



**HAL**  
open science

# Détection statistique d'informations cachées dans une image naturelle à partir d'un modèle physique

Rémi Cogranne

► **To cite this version:**

Rémi Cogranne. Détection statistique d'informations cachées dans une image naturelle à partir d'un modèle physique. Méthodologie [stat.ME]. Université de Technologie de Troyes, 2012. Français. NNT: . tel-00706171v2

**HAL Id: tel-00706171**

**<https://theses.hal.science/tel-00706171v2>**

Submitted on 11 Jun 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Détection statistique d'informations cachées dans une image naturelle à partir d'un modèle physique

Rémi COGRANNE



Manuscrit soumis à l'Université de Technologie de Troyes (UTT) en vue de l'obtention du titre de Docteur. Mention : Optimisation et Sécurité des Systèmes (OSS)

Directeurs de thèse : Lionel FILLATRE et Florent RETRAINT

Soutenance publique le 2 décembre 2011 à Troyes devant le jury suivant :

<b>Rapporteurs</b>	M. Jean-Marc CHASSERY	Directeur de recherche au CNRS, Gipsa-lab, Grenoble.
	M. Eric MOULINES	Professeur de Telecom-ParisTech, Paris.
<b>Examineurs</b>	M. Patrick BAS	Chargé de recherche au CNRS, Lagis, Lille.
	M. Lionel FILLATRE	Maître de conférence, co-directeur de thèse, UTT.
	M. Igor NIKIFOROV	Professeur des universités, UTT.
	M. William PUECH	Professeur des universités, LIRMM, Montpellier.
	M. Florent RETRAINT	Enseignant-chercheur, co-directeur de thèse, UTT.
<b>Invités</b>	M. Johann BARBIER	Novetix
	M. Didier NICHOLSON	Thalès Land & Joint Systems
	M. François DELOST	Ministère de l'intérieur



Rémi Cogranne : *Détection statistique d'informations cachées dans une image naturelle à partir d'un modèle physique*, Manuscrit soumis à l'Université de Technologie de Troyes (UTT) pour l'obtention du titre de Docteur, mention Optimisation et sûreté des systèmes.

Version préliminaire du 5 mars 2012

Ce mémoire ainsi que les images qu'il contient peuvent être utilisés librement ; les sources sont également disponibles sur simple demande.

Manuscrit rédigé sous  $\text{\LaTeX}_{\varepsilon}$  avec la classe *classicthesis*

*À Laure, ma muse, et à son admirable courage.*

*À Maelle, ma fille, et à l'avenir qui s'ouvre devant elle.*

*À ma mère, pour son soutien indéfectible.*

*«Moi qui pensais qu'il était bien ».  
Comme dans les poèmes de Richepin ».  
Moi qui croyais qu'il était sage ».  
D'être un oiseau de passage ».  
Au matin pris mon sac à dos ».  
Partis au fil de l'eau ».*

— Fred Burguière, 1976-...

## Remerciements

---

Ce travail a été réalisé au Laboratoire de Modélisation et de Sécurité des Systèmes (LM2S) de l'Université de Technologie de Troyes (UTT). L'ensemble des membres du LM2S m'ont réservé un accueil plus que chaleureux et ont maintenu un environnement de travail des plus agréables, propice au bon déroulement et à l'aboutissement de ce travail. Je tiens à les en remercier avant de commencer ce mémoire. Sans faire une liste exhaustive, je tiens à remercier particulièrement Madame Véronique Banse et Madame Marie-José Rousselet, nos Ô combien précieuses secrétaire de pôle.

Je tiens à remercier particulièrement Monsieur Florent Retrait, enseignant-chercheur à l'UTT qui m'a motivé à entamer ces études de doctorat. En tant que co-directeur de ces travaux de recherche ses conseils avisés et nos discussions sans fin m'ont été extrêmement bénéfique. Je tiens également à lui témoigner toute ma sympathie pour les qualités humaines, le soutien et la disponibilité dont il a fait preuve.

Je remercie également Monsieur Lionel Fillatre, maître de conférence à l'UTT, pour avoir co-dirigé cette thèse avec une grande efficacité en vertu de sa grande disponibilité et de ses avis pertinents ont été précieux. Qu'il soit assuré de ma reconnaissance.

Je tiens également à remercier Monsieur Igor Nikiforov, professeur à l'UTT et directeur scientifique du projet RIC dans lequel s'inscrivent ces travaux. Si nos rapports ont été moins réguliers qu'avec mes directeurs de thèses, il m'a apporté énormément au niveau personnel et professionnel ; il m'a transmis plus qu'une méthodologie, une véritable "philosophie" des statistiques.

J'exprime ma profonde gratitude envers Monsieur Jean-Marc Chasery, directeur de recherche au CNRS (GIPSA-Lab), et Monsieur Éric Moulines, professeur à Télécom-ParisTech, pour avoir accepté de rapporter le présent mémoire. Leur rapports témoignent du temps qu'ils y ont consacré, et je dois admettre que leur lecteur m'a profondément touché.

J'exprime ma reconnaissance envers Monsieur Patrick Bas, chargé de recherche au CNRS (LAGIS) et enseignant à l'école centrale de Lille, et Monsieur William Puech, professeur à l'université de Montpellier, d'avoir accepté d'examiner cette thèse.

De la même façon je remercie Monsieur Johann Barbier, anciennement chercheur au CELAR, DGA, créateur de la société Novetix, Monsieur François Delost, notre contact avec la Direction Générale de la Police Nationale (DGPN) dans le cadre de ce projet, et Monsieur Didier Nicholson de Thalès Land & Joint Systems d'avoir accepté de compléter la composition du jury de cette thèse.

Enfin, j'exprime ma plus sincère sympathie à l'ensemble des personnes qui m'ont accompagnés durant cette thèse. Il serait déplacé

de ma part de ne pas remercier très sincèrement Cathel, Sabrina, Laurent et Léopold qui ont partagés avec moi leur bureau et accessoirement des bons moments, ainsi que Mickael et Rémi S., squatter de bureau. De même, je me dois de remercier (pèle-mêle) Léa, Manu, Seb, Moumoune, Guillaume, Stéphanie, Raphael, Tatiana, Laurent, Manuel, Marion, pour les (très) heureux moments vécus ensemble. Je n'oublies pas non plus les équipes successives d'Ellidoc notamment, Matthieu, Jawad, Madhi, Amélie, Maud.

*Last, but not least*, je remercie très sincèrement ma famille, Laure et Maelle qui ont donné un sens une certaine tirade de Musset, Sylvie "Mama" (façon Italienne), Guillaume "frérot" et Bruno, Yves, Maryse et Frédo.

## Résumé

---

Avec l'avènement de l'informatique grand public, du réseau Internet et de la photographie numérique, de nombreuses images naturelles (acquises par un appareil photographique) circulent un peu partout dans le monde. Les images sont parfois modifiées de façon légitime ou illécite pour transmettre une information confidentielle ou secrète. Dans ce contexte, la stéganographie constitue une méthode de choix pour transmettre et dissimuler de l'information. Par conséquent, il est nécessaire de détecter la présence d'informations cachées dans des images naturelles. L'objectif de cette thèse est de développer une nouvelle approche statistique pour effectuer cette détection avec la meilleure fiabilité possible. Dans le cadre de ces travaux, le principal enjeu est la maîtrise des probabilités d'erreur de détection. Pour cela, un modèle paramétrique localement non-linéaire d'une image naturelle est développé. Ce modèle est construit à partir de la physique du système optique d'acquisition et de la scène imagée. Quand les paramètres de ce modèle sont connus, un test statistique théorique est proposé et ses propriétés d'optimalité sont établies. La difficulté principale de la construction de ce test repose sur le fait que les pixels de l'image sont toujours quantifiés. Lorsqu'aucune information sur l'image n'est disponible, il est proposé de linéariser le modèle tout en respectant la contrainte sur la probabilité de fausse alarme et en garantissant une perte d'optimalité bornée. De nombreuses expérimentations sur des images réelles ont confirmé la pertinence de cette nouvelle approche.

## Abstract

---

With the advent of mainstream computing, the Internet and digital photography, many natural images (acquired by a camera) circulate around the world. The images are sometimes altered by a legitimate or illegal information in order to transmit confidential or secret information. In this context, steganography is a method of choice to transmit and to hide information. Therefore, it is necessary to detect the presence of hidden information in natural images. The objective of this thesis is to develop a new statistical approach to perform this detection with the highest reliability possible. As part of this work, the main challenge is to control the probability of detection error. For this purpose, a parametric model locally non-linear of a natural image is developed. This model is built from the physics of optical acquisition system and from the imaged scene. When the parameters of this model are known, a statistical test is proposed and its theoretical optimality properties are established. The main difficulty in the construction of this test is based on the fact that image pixels are always quantified. When any information on the image is not available, it is proposed to linearize the model while respecting the constraint on the probability of false alarm and guaranteeing a loss of optimality bounded. Many experiments on real images have confirmed the relevance of this new approach.





# Table des matières

---

<b>1</b>	<b>INTRODUCTION GÉNÉRALE</b>	<b>1</b>
1.1	Contexte des travaux . . . . .	1
1.2	Description du problème abordé . . . . .	2
1.3	Organisation de ce manuscrit . . . . .	4
	<b>DÉTECTION STATISTIQUE DE LA PRÉSENCE D'INFORMATIONS CACHÉES DANS UNE IMAGE CONNUE</b>	<b>7</b>
<b>2</b>	<b>STÉGANOGRAPHIE ET STÉGANALYSE APPLIQUÉES AUX IMAGES NATURELLES</b>	<b>9</b>
	Introduction au chapitre . . . . .	9
2.1	Implémentations de la stéganographie moderne . . . . .	10
2.2	Méthodes usuelles d'insertion de messages . . . . .	13
2.3	La stéganalyse, détection de stéganographie : difficultés et solutions . . . . .	20
2.4	Apports de la théorie de la décision statistique à la stéganalyse . . . . .	33
	Conclusions du chapitre . . . . .	40
<b>3</b>	<b>DÉTECTION STATISTIQUE OPTIMALE DE L'INSERTION D'INFORMATION DANS LES LSB</b>	<b>41</b>
	Introduction au chapitre . . . . .	41
3.1	Décision statistique basée sur des observations quantifiées . . . . .	42
3.2	Test entre deux hypothèses simples : taux d'insertion connu . . . . .	44
3.3	Test entre deux hypothèses composites : taux d'insertion inconnu . . . . .	53
3.4	Test en présence de paramètres de nuisance : modèle régressif de l'image. . . . .	56
3.5	Détection optimale sous condition d'un faible pas de quantification . . . . .	60
3.6	Relation entre le test proposé et la méthode du WS . . . . .	65
3.7	Extension de la méthodologie proposée . . . . .	67
	Conclusions du chapitre . . . . .	71
3.8	Annexes du chapitre 3 : démonstration du théorème 3.7 . . . . .	73
	<b>ACQUISITION DES IMAGES NATURELLES ET MODÉLISATION LOCALE DU CONTENU</b>	<b>75</b>
<b>4</b>	<b>PROCESSUS DE FORMATION DES IMAGES : DESCRIPTION DE LA CHAÎNE D'ACQUISITION DES IMAGES</b>	<b>77</b>
	Introduction au chapitre . . . . .	77
4.1	Modélisation du processus de formation d'une image optique . . . . .	79
4.2	Modèle simplifié du processus optique de formation d'une image . . . . .	86
4.3	Modélisation du processus d'acquisition de l'image numérique . . . . .	88
	Conclusions du chapitre . . . . .	94
4.4	Annexes du chapitre 4 : De l'image "RAW" à l'image "TIFF" . . . . .	95
<b>5</b>	<b>VERS UNE MODÉLISATION DES IMAGES NUMÉRIQUES</b>	<b>101</b>
	Introduction au chapitre . . . . .	101
5.1	Modèle du contenu structuré des images naturelles . . . . .	102
5.2	Nécessité de la modélisation non-linéaire des discontinuités . . . . .	109
5.3	Discussions sur quelques modèles alternatifs des images . . . . .	112
	Conclusions du chapitre . . . . .	115
5.4	Annexes du chapitre 5 : démonstrations des théorèmes 5.1 et 5.2 . . . . .	116

