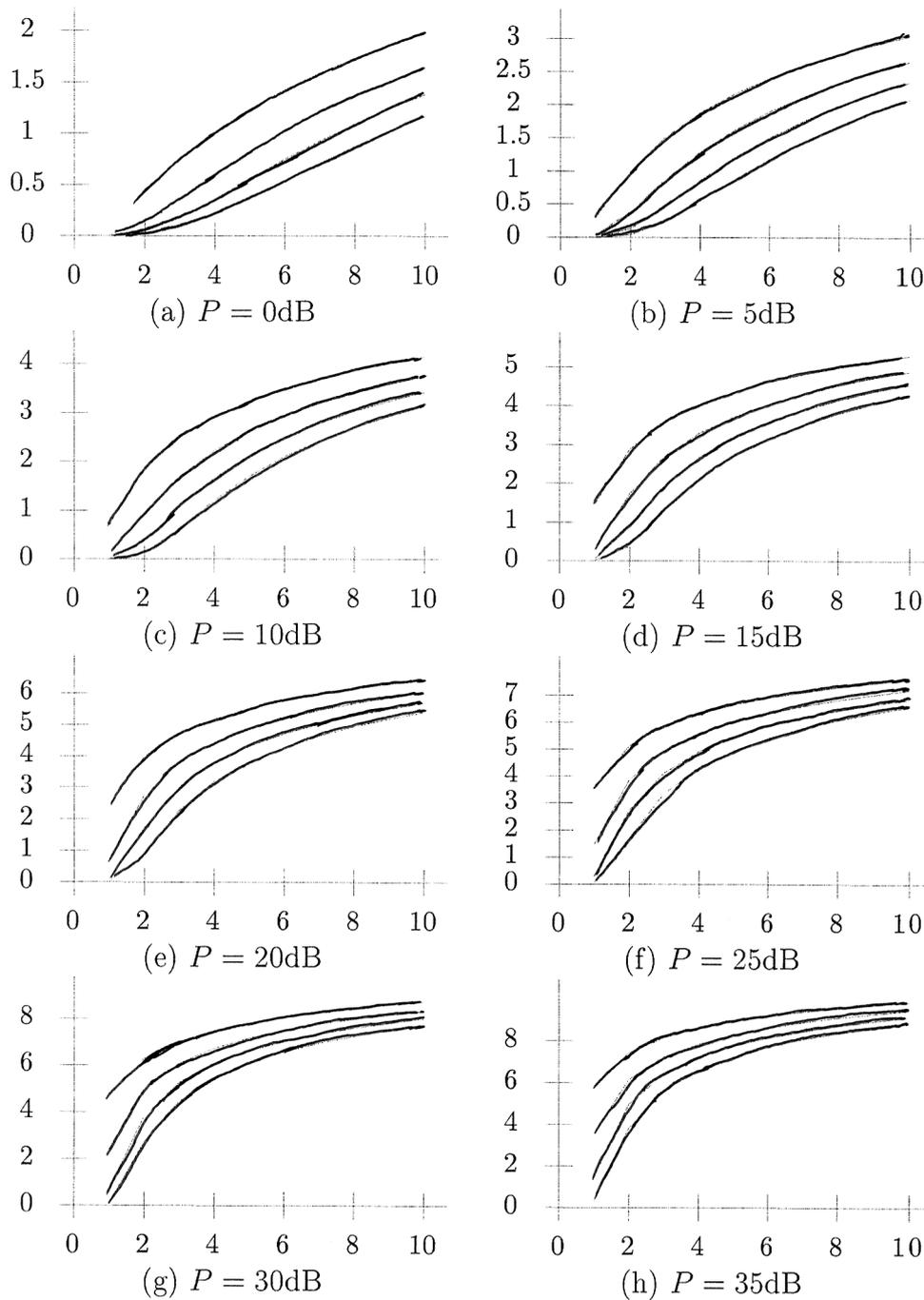
$$P_{out}(R, P) = \frac{\gamma(m, (e^R - 1)/P)}{\Gamma(m)} \quad (64)$$

où  $\gamma(a, x) = \int_0^x u^{a-1} e^{-u} du$  est la fonction incomplète de gamma. Supposons que  $\psi(P, \varepsilon)$  est la valeur de  $R$  qui satisfait

$$P(\Psi(P, \mathbf{H}) \leq R) = \varepsilon \quad (65)$$

La Figure 8 illustre  $\psi(P, \varepsilon)$  comme une fonction de  $m$  pour les différentes valeurs de  $\varepsilon$  et  $P$ . Rappelons-nous que  $\psi(P, \varepsilon)$  est le taux maximum pour lequel la probabilité d'outage est inférieur à  $\varepsilon$ . Chaque ensemble des courbes correspond à la puissance  $P$  indiquée dessous.

Dans chaque ensemble, les courbes correspondent, dans l'ordre descendant, à  $\varepsilon = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ .



**Figure 8 : La  $\varepsilon$ -capacité pour  $n = 1$  comme défini par (65)  $\psi(P, \varepsilon)$  en fonction de  $m$  pour les différentes valeurs de  $P$  et  $\varepsilon$**

On note d'après le Lemme 5 que la distribution de  $HU$  est la même que de  $H$  pour  $U$  unitaire. Nous pourrions donc conclure que :

$$\Psi(UQU^\dagger, H)$$

a la même distribution que  $\Psi(Q, H)$ . En choisissant  $U$  pour diagonaliser  $Q$ , on se concentre sur  $Q$  diagonale et la symétrie dans ce problème suggère la conjecture suivante.

**CONJECTURE.** La matrice optimale  $Q$  a la forme suivante.

$$\frac{P}{k} \text{diag}(\underbrace{1, \dots, 1}_{k \text{ uns}}, \underbrace{0, \dots, 0}_{n-k \text{ zéros}})$$

pour  $k = 1, \dots, n$ . La valeur de  $k$  dépend du taux : plus le taux de transmission est important (la probabilité d'outage est plus grande), plus  $k$  est petit.

Comme la puissance est identiquement partagée entre la majorité des émetteurs, l'espérance de  $\Psi$  augmente mais les queues de ces distributions se détériorent plus rapidement. Afin de minimiser la probabilité d'outage, il faut maximiser la densité de probabilité de  $\Psi$ . Si on veut avoir les taux plus grands que l'espérance de  $\Psi$ , il faudra utiliser un petit nombre d'émetteurs pour avoir l'avantage de la détérioration lente des queues de la distribution  $\Psi$ . Bien sur la probabilité d'outage correspondant sera encore large.

*Exemple 7.* Considérons  $m = 1$ . D'après la conjoncture ci-dessus, il suffit simplement de calculer  $P(\Psi((P/n)\mathbf{I}_n, H) \leq R)$  pour toutes les valeurs de  $n$ . Si le nombre actuel d'émetteurs est appelé  $\tau$ , la probabilité d'outage sera donc le minimum des probabilités pour  $n = 1, \dots, \tau$ . Comme dans l'exemple 6, on marque que  $HH$  a une statistique X2 avec  $2n$  degrés de liberté et une moyenne  $n$ . Donc on aura :

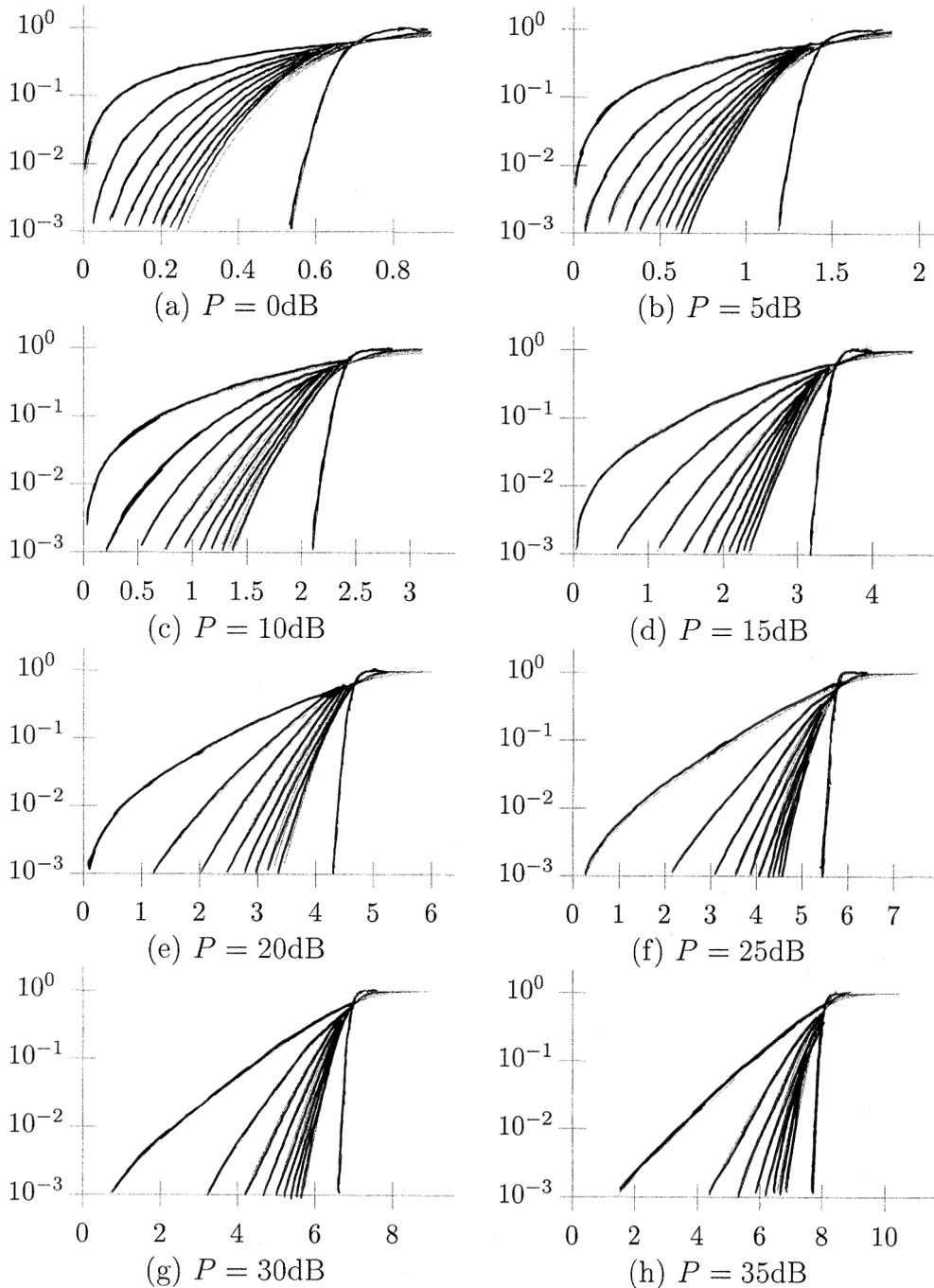
$$P(\Psi((P/n)\mathbf{I}_n, H) \leq R) = \frac{\gamma(n, n(e^R - 1)/P)}{\Gamma(n)} \quad (66)$$

La figure 9 illustre cette distribution pour les différentes valeurs de  $n$  et  $P$ . Chaque ensemble des courbes correspond à la puissance  $P$  indiquée dessous et dans chaque ensemble les courbes correspondent, dans l'ordre, de gauche à droite, à  $n = 1, 2, \dots, 10$  et 100. Il est clair, selon la figure, que les valeurs larges de  $n$  ont une meilleure performance pour  $R$  petit et les petites valeurs de  $n$  donnent une meilleure performance pour  $R$  large. Comme dans l'exemple 5, on suppose que  $\psi(P, \varepsilon)$  est la valeur de  $R$  qui satisfait l'expression suivante.

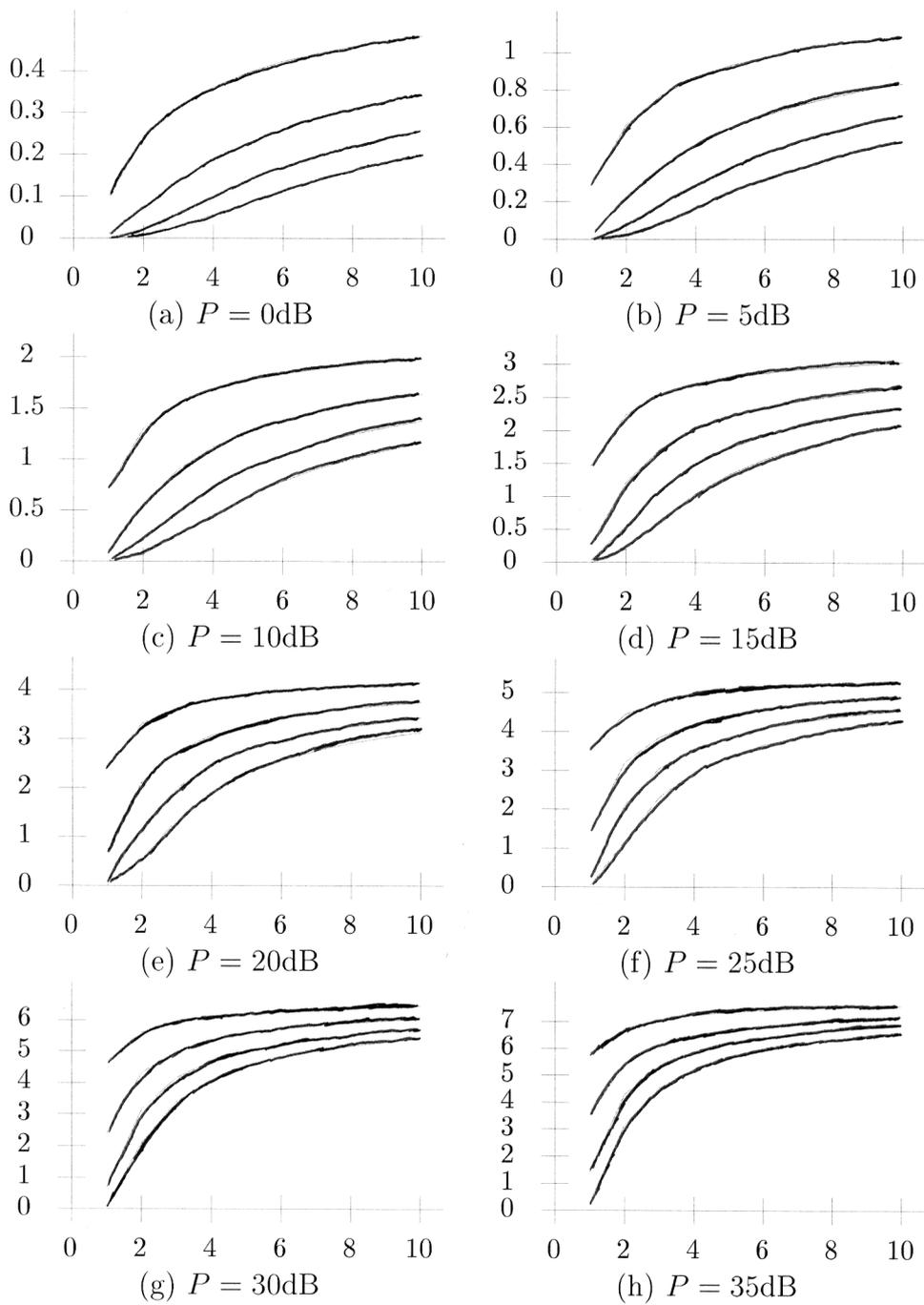
$$P(\Psi((P/n)\mathbf{I}_n, H) \leq R) = \varepsilon \quad (67)$$

La Figure 10 montre  $\psi(P, \varepsilon)$  en fonction de  $n$ , pour les différentes valeurs de  $P$  et  $\varepsilon$ . Rappelons que  $\psi(P, \varepsilon)$  est la taux maximum pour lequel la probabilité d'outage reste inférieur à  $\varepsilon$ . Chaque ensemble des courbes correspond à la puissance  $P$  indiquée dessous et dans chaque ensemble, les courbes correspondent, dans l'ordre descendant, à  $\varepsilon = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$ . On remarque d'après la figure que pour les valeur petites de  $\varepsilon$ ,

considérées dans la figure, l'utilisation de tous les émetteurs disponibles est toujours meilleure que l'utilisation d'un sous ensemble.



**Figure 9 : Distribution de  $P(\Psi((P/n)I_n, \mathbf{H}) \leq R)$  pour  $m=1$  en fonction de  $R$  pour les différentes valeurs de  $P$  et  $n$ .**



**Figure 10 :  $\varepsilon$ -capacité (comme définie par (18)) pour  $m=1$  par rapport à  $n$  et pour les différentes valeurs de  $P$  et  $\varepsilon$ .**

## 2-7 Canaux Accès Multiple

On considère maintenant un nombre d'émetteurs égal à  $M$ , chacun avec  $n$  antennes d'émission et disposant chacun d'une puissance maximale  $P$ . Nous considérons également le cas d'un unique récepteur avec  $m$  antennes. Le signal reçu s'écrit

$$y = [\mathbf{H}_1 \cdots \mathbf{H}_M] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} + n \quad (68)$$

où  $x_m$  est le signal transmis par l'émetteur  $m$ ,  $n$  est un bruit Gaussien et  $\mathbf{H}_m$   $m = 1, \dots, M$  sont des matrices complexes  $m \times n$ . Nous supposons que le récepteur connaît tous les  $\mathbf{H}_m$ 's et que chacune de ces variables suit une loi complexe Gaussienne de moyenne nulle et de variance unité. On peut alors généraliser le cas monoutilisateur au contexte multiutilisateurs décrit ici. Un ensemble de  $M$  débits de transmission  $(R_1, \dots, R_M)$  peut être réalisé si :

$$\sum_{i=1}^M R_{[i]} \leq C(m, in, iP) \quad \text{pour tout } i = 1, \dots, M \quad (69)$$

où  $(R_{[1]}, \dots, R_{[M]})$  est l'ensemble des taux de transmission rangés par ordre décroissant et  $C(a, b, P)$  correspond à la capacité simple utilisateur entre l'émetteur  $a$  et le récepteur  $b$  avec la contrainte de puissance  $P$ .

## 2-8 Conclusion

Dans ce chapitre nous avons présenté les expressions de la capacité de canaux MIMO pour les différents types de matrice de transfert (déterministe, processus aléatoire ergodique, processus aléatoire non ergodique). On s'est aperçu que l'utilisation des antennes multiples permettait d'augmenter considérablement le débit sur les canaux à fading si les paramètres du canal peuvent être bien estimés au récepteur et si les gains de trajet entre les différentes paires d'antennes sont indépendants. La deuxième contrainte peut être réalisée avec une facilité relative. La première est plus difficile à réaliser, et elle doit être en fonction du contexte étudié. En effet l'évaluation des paramètres d'un canal de propagation MIMO est un exercice particulièrement difficile dans un contexte multiutilisateurs-multiantennes où les signaux interférents sont très nombreux.

## **2-9 Références**

- [1] E. Telatar, "Capacity of multi-antenna Gaussian channels," AT&T-Bell Labs Internal Tech. Memo., June 1995.
- [2] W. C. Jakes, Ed., *Microwave Mobile Communication*. New York : IEEE Press, 1974.
- [3] G. W. Wornell and M. D. Trott, "Efficient Signal Processing Techniques for Exploiting Transmit Antenna Diversity on Fading Channels" *IEEE Trans. On Signal Process*, vol. 45, no. 1, pp. 191-205, Jan. 1997.
- [4] R. G. Gallager, *Information theory and Reliable Communication*. New York : John Wiley & Sons, 1968.
- [5] T. James, "Distribution of matrix variates and latent roots derived from normal samples," *Annals of Mathematical Statistics*, vol. 35, pp. 475-501, 1964.
- [6] Edelman, *Eigenvalues and Condition Numbers of Random Matrices*. Ph.D. thesis, Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 1989.
- [7] S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. New York: Academic Press, corrected and enlarged ed., 1980.
- [8] J. W. Silverstein, "Strong Convergence of the Empirical Distribution of Eigenvalues of Large Dimensional Random Matrices," *Journal of Multivariate Analysis*, vol. 55, pp. 331-339, 1995.
- [9] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. on Vehicular. Technology*, vol. 43, pp. 359-378, May 1994.
- [10] T. L. Marzetta and B. Hochwald, "Capacity of mobile multiple-antenna communication link in a Rayleigh flat-fading environment," *IEEE Trans. on Inform. Theory*, vol. 45, pp. 139-157, Jan. 1999.
- [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York : John Wiley & Sons, 1991.
- [12] G. J. Foschini and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 313-335, 1998.
- [13] A. Narula, M. D. Trott, and G. W. Wornell, "Information theoretic analysis of multiple-antenna transmission diversity," *IEEE Int. Symp. Inform. Theory and Appl.*, Victoria, Canada, Sept. 1996.
- [14] \_\_\_\_\_, "Performance Limits of Coded Diversity Methods for Transmitter Antenna Arrays," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2418-2433, Nov. 1999.



## Annexe A : Statistiques des vecteurs aléatoires complexes

Considérons le vecteur aléatoire complexe  $\mathbf{x}_{n \times 1} \in \mathbb{C}^n$

$$\mathbf{x} = \mathbf{x}_R + j \mathbf{x}_I$$

dont chaque partie réelle et imaginaire est également un vecteur aléatoire faisant partie de  $\mathfrak{R}^n$ . L'objectif est de trouver les statistiques du vecteur complexe  $\mathbf{x}$ , par rapport à celles des vecteurs  $\mathbf{x}_R$  et  $\mathbf{x}_I$ .

$$\boldsymbol{\mu}_x = \bar{\mathbf{x}} = E\{\mathbf{x}\} = E\{\mathbf{x}\} + jE\{\mathbf{x}\} \Rightarrow \boldsymbol{\mu}_x = \bar{\mathbf{x}} = \bar{\mathbf{x}}_R + j \bar{\mathbf{x}}_I \quad (\text{A-1-a})$$

$$\begin{aligned} \mathbf{Q} = \boldsymbol{\Sigma}_x &= E\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^\dagger\} \\ &= E\{[(\mathbf{x}_R + j \mathbf{x}_I) - (\bar{\mathbf{x}}_R + j \bar{\mathbf{x}}_I)][(\mathbf{x}_R^T - j \mathbf{x}_I^T) - (\bar{\mathbf{x}}_R^T + j \bar{\mathbf{x}}_I^T)]\} \\ &= E\{[(\mathbf{x}_R - \bar{\mathbf{x}}_R) + j(\mathbf{x}_I - \bar{\mathbf{x}}_I)][(\mathbf{x}_R^T - \bar{\mathbf{x}}_R^T) - j(\mathbf{x}_I^T - \bar{\mathbf{x}}_I^T)]\} \\ &= E\{(\mathbf{x}_R - \bar{\mathbf{x}}_R)(\mathbf{x}_R^T - \bar{\mathbf{x}}_R^T)\} + E\{(\mathbf{x}_I - \bar{\mathbf{x}}_I)(\mathbf{x}_I^T - \bar{\mathbf{x}}_I^T)\} \\ &\quad + j[E\{(\mathbf{x}_I - \bar{\mathbf{x}}_I)(\mathbf{x}_R^T - \bar{\mathbf{x}}_R^T)\} - E\{(\mathbf{x}_R - \bar{\mathbf{x}}_R)(\mathbf{x}_I^T - \bar{\mathbf{x}}_I^T)\}] \\ &\Rightarrow \mathbf{Q} = \boldsymbol{\Sigma}_x = (\boldsymbol{\sigma}_{x_R} + \boldsymbol{\sigma}_{x_I}) + j(\boldsymbol{\sigma}_{x_I x_R} - \boldsymbol{\sigma}_{x_R x_I}) \end{aligned} \quad (\text{A-1-b})$$

Où  $\boldsymbol{\sigma}_{x_R}$  et  $\boldsymbol{\sigma}_{x_I}$  sont respectivement les matrices d'auto-covariance des vecteurs  $\mathbf{x}_R$  et  $\mathbf{x}_I$  (matrices symétriques et non-négative définie) et  $\boldsymbol{\sigma}_{x_I x_R}$  et  $\boldsymbol{\sigma}_{x_R x_I}$  sont les matrices d'intercorrélations de ces vecteurs, en plus  $\boldsymbol{\sigma}_{x_I x_R} = \boldsymbol{\sigma}_{x_R x_I}^T$ . La matrice d'auto-covariance  $\mathbf{Q}$  sera également une matrice Hermitienne (symétrique en partie réelle et antisymétrique en partie imaginaire) et non-négative définie.

**Cas particuliers :**

$$\begin{aligned} 1. \mathbf{x} : \text{circulairement symétrique (CS)} &\Rightarrow \boldsymbol{\sigma}_{x_R} = \boldsymbol{\sigma}_{x_I} = \frac{1}{2} \text{Re}(\mathbf{Q}) \\ &\boldsymbol{\sigma}_{x_I x_R} = -\boldsymbol{\sigma}_{x_R x_I} = \frac{1}{2} \text{Im}(\mathbf{Q}) \\ &\boldsymbol{\sigma}_{x_I x_R} \text{ et } \boldsymbol{\sigma}_{x_R x_I} : \text{Antisymétrique} \end{aligned}$$

$$2. \mathbf{x}_R, \mathbf{x}_I \text{ indépendant} \Rightarrow \boldsymbol{\sigma}_{x_I x_R} = \boldsymbol{\sigma}_{x_R x_I} = \mathbf{0} \Rightarrow \mathbf{Q} = \boldsymbol{\Sigma}_x = \boldsymbol{\sigma}_{x_R} + \boldsymbol{\sigma}_{x_I}$$

$\mathbf{Q}$  : réelle et symétrique

$$3. \mathbf{x} : \text{circulairement symétrique et } \mathbf{x}_R, \mathbf{x}_I \text{ indépendant} : \Rightarrow \boldsymbol{\sigma}_{x_R} = \boldsymbol{\sigma}_{x_I} = \frac{1}{2} \mathbf{Q}$$

Pour  $n = 1$ , le vecteur  $x$  ne peut pas être CS, à moins que  $x_R$  et  $x_I$  soient indépendants.

Une autre façon de trouver les statistiques de vecteur complexe  $\mathbf{x}_{n \times 1} \in \mathbb{C}^n$ , en évitant les calculs complexes, est de définir le vecteur  $\hat{\mathbf{x}}_{2n \times 1} \in \mathbb{R}^{2n}$  qui contient la partie réelle et imaginaire du vecteur  $\mathbf{x}$ ,  $\hat{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix}$  et de se servir des statistiques de ce vecteur afin d'obtenir celles du vecteur  $\mathbf{x}$ .

$$\boldsymbol{\mu}_{\hat{\mathbf{x}}} = E\{\hat{\mathbf{x}}\} = \begin{bmatrix} \bar{\mathbf{x}}_R \\ \bar{\mathbf{x}}_I \end{bmatrix} = \hat{\boldsymbol{\mu}} \quad (\text{A-2-a})$$

$$\begin{aligned} \boldsymbol{\Sigma}_{\hat{\mathbf{x}}} &= E\left\{(\hat{\mathbf{x}} - E\{\hat{\mathbf{x}}\})(\hat{\mathbf{x}} - E\{\hat{\mathbf{x}}\})^T\right\} = E\left\{\begin{bmatrix} \mathbf{x}_R - \bar{\mathbf{x}}_R \\ \mathbf{x}_I - \bar{\mathbf{x}}_I \end{bmatrix} \begin{bmatrix} (\mathbf{x}_R - \bar{\mathbf{x}}_R)^T & (\mathbf{x}_I - \bar{\mathbf{x}}_I)^T \end{bmatrix}\right\} \\ &= \begin{bmatrix} E\{(\mathbf{x}_R - \bar{\mathbf{x}}_R)(\mathbf{x}_R^T - \bar{\mathbf{x}}_R^T)\} & E\{(\mathbf{x}_R - \bar{\mathbf{x}}_R)(\mathbf{x}_I^T - j\bar{\mathbf{x}}_I^T)\} \\ E\{(\mathbf{x}_I - \bar{\mathbf{x}}_I)(\mathbf{x}_R^T - j\bar{\mathbf{x}}_R^T)\} & E\{(\mathbf{x}_I - \bar{\mathbf{x}}_I)(\mathbf{x}_I^T - j\bar{\mathbf{x}}_I^T)\} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\sigma}_{\mathbf{x}_R} & \boldsymbol{\sigma}_{\mathbf{x}_R \mathbf{x}_I} \\ \boldsymbol{\sigma}_{\mathbf{x}_I \mathbf{x}_R} & \boldsymbol{\sigma}_{\mathbf{x}_I} \end{bmatrix}_{2n \times 2n} \end{aligned} \quad (\text{A-2-b})$$

En comparant (A-2-b) et (A-1-b), on peut trouver la matrice d'auto-covariance  $\mathbf{Q}$  à partir de  $\boldsymbol{\Sigma}_{\hat{\mathbf{x}}}$  qui est une matrice symétrique non-négative définie.

#### Cas particuliers :

$$1. \mathbf{x} : \text{circulairement symétrique} \quad \Rightarrow \quad \boldsymbol{\Sigma}_{\hat{\mathbf{x}}} = \frac{1}{2} \begin{bmatrix} \text{Re}(\mathbf{Q}) & -\text{Im}(\mathbf{Q}) \\ \text{Im}(\mathbf{Q}) & \text{Re}(\mathbf{Q}) \end{bmatrix} = \frac{1}{2} \hat{\mathbf{Q}}$$

où  $\hat{\mathbf{Q}}$  est la projection T (définie dans le lemme 1) de la matrice  $\mathbf{Q}$

$$2. \mathbf{x}_R, \mathbf{x}_I \text{ indépendant} \quad \Rightarrow \quad \boldsymbol{\Sigma}_{\hat{\mathbf{x}}} = \begin{bmatrix} \boldsymbol{\sigma}_{\mathbf{x}_R}^2 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma}_{\mathbf{x}_I}^2 \end{bmatrix}$$

$$3. \mathbf{x} : \text{circulairement symétrique et } \mathbf{x}_R, \mathbf{x}_I \text{ indépendant} : \quad \Rightarrow \quad \boldsymbol{\Sigma}_{\hat{\mathbf{x}}} = \frac{1}{2} \begin{bmatrix} \text{Re}(\mathbf{Q}) & \mathbf{0} \\ \mathbf{0} & \text{Re}(\mathbf{Q}) \end{bmatrix}$$

## **Annexe B : Preuve des propriétés de la projection**

$$z \rightarrow \hat{z}, \quad A \rightarrow \hat{A}$$

On va prouver les propriétés de la projection  $z \rightarrow \hat{z} = \begin{bmatrix} z_R \\ z_I \end{bmatrix}$  et  $A \rightarrow \hat{A} = \begin{bmatrix} A_R & -A_I \\ A_I & A_R \end{bmatrix}$ , présentées dans les équations (5). Les preuves des équations (5-b) et (5-f) sont évidentes.

$$(5-a) : C = AB = (A_R + jA_I)(B_R + jB_I) = \overbrace{(A_R B_R - A_I B_I)}^{C_R} + j \overbrace{(A_I B_R + A_R B_I)}^{C_I}$$

$$\hat{A}\hat{B} = \begin{bmatrix} A_R & -A_I \\ A_I & A_R \end{bmatrix} \begin{bmatrix} B_R & -B_I \\ B_I & B_R \end{bmatrix} = \begin{bmatrix} A_R B_R - A_I B_I & -(A_I B_R + A_R B_I) \\ A_I B_R + A_R B_I & A_R B_R - A_I B_I \end{bmatrix} = \begin{bmatrix} C_R & -C_I \\ C_I & C_R \end{bmatrix} = \hat{C}$$

$$(5-c) : C = A^\dagger = (A_R + jA_I)^\dagger = A_R^T - jA_I^T \Rightarrow C_R = A_R^T ; C_I = -A_I^T$$

$$\hat{A}^T = \begin{bmatrix} A_R & -A_I \\ A_I & A_R \end{bmatrix}^T = \begin{bmatrix} A_R^T & A_I^T \\ -A_I^T & A_R^T \end{bmatrix} = \begin{bmatrix} C_R & -C_I \\ C_I & C_R \end{bmatrix} = \hat{C}$$

Une suite de cette propriété est :  $A$  : Hermitienne, la matrice correspondant  $\hat{A}$  : symétrique

$$(5-d) : \text{La preuve est évidente en suivant (4-a) et le fait que } \hat{I}_n = I_{2n}.$$

$$(5-e) : \det(\hat{A}) = \det \left( \begin{bmatrix} I & jI \\ 0 & I \end{bmatrix} \begin{bmatrix} A_R & -A_I \\ A_I & A_R \end{bmatrix} \begin{bmatrix} I & -jI \\ 0 & I \end{bmatrix} \right) = \det \left( \begin{bmatrix} A_R + jA_I & 0 \\ A_I & A_R - jA_I \end{bmatrix} \right)$$

$$= \det \left( \begin{bmatrix} A & 0 \\ A_I & A^* \end{bmatrix} \right) = \det(A) \det(A^*) = \det(A) \det(A)^* = |\det(A)|^2$$

$$= \det(A) \det \left( (A^*)^T \right) = \det(A) \det(A^\dagger) = \det(AA^\dagger)$$

Dans cette preuve on s'est servi de :

$$E_{(m+n) \times (m+n)} = \begin{bmatrix} A_{n \times n} & 0_{n \times m} \\ C_{m \times n} & D_{m \times m} \end{bmatrix} \Rightarrow \det(E) = \det(A) \det(D)$$

$$(5-g) : y = Ax = (A_R + jA_I)(x_R + jx_I) = \overbrace{(A_R x_R - A_I x_I)}^{y_R} + j \overbrace{(A_I x_R + A_R x_I)}^{y_I}$$

$$\hat{A}\hat{x} = \begin{bmatrix} A_R & -A_I \\ A_I & A_R \end{bmatrix} \begin{bmatrix} x_R \\ x_I \end{bmatrix} = \begin{bmatrix} A_R x_R - A_I x_I \\ A_I x_R + A_R x_I \end{bmatrix} = \begin{bmatrix} y_R \\ y_I \end{bmatrix} = \hat{y}$$

$$(5-h) : \text{Re}(x^\dagger y) = \text{Re}((x_R^T - jx_I^T)(y_R + jy_I)) = x_R^T y_R + x_I^T y_I = \begin{bmatrix} x_R^T & x_I^T \end{bmatrix} \begin{bmatrix} y_R \\ y_I \end{bmatrix} = \hat{x}^T \hat{y}$$



## ***Annexe C : Distribution Normale (Gaussienne) complexe multi-dimensionnelle***

On dit que le vecteur aléatoire complexe  $\mathbf{x}_{n \times 1} \in C^n$  est gaussien, si chacun de ses éléments complexe, a une distribution gaussienne (la partie réelle et la partie imaginaire sont également gaussiennes). La corrélation entre les différents éléments ou entre la partie réelle et imaginaire est possible, de toute façon, les statistiques de ce vecteur complexe sont représentées par  $\boldsymbol{\mu}$  (la moyenne) et  $\mathbf{Q} = \boldsymbol{\Sigma}_x$  (la covariance). On peut regarder le vecteur complexe  $\mathbf{x}$  comme un vecteur réel de taille  $2n$ . Et comme la distribution normale réelle multi-dimensionnelle existe déjà dans la littérature mathématique, on utilise la projection  $\hat{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix}$  de ce vecteur complexe. Pour ce vecteur avec la moyenne  $\boldsymbol{\mu}_{\hat{\mathbf{x}}}$  et la matrice de covariance  $\boldsymbol{\Sigma}_{\hat{\mathbf{x}}}$ , la densité de probabilité s'écrit sous la forme suivante :

$$f_{\boldsymbol{\mu}, \boldsymbol{\Sigma}}(\mathbf{x}) = f_{\boldsymbol{\mu}_{\hat{\mathbf{x}}}, \boldsymbol{\Sigma}_{\hat{\mathbf{x}}}}(\hat{\mathbf{x}}) = \det(2\pi\boldsymbol{\Sigma}_{\hat{\mathbf{x}}})^{-1/2} \exp\left(-\frac{1}{2}(\hat{\mathbf{x}} - \hat{\boldsymbol{\mu}})^\dagger \boldsymbol{\Sigma}_{\hat{\mathbf{x}}}^{-1}(\hat{\mathbf{x}} - \hat{\boldsymbol{\mu}})\right) \quad (\text{C-1})$$

Dans le cas général, on ne peut pas écrire la densité de probabilité par rapport aux statistiques du vecteur complexe  $\mathbf{x}$  ( $\boldsymbol{\mu}$  et  $\boldsymbol{\Sigma}_x$ ), mais si ce vecteur est circulairement symétrique (Annexe A),  $\boldsymbol{\Sigma}_{\hat{\mathbf{x}}} = \hat{\mathbf{Q}}/2$  et l'équation (C-1) se met sous la forme suivante :

$$f_{\boldsymbol{\mu}, \boldsymbol{\Sigma}}(\mathbf{x}) = f_{\boldsymbol{\mu}_{\hat{\mathbf{x}}}, \boldsymbol{\Sigma}_{\hat{\mathbf{x}}}}(\hat{\mathbf{x}}) = \det(\pi\hat{\mathbf{Q}})^{-1/2} \exp\left(-(\hat{\mathbf{x}} - \hat{\boldsymbol{\mu}})^\dagger \hat{\mathbf{Q}}^{-1}(\hat{\mathbf{x}} - \hat{\boldsymbol{\mu}})\right) \quad (\text{C-2})$$

Maintenant, en utilisant les propriétés de la projection T (définie dans le lemme 1) et le corollaire 2, on arrive à :

$$f_{\boldsymbol{\mu}, \boldsymbol{\Sigma}}(\mathbf{x}) = \det(\pi\mathbf{Q})^{-1} \exp\left(-(\mathbf{x} - \boldsymbol{\mu})^\dagger \mathbf{Q}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) \quad (\text{C-3})$$



## **Annexe D : Décomposition valeur singulière (SVD)**

Quand, pour une matrice, il n'est pas possible de trouver la décomposition propre (Eigen Decomposition), quelle que soit la raison, on cherche à trouver la décomposition en valeur singulière (Singular Value Decomposition). Ce fait peut se produire, quand la matrice originale n'est pas carrée (pour laquelle on ne peut pas définir les valeurs et les vecteurs propres) ou quand sa matrice des vecteurs propres n'est pas inversible comme dans l'exemple suivant :

$$\text{Exemple 1 : } \mathbf{H} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}; \lambda_1 = \lambda_2 = 1 \Rightarrow \mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \mathbf{V} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Mais pour la même matrice  $\mathbf{H}$ , la décomposition en valeur singulière peut être effectuée comme :

$$\mathbf{H} = \begin{bmatrix} 1/\sqrt{5-\sqrt{5}} & 1/\sqrt{5+\sqrt{5}} \\ (\sqrt{5}-1)/2\sqrt{5-\sqrt{5}} & -(\sqrt{5}+1)/2\sqrt{5+\sqrt{5}} \end{bmatrix} \begin{bmatrix} \sqrt{5}+1 & 0 \\ 0 & \sqrt{5}-1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5+\sqrt{5}} & (\sqrt{5}+1)/2\sqrt{5+\sqrt{5}} \\ 1/\sqrt{5-\sqrt{5}} & -(\sqrt{5}-1)/2\sqrt{5-\sqrt{5}} \end{bmatrix}$$

Même si, nous avons une matrice carré pour laquelle la décomposition propre est possible, parfois la décomposition en valeurs propres est avantageuse, car elle donne certainement une matrice diagonale avec les éléments sont non-négatifs (nous allons montrer que les valeurs singulières sont toujours non-négatives). Ce qui est le cas de notre problème dans le calcul de la capacité de canal. Considérons l'exemple suivant :

$$\text{Exemple 2 : } \mathbf{H} = \begin{bmatrix} 2 & 3 \\ 0 & -2 \end{bmatrix} \Rightarrow \lambda_1 = 2; \lambda_2 = -2 \Rightarrow \mathbf{H} = \begin{bmatrix} 1 & 3/5 \\ 0 & -4/5 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 3/4 \\ 0 & -5/4 \end{bmatrix}$$

$$\text{Décomposition SVD : } \Rightarrow \mathbf{H} = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}$$

**Théorème :** Pour une matrice complexe  $\mathbf{H} \in C^{m \times n}$ , il existe des matrices unitaires  $\mathbf{U} \in C^{m \times m}$

et  $\mathbf{V} \in C^{n \times n}$ , telles que  $\mathbf{U}^\dagger \mathbf{H} \mathbf{V} = \mathbf{D}$ , où  $\mathbf{D}_{m \times n} = \begin{bmatrix} \mathbf{S}_{r \times r} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(m-r) \times r} & \mathbf{0}_{(m-r) \times (n-r)} \end{bmatrix}$  et  $\mathbf{S}$  est une matrice

diagonale réelle non-négative définie,  $\mathbf{S} = \text{diag}(d_1, d_2, \dots, d_r)$ ,  $d_1 \geq d_2 \geq \dots \geq d_r > 0$  et  $r = \text{rank}(\mathbf{H}) \leq \min(m, n)$ .

Comme les matrices  $\mathbf{U}$  et  $\mathbf{V}$  sont unitaires, cette diagonalisation peut être également interprétée comme une décomposition appelée SVD :

$$\mathbf{H} = \mathbf{U} \mathbf{D} \mathbf{V}^\dagger \quad (\text{D-1})$$

Plus particulièrement, on peut avoir une décomposition SVD minimisée de 'taille économique' :

$$\begin{aligned} \mathbf{H} &= [\mathbf{U}_1 \quad \mathbf{U}_2] \begin{bmatrix} \mathbf{S}_{r \times r} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(m-r) \times r} & \mathbf{0}_{(m-r) \times (n-r)} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^\dagger \\ \mathbf{V}_2^\dagger \end{bmatrix} \\ &= \mathbf{U}_1 \mathbf{S} \mathbf{V}_1^\dagger \end{aligned} \quad (\text{D-2})$$

où  $\mathbf{U}_1 \in C^{m \times r}$ ,  $\mathbf{U}_2 \in C^{m \times (m-r)}$ ,  $\mathbf{V}_1 \in C^{n \times r}$  et  $\mathbf{V}_2 \in C^{n \times (n-r)}$ .

**Preuve :** Si on accepte l'existence de cette décomposition, il est facile de montrer que  $\mathbf{U}$  et  $\mathbf{V}$  sont respectivement les matrices des vecteurs propres de  $\mathbf{H}\mathbf{H}^\dagger$  et  $\mathbf{H}^\dagger\mathbf{H}$  et  $d_1, \dots, d_r$  sont les racines carrées des valeurs propres non nulles de ces matrices.

$$\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^\dagger \Rightarrow \begin{cases} \mathbf{H}\mathbf{H}^\dagger = \mathbf{U}\mathbf{D}\mathbf{V}^\dagger\mathbf{V}\mathbf{D}^T\mathbf{U}^\dagger = \mathbf{U}\mathbf{D}\mathbf{D}^T\mathbf{U}^\dagger = \mathbf{U}\mathbf{A}^u\mathbf{U}^\dagger \\ \mathbf{H}^\dagger\mathbf{H} = \mathbf{V}\mathbf{D}^T\mathbf{U}^\dagger\mathbf{U}\mathbf{D}\mathbf{V}^\dagger = \mathbf{V}\mathbf{D}^T\mathbf{D}\mathbf{V}^\dagger = \mathbf{V}\mathbf{A}^v\mathbf{V}^\dagger \end{cases}$$

$$\text{où } \mathbf{A}_{m \times m}^u = \begin{bmatrix} \mathbf{S}_{r \times r}^2 & \mathbf{0}_{r \times (m-r)} \\ \mathbf{0}_{(m-r) \times r} & \mathbf{0}_{(m-r) \times (m-r)} \end{bmatrix} \text{ et } \mathbf{A}_{n \times n}^v = \begin{bmatrix} \mathbf{S}_{r \times r}^2 & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix}.$$

Mais il faut aussi prouver l'existence de cette décomposition. Comme la matrice  $\mathbf{H}^\dagger\mathbf{H}$  est hermitienne non-négative définie, ses valeurs propres sont réelles et non-négatives. On dénote les valeurs propres de  $\mathbf{H}^\dagger\mathbf{H}$  par  $\{d_i^2; i=1, \dots, n\}$ ,  $d_1 \geq \dots \geq d_r > 0 = d_{r+1} = \dots = d_n$  et les vecteurs propres orthonormaux correspondants par  $\{\mathbf{v}_i; i=1, \dots, n\}$ . On divise ces vecteurs propres en deux groupes, ceux qui correspondent aux valeurs propres non-zéros  $\mathbf{V}_1 = (\mathbf{v}_1, \dots, \mathbf{v}_r)$ , et ceux qui correspondent aux valeurs propres nuls  $\mathbf{V}_2 = (\mathbf{v}_{r+1}, \dots, \mathbf{v}_n)$ . En définissant la matrice  $\mathbf{S} = \text{diag}(d_1, \dots, d_r)$ , on peut écrire :

$$\mathbf{H}^\dagger\mathbf{H}[\mathbf{V}_1 \quad \mathbf{V}_2] = [\mathbf{V}_1 \quad \mathbf{V}_2] \begin{bmatrix} \mathbf{S}_{r \times r}^2 & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{bmatrix} \Rightarrow \begin{cases} \mathbf{H}^\dagger\mathbf{H}\mathbf{V}_1 = \mathbf{V}_1\mathbf{S}^2 \\ \mathbf{H}^\dagger\mathbf{H}\mathbf{V}_2 = \mathbf{0} \end{cases} \quad (\text{D-3})$$

Si on pré-multiplie le deuxième terme de (D-3) par  $\mathbf{V}_2^\dagger$ , on aura :

$$\mathbf{V}_2^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{V}_2 = \mathbf{0} \Rightarrow (\mathbf{H}\mathbf{V}_2)^\dagger\mathbf{H}\mathbf{V}_2 = \mathbf{0} \Rightarrow \mathbf{H}\mathbf{V}_2 = \mathbf{0} \quad (\text{D-4})$$

Et si on multiplie le premier terme de (D-3) d'abord par  $\mathbf{V}_1^\dagger$  à gauche et ensuite par  $\mathbf{S}^{-1}$  à gauche et à droite vu l'orthonormalité des vecteurs  $\mathbf{v}_i$ , on arrive à :

$$\begin{aligned} \mathbf{V}_1^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{V}_1 &= \mathbf{V}_1^\dagger\mathbf{V}_1\mathbf{S}^2 = \mathbf{S}^2 \\ \mathbf{S}^{-1}\mathbf{V}_1^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{V}_1\mathbf{S}^{-1} &= \mathbf{I}_r \Rightarrow (\mathbf{H}\mathbf{V}_1\mathbf{S}^{-1})^\dagger\mathbf{H}\mathbf{V}_1\mathbf{S}^{-1} = \mathbf{I}_r \Rightarrow \mathbf{U}_1^\dagger\mathbf{U}_1 = \mathbf{I}_r \end{aligned} \quad (\text{D-5})$$

où la matrice  $\mathbf{U}_1 \in C^{m \times r}$  est définie par  $\mathbf{U}_1 = \mathbf{H}\mathbf{V}_1\mathbf{S}^{-1}$ . Selon (D-4), cette matrice est orthonormale. On choisit n'importe quelle matrice  $\mathbf{U}_2 \in C^{m \times (m-r)}$  telle que  $\mathbf{U} = [\mathbf{U}_1 \quad \mathbf{U}_2]$  soit orthonormale aussi. Donc en utilisant (D-3) :

$$\mathbf{U}^\dagger\mathbf{H}\mathbf{V} = \begin{bmatrix} \mathbf{U}_1^\dagger \\ \mathbf{U}_2^\dagger \end{bmatrix} \mathbf{H}[\mathbf{V}_1 \quad \mathbf{V}_2] = \begin{bmatrix} \mathbf{U}_1^\dagger\mathbf{H}\mathbf{V}_1 & \mathbf{U}_1^\dagger\mathbf{H}\mathbf{V}_2 \\ \mathbf{U}_2^\dagger\mathbf{H}\mathbf{V}_1 & \mathbf{U}_2^\dagger\mathbf{H}\mathbf{V}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{U}_1^\dagger\mathbf{H}\mathbf{V}_1 & \mathbf{0} \\ \mathbf{U}_2^\dagger\mathbf{H}\mathbf{V}_1 & \mathbf{0} \end{bmatrix} \quad (\text{D-6})$$

A partir de la définition de  $U_1 = HV_1S^{-1}$  et de l'orthonormalité de  $U$ , on peut dire que :

$$U_1^\dagger HV_1 = S \quad ; \quad U_2^\dagger HV_1 = U_2^\dagger U_1 S = \mathbf{0}$$

Ainsi, l'équation (D-5) peut se mettre sous la forme suivante :

$$U^\dagger HV = \begin{bmatrix} S & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} = \mathbf{D} \quad (\text{D-7})$$

La définition de la matrice  $\mathbf{D}$  par cette dernière complète la preuve. On pouvait commencer la preuve en développant  $HH^\dagger$  au lieu de  $H^\dagger H$ .