EXPLOITATION DU MODÈLE DES IMAGES POUR LA DÉTECTION STATISTIQUE D'INFORMATIONS CACHÉES	121
6 EXPLOITATION DU MODÈLE NON-LINÉAIRE POUR LA DÉTECTION STATISTIQUE D'INFORMATIONS CACHÉES.	123
Introduction au chapitre	123
6.1 Introduction à la détection statistique en présence de paramètres non-linéaires	124
6.2 Principe de l'approche par linéarisation	127
6.3 Estimation du point de fonctionnement et biais des résidus	130
6.4 Détection statistique d'informations cachées : impact de la linéarisation sur les performances du test statistique	137
Conclusions du chapitre	140
6.5 Annexes du chapitre 5	141
7 RÉSULTATS NUMÉRIQUES	149
Introduction au chapitre	149
7.1 Discussion autour des hypothèses de travail	149
7.2 Résultats numériques sur des images naturelles : comparaisons avec l'état de l'art	151
7.3 Vérifications numériques des résultats théoriques	156
Conclusions du chapitre	161
8 SYNTHÈSE, CONCLUSION ET PERSPECTIVES OUVERTES	163
8.1 Synthèse des travaux présentés	163
8.2 Perspectives ouvertes à la suite des travaux menés	165
8.3 Conclusion	167
A ÉLÉMENTS DE LA THÉORIE DE LA DÉCISION ET DE L'ESTIMATION STATISTIQUE	169
A.1 Motivations et notations	169
A.2 Test entre des hypothèses simples	171
A.3 Test entre des hypothèses composites	174
A.4 Difficultés soulevées par la présence de paramètres de nuisance	176
A.5 Test entre deux hypothèses et classification par apprentissage	179
A.6 Éléments de la théorie de l'estimation statistique	184
B ÉLÉMENTS DE LA THÉORIE "CLASSIQUE" DE L'APPROXIMATION D'UNE FONCTION À UNE VARIABLE	189
C SÉCURITÉ D'UN STÉGOSYSTÈME	193
C.1 Critères qualitatifs de sécurité d'un stégosystème	193
RÉFÉRENCES BIBLIOGRAPHIQUES	197

# Table des figures

FIGURE 1.1	Exemple de la première image stéganographiée détectée sur Internet [1]. . . . .	2
FIGURE 2.1	Le problème des prisonniers; un scénario typique de stéganographie / stéganalyse. . . . .	10
FIGURE 2.2	Critères d'efficacité d'un stégosystème. . . . .	13
FIGURE 2.3	Illustration des différents plans de bits constituant une image en niveaux de gris. . . . .	14
FIGURE 2.4	Exemple simple d'insertion / d'extraction dans les premiers pixels d'une ligne. . . . .	15
FIGURE 2.5	Exemple de modifications engendrées par l'insertion d'informations cachées dans les LSB d'une image avec différentes méthodes. . . . .	17
FIGURE 2.6	Représentation simplifiée d'insertion par Quantification d'Index de Modulation (QIM) avec utilisation d'un quantificateur uniforme de pas $\Delta$ . . . . .	18
FIGURE 2.7	Décomposition des plans de bits d'une image en niveaux de gris avant et après stéganographie. . . . .	21
FIGURE 2.8	Diagramme état / transition illustrant les probabilités de passage d'un sous-ensemble à un autre pour chaque paire de valeurs. . . . .	25
FIGURE 2.9	Fréquence d'occurrence (en ordonnée) de valeurs (abscisse) des pixels avant et après stéganographie. . . . .	27
FIGURE 2.10	Illustration d'un test statistique d'hypothèses. . . . .	34
FIGURE 3.1	Graphique représentatif des valeurs du terme correctif $\epsilon(\theta_n; x)$ pour quelques valeurs de $\sigma$ avec $\theta = 128$ . . . . .	45
FIGURE 3.2	Représentations graphiques des moyennes $\mu_0, \mu_1$ et de leurs approximations, en négligeant la quantification, $\widehat{\mu}_0, \widehat{\mu}_1$ . . . . .	48
FIGURE 3.3	Représentations graphiques des variances $\sigma_0^2, \sigma_1^2$ et de leurs approximations, en négligeant la quantification, $\widehat{\sigma}_0^2, \widehat{\sigma}_1^2$ . . . . .	50
FIGURE 3.4	Représentations graphiques du seuil de décision $\tau_{\alpha_0}$ , et de la valeur approchée en négligeant la quantification $\widehat{\tau}_{\alpha_0}$ en fonction de $\theta_n$ . . . . .	50
FIGURE 3.5	Représentations graphiques de la fonction de puissance $\beta(\delta_R)$ , et de son approximation en négligeant la quantification $\widehat{\beta}(\delta_R)$ en fonction de $\theta_n$ . . . . .	51
FIGURE 3.6	Courbes COR illustrant les performances du test PP $\delta_R$ : représentation graphique de la fonction de puissance $\beta_{\delta_R}$ en fonction de la probabilité de fausse alarme $\alpha_0(\delta_R)$ . . . . .	52
FIGURE 3.7	Comparaison de la puissance entre le test PP $\delta_R$ (dans le cas d'hypothèses simples) et les tests LAUPP. Le taux d'insertion est ici $R = 0.05$ et les taux d'insertion considérés par les tests LAUPP sont $R^* = 0.1$ et $R^* = 1$ . . . . .	56
FIGURE 3.8	Illustration du découpage d'une image en segments $\mathbf{z}_k$ de $L = 16$ pixels et des approximations par $\widehat{\mathbf{z}}_k = \mathbf{H}\widehat{\mathbf{x}}_k$ . . . . .	57
FIGURE 3.9	Distribution des statistiques de décisions $\log(\Lambda_1(\mathbf{Z}))$ et $\log(\overline{\Lambda}(\mathbf{Z}))$ sous $\mathcal{H}_0$ et $\mathcal{H}_1$ avec $N = 1000$ , $R = 0.1$ , $\sigma_n = 2.25$ et $\theta_n = 127.5$ . . . . .	64
FIGURE 3.10	Courbes COR, $\beta(\alpha_0)$ , illustrant la puissance des tests $\delta_R$ et $\overline{\delta}$ selon la variance du bruit. . . . .	64

FIGURE 3.11	Représentations graphiques des moments (3.77) - (3.80) et de leurs approximations, en négligeant la quantification, en fonction de $\theta_n$ et pour $\sigma_n = 0.5$ . . . . .	70
FIGURE 3.12	Représentation graphique de la fonction de puissance du détecteur $\delta_R^\pm$ en fonction de la probabilité de fausse alarme $\alpha_0$ (courbes COR) pour $N = 1000$ pixels, $R = 0.1$ , $\sigma_n = 0.5$ et $\theta = \{127.5; 128\}$ . . . . .	71
FIGURE 4.1	De la scène à l'image numérique - illustration de la chaîne d'acquisition des images naturelles. . . . .	78
FIGURE 4.2	Illustration de l'effet de la PSF sur une image numérique naturelle simple. . . . .	88
FIGURE 4.3	Illustration du lobe de l'éclairage incident pour mettre en évidence le bruit quantique de comptage des photons. . . . .	91
FIGURE 4.4	Importance relative de chaque type de bruit en fonction du nombre de photo-électrons collectés dans un pixel. . . . .	93
FIGURE 4.5	Chaîne des traitements post-acquisition appliquée à une image numérique préalablement à son enregistrement par l'appareil. . . . .	95
FIGURE 4.6	Illustration de la colorisation des pixels d'une image et de la problématique du dématricage. . . . .	96
FIGURE 4.7	Illustration de la réponse non-linéaire des tubes cathodiques, de la fonction de correction $\gamma$ . . . . .	99
FIGURE 5.1	Illustration du modèle de scène local 1D pour une scène simple 2D (avant dégradation déterministes). . . . .	104
FIGURE 5.2	Illustration du modèle d'image local 1D pour une scène simple 2D (après dégradations déterministes). . . . .	105
FIGURE 5.3	Illustration du théorème 5.1 sur la décomposition d'une fonction $f$ continue par morceaux. . . . .	106
FIGURE 5.4	Illustration du théorème 5.1 : . . . . .	109
FIGURE 5.5	Valeur approchée de la meilleure erreur uniforme d'approximation $\mathcal{E}_n^*(\theta_{k,s})$ obtenue numériquement par l'algorithme de Remez. . . . .	111
FIGURE 5.6	Illustration d'un champ scalaire simple, $S(x, y)$ , continu par morceaux mais ne pouvant être décomposé tel (5.17). . . . .	118
FIGURE 6.1	Illustration en 3D de l'estimation d'un paramètre linéaire. . . . .	128
FIGURE 6.2	Illustration en 3D de la méthode proposée pour l'estimation d'un paramètre non-linéaire par linéarisation autour du point $\theta_0$ . . . . .	129
FIGURE 6.3	Illustration graphiques de la détection de contours. . . . .	132
FIGURE 6.4	Illustration graphique d'un exemple de signal à analyser (simulation), localisation des discontinuités convolution des données par $\varphi_{c_0}^{(j)}$ et par maximum de vraisemblance. . . . .	134
FIGURE 6.5	Diagramme représentant l'application pratique du test proposé. . . . .	138
FIGURE 7.1	Comparaison numérique de la performance des détecteurs : courbes COR, $\beta(\alpha_0)$ , en échelle logarithmique obtenues pour différents algorithmes avec $R = 0.05$ . . . . .	152
FIGURE 7.2	Comparaison numérique de la performance des détecteurs : courbes COR, $\beta(\alpha_0)$ , obtenues pour différents algorithmes avec $R = 0.1$ . . . . .	152
FIGURE 7.3	Distributions empiriques de la statistique de décision du test $\hat{\delta}$ proposé, du WS et du RS sous l'hypothèse $\mathcal{H}_0$ . . . . .	153
FIGURE 7.4	Espérance et écart-type du WS et du RV $\log(\hat{\Lambda})$ obtenus par simulation Monte-Carlo en utilisant des images texturées. . . . .	153
FIGURE 7.5	Comparaison numérique de la performance des détecteurs adaptés à la correspondance de LSB : courbes COR, $\beta(\alpha_0)$ pour différents algorithmes avec $R = 1$ (a) et $R = 0.5$ (b). . . . .	155
FIGURE 7.6	Comparaison numérique de la puissance des tests reposant sur différents estimateurs des paramètres $\theta$ et $\sigma$ : courbes COR obtenues avec $R = 1$ (a) et $R = 0.5$ (b). . . . .	156