**Définition :** Supposons que  $H = UDV^\dagger$  soit une SVD de  $H \in C^{m \times n}$ .

1. L'ensemble de  $\{d_1, \dots, d_r\}$  est appelé l'ensemble des **valeurs singulières** (non-zéros) de la matrice  $H$  il est dénoté par  $\mathbf{D}(H)$ . D'après la preuve du théorème, on peut remarquer que  $d_i(H) = \lambda_i^{1/2}(HH^\dagger) = \lambda_i^{1/2}(H^\dagger H)$ . On note qu'il existe respectivement  $m-r$  et  $n-r$  valeurs propres zéros pour les matrices  $HH^\dagger$  et  $H^\dagger H$  et  $\min(m,n) - r$  autres valeurs singulières zéros.
2. Les colonnes de  $U$  sont appelées les vecteurs singuliers gauches de  $H$  (et ils sont à la fois les vecteurs propres orthonormaux de  $HH^\dagger$ ).
3. Les colonnes de  $V$  sont appelées les vecteurs singuliers droits de  $H$  (et ils sont à la fois les vecteurs propres orthonormaux de  $H^\dagger H$ ).

**Remarque 1 :**  $\text{rank}(H) = r =$  le nombre des valeurs singulières non-zéros de  $H$ .

**Remarque 2 :** La décomposition en valeur singulière n'est pas unique. Par exemple, en suivant la preuve du théorème on peut remarquer que :

- N'importe quelle base orthonormale de  $N(H)$  (l'espace nulle de  $H$ ) peut être utilisée pour  $V_2$ .
- On peut utiliser n'importe quelle matrice  $U_2$  pourvu que  $[U_1 \quad U_2]$  soit orthonormale.
- Les colonnes de  $U$  et  $V$  peuvent être changées (en tandem) par le signe (ou en étant multipliées par  $e^{j\theta}$  en cas complexe).

**Remarque 3 :** Supposons que  $H \in C^{m \times m}$  soit une matrice hermitienne positive définie, et  $V$  soit une matrice des vecteurs propres orthonormale de  $H$ . Pour une telle matrice, les valeurs propres sont réelle et non-négative et sa décomposition propre ( $H = V\Lambda V^\dagger$ ) peut être à la fois une SVD de cette matrice, dans laquelle  $U = V$ .

**Exemple 3 :** A la fin, on présente un autre exemple pour montrer une SVD minimisée.

$$\begin{aligned}
\mathbf{H} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \\ 1 & 0 \end{bmatrix} &= \begin{bmatrix} 2/\sqrt{6} & 0 & 1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \\
&= \begin{bmatrix} 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}
\end{aligned}$$

## Annexe E : Preuve de l'expression (25)

Dans cette annexe on présente la preuve de l'équation suivante

$$E_0(\rho, q_x) = -\log\left(\int \left[\int q_x(\mathbf{x}) p(\mathbf{y}|\mathbf{x})^{1/(1+\rho)} d\mathbf{x}\right]^{1+\rho} d\mathbf{y}\right) = \rho \log\left(\det\left(\mathbf{I}_r + (1+\rho)^{-1} \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right)\right)$$

où  $p(\mathbf{y}|\mathbf{x}) = \det(\boldsymbol{\pi}\mathbf{I}_m)^{-1} \exp\left(-(\mathbf{y}-\mathbf{H}\mathbf{x})^\dagger(\mathbf{y}-\mathbf{H}\mathbf{x})\right)$  et  $f_\rho(\mathbf{x}) = \det(\boldsymbol{\pi}\mathbf{Q})^{-1} \exp\left(-\mathbf{x}^\dagger\mathbf{Q}^{-1}\mathbf{x}\right)$ .

Autrement dit, il faut démontrer que

$$I_2 = \underbrace{\int \left[\int q_x(\mathbf{x}) p(\mathbf{y}|\mathbf{x})^{1/(1+\rho)} d\mathbf{x}\right]^{1+\rho} d\mathbf{y}}_{I_1} = \left(\det\left(\mathbf{I}_m + (1+\rho)^{-1} \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger\right)\right)^{-\rho} \quad (\text{E-1})$$

Pour cela, on se servira de l'expression suivante qui vient de la définition d'une densité gaussienne, c'est à dire :

$$\int \exp\left(-\mathbf{z}^\dagger\mathbf{Q}^{-1}\mathbf{z}\right) d\mathbf{z} = \det(\boldsymbol{\pi}\mathbf{Q}) \quad (\text{E-2})$$

On commence par se concentrer sur le calcul de  $I_1$ .

$$\begin{aligned} I_1 &= \int \frac{1}{\det(\boldsymbol{\pi}\mathbf{Q})} e^{-\mathbf{x}^\dagger\mathbf{Q}^{-1}\mathbf{x}} \left(\frac{1}{\det(\boldsymbol{\pi}\mathbf{I}_m)}\right)^{1/(1+\rho)} e^{-\frac{(\mathbf{y}-\mathbf{H}\mathbf{x})^\dagger(\mathbf{y}-\mathbf{H}\mathbf{x})}{1+\rho}} d\mathbf{x} \\ &= \frac{1}{\det(\boldsymbol{\pi}\mathbf{Q})(\det(\boldsymbol{\pi}\mathbf{I}_m))^{1/(1+\rho)}} \int e^{-\mathbf{x}^\dagger\mathbf{Q}^{-1}\mathbf{x}} \cdot e^{-\frac{(\mathbf{y}-\mathbf{H}\mathbf{x})^\dagger(\mathbf{y}-\mathbf{H}\mathbf{x})}{1+\rho}} d\mathbf{x} \\ &= \underbrace{\left(\frac{1}{\pi^n \det(\mathbf{Q}) \pi^{m/(1+\rho)}}\right)}_K \int e^{-\mathbf{x}^\dagger\mathbf{Q}^{-1}\mathbf{x}} \cdot e^{-\frac{(\mathbf{y}^\dagger\mathbf{y} + \mathbf{x}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{x} - \mathbf{x}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{y} - \mathbf{y}^\dagger\mathbf{H}\mathbf{x})}{1+\rho}} d\mathbf{x} \\ &= K \int e^{-\left[\mathbf{x}^\dagger\left(\mathbf{Q}^{-1} + \frac{\mathbf{H}^\dagger\mathbf{H}}{1+\rho}\right)\mathbf{x} - \mathbf{x}^\dagger\left(\frac{\mathbf{H}^\dagger}{1+\rho}\right)\mathbf{y} - \mathbf{y}^\dagger\left(\frac{\mathbf{H}}{1+\rho}\right)\mathbf{x} + \frac{\mathbf{y}^\dagger\mathbf{y}}{1+\rho}\right]} d\mathbf{x} \\ &= Ke \left[ \frac{\mathbf{y}^\dagger\mathbf{y}}{1+\rho} - \frac{\mathbf{y}^\dagger\mathbf{H}\left(\mathbf{Q}^{-1} + \frac{\mathbf{H}^\dagger\mathbf{H}}{1+\rho}\right)^{-1}\mathbf{H}^\dagger\mathbf{y}}{(1+\rho)^2} \right] \int e^{-\left[\mathbf{x} - \left(\mathbf{Q}^{-1} + \frac{\mathbf{H}^\dagger\mathbf{H}}{1+\rho}\right)^{-1}\frac{\mathbf{H}^\dagger\mathbf{y}}{1+\rho}\right]^\dagger \left(\mathbf{Q}^{-1} + \frac{\mathbf{H}^\dagger\mathbf{H}}{1+\rho}\right) \left[\mathbf{x} - \left(\mathbf{Q}^{-1} + \frac{\mathbf{H}^\dagger\mathbf{H}}{1+\rho}\right)^{-1}\frac{\mathbf{H}^\dagger\mathbf{y}}{1+\rho}\right]} d\mathbf{x} \\ &= Ke \left[ \frac{\mathbf{y}^\dagger\mathbf{y} - \mathbf{y}^\dagger\mathbf{H}\left((1+\rho)\mathbf{Q}^{-1} + \mathbf{H}^\dagger\mathbf{H}\right)^{-1}\mathbf{H}^\dagger\mathbf{y}}{(1+\rho)} \right] \det\left(\pi\left(\mathbf{Q}^{-1} + \frac{\mathbf{H}^\dagger\mathbf{H}}{1+\rho}\right)^{-1}\right) \end{aligned} \quad (\text{E-3})$$

où la dernière égalité déduit de l'expression (E-2) et en remplaçant  $K$  par sa valeur (E-3) peut encore se simplifier comme suivant.

$$\begin{aligned}
I_1 &= \frac{1}{\pi^n \det(\mathbf{Q}) \pi^{m/(1+\rho)}} \times \frac{\pi^n}{\det\left(\mathbf{Q}^{-1} + \frac{\mathbf{H}^\dagger \mathbf{H}}{1+\rho}\right)} e^{-y^\dagger \left( \frac{\mathbf{I}_m - \mathbf{H}((1+\rho)\mathbf{Q}^{-1} + \mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger}{(1+\rho)} \right) y} \\
&= \frac{1}{\pi^{m/(1+\rho)} \det\left(\mathbf{I}_n + \frac{\mathbf{Q}\mathbf{H}^\dagger \mathbf{H}}{1+\rho}\right)} e^{-y^\dagger \left( \frac{\mathbf{I}_m - \mathbf{H}((1+\rho)\mathbf{Q}^{-1} + \mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger}{(1+\rho)} \right) y} \\
&= \frac{1}{\pi^{m/(1+\rho)} \det\left(\mathbf{I}_m + \frac{\mathbf{H}\mathbf{Q}\mathbf{H}^\dagger}{1+\rho}\right)} e^{-y^\dagger \left( \frac{\mathbf{I}_m - \mathbf{H}((1+\rho)\mathbf{Q}^{-1} + \mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger}{(1+\rho)} \right) y}
\end{aligned} \tag{E-4}$$

La dernière égalité vient de l'identité du déterminant (16). Maintenant on peut revenir au calcul de  $I_2 = \int I_1^{1+\rho} dy$  et on va encore utiliser (E-2) pour simplifier.

$$\begin{aligned}
I_2 &= \frac{1}{\pi^m \det\left(\mathbf{I}_m + \frac{\mathbf{H}\mathbf{Q}\mathbf{H}^\dagger}{1+\rho}\right)^{1+\rho}} \int e^{-y^\dagger \left( \frac{\mathbf{I}_m - \mathbf{H}((1+\rho)\mathbf{Q}^{-1} + \mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger}{(1+\rho)} \right) y} dy \\
&= \frac{1}{\pi^m \det\left(\mathbf{I}_m + \frac{\mathbf{H}\mathbf{Q}\mathbf{H}^\dagger}{1+\rho}\right)^{1+\rho}} \int e^{-y^\dagger \mathbf{C} y} dy = \frac{1}{\pi^m \det\left(\mathbf{I}_m + \frac{\mathbf{H}\mathbf{Q}\mathbf{H}^\dagger}{1+\rho}\right)^{1+\rho}} \times \det(\pi \mathbf{C}^{-1}) \\
&= \frac{1}{\det\left(\mathbf{I}_m + \frac{\mathbf{H}\mathbf{Q}\mathbf{H}^\dagger}{1+\rho}\right)^{1+\rho} \det(\mathbf{C})}
\end{aligned} \tag{E-5}$$

Il nous reste juste à calculer le terme de  $\det(\mathbf{C})$ , en utilisant (16) on aura :

$$\det(\mathbf{C}) = \det\left(\mathbf{I}_m - \mathbf{H}((1+\rho)\mathbf{Q}^{-1} + \mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger\right) = \det\left(\mathbf{I}_n - \mathbf{H}^\dagger \mathbf{H}((1+\rho)\mathbf{Q}^{-1} + \mathbf{H}^\dagger \mathbf{H})^{-1}\right) \tag{E-6}$$

On peut vérifier très simplement l'expression suivante en multipliant par  $((1+\rho)\mathbf{Q}^{-1} + \mathbf{H}^\dagger \mathbf{H})$  à droite et par  $(\mathbf{I}_n + \mathbf{H}^\dagger \mathbf{H}\mathbf{Q}/(1+\rho))$  à gauche.

$$\mathbf{I}_n - \mathbf{H}^\dagger \mathbf{H}((1+\rho)\mathbf{Q}^{-1} + \mathbf{H}^\dagger \mathbf{H})^{-1} = \left(\mathbf{I}_n + \frac{\mathbf{H}^\dagger \mathbf{H}\mathbf{Q}}{1+\rho}\right)^{-1} \tag{E-7}$$

En utilisant cette dernière et (16), (E-6) peut se mettre sous la forme suivante.

$$\det(\mathbf{C}) = \det\left(\mathbf{I}_n + \frac{\mathbf{H}^\dagger \mathbf{H}\mathbf{Q}}{1+\rho}\right)^{-1} = \det\left(\mathbf{I}_m + \frac{\mathbf{H}\mathbf{Q}\mathbf{H}^\dagger}{1+\rho}\right)^{-1} \tag{E-8}$$

En remplaçant  $\det(\mathbf{C})$  dans (E-5) par cette dernière, l'équation (E-2) sera justifiée.

$$I_2 = \det(\mathbf{I}_m + \mathbf{H}\mathbf{Q}\mathbf{H}^\dagger / (1 + \rho))^{-\rho} \quad (\text{E-9})$$



# **CHAPITRE 3**

## **DECODAGE SOUS-OPTIMAL DES SIGNAUX MIMO : BLAST**





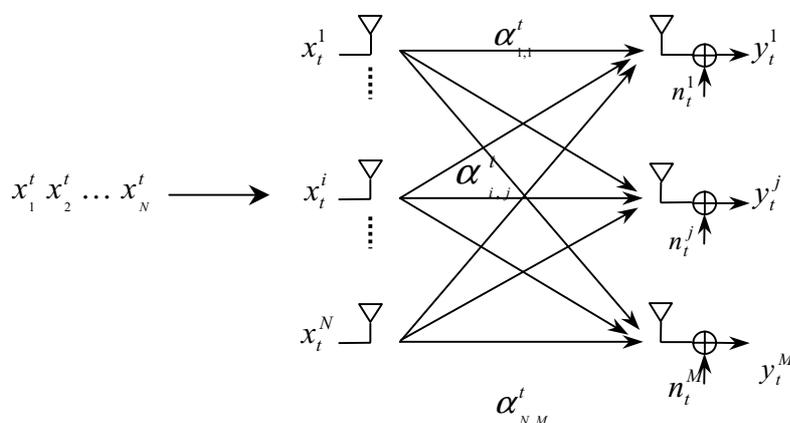
Cependant, l'algorithme D-BLAST présente une complexité relativement importante pour pouvoir être implanté en pratique, une simplification consiste en l'utilisation de l'algorithme V-BLAST [2] (Vertical BLAST) qui va faire l'objet de ce chapitre.

Dans cette structure simple, aucun codage n'est utilisé sur les signaux émis. Ici aussi comme pour le cas D-BLAST la détection est faite en  $N$  étapes en faisant la suppression successive d'interférences. L'ordre de détection des signaux se fait par SNR décroissant. La borne sur le taux de transmission pour l'architecture V-BLAST a été étudiée dans [3]. Il a été montré que la capacité augmente linéairement avec le nombre d'antennes, et que la borne du système D-BLAST peut être approchée à environ 72% près. K.C. Amis a montré dans [4] que la performance de la technique V-BLAST dépend considérablement de la différence  $(M - N)$ . Il est montré que la performance (en terme de probabilité d'erreur pour une modulation MDP-4) obtenue pour  $M = N$  est presque la même que pour un canal SISO. En augmentant  $(M - N)$  il y a une nette amélioration de la performance, qui s'approche de celle obtenue avec la méthode MAP. Ces résultats peuvent être expliqués de la façon suivante : dans le cas où  $M = N$ , l'ordre de diversité disponible pour le premier symbole à détecter est égal à zéro. Il est égal à un pour le deuxième, 2 pour le troisième, etc...La détection erronée du premier symbole peut engendrer une propagation d'erreur sur les décisions des autres symboles. En augmentant  $M - N$ , on peut profiter d'une diversité plus importante surtout pour les premiers symboles à détecter, et la performance sera meilleure.

Le principe de cet algorithme simplifié est le suivant : en émission, la séquence de données est démultiplexée en  $N$  sous-séquences (substreams), celles-ci sont ensuite modulées par  $N$  modulateurs indépendants et transmises sur chaque antenne d'émission. L'algorithme de détection en réception se compose principalement de trois parties 1) annulation d'interférences 2) suppression d'interférences 3) détection.

### 3-2 Modèle du système

La Figure 1 présente le schéma global de transmission d'un système multi-antennes.



**Figure 1- Modèle du système de transmission**

On considère un modèle de canal de fading quasi-statique (constant sur la période d'une trame pour chaque utilisation du canal) et les coefficients multiplicatifs varient de manière indépendante pour les utilisations suivantes. De plus, le canal est non sélectif en fréquence. Le système complet comprend  $N$  antennes d'émission et  $M$  antennes de réception.  $x_t^i$  représente le symbole émis par l'antenne d'émission  $i$  au temps  $t$ . Le coefficient  $\alpha_{i,j}^t$  est le gain complexe du trajet entre l'antenne d'émission  $i$  et l'antenne de réception  $j$  au temps  $t$ . Ces gains sont modélisés comme des variables aléatoires gaussiennes, centrées, complexes et indépendantes avec une variance égale à 0,5 pour chaque partie réelle et imaginaire.

Au temps  $t$ , le signal  $r_t^j$  reçu par l'antenne  $j$  s'écrit sous la forme suivante :

$$y_t^j = \sum_{i=1}^N \alpha_{i,j}^t x_t^i + n_t^j \quad (1)$$

où  $n_t^j$  est le bruit gaussien complexe additif sur l'antenne  $j$  au temps  $t$ . Les  $n_t^j$  sont considérés comme variables aléatoires gaussiennes centrées, complexes et indépendantes avec la variance  $\sigma^2 = 1/(2\text{SNR})$  par dimension complexe :  $E\{|n_t^j|^2\} = \sigma^2$ . L'index  $t$  sera sous-entendu pour une question de simplicité d'écriture. Les coefficients  $\alpha_{i,j}^t$  sont indépendants entre eux ce qui revient à supposer un espacement suffisant entre les antennes d'émission et de réception. On suppose que le récepteur connaît parfaitement les gains des trajets et la variance de bruit.

En utilisant une notation matricielle, la relation ( 1 ) peut s'écrire:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (2)$$

où :

$$\mathbf{H} = \begin{bmatrix} h^{11} & \dots & h^{N1} \\ \vdots & & \vdots \\ h^{1M} & \dots & h^{NM} \end{bmatrix}$$

$$\mathbf{y}_t = \begin{bmatrix} y_t^1 \\ \vdots \\ y_t^M \end{bmatrix}, \mathbf{x}_t = \begin{bmatrix} x_t^1 \\ \vdots \\ x_t^N \end{bmatrix}, \mathbf{n}_t = \begin{bmatrix} n_t^1 \\ \vdots \\ n_t^M \end{bmatrix}$$

La matrice  $\mathbf{H}$  contient les coefficients  $\alpha_{i,j}^t$  ( $h^{i,j} = \alpha_{i,j}^t$ ) du canal et le vecteur  $\mathbf{x}$  représente la sortie des antennes d'émission à l'instant  $t$ , le vecteur  $\mathbf{r}$  désigne les signaux reçus sur les antennes de réception, et le vecteur  $\mathbf{n}$  contient les échantillons de bruit du canal.

### 3-3 Architecture du récepteur

En supposant une synchronisation parfaite, nous allons maintenant expliquer le principe de l'algorithme V-BLAST. La Figure 2 illustre le schéma global du récepteur.

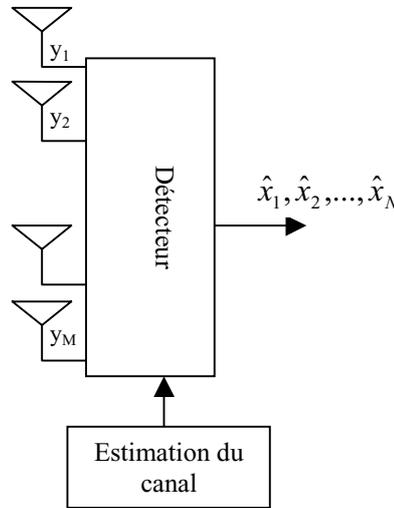


Figure 2- Schéma global du récepteur

Dans un premier temps, on considère le cas où le nombre d'antennes de réception est égal au nombre d'antennes d'émission :  $M = N$ . On généralisera les résultats obtenus au cas général où  $M \neq N$ .

Nous présentons maintenant une méthode itérative pour estimer simultanément toutes les signaux émis. L'idée de base est de transformer  $H$  en une matrice triangulaire supérieure  $R$  ce qui permettra d'estimer de façon itérative les symboles transmis.

On utilise une décomposition de la matrice  $H$  sous la forme:  $H = QR$ , où  $Q$  est une matrice unitaire :  $Q^\dagger Q = I$ , où  $R$  est une matrice triangulaire supérieure  $R = [r_{ij}]$  ( $r_{ij} = 0$  pour  $i > j$ ) [3] L'équation ( 2 ) peut alors s'écrire :

$$y = Hx + n = QRx + n \quad (3)$$

La transformation  $Q \rightarrow Q^\dagger$ , où  $Q^\dagger$  est la matrice hermitienne de  $Q$ , préserve la norme euclidienne. C'est à dire que pour chaque vecteur  $v$  on peut écrire :

$$|Q^\dagger v|^2 = (Q^\dagger v)^\dagger \cdot (Q^\dagger v) = v^\dagger (Q \cdot Q^\dagger) v = v^\dagger \cdot v = |v|^2$$

Ainsi si on multiplie un vecteur par  $Q^\dagger$ , sa norme demeure inchangée. Egalement si on multiplie les deux côtés d'une égalité par  $Q^\dagger$  l'égalité des normes demeure. Dans le contexte étudié ici on multiplie le signal reçu par  $Q^\dagger$ . En multipliant l'équation ( 3 ) par  $Q^\dagger$  on arrive à :

$$\begin{aligned} \tilde{y} &= Q^\dagger y & (4) \\ &= Q^\dagger (QRx + n) \\ &= Rx + \tilde{n} \end{aligned}$$

Les caractéristiques du bruit sont inchangées, il reste gaussien centré avec la même variance. Ceci est vérifié car l'opérateur  $\mathbf{Q}^\dagger$  est une matrice unitaire. On peut réécrire ( 4 ) en considérant la forme triangulaire de  $\mathbf{R}$  :

$$\begin{pmatrix} \tilde{\mathbf{y}}_1 \\ \tilde{\mathbf{y}}_2 \\ \vdots \\ \tilde{\mathbf{y}}_N \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1N} \\ 0 & r_{22} & \cdots & r_{2N} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{NN} \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_N \end{pmatrix} + \begin{pmatrix} \tilde{\mathbf{n}}_1 \\ \tilde{\mathbf{n}}_2 \\ \vdots \\ \tilde{\mathbf{n}}_N \end{pmatrix} \quad (5)$$

A chaque étape on suppose que la décision précédente est correcte. L'algorithme dans le domaine spatial (antenne) peut être décrit de la façon suivante :

$$\begin{aligned} \tilde{\mathbf{y}}_N &= r_{NN}\mathbf{x}_N + \tilde{\mathbf{n}}_N && \longrightarrow \text{on obtient } \hat{\mathbf{x}}_N \\ \tilde{\mathbf{y}}_{N-1} &= r_{N-1N-1}\mathbf{x}_{N-1} + r_{N-1N}\mathbf{x}_N + \tilde{\mathbf{n}}_{N-1} && \longrightarrow \text{on obtient } \hat{\mathbf{x}}_{N-1} \\ \vdots & && \vdots \\ \tilde{\mathbf{y}}_1 &= r_{11}\mathbf{x}_1 + \sum_{j=2}^N r_{1j}\mathbf{x}_j + \tilde{\mathbf{n}}_1 && \longrightarrow \text{on obtient } \hat{\mathbf{x}}_1 \end{aligned} \quad (6)$$

Donc, à partir de la première ligne de ( 6 ),  $\hat{\mathbf{x}}_N$  peut être obtenu. En supposant  $\hat{\mathbf{x}}_N$  correct, on peut donc mettre  $\mathbf{x}_N = \hat{\mathbf{x}}_N$  dans la deuxième ligne, alors on peut trouver l'estimation  $\hat{\mathbf{x}}_{N-1}$ . Egalement, en supposant qu'à la  $i^{\text{ème}}$  étape on connaît déjà les valeurs de  $\hat{\mathbf{x}}_N, \hat{\mathbf{x}}_{N-1}, \dots, \hat{\mathbf{x}}_{N-i+1}$ , (les décisions précédentes sont correctes), on a :

$$\begin{aligned} \tilde{\mathbf{y}}_{N-i+1} &= r_{N-i+1, N-i+1}\mathbf{x}_{N-i+1} + \sum_{j=N-i+2}^N r_{N-i+1, j}\hat{\mathbf{x}}_j + \tilde{\mathbf{n}}_{N-i+1} \longrightarrow \\ r_{N-i+1, N-i+1}\mathbf{x}_{N-i+1} + \tilde{\mathbf{n}}_{N-i+1} &= \tilde{\mathbf{y}}_{N-i+1} - \sum_{j=N-i+2}^N r_{N-i+1, j}\hat{\mathbf{x}}_j \end{aligned} \quad (7)$$

En répétant le même processus jusqu'à la dernière ligne toutes les valeurs,  $\hat{\mathbf{x}}_N, \hat{\mathbf{x}}_{N-1}, \dots, \hat{\mathbf{x}}_1$ , peuvent être obtenues.

On considère le cas  $M > N$  où le nombre d'antennes de réception est plus important que le nombre d'antenne d'émission. Ceci est équivalent à dire que le nombre d'équations est plus important que le nombre de variables inconnues. Pour l'instant, on laisse de côté le problème du bruit, il s'agit de trouver les solutions pour la forme matricielle suivante :

$$\mathbf{y} = \mathbf{H}\mathbf{x}$$

On choisit la méthode des moindres carrés pour ce système qui consiste à trouver  $\mathbf{x}_0$  qui minimise la norme du résidu :

$$\min |\mathbf{r}|^2 = F(\mathbf{x}) = |\mathbf{y} - \mathbf{H}\mathbf{x}|^2 \quad (8)$$

Afin de minimiser  $|\mathbf{r}|^2$ , on la note  $F(x)$ , on peut dire que :

$$\begin{aligned} F(\mathbf{x}) &= |\mathbf{r}|^2 = |\mathbf{y} - \mathbf{H}\mathbf{x}|^2 = (\mathbf{y} - \mathbf{H}\mathbf{x})^\dagger (\mathbf{y} - \mathbf{H}\mathbf{x}) \\ \frac{dF}{d\mathbf{x}} &= 2(\mathbf{H}^\dagger \mathbf{H}\mathbf{x} - \mathbf{H}^\dagger \mathbf{y})|_{\mathbf{x}=\tilde{\mathbf{x}}} = 0 \Rightarrow \mathbf{H}^\dagger \mathbf{H}\tilde{\mathbf{x}} = \mathbf{H}^\dagger \mathbf{y} \end{aligned} \quad (9)$$

On remarque que, si  $\mathbf{H}^\dagger \mathbf{H}$  est une matrice de rang complet, elle est inversible. Dans ce cas-là il y a une solution unique. A partir de (9) on arrive à :

$$\tilde{\mathbf{x}} = (\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{y} = \mathbf{H}^+ \mathbf{y}$$

où  $\mathbf{H}^+$  est la pseudo-inverse de matrice  $\mathbf{H}$  :

$$\mathbf{H}^+ = (\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \quad (10)$$

Pour que la matrice  $\mathbf{H}^\dagger \mathbf{H}$  soit inversible, il faut les conditions suivantes :

- $M \geq N$  ( $\mathbf{H}_{M \times N}$ ); ceci constitue une contrainte de réalisation de l'algorithme V-BLAST.
- les colonnes de  $\mathbf{H}$  doivent être linéairement indépendantes.

D'après la minimisation MMSE, on peut dire que le vecteur  $\mathbf{r}$  est orthogonal à l'espace colonne de  $\mathbf{H}$  car :

$$\mathbf{H}^\dagger \mathbf{H}\tilde{\mathbf{x}} = \mathbf{H}^\dagger \mathbf{y} \Rightarrow \mathbf{H}^\dagger (\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}) = \mathbf{H}^\dagger \mathbf{r} = 0 \quad (11)$$

Pour le cas où  $M > N$  on peut généraliser la factorisation  $\mathbf{QR}$  sous la forme suivante [5] :

$$\mathbf{H} = \mathbf{Q} \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}$$

où  $\mathbf{R}_{N \times N}$  est une matrice triangulaire supérieure,  $\mathbf{0}_{M-N \times M}$  et  $\mathbf{Q}_{M \times N}$ . Le problème devient alors :

$$\mathbf{Q}^\dagger \mathbf{H}\mathbf{x} = \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} = \mathbf{Q}^\dagger \mathbf{y} \quad (12)$$

Ce système conserve la même solution que le système précédent, puisque la norme du résidu est conservée. En remplaçant (12) dans (9) on arrive à :

$$|\mathbf{r}|^2 = \left| \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} - \mathbf{Q}^\dagger \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} \mathbf{x} \right|^2 = |\mathbf{c}_1 - \mathbf{R}\mathbf{x}|^2 + |\mathbf{c}_2|^2$$

On considère  $\mathbf{Q} = [\mathbf{Q}_1 \ \mathbf{Q}_2]$ , où  $\mathbf{Q}_1$  est de dimension  $M \times N$  et  $\mathbf{Q}_2$  est de dimension  $M \times (M-N)$ .

Donc minimiser le résidu revient à minimiser  $|\mathbf{c}_1 - \mathbf{R}\mathbf{x}|^2$  le seul terme de l'égalité qui dépende de  $\mathbf{x}$  et peut donc être minimisé. Puisqu'il s'agit d'un système carré  $N \times N$ , on peut obtenir une

solution exacte. Le problème des moindres carrés  $\mathbf{y} = \mathbf{H}\mathbf{x}$  se ramène donc à la solution d'un système carré. En considérant la factorisation faite, il s'agit d'un système triangulaire supérieur qui est encore plus simple à résoudre. Maintenant que nous avons expliqué l'idée mathématique de l'algorithme V-BLAST, nous allons détailler les différentes étapes de la détection.

### 3-3-1 Suppression d'interférences

On suppose que le récepteur a détecté les  $i-1$  premiers  $\mathbf{x}_i$ . On note que le signal reçu (cf (3)) peut encore s'écrire [6] :

$$\mathbf{y} = \mathbf{x}_1 \cdot h_1 + \mathbf{x}_2 \cdot h_2 + \dots + \mathbf{x}_i \cdot h_i + \dots + \mathbf{x}_N \cdot h_N + \mathbf{n} \quad (13)$$

où  $h_j$  est défini comme  $[h^{1j}, h^{2j}, \dots, h^{Nj}]$  et  $\mathbf{H} = [h_1, h_2, \dots, h_N]$ . On considère que les  $i-1$  premières décisions étaient sans erreurs. On peut alors supprimer l'interférence des composantes de  $\mathbf{x}$ . On réécrit (13)

$$\mathbf{y} = (\hat{\mathbf{x}}_1 \cdot h_1 + \hat{\mathbf{x}}_2 \cdot h_2 + \dots + \hat{\mathbf{x}}_{i-1} \cdot h_{i-1}) + \mathbf{x}_i \cdot h_i + (\mathbf{x}_{i+1} \cdot h_{i+1} + \dots + \mathbf{x}_N \cdot h_N) + \mathbf{n}.$$

Dans cette expression, la première parenthèse contient seulement les terme correctement détectés ; on les soustrait de  $\mathbf{y}$ . Le résultat est un  $n$ -D vecteur, on le note  $\mathbf{u}^i$

$$\mathbf{u}_i = (\mathbf{x}_i \cdot h_i) + (\mathbf{x}_{i+1} \cdot h_{i+1} + \dots + \mathbf{x}_N \cdot h_N) + \mathbf{n}.$$

soit

$$\mathbf{u}_i = \mathbf{y} - \sum_{j=1}^{i-1} \hat{\mathbf{x}}_j h_j \quad (14)$$

Considérant  $\mathbf{u}_i$  au lieu de  $\mathbf{y}$  on aura :

$$\mathbf{u}_i = [0 \ 0 \ \dots \ \mathbf{x}_i \cdot h_i \ \mathbf{x}_{i+1} \cdot h_{i+1} \ \dots \ \mathbf{x}_N \cdot h_N]^T + \mathbf{n} = \mathbf{H} [0 \ 0 \ \dots \ \mathbf{x}_i \ \mathbf{x}_{i+1} \ \dots \ \mathbf{x}_N]^T + \mathbf{n}$$

Dans chaque étape donc, on traite le vecteur  $\mathbf{u}_i$  au lieu de  $\mathbf{y}$ .

### 3-3-2 Annulation d'interférences

L'étape d'annulation d'interférences libère le processus de détection  $\mathbf{x}_i$  des autres symboles transmis simultanément,  $\mathbf{x}_{i+1}, \mathbf{x}_{i+2}, \dots, \mathbf{x}_N$ . Cette étape peut s'exécuter en choisissant des vecteurs de poids  $\mathbf{w}_i, i = 1, 2, \dots, N$  tels que [7] :

$$\mathbf{w}_i^T \mathbf{h}_j = \delta_{ij}$$

où  $\mathbf{h}_j$  est la  $j^{\text{ème}}$  colonne de  $\mathbf{H}$  et  $\delta$  est le delta Kronecker, où  $\delta_{ij} = \begin{cases} 0 & j \neq i \\ 1 & j = i \end{cases}$ , ce qui revient à

dire :

$$\mathbf{w}_i^T \mathbf{h}_j = \begin{cases} 0 & j \neq i \\ 1 & j = i \end{cases} \quad (15)$$

Si on considère (10) il est clair que  $\mathbf{w}_i$  peut-être la  $n^{\text{ème}}$  ligne de  $\mathbf{H}^+$  où  $+$  indique la matrice pseudo-inverse de Moore-Penrose, qui peut être obtenue à partir de l'équation (10).

En considérant l'équation ( 2 ) la décision pour la  $i$ -ème couche est

$$\begin{aligned}
 \mathbf{w}_i^T \mathbf{y} &= \mathbf{w}_i^T (\mathbf{H}\mathbf{x} + \mathbf{n}) \\
 &= (\mathbf{w}_i^T \mathbf{H})\mathbf{x} + \mathbf{w}_i^T \mathbf{n} \\
 (\mathbf{w}_i^T \mathbf{H})\mathbf{x} &= [0 \ 0 \ \dots \ 1 \ \dots \ 0] \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_i \\ \vdots \\ \mathbf{x}_N \end{bmatrix} = \mathbf{x}_i \\
 \mathbf{w}_i^T \mathbf{y} &= (\mathbf{w}_i^T \mathbf{H})\mathbf{x} + \mathbf{w}_i^T \mathbf{n} = \mathbf{x}_i + \tilde{\mathbf{n}}_i = \tilde{\mathbf{x}}_i
 \end{aligned} \tag{11}$$

Ceci explique que vecteur  $\mathbf{w}_i^T$  est orthogonal au sous-espace de  $\{\mathbf{h}_j\}$  où  $j \neq i$  (dimension  $N-1$ ). C'est une conséquence directe de la suppression.

$$(\mathbf{w}_i^T \mathbf{H})\mathbf{u}_i = \begin{bmatrix} id \downarrow & id \downarrow & \dots & 1 \downarrow & 0 & \dots & 0 \downarrow \\ j=1 & j=2 & & j=i & & & j=N \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \mathbf{x}_i \\ \vdots \\ \mathbf{x}_N \end{bmatrix} = \mathbf{x}_i$$

$$\mathbf{H}\mathbf{u}_i = [h_1 \ h_2 \ \dots \ h_i \ \dots \ h_N] \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \mathbf{x}_i \\ \vdots \\ \mathbf{x}_N \end{bmatrix} = [0 \ 0 \ \dots \ h_i \ \dots \ h_N] \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \mathbf{x}_i \\ \vdots \\ \mathbf{x}_N \end{bmatrix}$$

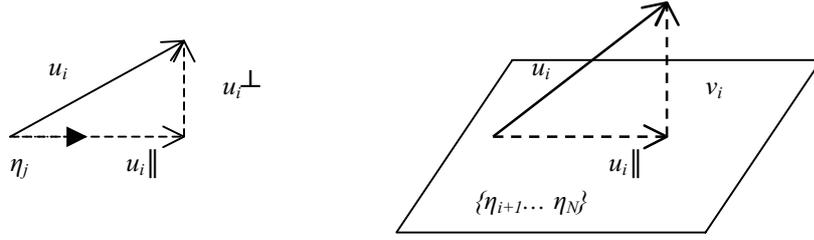
donc  $\mathbf{H}\mathbf{u}_i = \mathbf{H}_{i-1}^+ \mathbf{u}_i$ , où  $\mathbf{H}_{i-1}^+$  dénote la matrice obtenue en mettant les colonnes 1, 2, ...,  $i-1$  à zéro. Par conséquent  $\mathbf{w}_i^T$  vérifie l'équation suivante :

$$\mathbf{w}_i^T (\mathbf{H}_{i-1}^+)_j = \begin{cases} 0 & j \neq i \\ 1 & j = i \end{cases}$$

donc  $\mathbf{w}_i$  correspond à la  $j^{\text{ième}}$  ligne de  $\mathbf{H}_{i-1}^+$ .

### 3-3-3 Le point de vue géométrique

L'étape d'annulation d'interférence est basée sur le processus de Gram-Schmidt. Pour éviter des interférences, on projette  $\mathbf{u}_i$  sur le sous-espace engendré par  $\mathbf{h}_{i+1}, \mathbf{h}_{i+2}, \dots, \mathbf{h}_N$



**Figure 3- Illustration géométrique de l'étape d'annulation d'interférences. Le vecteur reçu (après la suppression d'interférences) est décomposé en sa composante orthogonale et sa composante parallèle**

Afin d'obtenir une modélisation mathématique nous construisons un ensemble de vecteurs orthogonaux  $\boldsymbol{\eta}_{i+1}, \boldsymbol{\eta}_{i+2}, \dots, \boldsymbol{\eta}_N$ . de vecteurs linéairement indépendants des  $\mathbf{h}_i$ .s. On utilise pour cela le procédé de Gram-Schmidt. Pour produire  $\boldsymbol{\eta}_i$   $i = 1, \dots, N$ , il faut dérouler les étapes suivantes [8]

Etape 1 :

$$\begin{aligned}\boldsymbol{\eta}_1 &= \mathbf{h}_1 \\ \boldsymbol{\eta}_1 &= \boldsymbol{\eta}_1 / \|\boldsymbol{\eta}_1\|\end{aligned}$$

Etape 2 :

$$\begin{aligned}\boldsymbol{\eta}_2 &= \mathbf{h}_2 - \langle \mathbf{h}_2, \boldsymbol{\eta}_1 \rangle \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 &= \boldsymbol{\eta}_2 / \|\boldsymbol{\eta}_2\|\end{aligned}$$

Etape 3 :

$$\begin{aligned}\boldsymbol{\eta}_3 &= \mathbf{h}_3 - \langle \mathbf{h}_3, \boldsymbol{\eta}_1 \rangle \boldsymbol{\eta}_1 - \langle \mathbf{h}_3, \boldsymbol{\eta}_2 \rangle \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 &= \boldsymbol{\eta}_3 / \|\boldsymbol{\eta}_3\|\end{aligned}$$

Etape  $i$  :

$$\begin{aligned}\boldsymbol{\eta}_i &= \mathbf{h}_i - \langle \mathbf{h}_i, \boldsymbol{\eta}_1 \rangle \boldsymbol{\eta}_1 - \dots - \langle \mathbf{h}_i, \boldsymbol{\eta}_j \rangle \boldsymbol{\eta}_j - \dots - \langle \mathbf{h}_i, \boldsymbol{\eta}_{i-1} \rangle \boldsymbol{\eta}_{i-1} \\ \boldsymbol{\eta}_i &= \boldsymbol{\eta}_i / \|\boldsymbol{\eta}_i\|\end{aligned}$$

On note le vecteur de projection  $\mathbf{v}_i$  égal à:

$$\mathbf{v}_i = \mathbf{u}_i - \langle \mathbf{u}_i, \boldsymbol{\eta}_{i+1} \rangle \boldsymbol{\eta}_{i+1} - \langle \mathbf{u}_i, \boldsymbol{\eta}_{i+2} \rangle \boldsymbol{\eta}_{i+2} \dots - \langle \mathbf{u}_i, \boldsymbol{\eta}_N \rangle \boldsymbol{\eta}_N$$

où  $\langle \mathbf{u}_i, \boldsymbol{\eta}_j \rangle$  représente le produit scalaire des vecteur  $\mathbf{u}_i$  et  $\boldsymbol{\eta}_j$ ,  $\langle \mathbf{u}_i, \boldsymbol{\eta}_j \rangle \boldsymbol{\eta}_j$  représente la projection de  $\mathbf{u}_i$  sur  $\boldsymbol{\eta}_j$ .

### 3-3-4 Effet de la suppression d'interférence sur les vecteurs poids $\mathbf{w}_i$

Afin de quantifier l'efficacité de la suppression d'interférences on utilise le rapport signal à bruit comme critère de qualité. On calcule ce SNR en revenant à l'équation ( 11 ), on a [8] :

$$SNR_{(i)} = \frac{E(|\mathbf{x}_i|^2)}{E(|\tilde{\mathbf{n}}_i|^2)}$$

avec

$$E(|\tilde{\mathbf{n}}_i|^2) = E(\langle \tilde{\mathbf{n}}_i, \tilde{\mathbf{n}}_i^* \rangle) = E(\langle \mathbf{w}_i^T \mathbf{n}_i, \mathbf{n}_i^* \mathbf{w}_i^{*T} \rangle) = |\mathbf{w}_i^T|^2 E(\langle \mathbf{n}_i, \mathbf{n}_i^* \rangle) = |\mathbf{w}_i^T|^2 \sigma^2$$

où  $\delta^2$  représente la puissance du bruit à chaque antenne de réception. On remplace la valeur obtenu dans l'équation précédente, on arrive à :

$$SNR_{(i)} = \frac{1}{N|\mathbf{w}_i|^2 \delta^2} \quad (16)$$

Une des conséquences de l'inégalité de Cauchy-Schwartz (cf Annexe) est que la norme de  $\mathbf{w}_i^T$  est d'autant plus grande que le nombre de lignes de  $\mathbf{H}$  orthogonales à  $\mathbf{w}_i^T$  est élevé. Ainsi, en utilisant la suppression d'interférences, les  $SNR$ 's obtenus seront meilleurs que leurs homologues obtenus lors de l'annulation sans suppression.

### 3-3-5 Ordre de détection

En utilisant la suppression d'interférence, l'ordre dans lequel les composants de  $\mathbf{x}$  sont détectés devient important ; cet ordre influence directement la qualité de la performance globale de décodage. Le détecteur désiré minimise la probabilité de faire une erreur de décision dans un bloc. Afin de réduire au minimal cette probabilité, il faut ranger les  $N$  composantes obtenues après suppression d'interférences selon les  $SNR$ 's décroissants. Le critère de rangement est donc le suivant : Maximiser le rapport  $[SNR_{(i)}, 1 \leq i \leq N]$  minimal où  $SNR_{(i)}$  est le  $SNR$  de la  $i$ -ème couche après détection. Il est alors aisé de montrer que ce critère équivaut à minimiser la probabilité d'erreur sur un bloc. En effet, si on considère que  $P_v$  est la probabilité qu'un symbole détecté soit erroné, on a [6] :

$$\begin{aligned} \text{Prob}[\text{bloc incorrect}] &= 1 - \text{Prob}[\text{bloc correct}] & (17) \\ &= 1 - \text{Prob}[\text{symbole correct}]^k = 1 - (1 - P_v)^k \\ &= P_v [1 + (1 - P_v) + \dots + (1 - P_v)^{k-1}] \end{aligned}$$

La dernière ligne correspond à la somme sur  $k$  événements indépendants. En considérant  $\lim(P_v) \rightarrow 0$ , l'équation (17) devient :

$$\text{Prob}[\text{bloc incorrect}] \approx k \cdot P_v \quad (18)$$

On peut pareillement obtenir  $P_v$  en fonction de la probabilité d'erreur par bit  $P_b$  :

$$P_v = 1 - [(1 - P_{b1})(1 - P_{b2}) \dots (1 - P_{bN})] \quad (19)$$

Si les erreurs faites à chaque niveau de détection sont statistiquement indépendantes, on peut réécrire l'équation (19)

$$P_v = \sum_{j=1}^N \left\{ P_b(SNR_{(j)}) \times \prod_{i=1}^j [1 - P_b(SNR_{(i)})] \right\}$$

A savoir que  $P_b(\cdot)$  est directement liée au SNR [5]. On peut écrire de plus dans le cas d'une modulation MAQ- $K$  :

$$P_b \approx \left[ \frac{(K^{1/2} - 1)}{(K^{1/2} \cdot \log_2 K)} \right] \times \pi^{-1/2} \int_a^{\infty} e^{(-\xi^2)} d\xi$$

$$(a = \left[ \frac{3 \cdot \text{SNR}_{(i)}}{2 \cdot (K - 1)} \right]^{1/2})$$

$P_b(\text{SNR}_{(i)})$  décroît exponentiellement avec  $\text{SNR}_{(i)}$ , pour les hauts SNRs et on peut considérer que  $1 - P_b(\text{SNR}_{(i)}) \approx 0$  donc

$$P_v \approx \sum_{j=1}^N P_b(\text{SNR}_{(j)})$$

$$\approx P_b(\text{SNR}_{(n)}) \left[ 1 + \sum_{j \neq n} P_b(\text{SNR}_{(j)}) / P_b(\text{SNR}_{(n)}) \right] \quad (20)$$

$$\approx P_b(\text{SNR}_{(n)})$$

où  $P_b(\text{SNR}_{(n)})$  est la valeur maximale des  $P_b(\text{SNR}_{(i)})$  et  $\text{SNR}_{(n)}$  est la valeur minimale des  $N$   $\text{SNR}_{(j)}$ s. Il est clair que l'hypothèse d'indépendance entre les erreurs, de par le processus d'annulation d'interférences utilisé, ne peut être complètement vérifiée. L'équation (20) donc n'est donc pas strictement correcte, mais elle est asymptotiquement correcte. En remplaçant les valeurs des probabilités obtenues dans [9] on aura :

$$\text{Prob}[\text{bloc incorrect}] \approx k \cdot P_b(\text{SNR}_{(n)})$$

Ce qui revient à dire la probabilité d'erreur dans un paquet de symboles est fonction du SNR minimal des symboles reçus et la plus faible « couche » peut donc dominer la performance globale du décodage, particulièrement la première « couche » décodée, car elle possède le moins de diversité. L'ordre de détection nous permet de choisir à la  $i^{\text{ème}}$  étape les  $N-i$  colonnes avec qui  $w_i^T$  doit être orthogonal. Les différents choix mènent à différents  $\text{SNR}_{(i)}$ .

### **3-4 Algorithme de programmation du détecteur VBLAST**

Soit  $S \equiv \{k_1, k_2, \dots, k_N\}$  l'ordre optimum de détection. Le processus se décompose en trois étapes.

#### **- Etape d'initialisation**

$$i = 1$$

$$u_1 = r$$

$$G_1 = H^+ :$$

$$k_1 = \arg \left\{ \min_j |(\mathbf{G}_1)_j|^2 \right\} \text{ où } (\mathbf{G}_1)_j \text{ est la } j^{\text{ième}} \text{ ligne de } \mathbf{G}_1$$

### - Annulation

On obtient la statistique de décision en utilisant le vecteur de poids :  $\tilde{\mathbf{x}}_{ki} = \mathbf{w}_{ki}^T \mathbf{u}_{ki}$

Détection :  $\hat{x}_{k1} = Q(\tilde{x}_{k1})$

Où  $Q(\cdot)$  désigne la métrique quadratique de maximum de vraisemblance. On cherche le symbole le plus proche de  $\tilde{x}_{k1}$  dans la constellation

$$\hat{x}_{ki} = \arg \min_{s \in A} |\tilde{x}_{ki} - s|^2$$

$A$  désigne les symboles de la constellation,  $|\cdot|$  est la distance d'Euclidienne .

### - Suppression d'interférences

On retranche l'interférence des composants détectés :

$$\mathbf{u}_{i+1} = \mathbf{u}_i - \hat{x}_{ki} (\mathbf{H})_{ki}$$

$$\mathbf{G}_{i+1} = \mathbf{H}_{i-1}^+$$

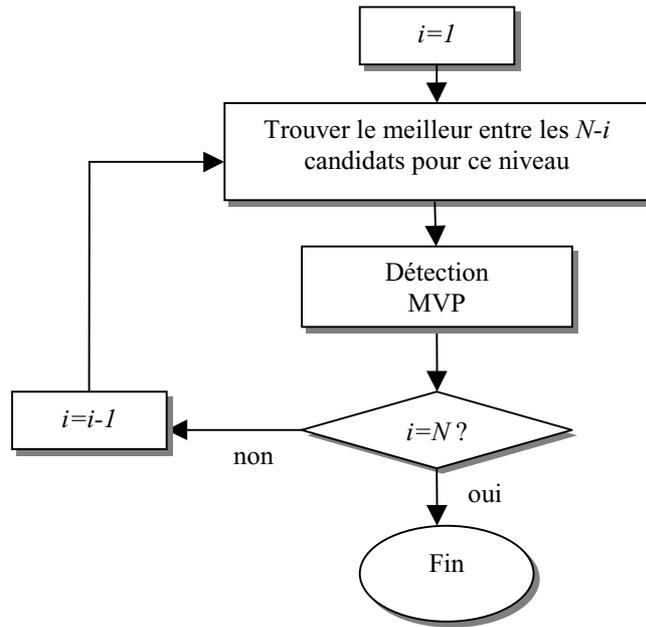
$\mathbf{H}_{i-1}^+$  : matrice obtenue en mettant les colonnes 1, 2, ..., i-1 à zéro

$$k_{i+1} = \arg \left\{ \min_{j \notin \{k_1, \dots, k_i\}} |(\mathbf{G}_{i+1})_j|^2 \right\}$$

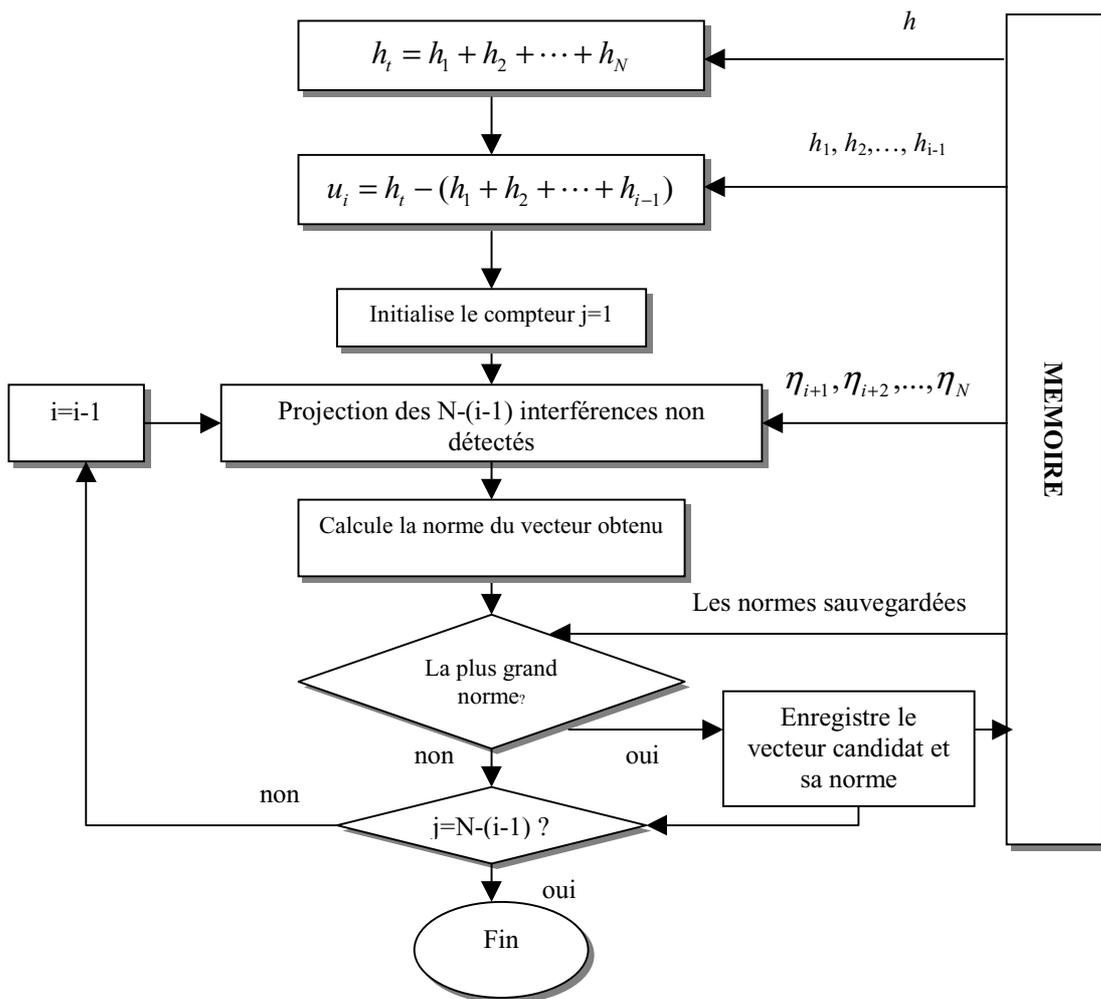
$$i \leftarrow i+1$$

### Simplification

On suppose que la matrice  $\mathbf{H}$  est constante sur la durée d'un bloc, on peut simplifier l'algorithme de détection. Dans la section précédente on a vu que les vecteurs de poids et l'ordre optimum dépendent de  $\mathbf{H}$ . On peut donc les calculer une seule fois pour un bloc et les enregistrer. Afin de calculer les vecteurs de poids  $\mathbf{w}_i$ , on propose la méthode suivante : on calcule les vecteurs orthonormaux à partir du processus de Gram-Schmidt. A chaque niveau on prend le vecteur le plus grand possible. Cette méthode consiste à détecter le vecteur  $\mathbf{1}$  pour chaque dimension (symbole 1 reçu pour chaque antenne de réception). On peut donc lancer l'algorithme pour ce vecteur, et à chaque étape la composante détectée est égale au vecteur de poids. Pour éviter de calculer l'inverse de la matrice  $\mathbf{H}$ , on peut chercher le vecteur de norme maximum entre tous les vecteurs possibles à chaque étape. Les Figure 4 - Figure 6 montrent les schéma de calcul des vecteurs de poids.



**Figure 4 : Algorithme V-BLAST**



**Figure 5 : Algorithme V-BLAST**

On peut décrire l'algorithme modifié simplifié sous la forme suivante :

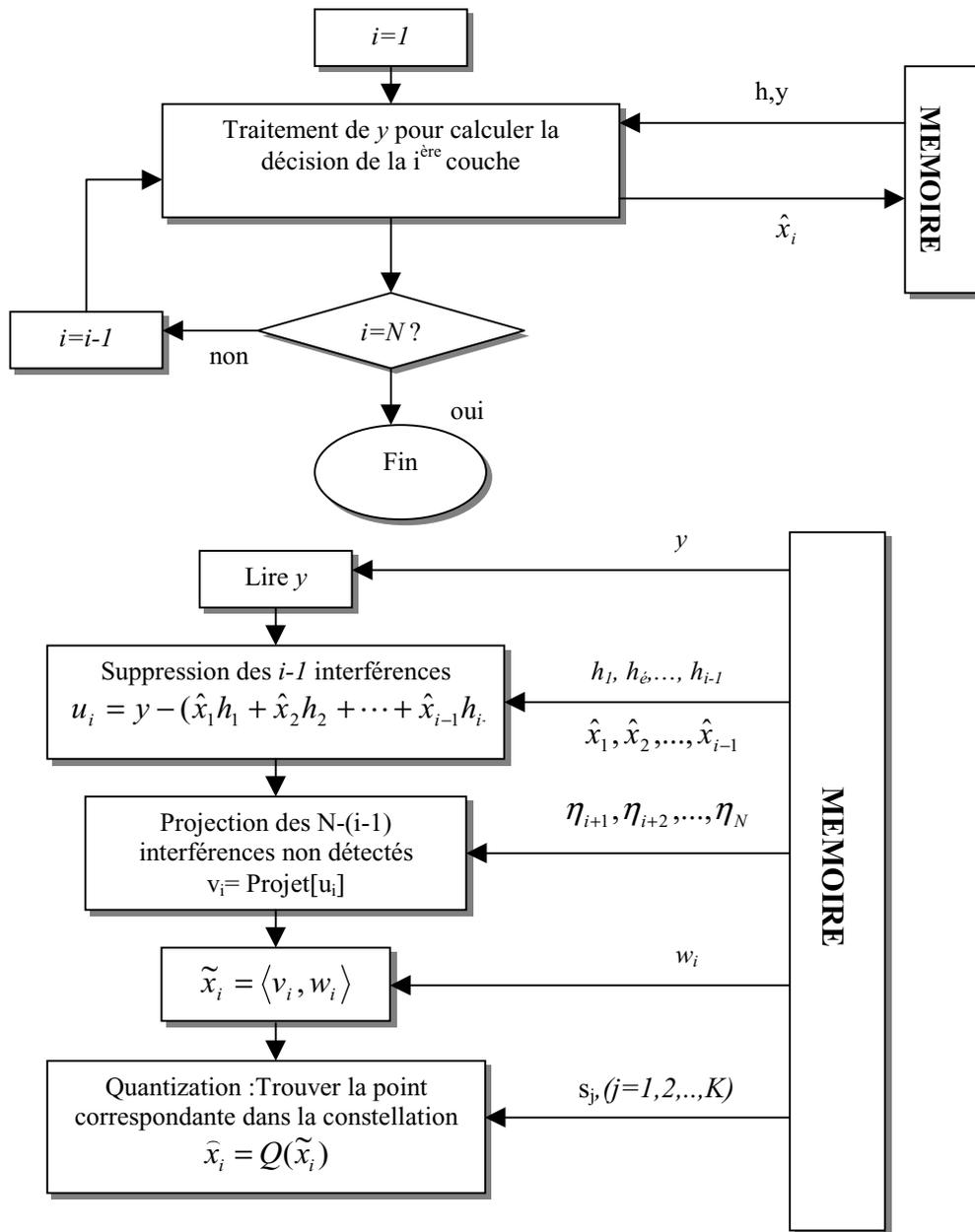


Figure 6 : Algorithme V-BLAST simplifié

### 3-5 Analyse de l'algorithme

Pour plus de simplicité, nous décrivons d'abord l'étape sans bruit ( $\mathbf{n} = \mathbf{0}$ ), puis nous tiendrons compte de celui-ci dans l'analyse suivante. On suppose que les  $(i-1)$  premiers symboles sont détectés sans erreurs. En mettant  $\mathbf{n} = \mathbf{0}$  dans l'équation ( 9 )

$$\mathbf{u}_i = \mathbf{y} - \sum_{j=1}^{i-1} \mathbf{x}_j h_j = \sum_{j=i}^N \mathbf{x}_j h_j$$

La forme analytique pour la projection de  $\mathbf{u}_i$  sur le sous-espace engendré par  $h_i, h_{i+1}, \dots, h_N$  est [7-9] :

$$\mathbf{u}_i = \frac{-1}{|\mathbf{R}^{[i+1,N]}|} \begin{vmatrix} & & & \theta_{i+1} \\ & & & \dots \\ & & & \theta_N \\ \langle \mathbf{u}_i \cdot \theta_{i+1}^* \rangle & \dots & \langle \mathbf{u}_i \cdot \theta_N^* \rangle & 0 \end{vmatrix}$$

où  $|\cdot|$  signifie le déterminant de matrice,  $\theta_j = h_j / \|h_j\|$ , et  $\mathbf{R}^{[i+1,N]}$  est la matrice de corrélation produit par  $(\theta_i, \dots, \theta_N)$  :

$$\mathbf{R}^{[i+1,N]} = \begin{bmatrix} 1 & \theta_{i+1,i+2} & \dots & \theta_{i+1,N} \\ \theta_{i+2,i+1} & 1 & \dots & \theta_{i+2,N} \\ \dots & \dots & \dots & \dots \\ \theta_{N,i+1} & \theta_{N,i+2} & \dots & 1 \end{bmatrix}$$

où  $\theta_{i,j} = \text{corre}(\theta_i, \theta_j) = \langle \theta_i \cdot \theta_j^* \rangle$  montre la corrélation entre  $\theta_i$  et  $\theta_j$ . La composante de  $\mathbf{u}_i$  orthogonale à  $(\theta_i, \dots, \theta_N)$  est  $\mathbf{v}_i$  :

$$\mathbf{v}_i = \mathbf{u}_{i\perp} = \mathbf{u}_i - \mathbf{u}_i = \frac{1}{|\mathbf{R}^{[i+1,N]}|} \begin{vmatrix} & & & \theta_{i+1} \\ & & & \dots \\ & & & \theta_N \\ \langle \mathbf{u}_i \cdot \theta_{i+1}^* \rangle & \dots & \langle \mathbf{u}_i \cdot \theta_N^* \rangle & \mathbf{u}_i \end{vmatrix} \quad (21)$$

En considérant :

$$\mathbf{u}_i = \sum_{j=i}^N \mathbf{x}_j h_j = \sum_{j=i}^N \mathbf{x}_j \|h_j\| \theta_j$$

Pour la dernière ligne de (21), nous obtenons la forme suivante :

$$\langle \mathbf{u}_i \cdot \theta_{i+n}^* \rangle = \sum_{j=i}^N \langle \mathbf{x}_j \|h_j\| \theta_j \cdot \theta_{i+n}^* \rangle = \sum_{j=i}^N \mathbf{x}_j \|h_j\| \langle \theta_j \cdot \theta_{i+n}^* \rangle = \sum_{j=i}^N \mathbf{x}_j \|h_j\| \theta_{j,i+n}$$

Cette ligne inclut des composantes proportionnelles à la première ligne ainsi que jusqu'à la  $(N-i)$ ème ligne qui s'annulent dans le calcul du déterminant. Ainsi le seul terme qui reste de (21) est  $\|h_j\| \mathbf{x}_i$ . Par conséquent (21) se réduit à :

$$\mathbf{v}_i = \frac{|x_i| |h_i|}{|\mathbf{R}^{[i+1,N]}|} \begin{vmatrix} & & & \theta_{i+1} \\ & \mathbf{R}^{[i+1,N]} & & \dots \\ & & & \theta_N \\ \theta_{i,i+1} & \dots & \theta_{i,N} & \theta_i \end{vmatrix}$$

En remplaçant la dernière ligne par la première ligne et la colonne de gauche par la colonne de droite on peut obtenir :

$$\mathbf{v}_i = \frac{|x_i| |h_i|}{|\mathbf{R}^{[i+1,N]}|} \begin{vmatrix} \theta_i & \theta_{i,i+1} & \dots & \theta_{i,N} \\ \theta_{i+1} & & & \\ \dots & & \mathbf{R}^{[i+1,N]} & \\ \theta_N & & & \end{vmatrix} \quad (22)$$

La puissance du signal est :

$$\|\mathbf{v}_i\|^2 = \langle \mathbf{v}_i \cdot \mathbf{v}_i \rangle$$

Nous ajoutons la valeur de  $\mathbf{v}_i$  de l'équation dans l'équation précédente on obtient :

$$\|\mathbf{v}_i\|^2 = \frac{|x_i|^2 |h_i|^2}{|\mathbf{R}^{[i+1,N]}|} \begin{vmatrix} \theta_i u_i & \theta_{i,i+1} & \dots & \theta_{i,N} \\ \theta_{i+1} u_i & & & \\ \dots & & \mathbf{R}^{[i+1,N]} & \\ \theta_N u_i & & & \end{vmatrix}$$

On note que la première colonne a des composants proportionnelles à la 2<sup>ème</sup> jusqu'au rang  $(N-i+1)$ , ces termes peuvent donc être soustraits. On arrive donc à la forme suivante :

$$\|\mathbf{v}_i\|^2 = \frac{|x_i|^2 |h_i|^2}{|\mathbf{R}^{[i+1,N]}|} \begin{vmatrix} 1 & \theta_{i,i+1} & \dots & \theta_{i,N} \\ \theta_{i+1,i} & & & \\ \dots & & \mathbf{R}^{[i+1,N]} & \\ \theta_{N,i} & & & \end{vmatrix} = |x_i|^2 |h_i|^2 \frac{|\mathbf{R}^{[i,N]}|}{|\mathbf{R}^{[i+1,N]}|}$$

Maintenant on considère l'effet du bruit en ajoutant celui-ci dans l'équation ( 21 ) :

$$\mathbf{v}_i = \mathbf{v}'_i + \mathbf{n}'_i$$

où  $\mathbf{v}'_i$  est le signal sans bruit, qui est obtenu à partir de l'équation ( 22 ). On peut alors généraliser ( 22 ) comme la formule suivante :

$$\mathbf{n}'_i = \frac{\mathbf{n}}{|\mathbf{R}^{[i+1,N]}|} \begin{vmatrix} 1 & \theta_{i+1} & \dots & \theta_N \\ \theta_{i+1} & & & \\ \dots & & \mathbf{R}^{[i+1,N]} & \\ \theta_N & & & \end{vmatrix}$$

La puissance de bruit après le traitement peut être exprimée par :

$$P_{n'_i} = E(|\mathbf{n}'_i|^2) = (M - N + i)\delta^2$$

où  $E(|\mathbf{n}'_i|^2)$  est l'espérance du bruit et  $\delta^2$  est la puissance du bruit avant le traitement.. Le *SNR* se calcule avec la formule suivante :

$$SNR_i = \frac{|\mathbf{x}_i|^2 |h_i|^2 |\mathbf{R}^{[i,N]}|}{(M - N + i) |\mathbf{R}^{[i+1,N]}|}$$

A partir de cette équation il est clair que la puissance totale du bruit est égale à  $M\delta^2$ , qui est donc supérieure à la puissance du bruit avant le traitement. C'est une des conséquences du processus de Gram-Schmidt. On remarque que la puissance du bruit augmente avec  $i$ . A la première étape elle est plus faible et à la dernière, elle est égale à la puissance totale.

### 3-5-1 La Diversité

Dans cette section nous prouvons que la diversité atteinte à la  $i^{\text{ème}}$  étape est égale à  $(M-N+i)$ . Sans perdre le caractère général de l'analyse, nous ignorons l'ordre de détection ainsi que le bruit additif. Pour l'instant on considère le cas le plus simple où  $N = M = 2$ , où  $\mathbf{H} = [h_1 \ h_2]$ . Nous supposons que l'on détecte d'abord le premier symbole transmis.

On peut montrer l'étape d'annulation d'interférence en considérant les équations (2)(7). La première étape consiste à calculer  $\tilde{\mathbf{r}}$  défini par (4)  $\tilde{\mathbf{y}} = \mathbf{Q}^\dagger \cdot \mathbf{y}$  où  $\mathbf{Q}^\dagger$  est une matrice qui projette  $\mathbf{r}$  dans le direction orthogonale à  $h_2$ , autrement dit,  $\tilde{\mathbf{y}} = \mathbf{y}_\perp$ . En remplaçant  $\mathbf{y}$  par l'équation (2) on peut arriver à la formule suivante :

$$\mathbf{y}_\perp = \mathbf{x}_1 \mathbf{Q}^\dagger \cdot h_1$$

Cela revient à dire que le signal après l'annulation d'interférence dépend uniquement de la partie de  $h_1$  qui est orthogonale à  $h_2$ , la puissance de signal est donc proportionnel à  $|h_{1\perp}|^2$ . La norme du vecteur ne change pas par rotation.

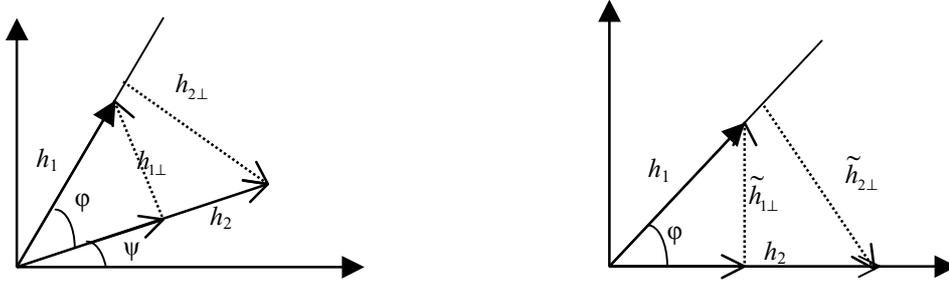


Figure 7 : Transformations des vecteurs reçus

On fait tourner  $[h_1 \ h_2]$  avec l'angle  $\psi$ ,  $h_2$  donc devient parallèle à  $e_2$  :  $h_{2\perp} = 0$ . On peut exprimer ceci sous la forme  $\tilde{h}_i = A \cdot h_i$  où  $A$  est la matrice de rotation, qui vérifie la propriété suivante :  $AA^+ = A^+A = I$

En utilisant l'équation précédente on peut obtenir :  $|h_{1\perp}| = |\tilde{h}_{1,1}|$ . On peut montrer que les composants de  $\tilde{h}_1$  ont la même distribution que  $h_1$  car l'angle  $\psi$  est indépendant de  $h_1$ . On se rappelle que les composants de  $h_i$  sont des variables aléatoires gaussiennes indépendantes de variance 0.5 pour chaque partie réelle et imaginaire. Donc  $\tilde{h}_1$  et également  $\tilde{h}_{1,1}$  sont gaussiens complexes.

Sachant que la somme carrée de  $N$  variables aléatoires gaussiennes  $(\delta, 0)$  sera une variable chi-squared avec  $n$  degré de liberté  $\chi_n^2$ . La pdf de  $y$  peut être obtenue sous la forme suivante :

$$y = \chi_N^2 = x_1^2 + x_2^2 + \dots + x_n^2$$

$$f_y(y) = \frac{1}{2^{n/2} \sigma^n \Gamma(n/2)} y^{(n-2)/2} e^{-y/2\sigma^2} \quad y \geq 0 \quad (23)$$

Par conséquent la distribution de  $|\tilde{h}_{1,1}|^2$  est chi-squared avec deux degré de liberté ainsi que  $|h_{1\perp}|^2 : |h_{1\perp}|^2 \sim \chi_2^2$ . A la première étape il n'y a donc aucune diversité. Dans le cas  $M > 2$  après la rotation,  $\tilde{h}_1$  aura  $M-1$  composants et donc  $|h_{1\perp}|^2$  sera la somme carrée de  $2(M-1)$  variables gaussiennes. Ainsi  $|h_{1\perp}|^2 \sim \chi_{2(M-1)}^2$  et la diversité est égale à  $M-2$ .

Pour le cas  $N > 2$ , on fait d'abord tourner ensemble de vecteurs  $[h_1 \ h_2 \dots \ h_N]$  afin que  $h_N$  soit orthogonal à  $e_N$ . A la deuxième étape, on garde la direction de  $h_N$ , et on fait tourner  $[h_1 \ h_2 \dots \ h_{N-1}]$  autour de  $h_N$ . Cette fois  $h_N$  est l'axe de rotation et on rend  $h_{N-1}$  parallèle à  $e_{N-1}$ . On continue la rotation jusqu'à ce que  $h_2$  soit placé sur  $[e_2 \ e_3 \dots \ e_N]$ .

A la fin de la rotation  $h_{1\perp}$  aura  $(M-N+1)$  composants non-nulles. Comme la rotation garde la distribution de la composante, donc  $|h_{1\perp}|^2 \sim \chi_{2(M-N+1)}^2$  avec  $2(M-N+1)$  degré de liberté, ce qui veut dire que la diversité est égale à  $(M-N)$ . Avec le même argument pour la  $i^{\text{ème}}$  étape on arrive à  $|h_{i\perp}|^2 \sim \chi_{2(M-N+i)}^2$ . On remarque qu'à la première étape ( $i = 1$ ) on aura la plus basse

diversité et qu'à la dernière étape ( $i = N$ ) elle sera la plus forte. Au cas où  $M = N$ , il n'y a aucune diversité à la première étape.

### **3-6 Choix de l'ordre optimal de détection**

Dans cette partie, pour la simplicité, nous supposons que tous les symboles transmises ont la puissance égale à un. Donc le vecteur reçu par antenne  $i$  de réception s'écrit  $r_i = h_i$ . On considère d'abord le cas  $2 \times N$  V-BLAST.

#### **Le cas $2 \times N$**

On considère le cas de la Figure 7 car la rotation ne change ni la norme ni l'angle entre deux vecteurs. Comme on l'a vu au sous-chapitre précédent, les vecteurs ayant tourné ont les mêmes distributions que les vecteurs sans rotation :  $|h_{1\perp}|^2 \sim \chi_{2N-2}^2$ ,  $|h_{1\parallel}|^2 \sim \chi_2^2$ ,  $|h_1|^2 \sim \chi_{2N}^2$ . Les composants de  $h_2$  et  $h_1$  ont alors les mêmes distributions. La fonction de répartition d'une variable de Rayleigh est :

$$F_y(x) = 1 - e^{-x} \sum_{k=0}^{N-1} \frac{x^k}{k!} \quad (24)$$

L'ordre optimal de détection peut se décrire sous la forme suivante :

$$s_1 = \max\{|h_{1\perp}|^2, |h_{2\perp}|^2\} = (\sin \phi)^2 \max\{|h_1|^2, |h_2|^2\}$$

où  $s_1$  est la puissance de signal après la première étape de détection. Pour choisir  $s_1$  on a deux possibilités,  $|h_{1\perp}|^2$  et  $|h_{2\perp}|^2$ , afin d'avoir l'ordre optimal on prend la valeur maximale entre les deux. En choisissant  $|h_{1\perp}|^2 = (\sin \phi)^2 |h_1|^2$  ou  $|h_{2\perp}|^2 = (\sin \phi)^2 |h_2|^2$ , on aura toujours le même ordonnancement. On peut donc laisser de côté le terme  $(\sin \phi)^2$  et on compare les distributions de  $\max\{|h_1|^2, |h_2|^2\}$ ,  $|h_1|^2$  et  $|h_2|^2$ . On calcule d'abord la probabilité de  $\max\{|h_1|^2, |h_2|^2\}$  :

$$\max\{|h_1|^2, |h_2|^2\} < x \Leftrightarrow |h_1|^2 < x, |h_2|^2 < x$$

donc on peut dire que :

$$\Pr\{\max\{|h_1|^2, |h_2|^2\} < x\} = \Pr\{|h_1|^2 < x, |h_2|^2 < x\}$$

Si  $|h_1|^2$  et  $|h_2|^2$  sont des variables indépendantes :

$$\Pr\{\max\{|h_1|^2, |h_2|^2\} < x\} = \Pr\{|h_1|^2 < x\} \cdot \Pr\{|h_2|^2 < x\} = F_h^2(x)$$

En considérant (20), le comportement asymptotique de  $F_h(x)$  est :

$$\lim_{x \rightarrow 0} [F_h(x)] = \lim_{x \rightarrow 0} \left[ 1 - e^{-x} \sum_{k=0}^{N-1} \frac{x^k}{k!} \right] \approx x^N / N! \quad (25)$$

Maintenant on passe à  $F_h^2$  :

$$\lim_{x \rightarrow 0} [F_h^2(x)] \approx x^{2N} / (N!)^2 \quad (26)$$

En fait, l'ordre optimal augmente le SNR et il baisse la probabilité d'erreur lors de la première étape.

Dans la seconde étape on va prendre le vecteur avec la largeur minimale, car le vecteur avec la largeur la plus grande est déjà choisi dans l'étape précédente. On remarque que dans le cas considéré, deux antennes d'émission, la seconde étape est la dernière, donc on ne fait pas l'annulation d'interférences. Ce qui revient à dire :

$$s_2 = \min[|h_1|^2, |h_2|^2]$$

donc on peut dire que :

$$\begin{aligned} F_2(x) &= \Pr\{s_2 < x\} = \Pr\{\min[|h_1|^2, |h_2|^2] < x\} \\ &= 1 - \Pr\{|h_1|^2 > x, |h_2|^2 > x\} \\ &= 1 - [(1 - F_h(x))(1 - F_h(x))] = F_h(x)[2 - F_h(x)] \end{aligned}$$

Le comportement asymptotique de  $F_2(x)$  est :

$$\lim_{x \rightarrow 0} [F_2(x)] \approx \lim_{x \rightarrow 0} [2F_h(x)] \approx 2x^N / N! \quad (27)$$

En comparant (21) et (23), on voit que la probabilité d'erreur lors de la seconde étape est 3dB plus haute par rapport au cas sans ordre optimal.

### 3-6-1 Capacité du canal

On a vu que le processus de détection V-BLAST se décompose en  $N$  parties pour chaque couche. La capacité pour la  $i^{\text{ème}}$  couche peut s'écrire sous la forme suivante :

$$C_i = \log_2 \left( 1 + \frac{1}{N |w_i|^2 \delta^2} \right)$$

Donc la capacité totale de système peut être obtenue par la somme des capacités de chaque couche :

$$C_{V-BLAST} = \sum_{i=1}^N C_i = \sum_{i=1}^N \log_2 \left( 1 + \frac{1}{N |w_i|^2 \delta^2} \right)$$

### 3-6-2 Exemples de performances

Les exemples de performances donnés ci-dessous sont calculés pour la technique V-BLAST dans le cas d'une modulation MDP-4 sans codage de canal et avec un nombre d'antennes d'émission et de réception variable et en supposant le canal parfaitement estimé. Le premier chiffre correspond au nombre d'antennes d'émission et le second au nombre d'antennes de réception. Malgré des capacités théoriques de canal plus élevées, on constate par exemple que la technique V-BLAST donne des performances moins élevées pour un système (5-5) que

pour un système (2,2). Ceci montre bien la sensibilité de ce type de détection aux interférences et justifie de la nécessité de codage spatio-temporel pour mieux exploiter cette diversité.

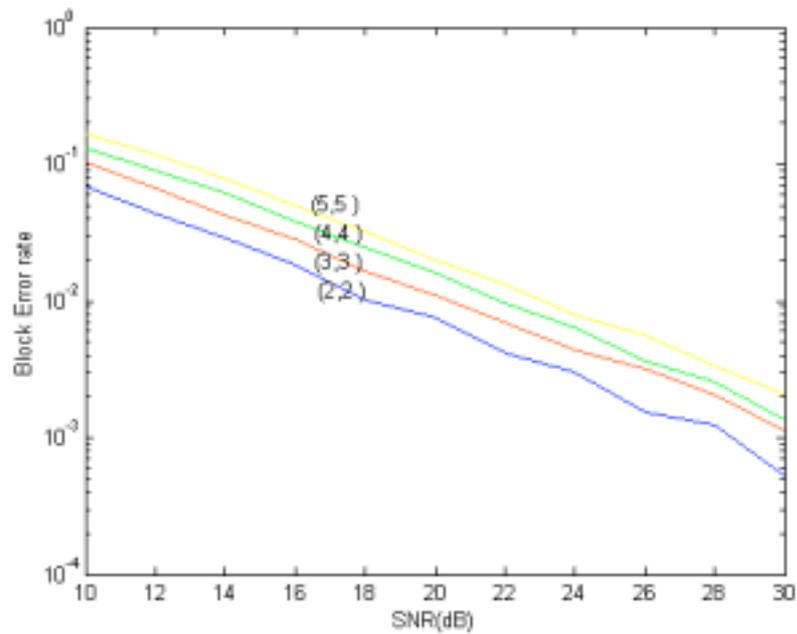


Figure 8. Taux d'erreur de Blocs pour différents nombre d'antennes  $(N,M)=(2,2),(3,3),(4,4)$  et  $(5,5)$

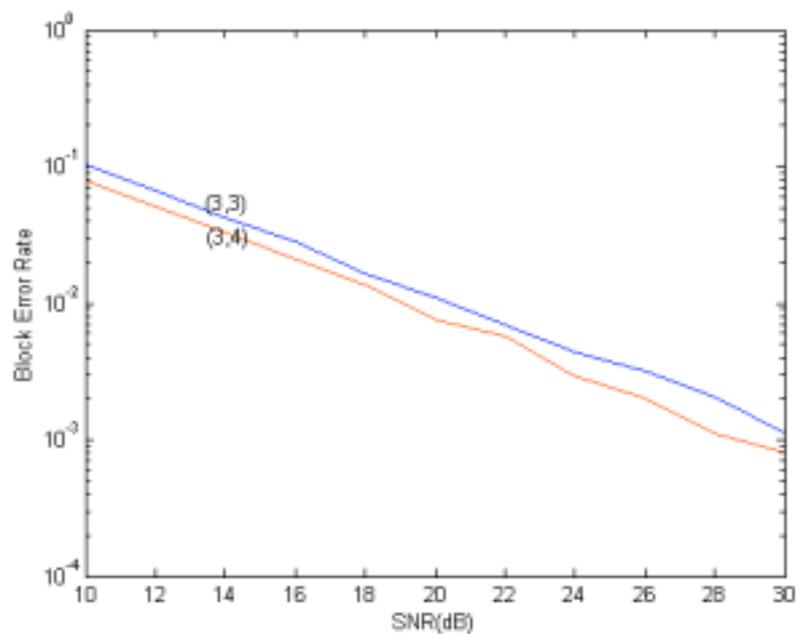
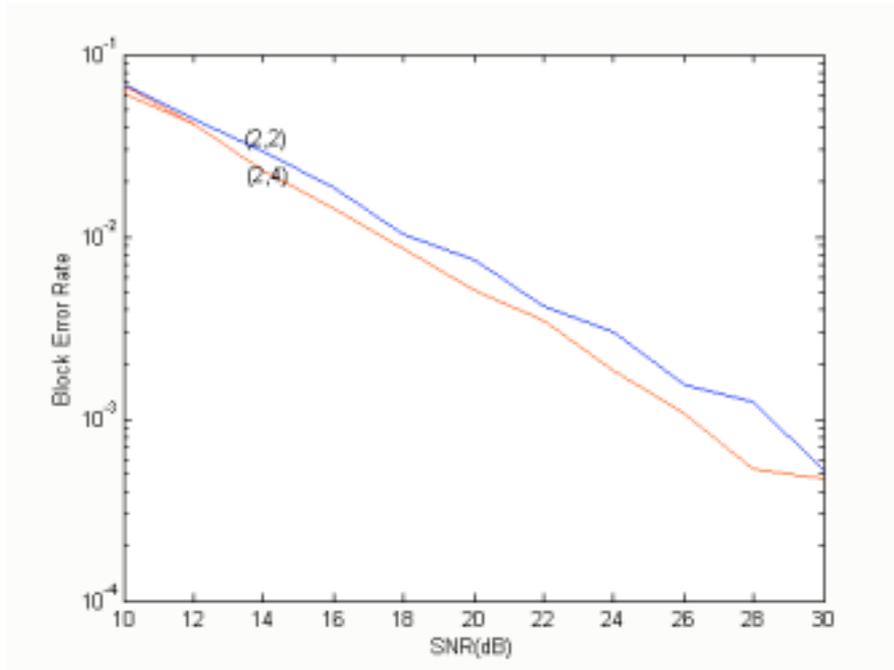


Figure 9. Taux d'erreur de Blocs pour différents nombres d'antennes de réception pour un nombre d'antennes d'émission fixé à 3.



**Figure 10. Taux d'erreur de Blocs pour différents nombres d'antennes de réception pour un nombre d'antennes d'émission fixé à 2.**

Les Figure 9 et Figure 10 mettent bien en évidence l'amélioration de la qualité de détection lorsque  $M - N$  augmente, l'effet de diversité pour les premiers symboles décodés augmentant de façon conséquente.

### **Conclusion**

Nous avons abordé dans ce chapitre une technique de détection sous-optimale pour les systèmes MIMO, à savoir la technique V-BLAST. Cette technique qui se base sur une élimination successive des interférences par ordre d'énergie décroissante ne permet cependant pas d'utiliser toute la diversité spatiale d'un canal MIMO. En effet, les performances obtenues ne s'améliorent pas avec le nombre d'antennes d'émission et de réception contrairement à la capacité théorique du canal. On trouve ici une justification à l'utilisation des codes spatio-temporels.

### **3-7 Références**

- [1] Foschini, G.J “Layered space-time architecture for wireless communication in a fading environment when using multiple antennas”, Bell Lab. Tech. J.,vol. 1,N. 2, pp. 41-59, 1996.
- [2] Wolniansky, P.W.; Foschini, G.J.; Golden, G.D.; R.A. Valenzuela, R.A.; “V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel”, *ISSSE 1998, URSI International Symposium*, pp. 295 –300
- [3] G.J. Foschini, G.D. Golden, R.A. Valenzuela and P.W. Wolniansky, “Simplified processing for high spectral efficiency communication employing multi-element arrays”, *IEEE Journal on Selected Areas in Communications*, vol. 17, n° 11, pp.1841-1852, Nov. 1999.
- [4] K. Cavalec-Amis, “Optimisation des Turbo Codes pour les systèmes à grande efficacité spectrale”, Ph.D Dissertation, Université de Rennes 1, Brest, France Dec. 2001.
- [5] Won-Joon Choi; Negi, R.; Cioffi, J.M “Combined ML and DFE decoding for the V-BLAST system”. ICC 2000, vol. 3 , pp. 1243 –1248.
- [6] Rudolf K. Bock, 7 April 1998 , <http://rkb.home.cern.ch/rkb/AN16pp/node224.html>
- [7] Foschini G.J., Golden G.D, Valenzuela R.A., Wolniansky P.W., “Simplified processing for high spectral efficiency wireless communication employing multi-element arrays” *IEEE Journal on Selected Areas in Communications*, vol 17, issue 11, Nov. 1999 pp. 1841 –1852.
- [8] Gene H. Golub and Charles F. van Loan: *Matrix Computations*, 2nd edn., The John Hopkins University Press, 1989
- [9] Loyka S., Gagnon F., “Performance analysis of the V-BLAST algorithm: an analytical approach”, *International Zurich Seminar on Broadband Communications, Access, Transmission, Networking*. 2002, Feb. 2002 Page(s): 5-1 -5-6.



## *Annexe : Programme V-BLAST*

```

function a_hat=vblast(r_rec,H,ss);
%*****
%*Entrées: r_rec: les signaux reçus          *
%*      : H   : la matrice du canal         *
%*      : ss  : les points de constellation *
%*Sorties :a_hat : les vecteur de détection *
%*****
[temp,L_total]=size(r_rec);
inf=1e5;
[temp,Q]=size(ss);
[m,n]=size(H);
%n: antennes de réceptions
%m: antennes d'émissions
Hk_=H;
Gi=Ps_inv(Hk_);
sort_index=sort_min_index(Gi);% l'ordre optimal de détection
new_index=sort_index;
for i=1:n
    ki=new_index(i);
    Wk(i,:)=Gi(ki,:);
    ii=1;
    HH=zeros(m,n-i);
    for ih=1:n-i+1
        if ~(ih==ki)
            HH(:,ii)=Hk_(:,ih);
            ii=ii+1;
        end
    end
    Hk_=HH;
    clear HH
    Gi=Ps_inv(Hk_);
    for si=i:n
        if new_index(si)>ki
            new_index(si)=new_index(si)-1;
        end
    end;
end
a_hat=zeros(1,n*L_total);
for t=1:L_total
    Ri=r_rec(:,t);
    for i=1:n
        ki=sort_index(i);
        Yk=Wk(i,:)*(Ri);% annulation d'interférence
        metric=inf;
        am_hat=ss(1);
        for q=1:Q
            new_metric=abs(Yk-ss(q));
            if new_metric<metric
                metric=new_metric;
                a_hat(ki+((t-1)*n))=q-1;
                am_hat=ss(q);
            end
        end
        ht=H(:,ki);
        Ri=Ri-(am_hat*ht);% suppression d'interférence
    end
end

```

end

# **CHAPITRE 4**

**DECODAGE DES CODES STBC  
ORTHOGONAUX**

**ALGORITHME SYSTEMATIQUE DE  
DETECTION & DERIVATION DES BORNES  
SUPERIEURES PRECISES**



## **4- Décodage des Codes STBC Orthogonaux**

### **Algorithme Systématique de Détection & Dérivation des Bornes Supérieures Précises**

#### **4-1 Introduction**

La diversité de transmission a été largement étudiée comme une méthode efficace pour les canaux à fading [1][2][3][4][5]. Elle est particulièrement intéressante pour sa simplicité relative d'implantation et la faisabilité des antennes multiples à la station de base. STTC (Space-Time Trellis Coding) combine le codage de canal de type TCM avec les techniques appropriées aux antennes multiples à l'émetteur. Dans [6][7][8], les auteurs atteignent une très bonne performance pour 2 à 4 antennes d'émission et pour les canaux slow-fading en s'approchant de la capacité d'outage du canal. Cependant, le désavantage de tels codes est la complexité de l'algorithme de décodage qui augmente de manière exponentielle par rapport au taux de transmission quand le nombre d'antennes d'émission est fixe.

Pour éviter la complexité prohibitive de STTC, les schémas de la simple diversité de transmission basés sur la conception de matrices orthogonales ont été proposés [9][10][11]. L'utilisation d'un tel schéma de diversité, en combinant avec un code FEC (Forward Error Correction) qui peut être de type TCM afin de pouvoir comparer les deux cas, aboutit à une performance presque similaire obtenue pour STTC. La construction des matrices orthogonales aboutit aux techniques de codage STBC, qui pourront être planifiées, par exemple, pour la nouvelle génération des systèmes radio-mobiles cellulaires.

Dans ce chapitre, on propose, d'abord une structure générale de réception pour le décodage ML des codes STBC. Utilisant les outils algébriques, on propose donc un modèle matriciel particulièrement approprié à la description des systèmes MIMO. Le modèle obtenu a l'avantage de donner un simple algorithme de décodage ML, basé sur un traitement linéaire au récepteur. Grâce à ce modèle matriciel, on va trouver une nouvelle borne supérieure très précise pour le TEB du code STBC. Ces résultats sont finalement vérifiés avec une simulation Monte-Carlo.

## 4-2 Description du Système

La Figure 1 illustre un système multi-antenne de type STBC. Il y a donc  $n$  antennes d'émission,  $m$  antennes de réception et un rendement de  $k/l$  qui décrit la redondance du code. Pour un paquet de  $k$  symboles d'information en entrée du codeur, on aura  $l$  symboles sur chaque antenne d'émission. Ces  $l$  symboles peuvent être une combinaison linéaire des symboles d'entrée  $(x_1, x_2, \dots, x_k)$  et de leur conjugué.  $C_t^i$  représente le symbole émis par l'antenne d'émission  $i$  dan le temps  $t$ . Le coefficient  $\alpha_{i,j}$  est le gain de trajet entre l'antenne d'émission  $i$  et l'antenne de réception  $j$  ; ces atténuations du canal sont modélisées comme des variables aléatoires gaussiennes complexes et indépendantes avec une variance de 0,5 pour les parties réelle et imaginaire.