FIGURE 7.7	Distributions théoriques et empiriques des LR $\bar{\delta}$ (+) et du GLR $\hat{\delta}$ ( $\Delta$ ). . . . .	157
FIGURE 7.8	Représentation graphique, sous la forme d'une courbe COR de la puissance des tests $\bar{\delta}(\mathbf{Z})$ (vert), $\bar{\delta}^{\text{TVG}}$ (bleu) et $\hat{\delta}(\mathbf{Z})$ (rouge). . . . .	157
FIGURE 7.9	Représentation graphique du RV $\hat{\Lambda}(\mathbf{z}_k)$ illustrant les résultats (6.24) - (6.28). . . . .	158
FIGURE 7.10	Vérification des résultats de la section 6.4 au travers de simulations numérique basée sur une image naturelle. . . . .	159
FIGURE A.1	Illustration d'un test statistique d'hypothèses. . . . .	169
FIGURE A.2	Courbes COR du test le plus puissant entre les deux hypothèses simples $\mathcal{H}_0 = \{\mathbf{Z} \sim \mathcal{N}(0, 1)\}$ et $\mathcal{H}_1 = \{\mathbf{Z} \sim \mathcal{N}(\bar{m}, 1)\}$ , en fonction de $\bar{m}$ (figure A.2a) et en fonction de $N$ (figure A.2b). . . . .	172
FIGURE A.3	Illustration de la fonction de puissance d'un test uniformément le plus puissant (UPP) entre les deux hypothèses composites $\mathcal{H}_0 = \{\theta < 1\}$ et $\mathcal{H}_1 = \{\theta > 1\}$ , avec évidemment $\theta \in \mathbb{R}$ . . . . .	175
FIGURE A.4	Illustration du principe de fonctionnement des méthodes de classification. . . . .	180
FIGURE A.5	Illustration d'un classifieur à vecteurs de support utilisant l'astuce du noyau : la classification linéaire dans $\mathcal{H}$ est non-linéaire dans $\mathcal{X}$ . . . . .	184

## Liste des tableaux

---

TABLE 3.3	Rappels des différents tests proposés dans cette section ainsi que leur notation et leur caractère d'optimalité éventuel. . . . .	66
TABLE 4.2	Bruits d'acquisition pris en compte dans le modèle statistique d'une image. . . . .	92
TABLE 6.2	Normes des composantes tangentielles des dérivées partielles de second-ordre. . . . .	136
TABLE 6.4	Biais des estimateurs du paramètre non linéaire $\eta$ par maximum de vraisemblance (calculé par la méthode de Newton) et par l'algorithme utilisé. . . . .	142
TABLE 6.5	Variance des estimateurs du paramètre non linéaire $\eta$ par maximum de vraisemblance (calculé par la méthode de Newton) et par l'algorithme utilisé. . . . .	142
TABLE 7.2	Temps de calcul (en secondes) nécessaire à l'exécution de l'algorithme proposé en comparaison avec le WS sur des images de $1024 \times 1024$ pixels. . . . .	154

## Glossaire des notations

De manière générale, les lettres calligraphiées  $\mathcal{A}, \mathcal{B}, \dots, \mathcal{Z}$  sont utilisées pour représenter des ensembles, les caractères gras en majuscules  $\mathbf{H}, \mathbf{F}, \mathbf{P}, \dots$  représentent des matrices, les caractères gras minuscules  $\mathbf{v}, \boldsymbol{\eta}, \boldsymbol{\zeta}, \dots$  représentent des vecteurs, enfin la police classique est utilisée pour représenter des variables quelles soient réelles,  $x \in \mathbb{R}$ , ou entières  $n \in \mathbb{Z}$ .

Notation	Signification
$\mathbf{C}$	Support (Image) de couverture dans lequel le stéganographe dissimule le message $\mathbf{M}$ .
$\mathbf{M}$	Message dissimulé par le stéganographe dans l'image de couverture $\mathbf{C}$ .
$\mathbf{S}$	Stego-objet (stégo-image) obtenu par dissimulation du message $\mathbf{M}$ dans le support $\mathbf{C}$ .
$\mathbf{Z}$	Image numérique inspectée.
$\mathbf{Y}$	Image numérique discrétisée avant quantification.
$N_x, N_y$	Nombre de colonnes, $N_x$ , et de lignes $N_y$ de l'image numérique $\mathbf{Z}$ .
$N$	Nombre total de pixels de l'image numérique $\mathbf{Z}$ , $N = N_x \times N_y$ .
$L$	Nombre de pixels dans les segments utilisés pour l'analyse de l'image $\mathbf{Z}$ .
$K$	Nombre de segments dans l'image numérique $\mathbf{Z}$ , $K \times L \approx N$ .
$\ell$	Nombre de bits d'information cachés dans l'image $\mathbf{Z}$ .
$R$	Taux d'insertion des informations cachées, $R = \ell/N$ .
$\mathbf{z}$	Vecteur représentant l'image $\mathbf{Z}$ parcourue dans le sens lexicographique.
$\mathbf{y}$	Vecteur représentant l'image $\mathbf{Y}$ (avant discrétisation) parcourue dans le sens lexicographique.
$\theta_n$	Espérance mathématique du $n$ -ième pixels de l'image.
$\boldsymbol{\theta}$	Espérance mathématique de l'ensemble des pixels de l'image.
$\hat{\boldsymbol{\theta}}$	Estimation statistique l'ensemble des pixels de l'image.
$\Theta$	Espace des paramètres $\boldsymbol{\theta} \in \Theta$ .
$\xi$	Réalisation de la variable aléatoire $\Xi$ .
$\Xi$	Variable aléatoire modélisant les phénomènes stochastiques intervenant dans le processus d'acquisition.
$\mathcal{Y}$	Espace des pixels discrétisés avant numérisation, $\mathbf{Y} \in \mathcal{Y} \subset \mathbb{R}^{MN}$ .
$\mathcal{Z}$	Espace des pixels numériques (après discrétisation) $\mathcal{Z} = \mathbb{F}_{2^b}$ .
$\mathcal{I}$	Ensemble discret $\mathcal{I} \subset \mathbb{Z}^2$ définissant les pixels de l'image.
$P_{\boldsymbol{\theta}}$	Loi de distribution paramétrée par le paramètre vectoriel $\boldsymbol{\theta}$ .
$P_{\theta_n}$	Loi de probabilité du $n$ -ième pixel pour une image "saine".
$Q_{\theta_n}^R$	Loi de probabilité du $n$ -ième pixel pour une image contenant des informations cachées avec un taux $R$ par substitution de LSB.

Notation	Signification
$\mathcal{H}_0, \mathcal{H}_1$	Hypothèse nulle et hypothèse alternative (respectivement média sain et média stéganographié).
$\alpha_\delta$	Probabilité, ou taux, de fausse alarme du test $\delta$ .
$\alpha_0$	Contrainte fixée par l'utilisateur sur la probabilité de fausse alarme.
$\mathcal{K}_{\alpha_0}$	Classe des test dont la probabilité de fausse alarme est bornée par $\alpha_0$
$\varphi(\cdot)$	Densité de probabilité de la loi de distribution normale (Gaussienne) centrée réduite $\mathcal{N}(0, 1)$ .
$\Phi(\cdot)$	Fonction de répartition de la loi de distribution normale (Gaussienne) centrée réduite $\mathcal{N}(0, 1)$ .
$S(x, y, \lambda, \omega, \phi)$	Luminance rayonnée par la scène à la longueur d'onde $\lambda$ dans la direction données par les angles $\omega, \phi$ .
$S(x, y)$	Luminance rayonnée par la scène imagée dans l'angle solide sous-tendu par le système optique et sur l'ensemble du spectre de longueur d'onde.
$S_k(x, y)$	Luminance rayonnée par la scène sur la droite $y = y_k$ .
$E(x, y)$	Éclairement incident sur le plan du photo-détecteur.
$\theta(x, y)$	Signal "théorique" enregistré par le système d'acquisition, $\theta(x, y) \propto E(x, y)$ .
$\theta_k(x)$	Signal enregistré par le système d'acquisition sur la droite $y = y_k$ .
$\theta_{k,s}(x)$	Partie représentant les discontinuités présente dans le signal $\theta_k(x)$ .
$\theta_{k,c}(x)$	Partie continue du signal $\theta_k(x)$ .
$\eta_k$	Paramètres décrivant les éventuelles discontinuités présentes dans le $k$ -ième segment $\mathbf{z}_k$ .
$\nu_{k,i}$	Localisation de la $i$ -ième discontinuité présente dans le $k$ -ième segment.
$\zeta_{k,i}$	Paramètre local du noyau de convolution (PSF) autour de la discontinuité de localisation $\nu_{k,i}$ .
PP	Plus Puissant (test le -)
AUPP	Uniformément le Plus Puissant.
AUPP	Asymptotiquement Uniformément le Plus Puissant.
LAUPP	Localement Asymptotiquement Uniformément le Plus Puissant.
RV	Rapport de Vraisemblance.
RVG	Rapport de Vraisemblance Généralisé
MV	Maximum de Vraisemblance.
LSB	<i>Least Significant Bits</i> (bits de poids faible).
PSF	<i>Point Spread Function</i> (fonction d'étalement du point).
1D	Une dimension.
2D	Deux dimensions.
CCD	Charged Coupled Device





---

# Introduction générale

---



«*There are those who look at things the way they are, and ask why...  
I dream of things that never were, and ask why not ?* ».

— George Bernard Shaw, 1856–1950

«*L'élégance, la science, la violence !* ».

— Arthur Rimbaud, 1854–1891

Comme Régis Lengellé, directeur de l'école doctorale SSTO rattachée à l'UTT, l'a souvent répété, une thèse est une mayonnaise dont les ingrédients sont un (des) directeur(s) de thèse, un doctorant et un sujet. Dans le présent cas, cette allégorie est assez révélatrice du déroulement de cette thèse et, à titre personnel, mon plus grand souhait est que le lecteur parcourant ces pages ait l'impression que cette mayonnaise «ait pris». Ce souhait est intrinsèquement lié au contenu quelque peu hétéroclite des travaux présentés ; positionnés entre la dissimulation d'informations, la modélisation des images naturelles et la détection statistique.

## 1.1 Contexte des travaux

La stéganographie est la science de la communication cachée ou dissimulée. Comme le laisse supposer l'étymologie grecque du mot stéganographie, "*stego*" le secret et "*graphia*" l'écriture, les origines de cette science remontent à l'antiquité. En effet, on attribue souvent le premier exemple d'utilisation de la stéganographie à Hérodote (480- 420 av. J.-C.). Dans son unique ouvrage, *Enquêtes*, Hérodote relate comment Démarate pu prévenir les grecs d'une invasion perse en faisant parvenir un message gravé dans une tablette en bois, ensuite recouverte d'une cire d'apparence vierge [2]. Plus technique mais presque aussi ancienne, l'invention de l'encre sympathique, encore utilisée de nos jours, est attribuée à Pline l'Ancien au premier siècle [3].

Depuis l'antiquité, la stéganographie n'a cessé d'évoluer. Parmi les quelques ouvrages publiés sur le sujet, citons les plus marquants tels que celui de l'abbé Jean Trithème *Steganographia* (1499), qui ne livrera ses derniers secrets qu'en 1998 [4, 5], ainsi que *Schola steganographica* (1665) de Gaspart Schott dans lequel l'auteur explique notamment comment dissimuler un texte dans une partition musicale.