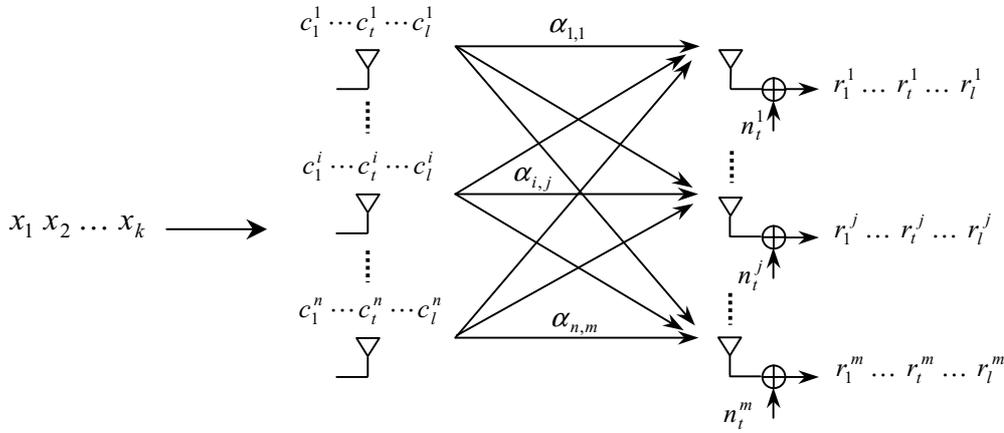


Figure 1 : Modèle du système MIMO-STBC

Dans le temps  $t$ , le signal  $r_t^j$  reçu par l'antenne  $j$  s'écrit sous la forme suivante :

$$r_t^j = \sum_{i=1}^n \alpha_{i,j} c_t^i + n_t^j \quad (1)$$

où  $n_t^j$  est le bruit gaussien complexe et additif sur l'antenne  $j$  pour le temps  $t$  ; les  $n_t^j$  sont les variables aléatoires gaussiennes complexes et indépendantes à moyenne nulle et variance  $\sigma^2/2 = 1/(2\text{SNR})$  par dimension complexe :  $E\{|n_t^j|^2\} = \sigma^2$ . Pour obtenir la totalité d'énergie émise égale à 1, l'énergie moyenne des symboles transmis par chaque antenne est normalisée par  $1/n$ . C'est à dire que la totalité d'énergie sur toutes les antennes d'émission est égale à 1. Ainsi une comparaison judicieuse par rapport à un système monoantenne est possible.

Afin d'avoir une forme matricielle pour l'équation ( 1 ), on définit le vecteur  $\mathbf{r}^j = (r_1^j \ \dots \ r_t^j \ \dots \ r_l^j)$  représentant le signal reçu sur l'antenne  $j$ . On arrive ainsi à l'équation suivante :

$$\mathbf{r}^j = \underbrace{\begin{bmatrix} \alpha_{1,j} & \cdots & \alpha_{i,j} & \cdots & \alpha_{n,j} \end{bmatrix}}_{\mathbf{h}^j} \underbrace{\begin{bmatrix} c_1^1 & \cdots & c_t^1 & \cdots & c_l^1 \\ \vdots & & & & \vdots \\ c_1^i & & c_t^i & & c_l^i \\ \vdots & & & & \vdots \\ c_1^n & \cdots & c_t^n & \cdots & c_l^n \end{bmatrix}}_{\mathbf{C}} + \mathbf{n}^j = \mathbf{h}^j \mathbf{C} + \mathbf{n}^j \quad (2)$$

où le vecteur  $\mathbf{h}^j$  contient les gains des trajets reçus par l'antenne  $j$ ,  $\mathbf{C}$  est la matrice de codes en bloc contenant les symboles émis par toutes les antennes et pour tous les temps d'un paquet et  $\mathbf{n}^j = (n_1^j \ \cdots \ n_t^j \ \cdots \ n_l^j)$  est le vecteur de bruit avec la matrice d'auto-corrélation suivante :

$$E\left\{(\mathbf{n}^j)^\dagger \mathbf{n}^j\right\} = \sigma^2 \mathbf{I}_{l \times l}$$

Maintenant en définissant la matrice  $\mathbf{r}$  sous la forme suivante, on trouve une forme matricielle pour tous les signaux reçus par toutes les antennes.

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}^1 \\ \vdots \\ \mathbf{r}^j \\ \vdots \\ \mathbf{r}^m \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{h}^1 \\ \vdots \\ \mathbf{h}^j \\ \vdots \\ \mathbf{h}^m \end{bmatrix}}_{\mathbf{H}} \cdot \mathbf{C} + \underbrace{\begin{bmatrix} \mathbf{n}^1 \\ \vdots \\ \mathbf{n}^j \\ \vdots \\ \mathbf{n}^m \end{bmatrix}}_{\mathbf{n}} = \mathbf{H} \cdot \mathbf{C} + \mathbf{n} \quad (3)$$

Supposons que les informations du canal sont disponibles, le récepteur ML minimise la métrique de décision ci-dessous sur tous les mots du code  $\mathbf{C}$  :

$$\text{ML : } \underset{\mathbf{C}}{\text{Min}} \ \Omega = \sum_{j=1}^m \sum_{t=1}^l \left| r_t^j - \sum_{i=1}^n \alpha_{i,j} c_t^i \right|^2 \quad (4)$$

En considérant les définitions matricielles présentées ci-dessus, on peut réécrire la métrique  $\Omega$  sous la forme suivante :

$$\begin{aligned} \text{Min} \ \Omega &= \sum_{j=1}^m (\mathbf{r}^j - \mathbf{h}^j \mathbf{C})(\mathbf{r}^j - \mathbf{h}^j \mathbf{C})^\dagger \\ \text{Min} \ \sum_{j=1}^m & \left( -\mathbf{r}^j \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger - \mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger + \mathbf{h}^j \mathbf{C} \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger \right) \end{aligned} \quad (5)$$

Pour pouvoir simplifier cette relation, il faut connaître le code et ses propriétés. Cependant, il est toujours possible de décoder en calculant la métrique de l'équation (5) pour tous les mots possibles du code et choisir celui qui minimise cette équation. Par exemple, pour une modulation M-aire et pour les séquences de  $k$  symboles en entrée du codeur, la métrique ci-dessus est calculée  $M^k$  fois. Le code qui donne la métrique minimale parmi toutes ces métriques, est annoncé en tant que la séquence décodée suivant le critère ML.

### **4-3 Codes STBC Orthogonaux**

Dans ce paragraphe, on cherche à simplifier l'équation ( 5 ) en utilisant une matrice de code orthogonale. On considère deux cas séparément :

- Le cas du code complexe où la constellation des symboles envoyés est défini dans le domaine complexe comme des modulation MPSK ou QAM.
- Le cas du code réel où la constellation des symboles envoyés est définie uniquement sur l'axe réel comme les modulation PAM. C'est en fait un cas particulier des codes complexes.

La matrice de code  $C_{n \times l}$  (complexe en général) est de taille  $n$  et de rendement  $k/l$ . Le code est orthogonal si les conditions suivantes sont satisfaites :

1. Les éléments de la matrice  $C$  sont les combinaisons linéaires de  $k$  symboles d'entrées  $(x_1, x_2, \dots, x_k)$  et de leur conjugué.
2.  $CC^\dagger = D_c$  où  $D_c$  est une matrice diagonale avec les éléments  $(i,i)$  de la forme suivante :

$$\left( l_1^i |x_1|^2 + l_2^i |x_2|^2 + \dots + l_k^i |x_k|^2 \right)$$

et les coefficients  $l_1^i, l_2^i, \dots, l_k^i$  sont strictement positifs. Il est clair que pour cela, il faut absolument que  $l \geq n$ . Car le rang de la matrice  $C$  est inférieur ou égal à  $\min(n,l)$  et le fait que  $CC^\dagger$  est une matrice diagonale de taille  $n \times n$  avec les éléments de diagonal non nuls, nécessite que le rang de  $C$  soit égal à  $n$  et par conséquent  $n \leq l$ .

Tarokh a montré [11] que s'il existe un code complexe et orthogonal  $n \times l$  pour les variables  $(x_1, x_2, \dots, x_k)$ , il existera un code complexe et orthogonal  $C$  de la même taille et pour les mêmes variables telle que :

$$CC^\dagger = \left( |x_1|^2 + |x_2|^2 + \dots + |x_k|^2 \right) I_{n \times n}$$

Dans la définition du code orthogonal, on peut donc remplacer la deuxième condition par cette dernière relation.

Pour la première condition et dans la plupart des cas, on peut trouver un code orthogonal dont les éléments sont

$$0, \pm x_1, \pm x_1^*, \pm x_2, \pm x_2^*, \dots, \pm x_k, \pm x_k^*$$

et non pas la combinaison linéaire des  $x_i$  et  $x_i^*$ . Nous développons ci-dessous le cas particulier du code réel .

#### **Cas des codes réels :**

Pour un code réel et orthogonal,  $C$  est une matrice réelle dont les éléments sont les combinaisons linéaires des symboles d'entrées  $(x_1, x_2, \dots, x_k)$ . Les  $x_k$  sont choisis dans une constellation réelle, ce qui est le cas des modulations PAM. Dans ce cas  $C^\dagger = C^T$  et pour un code réel et orthogonal, on aura :

$$CC^T = (x_1^2 + x_2^2 + \dots + x_k^2) \mathbf{I}_{n \times n}$$

Dans la majorité des cas, on peut trouver le code orthogonal réel dont les éléments sont

$$0, \pm x_1, \pm x_2, \dots, \pm x_k$$

### Exemples des codes orthogonaux:

Quel que soit le type du code (complexe ou réel), la définition du code orthogonal que l'on vient de présenter n'impose aucune condition sur les éléments de la matrice  $C$ , et cette définition peut comprendre les exemples ci-dessous.

**Codes réels :**

$$\begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix} \quad \begin{pmatrix} x_1 & -x_2 & -x_3 & -x_4 \\ x_2 & x_1 & x_4 & -x_3 \\ x_3 & -x_4 & x_1 & x_2 \end{pmatrix} \quad \begin{pmatrix} x_1 & -x_2 & -x_3 & -x_4 \\ x_2 & x_1 & x_4 & -x_3 \\ x_3 & -x_4 & x_1 & x_2 \\ x_4 & x_3 & -x_2 & x_1 \end{pmatrix}$$

**Codes complexes :**

$$\begin{pmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{pmatrix}$$

$$\begin{pmatrix} x_1 & -x_2^* & \frac{x_3^*}{\sqrt{2}} & \frac{x_3^*}{\sqrt{2}} \\ x_2 & x_1^* & \frac{x_3^*}{\sqrt{2}} & -\frac{x_3^*}{\sqrt{2}} \\ \frac{x_3}{\sqrt{2}} & \frac{x_3}{\sqrt{2}} & \frac{(-x_1 - x_1^* + x_2 - x_2^*)}{2} & \frac{(x_1 - x_1^* + x_2 + x_2^*)}{2} \end{pmatrix}$$

$$\begin{pmatrix} x_1 & -x_2 & -x_3 & -x_4 & x_1^* & -x_2^* & -x_3^* & -x_4^* \\ x_2 & x_1 & x_4 & -x_3 & x_2^* & x_1^* & x_4^* & -x_3^* \\ x_3 & -x_4 & x_1 & x_2 & x_3^* & -x_4^* & x_1^* & x_2^* \\ x_4 & x_3 & -x_2 & x_1 & x_4^* & x_3^* & -x_2^* & x_1^* \end{pmatrix}$$

#### **4-3-1 Développement et mise en équation de la matrice du code**

D'après la définition du code orthogonal, la matrice  $C$  peut s'écrire comme la somme de deux termes. Le premier ne contient que les  $x_i$  et le second les  $x_i^*$ . La matrice  $C$  peut s'écrire alors sous la forme :

$$\mathbf{C}_{n \times l} = (x_1 \mathbf{A}_1 + \dots + x_k \mathbf{A}_k) + (x_1^* \mathbf{B}_1 + \dots + x_k^* \mathbf{B}_k) = \sum_{i=1}^k (x_i \mathbf{A}_i + x_i^* \mathbf{B}_i) \quad (6)$$

où les matrices  $\mathbf{A}_i$  et  $\mathbf{B}_i$  sont les matrices constantes  $n \times l$  et avec les éléments  $\{0, \pm 1\}$  (nous avons supposé que les éléments du  $\mathbf{C}$  ne sont que des  $x_i$  et  $x_i^*$ , qui est vrai dans la majorité des cas). Pour vérifier l'orthogonalité, ces matrices doivent satisfaire les conditions suivantes (preuve dans l'annexe A):

$$\begin{cases} \mathbf{A}_i \mathbf{A}_i^\dagger + \mathbf{B}_i \mathbf{B}_i^\dagger = \mathbf{I}_{n \times n} \\ \mathbf{A}_i \mathbf{A}_j^\dagger + \mathbf{B}_j \mathbf{B}_i^\dagger = \mathbf{O}_{n \times n} \\ \mathbf{A}_i \mathbf{B}_i^\dagger = \mathbf{O}_{n \times n} \\ \mathbf{A}_i \mathbf{B}_j^\dagger + \mathbf{A}_j \mathbf{B}_i^\dagger = \mathbf{O}_{n \times n} \end{cases} \quad \text{pour } i, j = 1, \dots, k \text{ et } i \neq j \quad (7)$$

Pour les codes orthogonaux réels les matrices  $\mathbf{B}_i = \mathbf{O}_{n \times l}$  et les matrices  $\mathbf{A}_i$  sont telles que (preuve dans l'annexe A):

$$\begin{cases} \mathbf{A}_i \mathbf{A}_i^T = \mathbf{I}_{n \times n} \\ \mathbf{A}_i \mathbf{A}_j^T = -\mathbf{A}_j \mathbf{A}_i^T \end{cases} \quad \text{pour } i, j = 1, \dots, k \text{ et } i \neq j \quad (8)$$

Nous utiliserons ces propriétés des matrices A et B pour le décodage.

#### 4-4 Décodage ML des codes STBC orthogonaux

Pour expliquer l'algorithme de décodage on revient à l'équation ( 5 )

$$\text{Min } \Omega = - \underbrace{\sum_{j=1}^m \left( \mathbf{r}^j \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger + \mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger \right)}_{\Omega_1} + \underbrace{\sum_{j=1}^m \mathbf{h}^j \mathbf{C} \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger}_{\Omega_2}$$

Dans cette expression  $\Omega$  est un scalaire réel, les deux premiers termes ( $\Omega_1$ ) sont linéaires par rapport aux symboles envoyés et le troisième terme ( $\Omega_2$ ) est un polynôme de deuxième degré qui contient les termes  $|x_i|^2, x_i x_j^*, x_i x_j, x_i^* x_j^*$ .

L'idée du code orthogonal est la suivante: si on pouvait minimiser le  $\Omega$ , symbole par symbole et non pas sur la séquence entière de taille  $l$  symboles, l'algorithme de décodage se simplifierait énormément. Si le code  $\mathbf{C}$  est de telle sorte que dans  $\Omega_2$ , il ne reste que les termes  $|x_i|^2$  et non pas les termes croisés  $x_i x_j$ , la minimisation de  $\Omega$  (ou  $\Omega_2$ ) sur  $k$  symboles envoyés peut être remplacée par la minimisation de chacune des  $k$  métriques symbole par symbole. Ceci peut être obtenu si le code est orthogonal. Par la suite, on ne considère que les codes orthogonaux  $\mathbf{C}_{n \times l}$  et on se concentre sur le terme  $\Omega_2$ .

$$\begin{aligned} \Omega_2 &= \sum_{j=1}^m \mathbf{h}^j \mathbf{C} \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger = \sum_{j=1}^m \mathbf{h}^j \left( \sum_{q=1}^k |x_q|^2 \right) \mathbf{I} (\mathbf{h}^j)^\dagger = \sum_{q=1}^k |x_q|^2 \cdot \sum_{j=1}^m \mathbf{h}^j (\mathbf{h}^j)^\dagger \\ &= \sum_{q=1}^k |x_q|^2 \cdot \sum_{j=1}^m (\alpha_{1,j} \quad \dots \quad \alpha_{n,j}) \cdot \begin{pmatrix} \alpha_{1,j}^* \\ \vdots \\ \alpha_{n,j}^* \end{pmatrix} = \sum_{q=1}^k |x_q|^2 \cdot \left( \sum_{i,j} |\alpha_{i,j}|^2 \right) \end{aligned} \quad (9)$$

Maintenant on passe au terme  $\Omega_1$ , étant donné la linéarité de  $\Omega_1$  par rapport à la matrice  $\mathbf{C}$  dont chaque élément est une combinaison linéaire de  $x_1, \dots, x_k$  (et leur conjugué) et en notant que  $\Omega_1$  est un scalaire réel, on pourra dire que :

$$\Omega_1 = \sum_{j=1}^m - \left( \mathbf{r}^j \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger + \mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger \right) = \sum_{q=1}^k - \left( p_q^* x_q + p_q x_q^* \right) \quad (10)$$

où  $p_q$  est un scalaire complexe. Nous aborderons le calcul des  $p_q$  plus tard et on se concentre sur la métrique à minimiser. Utilisant les équations ( 9 ) et ( 10 ), on obtient :

$$\begin{aligned} \Omega &= \sum_{q=1}^k \Omega_q \\ \text{où } \Omega_q &= \left( \sum_{i,j} |\alpha_{i,j}|^2 \right) |x_q|^2 - \left( p_q^* x_q + p_q x_q^* \right). \end{aligned} \quad (11)$$

Comme  $\Omega_q$  ne dépend que de  $x_q$ , minimiser  $\Omega$  est équivalent à minimiser toutes les sous-métriques  $\Omega_q$ .

$$\text{Min}_{(x_1, \dots, x_k)} \Omega \Leftrightarrow \text{Min}_{x_q} \Omega_q ; \forall q = 1, \dots, k$$

Grâce à cette propriété, l'algorithme de décodage s'effectuera beaucoup plus simplement : pour une modulation M-aire la complexité est dimensionnée par  $kM$  et non plus par  $M^k$ . Par la suite, on se concentre sur la minimisation des sous-métriques  $\Omega_q$ . Utilisant l'équation ( 11 ) on peut écrire :

$$\begin{aligned} \text{Min}_{x_q} \Omega_q &= \left( \underbrace{\sum_{i,j} |\alpha_{i,j}|^2}_{A>0} \right) |x_q|^2 - p_q^* x_q - p_q x_q^* = A|x_q|^2 - p_q^* x_q - p_q x_q^* \\ &= A \left( |x_q|^2 - \frac{p_q^*}{A} x_q - \frac{p_q}{A} x_q^* + \frac{|p_q|^2}{A^2} \right) - \frac{|p_q|^2}{A} \quad (12) \\ \text{Min}_{x_q} \Omega_q &\equiv \text{Min}_{x_q} \left| x_q - \frac{p_q}{A} \right|^2 = |x_q - p'_q|^2 ; \forall q = 1, 2, \dots, k \end{aligned}$$

Dans l'expression ( 12 ),  $A$  est un scalaire réel qui dépend des gains du canal, et  $p_q$  se calcule à partir des signaux reçus par toutes les antennes de réception et les informations du canal. on appelle  $p'_q = \frac{p_q}{A}$  le paramètre de décision du symbole  $q$ . Ainsi, le décodage se simplifie et consiste à trouver pour chaque symbole, le point de constellation le plus proche (en terme de la distance euclidienne) au paramètre de décision de ce symbole.

Le résultat obtenu nous permet d'avoir non seulement une décision sur les données envoyées, mais aussi une mesure sur la fiabilité de ce choix : plus le paramètre de décision est proche d'un point de constellation, plus le symbole détecté est fiable. On peut donc envisager des systèmes mixtes TCM-STBC avec un décodage type "turbo".

#### **4-4-1 Calcul des paramètres de décision**

Dans ce paragraphe nous calculons les paramètres de décision  $p_q$ . On revient à l'équation ( 10 ). Dans cette équation la somme est composée de deux termes l'un conjugué de l'autre. On commence par le deuxième terme et on se focalise sur  $\mathbf{h}^j \mathbf{C}$ . Utilisant l'équation ( 6 ) on peut écrire :

$$\begin{aligned}
\mathbf{h}^j \mathbf{C} &= \mathbf{h}^j \cdot \sum_{i=1}^k (x_i \mathbf{A}_i + x_i^* \mathbf{B}_i) = \sum_{i=1}^k x_i \mathbf{h}^j \mathbf{A}_i + \sum_{i=1}^k x_i^* \mathbf{h}^j \mathbf{B}_i \\
&= \underbrace{(x_1 \quad \cdots \quad x_i \quad \cdots \quad x_k)}_{\mathbf{x}} \cdot \underbrace{\begin{pmatrix} \mathbf{h}^j \mathbf{A}_1 \\ \vdots \\ \mathbf{h}^j \mathbf{A}_i \\ \vdots \\ \mathbf{h}^j \mathbf{A}_k \end{pmatrix}}_{\mathbf{H}_R^j} + \underbrace{(x_1^* \quad \cdots \quad x_i^* \quad \cdots \quad x_k^*)}_{\mathbf{x}^*} \cdot \underbrace{\begin{pmatrix} \mathbf{h}^j \mathbf{B}_1 \\ \vdots \\ \mathbf{h}^j \mathbf{B}_i \\ \vdots \\ \mathbf{h}^j \mathbf{B}_k \end{pmatrix}}_{\mathbf{H}_I^j} \quad (13)
\end{aligned}$$

$$\mathbf{h}^j \mathbf{C} = \mathbf{x} \cdot \mathbf{H}_R^j + \mathbf{x}^* \cdot \mathbf{H}_I^j$$

On remplace  $\mathbf{h}^j \mathbf{C}$  ainsi obtenu dans l'équation ( 10 ) on obtient :

$$\begin{aligned}
\sum_{j=1}^m \mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger &= \sum_{j=1}^m (\mathbf{x} \cdot \mathbf{H}_R^j + \mathbf{x}^* \cdot \mathbf{H}_I^j) (\mathbf{r}^j)^\dagger = \mathbf{x} \cdot \sum_{j=1}^m \mathbf{H}_R^j (\mathbf{r}^j)^\dagger + \mathbf{x}^* \cdot \sum_{j=1}^m \mathbf{H}_I^j (\mathbf{r}^j)^\dagger \\
&= \mathbf{x} \cdot \mathbf{H}_R \cdot \mathbf{R}^\dagger + \mathbf{x}^* \cdot \mathbf{H}_I \cdot \mathbf{R}^\dagger \quad (14)
\end{aligned}$$

où  $\mathbf{H}_R = (\mathbf{H}_R^1 \quad \cdots \quad \mathbf{H}_R^j \quad \cdots \quad \mathbf{H}_R^m)_{k \times ml}$ ,  $\mathbf{H}_I = (\mathbf{H}_I^1 \quad \cdots \quad \mathbf{H}_I^j \quad \cdots \quad \mathbf{H}_I^m)_{k \times ml}$  et  $\mathbf{R} = (\mathbf{r}^1 \quad \cdots \quad \mathbf{r}^j \quad \cdots \quad \mathbf{r}^m)_{l \times ml}$ . Nous allons remplacer ce résultat dans l'équation ( 10 ) pour arriver à :

$$\begin{aligned}
\Omega_1 &= \sum_{j=1}^m -(\mathbf{r}^j \mathbf{C}^\dagger (\mathbf{h}^j)^\dagger + \mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger) = \\
&= -\left( (\mathbf{x} \cdot \mathbf{H}_R \mathbf{R}^\dagger + \mathbf{x}^* \cdot \mathbf{H}_I \mathbf{R}^\dagger)^* + (\mathbf{x} \cdot \mathbf{H}_R \mathbf{R}^\dagger + \mathbf{x}^* \cdot \mathbf{H}_I \mathbf{R}^\dagger) \right) \quad (15) \\
&= -\left( (\mathbf{x}^* \cdot \mathbf{H}_R^* \mathbf{R}^\top + \mathbf{x} \cdot \mathbf{H}_I^* \mathbf{R}^\top) + (\mathbf{x} \cdot \mathbf{H}_R \mathbf{R}^\dagger + \mathbf{x}^* \cdot \mathbf{H}_I \mathbf{R}^\dagger) \right) \\
&= -\left( \mathbf{x} \cdot \underbrace{(\mathbf{H}_R \mathbf{R}^\dagger + \mathbf{H}_I^* \mathbf{R}^\top)}_{\mathbf{P}_{k \times l}^\dagger} + \mathbf{x}^* \cdot \underbrace{(\mathbf{H}_I \mathbf{R}^\dagger + \mathbf{H}_R^* \mathbf{R}^\top)}_{\mathbf{P}_{k \times l}^\top} \right) = -(\mathbf{x} \cdot \mathbf{P}^\dagger + \mathbf{x}^* \cdot \mathbf{P}^\top)
\end{aligned}$$

Où

$$\mathbf{P} = \mathbf{R} \cdot \mathbf{H}_R^\dagger + \mathbf{R}^* \cdot \mathbf{H}_I^\top \quad (16)$$

En comparant l'équation ( 15 ) avec ( 10 ), le vecteur  $\mathbf{P}' = (p_1 \quad \cdots \quad p_q \quad \cdots \quad p_k) / A$  contient les paramètres de décision. Maintenant, nous avons tous les éléments nécessaires pour le décodage d'un code STBC orthogonal sous forme matricielle. Cette présentation est très utile pour une programmation en Matlab. De plus, grâce à cette démonstration matricielle, on peut trouver les paramètres de décision de tous les symboles d'un paquet d'information d'un seul coup (en calculant le vecteur  $\mathbf{P}$  d'après les signaux reçus et les informations du canal et du code).

### Cas des Codes réels :

On peut déduire l'algorithme de décodage pour les codes réels à partir de celui des codes complexes. Pour ces codes  $\mathbf{B}_i = \mathbf{0}_{n \times l}$ ,  $\mathbf{C}^\dagger = \mathbf{C}^T$  et  $x^* = x$ .  $\Omega_1$  et  $\Omega_2$  se mettent alors sous les formes suivantes :

$$\begin{aligned} \Omega_1 &= \sum_{j=1}^m -(\mathbf{r}^j \mathbf{C}^T (\mathbf{h}^j)^\dagger + \mathbf{h}^j \mathbf{C} (\mathbf{r}^j)^\dagger) = \\ &= \sum_{q=1}^k -(p_q^* x_q + p_q x_q) = \sum_{q=1}^k -2\text{real}(p_q) \cdot x_q \end{aligned} \quad (17)$$

$$\Omega_2 = \sum_{q=1}^k x_q^2 \cdot \left( \sum_{i,j} |\alpha_{i,j}|^2 \right) \quad (18)$$

Et on peut également trouver les sous-métriques  $\Omega_q$  comme suivant :

$$\begin{aligned} \text{Min}_{x_q} \Omega_q &= \underbrace{\left( \sum_{i,j} |\alpha_{i,j}|^2 \right)}_{A > 0} \cdot x_q^2 - 2\text{real}(p_q) \cdot x_q = Ax_q^2 - 2\text{real}(p_q) \cdot x_q \\ &= A \left( x_q^2 - 2\text{real}(p_q) \cdot x_q + \frac{(\text{real}(p_q))^2}{A^2} \right) - \frac{(\text{real}(p_q))^2}{A} \\ \text{Min}_{x_q} \Omega_q &\equiv \text{Min}_{x_q} \left( x_q - \frac{\text{real}(p_q)}{A} \right)^2 \end{aligned} \quad (19)$$

Et pour le cas particulier de la modulation MDP2 la prise de décision s'effectue par utilisation d'un seuil :

$$\hat{x}_q = \text{sign} \left( \frac{\text{real}(p_q)}{A} \right) \quad (20)$$

Pour trouver le vecteur  $\mathbf{P}$ , on réécrit les expressions de (13) à (15) pour les codes réels. On obtient les matrices équivalentes pour l'antenne de réception  $j$  :

$$\mathbf{h}^j \mathbf{C} = \mathbf{h}^j \cdot \sum_{i=1}^k x_i \mathbf{A}_i = \sum_{i=1}^k x_i \mathbf{h}^j \mathbf{A}_i = \underbrace{(x_1 \quad \dots \quad x_i \quad \dots \quad x_k)}_{\mathbf{x}} \cdot \underbrace{\begin{pmatrix} \mathbf{h}^j \mathbf{A}_1 \\ \vdots \\ \mathbf{h}^j \mathbf{A}_i \\ \vdots \\ \mathbf{h}^j \mathbf{A}_k \end{pmatrix}}_{\mathbf{H}_R^j} = \mathbf{x} \cdot \mathbf{H}_R^j \quad (21)$$

En utilisant cette forme équivalente, on calcule  $\Omega_1$ .

$$\sum_{j=1}^m \mathbf{h}^j \mathbf{C}(\mathbf{r}^j)^\dagger = \sum_{j=1}^m \mathbf{x} \cdot \mathbf{H}_R^j \cdot (\mathbf{r}^j)^\dagger = \mathbf{x} \cdot \sum_{j=1}^m \mathbf{H}_R^j (\mathbf{r}^j)^\dagger = \mathbf{x} \cdot \mathbf{H}_R \cdot \mathbf{R}^\dagger \quad (22)$$

Avec les mêmes définitions pour  $\mathbf{H}_R$  et  $\mathbf{R}$ .

$$\begin{aligned} \Omega_1 &= \sum_{j=1}^m -\left(\mathbf{r}^j \mathbf{C}^T(\mathbf{h}^j)^\dagger + \mathbf{h}^j \mathbf{C}(\mathbf{r}^j)^\dagger\right) = -\left(\left(\mathbf{x} \cdot \mathbf{H}_R \mathbf{R}^\dagger\right)^* + \left(\mathbf{x} \cdot \mathbf{H}_R \mathbf{R}^\dagger\right)\right) \\ &= -\mathbf{x} \cdot \left( \underbrace{\mathbf{H}_R^* \mathbf{R}^T}_{\mathbf{P}_{k \times l}^T} + \underbrace{\mathbf{H}_R \mathbf{R}^\dagger}_{\mathbf{P}_{k \times l}^\dagger} \right) = -2\mathbf{x} \cdot \text{real}(\mathbf{P}^T) \end{aligned} \quad (23)$$

Pour les codes réels le vecteur  $\mathbf{P}$  simplifie alors sous la forme suivante :

$$\mathbf{P} = \mathbf{R} \mathbf{H}_R^\dagger \quad (24)$$

En comparant l'équation (23) avec (17), le vecteur  $\mathbf{P}' = \text{real}(p_1 \dots p_q \dots p_k)/A$  contient les paramètres de décision.

#### 4-4-2 Algorithme de Décodage :

La Figure 2 illustre le principe du décodage ML pour les codes STBC orthogonaux et complexes. Nous développons par la suite l'algorithme de décodage pour les codes complexes ainsi que pour les codes réels.

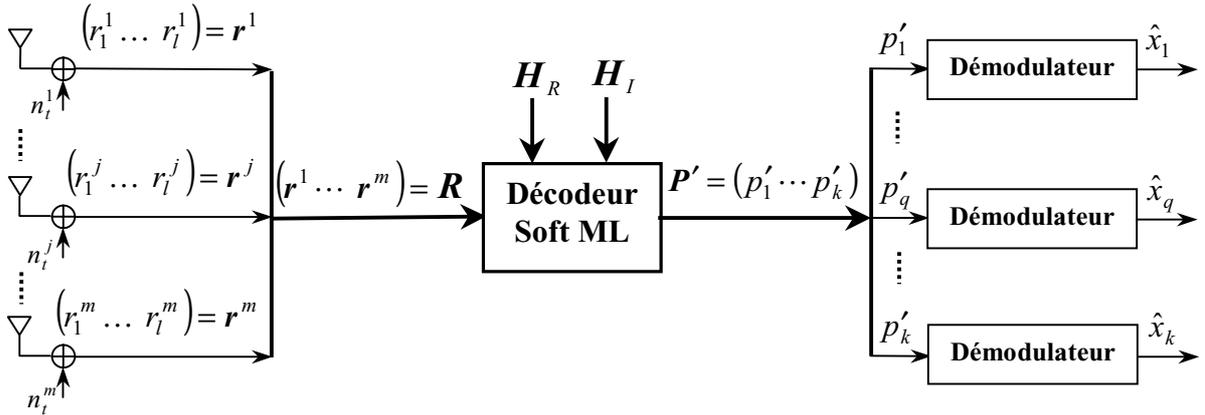


Figure 2 : Structure du décodeur ML pour les codes STBC orthogonaux complexes

Les données du programme :

- Le code, autrement dit les matrices  $\mathbf{A}_i$  et  $\mathbf{B}_i$  pour  $i = 1, \dots, k$ .
- Les signaux reçus  $r_i^j$  pour  $j = 1, \dots, m$  ;  $t = 1, \dots, l$  (Donc le vecteur  $\mathbf{R}$ ).

- Les informations du canal  $\alpha_{i,j}$  pour  $\forall i, j$  (Donc les vecteurs  $\mathbf{h}^j$  et la matrice  $\mathbf{H}$ ).

L'algorithme de décodage se résume alors comme suit :

- 1- Trouver les matrices équivalentes  $\mathbf{H}_R^j$  et  $\mathbf{H}_I^j$  pour  $j = 1, \dots, m$  (à partir de l'équation ( 13 )) et par conséquent les matrices  $\mathbf{H}_R$  et  $\mathbf{H}_I$  (définies dans l'équation ( 14 )).
- 2- Calculer le vecteur  $\mathbf{P}$  d'après l'équation ( 16 ) et obtenir les paramètres de décision  $p'_q = \frac{P_q}{A}$  pour tous les symboles d'un paquet d'information ( $q = 1, \dots, k$ ).
- 3- Prise de décision sur tous les symboles selon l'équation ( 12 ).

### Cas des codes réels :

Les données du programme :

- Le code, autrement dit les matrices  $A_i$  pour  $i = 1, \dots, k$ .
- Les signaux reçus  $r_t^j$  pour  $j = 1, \dots, m$  ;  $t = 1, \dots, l$  (Donc le vecteur  $\mathbf{R}$ ).
- Les informations du canal  $\alpha_{i,j}$  pour  $\forall i, j$  (Donc les vecteurs  $\mathbf{h}^j$  et la matrice  $\mathbf{H}$ ).

L'algorithme de décodage se résume alors comme suit :

- 1- Trouver les matrices équivalentes  $\mathbf{H}_R^j$  pour  $j = 1, \dots, m$  (à partir de l'équation ( 21 )) et par conséquent les matrices  $\mathbf{H}_R$  (définies dans l'équation ( 22 )).
- 2- Calculer le vecteur  $\mathbf{P}$  d'après l'équation ( 24 ) et obtenir les paramètres de décision  $p'_q = \frac{\text{real}(p_q)}{A}$  pour tous les symboles d'un paquet d'information ( $q = 1, \dots, k$ ).
- 3- Prise de décision sur tous les symboles selon l'équation ( 19 ).

## 4-5 Etude Analytique (Dérivation des Bornes Supérieurs de TEB)

Dans ce paragraphe, on aborde le calcul analytique du TEB pour un décodage ML présenté dans la section précédente. Le calcul se base à partir des paramètres de décision : le vecteur  $\mathbf{P}$  (les équations ( 16 ) et ( 24 )). On verra que les paramètres de décision contiennent deux termes, l'un qui est exactement égal au symbole d'information envoyé et l'autre qui correspond au bruit (bruit en sortie du décodeur). Le calcul du TEB nécessite de trouver les statistiques de ce bruit.

Pour commencer, on applique la définition de la matrice  $\mathbf{R}$  (utilisée dans l'équation ( 14 )) en utilisant les équations ( 2 ) et ( 13 ). Avec cette définition, les signaux reçus sur l'ensemble des antennes de réception et dans tous les temps d'un paquet d'information s'écrivent sous la forme suivante :

$$\mathbf{R} = \mathbf{x}.\mathbf{H}_R + \mathbf{x}^*.\mathbf{H}_I + \mathbf{N} \quad (25)$$

où  $\mathbf{N} = (\mathbf{n}^1 \quad \dots \quad \mathbf{n}^j \quad \dots \quad \mathbf{n}^m)_{l \times ml}$ , et  $\mathbf{n}^j$  le vecteur de bruit à l'antenne  $j$  a déjà été défini dans l'équation ( 2 ).

Avant d'appliquer ( 25 ) dans l'expression ( 16 ), il faut parler de la propriété des matrices  $\mathbf{H}_R$  et  $\mathbf{H}_I$  (composées des matrices  $\mathbf{H}_R^j$  et  $\mathbf{H}_I^j$ ) qui contiennent les informations du canal ainsi que le code. Pour cela, nous utiliserons les propriétés des matrices de code présentées aux équations ( 7 ). En utilisant les définitions des matrices  $\mathbf{H}_R^j$  et  $\mathbf{H}_I^j$  dans l'équation ( 13 ), nous pouvons démontrer que :

$$\begin{aligned} \mathbf{H}_R^j \cdot (\mathbf{H}_I^j)^\dagger + (\mathbf{H}_R^j \cdot (\mathbf{H}_I^j)^\dagger)^T &= \mathbf{0}_{k \times k} & (a) \\ \mathbf{H}_R^j \cdot (\mathbf{H}_R^j)^\dagger + (\mathbf{H}_I^j)^* \cdot (\mathbf{H}_I^j)^T &= \sum_{i=1}^n |\alpha_{i,j}|^2 \cdot \mathbf{I}_{k \times k} & (b) \end{aligned} \quad (26)$$

Les preuves de ces particularités des matrices  $\mathbf{H}_R^j$  et  $\mathbf{H}_I^j$  se trouvent dans l'annexe B.

L'équation ( 26 )-a signifie que la production matricielle  $(\mathbf{H}_R^j \cdot (\mathbf{H}_I^j)^\dagger)$  est antisymétrique avec des éléments zéros sur la diagonale. La particularité de ce type de matrice est que sa multiplication par un vecteur et son transposé de deux cotés est zéro (l'Annexe C).

A partir de l'équation ( 26 ) et en utilisant les définitions des matrices  $\mathbf{H}_R$  et  $\mathbf{H}_I$ , on peut trouver les expressions semblables pour ces deux matrices :

$$\mathbf{H}_R \cdot \mathbf{H}_I^\dagger + (\mathbf{H}_R \cdot \mathbf{H}_I^\dagger)^T = \mathbf{0}_{k \times k} \quad (a) \quad (27)$$

$$\mathbf{H}_R \cdot \mathbf{H}_R^\dagger + \mathbf{H}_I^* \cdot \mathbf{H}_I^T = \sum_{i=1}^n \sum_{j=1}^m |\alpha_{i,j}|^2 \cdot \mathbf{J}_{k \times k} \quad (\text{b})$$

Maintenant, après avoir présenté ces particularités des matrices  $\mathbf{H}_R$  et  $\mathbf{H}_I$ , on peut se focaliser sur les paramètres de décision (le vecteur  $\mathbf{P}$ ). Pour cela on remplace le vecteur  $\mathbf{R}$  dans l'équation ( 16 ) par celui de l'équation ( 25 ) et en utilisant les équations ( 27 ). On arrive donc à :

$$\begin{aligned} \mathbf{P} &= \mathbf{R} \cdot \mathbf{H}_R^\dagger + \mathbf{R}^* \cdot \mathbf{H}_I^T \\ &= [\mathbf{x} \cdot \mathbf{H}_R + \mathbf{x}^* \cdot \mathbf{H}_I + \mathbf{N}] \cdot \mathbf{H}_R^\dagger + [\mathbf{x}^* \cdot \mathbf{H}_R^* + \mathbf{x} \cdot \mathbf{H}_I^* + \mathbf{N}^*] \cdot \mathbf{H}_I^T \\ &= \mathbf{x} \cdot [\mathbf{H}_R \cdot \mathbf{H}_R^\dagger + \mathbf{H}_I^* \cdot \mathbf{H}_I^T] + \mathbf{x}^* \cdot [\mathbf{H}_I \cdot \mathbf{H}_R^\dagger + \mathbf{H}_R^* \cdot \mathbf{H}_I^T] + [\mathbf{N} \cdot \mathbf{H}_R^\dagger + \mathbf{N}^* \cdot \mathbf{H}_I^T] \quad (28) \\ &= \left( \sum_{i=1}^m \sum_{j=1}^n |\alpha_{i,j}|^2 \right) \cdot \mathbf{x} + [\mathbf{N} \cdot \mathbf{H}_R^\dagger + \mathbf{N}^* \cdot \mathbf{H}_I^T] \end{aligned}$$

Donc

$$\mathbf{P}' = \mathbf{x} + \frac{1}{\sum_{i,j} |\alpha_{i,j}|^2} [\mathbf{N} \cdot \mathbf{H}_R^\dagger + \mathbf{N}^* \cdot \mathbf{H}_I^T] = \mathbf{x} + \boldsymbol{\eta} \quad (29)$$

où  $\boldsymbol{\eta}$  est un vecteur de bruit complexe (et gaussien si les gains du canal sont considérés constants).

$$\boldsymbol{\eta} = \frac{1}{\sum_{i,j} |\alpha_{i,j}|^2} [\mathbf{N} \cdot \mathbf{H}_R^\dagger + \mathbf{N}^* \cdot \mathbf{H}_I^T]$$

Les paramètres de décision que l'on obtient pour le décodeur ML, donnent exactement les symboles d'information contaminés par un bruit additif en sortie de décodeur. On peut donc imaginer que le système est équivalent à un système mono-antenne avec le même type de modulation utilisée dans le système multi-antennes. Par contre, les statistiques du bruit additif équivalent dépendent du canal MIMO.

Afin de trouver une expression analytique pour le TEB, il nous faut calculer les statistiques de ce bruit. On considère dans un premier temps que les gains du canal ( $\alpha_{i,j}$ ) sont constants. Ensuite, on les considère comme les variables aléatoires gaussiennes et complexes, ce qui est généralement le cas dans la pratique.

#### 4-5-1 Canal constant

Pour calculer la matrice d'autocorrélation du vecteur  $\boldsymbol{\eta}$ , on a besoin des statistiques du vecteur  $\mathbf{N}$ . A partir des statistiques des bruits  $n_i^j$ , on peut vérifier très facilement que :

$$\begin{aligned}
E\{\mathbf{N}^\dagger \mathbf{N}\} &= E\{\mathbf{N}^T \mathbf{N}^*\} = \sigma^2 \mathbf{I}_{ml \times ml} \\
E\{\mathbf{N}^T \mathbf{N}\} &= E\{\mathbf{N}^\dagger \mathbf{N}^*\} = \mathbf{0}_{ml \times ml}
\end{aligned} \tag{30}$$

En utilisant les équations (30) et (27-b), on calcule la matrice d'autocorrélation du vecteur bruit  $\boldsymbol{\eta}$ .

$$\begin{aligned}
E\{\boldsymbol{\eta}^\dagger \boldsymbol{\eta}\} &= \frac{1}{\left(\sum_{i,j} |\alpha_{i,j}|^2\right)^2} E\left\{ \left[ \mathbf{H}_R \cdot \mathbf{N}^\dagger + \mathbf{H}_I^* \cdot \mathbf{N}^T \right] \left[ \mathbf{N} \cdot \mathbf{H}_R^\dagger + \mathbf{N}^* \cdot \mathbf{H}_I^T \right] \right\} \\
&= \frac{1}{\left(\sum_{i,j} |\alpha_{i,j}|^2\right)^2} \left[ \underbrace{\mathbf{H}_R E\{\mathbf{N}^\dagger \mathbf{N}\} \mathbf{H}_R^\dagger}_{\sigma^2 \mathbf{I}_{ml \times ml}} + \underbrace{\mathbf{H}_I^* E\{\mathbf{N}^T \mathbf{N}^*\} \mathbf{H}_I^T}_{\sigma^2 \mathbf{I}_{ml \times ml}} + \underbrace{\mathbf{H}_R E\{\mathbf{N}^\dagger \mathbf{N}^*\} \mathbf{H}_I^T}_{\mathbf{0}_{ml \times ml}} + \underbrace{\mathbf{H}_I^* E\{\mathbf{N}^T \mathbf{N}\} \mathbf{H}_R^\dagger}_{\mathbf{0}_{ml \times ml}} \right] \\
&= \frac{\sigma^2}{\left(\sum_{i,j} |\alpha_{i,j}|^2\right)^2} \left[ \mathbf{H}_R \mathbf{H}_R^\dagger + \mathbf{H}_I^* \mathbf{H}_I^T \right] = \frac{\sigma^2}{\sum_{i,j} |\alpha_{i,j}|^2} \mathbf{I}_{k \times k}
\end{aligned} \tag{31}$$

Cette relation démontre que le bruit du système multi-antennes reste gaussien et indépendant avec une variance atténuée (si  $A = \sum_{i,j} |\alpha_{i,j}|^2$  est supérieur à 1, ce qui est souvent le cas).

Comme  $\sigma^2 = 1/\text{SNR}$  et que l'énergie de chaque symbole est égale à  $1/n$ , le rapport équivalent signal à bruit en sortie de décodeur pour chaque symbole se calcule comme suivant.

$$\text{SNR}_q = \frac{1/n}{\sigma^2 / \sum_{i,j} |\alpha_{i,j}|^2} = \left( \sum_{i,j} |\alpha_{i,j}|^2 / n \right) \text{SNR} \tag{32}$$

Le décodeur soft de la Figure 2, calcule les paramètres de décision à partir des sorties des récepteurs et les paramètres de canal. Ces paramètres de décision vont être appliqués au démodulateur qui cherche le symbole le plus proche à ce paramètre dans la constellation envoyée. Cette prise de décision s'effectue séparément pour chaque symbole  $x_1, x_2, \dots, x_k$ . Vue l'indépendance des symboles détectés, on se focalise sur la détection d'un symbole pour calculer le taux d'erreur.

Le signal en entrée du démodulateur est composé donc du symbole  $x_q$  avec l'énergie  $1/n$  et du bruit gaussien et complexe  $\eta_q$  avec la moyenne nulle et la variance  $\sigma^2 / \sum_{i,j} |\alpha_{i,j}|^2$ . C'est un

cas classique qui se trouve dans la littérature [12]-[13] pour lequel il existe des expressions (exacte ou approximatives) pour le taux d'erreur de symbole. Suivant la constellation du signal envoyé, on peut donc donner une expression analytique du taux d'erreur de symbole pour un récepteur ML. C'est une fonction de la distance euclidienne minimale entre les points

de constellation  $d$ . Nous calculons le taux d'erreur de symbole  $p_e$  pour les modulations QPSK, MAQ-16 et MPSK. On définit d'abord les relations suivantes :

$$w = \frac{1}{n} \quad N_0 = \frac{\sigma^2}{\sum_{i,j} |\alpha_{i,j}|^2} \quad p = \frac{1}{2} \operatorname{erfc} \left( \frac{d}{2\sqrt{N_0}} \right) = Q \left( \frac{d}{\sqrt{2N_0}} \right)$$

On obtient donc :

$$\begin{aligned} \text{QPSK :} \quad P_e &= 2p - p^2 \cong 2p & ; d &= \sqrt{2w} \\ &\Rightarrow P_e = 2Q(\sqrt{SNR_q}) - Q^2(\sqrt{SNR_q}) \cong 2Q(\sqrt{SNR_q}) \end{aligned}$$

$$\begin{aligned} \text{M-Aire PSK :} \quad P_e &\cong 2p & ; d &= 2\sqrt{w} \sin(\pi/M) \\ &\Rightarrow P_e \cong 2Q(\sqrt{2SNR_q} \sin(\pi/M)) \end{aligned}$$

$$\begin{aligned} \text{QAM-16 :} \quad P_e &= 3p - \frac{9}{4}p^2 \cong 3p & ; d &= 2(\sqrt{10} - \sqrt{2})\sqrt{w} \\ &\Rightarrow P_e = 3Q(2(\sqrt{5} - 1)\sqrt{SNR_q}) - \frac{9}{4}Q^3(2(\sqrt{5} - 1)\sqrt{SNR_q}) \cong 3Q(2(\sqrt{5} - 1)\sqrt{SNR_q}) \end{aligned}$$

#### 4-5-2 Canal à variation aléatoire

Maintenant on passe au cas plus réaliste où les gains du canal sont des variables aléatoires. Dans ce cas, pour calculer la matrice d'autocorrélation de bruit, la moyenne s'effectue à la fois sur le bruit et sur les atténuations du canal. Sachant qu'ils sont indépendants, on peut séparer l'opérateur de l'espérance sur chacun des deux variables aléatoires. D'abord, on le fera sur le bruit et ensuite sur les matrices  $\mathbf{H}_R$  et  $\mathbf{H}_I$ , autrement dit sur les gains  $\alpha_{i,j}$ .

$$\begin{aligned} E\{\boldsymbol{\eta}^\dagger \boldsymbol{\eta}\} &= E_\alpha \left\{ E_N \left\{ \boldsymbol{\eta}^\dagger \boldsymbol{\eta} \right\} \right\} \\ &= E_\alpha \left\{ \sigma^2 \mathbf{I}_{k \times k} / \left( \sum_{i,j} |\alpha_{i,j}|^2 \right)^2 \right\} = \sigma^2 \mathbf{I}_{k \times k} E_\alpha \left\{ 1 / \sum_{i,j} |\alpha_{i,j}|^2 \right\} \end{aligned} \quad (33)$$

Il nous faut donc l'espérance de l'inverse de  $\sum |\alpha_{i,j}|^2$ . On sait que les gains du canal  $\alpha_{i,j}$  sont les variables aléatoires gaussiens, complexes et indépendantes à variance 0,5 pour chaque partie réelle et imaginaire et à moyenne nulle. Dans ce cas, le module au carré de  $a_{ij}$  ( $|\alpha_{i,j}|^2$ ) suit la loi de Gamma de zéro degré (Gamma(1,0)) avec les paramètres  $c=1$  et  $b=0$  (voir l'Annexe D). La densité de distribution de Gamme est la suivante :

$$y \sim \text{Gamma}(c, N) \Leftrightarrow f_Y(y) = \frac{c^{N+1}}{N!} y^N e^{-cy} U(y)$$

On peut également démontrer que la somme de  $N$  variables aléatoire de type Gamme(c,0) suit une densité de probabilité de Gamma de  $(N-1)^{\text{ième}}$  degré, Gamma(c,N-1) (Annexe D). Par conséquent,  $z = \sum_{i=1}^n \sum_{j=1}^m |\alpha_{i,j}|^2$  a une densité de probabilité de Gamma(1, nm-1). Le problème se réduit alors à trouver l'espérance de  $E_Z \left\{ \frac{1}{z} \right\}$ . On montre dans l'annexe D que cette espérance est égale à  $1/nm$ . Utilisant ce résultat, l'équation (33) se met sous la forme suivante :

$$E\{\boldsymbol{\eta}^\dagger \boldsymbol{\eta}\} = \sigma^2 E_\alpha \left\{ 1 / \sum_{i,j} |\alpha_{i,j}|^2 \right\} = \frac{\sigma^2}{nm} \quad (34)$$

Bien que cette relation donne la matrice d'auto-corrélation du bruit, il ne donne pas sa fonction de distribution. De toute façon le rapport signal à bruit peut être calculé et on remarque une atténuation de la puissance de bruit de  $mn$ . Le rapport signal à bruit en sortie de décodeur soft par rapport à un système mono-antenne est donc le suivant.

$$\text{SNR}_q = \frac{1/n}{\sigma^2/mn} = m \times \text{SNR} \quad (35)$$

Nous avons donc réussi à calculer la variance du bruit en sortie du décodeur soft. Mais l'expression de la fonction de distribution de ce bruit s'avère complexe sans quoi, la probabilité d'erreur ne peut pas être estimée. Pour pouvoir débloquent la situation, nous utilisons le théorème de la limite centrale pour le cas où le nombre d'antennes de réception et d'émission ( $mn$ ) est assez important. En examinant la relation ( 29 ), on remarque que le bruit  $\eta$  est une somme de  $ml$  variables aléatoires. Supposant  $ml$  important, on peut modéliser ce bruit avec une distribution gaussienne. Cette condition n'est malheureusement pas vérifiée pour des cas relativement courants où le nombre d'antennes reste assez limité.

Nous utilisons donc la probabilité d'erreur que nous avons calculée pour le cas d'un canal constant. Ensuite, en faisant varier les paramètres du canal, on calcule l'espérance de cette probabilité d'erreur conditionnelle. Autrement dit, on calcule :

$$\bar{P}_e = E_{\alpha} \{P(e|\alpha)\} \quad (36)$$

où  $|\alpha|$  représente l'ensemble des modules de tous les gains complexes de canal :  $|\alpha_{i,j}|$  ( $i = 1, \dots, n; j = 1, \dots, m$ ) et  $P(|\alpha|)$  est la densité de probabilité conjointe de ces derniers. Comme les gains de canal sont gaussiens et indépendants, chaque  $|\alpha_{i,j}|$  aura une densité de probabilité de Rayleigh et la densité de probabilité conjointe sera le produit de toutes ces densités de probabilité.

$$\begin{aligned} \alpha_{i,j} = (\alpha_{i,j}^R, \alpha_{i,j}^I) \sim N(0, 1/2) &\Rightarrow P(|\alpha_{i,j}|) = 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} \\ &\Rightarrow P(|\alpha|) = \prod_{i,j} P(|\alpha_{i,j}|) \end{aligned} \quad (37)$$

Utilisant ces deux dernières équations, on va pouvoir calculer  $P_e$ .

$$\bar{P}_e = E_{\alpha} \{P(e|\alpha)\} = \int_{\alpha} P(e|\alpha) P(|\alpha|) d|\alpha| = \int_{\alpha} P(e|\alpha) \prod_{i,j} P(|\alpha_{i,j}|) \prod_{i,j} d|\alpha_{i,j}| \quad (38)$$

Nous allons donc calculer cette probabilité pour un certain nombre de modulations classiques. Ceci se fait en remplaçant le  $P(e|\alpha)$  dans la relation ci-dessus par les valeurs trouvées pour le cas du canal constant. D'après nos calculs donnés sur la page 16, ces probabilités d'erreur sont des polynômes de  $Q(G\sqrt{SNR_q})$  où  $G$  n'est qu'un facteur d'échelle. Afin de trouver des expressions simples, on utilise des fonctions approximatives pour la fonction  $Q(x)$ . [12]-[13] :

$$\begin{aligned} Q(x) &< \frac{1}{2} e^{-\frac{x^2}{2}} & (a) \\ Q(x) &\cong \frac{1}{\sqrt{2\pi x}} e^{-\frac{x^2}{2}} & (b) \end{aligned} \quad (39)$$

L'expression ( 39-a ) s'utilise afin de trouver une borne supérieure pour le taux d'erreur tandis que l'équation ( 39-b ) est une bonne approximation pour  $Q(x)$  (toujours supérieur mais plus proche de  $Q(x)$ ). Afin de trouver une seule expression pour toutes les modulations présentées, on utilise les notations suivantes.

$$P_e = P(e|\alpha) \cong G_1 Q(G_2 \sqrt{SNR_q}) ; \begin{cases} \text{QPSK :} & G_1 = 2 ; G_2 = 1 \\ \text{M - Aire :} & G_1 = 2 ; G_2 = \sqrt{2} \sin(\pi/M) \\ \text{QAM-16 :} & G_1 = 3 ; G_2 = 2(\sqrt{5} - 1) \end{cases} \quad (40)$$

On peut donc regrouper les équations ( 32 ) en utilisant la notation donnée en ( 40 ) pour arriver au résultat suivant (dans le quelle l'approximation ( 39-a ) a été utilisée).

$$\begin{aligned} \bar{P}_e &= \int_{|\alpha|} P(e|\alpha) \prod_{i,j} P(|\alpha_{i,j}|) \prod_{i,j} d|\alpha_{i,j}| = \int_{|\alpha|} G_1 Q(G_2 \sqrt{SNR_q}) \prod_{i,j} \left( 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} \right) d|\alpha_{i,j}| \\ &< G_1 \int_{|\alpha|} \frac{1}{2} e^{-\frac{G_2^2 SNR_q}{2}} \prod_{i,j} \left( 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} \right) d|\alpha_{i,j}| \\ &< \frac{G_1}{2} \int_{|\alpha|} e^{-G_2^2 SNR (\sum |\alpha_{i,j}|^2 / 2n)} \prod_{i,j} \left( 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} \right) d|\alpha_{i,j}| \\ &< \frac{G_1}{2} \int_{|\alpha|} \prod_{i,j} 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} e^{-G_2^2 SNR |\alpha_{i,j}|^2 / 2n} d|\alpha_{i,j}| \\ &< \frac{G_1}{2} \prod_{i,j} \left( \int_{|\alpha_{i,j}|=0}^{\infty} 2|\alpha_{i,j}| \exp\left(-|\alpha_{i,j}|^2 \left(1 + G_2^2 \frac{SNR}{2n}\right)\right) d|\alpha_{i,j}| \right) \quad (41) \\ &< \frac{G_1}{2} \left( \int_0^{\infty} 2\alpha \exp\left(-\alpha^2 \left(1 + G_2^2 \frac{SNR}{2n}\right)\right) d\alpha \right)^{mn} \\ &< \frac{G_1}{2} \left( \left[ -\exp\left(-\alpha^2 \left(1 + G_2^2 \frac{SNR}{2n}\right)\right) \right] / \left[ \left(1 + G_2^2 \frac{SNR}{2n}\right) \right]_0^{\infty} \right)^{mn} \\ &\Rightarrow \bar{P}_e < \frac{G_1}{2 \left(1 + G_2^2 \frac{SNR}{2n}\right)^{mn}} \end{aligned}$$

Cette relation simple donne la borne supérieure du taux d'erreur pour le récepteur ML et pour le code STBC orthogonal et complexe. On remarque que  $\bar{P}_e$  diminue de manière exponentielle par rapport au nombre d'antennes.

Pour améliorer l'approximation de la borne supérieure, on recalcule la somme de l'équation ( 38 ) en utilisant l'équation l'approximation ( 39-b ) au lieu de ( 39-a ). On a donc :

$$\begin{aligned}
\bar{P}_e &= \int_{|\alpha|} P(e|\alpha) \prod_{i,j} P(|\alpha_{i,j}|) \prod_{i,j} d|\alpha_{i,j}| = \int_{|\alpha|} G_1 Q(G_2 \sqrt{SNR_q}) \prod_{i,j} \left( 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} \right) d|\alpha_{i,j}| \\
&= \int_{|\alpha|} \frac{G_1}{\sqrt{2\pi} G_2 \sqrt{SNR_q}} e^{-\frac{G_2^2 SNR_q}{2}} \prod_{i,j} \left( 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} \right) d|\alpha_{i,j}| \\
&= \frac{G_1 \sqrt{n}}{\sqrt{2\pi} G_2 \sqrt{SNR}} \int_{|\alpha|} \frac{1}{\sqrt{\sum_{i,j} |\alpha_{i,j}|^2}} e^{-G_2^2 SNR (\sum_{i,j} |\alpha_{i,j}|^2 / 2n)} \prod_{i,j} \left( 2|\alpha_{i,j}| e^{-|\alpha_{i,j}|^2} \right) d|\alpha_{i,j}| \quad (42) \\
&= \frac{G_1 \sqrt{n}}{\sqrt{2\pi} G_2 \sqrt{SNR}} \int_{|\alpha|} \frac{1}{\sqrt{\sum_{i,j} |\alpha_{i,j}|^2}} \exp\left(-\left(1 + G_2^2 \frac{SNR}{2n}\right) \sum_{i,j} |\alpha_{i,j}|^2\right) \prod_{i,j} 2|\alpha_{i,j}| d|\alpha_{i,j}|
\end{aligned}$$

En comparant cette équation avec l'équation ( 41 ), on remarque le terme supplémentaire  $1/\sqrt{\sum |\alpha_{i,j}|^2}$  qui va largement compliquer le calcul de l'intégrale car il n'est plus possible de le séparer en produit de  $mn$  intégrales. Pour résoudre cette intégrale, on se sert d'un changement de repère du cartésien au polaire de dimension  $mn$ . Pour cela, et avant tout, on définit les notations suivantes :

$$\begin{aligned}
M &= mn \quad ; \quad \alpha_d = |\alpha_{i,j}| \quad ; \quad d = 1, 2, \dots, mn \quad ; \quad i = 1, 2, \dots, n \quad ; \quad j = 1, 2, \dots, m \\
i &= \left\lfloor \frac{d-1}{m} \right\rfloor + 1 \quad , \quad j = (d-1) \% m \quad ; \quad i = (i-1)m + j
\end{aligned}$$

où  $\lfloor y \rfloor$  indique l'entier inférieur le plus proche de  $y$  et  $\%$  le module de la division de deux entiers.

Avec ces notations, on applique le changement de repère du cartésien au polaire comme suivant :

$$\begin{aligned}
\alpha_1 &= \rho \cos \varphi_1 \\
\alpha_i &= \rho \left( \prod_{j=1}^{i-1} \sin \varphi_j \right) \cos \varphi_i \quad ; \quad i = 2, \dots, M-1 \\
\alpha_M &= \rho \prod_{j=1}^{M-1} \sin \varphi_j
\end{aligned} \quad (43)$$

La limite de variations de toutes les  $\varphi_i$  est de 0 à  $\pi/2$  car les gains  $\alpha_d$  sont positifs. On peut vérifier très facilement que  $\sum \alpha_d^2 = \rho^2$  et c'est aussi grâce à cela que l'intégrale ( 42 ) peut se simplifier. Dans ce changement de repère on peut montrer que [Annexe E] :

$$d\alpha_1 d\alpha_2 \dots d\alpha_M = \rho^{M-1} (\sin \varphi_1)^{M-2} \dots (\sin \varphi_i)^{M-1-i} \dots \sin \varphi_{M-2} d\rho d\varphi_1 d\varphi_2 \dots d\varphi_{M-1} \quad (44)$$

Maintenant en appliquant ce changement du repère l'intégrale ( 42 ) se réécrit sous la forme suivante.

$$\begin{aligned}
\bar{P}_e &= \frac{G_1 \sqrt{n}}{\sqrt{2\pi} G_2 \sqrt{SNR}} \int_{|\alpha|} \frac{1}{\sqrt{\sum_{i,j} |\alpha_{i,j}|^2}} \exp\left(-\left(1 + G_2^2 \frac{SNR}{2n}\right) \sum_{i,j} |\alpha_{i,j}|^2\right) \prod_{i,j} 2 |\alpha_{i,j}| d|\alpha_{i,j}| \\
&= \frac{2^{mn} G_1 \sqrt{n}}{\sqrt{2\pi} G_2 \sqrt{SNR}} \int_{\rho, \varphi} \frac{1}{\rho} e^{-\left(1 + G_2^2 \frac{SNR}{2n}\right) \rho^2} \left( \rho^M \prod_{i=1}^{M-1} (\sin \varphi_i)^{M-i} \cos \varphi_i \right) \left( \rho^{M-1} d\rho \prod_{i=1}^{M-1} (\sin \varphi_i)^{M-1-i} d\varphi_i \right) \\
&= \frac{2^{mn} G_1 \sqrt{n}}{\sqrt{2\pi} G_2 \sqrt{SNR}} \int_{\rho=0}^{\infty} \rho^{2M-2} e^{-\left(1 + G_2^2 \frac{SNR}{2n}\right) \rho^2} d\rho \prod_{i=1}^{M-1} \int_{\varphi=0}^{\pi/2} (\sin \varphi_i)^{2M-2i-1} \cos \varphi_i d\varphi_i \quad (45)
\end{aligned}$$

Dans cette expression, les intégrales concernant les  $\varphi_i$  se calculent facilement alors que pour l'intégrale concernant  $\rho$ , on utilise la relation suivante (preuve dans l'annexe E).

$$\int_0^{\infty} \rho^{2K} e^{-\rho^2} d\rho = \frac{(2K)!}{K! 2^{2K+1}} \sqrt{\pi} \quad (46)$$

Pour pouvoir utiliser cette relation, on applique le changement de variable suivant :

$$\rho' = \sqrt{1 + G_2^2 \frac{SNR}{2n}} \rho \Rightarrow d\rho' = \sqrt{1 + G_2^2 \frac{SNR}{2n}} d\rho$$

On arrive finalement à :

$$\begin{aligned}
\bar{P}_e &\cong \frac{2^{mn} G_1 \sqrt{n}}{\sqrt{2\pi} G_2 \sqrt{SNR} \sqrt{\left(1 + G_2^2 \frac{SNR}{2n}\right)^{2M-1}}} \int_{\rho=0}^{\infty} \rho^{2mn-2} e^{-\rho^2} d\rho \prod_{i=1}^{mn-1} \left[ \frac{(\sin \varphi_i)^{2mn-2i}}{2mn-2i} \right]_0^{\pi/2} \\
&= \frac{2^{mn} G_1 \sqrt{n}}{\sqrt{2\pi} G_2 \sqrt{SNR} \sqrt{\left(1 + G_2^2 \frac{SNR}{2n}\right)^{2mn-1}}} \left( \frac{(2mn-2)! \sqrt{\pi}}{(mn-1)! 2^{2mn-1}} \right) \prod_{j=1}^{mn-1} \left( \frac{1}{2j} \right) \quad (47) \\
&= \frac{G_1}{\left(1 + G_2^2 \frac{SNR}{2n}\right)^{mn}} \sqrt{\frac{1 + G_2^2 \frac{SNR}{2n}}{G_2^2 \frac{SNR}{2n}}} \left( \frac{(2mn-2)!}{((mn-1)!)^2 2^{2mn-1}} \right)
\end{aligned}$$

Cette relation simple donne une borne supérieure très proche de la probabilité d'erreur du système STBC.

En comparant cette expression avec celle donnée en ( 41 ), on remarque le terme commun  $G_1 / \left(1 + G_2^2 \text{SNR} / 2n\right)^{mn}$  qui diminue de manière exponentielle par rapport au nombre d'antennes ( $mn$ ). Le terme  $\sqrt{\left(1 + G_2^2 \frac{\text{SNR}}{2n}\right)} / G_2^2 \frac{\text{SNR}}{2n}$  s'approche à un pour les rapports SNR assez important. En revanche pour les SNR très faibles, il devient trop important, ce qui peut se voir aussi par l'approximation ( 39-b) qui n'est plus une bonne approximation de fonction  $Q$ . Cela veut dire que pour les SNR faibles, la précision de cette borne va chuter et on préfère utiliser la borne de l'équation ( 41 ). Pour les grands SNR, on peut encore simplifier l'équation ( 47 ) pour arriver à :

$$\bar{P}_e \cong \frac{G_1}{\left(1 + G_2^2 \frac{\text{SNR}}{2n}\right)^{mn}} \left( \frac{(2mn - 2)!}{((mn - 1)!)^2 2^{2mn-1}} \right) \quad (48)$$

### 4-5-3 Cas des codes réels

Dans ce paragraphe nous essayons de simplifier les résultats obtenus pour le cas d'une constellation réelle. En utilisant les équations ( 2 ) et ( 21 ) et la définition de la matrice  $\mathbf{R}$  (voir page 9), les signaux reçus sur toutes les antennes de réception et pour tous les temps d'un paquet s'écrivent sous la forme suivante :

$$\mathbf{R} = \mathbf{x}.\mathbf{H}_R + \mathbf{N} \quad (49)$$

avec la même définition du vecteur  $\mathbf{N}$  donnée par l'équation ( 25 ). En utilisant les propriétés de la matrices du code réel (les équations ( 8 )) et la définition des matrices  $\mathbf{H}_R^j$  dans l'équation ( 21 ), nous pouvons démontrer que (ref. Annexe B) :

$$real\left(\mathbf{H}_R^j \cdot (\mathbf{H}_R^j)^\dagger\right) = \sum_{i=1}^n |\alpha_{i,j}|^2 \cdot \mathbf{I}_{k \times k} \quad (50)$$

Et à partir de cette dernière et la définition des matrices  $\mathbf{H}_R$ , on arrive à :

$$real\left(\mathbf{H}_R \cdot (\mathbf{H}_R)^\dagger\right) = (\mathbf{H}_R \mathbf{H}_R^\dagger + \mathbf{H}_R^* \mathbf{H}_R^T) / 2 = \sum_{i=1}^n \sum_{j=1}^m |\alpha_{i,j}|^2 \cdot \mathbf{I}_{k \times k} \quad (51)$$

Maintenant on se concentre sur les paramètres de décision et on remplace le vecteur  $\mathbf{R}$  dans l'équation ( 24 ) par celui de ( 49 ) et en utilisant l'équation ( 51 ) on arrive à :

$$\mathbf{P} = \mathbf{R} \cdot \mathbf{H}_R^\dagger = [\mathbf{x}.\mathbf{H}_R + \mathbf{N}] \cdot \mathbf{H}_R^\dagger = \mathbf{x} \cdot [\mathbf{H}_R \cdot \mathbf{H}_R^\dagger] + \mathbf{N} \cdot \mathbf{H}_R^\dagger \quad (52)$$

Donc

$$\mathbf{P}' = \frac{real(\mathbf{P})}{\sum_{i,j} |\alpha_{i,j}|^2} = \mathbf{x} + \frac{1}{\sum_{i,j} |\alpha_{i,j}|^2} real(\mathbf{N} \cdot \mathbf{H}_R^\dagger) = \mathbf{x} + \boldsymbol{\eta} \quad (53)$$

où

$$\boldsymbol{\eta} = \frac{1}{\sum_{i,j} |\alpha_{i,j}|^2} real(\mathbf{N} \cdot \mathbf{H}_R^\dagger) = \frac{1}{2 \sum_{i,j} |\alpha_{i,j}|^2} (\mathbf{N} \cdot \mathbf{H}_R^\dagger + \mathbf{N}^* \cdot \mathbf{H}_R^T)$$

En ce qui concerne les statistiques de bruit additif du canal  $\mathbf{N}$ , l'équation ( 30 ) est toujours valable. On va alors calculer la matrice d'autocorrélation de vecteur de bruit réel  $\boldsymbol{\eta}$ , d'abord pour le cas où les gains du canal seraient constants.

$$\begin{aligned}
E\{\boldsymbol{\eta}^\dagger \boldsymbol{\eta}\} &= \frac{1}{4 \left( \sum_{i,j} |\alpha_{i,j}|^2 \right)^2} E\left\{ \left[ \mathbf{H}_R \cdot \mathbf{N}^\dagger + \mathbf{H}_R^* \cdot \mathbf{N}^T \right] \left[ \mathbf{N} \cdot \mathbf{H}_R^\dagger + \mathbf{N}^* \cdot \mathbf{H}_R^T \right] \right\} \\
&= \frac{1}{4 \left( \sum_{i,j} |\alpha_{i,j}|^2 \right)^2} \left[ \underbrace{H_R E\{N^\dagger N\} H_R^\dagger}_{\sigma^2 I_{m \times m}} + \underbrace{H_R^* E\{N^T N^*\} H_R^T}_{\sigma^2 I_{m \times m}} + \underbrace{H_R E\{N^\dagger N^*\} H_R^T}_{0_{m \times m}} + \underbrace{H_R^* E\{N^T N\} H_R^\dagger}_{0_{m \times m}} \right] \\
&= \frac{\sigma^2}{4 \left( \sum_{i,j} |\alpha_{i,j}|^2 \right)^2} \left[ H_R H_R^\dagger + H_R^* H_R^T \right] = \frac{\sigma^2}{2 \sum_{i,j} |\alpha_{i,j}|^2} I_{k \times k}
\end{aligned} \tag{54}$$

Donc l'énergie de chaque élément de bruit est identique à l'énergie de bruit additif du système multi-antennes divisée par  $2A$ . Le rapport signal à bruit en sortie de décodeur soft pour chaque symbole sera alors :

$$\text{SNR}_k = \frac{1/n}{\sigma^2 / 2 \sum_{i,j} |\alpha_{i,j}|^2} = \left( 2 \sum_{i,j} |\alpha_{i,j}|^2 / n \right) \text{SNR} \tag{55}$$

Alors pour le cas de la modulation PAM et pour un canal aux coefficients constants, on peut écrire :

$$w = \frac{1}{n} \quad \frac{N_0}{2} = \frac{\sigma^2}{2 \sum_{i,j} |\alpha_{i,j}|^2} \quad p = \frac{1}{2} \text{erfc} \left( \frac{d}{2\sqrt{N_0}} \right) = Q \left( \frac{d}{\sqrt{2N_0}} \right)$$

$$\text{BPSK :} \quad P_e = p \quad ; \quad d = 2\sqrt{w} \quad \Rightarrow \quad P_e = Q(\sqrt{\text{SNR}_q})$$

$$\text{M-Aire PAM :} \quad P_e = \frac{2(M-1)}{M} p \quad ; \quad d = \frac{4}{M} \sqrt{w} \quad \Rightarrow \quad P_e = \frac{2(M-1)}{M} Q \left( \sqrt{\frac{2}{M} \text{SNR}_q} \right)$$

Dans le cas où les gains du canal sont considérés aléatoires, on trouve une expression semblable à l'équation ( 33 ).

$$E\{\boldsymbol{\eta}^\dagger \boldsymbol{\eta}\} = E_\alpha \left\{ E_N \left\{ \boldsymbol{\eta}^\dagger \boldsymbol{\eta} \right\} \right\} = E_\alpha \left\{ \sigma^2 \mathbf{I}_{k \times k} / 2 \left( \sum_{i,j} |\alpha_{i,j}|^2 \right)^2 \right\} = \frac{\sigma^2}{2} \mathbf{I}_{k \times k} \overbrace{E_\alpha \left\{ 1 / \sum_{i,j} |\alpha_{i,j}|^2 \right\}}^{1/nm} \tag{56}$$

$$E\{\boldsymbol{\eta}^\dagger \boldsymbol{\eta}\} = \frac{\sigma^2}{2nm}$$

Alors, l'énergie du bruit équivalent est la même que l'énergie de bruit du système MIMO mais atténuée de  $2mn$ . Le rapport signal à bruit en sortie du décodeur soft se calcule alors :

$$\text{SNR}_q = \frac{1/n}{\sigma^2/2mn} = 2m \times \text{SNR} \quad (57)$$

Mais comme on a vu pour le cas des codes complexes, cette expression n'est pas directement exploitable car le bruit  $\boldsymbol{\eta}$  n'est plus gaussien. On peut donc calculer la probabilité d'erreur de symbole comme on a fait pour le cas des codes complexes. C'est-à-dire en calculant pour le cas du canal constant et en évaluant son espérance sur la distribution de probabilité des coefficients du canal. Sachant que les équations sont identiques à ceux que nous avons écrites pour le cas complexe à condition de redéfinir les paramètres  $G_1$  et  $G_2$ .

$$P_e = P(e|\alpha) \cong G_1 Q(G_2 \sqrt{\text{SNR}_q}) ; \begin{cases} \text{BPSK : } G_1 = 1 ; G_2 = 1 \\ \text{PAM : } G_1 = \frac{2(M-1)}{M} ; G_2 = \frac{2}{M} \end{cases} \quad (58)$$

Dans ce cas les équations ( 41 ) et ( 47 ) sont toujours utilisables.

## 4-6 Résultats des Simulations

Afin d'éprouver la précision des approximations que l'on a trouvées dans la section précédente (équations ( 41 ) et ( 47 )), on utilise deux codes STBC orthogonaux donnés par la suite et en faisant varier le nombre d'antennes de réception. Le premier code, connu sous le nom du code d'Alamouti [9], est un code très simple (2×2), avec deux antennes d'émission et un rendement de code égal à 1. Dans le deuxième code, nous avons un code plus compliqué (4×8), avec quatre antennes d'émission et un rendement de code 0,5. Les matrices de code pour ces deux exemples sont les suivantes :

$$G_1 : \begin{pmatrix} x_1 & -x_2^* \\ x_2 & x_1^* \end{pmatrix} \quad G_2 : \begin{pmatrix} x_1 & -x_2 & -x_3 & -x_4 & x_1^* & -x_2^* & -x_3^* & -x_4^* \\ x_2 & x_1 & x_4 & -x_3 & x_2^* & x_1^* & x_4^* & -x_3^* \\ x_3 & -x_4 & x_1 & x_2 & x_3^* & -x_4^* & x_1^* & x_2^* \\ x_4 & x_3 & -x_2 & x_1 & x_4^* & x_3^* & -x_2^* & x_1^* \end{pmatrix}$$

Une série de simulations a été faite pour ces deux cas dont les résultats sont présentés par la suite.

La Figure 3 présente la performance du code Alamouti ( $G_1$ ) en comparant nos bornes supérieures avec les résultats d'une simulation Monte-Carlo. La simulation a été faite pour différents nombres d'antennes de réception.

Le même principe a été effectué pour la Figure 4 mais utilisant le code  $G_2$ . Il faut bien noter que le nombre d'antennes d'émission est fixé par la matrice du code, ce qui n'est pas le cas pour le nombre d'antennes de réception. La précision de la borne supérieure améliorée (équation ( 47 )) est telle que la différence avec les résultats des simulations Monte-Carlo est inférieure de 0,2 dB pour les forts SNRs.

Figure 5 trace la probabilité d'erreur de symbole en fonction du nombre d'antennes  $m$  de réception pour un SNR donné (8 dB). L'influence du nombre d'antennes de réception est clairement démontrée et la précision de nos bornes est également validée.

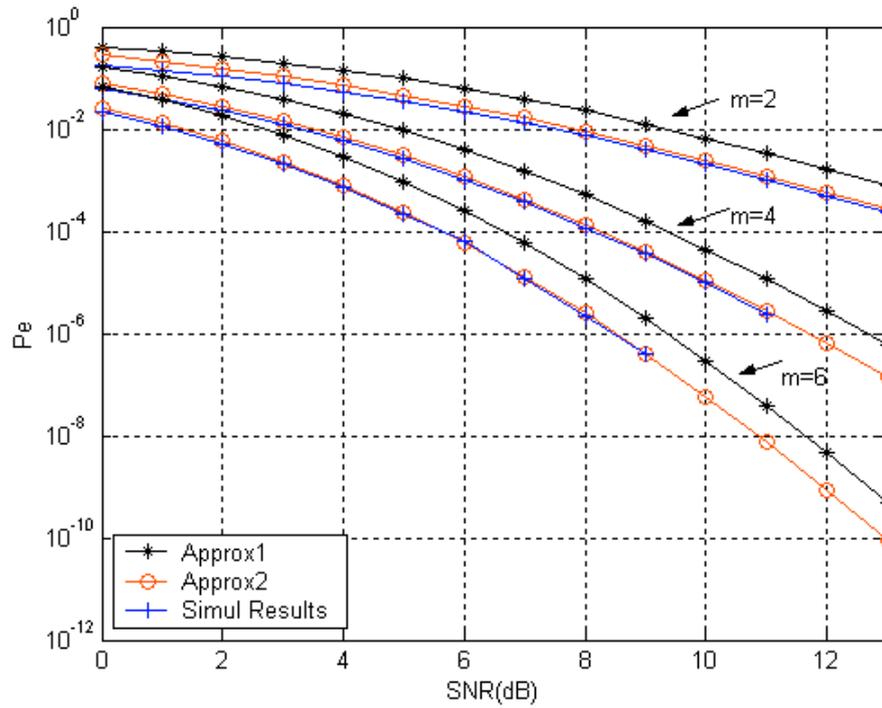


Figure 3 : Probabilité d'erreur de symbole en fonction du SNR ( $E_s/N_0$ ) pour le code G1

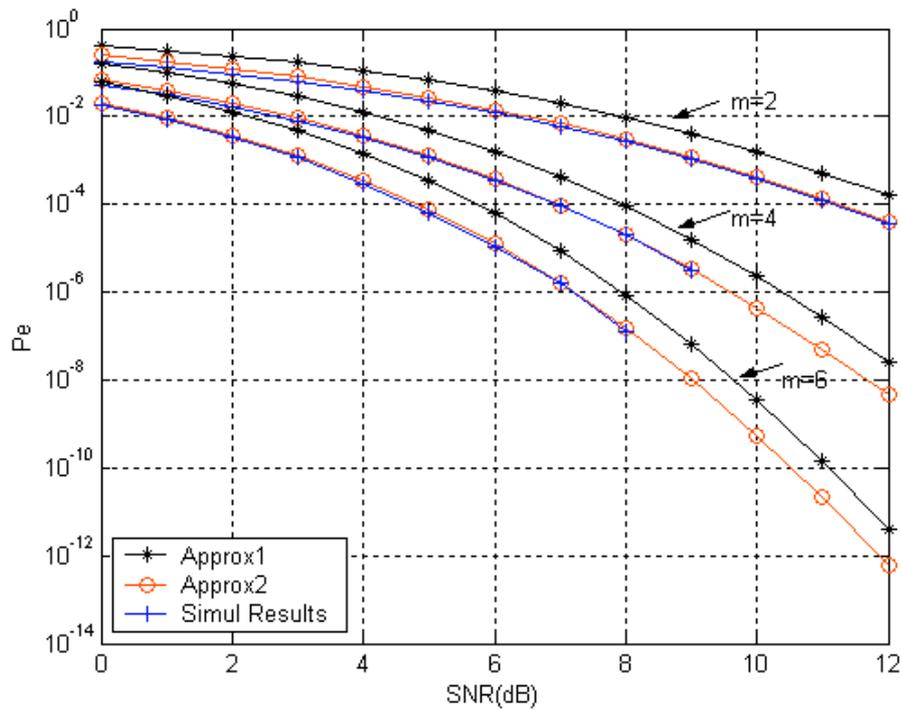
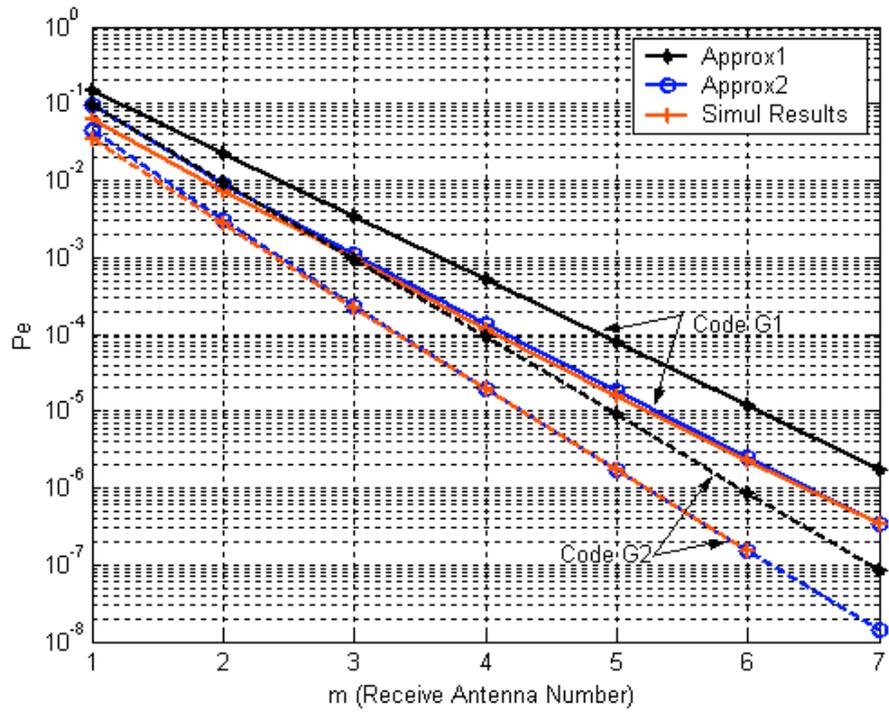


Figure 4 : Probabilité d'erreur de symbole en fonction du SNR ( $E_s/N_0$ ) pour le code G2



**Figure 5 : Probabilité d'erreur de symbole en fonction du nombre d'antennes de réception pour SNR=8dB**

## **4-7 Conclusion**

Les principes des codes STBC sont déjà présents dans la littérature, mais l'objectif de ce travail était de trouver d'abord une démonstration systématique matricielle pour le décodeur ML des codes STBC orthogonaux (complexes ou réels). Cette forme systématique a abouti à un algorithme de décodage très facile à implanter. De plus, nous n'imposons aucune limite sur le nombre d'antennes d'émission, de réception ou rendement de code. La seule condition est que le code doit être orthogonal.

Nous avons considéré le code STBC orthogonal typique, avec  $n$  antennes d'émission,  $m$  antennes de réception et un rendement de  $k/l$ . Ce code peut être représenté par une matrice  $C_{n \times l}$  qui satisfait l'orthogonalité définie dans la section 4-3.

Ensuite, nous avons fait une étude analytique sur la probabilité d'erreur du décodeur ML pour les codes STBC orthogonaux. Pour le cas du canal aux gains constants, nous avons trouvé les valeurs exactes de la probabilité d'erreur pour certaines modulations. Pour le cas plus réaliste où les gains du canal seraient aléatoires, et à partir du taux d'erreur du canal constant, nous avons obtenu deux bornes supérieures pour la probabilité d'erreur symbole. Nous avons ensuite comparé ces deux bornes et vu l'utilité de chacune.

## **4-8 Références**

- [1] N. Balaban and J. Salz, "Dual diversity combining and equalization in digital cellular mobile radio", IEEE Trans. Veh. Technol., vol. 40, pp. 342-354, May 1991.
- [2] G. J. Foschini Jr. and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas", *Wireless Personal Commun.*, Mar. 1998.
- [3] J. C. Guey, M. P. Fitz, M. R. Bell, and W. Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels", in Proc. VTC'96, 1996, pp. 136-140.
- [4] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity", Int. J. Wireless Inform. Networks, vol. 1, n° 1, 1994.
- [5] E. Telatar, "Capacity of multi-antenna Gaussian channels", AT&T-Bell Laboratories Internal Tech. Memo., June 1995.
- [6] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communications : Performance analysis and code construction", IEEE trans. Inform. Theory, vol. 44, pp. 744-765, Mar. 1998.
- [7] V. Tarokh, A. Naguib, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communications : Performance criteria in the presence of channel estimation errors, mobility and multipaths", IEEE trans. Commun., vol. 47, pp. 199-207, Feb. 1999.
- [8] S. Baro, G. Bauch and A. Hansmann, "Improved codes for space-time trellis-coded modulation", IEEE Commun. Lett., vol. 4, pp. 20-22, Jan. 2000.
- [9] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications", IEEE J. Select. Areas. Commun., vol. 16, pp. 1451-1458, Oct. 1998.
- [10] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "Space-Time Block Codes from Orthogonal Designs", IEEE trans. Inform Theory, vol. 45, pp. 1456-1467, July 1999.
- [11] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "Space-Time Block Coding for wireless communications : Performance results", IEEE J. Select. Areas Commun., vol. 17, pp. 451-460, Mar. 1999.
- [12] S. Benedetto and E. Biglieri, Principles of Digital Transmission With Wireless Applications, Plenum Publishers, 1999.
- [13] J. G. Proakis, Digital Communications, Third Edition, McGraw-Hill, Inc., 1995.
- [14] Papoulis, Probability, Random Variable, and Stochastic Processes,

## ***Annexe A : Preuve des particularités des matrices de codes***

En appliquant l'équation (6) dans la définition de l'orthogonalité de code on aura :

$$\begin{aligned}
 \mathbf{C}\mathbf{C}^\dagger &= \left( \sum_{i=1}^k (x_i \mathbf{A}_i + x_i^* \mathbf{B}_i) \right) \left( \sum_{j=1}^k (x_j \mathbf{A}_j + x_j^* \mathbf{B}_j) \right)^\dagger \\
 &= \left( \sum_{i=1}^k (x_i \mathbf{A}_i + x_i^* \mathbf{B}_i) \right) \left( \sum_{i=1}^k (x_j^* \mathbf{A}_j^\dagger + x_j \mathbf{B}_j^\dagger) \right) \\
 &= \sum_{i=1}^k \sum_{j=1}^k (x_i x_j^* \mathbf{A}_i \mathbf{A}_j^\dagger + x_i x_j \mathbf{A}_i \mathbf{B}_j^\dagger + x_i^* x_j^* \mathbf{B}_i \mathbf{A}_j^\dagger + x_i^* x_j \mathbf{B}_i \mathbf{B}_j^\dagger) \\
 &= \sum_{i=1}^k |x_i|^2 (\mathbf{A}_i \mathbf{A}_i^\dagger + \mathbf{B}_i \mathbf{B}_i^\dagger) + \sum_{i=1}^k x_i^2 (\mathbf{A}_i \mathbf{B}_i^\dagger) + \sum_{i=1}^k (x_i^*)^2 (\mathbf{B}_i \mathbf{A}_i^\dagger) \\
 &\quad + \sum_{i=1}^k \sum_{j=i+1}^k (x_i x_j^* (\mathbf{A}_i \mathbf{A}_j^\dagger + \mathbf{B}_j \mathbf{B}_i^\dagger) + x_i^* x_j (\mathbf{A}_j \mathbf{A}_i^\dagger + \mathbf{B}_i \mathbf{B}_j^\dagger)) \\
 &\quad + \sum_{i=1}^k \sum_{j=i+1}^k (x_i x_j (\mathbf{A}_i \mathbf{B}_j^\dagger + \mathbf{A}_j \mathbf{B}_i^\dagger) + x_i^* x_j^* (\mathbf{B}_i \mathbf{A}_j^\dagger + \mathbf{B}_j \mathbf{A}_i^\dagger))
 \end{aligned}$$

Pour que les matrices  $\mathbf{A}_i$  et  $\mathbf{B}_i$  satisfassent l'orthogonalité de code ( $\mathbf{C}\mathbf{C}^\dagger = \sum_{i=1}^k |x_i|^2 \mathbf{I}_{n \times n}$ ) pour n'importe quels symboles d'informations  $(x_1, x_2, \dots, x_k)$ , les coefficients de  $x_i^2$ ,  $x_i x_j^*$  et  $x_i x_j$  doivent être égaux à zéros et les coefficients de  $|x_i|^2$  égaux à la matrice unité. Donc les matrices  $\mathbf{A}_i$  et  $\mathbf{B}_i$  doivent satisfaire l'équation (7).

### **Cas des codes réels :**

On applique encore la matrice de codes dans la définition de l'orthogonalité :

$$\begin{aligned}
 \mathbf{C}\mathbf{C}^\mathbf{T} &= \left( \sum_{i=1}^k x_i \mathbf{A}_i \right) \left( \sum_{j=1}^k x_j \mathbf{A}_j \right)^\mathbf{T} = \left( \sum_{i=1}^k x_i \mathbf{A}_i \right) \left( \sum_{i=1}^k x_j \mathbf{A}_j^\mathbf{T} \right) \\
 &= \sum_{i=1}^k \sum_{j=1}^k (x_i x_j \mathbf{A}_i \mathbf{A}_j^\mathbf{T}) = \sum_{i=1}^k (x_i^2 \mathbf{A}_i \mathbf{A}_i^\mathbf{T}) + \sum_{i=1}^k \sum_{j=i+1}^k (x_i x_j (\mathbf{A}_i \mathbf{A}_j^\mathbf{T} + \mathbf{A}_j \mathbf{A}_i^\mathbf{T}))
 \end{aligned}$$

Pour que les matrices  $\mathbf{A}_i$  satisfassent l'orthogonalité de code ( $\mathbf{C}\mathbf{C}^\mathbf{T} = \sum_{i=1}^k x_i^2 \mathbf{I}_{n \times n}$ ) pour n'importe quels symboles d'informations  $(x_1, x_2, \dots, x_k)$ , les coefficients de  $x_i x_j$  soient égaux à zéros et les coefficients de  $x_i^2$  égaux à la matrice unité. Donc les matrices  $\mathbf{A}_i$  doivent satisfaire l'équation (8).