Les méthodes de dissimulation d'informations se sont complexifiées au fil des évolutions techniques. L'invention du micro-film a, par exemple, ouvert la porte à une utilisation de la dissimulation d'information lors la seconde guerre mondiale : ainsi réduit, un document (photographie ou texte) pouvait alors être intégré comme un «●» dans un texte anodin en apparence [6].

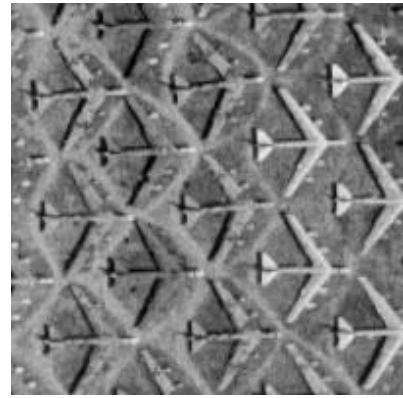
Enfin, un dernier exemple historique fort intéressant est celui de Margaret Thatcher qui, dans les années 1980, alors qu'elle était Premier ministre du Royaume-Uni, a vu des documents de son propre cabinet être rendus publiques par voie de presse. Il semblerait alors que Margaret Thatcher décida de faire programmer les outils de traitements de textes de sorte que l'espacement des caractères soit imperceptiblement modifié afin de révéler l'identité de l'auteur [7, §1]. Dans cet exemple (sujet à caution), la dissimulation de l'information a été, en un certain sens, détournée. Le but n'étant plus alors de cacher une information, mais d'identifier la machine à l'origine de la fuite ; on parle alors de "tatouage" ou "d'empreinte numérique", concept brièvement abordé dans la

*Steganographia de Jean Trithème est (en partie) consultable à l'adresse suivante : <http://www.apprendre-en-ligne.net/crypto/bibliotheque/tritheim/stegano.htm>*

*Les premières pages de l'ouvrage Schola steganographica sont consultables sur le site de Fabien Petitcolas : <http://www.petitcolas.net/fabien/steganography/history.html>*



(a) Image stéganographiée



(b) Information dissimulée, puis extraite

FIGURE 1.1: Exemple de la première image stéganographiée détectée sur Internet [1].

section 2.1.1.

Le renouveau de l'étude de la stéganographie a fait irruption dans le domaine de la recherche académique dans le milieu des années 1990, avec une première thèse [8] soutenue sur le sujet en 1997 et la tenue en 1996 à Cambridge de la première édition de la conférence *Information Hiding* ("dissimulation d'information" en français). La création de journaux, dont la ligne éditoriale est bâtie autour de la criminalistique et de la sécurité des informations, chez les principaux éditeurs a finalement permis à cette thématique d'acquiescer ses lettres de noblesse ; citons notamment la première publication de *l'International Journal of Information Security* (Springer) en août 2001, de *IEEE transaction on Information Forensics and Security* en mars 2006 et de *EURASIP Journal on Information Security* en mars 2011.

## 1.2 Description du problème abordé

Il est simple d'imaginer des usages légitimes de la stéganographie, comme des usages détournés à des fins malveillantes. Dans les deux cas, l'usage de la stéganographie comme une technique sécurisée de transmission d'informations est sujette aux deux questions suivantes :

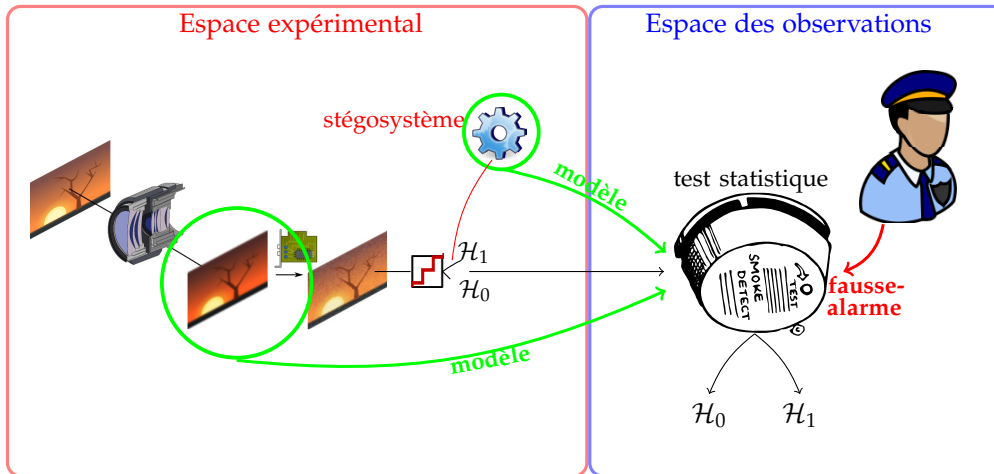
1. Disposant d'un média  $Z$  donné, quelle est la taille maximale du message qu'il est possible d'insérer tout en assurant la sécurité de la transmission ?
2. Inversement, supposant qu'une image  $Z$  est analysée, quelle est la performance attendue du test utilisé ; comment prendre une décision tout en respectant des critères d'erreur préalablement fixés ?

C'est à ce type de questions qu'il est proposé d'apporter une première réponse au travers des travaux présentés. En effet, pour répondre aux deux questions précédentes, il est nécessaire de pouvoir caractériser statistiquement la performance du test statistique proposé. Or, parmi les nombreuses méthodes qui ont été proposées pour détecter d'éventuelles informations cachées dans un média numérique, il est notable que l'usage de la théorie de la détection statistique est resté anecdotique.

Illustration graphique de la problématique abordée : procédure d'analyse d'une image depuis son acquisition jusqu'à la prise de décision statistique.

Le problème abordé dans ce manuscrit est présenté à travers un exemple de la procédure typique de stéganographie (et de stéganalyse) d'une photographie numérique. Le déroulement de la procédure de stéganographie et de stéganalyse d'une image

*Les termes de sécurité de la transmission et de performance du test sont volontairement non-définis.*



naturelle, illustré par la figure 1.2, est typiquement résumé par les quatre étapes suivantes :

1. Le système optique de l'appareil photographique permet la formation d'une image  $I$  à partir de la scène imagée  $S$ , décrite par ses propriétés radiométriques.
2. Le système numérique d'acquisition enregistre l'image  $I$  pour donner naissance à l'image numérique de couverture  $C$ .
3. L'image numérique de couverture  $C$  est utilisée par le stégo-système pour produire une stégo-image  $S$  contenant le message caché  $M$ .
4. Enfin, le stéganalyste doit décider à partir de la seule image  $Z$  qu'il inspecte si cette dernière contient une information cachée ou non.

Cette présentation simpliste de la procédure de stéganographie et de stéganalyse décrit schématiquement le problème général de toute méthode de stéganalyse d'une image. Plus précisément, cette thèse s'inscrit dans le cadre du projet ANR RIC (Recherche d'Informations Cachées) dont la finalité applicative est, en collaboration avec des partenaires étatique, le développement d'une méthodologie permettant d'analyser un grand nombre de fichiers images dans un cadre policier et/ou judiciaire. Dans un tel contexte opérationnel, compte-tenu des enjeux et du nombre potentiellement très important de fichiers "suspects", deux contraintes fondamentales ont été fixées dès le début des travaux :

1. le respect d'une contrainte sur la probabilité de fausse alarme est nécessaire, décider qu'une image contient un message caché peut avoir des conséquences importantes ;
2. le temps d'analyse d'une image doit être aussi faible que possible, la législation n'autorisant pas de disposer des données d'une personne pendant une durée illimitée.

Aucun des travaux de recherches publiés dans la littérature ne prend en compte une de ces deux contraintes. La prise en compte de ces contraintes a donc amené à développer une nouvelle méthodologie dont les deux principales originalités sont les suivantes :

- d'une part, afin de respecter une contrainte sur la probabilité de fausse alarme l'utilisation de la théorie de la décision statistique semble incontournable ; cet aspect est directement lié aux contraintes du projet ;
- d'autre part, l'image numérique  $Z$  inspectée étant généralement issue d'un appareil photographique, il est proposé d'exploiter certaines propriétés de ce type "d'image naturelle" lors de la prise de décision.

### 1.3 Organisation de ce manuscrit

Outre le chapitre 2 présentant brièvement les principales techniques de stéganographie et un état de l'art critique des méthodes de détection d'informations cachées, le présent mémoire est constitué des quatre principaux chapitres suivants :

- Le chapitre 3 est dévolu à la formalisation statistique du problème de la détection d'informations cachées. Les pixels constituant une image sont modélisés comme la réalisation de variables aléatoires indépendantes et distribuées suivant une loi de distribution dont les paramètres sont connus (pour tous les pixels); en ce sens, cette approche statistique est donc davantage théorique. Dans le cas où en outre, le nombre de bits dissimulés est supposé connu, l'approche de Neyman-Pearson est utilisée pour construire le test statistique le plus puissant respectant une contrainte sur la probabilité de fausse alarme. L'expression de la puissance asymptotique du test est établie et met en évidence l'impact de la quantification des données sur la performance du test. Dans le cas où le nombre moyen de bits dissimulés n'est pas connu, le problème est celui d'un test statistique entre hypothèses composites. En utilisant la normalité asymptotique locale de L. Le Cam, un test localement asymptotiquement le plus puissant (LAUPP) est proposé. Enfin, sous des conditions de quantification avec un pas faible, un test statistique maximisant la puissance de détection, uniformément par rapport au nombre de bits insérés (AUPP) est proposé.
- L'exploitation pratique des tests statistiques proposés dans le chapitre 3 est sujette à la connaissance des paramètres de distribution de chacun des pixels; or, l'estimation de ces paramètres n'est possible que si des informations sur ces derniers sont disponibles. L'approche proposée dans le chapitre 4 consiste à modéliser ces paramètres en analysant et en modélisant le processus physique permettant l'acquisition d'une image numérique, à partir de la scène imagée. La chaîne d'acquisition est conceptuellement considérée et modélisée sous deux aspects distincts. D'une part, le système optique d'acquisition permettant la formation d'une image sur le photo-détecteur est modélisée. Les éléments optiques intervenant de façon déterministe, une relation analytique entre la scène imagée et l'image optique est établie. D'autre part, différents phénomènes stochastiques interviennent lors de l'acquisition de l'image numérique. En raison de la nature aléatoire de ces phénomènes de dégradation d'origine électronique, une modélisation statistique de la valeur des pixels est proposée.  
Il ressort finalement de cette modélisation que le système d'acquisition d'une image naturelle présente deux particularités que sont la non-stationnarité du système optique, ce dernier n'est notamment pas invariant par translation, et la non-stationnarité du bruit d'acquisition, la variance de ce dernier n'est pas constante sur l'ensemble des pixels d'une image.
- Afin de permettre une modélisation des images naturelles fondée sur les phénomènes intervenant lors du processus d'acquisition, une analyse des propriétés physiques de la scène est nécessaire. Le chapitre 5 propose donc une modélisation simple et générale de la scène imagée : la scène est physiquement décrite comme la juxtaposition de différents objets au sein desquels les propriétés radiométriques varient peu. L'utilisation conjointe du modèle d'une scène et du processus d'acquisition d'une image permet l'obtention d'un modèle local d'une image numérique capturée par un appareil photographique. En raison de son fondement physique, ce modèle présente l'originalité d'être précisément adapté pour décrire la structure particulière des images. En outre, ce modèle présente l'intérêt de décrire

localement le contenu d'une image en utilisant un nombre limité de paramètres (qui sont en outre dotés d'une signification physique). Malheureusement, ce modèle local présente l'inconvénient majeur d'être non-linéaire vis-à-vis de certains paramètres décrivant physiquement les discontinuités séparant les objets constituant la scène imagée.

- Enfin, il est proposé dans le chapitre 6 une exploitation du modèle non-linéaire du contenu des images pour la détection statistique d'informations cachées. Compte-tenu du contexte opérationnel dans lequel s'inscrivent ces travaux, une attention particulière est portée, d'une part, à la simplicité calculatoire de l'algorithme proposé et, d'autre part, à la description statistique des performances du test, en terme de probabilités des erreurs notamment. Afin de permettre une estimation des paramètres de distribution des pixels avec un coût calculatoire réduit, une approche par linéarisation du modèle est proposée. En utilisant cette approche, les performances statistiques du test proposé sont établies, ce qui permet le respect d'une contrainte sur la probabilité de fausse alarme. Il est en outre souligné que cette procédure d'estimation préalable engendre une perte de puissance bornée, comparativement aux tests optimaux proposés dans le cas où les paramètres de distribution sont connus.

Les travaux présentés se veulent avoir une finalité applicative, aussi quelques résultats numériques validant les résultats établis sont présentés et discutés dans le chapitre 7.



Première partie

DÉTECTION STATISTIQUE DE LA PRÉSENCE  
D'INFORMATIONS CACHÉES DANS UNE IMAGE  
CONNUE





---

# Stéganographie et stéganalyse appliquées aux images naturelles

---

# 2

«*If I have seen further it is only by standing on the shoulders of giants.*»  
— Isaac Newton, 1643 – 1727

«*Ce sont rarement les réponses qui apportent la vérité mais l'enchaînement des questions.*»  
— Daniel Pennac, 1954–..., La fée carabine, 1987

## Sommaire

---

Introduction au chapitre . . . . .	9
2.1 Implémentations de la stéganographie moderne . . . . .	10
2.1.1 Stéganographie dans les images . . . . .	10
2.2 Méthodes usuelles d'insertion de messages . . . . .	13
2.2.1 Stéganographie dans le domaine spatial . . . . .	13
2.2.2 Insertion d'informations dans un domaine transformé . . . . .	19
2.3 La stéganalyse, détection de stéganographie : difficultés et solutions . . . . .	20
2.3.1 Stéganalyses universelles ou aveugles . . . . .	22
2.3.2 Stéganalyses adaptées à un stégosystème . . . . .	24
2.3.3 Stéganalyses adaptées à la stéganographie par substitution de LSB . . . . .	24
2.3.4 Stéganalyses adaptées à d'autres méthodes de stéganographie . . . . .	32
2.4 Apports de la théorie de la décision statistique à la stéganalyse . . . . .	33
2.4.1 Motivations et notations . . . . .	33
2.4.2 Approche proposée et critère d'optimalité utilisé . . . . .	34
2.4.3 Détection en présence de paramètres de nuisance . . . . .	37
2.4.4 Apprentissage supervisé pour la décision statistique . . . . .	38
2.4.5 Applications à la stéganographie . . . . .	38
Conclusions du chapitre . . . . .	40

---

### Introduction au chapitre :

Ce second chapitre se veut être une introduction à la dissimulation d'informations adaptée aux travaux présentés ; aussi, les enjeux de la stéganographie et de la stéganalyse sont présentés dans la section 2.1. Sans prétention à l'exhaustivité, les principales méthodes utilisées pour l'insertion d'informations cachées dans un média numérique sont présentées dans la section 2.2. Compte-tenu du problème abordé dans les présents travaux, la section 2.3 dresse un état de l'art des principales méthodes de détection d'informations cachées proposées dans la littérature. L'objectif de cet état de l'art critique est principalement de pointer l'exploitation très limitée de la théorie de la décision statistique. L'intérêt de cette approche est ensuite présenté dans la section 2.4 en portant une attention toute particulière aux difficultés que présente l'application envisagée.

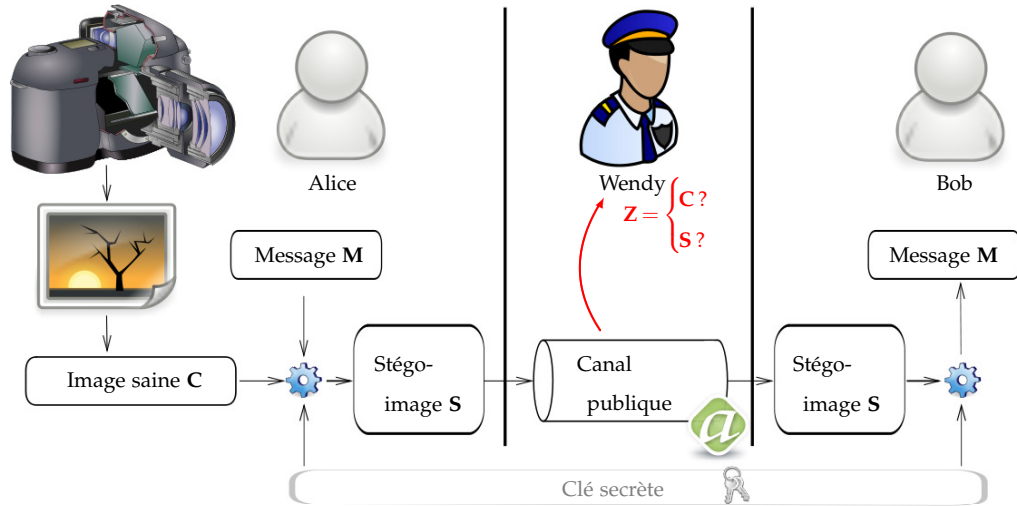


FIGURE 2.1: Les problème des prisonniers ; un scénario typique de stéganographie / stéganalyse.

## 2.1 Implémentations de la stéganographie moderne

La formulation de la stéganographie moderne (*i.e.* exploitant des supports d'informations numériques) est souvent associée au problème du prisonnier [9], illustré par la figure 2.1 : Alice et Bob sont deux prisonniers qui souhaitent échanger des informations pour préparer leur évasion sans que leurs communications ne soient interceptées par la gardienne, Wendy. Plus précisément, Alice souhaite envoyer un message (binaire) de  $L$  bits  $M = \{0, 1\}^L$  à Bob. Pour dissimuler ce message, elle l'incorpore dans un média de couverture  $C$  pour obtenir le stégo-objet  $S$  qui est transmis à Bob via un canal de communication public. D'un point de vue très général, la technique d'incorporation du message caché est inconnue de Wendy. D'un point de vue plus pratique, la technique peut être connue par Wendy mais l'incorporation dépend d'une clé de stéganographie partagée uniquement par Alice et Bob. Cette clé peut par exemple être un mot de passe utilisé pour générer une séquence numérique pseudo-aléatoire servant à sélectionner la position des pixels où sera caché le message secret  $M$ . Dans le cas d'une stéganographie à clé publique, Alice incorpore le message  $M$  à l'aide d'une clé publique connue par tous mais, seul Bob possède la clé privée permettant d'extraire le message. Du côté de la gardienne, Wendy est libre d'examiner les messages de manière active ou passive. Une inspection passive consiste simplement à inspecter le média  $Z$  pour décider s'il contient ou non un message secret  $M$ . Lorsque Wendy détecte un message caché, elle peut alors intervenir par exemple en coupant la communication. Une inspection active consiste à altérer suffisamment, par des moyens adaptés, les objets inspectés  $Z$  afin de rendre impossible la lecture d'un éventuel message caché  $M$ .

Dans la suite de ce manuscrit, les notations  $C$  et  $S$  sont utilisées lorsque la nature du média doit être explicitement précisée, média de couverture ou stégo-média. Lorsque la nature du média n'a pas d'intérêt ou que cette dernière n'est pas connue, la notation  $Z$  est utilisée.

### 2.1.1 Stéganographie dans les images

La stéganographie est potentiellement applicable à de nombreux supports numériques différents [10], allant du protocole de communication au système de gestion de

*La nature du message échangé entre Alice et Bob est quelconque, il peut s'agir de texte, d'image, de son, ou d'un programme exécutable par exemple.*

fichiers en passant par différents média numériques : vidéo, audio, textes, etc. ... Une grande majorité des stégosystèmes sont spécifiquement dédiés à l'insertion d'informations cachées dans les fichiers d'images. Par ailleurs, ce type de fichier étant très couramment échangé sur Internet, la plupart des travaux de recherches se focalisent également sur les images numériques.

Afin de rester le plus général possible, un média numérique est, dans un premier temps, défini comme un vecteur d'entiers naturels (la nature des observations sera explicitement précisée lorsqu'il ne s'agit pas des entiers naturels). Plus formellement un média numérique  $\mathbf{Z}$  est constitué de  $N$  échantillons. Afin de simplifier les notations, il est considéré que les échantillons sont indexés par un indice  $n \in \mathcal{I}$  avec  $\mathcal{I} = \{1, \dots, N\}$ . En notant  $\mathbb{F}_p$  l'anneau fini des congruences modulo  $p$  (parfois aussi noté  $\mathbb{Z}/p\mathbb{Z}$ ), chacun de ces échantillons,  $z_n$ , appartient à l'ensemble fini  $\mathcal{Z} = \mathbb{F}_p = \{0; \dots; p-1\}$ . Ainsi un média numérique appartient, en toute généralité, à l'ensemble fini  $\mathcal{Z}^N$ .

Dans le cas d'une image en niveaux de gris, les échantillons sont des pixels et forment un matrice de taille  $N_y \times N_x$ , constituée de  $N = N_x N_y$  pixels. Dans la plupart des cas, les pixels sont notés  $z_{m,n}$  avec  $(m, n) \in \mathcal{I} = \{1, \dots, N_y\} \times \{1, \dots, N_x\}$  mais, encore une fois dans un souci de clarté, lorsque la structure de la matrice d'une image n'a pas d'importance les pixels sont notés  $z_n$  où  $n \in \mathcal{I}$  désignera un des élément de  $\mathcal{I}$ . Pour tout  $n \in \mathcal{I}$ , les pixels  $z_n$  sont, sauf précision contraire, des entiers naturels codés sur  $b$  bits, aussi  $z_n \in \mathcal{Z} = \mathbb{F}_{2^b}$ .

De manière analogue, une image en couleurs est représentée par  $N_c$  matrices (semblables à des images en niveaux de gris) correspondant chacune à un canal de couleur distinct; en général les trois canaux de couleurs utilisés sont le rouge, le vert et le bleu. Par extension, une image couleur est donc un élément  $\mathbf{Z} \in \mathbb{F}_{2^b}^N$  avec désormais  $N = N_c N_x N_y$ .

Enfin, l'information cachée est constituée de  $L$  bits; le message  $\mathbf{M}$  est donc en toute généralité considéré comme un élément de  $\mathbb{F}_2^L = \{0; 1\}^L$ . En pratique, le taux d'insertion est préféré pour quantifier le nombre de bits inséré, ce taux noté  $R$  correspond au rapport de  $L$  par le nombre maximal de bits qu'il est possible d'insérer dans un média de couverture  $\mathbf{C}$ .