## Annexe B : Preuve des particularités des matrices $H_R^j$ et $H_I^j$

D'abord on va montrer que la matrice  $H_R^j \cdot (H_I^j)^\dagger$  est antisymétrique (et avec les éléments zéros sur le diagonal principal)

$$H_R^j \cdot (H_I^j)^\dagger = \begin{pmatrix} \mathbf{h}^j A_1 \\ \vdots \\ \mathbf{h}^j A_i \\ \vdots \\ \mathbf{h}^j A_k \end{pmatrix}_{k \times l} \cdot \begin{pmatrix} \mathbf{h}^j B_1 \\ \vdots \\ \mathbf{h}^j B_q \\ \vdots \\ \mathbf{h}^j B_k \end{pmatrix}_{l \times k}^\dagger = \begin{pmatrix} \mathbf{h}^j A_1 \\ \vdots \\ \mathbf{h}^j A_i \\ \vdots \\ \mathbf{h}^j A_k \end{pmatrix}_{k \times l} \cdot \left( B_1^\dagger(\mathbf{h}^j)^\dagger \quad \dots \quad B_q^\dagger(\mathbf{h}^j)^\dagger \quad \dots \quad B_k^\dagger(\mathbf{h}^j)^\dagger \right)_{l \times k}$$

Donc l'élément  $(i,q)$  de cette production matricielle est égale à :  $\mathbf{h}^j A_i \cdot B_q^\dagger(\mathbf{h}^j)^\dagger$ , et d'après l'équation (7) on peut apercevoir que  $H_R^j \cdot (H_I^j)^\dagger$  est antisymétrique. Maintenant on passe à l'équation (26-b).

$$H_R^j (H_R^j)^\dagger + (H_I^j)^* (H_I^j)^\dagger = \begin{pmatrix} \mathbf{h}^j A_1 \\ \vdots \\ \mathbf{h}^j A_i \\ \vdots \\ \mathbf{h}^j A_k \end{pmatrix} \left( A_1^\dagger(\mathbf{h}^j)^\dagger \dots A_k^\dagger(\mathbf{h}^j)^\dagger \right) + \begin{pmatrix} (\mathbf{h}^j)^* B_1^* \\ \vdots \\ (\mathbf{h}^j)^* B_i^* \\ \vdots \\ (\mathbf{h}^j)^* B_k^* \end{pmatrix} \left( B_1^\dagger(\mathbf{h}^j)^\dagger \dots B_k^\dagger(\mathbf{h}^j)^\dagger \right)$$

Donc d'après l'équation (7) l'élément  $(i,i)$  de la matrice résultat est :

$$\begin{aligned} \mathbf{h}^j A_i A_i^\dagger (\mathbf{h}^j)^\dagger + \overbrace{(\mathbf{h}^j)^* B_i^* B_i^\dagger (\mathbf{h}^j)^\dagger}^{\text{scalaire réel}} &= \mathbf{h}^j A_i A_i^\dagger (\mathbf{h}^j)^\dagger + \mathbf{h}^j B_i B_i^\dagger (\mathbf{h}^j)^\dagger \\ &= \mathbf{h}^j (A_i A_i^\dagger + B_i B_i^\dagger) (\mathbf{h}^j)^\dagger = \mathbf{h}^j \cdot \mathbf{I}_{n \times n} \cdot (\mathbf{h}^j)^\dagger = \mathbf{h}^j \cdot (\mathbf{h}^j)^\dagger = \sum_{i=1}^n |\alpha_{i,j}|^2 \end{aligned}$$

Et l'élément  $(i,q)$  sera comme suivant :

$$\begin{aligned} \mathbf{h}^j A_i A_q^\dagger (\mathbf{h}^j)^\dagger + \overbrace{(\mathbf{h}^j)^* B_i^* B_q^\dagger (\mathbf{h}^j)^\dagger}^{\text{scalaire réel}} &= \mathbf{h}^j A_i A_q^\dagger (\mathbf{h}^j)^\dagger + \mathbf{h}^j B_q B_i^\dagger (\mathbf{h}^j)^\dagger \quad i \neq q \\ &= \mathbf{h}^j (A_i A_q^\dagger + B_q B_i^\dagger) (\mathbf{h}^j)^\dagger = \mathbf{h}^j \cdot \mathbf{0}_{n \times n} \cdot (\mathbf{h}^j)^\dagger = 0 \end{aligned}$$

$$\text{Donc } H_R^j (H_R^j)^\dagger + (H_I^j)^* (H_I^j)^\dagger = \sum_{i=1}^n |\alpha_{i,j}|^2 \cdot \mathbf{I}_{k \times k}$$

### Cas des codes réels :

D'abord on va montrer que la matrice  $\mathbf{H}_R^j \cdot (\mathbf{H}_R^j)^\dagger$  est antisymétrique (mais avec  $\sum_i |\alpha_{i,j}|^2$  sur le diagonal principal).

$$\mathbf{H}_R^j \cdot (\mathbf{H}_R^j)^\dagger = \begin{pmatrix} \mathbf{h}^j \mathbf{A}_1 \\ \vdots \\ \mathbf{h}^j \mathbf{A}_i \\ \vdots \\ \mathbf{h}^j \mathbf{A}_k \end{pmatrix}_{k \times l} \cdot \begin{pmatrix} (\mathbf{h}^j \mathbf{A}_1)^\dagger \\ \vdots \\ (\mathbf{h}^j \mathbf{A}_q)^\dagger \\ \vdots \\ (\mathbf{h}^j \mathbf{A}_k)^\dagger \end{pmatrix}_{l \times k} = \begin{pmatrix} \mathbf{h}^j \mathbf{A}_1 \\ \vdots \\ \mathbf{h}^j \mathbf{A}_i \\ \vdots \\ \mathbf{h}^j \mathbf{A}_k \end{pmatrix}_{k \times l} \cdot \left( \mathbf{A}_1^\dagger (\mathbf{h}^j)^\dagger \quad \dots \quad \mathbf{A}_q^\dagger (\mathbf{h}^j)^\dagger \quad \dots \quad \mathbf{A}_k^\dagger (\mathbf{h}^j)^\dagger \right)_{l \times k}$$

Donc d'après l'équation (8), l'élément diagonal  $(i,i)$  de matrice résultat est :

$$\mathbf{h}^j \mathbf{A}_i \cdot \mathbf{A}_i^\dagger (\mathbf{h}^j)^\dagger = \mathbf{h}^j \cdot \mathbf{I}_{n \times n} \cdot (\mathbf{h}^j)^\dagger = \mathbf{h}^j \cdot (\mathbf{h}^j)^\dagger = \sum_{i=1}^n |\alpha_{i,j}|^2$$

On peut également vérifier, encore en utilisant l'équation (8), que la somme de l'élément  $(i,q)$  et  $(q,i)$  est égale à zéro (Antisymétrie).

$$\mathbf{h}^j \mathbf{A}_i \cdot \mathbf{A}_q^\dagger (\mathbf{h}^j)^\dagger + \mathbf{h}^j \mathbf{A}_q \cdot \mathbf{A}_i^\dagger (\mathbf{h}^j)^\dagger = \mathbf{h}^j \cdot \underbrace{(\mathbf{A}_i \cdot \mathbf{A}_q^\dagger + \mathbf{A}_q \cdot \mathbf{A}_i^\dagger)}_{\mathbf{0}_{n \times n}} \cdot (\mathbf{h}^j)^\dagger = 0$$

Donc la matrice  $\mathbf{H}_{RR} = \mathbf{H}_R^j \cdot (\mathbf{H}_R^j)^\dagger$  est antisymétrique avec  $\sum_i |\alpha_{i,j}|^2$  sur le diagonal.

D'ailleurs cette matrice est hermitien, donc à part les éléments diagonaux, les autres ont des valeurs imaginaires, ce qui peut être trouvé très facilement. Pour cela, on montre l'élément  $(i,q)$  de la matrice  $\mathbf{H}_{RR}$  par  $h_{RR}(i,q)$ .

D'après l'hermécité  $\Rightarrow h_{RR}(q,i) = h_{RR}^*(i,q)$

$$i \neq q \quad \Rightarrow h_{RR}^*(i,q) = -h_{RR}(i,q) \quad \Rightarrow \text{real}(h_{RR}(i,q)) = 0$$

D'après l'anti-symétrie  $\Rightarrow h_{RR}(q,i) = -h_{RR}(i,q)$

Donc l'élément non diagonaux sont imaginaires et par conséquent.

$$\text{real}(\mathbf{H}_R \cdot (\mathbf{H}_R)^\dagger) = \sum_{i=1}^n \sum_{j=1}^m |\alpha_{i,j}|^2 \cdot \mathbf{I}_{k \times k}$$

## *Annexe C : Particularité des matrices antisymétriques*

Dans cette annexe on présente la définition et propriétés des matrices antisymétriques. Une matrice antisymétrique est une matrice carrée qui satisfait la condition suivante :

$$\mathbf{G}_{n \times n} : \text{Antisymétrique} \Leftrightarrow \mathbf{G}^T = -\mathbf{G}$$

Il est clair que les éléments de diagonales sont égaux à zéros. Certaines propriétés de ces matrices sont les suivantes :

1.  $\mathbf{G}_{n \times n} : \text{Antisymétrique} \Leftrightarrow \mathbf{G}^T, \mathbf{G}^\dagger, \mathbf{G}^* : \text{Antisymétrique}$
2.  $\mathbf{G}_{n \times n}$  et  $\mathbf{J}_{n \times n} : \text{Antisymétrique} \Rightarrow \mathbf{G} \pm \mathbf{J} : \text{Antisymétrique}$
3. Si on multiplie une matrice antisymétrique par une matrice quelconque d'un côté et par sa transposée de cette matrice de l'autre côté, le résultat sera encore une matrice antisymétrique.

$$\mathbf{G}_{n \times n} : \text{Antisymétrique} ; \mathbf{Q}_{m \times m} = \mathbf{S}_{m \times n} \mathbf{G}_{n \times n} \mathbf{S}_{n \times m}^T \Rightarrow \mathbf{Q}_{m \times m} : \text{Antisymétrique}$$

Il suffit juste de montrer que  $\mathbf{Q}^T = -\mathbf{Q}$ .

$$\mathbf{Q} = \mathbf{S} \cdot \mathbf{G} \cdot \mathbf{S}^T \Rightarrow \mathbf{Q}^T = \mathbf{S} \cdot \mathbf{G}^T \cdot \mathbf{S}^T = -\mathbf{S} \cdot \mathbf{G} \cdot \mathbf{S}^T = -\mathbf{Q}$$

Si  $\mathbf{S}$  est un vecteur de taille  $1 \times n$ ,  $\mathbf{Q}$  sera scalaire et  $\mathbf{Q}^T = \mathbf{Q}$ , d'ailleurs  $\mathbf{Q}^T = -\mathbf{Q}$ , donc  $\mathbf{Q}$  sera égale à zéro.



## ***Annexe D : Densité de probabilité de Gamma***

La densité de Gamma se définit comme suivante [14] :

$$y \sim \text{Gamma}(c, b) \Leftrightarrow f_Y(y) = \frac{c^{b+1}}{\Gamma(b+1)} y^b e^{-cy} U(y)$$

où des constants  $b \geq 0$  et  $c > 0$  sont des paramètres de la densité et  $\Gamma(b+1)$  est connu sous le nom de la fonction de Gamma.

$$\Gamma(b+1) = \int_0^\infty x^b e^{-x} dx$$

On peut montrer que pour cette densité les moments  $m_k$  s'écrivent sous la forme suivante [14] :

$$m_k = E\{y^k\} = \int_0^\infty y^k \frac{c^{b+1}}{\Gamma(b+1)} y^b e^{-cy} dy = \frac{c^{b+1}}{\Gamma(b+1)} \int_0^\infty y^{b+k} e^{-cy} dy$$

Maintenant avec un changement de variable, pour  $k \geq -b$  on aura :

$$m_k = E\{y^k\} = \frac{c^{b+1}}{\Gamma(b+1) c^{b+k+1}} \int_0^\infty x^{b+k} e^{-x} dx = \frac{\Gamma(b+k+1) c^k}{\Gamma(b+1)}$$

Si  $b = n$  est un entier,  $\Gamma(n+1) = n!$  et on appelle cette densité de probabilité, la densité Gamma de  $n^{\text{ième}}$  degré.

$$y \sim \text{Gamma}(c, n) \Leftrightarrow f_Y(y) = \frac{c^{n+1}}{n!} y^n e^{-cy} U(y)$$

Donc pour une densité  $\text{Gamma}(c, n)$ , les moments  $m_k$  seront comme suivant :

$$m_k = E\{y^k\} = \frac{(b+k)!}{b!} c^k \quad \text{pour } k \geq -b$$

Maintenant on montre que si  $x$  est une variable gaussienne complexe avec la moyenne zéro et la variance  $\sigma^2/2$  pour chaque partie réelle et imaginaire (ces deux-là sont indépendantes),  $|x|^2$  aura une densité de Gamma de zéro degré.

$$\begin{aligned} x_R \sim N(0, \sigma^2/2) \\ x_I \sim N(0, \sigma^2/2) \end{aligned} ; x_R \text{ et } x_I \text{ indépendants} \quad \Rightarrow \quad y = x_R^2 + x_I^2 \sim \text{Gamma}(\sigma^2, 0)$$

Preuve :

$$y = x_R^2 + x_I^2$$

$$\theta = \arctg\left(\frac{x_I}{x_R}\right) \Rightarrow f_{Y,\Theta}(y,\theta) = \frac{f_{X_R,X_I}(x_R,x_I)}{\begin{vmatrix} 2x_R & 2x_I \\ x_I & x_R \\ x_R^2 + x_I^2 & x_R^2 + x_I^2 \end{vmatrix}} = \frac{1}{4\pi(\sigma^2/2)} e^{-\frac{x_R^2+x_I^2}{2(\sigma^2/2)}} = \frac{1}{2\pi\sigma^2} e^{-\frac{y}{\sigma^2}}$$

$$f_Y(y) = \int f_{Y,\Theta}(y,\theta) d\theta \Rightarrow f_Y(y) = \frac{1}{\sigma^2} e^{-\frac{y}{\sigma^2}}$$

Maintenant on montre par récurrence que la somme de  $n$  variables aléatoires indépendantes de type  $Gamma(c,0)$  sera une variable aléatoire de type  $Gamma(c,n-1)$ .

On suppose que  $x_i \sim Gamma(c,0)$  ; pour  $i = 1, \dots, n$ .

Pour  $i = 1$  :  $y_1 = x_1 \Rightarrow y_1 \sim Gamma(c,0)$ .

Si pour  $i = n-1$  :  $y_{n-1} = \sum_{j=1}^{n-1} x_j \sim Gamma(c,n-2)$

On va montrer que pour  $i = n$  :  $y_{n-1} + x_n = y_n \sim Gamma(c,n-1)$ .

Preuve :  $y = y_{n-1}$  ;  $x = x_n$  ;  $z = y_n = x + y$

$y \sim Gamma(c,n-2)$  ;  $y$  et  $x$  indépendants  $\Rightarrow f_{X,Y}(x,y) = \frac{c^{n-1}}{(n-2)!} y^{n-2} e^{-cy} c e^{-cx} U(x)U(y)$   
 $x \sim Gamma(c,0)$

$$f_{X,Y}(x,y) = \frac{c^{n-1}}{(n-2)!} y^{n-2} e^{-c(x+y)} U(x)U(y) ; \quad f_Z(z) = \int_{-\infty}^{+\infty} f_{X,Y}(z-y,y) dy$$

$$f_Z(z) = \int_0^z \frac{c^n}{(n-2)!} y^{n-2} e^{-cz} dy = \frac{c^n}{(n-2)!} e^{-cz} \int_0^z y^{n-2} dy = \frac{c^n}{(n-1)!} z^{n-1} e^{-cz} U(z)$$

## Annexe E : Repère polaire de dimension $M$

On va montrer que dans un changement de repère du cartésien  $(y_1, y_2, \dots, y_M)$  au polaire de  $M$  dimension  $(\rho, \varphi_1, \varphi_2, \dots, \varphi_{M-1})$  comme suivant :

$$y_1 = \rho \cos \varphi_1 \quad ; \quad y_M = \rho \prod_{j=1}^{M-1} \sin \varphi_j$$

$$y_i = \rho \left( \prod_{j=1}^{i-1} \sin \varphi_j \right) \cos \varphi_i \quad ; \quad i = 1, \dots, M-1$$

On aura :

$$dy_1 dy_2 \dots dy_M = \rho^{M-1} (\sin \varphi_1)^{M-2} \dots (\sin \varphi_i)^{M-1-i} \dots \sin \varphi_{M-2} d\rho d\varphi_1 d\varphi_2 \dots d\varphi_{M-1}$$

Pour cela on doit développer la formule suivante.

$$dy_1 dy_2 \dots dy_M = \underbrace{\begin{vmatrix} \frac{\partial y_1}{\partial \rho} & \frac{\partial y_1}{\partial \varphi_1} & \dots & \frac{\partial y_1}{\partial \varphi_{M-1}} \\ \frac{\partial y_2}{\partial \rho} & \frac{\partial y_2}{\partial \varphi_1} & \dots & \frac{\partial y_2}{\partial \varphi_{M-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_M}{\partial \rho} & \frac{\partial y_M}{\partial \varphi_1} & \dots & \frac{\partial y_M}{\partial \varphi_{M-1}} \end{vmatrix}}_{\Delta_M(\rho, \varphi_1, \dots, \varphi_{M-1})} d\rho d\varphi_1 d\varphi_2 \dots d\varphi_{M-1}$$

Donc on va montrer par récurrence que  $\Delta_M(\rho, \varphi_1, \dots, \varphi_{M-1}) = \rho^{M-1} \prod_{i=1}^{M-1} (\sin \varphi_i)^{M-1-i}$ .

Pour  $i = 2$  :  $dy_1 dy_2 = \rho d\rho d\varphi_1 \Rightarrow \Delta_2(\rho, \varphi_1) = \begin{vmatrix} \cos \varphi_1 & -\rho \sin \varphi_1 \\ \sin \varphi_1 & \rho \cos \varphi_1 \end{vmatrix} = \rho$

Pour  $i = M-1$  :  $\Delta_{M-1}(\rho, \varphi_2, \dots, \varphi_{M-1}) = \rho^{M-2} \prod_{i=2}^{M-1} (\sin \varphi_i)^{M-1-i}$

Pour  $i = M \geq 2$  :

$$\Delta_M = \begin{vmatrix} \cos \varphi_1 & -\rho \sin \varphi_1 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \sin \varphi_1 \cos \varphi_2 & \rho \cos \varphi_1 \cos \varphi_2 & -\rho \sin \varphi_1 \sin \varphi_2 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & & & & & & \vdots \\ \prod_{j=1}^{i-1} \sin \varphi_j \cdot \cos \varphi_i & \rho \cos \varphi_1 \cdot \prod_{j=2}^{i-1} \sin \varphi_j \cdot \cos \varphi_i & \dots & \dots & \rho \cos \varphi_k \cdot \prod_{\substack{j=1 \\ j \neq k}}^{i-1} \sin \varphi_j \cdot \cos \varphi_i & \dots & -\rho \prod_{j=1}^i \sin \varphi_j & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & & \ddots & & & \vdots \\ \prod_{j=1}^{M-2} \sin \varphi_j \cdot \cos \varphi_{M-1} & \rho \cos \varphi_1 \cdot \prod_{j=2}^{M-2} \sin \varphi_j \cdot \cos \varphi_{M-1} & \dots & \dots & \rho \cos \varphi_k \cdot \prod_{\substack{j=1 \\ j \neq k}}^{M-2} \sin \varphi_j \cdot \cos \varphi_{M-1} & \dots & \dots & \dots & \dots & -\rho \prod_{j=1}^{M-1} \sin \varphi_j \\ \prod_{j=1}^{M-1} \sin \varphi_j & \rho \cos \varphi_1 \cdot \prod_{j=2}^{M-1} \sin \varphi_j & \dots & \dots & \rho \cos \varphi_k \cdot \prod_{\substack{j=1 \\ j \neq k}}^{M-1} \sin \varphi_j & \dots & \dots & \dots & \dots & \rho \cos \varphi_{M-1} \cdot \prod_{j=1}^{M-1} \sin \varphi_j \end{vmatrix}$$

Pour calculer ce déterminant on le développe sur la première ligne. Dans les déterminants de  $\Delta_{1,1}$  et  $\Delta_{1,2}$ , il existe respectivement des termes de  $\rho \cos \varphi_1$  et  $\sin \varphi_1$  à factoriser et après ces factorisations, on arrive à une matrice identique.

$$\Delta_M = \cos \varphi_1 \Delta_{1,1} + \rho \sin \varphi_1 \Delta_{1,2} = \rho (\cos \varphi_1)^2 |\Lambda| + \rho (\sin \varphi_1)^2 |\Lambda| = \rho |\Lambda|$$

où

$$|\Lambda| = \begin{vmatrix} \cos \varphi_2 & -\rho \sin \varphi_1 \sin \varphi_2 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & \vdots \\ \prod_{j=2}^{i-1} \sin \varphi_j \cdot \cos \varphi_i & \dots & \dots & \rho \cos \varphi_k \cdot \prod_{\substack{j=1 \\ j \neq k}}^{i-1} \sin \varphi_j \cdot \cos \varphi_i & \dots & -\rho \prod_{j=1}^i \sin \varphi_j & 0 & \dots & 0 \\ \vdots & \vdots \\ \prod_{j=2}^{M-2} \sin \varphi_j \cdot \cos \varphi_{M-1} & \dots & \dots & \rho \cos \varphi_k \cdot \prod_{\substack{j=1 \\ j \neq k}}^{M-2} \sin \varphi_j \cdot \cos \varphi_{M-1} & \dots & \dots & \dots & \dots & -\rho \prod_{j=1}^{M-1} \sin \varphi_j \\ \prod_{j=2}^{M-1} \sin \varphi_j & \dots & \dots & \rho \cos \varphi_k \prod_{\substack{j=1 \\ j \neq k}}^{M-1} \sin \varphi_j & \dots & \dots & \dots & \dots & \rho \cos \varphi_{M-1} \prod_{j=1}^{M-1} \sin \varphi_j \end{vmatrix}$$

Et dans ce déterminant dans toutes les colonnes (à part le premier), il existe un terme de  $\sin \varphi_1$  à factoriser et après cette factorisation on arrive à une matrice de la même structure que  $\Delta_M$ , mais avec la dimension  $M - 1$  et par rapport aux variables  $(\rho, \varphi_2, \dots, \varphi_{M-1})$ .

$$|\Lambda| = (\sin \varphi_1)^{M-2} \begin{vmatrix} \cos \varphi_2 & -\rho \sin \varphi_2 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & \vdots \\ \prod_{j=2}^{i-1} \sin \varphi_j \cdot \cos \varphi_i & \dots & \dots & \rho \cos \varphi_k \cdot \prod_{\substack{j=2 \\ j \neq k}}^{i-1} \sin \varphi_j \cdot \cos \varphi_i & \dots & -\rho \prod_{j=2}^i \sin \varphi_j & 0 & \dots & 0 \\ \vdots & \vdots \\ \prod_{j=2}^{M-2} \sin \varphi_j \cdot \cos \varphi_{M-1} & \dots & \dots & \rho \cos \varphi_k \cdot \prod_{\substack{j=2 \\ j \neq k}}^{M-2} \sin \varphi_j \cdot \cos \varphi_{M-1} & \dots & \dots & \dots & \dots & -\rho \prod_{j=2}^{M-1} \sin \varphi_j \\ \prod_{j=2}^{M-1} \sin \varphi_j & \dots & \dots & \rho \cos \varphi_k \prod_{\substack{j=2 \\ j \neq k}}^{M-1} \sin \varphi_j & \dots & \dots & \dots & \dots & \rho \cos \varphi_{M-1} \prod_{j=2}^{M-1} \sin \varphi_j \end{vmatrix}$$

$$|\Lambda| = (\sin \varphi_1)^{M-2} \Delta_{M-1}(\rho, \varphi_2, \dots, \varphi_{M-1})$$

On peut donc réécrire  $\Delta_M$  sous la forme suivante.

$$\begin{cases} \Delta_M = \rho |\Lambda| = \rho (\sin \varphi_1)^{M-2} \Delta_{M-1}(\rho, \varphi_2, \dots, \varphi_{M-1}) \\ \Delta_{M-1}(\rho, \varphi_2, \dots, \varphi_{M-1}) = \rho^{M-2} \prod_{i=2}^{M-1} (\sin \varphi_i)^{M-1-i} \end{cases} \Rightarrow \Delta_M(\rho, \varphi_1, \dots, \varphi_{M-1}) = \rho^{M-1} \prod_{i=1}^{M-1} (\sin \varphi_i)^{M-1-i}$$

On a finalement trouvé  $\Delta_M(\rho, \varphi_1, \dots, \varphi_{M-1})$ , et on peut maintenant utiliser le repère polaire en

remplaçant  $dy_1 dy_2 \dots dy_M$  par  $\rho^{M-1} \prod_{j=1}^{M-1} (\sin \varphi_j)^{M-1-i} \cdot d\rho d\varphi_1 d\varphi_2 \dots d\varphi_{M-1}$ .

**Preuve de l'équation (46) :**  $I = \int_0^\infty \rho^{2K} e^{-\rho^2} d\rho = \frac{(2K)!}{K! 2^{2K+1}} \sqrt{\pi}$

Avec un changement de variable  $\rho = y/\sqrt{2}$  on aura :

$$I = \frac{1}{2^K \sqrt{2}} \int_0^\infty y^{2K} e^{-y^2/2} dy = \frac{1}{2^K \sqrt{2}} I_K$$

Donc on va montrer par récurrence que :

$$I_K = \int_0^\infty y^{2K} e^{-y^2/2} dy = \frac{(2K)!}{K! 2^{K+1}} \sqrt{2\pi}$$

Pour  $i = 0$  :  $I_0 = \int_0^\infty e^{-y^2/2} dy = \frac{\sqrt{2\pi}}{2}$

Pour  $i = K - 1$ , nous avons  $I_{K-1} = \int_0^\infty y^{2K-2} e^{-y^2/2} dy = \frac{(2K-2)!}{(K-1)! 2^K} \sqrt{2\pi}$ .

Pour  $i = K \geq 1$  :

$$I_K = \int_0^\infty y^{2K} e^{-y^2/2} dy = \int_0^\infty \underbrace{(y^{2K-1})}_u \cdot \underbrace{y e^{-y^2/2}}_{dv} dy = \left[ -y^{2K-1} e^{-y^2/2} \right]_0^\infty + (2K-1) \int_0^\infty y^{2K-2} e^{-y^2/2} dy$$

$$I_K = (2K-1)I_{K-1} = (2K-1) \frac{(2K-2)!}{(K-1)! 2^K} \sqrt{2\pi} = \frac{2K}{2K} \frac{(2K-1)!}{(K-1)! 2^K} \sqrt{2\pi} = \frac{(2K)!}{K! 2^{K+1}} \sqrt{2\pi}$$



# **CHAPITRE 5**

## **CODAGE TEMPS-ESPACE EN TREILLIS (SPACE-TIME-TRELLIS-CODED MODULATION)**

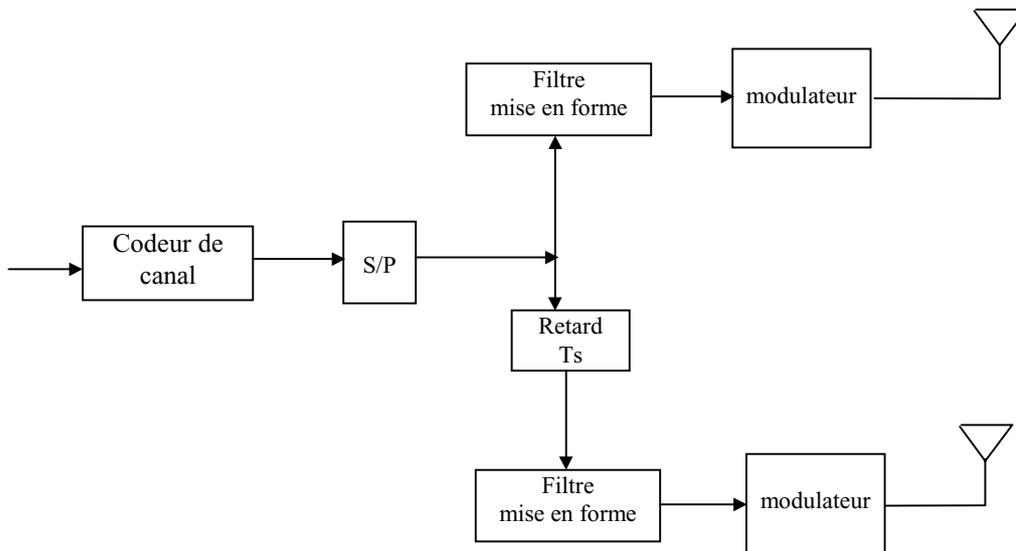


## **5- Codage Temps-Espace en Treillis (Space-Time-Trellis-Coded Modulation)**

Dans ce chapitre, nous présentons un codage temps-espace qui associe le principe de la diversité spatio-temporelle et le codage convolutif (code en treillis). Ce type de codage est généralement considéré comme un concurrent du codage spatio-temporel en bloc (STBC : Space-Time Block Code). Cependant, contrairement au cas des codes STBC qui permettent d'obtenir une diversité optimale, le codage STTC introduit de la corrélation au niveau des symboles transmis entre les différentes antennes. Ceci explique que les performances de ce type de code sont en général supérieures à celles des codes STBC, au prix d'une complexité supérieure. Le but de ce chapitre est de présenter les critères de construction des codes STTC pour obtenir les meilleures performances en termes de FER (Frame Error Rate). Dans la première partie, nous donnons une première approche intuitive du codage STTC. La seconde partie traite des critères de construction des codes STTC. La partie 3 montre des exemples de codes STTC et leurs performances par comparaison à la capacité Outage du canal. Enfin, en partie 4, s'inspirant des techniques de construction des modulations MTCM à partir des critères de Ungerboeck, nous étudions une nouvelle catégorie de codes (SOSTTC) capables de combiner le gain de codage des codes STTC et le gain en diversité optimal des codes STBC.

### **5-1 Une première approche du codage temps-espace en treillis (STTC)**

Une première approche de ce type de codage a été donnée dans [1] par Wittneben. Le principe de transmission proposé peut être résumé par la figure ci-dessous (Fig.1). L'information est codée par un codeur de canal (ici il s'agit d'un simple code à répétition de longueur 2). La sortie du code à répétition est ensuite partagée en deux trains de données parallèles avec un retard de un symbole entre les deux trains. Il n'y a aucune perte en termes de bande passante due à l'utilisation d'un code à répétition puisque deux symboles sont transmis en sortie à chaque intervalle de temps.



**Fig. 1: schéma d'un émetteur à diversité d'émission**

Seshadri et Winters ont montré dans [2] que ce schéma de transmission avait pour effet de transformer un canal de propagation non sélectif en fréquence en un canal sélectif en fréquence. Un récepteur à maximum de vraisemblance est alors capable de parfaitement exploiter la diversité créée en émission. A la lueur de cet exemple simple, il est naturel de se poser la question de savoir s'il est possible de choisir un codeur de canal plus performant que le simple code à répétition de rendement  $\frac{1}{2}$  mentionné ci-dessus.

Tarokh a alors montré en partant de cet exemple simple qu'il était effectivement possible de construire des schémas de transmission avec diversité spatiale, autres que des codes STBC, utilisant des codeurs de canal performants et ne nécessitant plus d'introduire un délai entre les branches des antennes émettrices. Il a proposé des structures de codes spatio-temporels en treillis pour des modulations de type MDP-4, MDP-8 ou MAQ-16 qui permettent de travailler à 2-3 dB de la capacité de coupure ("Outage Capacity") du canal [3]. Nous avons réalisé et simulé la plupart des codes proposés par Tarokh dans [3]. De plus, nous proposons dans le chapitre 4, des codes originaux appelés SOSTTC qui combinent le gain de codage des codes STTC avec le gain en diversité optimal des codes STBC. La construction de ces codes est basée sur une règle de partitionnement de la constellation initiale en sous-constellations comme pour la construction des modulations codées en treillis de Ungerboeck. Le critère de partitionnement utilisé est basé sur le calcul du déterminant de la matrice des différences de mots de code. On affecte alors aux différentes sous-constellations ainsi définies des codes STBC différents qui vont permettre de définir les nouveaux états du treillis. La construction de ces codes est assez semblable à celle des M-TCM (Multidimensional-TCM). Les performances obtenues montrent une amélioration sensible par rapport aux codes STTC.

## 5-2 Critères de construction des codes STTC

On considère ici un système de radiocommunication mobile cellulaire dans lequel la station de base est équipée de  $n$  antennes et le portable est équipé de  $m$  antennes. Le schéma de transmission est conforme à celui de la figure 2. Les données sont codées par un codeur de canal puis passent par un convertisseur série/parallèle pour obtenir  $n$  trains de données. Chaque train après filtrage de mise en forme est modulé puis transmis par une antenne d'émission. Ainsi, à chaque créneau temporel  $t$ , la sortie du modulateur  $i$  correspond à un signal  $c_i^t$  transmis sur l'antenne  $i$  avec  $1 \leq i \leq n$ . Les  $n$  signaux sont donc transmis simultanément chacun ayant la même période  $T$ .

Le signal à l'antenne de réception est la superposition des  $n$  signaux transmis en tenant compte du bruit additif et du fading de Rayleigh sur la liaison considérée (entre les  $n$  antennes d'émission et l'antenne de réception  $j$ ,  $1 \leq j \leq m$ ). Dans les calculs qui suivent il est supposé que l'énergie moyenne de la constellation est égale à un (ceci revient à diviser tous les signaux transmis par un facteur d'échelle  $\sqrt{E_s}$ ).

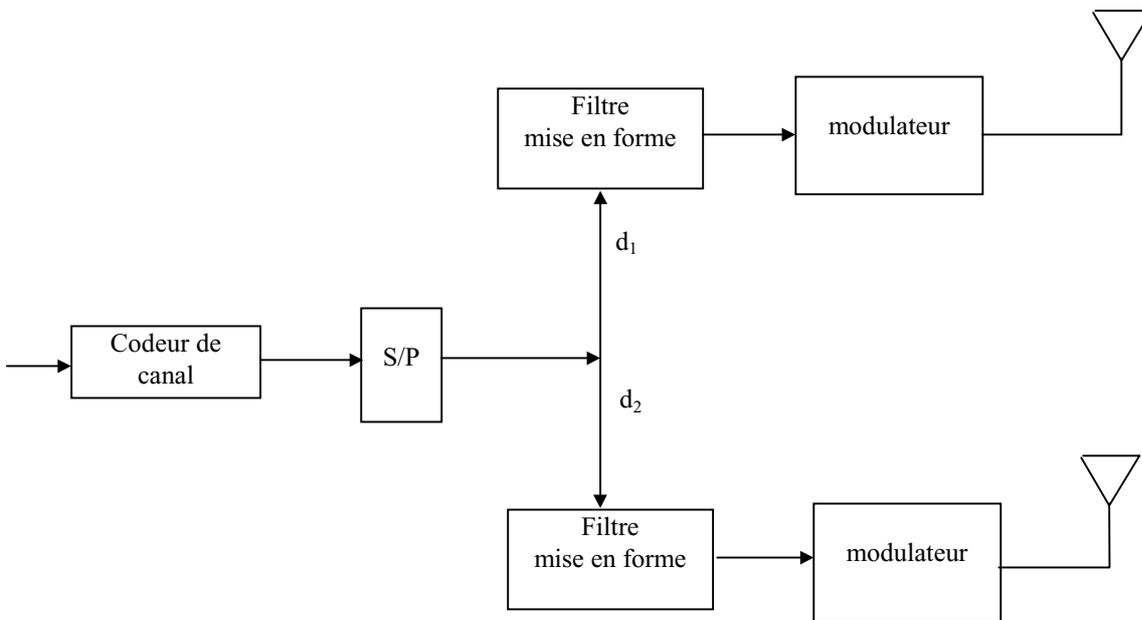


Fig. 2: schéma d'émission considéré

Au récepteur, le démodulateur utilise une statistique basée sur le critère du maximum de vraisemblance à postériori, cette statistique est construite à partir du signal reçu qui se met sous la forme :

$$d_t^j = \sum_{i=1}^n \alpha_{i,j} c_i^t \cdot \sqrt{E_s} + \eta_j^t \quad (1)$$

Le bruit additif  $\eta_j^t$  est un bruit Gaussien complexe de moyenne nulle et de variance  $N_0/2$  sur chaque partie réelle ou imaginaire. Les coefficients  $\alpha_{i,j}$  correspondent au gain du trajet entre

l'antenne de transmission  $i$  et l'antenne de réception  $j$ . On suppose ici que les coefficients  $\alpha_{i,j}$  restent constants à l'intérieur d'une trame mais peuvent varier d'une trame à l'autre (hypothèse de fading quasi-statique).

### 5-2-1 Cas où les coefficients $\alpha_{i,j}$ sont tous indépendants

Dans ce sous-paragraphe, nous supposons que les gains de trajet sont modélisés par des échantillons de bruit complexes Gaussiens de moyenne éventuellement non-nulle (cas d'un canal de Rice) et de variance 0.5 sur chaque partie réelle ou imaginaire. Ceci revient à dire que les signaux transmis depuis des antennes différentes sont confrontés à des "fadings" indépendants.

Utilisant cette hypothèse, nous pouvons alors concevoir des codes optimisés pour ce type de scénario. Cependant, avant d'établir les critères de construction, il est nécessaire de donner quelques rappels d'algèbre matricielle.

Si nous définissons  $\mathbf{x} = (x_1, x_2, \dots, x_k)$  et  $\mathbf{y} = (y_1, y_2, \dots, y_k)$  deux vecteurs de nombres complexes sur l'espace vectoriel  $C^k$ , le produit scalaire de ces deux vecteurs est défini par:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^k x_i \overline{y_i} \quad (2)$$

où la quantité  $\overline{y_i}$  désigne le complexe conjugué de  $y_i$ . Pour une matrice  $A$  quelconque, sa transformée Hermitienne (conjuguée transposée) sera notée  $A^*$ .

A partir des considérations algébriques donnée en annexe A, nous commençons par le calcul de la probabilité de l'événement d'erreur suivant (PEP: Pairwise Error Probability). Nous calculons la probabilité que le récepteur à maximum de vraisemblance décide que le signal reçu est égal à

$$\mathbf{e} = e_1^1 e_1^2 \dots e_1^n e_2^1 e_2^2 \dots e_2^n \dots e_l^1 e_l^2 \dots e_l^n$$

en supposant que le signal réellement transmis s'écrit

$$\mathbf{c} = c_1^1 c_1^2 \dots c_1^n c_2^1 c_2^2 \dots c_2^n \dots c_l^1 c_l^2 \dots c_l^n$$

En supposant l'estimation de canal parfaite (les coefficients  $\alpha_{i,j}$  sont parfaitement estimés au récepteur), on peut approximer le calcul de la probabilité de l'événement d'erreur sous la forme

$$P(\mathbf{c} \rightarrow \mathbf{e} | \alpha_{i,j}, i = 1, 2, \dots, n, j = 1, 2, \dots, m) \leq \exp(-d^2(\mathbf{c}, \mathbf{e}) \cdot E_s / 4 \cdot N_0) \quad (3)$$

où  $N_0/2$  désigne la variance du bruit additif sur chaque partie réelle ou imaginaire. La distance  $d^2(\mathbf{c}, \mathbf{e})$  s'écrit

$$d^2(\mathbf{c}, \mathbf{e}) = \sum_{j=1}^m \sum_{t=1}^l \left| \sum_{i=1}^n \alpha_{i,j} (c_t^i - e_t^i) \right|^2 \quad (4)$$

En appelant  $\Omega_j$  l'ensemble  $\Omega_j = (\alpha_{1,j}, \alpha_{2,j}, \dots, \alpha_{n,j})$ , on peut réécrire (4) sous la forme

$$d^2(\mathbf{c}, \mathbf{e}) = \sum_{j=1}^m \sum_{i=1}^n \sum_{i'=1}^n \alpha_{i,j} \overline{\alpha_{i',j}} \sum_{t=1}^l (c_t^i - e_t^i) \overline{(c_t^{i'} - e_t^{i'})} \quad (5)$$

En notation matricielle (5) peut encore s'écrire

$$d^2(\mathbf{c}, \mathbf{e}) = \sum_{j=1}^m \Omega_j \cdot A \cdot \Omega_j^* \quad (6)$$

avec  $A_{p,q} = \mathbf{x}_p \cdot \mathbf{x}_q$  et  $\mathbf{x}_p = (c_1^p - e_1^p, c_2^p - e_2^p, \dots, c_l^p - e_l^p)$  pour  $1 \leq p, q \leq n$ . Ainsi, la probabilité d'erreur de décider le signal  $\mathbf{e}$  à la place du signal  $\mathbf{c}$  s'écrit finalement :

$$P(\mathbf{c} \rightarrow \mathbf{e} | \alpha_{i,j}, i = 1, 2, \dots, n, j = 1, 2, \dots, m) \leq \exp(-\Omega_j \cdot A(\mathbf{c}, \mathbf{e}) \cdot \Omega_j^* \cdot E_s / 4 \cdot N_0) \quad (7)$$

avec  $A_{p,q} = \sum_{t=1}^l (c_t^p - e_t^p) \overline{(c_t^q - e_t^q)}$ . Puisque  $A(\mathbf{c}, \mathbf{e})$  est une matrice Hermitienne, il existe une matrice unitaire  $V$  et une matrice diagonale réelle  $D$  telle que  $V \cdot A(\mathbf{c}, \mathbf{e}) \cdot V^* = D$ . Les lignes de  $V$ ,  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  constituent une base orthonormée complète de  $C^n$  et correspondent aux vecteurs propres de  $A$  en tenant compte des ordres de multiplicité. Par construction, la matrice  $B(\mathbf{c}, \mathbf{e})$  définie ci-dessous est alors une racine carrée de  $A(\mathbf{c}, \mathbf{e})$ .

$$B(\mathbf{c}, \mathbf{e}) = \begin{pmatrix} e_1^1 - c_1^1 & e_2^1 - c_2^1 & \dots & \dots & e_l^1 - c_l^1 \\ e_1^2 - c_1^2 & e_2^2 - c_2^2 & \dots & \dots & e_l^2 - c_l^2 \\ e_1^3 - c_1^3 & e_2^3 - c_2^3 & \ddots & \vdots & e_l^3 - c_l^3 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ e_1^n - c_1^n & e_2^n - c_2^n & \dots & \dots & e_l^n - c_l^n \end{pmatrix} \quad (8)$$

Ainsi les valeurs propres de  $A(\mathbf{c}, \mathbf{e})$  sont toutes positives ou nulles. Nous pouvons alors exprimer  $d^2(\mathbf{c}, \mathbf{e})$  en fonction des valeurs propres de la matrice  $A(\mathbf{c}, \mathbf{e})$ . En posant  $(\beta_{1,j}, \dots, \beta_{n,j}) = \Omega_j V^*$ , il vient

$$\Omega_j \cdot A(\mathbf{c}, \mathbf{e}) \cdot \Omega_j^* = \sum_{i=1}^n \lambda_i |\beta_{i,j}|^2 \quad (9)$$

Maintenant, en se rappelant que les coefficients  $\alpha_{i,j}$  sont des échantillons d'un bruit complexe Gaussien de moyenne  $E\alpha_{i,j}$ , nous utilisons le vecteur :

$$\mathbf{K}^j = (E\alpha_{1,j}, E\alpha_{2,j}, \dots, E\alpha_{n,j})$$

Comme la matrice  $V$  est unitaire,  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  est une base orthonormale de  $C^n$  et les coefficients  $\beta_{i,j}$  sont des variables aléatoires complexes Gaussiennes de variance 0.5 sur chaque partie réelle ou imaginaire et de moyenne  $\mathbf{K}^j \cdot \mathbf{v}_i$ . Avec  $K_{i,j} = |E \cdot \beta_{i,j}|^2 = |\mathbf{K}^j \cdot \mathbf{v}_i|^2$ , les variables  $|\beta_{i,j}|$  sont toutes indépendantes et ont une distribution de probabilité égale à :

$$p(|\beta_{i,j}|) = 2 \cdot |\beta_{i,j}| \cdot \exp(-|\beta_{i,j}|^2 - K_{i,j}) \cdot I_0(2 \cdot |\beta_{i,j}| \cdot \sqrt{K_{i,j}})$$

où  $I_0$  désigne la fonction de Bessel modifiée d'ordre zéro. Le but étant d'obtenir une probabilité PEP moyennée sur l'ensemble des erreurs d'événements, on moyenne la probabilité calculée en (7) en utilisant la distribution des variables  $\beta_{i,j}$ . On obtient alors

$$P(\mathbf{c} \rightarrow \mathbf{e}) \leq \prod_{j=1}^m \left( \prod_{i=1}^n \frac{1}{1 + \frac{E_s}{4 \cdot N_0} \cdot \lambda_i} \cdot \exp \left( - \frac{K_{i,j} \frac{E_s}{4 \cdot N_0} \cdot \lambda_i}{1 + \frac{E_s}{4 \cdot N_0} \cdot \lambda_i} \right) \right) \quad (10)$$

Si on considère alors le cas du fading de Rayleigh, on a  $E(\alpha_{i,j}) = 0$  et à fortiori  $K_{i,j} = 0$  pour tout  $i$  et  $j$ . Ainsi, on peut réécrire l'inégalité (10) sous la forme

$$P(\mathbf{c} \rightarrow \mathbf{e}) \leq \left( \frac{1}{\prod_{i=1}^n (1 + \lambda_i E_s / 4 \cdot N_0)} \right)^m \quad (11)$$

En notant  $r$  le rang de la matrice  $A$ , le noyau de  $A$  a pour dimension  $n-r$  et a exactement  $n-r$  valeurs propres égales à zéro. Si les valeurs propres non nulles sont  $\lambda_1, \lambda_2, \dots, \lambda_r$ , on déduit de l'inégalité (11) que

$$P(\mathbf{c} \rightarrow \mathbf{e}) \leq \left( \prod_{i=1}^r \lambda_i \right)^{-m} (E_s / 4 \cdot N_0)^{-rm} \quad (12)$$

On obtient alors un avantage en diversité de  $m \cdot r$  et un avantage en gain de codage égal à  $(\lambda_1 \lambda_2 \dots \lambda_r)^{1/r}$  est réalisé.  $\lambda_1 \lambda_2 \dots \lambda_r$  est la valeur absolue de la somme des déterminants de tous les principaux cofacteurs de  $A$  de taille  $r \times r$  et il est aisé de voir que les rangs de  $A(\mathbf{c}, \mathbf{e})$  et de  $B(\mathbf{c}, \mathbf{e})$  sont égaux.

*Remarque:* on peut noter que l'avantage en diversité correspond à la puissance du SNR dans le dénominateur de l'expression de la probabilité d'erreur d'événements (PEP) calculée

précédemment. Le gain de codage est une mesure approximative du gain obtenu par rapport à un système non codé opérant avec le même avantage en diversité.

Ces premières analyses nous permettent d'arriver aux critères de construction suivants pour les codes STTC.

Critère de construction pour les canaux de Rayleigh à "fadings" indépendants

- *Le critère du rang*: Afin d'obtenir une diversité maximale égale à  $m.n$ , la matrice  $B(\mathbf{c}, \mathbf{e})$  doit être de rang complet (son noyau est égal au vecteur nul) pour chaque paire de mots de code  $\mathbf{c}$  et  $\mathbf{e}$ . Si  $B(\mathbf{c}, \mathbf{e})$  a un rang minimum égal à  $r$  sur toutes les paires possibles de mots de code alors un gain en diversité de  $r.m$  est réalisé. Ce critère est aussi évoqué dans [4].

- *Le critère du déterminant*: Supposons que l'on veuille obtenir un gain en diversité de  $r.m$ . Le minimum de la somme des déterminants de tous les  $r \times r$  cofacteurs principaux de  $A(\mathbf{c}, \mathbf{e}) = B(\mathbf{c}, \mathbf{e}).B^*(\mathbf{c}, \mathbf{e})$  pris sur toutes les paires de mots de code distincts  $\mathbf{e}$  et  $\mathbf{c}$  correspond au gain de codage, où  $r$  désigne le rang de  $A(\mathbf{c}, \mathbf{e})$ . Le but de la conception des codes STTC est de rendre cette somme aussi grande que possible. Si on veut obtenir une diversité égale à  $n.m$ , c'est-à-dire une diversité maximale, alors le minimum du déterminant de  $A(\mathbf{c}, \mathbf{e})$  pris sur toutes les paires de mots de code distincts  $\mathbf{e}$  et  $\mathbf{c}$  doit être maximisé.