Le terme "stégosystème" désigne l'ensemble des algorithmes nécessaires pour produire un stégo-objet et en extraire le message secret. La définition générale d'un stégosystème utilisée dans ce manuscrit, largement inspirée de la définition proposée par C. Cachin [11], est la suivante :

**Définition 2.1** (stégosystème à clé privée). *Un stégosystème est un ensemble de deux algorithmes  $(E, D)$  avec les fonctionnalités suivantes :*

- L'algorithme d'encodage  $E$  permet de dissimuler, à partir d'une clé  $K$ , le message  $M$  dans l'image de couverture  $\mathbf{C}$  pour obtenir la stégo-image  $\mathbf{S} = E(\mathbf{M}, \mathbf{C}, K)$ .
- L'algorithme de décodage  $D$  permet de fournir, à partir d'une clé  $K$  et de la stégo-image  $S$ , le message  $\hat{\mathbf{M}} = D(\mathbf{S}, K)$ .

Cette définition fonctionnelle peut être complétée par les deux éléments suivants :

- Pour des raisons de simplicité, un tierce algorithme  $G$  est souvent utilisé pour générer la clé  $K$  en fonction d'un mot de passe  $pwd$  :  $K = G(pwd)$ . La clé utilisée est alors typiquement une séquence binaire pseudo-aléatoire d'une longueur fixe de  $l_k$  bits :  $K = \mathbb{F}_2^{l_k}$ .
- Il semble par ailleurs important d'exiger de la part du stégosystème que si l'image n'est pas modifiée entre l'émetteur et le récepteur, le message extrait soit identique à celui inséré. Autrement dit, il est souhaité que le stégosystème vérifie la propriété suivante :  $\forall K \in \mathbb{F}_2^{l_k}, \forall \mathbf{M} \in \mathbb{F}_2^L, \forall \mathbf{C} \in \mathbb{F}_{2^b}^N, D(E(\mathbf{M}, \mathbf{C}, K), K) = \mathbf{M}$ . Bien que nullement nécessaire, cette propriété assure aux utilisateurs la possibilité d'extraire le message pour peu que l'intégrité de l'image reçue par Bob soit assurée.

Cette définition fonctionnelle d'un stégosystème reste très générale mais elle met en évidence les éléments principaux d'un stégosystème. La suite de cet état de l'art

L'article La cryptographie militaire de A. Kerckhoffs est consultable sur le site de Fabien Petitcolas : <http://www.petitcolas.net/fabien/kerckhoffs/>

Les stégosystèmes à clé publique ont déjà été abordés de façon simpliste dans l'un des premiers articles de Anderson et Petitcolas [13] en utilisant la cryptographie à clé publique.

sur la stéganographie se focalise sur le point crucial du stégosystème : l'algorithme de l'encodeur.

Soulignons qu'un stégosystème efficace n'a, dans la pratique, nullement besoin d'une clé  $K$ , mais, cela reviendrait alors à faire reposer la sécurité du stégosystème sur l'obscurité de son fonctionnement. L'introduction de la clé  $K$  est une réponse au principe de Kerckhoffs [12] stipulant que la sécurité ne doit pas tenir dans le fonctionnement du système mais dans la clé uniquement afin «*qu'il puisse sans inconvénient tomber entre les mains de l'ennemi*».

Enfin, notons que des stégosystèmes asymétriques / à clé publique [14] ont été proposés mais cette considération sort légèrement de la problématique de détection statistique d'information cachée, abordée dans ce manuscrit.

La stéganographie n'est pas la seule discipline dans le domaine, plus large, de la dissimulation d'information. Plus précisément, la stéganographie, telle que décrite dans la définition 2.1, peut avoir d'autres objectifs que l'établissement d'une communication dissimulée. Les autres applications de la dissimulation d'informations sont notamment :

- Le tatouage numérique, ou (*digital watermarking*), dont le but principal est l'insertion d'une marque (tatouage) servant à identifier l'auteur du média et donc, de lui assurer la paternité des documents tatoués [15, 16].
- Le tatouage fragile, ou (*fragil watermarking*), permettant de contrôler l'intégrité ou l'authenticité d'un document en y dissimulant un tatouage qui ne pourra être extrait en cas de modification (même mineure) du document [16].
- L'empreinte numérique, ou (*fingerprinting*), est une utilisation à des fins davantage commerciales, dont le but est d'insérer un identifiant unique d'utilisateur dans un document ; la finalité est alors de pouvoir identifier individuellement une source [17], lors d'une divulgation frauduleuse par exemple.

Bien que similaires dans l'approche, les multiples techniques de dissimulation d'informations diffèrent par leurs objectifs (les attaques envisageables n'étant pas identiques) et donc par les méthodes mises en œuvre. Imaginons, pour l'exemple, un photographe cherchant à insérer un tatouage pour préserver ses droits d'auteur. L'information qu'il souhaiterait dissimuler devrait pouvoir être extraite même après que l'image ait été modifiée (afin de se prémunir d'une utilisation de ses photographies légèrement retouchées). Par contre, la taille du message à insérer est faible, théoriquement un bit d'information est suffisant pour représenter la présence ou non du tatouage.

Dans ce manuscrit, la stéganographie est uniquement étudiée comme l'utilisation des méthodes de dissimulation d'informations en vue d'établir une communication entre l'émetteur et le(s) récepteur(s). L'objectif est alors double : il s'agit d'une part de pouvoir insérer une grande quantité d'informations ; d'autre part, puisqu'il est possible qu'une tierce personne essaie de détecter ces communications, cf. figure 2.1, il est important de minimiser l'impact de l'insertion d'informations. Ce double objectif permet de formaliser les critères usuels quantifiant la qualité d'un stégosystème [6, 18]. Ces critères sont les suivants :

- La capacité de dissimulation mesure la quantité d'information qu'il est possible de dissimuler dans une image. Cette capacité se mesure en général comme le nombre total de bits d'information qu'il est possible d'insérer par pixel de l'image.
- La discrétion (ou l'invisibilité) d'un stégosystème quantifie la capacité à mettre en échec les méthodes de stéganalyse passive. La discrétion est assez délicate à mesurer, comme il est décrit dans la section C.1 de l'annexe C, plusieurs approches amènent à définir plusieurs critères qualitatifs de la sécurité d'un stégosystème.
- Enfin, la robustesse quantifie la capacité que possède un stégosystème à résister aux méthodes de stéganalyse active ; autrement dit de pouvoir extraire l'information d'un média dont le contenu a été (légèrement) modifié. Cette robustesse est en général mesurée vis-à-vis d'un type particulier de modification (rotation, filtrage, réduction/augmentation de la taille, compression, etc. ...).

Comme l'illustre la figure 2.2, la qualité d'un stégosystème se mesure différemment suivant le type de dissimulation d'informations. La suite de ce manuscrit est consacrée à la stéganographie; le lecteur intéressé par les autres disciplines de la dissimulation d'informations pourra se référer aux ouvrages [16, 19].

## 2.2 Méthodes usuelles d'insertion de messages

Les travaux de recherches présentés dans ce manuscrit se focalisent sur les images naturelles, c'est à dire acquises à l'aide d'un appareil d'enregistrement type appareil photographique ou scanner. Aussi, seules les images matricielles sont considérées (voir section 2.1.1).

Nous distinguerons deux familles de schémas stéganographiques suivant la nature du fichier image considéré. En effet, la compression du média de couverture  $C$  (souvent utilisée pour la sauvegarde des images) doit être prise en compte dans la mise au point d'un stégosystème pour plusieurs raisons. D'une part, les méthodes les plus courantes de compression sont dites avec perte. L'information originale (valeur des pixels de l'image dans notre cas) n'est plus accessible après compression; il est alors fort peu probable que le message inséré, dans une image non compressée, demeure lisible après l'étape de compression avec perte. D'autre part, il convient de noter que les méthodes de compressions réduisent, en général, considérablement l'entropie des données (via l'utilisation d'une transformation adaptée au support à compresser). Or, cette réduction d'entropie réduit également la capacité d'insertion d'informations cachées. Enfin, les informations compressées possèdent des propriétés statistiques particulières, le fait de ne pas en tenir compte peut rendre le stégosystème aisément détectable.

L'exemple de ces quelques facteurs montre que le fait de travailler, ou pas, sur des images préalablement compressées peut modifier considérablement la conception d'un stégosystème qui doit impérativement tenir compte de la spécificité du format d'image auquel il est dédié.

*Par opposition les images numériques de type vectoriel ne sont pas considérées dans la suite de ce manuscrit.*

*Dans le cas des images JPEG par exemple, la distribution des coefficients DCT est usuellement modélisée par une loi Laplacienne discrétisée [20] qui peut être exploitée pour la détection d'informations cachées [21].*

### 2.2.1 Stéganographie dans le domaine spatial

Comme nous l'avons déjà précisé, pour chaque canal de couleur la valeur de chacun des pixels est représentée numériquement par un entier non-signé (positif ou nul) codé sur  $b$  bits. En adoptant la notation de l'image numérique  $\mathbf{Z} = \{z_n\}_{n=1}^N$  avec pour tout  $n = \{1, \dots, N\}$ ,  $z_n \in \mathbb{F}_{2^b} = \{0, \dots, 2^b - 1\}$ , la représentation binaire de  $z_n$  est donnée par :

$$z_n = \sum_{i=0}^{b-1} b_{n,i} 2^i \quad (2.1)$$

où  $b_{n,i} \in \{0, 1\}$  représente le  $i$ -ème bit codant le  $n$ -ième pixel. On constatera que les bits n'ont pas tous la même importance dans le codage de la valeur  $z_n$ ; en effet le premier bit  $b_{n,0}$  est pondéré par  $2^0 = 1$  alors que le dernier bit  $b_{n,b-1}$  est pondéré par

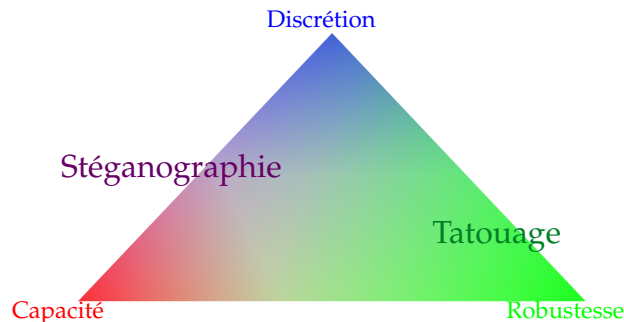


FIGURE 2.2: Critères d'efficacité d'un stégosystème.

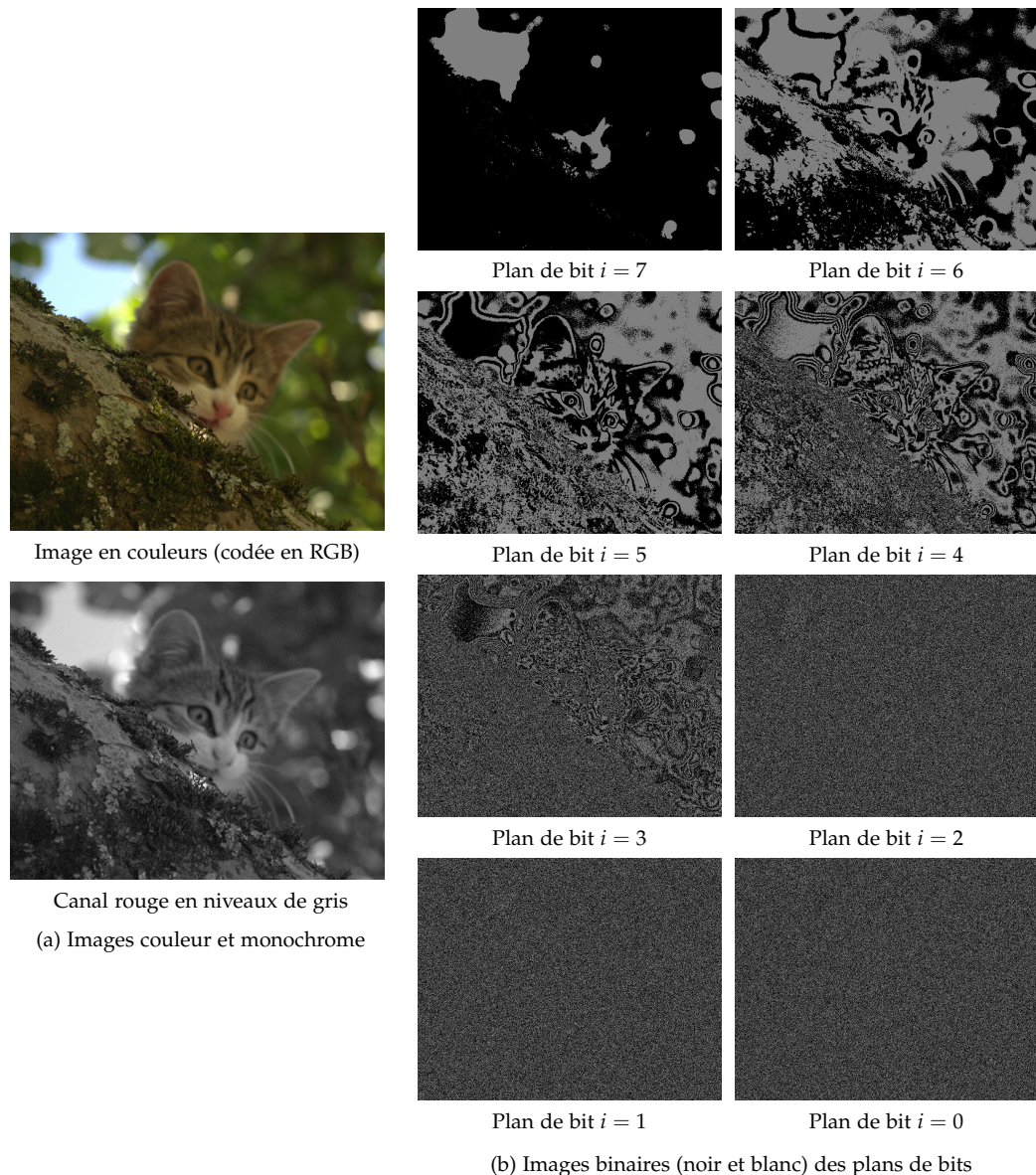


FIGURE 2.3: Illustration des différents plans de bits constituant une image en niveaux de gris.

$2^{b-1}$ . Cette observation est à la base des deux premiers stégosystèmes présentés par la suite.

**Remarque 2.1.** La très grande majorité des méthodes d'insertion considère l'image comme une succession d'échantillons (les pixels de chaque canal de couleur). Le message à insérer est décomposé en bits,  $\mathbf{M} \in \mathbb{F}_2^L = \{0; 1\}^L$  (ou plus rarement en symboles  $q$ -aires), puis ces bits sont ensuite insérés indépendamment dans les échantillons. Ce schéma usuel s'explique notamment par sa simplicité de mise en œuvre (insertion et extraction), par la possibilité d'utiliser les méthodes usuelles de codage, préalablement à l'insertion, et par sa possibilité de "dispenser" les bits du message  $\mathbf{M}$  dans l'image de couverture  $\mathbf{C}$ .

### Stéganographie par substitution de LSB ("LSB Replacement")

Le bit,  $b_{n,0}$ , de moindre influence sur la valeur du pixel  $z_n$  est appelé bit de poids faible ou, en anglais, "Least Significant Bit" LSB. L'équation (2.1) met en évidence que si l'on

modifie  $b_{n,0}$  la valeur du pixel  $z_n$  n'est modifiée que de  $\pm 1$ . Or la modification d'un bit de poids faible d'un pixel dans une image numérique naturelle sera imperceptible compte-tenu de la sensibilité de la vision humaine. Par ailleurs, comme l'illustre la figure 2.3, si l'on observe les différents plans de bits d'une image (les  $i$ -ème bits de chacun des pixels), on constatera que les plans des bits de poids faibles sont nettement moins structurés que ceux de poids plus forts. Il est donc tout à fait possible d'insérer un message quelconque, sous sa forme binaire, en substituant les LSB d'autant de pixels que nécessaire. L'émetteur et le récepteur de l'image stéganographiée doivent préalablement échanger un sens de parcours des pixels de l'image. Comme évoqué dans la définition 2.1, pour éviter aux protagonistes d'avoir à échanger une quantité d'information supérieure à celle qu'ils souhaitent dissimuler, il est usuel d'utiliser une permutation pseudo-aléatoire  $\mathbf{p}_K$  construite à partir d'une clé  $K$  (également appelée graine). Plus formellement, la permutation  $\mathbf{p}_K$  utilisée est une bijection de l'ensemble  $\{1, \dots, MN\}$  sur lui-même (automorphisme) :  $\mathbf{p}_K : \{1, \dots, MN\} \mapsto \{1, \dots, MN\}$ . Ainsi le LSB du  $\ell$ -ième pixel parcouru  $b_{\mathbf{p}(\ell),0} = b_{\mathbf{p}_K(\ell),0}$  prendra la valeur du  $\ell$ -ième bit à insérer  $m_\ell$  (l'indice  $K$  est omis dans la suite pour alléger les notations). Schématiquement, considérons le pixel de couverture  $c_{\mathbf{p}(\ell)}$  dans lequel on souhaite insérer le  $\ell$ -ième bit d'information  $m_\ell$ . Le pixel stéganographié  $s_{\mathbf{p}(\ell)}$  s'écrit alors de manière équivalente [18, 22, 23] :

$$s_{\mathbf{p}(\ell)} = 2 \lfloor c_{\mathbf{p}(\ell)} / 2 \rfloor + m_\ell = \mathcal{Q}_2[c_{\mathbf{p}(\ell)}] + m_\ell \quad (2.2)$$

où  $\lfloor \cdot \rfloor$  représente l'opération d'arrondi par défaut et  $\mathcal{Q}_2[\cdot]$  représente la quantification par "suppression du bit de poids" faible. Notons que la valeur du pixel peut ne pas être modifiée  $s_{\mathbf{p}(\ell)} = c_{\mathbf{p}(\ell)}$  si avant insertion on a  $\text{lsb}(c_{\mathbf{p}(\ell)}) = b_{0,\mathbf{p}(\ell)} = m_\ell$ . À l'inverse, lorsqu'une modification de la valeur  $c_\ell$  est nécessaire on constatera que l'on a, de manière équivalente :

$$\text{lsb}(c_{\mathbf{p}(\ell)}) \neq m_\ell \Rightarrow s_{\mathbf{p}(\ell)} = c_{\mathbf{p}(\ell)} + 1 - 2\text{lsb}(c_{\mathbf{p}(\ell)}) = c_{\mathbf{p}(\ell)} + (-1)^{c_{\mathbf{p}(\ell)}}. \quad (2.3)$$

La formulation (2.3) permet, comparativement à (2.2), de mettre en évidence l'impact de la substitution de LSB : si la valeur d'un pixel est paire cette dernière ne peut être qu'incrémentée alors qu'à l'inverse si la valeur d'un pixel est impaire cette dernière ne peut être que décrétementée.

Extraire le message de la stégo-image se fait alors simplement, en considérant que le mot de passe générateur de la permutation pseudo-aléatoire est connu. Comme l'illustre la figure 2.4, le message extrait  $\widehat{\mathbf{M}}$  est alors recomposé en concaténant les LSB des pixels parcourus :

$$\forall \ell \in \{1, \dots, L\}, \widehat{m}_\ell = s_{\mathbf{p}(\ell)}.$$

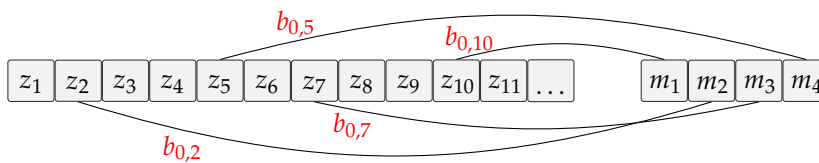


FIGURE 2.4: Exemple simple d'insertion / d'extraction dans les premiers pixels d'une ligne ; le sens de parcours obtenu dans ce cas est défini par le vecteur :  $\mathbf{p} = (10, 2, 7, 5, \dots)$ .

Historiquement, la stéganographie par substitution des LSB fut la première à voir le jour. Elle demeure encore souvent implémentée dans les logiciels téléchargeables sur Internet pour sa simplicité de mise en œuvre. Il est d'ailleurs possible d'implémenter cette méthode de stéganographie en une seule ligne de commande PERL comme suit [24, p.99] [18, p.41] :

*L'insertion est possible pour peu que le nombre de bits à insérer ne dépasse pas le nombre de pixels :  $L \leq N$ .*



```
perl -n0777e '$_ = unpack "b*", $_; split /(\s+)/, <STDIN>, 5;
uuuu@_ [8] = ~ s{.} {&&&v254 | chop () &v1} ge;
uuuu print @_' <image.pgm > stego_image.pgm secrettextfile
```

De nombreuses améliorations ont ensuite été proposées avec notamment pour but de réduire l'impact de l'insertion d'informations cachées (ou le nombre de modifications nécessaires à quantité équivalente d'informations cachées) [25, 26, 27, 28] ou de choisir la localisation des pixels à modifier dans les régions les plus structurées des images [29, 30, 31, 32].

Parmi les logiciels de stéganographie couramment répertoriés, BMPsteg, Hide and Seek, StegoDos et Steganos notamment reposent sur le schéma d'insertion par substitution de LSB.

### Stéganographie par correspondance de LSB ("*LSB Matching*")

L'amélioration la plus courante et la plus étudiée de la stéganographie par substitution de LSB est la méthode d'insertion par correspondance des LSB ou, en "*LSB matching*" anglais. Cet algorithme de stéganographie, très proche de celui de substitution de LSB, insère également le message  $\mathbf{M}$  dans les LSB des pixels. Ici encore le sens de parcours des pixels est usuellement choisi en générant une permutation pseudo-aléatoire. La principale évolution introduite par la correspondance des LSB est que si la valeur du LSB du pixel considéré ne correspond pas à celle qui doit y être insérée, le stéganographe s'autorise à incrémenter ou décrémenter la valeur de ce pixel de façon aléatoire. Dans les faits, incrémentation et décrémentation de la valeur du pixel (lorsque cela est nécessaire) sont généralement équiprobables.