Il est intéressant de regarder le comportement du membre de droite de l'égalité (10) pour les rapports signal à bruit élevés. Pour des valeurs élevées de  $E_s / N_0$  on peut réécrire (10) sous la forme

$$P(\mathbf{c} \rightarrow \mathbf{e}) \leq \left( \frac{E_s}{4.N_0} \right)^{-rm} \cdot \left( \prod_{i=1}^r \lambda_i \right)^{-m} \cdot \left[ \prod_{j=1}^m \prod_{i=1}^r \exp(-K_{i,j}) \right] \quad (13)$$

Ainsi, on obtient un gain en diversité égal à  $r.m$  et un gain de codage égal à

$$(\lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_r)^{1/r} \left[ \prod_{j=1}^m \prod_{i=1}^r \exp(-K_{i,j}) \right]^{1/r.m}$$

est réalisé. On en déduit les règles de construction suivantes pour les codes STTC dans le cas des canaux de Rice et pour des valeurs élevées de  $E_s/N_0$ .

- *Le critère du rang*: Ce critère est le même que celui déjà énoncé en page 6 pour le canal de Rayleigh à "fadings" indépendants.

- *Le critère du gain de codage*: En appelant  $\Lambda(\mathbf{c}, \mathbf{e})$  la somme de tous les déterminants des principaux cofacteurs  $r \times r$  de  $A(\mathbf{c}, \mathbf{e})$  où  $r$  désigne le rang de  $A(\mathbf{c}, \mathbf{e})$ . Le minimum des produits

$$\Lambda(\mathbf{c}, \mathbf{e})^{1/r} \left[ \prod_{j=1}^m \prod_{i=1}^r \exp(-K_{i,j}) \right]^{1/r.m}$$

pris sur toutes les paires de mots de code distincts  $\mathbf{c}$  et  $\mathbf{e}$  doit être maximisé.

### 5-2-2 a- Cas où les coefficients $\alpha_{i,j}$ sont corrélés

Dans ce paragraphe, nous supposons que les coefficients  $\alpha_{i,j}$  sont des échantillons corrélés de variables aléatoires Gaussiennes de moyenne nulle et de variance 0.5 sur chaque partie réelle ou imaginaire. Nous traitons le cas du canal de Rayleigh mais l'extension des calculs au cas du canal de Rice est immédiate. Nous considérons d'abord la matrice  $mn \times mn$

$$Y(\mathbf{c}, \mathbf{e}) = \begin{pmatrix} A(\mathbf{c}, \mathbf{e}) & 0 & \dots & \dots & 0 & 0 \\ 0 & A(\mathbf{c}, \mathbf{e}) & \dots & \dots & 0 & 0 \\ 0 & 0 & A(\mathbf{c}, \mathbf{e}) & \ddots & \vdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & A(\mathbf{c}, \mathbf{e}) \end{pmatrix}$$

0 représente ici la matrice nulle de taille  $n \times n$ . Avec  $\Omega = (\Omega_1, \dots, \Omega_m)$  on peut réécrire l'équation (7) sous la forme

$$P(\mathbf{c} \rightarrow \mathbf{e} | \alpha_{i,j}, i = 1, 2, \dots, n, j = 1, 2, \dots, m) \leq \exp(-\Omega_j \cdot Y(\mathbf{c}, \mathbf{e}) \cdot \Omega_j^* \cdot E_s / 4 \cdot N_0) \quad (14)$$

On note alors  $\Theta = E \Omega \Omega^*$  la matrice de corrélation de  $\Omega$ . On supposera dans la suite des calculs que  $\Theta$  est de rang complet (son noyau est le vecteur nul). Comme la matrice  $\Theta$  est une matrice définie positive, on sait qu'elle admet une racine carrée  $C$  qui est une matrice triangulaire inférieure de taille  $nm \times nm$ . En utilisant la transformation  $W = \Omega(C^*)^{-1}$ , il est aisé de voir que les composantes de  $W$  deviennent des variables Gaussiennes non-corrélées de variance 0.5 sur chaque partie réelle ou imaginaire. La moyenne des composantes de  $W$  peut être facilement calculée à partir de la moyenne des coefficients  $\alpha_{i,j}$  et de la matrice  $C$ . En particulier, cette moyenne est nulle si la moyenne des  $\alpha_{i,j}$  est nulle. En utilisant (14), on arrive finalement à :

$$P(\mathbf{c} \rightarrow \mathbf{e} | \alpha_{i,j}, i = 1, 2, \dots, n, j = 1, 2, \dots, m) \leq \exp(-W C^* \cdot Y(\mathbf{c}, \mathbf{e}) \cdot C W^* \cdot E_s / 4 \cdot N_0) \quad (15)$$

On peut alors suivre la même démarche d'optimisation que dans le cas des "fadings" indépendants en remplaçant la matrice  $A(\mathbf{c}, \mathbf{e})$  par la matrice  $C^* \cdot Y(\mathbf{c}, \mathbf{e}) \cdot C$ . On en déduit que le rang de  $C^* \cdot Y(\mathbf{c}, \mathbf{e}) \cdot C$  doit être maximisé. Comme on a déjà supposé que  $C$  était de rang plein, ceci revient à maximiser

$$\text{rang}[Y(\mathbf{c}, \mathbf{e})] = m \cdot \text{rang}[A(\mathbf{c}, \mathbf{e})]$$

On voit que le critère du rang énoncé pour le cas des "fadings" indépendants est toujours valable.

Comme les coefficients  $\alpha_{i,j}$  ont une moyenne nulle, il en va de même pour les composantes de  $W$ . On peut alors, en suivant la même démarche que dans le paragraphe précédent (cas des “fadings” indépendants), en déduire que le déterminant de  $C^* . Y(c, e) . C$  doit être maximisé. Or, ce déterminant est égal à :

$$\det(\Theta) . \det(Y(c, e)) = \det(\Theta) . [\det(A(c, e))]^m \quad (16)$$

Là encore, on s’aperçoit que le critère du déterminant énoncé pour le cas des “fadings” indépendants est toujours valable. De plus, en comparant cette expression du déterminant avec celle obtenue dans le cas a- (“fadings” indépendants) on peut quantifier la perte en gain de codage due à la corrélation des coefficients  $\alpha_{i,j}$ . Cette perte au vue de (16) est approximativement égale à  $(10 / n.m) . \log_{10}(\det(\Theta))$  décibels (db).

De plus, en utilisant le même type d’argumentation, on se rend compte que le critère du rang est toujours valable pour le cas des canaux de Rice et qu’un code optimisé pour le canal de Rayleigh le sera également pour un canal de Rice même si les coefficients  $\alpha_{i,j}$  sont corrélés. Pour obtenir le critère du gain de codage, on doit calculer la moyenne des composantes de  $W$  et appliquer le critère du gain de codage obtenu dans le cas des “fadings” indépendants.

### 5-2-3 Cas des évanouissements “fadings” rapides

Lorsque les variations du canal de propagation sont rapides, nous utilisons le modèle de canal suivant :

$$d_i^j = \sum_{i=1}^n \alpha_{i,j}(t) . c_i^i . \sqrt{E_s} + \eta_i^j \quad (17)$$

Les coefficients  $\alpha_{i,j}(t)$  sont modélisés pour  $t = 1, 2, \dots, l$ ,  $i = 1, 2, \dots, n$  et  $j = 1, 2, \dots, m$  par des échantillons complexes Gaussiens indépendants de moyenne nulle et de et de variance 0.5 sur chaque partie réelle ou imaginaire. Ces hypothèses correspondent au cas du canal de Rayleigh mais, là encore, l’extension au cas du canal de Rice est immédiate. De plus, les  $\eta_i^j$  sont des échantillons de bruit Gaussiens complexes de variance  $N_0/2$  sur chaque partie réelle ou imaginaire. Comme dans les paragraphes précédents nous supposons que les coefficients  $\alpha_{i,j}(t)$  sont parfaitement estimés au récepteur. On calcule la probabilité de décider la suite de symboles décodés

$$e = e_1^1 e_1^2 \dots e_1^n e_2^1 e_2^2 \dots e_2^n \dots e_l^1 e_l^2 \dots e_l^n$$

alors que la suite de symboles réellement transmis s’écrit

$$c = c_1^1 c_1^2 \dots c_1^n c_2^1 c_2^2 \dots c_2^n \dots c_l^1 c_l^2 \dots c_l^n$$

Ce calcul de probabilité conduit à l'expression :

$$P(\mathbf{c} \rightarrow \mathbf{e} | \alpha_{i,j}(t), i, j, t) \leq \exp(-d^2(\mathbf{c}, \mathbf{e}) \cdot E_s / 4 \cdot N_0) \quad (18)$$

avec

$$d^2(\mathbf{c}, \mathbf{e}) = \sum_{j=1}^m \sum_{t=1}^l \left| \sum_{i=1}^n \alpha_{i,j}(t) \cdot (c_t^i - e_t^i) \right|^2$$

En notant :  $\Omega_j(t) = (\alpha_{1,j}(t), \alpha_{2,j}(t), \dots, \alpha_{n,j}(t))$  et  $C(t)$  la matrice  $n \times n$  dont la  $p^{\text{ème}}$  ligne et la  $q^{\text{ème}}$  colonne admet pour élément  $(c_t^p - e_t^p) \cdot \overline{(c_t^q - e_t^q)}$ . On peut alors réécrire la distance  $d^2(\mathbf{c}, \mathbf{e})$  sous la forme :

$$d^2(\mathbf{c}, \mathbf{e}) = \sum_{j=1}^m \sum_{t=1}^l \Omega_j(t) C(t) \cdot \Omega_j^*(t) \quad (19)$$

La matrice  $C(t)$  est hermitienne, il existe donc une matrice unitaire  $V(t)$  et une matrice diagonale  $D(t)$  telles que  $C(t) = V(t) \cdot D(t) \cdot V^*(t)$ . Les éléments de la diagonale de  $D(t)$ , notés ici  $D_{ii}(t)$ ,  $1 \leq i \leq n$ , sont les valeurs propres de  $C(t)$  avec leurs ordres de multiplicité. Comme  $C(t)$  est hermitienne, ces valeurs propres sont réelles. En posant

$$(\beta_{1,j}(t), \dots, \beta_{n,j}(t)) = \Omega_j(t) \cdot V(t)$$

alors les variables  $\beta_{i,j}(t)$  pour  $i = 1, 2, \dots, n, j = 1, 2, \dots, m$  et  $t = 1, 2, \dots, l$  sont des variables Gaussiennes complexes de moyenne nulle et de variance 0.5 sur chaque partie réelle ou imaginaire. On a alors :

$$\Omega_j(t) C(t) \cdot \Omega_j^*(t) = \sum_{i=1}^n |\beta_{i,j}(t)|^2 \cdot D_{i,i}(t)$$

En réitérant les mêmes étapes que dans le cas a- on arrive alors, après moyennage sur la distribution de Rayleigh, à :

$$P(\mathbf{c} \rightarrow \mathbf{e}) \leq \prod_{i,t} \left( 1 + D_{i,i}(t) \cdot \frac{E_s}{4 \cdot N_0} \right)^{-m} \quad (20)$$

Pour aller plus loin, il nous faut maintenant examiner la structure de la matrice  $C(t)$ . Les colonnes de  $C(t)$  sont toutes multiples de

$$\mathbf{c}_t - \mathbf{e}_t = (c_t^1 - e_t^1, c_t^2 - e_t^2, \dots, c_t^n - e_t^n)$$

Ainsi,  $C(t)$  a un rang égal à un si  $c_t^1 c_t^2 \dots c_t^n \neq e_t^1 e_t^2 \dots e_t^n$  et un rang nul dans le cas contraire. Il s'ensuit que  $(n-1)$  éléments dans la liste  $D_{11}(t), D_{22}(t), \dots, D_{nn}(t)$  sont égaux à zéro et que le seul élément non nul dans cette liste est  $|c_t - e_t|^2$ . En utilisant (20), nous pouvons maintenant conclure que

$$P(\mathbf{c} \rightarrow \mathbf{e} | \alpha_{i,j}(t), i, j, t) \leq \prod_{t=1}^l \left( 1 + |c_t - e_t|^2 \frac{E_s}{4N_0} \right)^{-m} \quad (21)$$

Si on note  $\nu(\mathbf{c}, \mathbf{e})$  l'ensemble des indices  $t$ ,  $1 \leq t \leq l$  tels que  $|c_t - e_t| \neq 0$  et  $|\nu(\mathbf{c}, \mathbf{e})|$  le nombre d'éléments dans l'ensemble  $\nu(\mathbf{c}, \mathbf{e})$ . L'expression (21) entraîne alors

$$P(\mathbf{c} \rightarrow \mathbf{e}) \leq \prod_{t \in \nu(\mathbf{c}, \mathbf{e})} \left( |c_t - e_t|^2 \cdot \frac{E_s}{N_0} \right)^{-m} \quad (22)$$

Il s'ensuit qu'une diversité de  $m|\nu(\mathbf{c}, \mathbf{e})|$  est réalisée. Les critères de construction du code STTC se déduisent de l'examen du coefficient de  $(E_s / 4N_0)^{-m\nu(\mathbf{c}, \mathbf{e})}$

- *Le critère de la distance*: Pour obtenir un ordre de diversité égal à  $\mu m$  dans un environnement à évanouissements rapides, les quantités  $c_t^1 c_t^2 \dots c_t^n$  et  $e_t^1 e_t^2 \dots e_t^n$  doivent être différentes au moins pour  $\mu$  valeurs de  $t$ .

- *Le critère du produit* : En notant  $\nu(\mathbf{c}, \mathbf{e})$  l'ensemble des indices  $t$ ,  $1 \leq t \leq l$  tels que  $c_t^1 c_t^2 \dots c_t^n \neq e_t^1 e_t^2 \dots e_t^n$  et en notant

$$|\mathbf{c}_t - \mathbf{e}_t|^2 = \sum_{i=1}^n |c_t^i - e_t^i|^2$$

on trouve alors que le maximum de gain en codage est obtenu lorsque le minimum des produits

$$\prod_{t \in \nu(\mathbf{c}, \mathbf{e})} |\mathbf{c}_t - \mathbf{e}_t|^2$$

pris sur les différents mots de code  $\mathbf{e}$  et  $\mathbf{c}$  doit être maximisé.

### **5-3 Construction des codes**

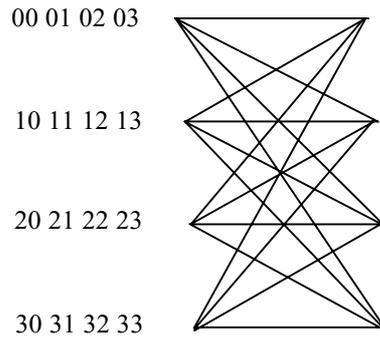
#### ***a- Limites fondamentales basées sur le critère de la capacité de coupure ("Outage Capacity") du canal***

On considère un système de communication employant  $n$  antennes de transmission et une (ou deux) antenne de réception avec un fading non sélectif en fréquence et quasi-statique. Il est logique de penser qu'il existe une limite au-delà laquelle rajouter des antennes d'émission

n'apporte plus qu'un gain marginal en termes de capacité Outage. Foschini et Gans [5] ont prouvé que la capacité Outage d'un tel système est une variable aléatoire de la forme  $\log_2(1 + (\chi_{2n}^2)SNR)$  où  $\chi_{2n}^2$  est une variable aléatoire obtenue en sommant les carrés de  $2n$  variables gaussiennes indépendantes de moyenne zéro et de variance un. Ceci entraîne d'après la loi des grands nombres que  $\chi_{2n}^2/2n \rightarrow 1$  en loi de probabilité. De façon concrète, dans le cas  $n = 4$ , on a  $\chi_{2n}^2/2n \cong 1$  et on retrouve le calcul de capacité classique pour un canal Gaussien complexe, à savoir  $\log_2(1 + SNR)$  par dimension complexe. Ainsi, avec une antenne de réception, on montre qu'au delà de quatre antennes d'émission le gain en capacité obtenu est marginal. De la même façon, on peut démontrer qu'avec deux antennes de réception on obtient la quasi-totalité de la capacité du canal avec  $n = 6$  antennes d'émission. Ces observations corroborent complètement les résultats donnés par Telatar dans [6]. Dans ce chapitre nous considérons des systèmes de transmission avec une ou deux antennes de réception et ainsi le nombre d'antennes d'émission est toujours inférieur à six. Si un nombre supérieur d'antennes d'émission et de réception est considéré alors les méthodes présentées dans [7] où l'on combine le traitement d'antenne (changements du diagramme de rayonnement par la commande des formateurs de faisceaux de l'antenne) et le codage spatio-temporel peuvent s'appliquer. De plus, nous considérons ici des transmissions hauts débits par paquets de petite taille, il est alors naturel d'introduire le codage spatio-temporel à l'intérieur de chaque trame mais sans introduire de corrélation entre les trames. La capacité de coupure ("Outage Capacity") représente alors un excellent paramètre de mesure des performances. Ceci explique pourquoi les courbes de résultats seront présentées en termes de FER: Frame Error Rate.

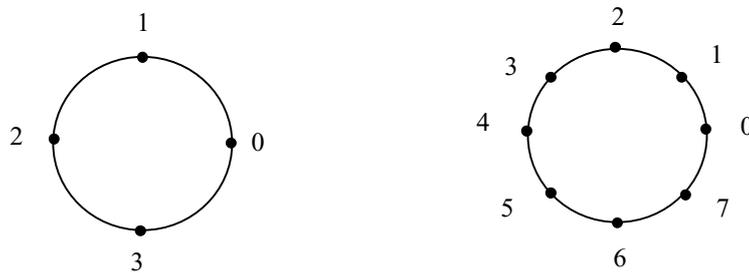
### ***b- Construction des codes pour le cas des "fading" non sélectif en fréquence et quasi-statiques***

Nous procédons à la construction des codes en utilisant les critères présentés dans la partie 2. Les treillis présentés ci-dessous sont simples à interpréter: les symboles sont d'abord groupés  $n$  par  $n$  (pour  $n$  antennes de transmission) puis, dépendant de l'état actuel du codeur, une branche de transition est sélectionnée dans le treillis conduisant à un nouvel état. Par exemple, dans le cas du codeur ci-dessous (figure 3), utilisant deux antennes d'émission, le symbole transmis sur l'antenne 1 est égal à zéro dans le premier état. On passe alors de l'état un à l'état un si le second symbole transmis sur l'antenne 2 est un zéro. On passe de l'état un à l'état deux si le second symbole transmis sur l'antenne 2 est un un. On passe de l'état un à l'état trois si le second symbole transmis sur l'antenne 2 est un deux. On passe de l'état un à l'état quatre si le second symbole transmis sur l'antenne 2 est un trois. Le même raisonnement se répète avec les états deux, trois et quatre.



**Fig. 3: STTC avec deux antennes d'émission, MDP-4, quatre états, 2 bits/s/Hz**

Les constellations utilisées en MDP-4 et MDP-8 sont présentées sur la figure 4 ci-dessous.



**Fig.4: constellations MDP-4 et MDP-8**

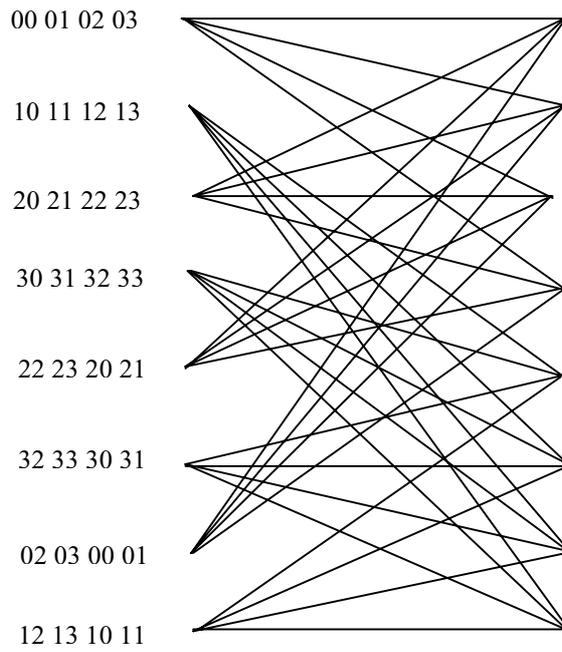
Pour la MDP-4, outre, le code présenté en figure 3, nous avons également testé les codes présentés en figure 5 (MDP-4, 2 antennes d'émission, huit et seize états).

D'autres exemples de codes sont présentés dans [3] page 752. En ce qui concerne la MDP-8, nous avons testé le code de la figure 6, ce code présente une efficacité spectrale de 3 bits/s/Hz

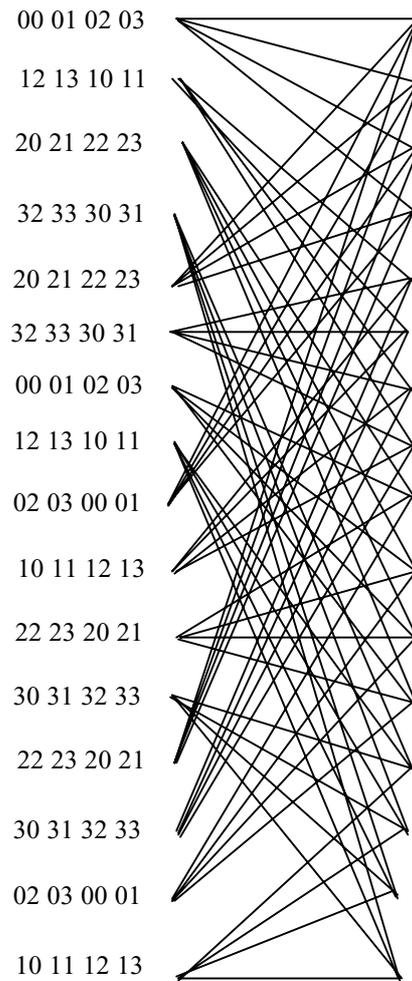
En ce qui concerne les algorithmes de décodage des codes STTC, l'algorithme de Viterbi ou l'algorithme de décodage par shères récemment introduit par Boutros [8] sont utilisables. Dans le cas de l'algorithme de Viterbi, la métrique de branche à calculer lors d'une transition notée  $q_t^1 q_t^2 \dots q_t^n$  est donnée par :

$$\sum_{j=1}^m \left| r_t^j - \sum_{i=1}^n \alpha_{i,j} \cdot q_t^i \right|^2$$

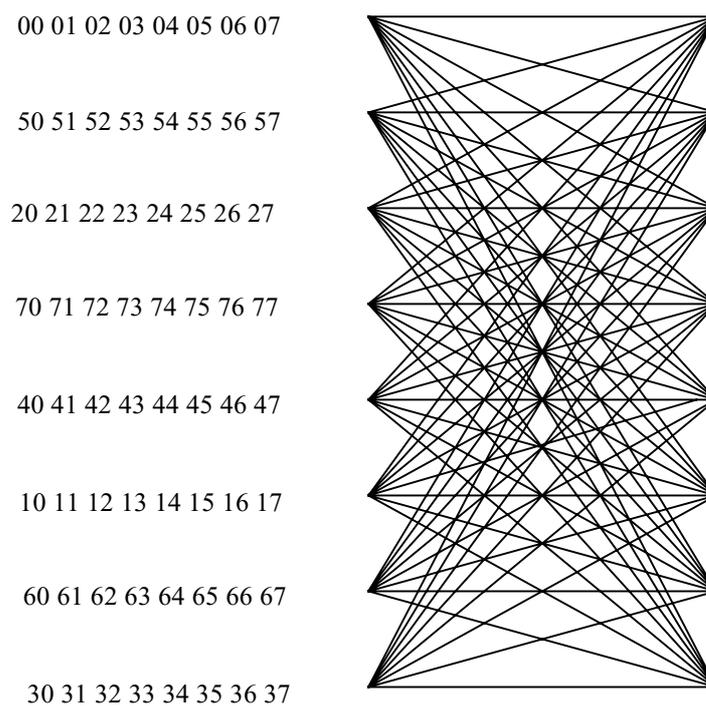
L'algorithme de Viterbi sert alors à sélectionner le chemin ayant la métrique cumulée la plus faible. Les performances de ces codes sont illustrées sur les figures 7-10 dans le cas de deux antennes d'émission et de un ou deux antennes de réception. Les paquets transmis ont une taille de 260 symboles (130 sur chaque antenne). Les codes utilisés sont ceux présentés dans [3] page 752. Les coefficients  $\alpha_{i,j}$  sont simulés par des échantillons de bruit complexes Gaussiens de moyenne éventuellement nulle et de variance 0.5 sur chaque partie réelle ou imaginaire.



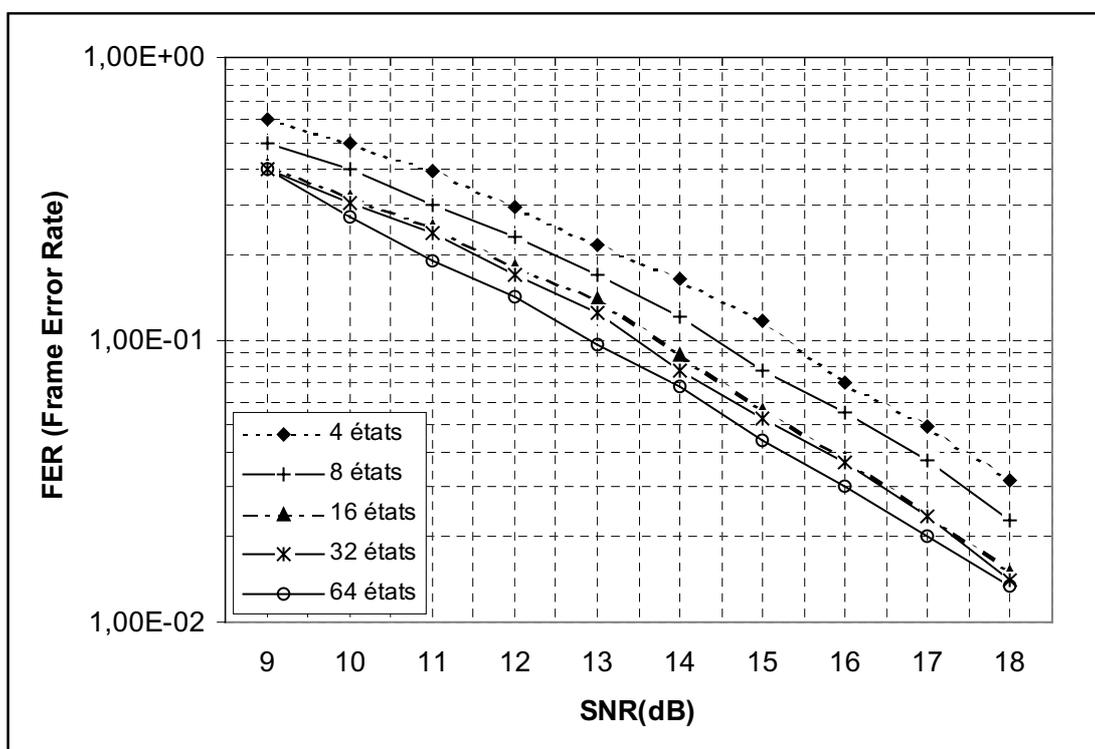
**Fig. 5-a : STTC avec deux antennes d'émission, MDP-4, huit états, 2 bits/s/Hz**



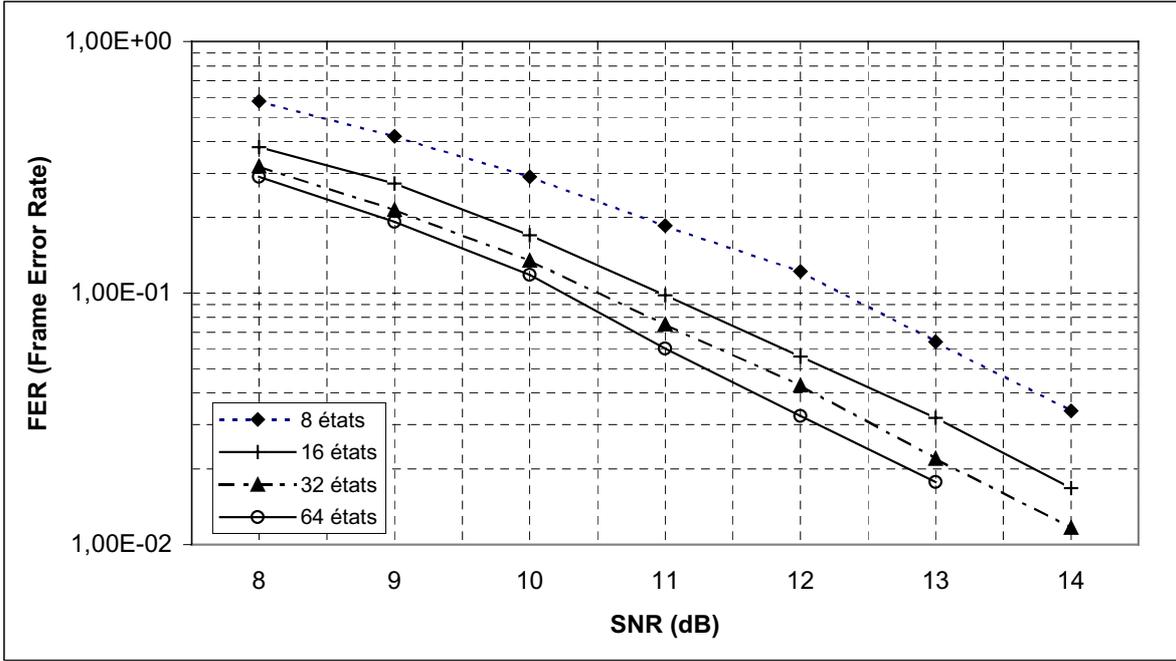
**Fig. 5-b: STTC avec deux antennes d'émission, MDP-4, 16 états, 2 bits/s/Hz**



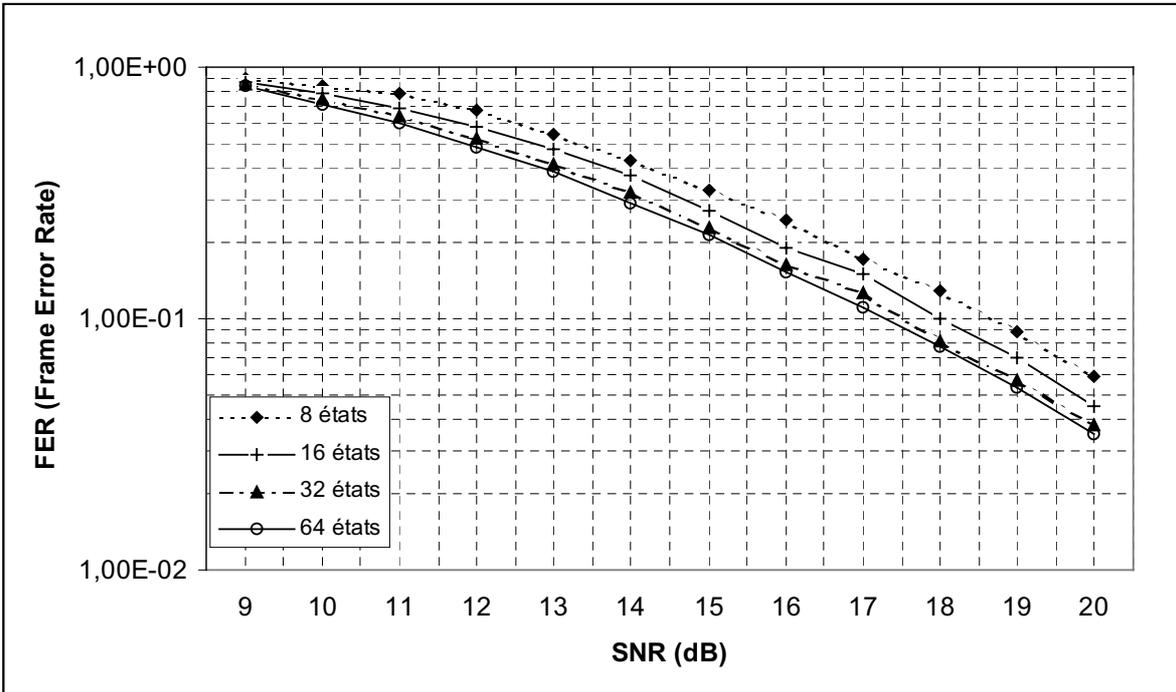
**Fig. 6: STTC avec deux antennes d'émission, MDP-8, 8 états, 3 bits/s/Hz**



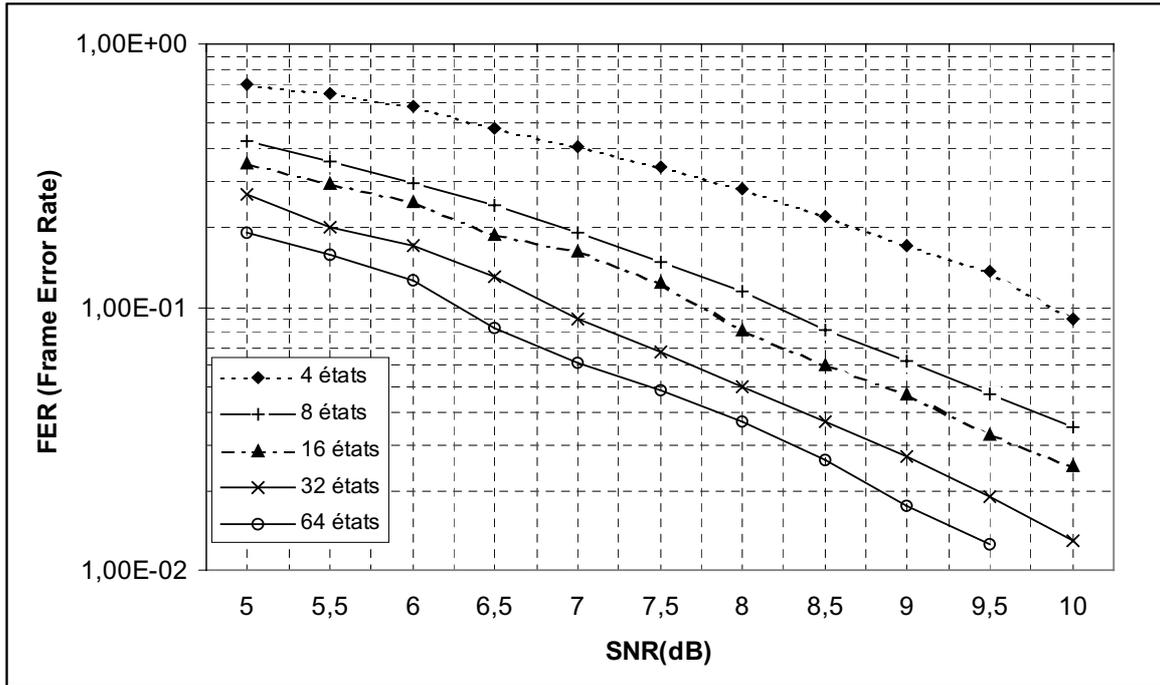
**Fig.7: Performances des codes STTC en MDP-4 avec une antenne de réception et deux antennes d'émission (2 bits/s/Hz)**



**Fig.8: Performances des codes STTC en MDP-4 avec deux antennes de réception et deux antennes d'émission (2 bits/s/Hz)**

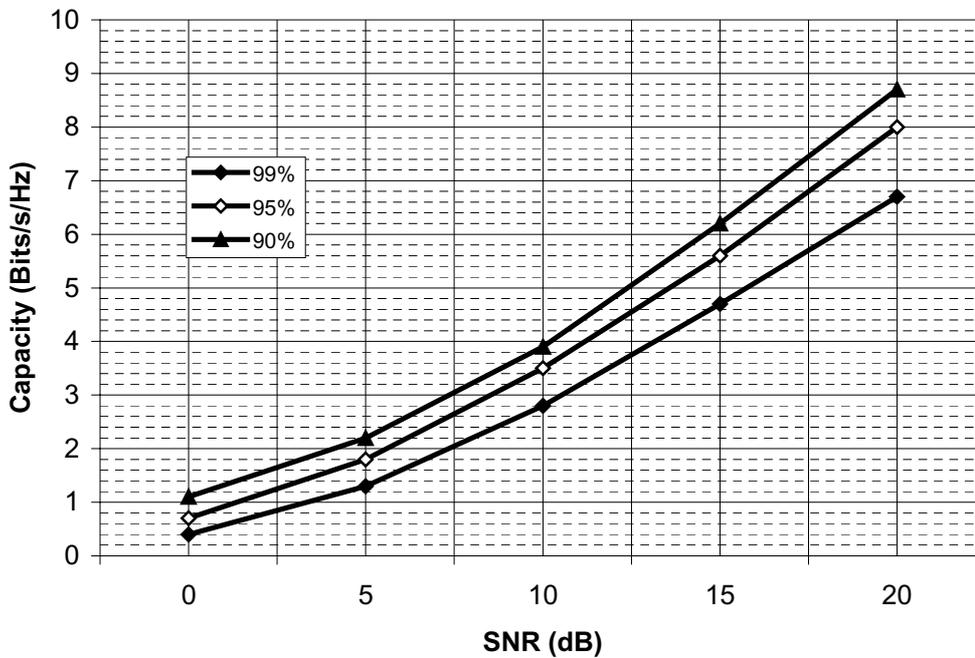


**Fig.9: Performances des codes STTC en MDP-8 avec une antenne de réception et deux antennes d'émission (3 bits/s/Hz)**

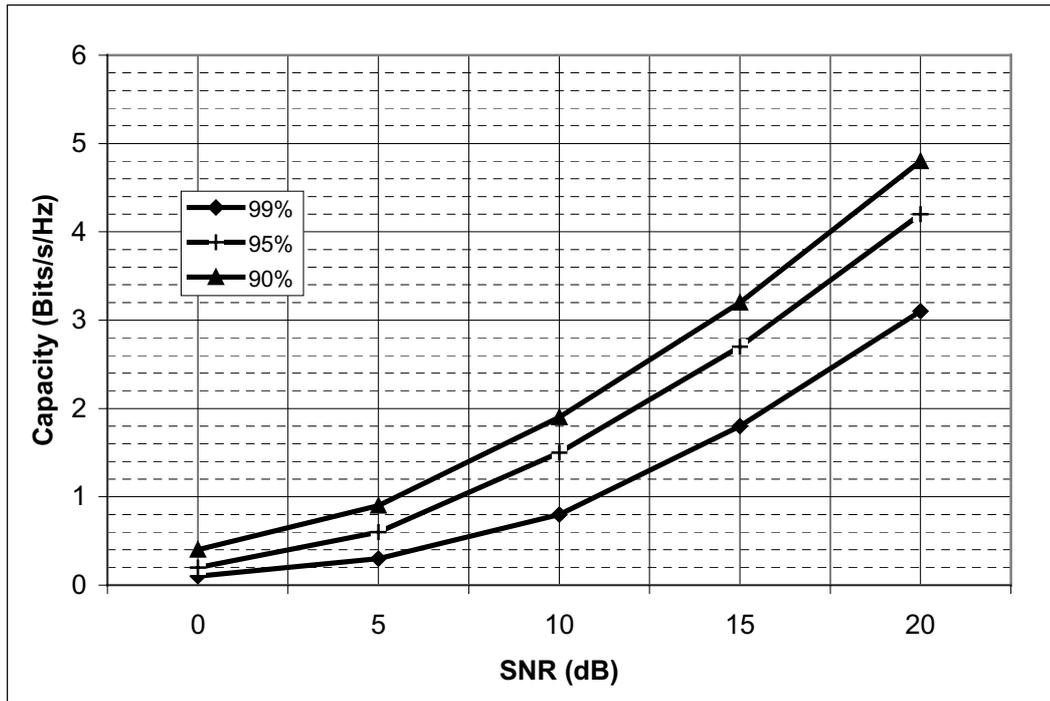


**Fig.10: Performances des codes STTC en MDP-8 avec deux antennes de réception et deux antennes d'émission (3 bits/s/Hz)**

A titre de comparaison nous donnons également les courbes de capacité Outage extraites de [5] (cf figures 11 et 12).



**Figure 11: Capacité Outage pour deux antennes d'émission et deux antennes de réception**



**Figure 12: Capacité Outage pour deux antennes d'émission et une antenne de réception**

On peut alors observer que pour un Frame Error Rate égal à 0.1, les codes réalisés et simulés sont à 2.5 dB de la capacité Outage. Ceci montre bien que les codes STTC permettent d'exploiter une grande partie de la capacité des canaux MIMO. On peut également observer que la performance des codes (le gain de codage) augmente avec le nombre d'antennes de réception. C'est ainsi que sur la figure 9 (cas d'une antenne de réception), on voit que l'on améliore peu les performances en passant de 4 à 64 états (le gain est de l'ordre de 1 dB à FER =  $10^{-1}$ ). Par contre, sur la figure 10, le gain pour FER =  $10^{-1}$  devient supérieur à 2 dB. On peut également constater que le gain de codage par rapport au code à quatre états n'est pas aussi important que ce qui avait été prévu par le critère du déterminant. Ceci n'est pas très étonnant car le critère du déterminant n'est qu'un critère approché qui ne tient pas compte, par exemple, de la diversité. Cependant, malgré les bonnes performances réalisées par les codes STTC, on peut se poser la question de leur optimalité: en d'autres termes, existent-ils des schémas de codage permettant de transmettre à des débits plus élevés en MDP-4 ou MDP-8 ou, pour les débits atteints par les codes des figures 7-10, existent-ils des schémas de codage plus simples? Le prochain paragraphe tente d'apporter une réponse à ces questions en énonçant certaines règles de conception qui permettent d'obtenir des débits élevés avec une bonne diversité en tenant compte de la taille de la constellation et du nombre d'états du treillis.

## 5-4 Règles générales de conception des codes STTC

Le but ici est de proposer pour un taux de transmission donné (nombre de bits par seconde et par Hz), une taille de constellation donnée et un nombre d'états du treillis fixé, un code STTC qui maximise le gain de codage donné par le critère du déterminant.

Ainsi, avec un système à  $n$  antennes d'émission et  $m$  antennes de réception on a vu que l'on pouvait espérer une diversité maximale égale à  $m.n$ . On cherche alors à réaliser le taux de transmission le plus élevé possible avec un avantage en diversité égal à  $r.m$ .

La construction du code utilise les théorèmes suivants:

**Théorème 1**: Si on considère un système de radiocommunication mobile avec  $n$  antennes de transmission et  $m$  antennes de réception avec un canal de Rice sur chaque trajet équivalent. On appelle  $r.m$  l'avantage en diversité du système. En supposant que l'on traite une constellation  $Q$  à  $2^b$  éléments, le taux de transmission satisfait à l'inégalité suivante

$$R \leq \frac{\log[A_{2^{bl}}(n, r)]}{l} \quad (23)$$

avec  $R$  exprimé en bits par seconde et par Hertz.  $A_{2^{bl}}(n, r)$  est la taille maximale d'un code de longueur  $n$  et de distance de Hamming minimale  $r$  défini sur un alphabet de taille  $2^{bl}$ .

**Preuve**: En appelant  $l$  la longueur de la trame, on considère le superalphabet  $Q^l = Q \times Q \dots \times Q$  défini par le  $l$  produit cartésien de  $Q$ . L'application (ou la correspondance)  $f: Q^{ln} \rightarrow [Q^l]^n$  qui associe à  $e_1^1 e_1^2 \dots e_1^n e_2^1 e_2^2 \dots e_2^n \dots e_l^1 e_l^2 \dots e_l^n$  l'élément :

$\left[ (e_1^1, e_1^2, \dots, e_1^n), (e_2^1, e_2^2, \dots, e_2^n), \dots, (e_l^1, e_l^2, \dots, e_l^n) \right]$  est une application bijective. En utilisant le critère du rang, la matrice  $B(c, e)$  donnée en (8) a au moins un rang égal à  $r$  pour chaque paire de mots distincts  $c$  et  $e$ . Dans ce cas,  $r$  lignes de  $B(c, e)$  sont non-nulles. Il s'ensuit que  $f(c)$  et  $f(e)$  ont une distance de Hamming au moins égale à  $r$  en tant que mots de code définis sur  $Q^l$ . L'alphabet  $Q^l$  a une taille égale à  $2^{bl}$  et ainsi le nombre de mots de code est majoré par  $A_{2^{bl}}(n, r)$ . Le taux de transmission maximum est bien alors celui donné par (23).

**Corollaire** : Toujours en considérant une transmission sur un canal de Rice avec  $n$  antennes de transmission et  $m$  antennes de réception, on trouve que si l'avantage en diversité est égal à  $n.m$ , alors le taux de transmission est au moins égal à  $b$  bits par seconde et par Hertz.

**Preuve**: on a évidemment  $A_{2^{bl}}(n, n) = 2^{bl}$  et ceci peut être atteint par un simple code à répétition.

**Remarque**: Pour des modulations MDP-4, MDP-8 ou MAQ-16, un avantage en diversité de  $nm$  donne une borne supérieure de respectivement 2, 3 et 4 bits/s/Hz pour le taux de transmission.

L'ensemble de ces premiers résultats conduit à la conclusion qu'il existe un compromis entre la taille de la constellation, l'ordre de diversité et le taux de transmission. Ce compromis est relié directement à la complexité du treillis réalisé.

**Lemme 1**: La longueur de contrainte d'un code STTC avec un avantage en diversité égal à  $r$  est au moins égale à  $r-1$ .

**Preuve:** Si on considère deux transitions (deux branches parallèles) correspondant à la longueur de contrainte  $v$  dans le treillis du code. Sans perdre en généralité, on peut supposer que l'une de ces transitions correspond au chemin nul  $0000\dots 0$  et l'autre correspond à

$$e_1^1, e_2^2, \dots, e_1^n, e_2^2, \dots, e_2^n \dots e_{v+1}^1, e_{v+1}^2, \dots, e_{v+1}^n 00\dots 0000$$

Si  $v < r - 1$  le critère du rang n'est plus vérifié.

**Lemme 2:** En appelant  $b$  le taux de transmission d'un système multiantennes employant un codage STTC avec un avantage égal à  $r$ , on montre que la complexité du treillis associé à ce code est au moins égal à  $2^{b(r-1)}$ .

**Preuve:** Comme le taux de transmission est égal à  $b$  bits par seconde et par Hertz, le nombre de branches qui partent de chaque état est égal à  $2^b$ . Ainsi, à l'instant  $r - 1$ , il y a  $2^{b(r-1)}$  chemins qui ont divergé depuis l'état zéro à l'instant  $t = 0$ . A cause du lemme 1, on sait qu'aucun de ces chemins ne reconverge vers le même état. Ainsi, il y a au moins  $2^{b(r-1)}$  états dans le treillis.

Les codes présentés en figures 3 et 6, par exemple, atteignent cette borne supérieure. Ainsi la borne du théorème 1 est une borne exacte (qui peut être atteinte par un code STTC optimal). Cela prouve également que les codes des figures 5-6 présentent le meilleur compromis entre le taux de transmission, le gain en diversité, la complexité du treillis et la taille de la constellation.

## **5-5 Les codes SOSTTC (Super-Orthogonal Space-Time Trellis Code)**

Depuis les travaux initiés par Tarokh & al dans [3], de nombreux schémas alternatifs ont été proposés pour améliorer les codes STTC [9-13]. En effet, un des principaux inconvénients des codes STTC réside dans le compromis entre le taux de transmission, le gain en diversité, la complexité du treillis et la taille de la constellation (cf partie ci-dessus), compromis qui empêche souvent le code STTC d'atteindre un gain en diversité maximum pour une complexité de treillis donnée. Cependant, on peut enlever cet inconvénient en utilisant les règles de partitionnement de Ungerboeck. Nous allons d'abord montrer comment utiliser les règles de partitionnement en sous-constellations [14]. Dans la suite de cette partie nous considérerons un système à deux antennes de transmission. Les matrices de transmission utilisées seront construites à partir du schéma d'Alamouti et prendront la forme suivante:

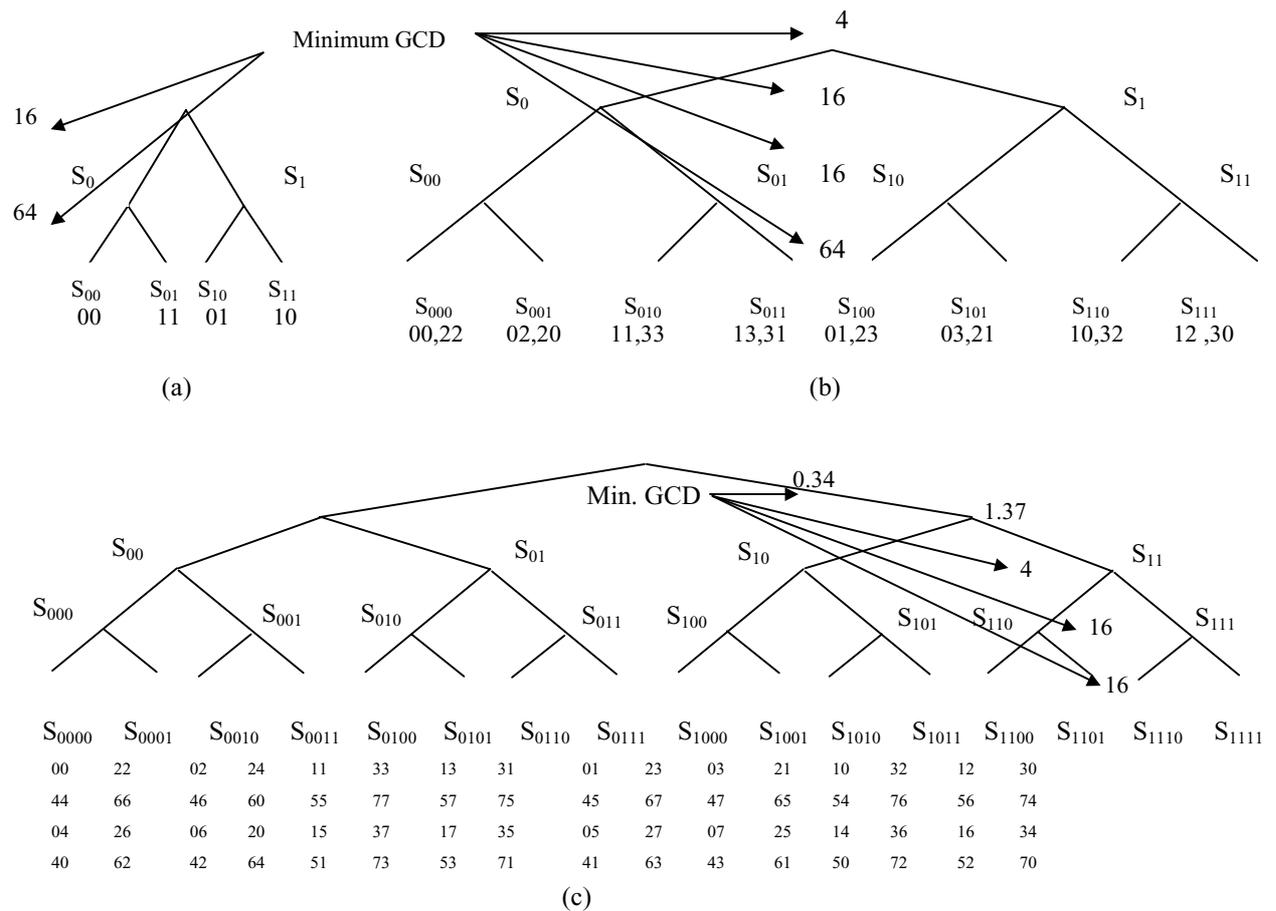
$$C(x_1, x_2, \theta) = \begin{pmatrix} x_1 \cdot e^{j\theta} & x_2 \\ -x_2^* \cdot e^{j\theta} & x_1^* \end{pmatrix} \quad (24)$$

La valeur de  $\theta$  est choisie de telle sorte que, pour chaque paire de mots  $x_1, x_2$  de la constellation initiale, les signaux transmis appartiennent toujours à cette même constellation. Par exemple, pour une modulation MDP- $L$ , les signaux  $x_1, x_2$  peuvent être mis sous la forme

$e^{j \frac{2\pi l}{L}}, l = 0, 1, \dots, L - 1$ . On peut alors prendre  $\theta = 2\pi l' / L, l' = 0, 1, \dots, L - 1$ . Dans ce cas les

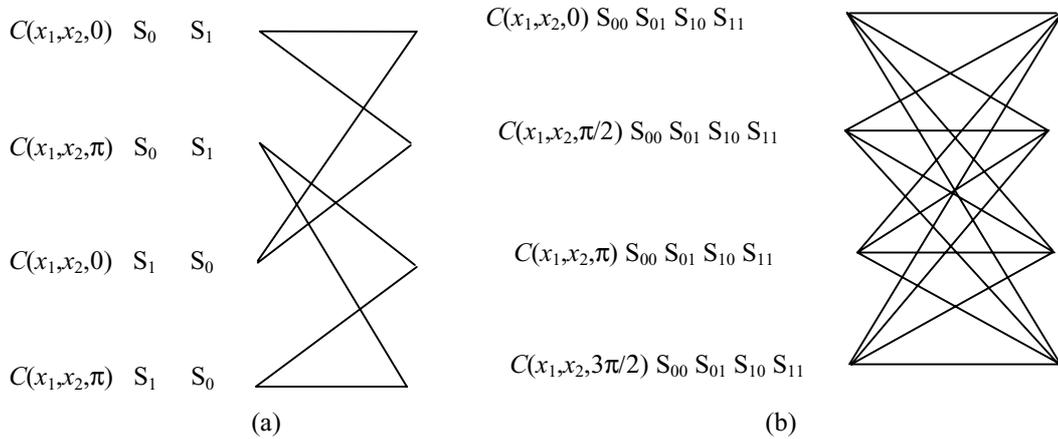
signaux transmis par le code STBC appartiendront toujours à la constellation de départ. Pour la modulation MDP-2, on peut utiliser  $C(x_1, x_2, 0)$  et  $C(x_1, x_2, \theta)$  et  $C(x_1, x_2, 0)$ ,  $C(x_1, x_2, \theta/2)$ ,  $C(x_1, x_2, \theta)$  et  $C(x_1, x_2, 3 \cdot \theta/2)$  pour la modulation MDP-4 et avec  $\theta = \pi/2$ .

Nous appelons  $B(c_1, c_2)$  la matrice des différences des mots de code  $c_1, c_2$  (cf équation 8) et  $B^H(c_1, c_2)$  sa transposée conjuguée. En reprenant les définitions du paragraphe 2, la diversité du code STTC est définie par le rang minimum de la matrice  $B(c_1, c_2)$ . Pour un code de diversité maximal, le minimum du déterminant de la matrice  $A(c_1, c_2) = B(c_1, c_2) \cdot B^H(c_1, c_2)$ , pour toutes les paires de mots de code ( $c_1, c_2$ ), correspond au gain de codage. On définit la distance en gain de codage (Coding Gain Distance: CGD) entre deux mots de code  $c_1$  et  $c_2$  par:  $d^2(c_1, c_2) = \det(A(c_1, c_2))$ . En général, lorsque la diversité du code est inférieure à  $r$  ( $r$ : nombre d'antennes d'émission), la distance peut être définie comme la moyenne harmonique des valeurs propres non nulles de  $A(c_1, c_2)$ . Le CGD nous permet alors de partitionner une constellation en sous-constellation. Un exemple très simple est d'abord donné dans le cas de la MDP-2. Il y a quatre façons de partitionner le code STBC de transmission comme le montre la figure 13.



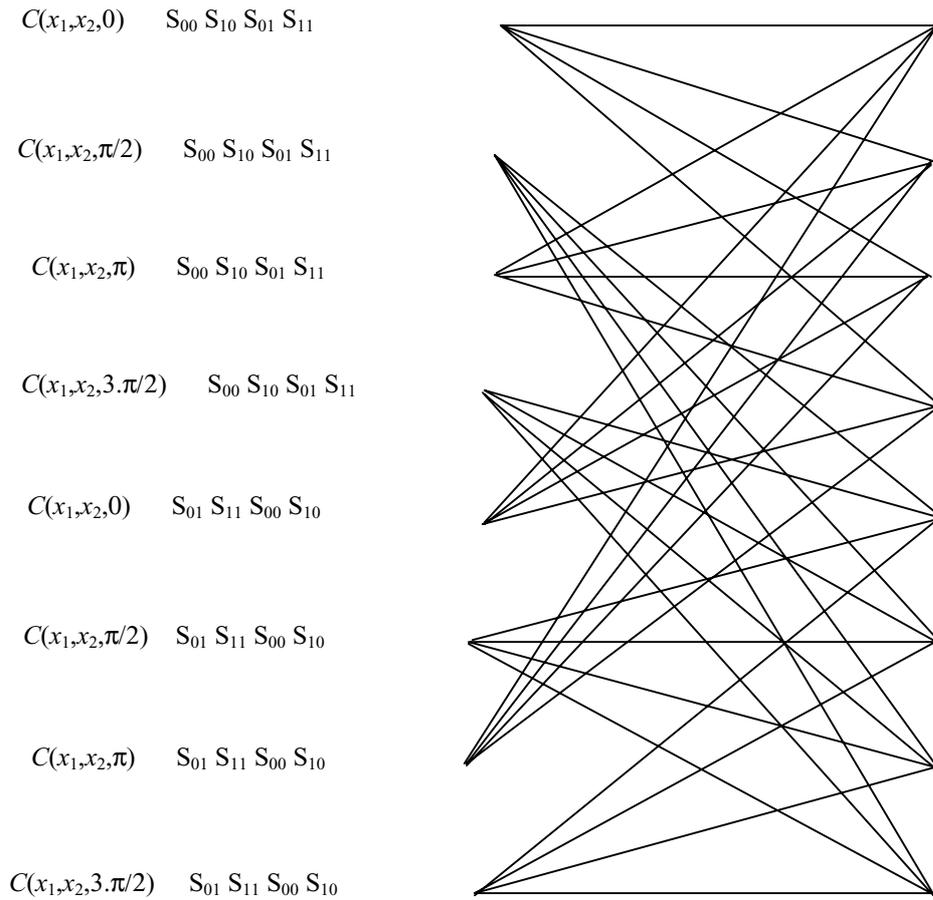
**Fig. 13: Principe de partitionnement d'après le CGD a) cas de la MDP-2, b) cas de la MDP-4, c) cas de la MDP-8**

Nous allons maintenant montrer comment maximiser le gain de codage sans sacrifier le taux de transmission. Nous affectons un code STBC du type de l'équation (24) à chaque transition depuis un état. Les états adjacents du treillis sont affectés chacun à un code STBC différent. Par contre, on peut affecter le même code STBC aux branches du treillis qui convergent vers un même état. Cette règle garantit que les chemins qui divergent (ou qui convergent) vers le bon chemin diffèrent tous entre eux d'un rang au moins égal à deux. En d'autres termes, chaque paire de mots de code convergeant ou divergeant vers un état réalise une diversité maximale car les mots appartiennent au même code STBC. Les codes réalisés portent le nom de SOSTTC (Super Orthogonal Space-time Trellis Codes). La figure 14 a) montre un exemple de code SOSTTC à quatre états  $r = 2$  bits/s/Hz (MDP-4) et la figure 14 b) un exemple de code SOSTTC à quatre états  $r = 3$  bits/s/Hz (MDP-8). Le CGD minimum du code figure 14 a) est égal à 16 ce qui est supérieur au CGD du code correspondant présenté par Tarokh dans [3]. Le CGD minimum du code figure 14 b) est égal à 2.69 et il n'y a pas de code STTC correspondant à quatre états en MDP-8 dans [3].



**Fig. 14: Exemples de codes SOSTTC à 4 états:**  
 **$r = 2$  bits/s/Hz (MDP-4) b)  $r = 3$  bits/s/Hz (MDP-8)**

La figure 15 présente un code SOSTTC à huit états en MDP-8  $r = 3$  bits/s/Hz.



**Fig. 15: Exemple de codes SOSTTC à huit états en MDP-8  $r = 3$  bits/s/Hz.**

Avant de montrer les performances obtenues avec les codes SOSTTC et l'amélioration apportée par rapport aux codes STTC classiques, nous donnons quelques outils mathématiques nécessaires à la partition de la constellation de départ en sous-constellations, le but étant, bien évidemment, comme dans le cas des modulations TCM de Ungerboeck, de maximiser la distance entre sous-constellations. Si on considère une modulation MDP- $L$  chaque signal peut se mettre sous la forme  $s = e^{j.2.\pi.l/L}$ ,  $l = 0, 1, \dots, L-1$  ou encore  $s = e^{j.\omega}$  en posant  $\omega = 2.\pi/L$ . Sachant que nous travaillons sur un système à deux antennes d'émission, nous devons alors considérer les paires de symboles distincts dans la constellation transmise (cf équation 24). Nous avons alors

$$(s_1^1 = e^{j.k_1.\omega}, s_2^1 = e^{j.l_1.\omega}) \quad \text{et} \quad (s_1^2 = e^{j.k_2.\omega}, s_2^2 = e^{j.l_2.\omega})$$

pour les symboles transmis par le code STBC utilisé ( $\theta = 0$ ). Nous calculons alors les matrices  $A$  et  $B$  définies au début de ce paragraphe et dans le paragraphe sur les critères de construction des codes STTC (cf paragraphe 2). Pour les branches parallèles dans le treillis, nous obtenons

$$B = \begin{pmatrix} e^{j.k_1.\omega} - e^{j.k_2.\omega} & e^{j.l_1.\omega} - e^{j.l_2.\omega} \\ e^{-j.l_2.\omega} - e^{-j.l_1.\omega} & e^{-j.k_1.\omega} - e^{-j.k_2.\omega} \end{pmatrix} \quad (25 - 26)$$

$$A = \begin{pmatrix} 4 - 2.\cos[\omega.(k_2 - k_1)] - 2.\cos[\omega.(l_2 - l_1)] & 0 \\ 0 & 4 - 2.\cos[\omega.(k_2 - k_1)] - 2.\cos[\omega.(l_2 - l_1)] \end{pmatrix}$$

En utilisant l'expression de la matrice  $A$  (équation 26), on obtient alors:

$$\det(A) = [4 - 2.\cos[\omega.(k_2 - k_1)] - 2.\cos[\omega.(l_2 - l_1)]]^2 \quad (27)$$

Ainsi, si pour le premier mot de code transmis on appelle l'ensemble des symboles de la sous-constellation

$$(s_1^1, s_2^1)^p = (e^{j.k_1^p.\omega}, e^{j.l_1^p.\omega}), p = 1, 2, \dots, P$$

et pour le second mot de code transmis

$$(s_1^2, s_2^2)^p = (e^{j.k_2^p.\omega}, e^{j.l_2^p.\omega}), p = 1, 2, \dots, P$$

on obtient alors:

$$\det(A) = \left\{ \sum_{p=1}^P [4 - 2.\cos[\omega.(k_2^p - k_1^p)] - 2.\cos[\omega.(l_2^p - l_1^p)]] \right\}^2 \quad (28)$$

L'égalité (28) regroupe une somme de  $P$  termes tous positifs, on a donc l'inégalité suivante:

$$\begin{aligned} \det(A) &= \left\{ \sum_{p=1}^P [4 - 2.\cos[\omega.(k_2^p - k_1^p)] - 2.\cos[\omega.(l_2^p - l_1^p)]] \right\}^2 \\ &\geq \sum_{p=1}^P \left\{ [4 - 2.\cos[\omega.(k_2^p - k_1^p)] - 2.\cos[\omega.(l_2^p - l_1^p)]] \right\}^2 \end{aligned} \quad (29)$$

L'inégalité (29) permet alors de construire les sous-constellations définies sur la figure 13.

#### - Performances obtenues

Nous nous plaçons dans le même contexte que pour l'étude des codes STTC. Les performances des codes SOSTTC sont illustrées sur les figures 16-17 dans le cas de deux antennes d'émission et de un ou deux antennes de réception. Les paquets transmis ont une taille de 260 symboles (130 sur chaque antenne). Le code utilisé est celui présenté en figure 14 a) en MDP-4. Les coefficients  $\alpha_{i,j}$  sont simulés par des échantillons de bruit complexes Gaussiens de moyenne éventuellement nulle et de variance 0.5 sur chaque partie réelle ou imaginaire. A titre de comparaison, nous donnons aussi les performances obtenues par les codes donnés en [10-12].

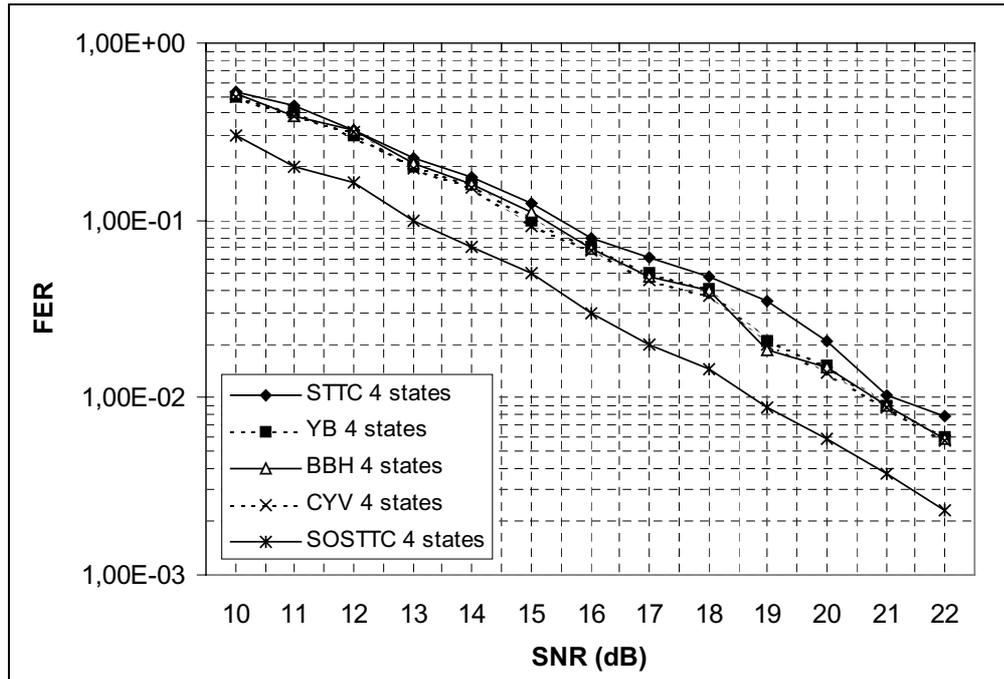


Fig 16 : Performances des codes SOSTTC, cas de deux antennes d'émission et une antenne de réception (2 bits/s/Hz-MDP-4)

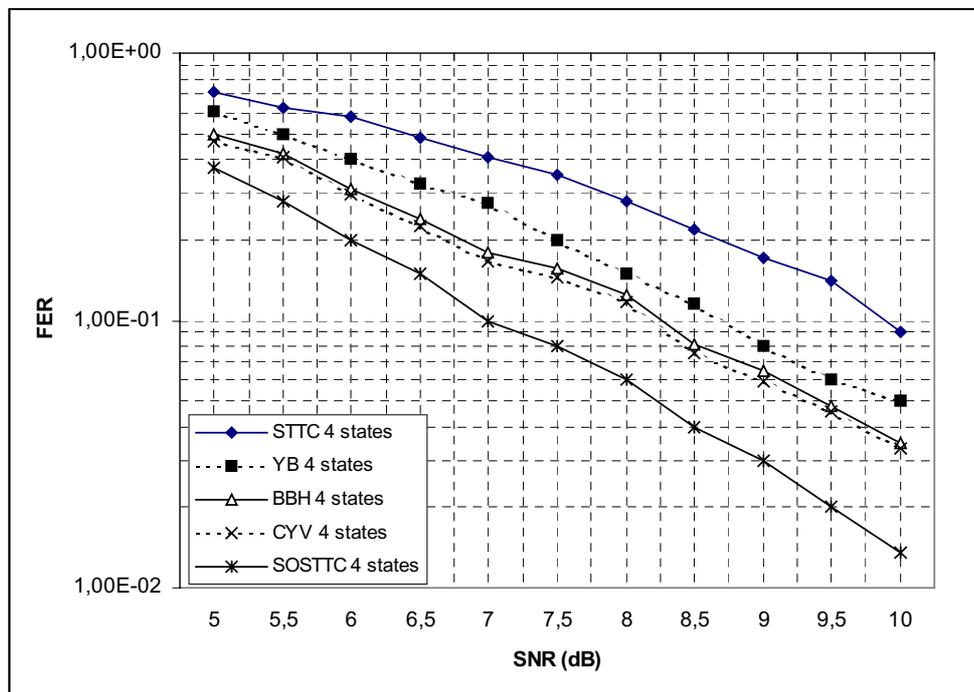


Fig 17 : Performances des codes SOSTTC, cas de deux antennes d'émission et de deux antennes de réception (2 bits/s/Hz-MDP-4)

Les codes qui servent de comparaison, outre les codes STTC présentés dans [3], sont les codes de Q. Yan et R. S. Blum [11] notés YB sur les courbes, les codes de S. Baro, G. Bauch et A.

Hansmann [12] notés BBH et les codes de Z. Chen, J. Yuan et B. Vucetic [13] notés CYV. Il s'agit, dans la littérature, des codes STTC les plus performants réalisés jusqu'alors.

Il est clair d'après les résultats des figures 16 et 17 que les codes SOSTTC présentent un avantage conséquent par rapport aux codes STTC, cet avantage, en termes de FER, étant de l'ordre de 1-1,5 dB (cas des FER à  $10^{-2}$ ) pour le cas où l'on dispose de deux antennes de réception (figure 17) et de 2 dB dans le cas où l'on utilise une seule antenne de réception (figure 16). Ainsi, l'avantage sur les codes STTC diminue lorsque le nombre d'antennes de réception augmente. Il est à signaler de plus que les performances des codes YB, BBH et CYV sont très proches de celles des codes STTC de Tarokh dans le cas d'une seule antenne de réception (figure 16). Leurs avantages, par rapport aux STTC, n'apparaît que dans le cas de deux antennes de réception (figure 17).

**Remarque:** l'algorithme de décodage fait appel à la métrique de maximum de vraisemblance donnée dans le chapitre sur les codes STBC. Pour chaque sous-constellation on calcule la métrique STBC la plus petite (la plus vraisemblable) et on la cumule dans le treillis de Viterbi comme pour n'importe quel code convolutif conventionnel. L'algorithme de décodage est présenté sous forme d'un fichier Matlab à la fin de ce chapitre pour le code SOSTTC de la figure 14 a).

## **5-6 Conclusion et perspectives**

Nous avons présenté dans ce chapitre une nouvelle catégorie de codes spatio-temporels qui s'apparentent à la classe des codes correcteurs convolutifs, ces codes sont appelés codes spatio-temporels en treillis (Space-Time Trellis Codes). En effet, les symboles sont transmis sur les différentes antennes d'émission suivant un schéma en treillis et le décodage de ce type de code est réalisé avec l'algorithme de Viterbi. Leur gain de codage est supérieur aux codes STBC qui ne sont que des schémas de diversité optimaux, cependant le gain en diversité n'est pas toujours maximum. Après avoir examiné les principales règles de construction de ces codes, en particulier le critère du rang et le critère du déterminant, nous avons exposé des schémas de réalisation pour des modulations MDP-4, MDP-8 avec des taux de transmission de 2 et 3 bits/s/Hz. Les simulations réalisées ont permis de constater que les codes STTC approchaient la borne de la capacité Outage à 2.5 dB près. Nous avons alors proposé de combiner les avantages des codes STTC (gain de codage) avec les avantages des codes STBC (diversité optimale) pour obtenir un nouveau type de codes: les codes SOSTTC (Super Orthogonal Space-Time Trellis Codes). Ce type de code permet d'améliorer sensiblement les performances des codes STTC. En effet, des gains de 1 à 2 dB sont observés par rapport aux codes STTC de Tarokh.

Pour compléter cette étude, il est évident qu'une extension des codes SOSTTC au cas de schémas de transmission utilisant plus de deux antennes doit être envisagé. Le compromis complexité du treillis-gain de codage doit être également mieux étudié. Enfin, l'influence d'un algorithme d'estimation des paramètres du canal (détermination de la réponse impulsionnelle) doit être prise en compte dans l'estimation des performances.

## **5-7 Références**

- [1] A. Wittneben, "Base station modulation diversity for digital SIMULCAST", in *Proc IEEE VTC*, May 1993, pp. 505-511.
- [2] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity", *Int. J. Wireless Inform. Networks*, vol. 1, n°1, 1994.
- [3] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-Time codes for high data rate wireless communication: performance criterion and code construction", *IEEE Trans. Inform. Theory*, vol. 44, n°2, march 1998.
- [4] J. C. Guey, M. P. Fitz, M. R. Bell and W. Y. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels", in *Proc IEEE VTC'96*, pp. 136-140.
- [5] G. J. Foschini, Jr and M. J. Gans, "On limits of wireless communication in a fading environment when using multiple antennas", *Int J. Wireless Person. Comm.*, Sept. 99.
- [6] E. Telatar, "Capacity of multi-antenna Gaussian channels", *AT&T-Bell Labs Internal Tech. Memo.*, june 1995.
- [7] V. Tarokh, A. Naguib, N. Seshadri and A. R. Calderbank, "Array processing and space-time coding for very high data rate wireless communication", *IEEE Trans. Inform. Theory*, vol. 45, n°8, September 1999.
- [8] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels", *IEEE Trans. Inform. Theory*, vol. 45, pp. 1639-1642, July 1999.
- [9] H. E. Gamal and A. R. Hammons, "On the design and performance of algebraic Space-Time codes for BPSK and QPSK modulation", *IEEE Trans. Commun.*, vol. 50, n°6, pp. 907-913, June 2002.
- [10] J. Grimm, M. P. Fitz and J. V. Krogmeier, "Further results on space-time codes for Rayleigh fading", in *Proc. 36th Annu. Allerton Conf.*, Sept. 1998, pp. 391-400.
- [11] Q. Yan and R. S. Blum, "Optimum space-time convolutional codes for quasistatic slow fading channels", in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, vol. 3, Sept. 2000, pp. 1351-1355.
- [12] S. Baro, G. Bauch and A. Hansmann, "Improved codes for space-time trellis-coded modulation", *IEEE Commun. Letters*, vol. 4, pp. 20-22, Jan. 2000.
- [13] Z. Chen, J. Yuan and B. Vucetic, "Improved space-time trellis coded modulation scheme on slow fading channels", *Electron. Letters*, vol. 37, pp. 440-441, March 2001.
- [14] H. Jafarkhani, N. Seshadri, "Super-Orthogonal space-time trellis codes", *IEEE Trans. Inform. Theory*, vol. 49, n°4, pp. 937-950, April 2003.



## *Annexe A : Quelques propriétés matricielles*

Une matrice  $A_{n \times n}$  est dite Hermitienne si et seulement si elle vérifie  $A = A^*$ . La matrice  $A$  est dite définie positive si pour tout vecteur complexe  $x$  de taille  $1 \times n$ , on a  $x.Ax^* \geq 0$ . Une matrice complexe  $V$  est dite unitaire si on a la relation  $V.V^* = I$  où  $I$  désigne évidemment la matrice identité d'ordre  $n$ . Une matrice  $B$  de taille  $n \times l$  est dite racine carré de  $A$  si on a la relation  $B.B^* = A$ . Nous utiliserons également par la suite les résultats suivants :

- Un vecteur propre  $v$  d'une matrice  $A_{n \times n}$  est dit associé à la valeur propre  $\lambda$  de  $A$  lorsqu'il vérifie  $vA = \lambda.v$ . L'espace engendré par les vecteurs propres de  $A$  associés à la valeur propre 0 correspond au noyau de  $A$  et a pour dimension  $n-r$  où  $r$  est le rang de  $A$ .
- Toute matrice  $A$  qui admet une racine carré  $B$ , est une matrice définie positive.
- Pour chaque matrice  $A_{n \times n}$  hermitienne définie positive il existe une matrice  $B_{n \times n}$  triangulaire inférieure telle que  $B.B^* = A$ .
- Pour une matrice Hermitienne  $A$ , l'espace engendré par ses vecteurs propres est égal à  $C^n$ , espace de dimension  $n$ . Il est alors possible en utilisant le procédé de Gram-Schmidt de construire une base orthonormée de  $C^n$  en utilisant les vecteurs propres de  $A$ . De plus, dans ce cas, il existe une matrice  $V$  unitaire telle que  $V.A.V^* = D$ . Les lignes de  $V$  constituent une base orthonormale de  $C^n$  donnée par les vecteurs propres de  $A$ .  $D$  est une matrice diagonale dont les éléments correspondent aux valeurs propres  $\lambda_i$  de  $A$ , comptées avec leurs ordres de multiplicité.
- Les valeurs propres d'une matrice Hermitienne sont réelles.
- Les valeurs propres d'une matrice Hermitienne définie positive sont toutes positives ou nulles.



## *Annexe B : liste des programmes Matlab*

### 1-Algorithme de codage des codes SOSTTC

```
function [output,next_state] = encodeTTCsymbol(input, state)

if (state==1) % on est dans le cas de la matrice d'Alamouti avec angle == 0
    if (all(input==[1 1]) | all(input==[-1 -1]) | all(input==[1 -1]) |
all(input==[-1 1]))...
        | all(input==[exp(j*pi/2) exp(j*pi/2)]) |
all(input==[exp(j*3*pi/2) exp(j*3*pi/2)])...
        | all(input==[exp(j*pi/2) exp(j*3*pi/2)]) |
all(input==[exp(j*3*pi/2) exp(j*pi/2)]))
        output = [input(1) input(2); -conj(input(2)) conj(input(1))];
        next_state=1; % on reste dans le même état
    else
        output = [input(1) input(2); -conj(input(2)) conj(input(1))];
        next_state=2; % on change d'état
    end;

elseif (state==2) % on est dans le cas de la matrice d'Alamouti avec angle
== pi
    if(all(input==[1 1]) | all(input==[-1 -1]) | all(input==[1 -1]) |
all(input==[-1 1]))...
        | all(input==[exp(j*pi/2) exp(j*pi/2)]) |
all(input==[exp(j*3*pi/2) exp(j*3*pi/2)])...
        | all(input==[exp(j*pi/2) exp(j*3*pi/2)]) |
all(input==[exp(j*3*pi/2) exp(j*pi/2)]))
        output = [-input(1) input(2); conj(input(2)) conj(input(1))];
        next_state=3; % on change d'état
    else
        output = [-input(1) input(2); conj(input(2)) conj(input(1))];
        next_state=4; % on change d'état
    end;

elseif (state==3) % on est dans le cas de la matrice d'Alamouti avec angle
== 0
    if(all(input==[1 1]) | all(input==[-1 -1]) | all(input==[1 -1]) |
all(input==[-1 1]))...
        | all(input==[exp(j*pi/2) exp(j*pi/2)]) |
all(input==[exp(j*3*pi/2) exp(j*3*pi/2)])...
        | all(input==[exp(j*pi/2) exp(j*3*pi/2)]) |
all(input==[exp(j*3*pi/2) exp(j*pi/2)]))
        output = [input(1) input(2); -conj(input(2)) conj(input(1))];
        next_state=2; % on change d'état Fig. 5
    else
        output = [input(1) input(2); -conj(input(2)) conj(input(1))];
        next_state=1; % on change d'état
    end;

elseif (state==4) % on est dans le cas de la matrice d'Alamouti avec angle
== pi
    if( all(input==[1 1]) | all(input==[-1 -1]) | all(input==[1 -1]) |
all(input==[-1 1]) ...
        | all(input==[exp(j*pi/2) exp(j*pi/2)]) |
all(input==[exp(j*3*pi/2) exp(j*3*pi/2)]) ...
        | all(input==[exp(j*pi/2) exp(j*3*pi/2)]) |
all(input==[exp(j*3*pi/2) exp(j*pi/2)]))
        output = [-input(1) input(2); conj(input(2)) conj(input(1))];
        next_state=4; % on reste dans le même état
    else
        output = [-input(1) input(2); conj(input(2)) conj(input(1))];
```

```

        next_state=3; % on change d'état
    end;
end;

```

## 2-Algorithmme de décodage des codes SOSTTC

```

function x_hat = decodeSOTTCbis(z,alpha1,alpha2)

[m,K] = size(z);
max_state = 4;
n=2;
L_total = m
s0=1;
s1=exp(j*pi/2);
s2=-1;
s3=exp(j*3*pi/2);;
% construction de la matrice de test pour le maximum de vraisemblance %
% première sous-constellation
MV(1,1)=s0; MV(1,2)=s0;
MV(2,1)=s2; MV(2,2)=s2;
MV(3,1)=s0; MV(3,2)=s2;
MV(4,1)=s2; MV(4,2)=s0;
MV(5,1)=s1; MV(5,2)=s1;
MV(6,1)=s3; MV(6,2)=s3;
MV(7,1)=s1; MV(7,2)=s3;
MV(8,1)=s3; MV(8,2)=s1;
% deuxième sous-constellation
MV(9,1)=s0; MV(9,2)=s1;
MV(10,1)=s2; MV(10,2)=s3;
MV(11,1)=s0; MV(11,2)=s3;
MV(12,1)=s2; MV(12,2)=s1;
MV(13,1)=s1; MV(13,2)=s0;
MV(14,1)=s3; MV(14,2)=s2;
MV(15,1)=s1; MV(15,2)=s2;
MV(16,1)=s3; MV(16,2)=s0;

st_metric10(1:L_total/n)=10^2;
st_metric20(1:L_total/n)=10^2;
st_metric1pi(1:L_total/n)=10^2;
st_metric2pi(1:L_total/n)=10^2;

% set infinity to an arbitrarily large value
inf = 10^5;

% initialize trellis and path matrices
trellis = inf*ones(max_state,L_total/n);
path = zeros(max_state,L_total);
new_path = path;

% determine trellis and path matrices at time 1
y_segment1 = z(1:n,:);
P0 = [y_segment1(1) y_segment1(2)]*[conj(alpha1(1)) conj(alpha2(1)); 0 0] +
...
[conj(y_segment1(1)) conj(y_segment1(2))]*[0 0 ; alpha2(2) -alpha1(2)];
P0 = (1/((abs(alpha1(1)))^2+(abs(alpha2(1)))^2))*P0;

% choose the most likely transmitted symbols using P vector components
% for the points belonging to the first subset
for i=1:8
    ecart(i) = abs(MV(i,1)-P0(1))^2+abs(MV(i,2)-P0(2))^2;
    if (ecart(i)<st_metric10(1))
        st_metric10(1) = ecart(i);
        min10=i;
    end
end

```

```

    end;
end;
% in this case, the next state is state 1
next_state=1;
trellis(next_state,1)=st_metric10(1);
path(next_state,1:2)=[MV(min10,1) MV(min10,2)];

% choose the most likely transmitted symbols using P vector components
% for the points belonging to the second subset
for i=9:16
    ecart(i) = abs(MV(i,1)-P0(1))^2+abs(MV(i,2)-P0(2))^2;
    if (ecart(i)<st_metric20(1))
        st_metric20(1) = ecart(i);
        min20=i;
    end;
end;
% in this case, the next state is state 2
next_state=2;
trellis(next_state,1)=st_metric20(1);
path(next_state,1:2)=[MV(min20,1) MV(min20,2)];

% determine trellis and path matrices from time 2 to time L_total
counter=n+1;
for time=2:(L_total/n)
    y_segment1 = z(counter:counter+n-1,:);
    P0 = [y_segment1(1) y_segment1(2)]*[conj(alpha1(counter))
conj(alpha2(counter)) ; 0 0] + ...
        [conj(y_segment1(1)) conj(y_segment1(2))]*[0 0; alpha2(counter+1) -
alpha1(counter+1)];
    P0= (1/((abs(alpha1(counter)))^2+(abs(alpha2(counter)))^2))*P0;
    for i=1:8
        ecart(i) = abs(MV(i,1)-P0(1))^2+abs(MV(i,2)-P0(2))^2;
        if (ecart(i)<st_metric10(time))
            st_metric10(time) = ecart(i);
            min10=i;
        end;
    end;
    for i=9:16
        ecart(i) = abs(MV(i,1)-P0(1))^2+abs(MV(i,2)-P0(2))^2;
        if (ecart(i)<st_metric20(time))
            st_metric20(time) = ecart(i);
            min20=i;
        end;
    end;
    Ppi = [y_segment1(1) y_segment1(2)]*[-conj(alpha1(counter))
conj(alpha2(counter)) ; 0 0] + ...
        [conj(y_segment1(1)) conj(y_segment1(2))]*[0 0; alpha2(counter+1)
alpha1(counter+1)];
    Ppi= (1/((abs(alpha1(counter)))^2+(abs(alpha2(counter)))^2))*Ppi;
    for i=1:8
        ecart(i) = abs(MV(i,1)-Ppi(1))^2+abs(MV(i,2)-Ppi(2))^2;
        if (ecart(i)<st_metric1pi(time))
            st_metric1pi(time) = ecart(i);
            min1pi=i;
        end;
    end;
    for i=9:16
        ecart(i) = abs(MV(i,1)-Ppi(1))^2+abs(MV(i,2)-Ppi(2))^2;
        if (ecart(i)<st_metric2pi(time))
            st_metric2pi(time) = ecart(i);
            min2pi=i;
        end;
    end;
end;
for state=1:max_state
    if state==1
        % in this case, Alamouti's scheme with theta = 0

```

```

    % in this case, the next state is state 1
    next_state=1;
    path_metric = st_metric10(time) + trellis(1,time-1);
    if path_metric < trellis( next_state, time )
        trellis( next_state, time ) = path_metric;
        new_path(next_state,1:n*time)= [path(state,1:n*(time-1))
MV(min10,:)]];
    end;
    % in this case, the next state is state 2
    next_state=2;
    path_metric = st_metric20(time) + trellis(1,time-1);
    if path_metric < trellis( next_state, time )
        trellis( next_state, time ) = path_metric;
        new_path(next_state,1:n*time)= [path(state,1:n*(time-1))
MV(min20,:)]];
    end;

elseif state==2
    % in this case, Alamouti's scheme with  $\theta = \pi$ 
    % in this case, the next state is state 3
    next_state=3;
    path_metric = st_metric1pi(time) + trellis(2,time-1);
    if path_metric < trellis( next_state, time )
        trellis( next_state, time ) = path_metric;
        new_path(next_state,1:n*time)= [path(state,1:n*(time-1))
MV(min1pi,:)]];
    end;
    % in this case, the next state is state 4
    next_state=4;
    path_metric = st_metric2pi(time) + trellis(2,time-1);
    if path_metric < trellis( next_state, time )
        trellis( next_state, time ) = path_metric;
        new_path(next_state,1:n*time)= [path(state,1:n*(time-1))
MV(min2pi,:)]];
    end;

elseif state==3
    % in this case, Alamouti's scheme with  $\theta = 0$ 
    % in this case, the next state is state 2
    next_state=2;
    path_metric = st_metric10(time) + trellis(3,time-1);
    if path_metric < trellis( next_state, time )
        trellis( next_state, time ) = path_metric;
        new_path(next_state,1:n*time)= [path(state,1:n*(time-1))
MV(min10,:)]];
    end;
    % in this case, the next state is state 1
    next_state=1;
    path_metric = st_metric20(time) + trellis(3,time-1);
    if path_metric < trellis( next_state, time )
        trellis( next_state, time ) = path_metric;
        new_path(next_state,1:n*time)= [path(state,1:n*(time-1))
MV(min20,:)]];
    end;

else
    % in this case, Alamouti's scheme with  $\theta = \pi$ 
    % in this case, the next state is state 4
    next_state=4;
    path_metric = st_metric1pi(time) + trellis(4,time-1);
    if path_metric < trellis( next_state, time )
        trellis( next_state, time ) = path_metric;
        new_path(next_state,1:n*time)= [path(state,1:n*(time-1))
MV(min1pi,:)]];
    end;
    % in this case, the next state is state 3

```

```

        next_state=3;
        path_metric = st_metric2pi(time) + trellis(4,time-1);
        if path_metric < trellis( next_state, time )
            trellis( next_state, time ) = path_metric;
            new_path(next_state,1:n*time)= [path(state,1:n*(time-1))
MV(min2pi, :)]];
        end;
    end;
end;
path=new_path;
counter=counter+n;
end;
%selection of the best surviving paths among the remaining max_states ones
trellis_min=10^5;;
for i=1:max_state
    if trellis(i,L_total/n)<trellis_min
        trellis_min=trellis(i,L_total/n);
        MVstate=i;
    end;
end;
x_hat = path(MVstate,1:L_total);
return

```

### 3-Programme de test des codes SOSTTC

```

function [SR,SER] = SOTTC
%M nombre d'états de la modulation, ici une MDP4 %
M=2;
n=50; % taille d'un paquet
P=2; % nombre d'antennes d'émission
fd=100; % frequence Doppler
coding_rate=1;
Nbpaket=10;
SNR=[6];

% initialisation des SER par paquet et par SNR
for da=1:length(SNR)
    SER(da,1)=0;
    for jj=1:Nbpaket
        num_error(da,jj)=0;
        SR(da,jj)=0;
    end;
end;
% création de la modulation MDP4 %
for da=1:length(SNR)
    for num_packet=1:Nbpaket
        seq(num_packet,:)=(rand(1,n*M)>0.5);
        mseq=seq(num_packet,:);
        for i=1:n
            sn=mseq(M*(i-1)+1:M*i);
            if(sn==[0 0])
                vn(i)=1;
            elseif(sn==[0 1])
                vn(i)=exp(j*pi/2);
            elseif(sn==[1 1])
                vn(i)=-1;
            else
                vn(i)=exp(j*3*pi/2);
            end;
        end;
    end;
end;
% création du codage SOTTC : les symboles arrivent

```

```

% deux par deux sur l'antenne d'émission %
state = 1; % on initialise l'état dans lequel on part à 1
for m=1:(n/P)
    en=vn(P*(m-1)+1:P*m);
    [zn, next_state]=encodeTTCsymbol(en, state);
    A(P*(m-1)+1:P*m,1)=zn(:,1);
    A(P*(m-1)+1:P*m,2)=zn(:,2);
    state = next_state;
end;
x1=fadingtaps(fd,1,1,n,512);
x2=fadingtaps(fd,1,1,n,512);
A(:,1)=A(:,1).*x1;
A(:,2)=A(:,2).*x2;
z=A(:,1)+A(:,2);
[zz,d]=awgncomplex(coding_rate,n,SNR(da),z);
x_hat = decodeSOTTCbis(zz, x1, x2);
error = x_hat-vn;
for k=1:n
    if abs(error(:,k))>0.1
        num_error(da,num_packet)=num_error(da,num_packet)+1;
    end;
end;
SR(da,num_packet)=num_error(da,num_packet)/n;
end;
SER(da)=sum(SR(da,:),2)/Nbpaket;
end;
return

```

**CONCLUSION GENERALE ET  
PERSPECTIVES**



## CONCLUSION GENERALE ET PERSPECTIVES

Les différents chapitres de ce travail de thèse visent à étudier les performances des systèmes MIMO. De nombreux résultats peuvent être déduits des performances présentées. En ce qui concerne les systèmes BLAST nous avons entre autres, montré que leurs performances dépendait surtout de l'écart entre le nombre d'antennes de réception et d'émission. De plus, ces systèmes étant très sensibles aux erreurs de propagation, il faut utiliser des systèmes de codage performants (tels que les turbo codes) à inclure dans une structure telle que le D-BLAST. On aboutit ainsi à un système de type turbo-BLAST que le groupe de recherche GESTE étudie actuellement. Pour les systèmes à codage spatio temporels de type STBC nous avons redéfini les règles matricielles de codage et de décodage pour un nombre quelconque d'antennes d'émission et de réception et nous avons calculé une borne très précise des performances en termes de TEB sur canaux de Rayleigh non-sélectifs en fréquence. Cette borne coïncide avec les résultats de simulation à moins de 0.5 dB près ce qui est remarquable et bien plus précis que la borne classique de Chernoff. Pour les codes STTC nous avons d'abord repris les schémas proposés initialement par Tarokh, nous avons vérifié qu'ils permettaient d'atteindre la capacité "outage" à 2-2.5 dB près et nous avons proposé un schéma de codage original qui combine l'avantage en diversité des codes STBC avec le gain de codage des codes STTC. Ce type de code appelé SOSTTC (Super Orthogonal Space Time Trellis Code) permet d'améliorer sensiblement les performances des codes STTC. En effet, des gains de 1 à 2 dB sont observés par rapport aux codes STTC de Tarokh. Pour compléter cette étude, il est évident qu'une extension des codes SOSTTC au cas de schémas de transmission utilisant plus de deux antennes doit être envisagée. Le compromis complexité du treillis-gain de codage doit être également mieux étudié.

Il est à signaler que le canal de propagation a été supposé non-sélectif en fréquence dans chacun des chapitres. La généralisation des résultats au cas de canaux sélectifs en fréquence devra être abordé. Enfin, et ceci est valable pour l'ensemble des chapitres, l'influence d'un algorithme d'estimation des paramètres du canal (détermination de la réponse impulsionnelle) doit être prise en compte dans l'estimation des performances.

Nous étudions actuellement la combinaison des codes STTC, STBC et SOSTTC avec des codages turbo à haute efficacité spectrale du type de ceux présentés dans le chapitre 1. De plus, l'utilisation d'algorithmes d'apprentissage du canal en conjonction avec l'utilisation de turbo codes permet d'aboutir à des structures de turbo-égaliseurs qui restent délicates à mettre au point pour les systèmes MIMO dans le cas de canaux sélectifs en fréquence. Ceci constitue certainement une voie de recherche d'avenir.