Formellement, considérons le pixel de couverture de valeur  $c_k = c_{\mathbf{p}(\ell)}$  dans lequel on souhaite insérer le  $\ell$ -ième bit d'information  $m_\ell$  par correspondance de LSB. Le pixel stéganographié  $z_k$  s'écrit alors de manière équivalente [18, 22] :

$$z_k = c_k + u_k \cdot \text{xor}(\text{lsb}(c_k), m_\ell) \quad \text{où} \quad \frac{u_k + 1}{2} \sim B(1, p) \quad (2.4)$$

et  $\text{xor}(\cdot)$  représente l'opérateur logique "ou exclusif" modélisant le fait que  $z_k = c_k$  si  $\text{lsb}(c_k) = m_\ell$ , en raison du fait que  $\text{xor}(\text{lsb}(c_k), m_\ell) = 0$ . À l'instar de la relation (2.3), l'impact de la correspondance de LSB peut être mis en évidence, lorsqu'une modification de la valeur  $c_k$  est nécessaire, en constatant que :

$$\text{lsb}(c_k) \neq m_\ell \Rightarrow z_k = c_k + u_k = \mathcal{Q}_2[c_k + u_k] + m_\ell. \quad (2.5)$$

L'algorithme D de décodage du message est identique à celui présenté dans le cas de la stéganographie par substitution de LSB et illustré dans la figure 2.4. En effet, puisque le principe demeure de faire correspondre la valeur du LSB des pixels parcourus aux bits du message que l'on souhaite insérer, le décodage n'est en rien modifié.

Comme nous le verrons dans la section 2.3, et avec plus de détails statistiques dans le chapitre 3, la légère modification introduite par la méthode de stéganographie par correspondance des LSB modifie considérablement l'impact d'insertion; détecter la présence d'informations cachées par correspondance des LSB est nettement plus délicat que lorsque l'information est cachée par substitution de LSB.

La méthode de stéganographie par correspondance des LSB a été introduite pour la première fois dans [34]. En effet, ayant été observé que la stéganographie par substitution de LSB introduit des artefacts caractéristiques, de nombreuses méthodes (présentées dans la section 2.3) permettant de détecter la présence de ces artefacts ont vu le jour. La stéganographie par correspondance de LSB n'introduisant pas ces artefacts caractéristiques, les méthodes spécifiquement dédiées à la détection de stéganographie par substitution de LSB sont inefficaces pour détecter la correspondance de LSB.

Les diagrammes de la figure 2.5 illustrent les modifications engendrées par l'insertion d'informations cachées selon quelques méthodes basées sur la substitution et la correspondance de LSB.

*Intuitivement, cette construction correspond au pire, c'est à dire ne créant pas de modification particulièrement identifiable. Dans un jargon plus proche de la théorie de l'information, la configuration d'équiprobabilité de l'incrémenta-tion et de la décrémentation est celle minimisant l'entropie relative des distributions avant et après insertion [33]*

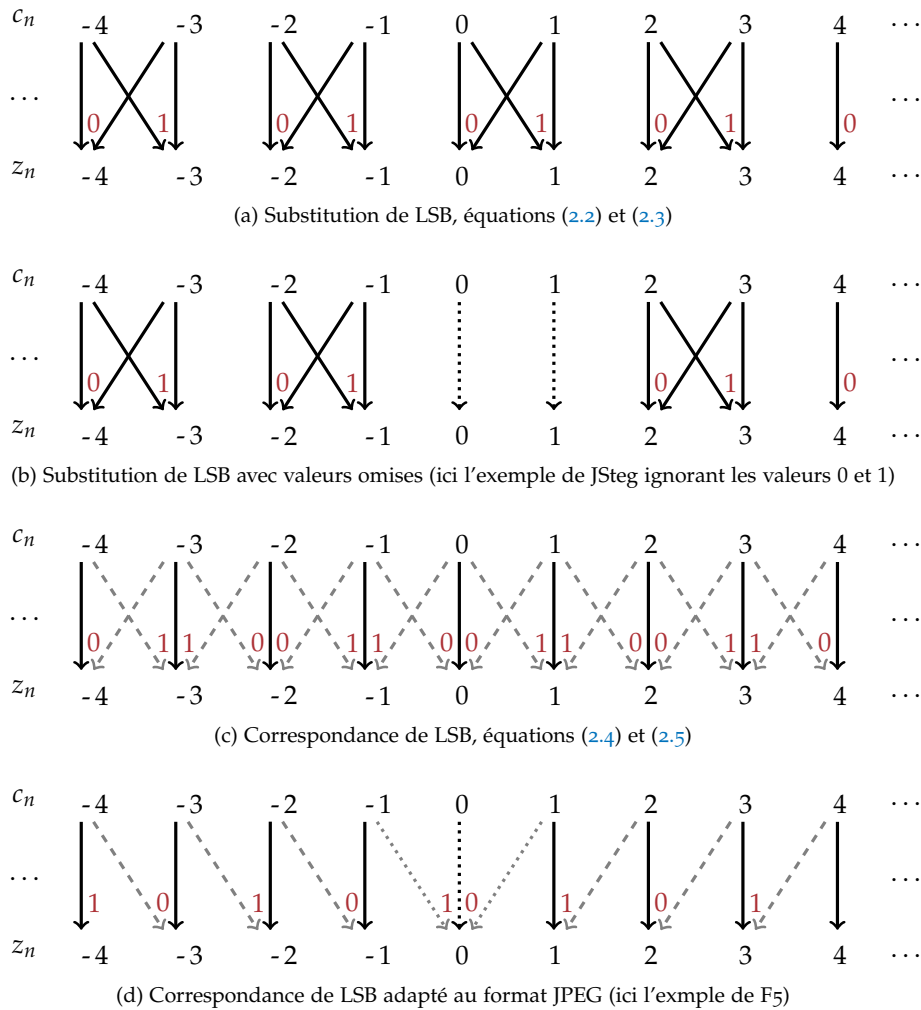


FIGURE 2.5: Exemple de modifications engendrées par l'insertion d'informations cachées dans les LSB d'une image avec différentes méthodes (intensité du pixels  $c_n$  avant et après  $z_n$  insertion).

**Remarque 2.2.** Il est enfin notable que, pour les méthodes d'insertion d'informations dans le LSB, le taux d'insertion est simplement défini par :  $R = L/N$ .

### Stéganographie par étalement de spectre

Les deux méthodes de stéganographie présentées précédemment n'ont pas pour objectif d'être robustes à une modification de la stégo-image. Il est aisé de comprendre que l'effacement de tous les LSB d'une image, par exemple, supprimera de fait le message en rendant l'extraction impossible. Plus proche du watermarking, les méthodes de stéganographie par étalement de spectre et par modulation d'index de quantification, ou en anglais "*Spread-spectrum*" et QIM "*Quantized Index Modulation*", offrent la possibilité de dissimuler des informations avec une robustesse paramétrable. Comme expliqué dans [35, 36, 37, 38], l'étalement de spectre consiste à insérer le  $\ell$ -ième bit du message  $m_\ell$  de façon additive dans un bloc  $c_k$  de pixels pour obtenir le bloc de pixels stéganographiés ainsi :

$$\mathbf{s}_k(c_k, m_\ell) = c_k + w(m_\ell)$$

où  $w(m_\ell)$  est typiquement un vecteur pseudo-aléatoire [39, 38] de la forme :

$$w(m_\ell) = a(m_\ell)v$$

avec ici  $v$  un vecteur (usuellement d'énergie unitaire, *i.e.*  $\|v\| = 1$ ) prédéfini dans un dictionnaire et  $a(m_\ell)$  une fonction de modulation selon la valeur du bit  $m_\ell$  du message que l'on souhaite insérer. Le décodage du  $k$ -ième bit du message se fait simplement par projection [39, 38] du bloc de pixels sur le vecteur  $v$  :

$$\tilde{m}_k = v^T \mathbf{s}_k(\mathbf{c}_k, m_\ell) = v^T (\mathbf{c}_k + a(m_\ell)v) = v^T \mathbf{c}_k + a(m_\ell)\|v\|_2^2. \quad (2.6)$$

La valeur du bit décodé donnée par la relation (2.6) met en évidence le fait que cette technique de dissimulation ne permet pas d'assurer, en soit, l'intégrité du message puisque, l'interférence créée par le signal hôte (le bloc de pixels de couverture  $\mathbf{c}_k$ ) peut modifier la valeur du bit décodé  $\tilde{m}_k$ . Si par exemple on considère les fonctions d'encodage et de décodage suivantes :

$$a(m_\ell) = \begin{cases} \Delta & \text{si } m_\ell = 1 \\ -\Delta & \text{si } m_\ell = 0 \end{cases} \quad \text{et} \quad \tilde{m}_k = \begin{cases} 1 & \text{si } v^T \mathbf{s}_k > 0 \\ -1 & \text{si } v^T \mathbf{s}_k \leq 0 \end{cases}, \quad (2.7)$$

il est facile de se rendre compte avec l'équation (2.6) que si  $|v^T \mathbf{c}_k| > |\Delta|\|v\|_2^2$  le décodage du  $k$ -ième bit du message est erroné. Certes, les codes correcteurs linéaires permettent simplement et efficacement de détecter et de corriger une éventuelle erreur de décodage (dont la probabilité d'occurrence est *a priori* très faible). En revanche, on comprend que l'utilisation d'un bloc  $\mathbf{c}_k$  pour la dissimulation d'un unique bit d'information engendre une perte conséquente dans la capacité du stégosystème : l'utilisation, en sus, d'un codage introduisant de la redondance n'est pas forcément souhaitable.

On notera par ailleurs que dans la fonction d'encodage telle que présentée dans (2.7), le paramètre  $\Delta$  est un facteur influant sur la robustesse et la discrétion du stégosystème. Augmenter la valeur de  $\Delta$  rend le message plus robuste aux éventuelles modifications de l'image mais facilite la détection de la présence d'informations cachées.

La modification apportée par [39, 40] consiste à utiliser la valeur  $v^T \mathbf{c}_k$  lors de l'insertion du  $k$ -ième bit d'information. La figure 2.6 illustre un schéma simple d'insertion par QIM. Plus précisément, le schéma d'insertion par QIM peut s'écrire de manière générale à l'aide d'une fonction de quantification de  $FQ_\Delta(\cdot)$ , de pas  $\Delta$ , sous la forme :

$$\mathbf{s}_k(\mathbf{c}_k, m_\ell) = FQ_\Delta(\mathbf{c}_k + d(m_\ell)) - d(m_\ell)$$

où le vecteur de distorsion  $d(m_\ell)$  est un code binaire, ne dépendant que du bit que l'on souhaite insérer, respectant notamment [40] :

$$d(m_\ell=1) = \begin{cases} d(m_\ell=0) + \Delta/2 & \text{si } d(m_\ell=0) < 0 \\ d(m_\ell=0) - \Delta/2 & \text{si } d(m_\ell=0) \geq 0. \end{cases}$$

Les schémas d'insertion par QIM sont notamment étudiés pour leur propriété de robustesse face à des modifications de l'image et d'indétectabilité [41, 42, 43].

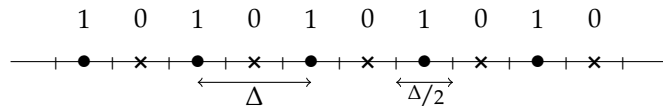


FIGURE 2.6: Représentation simplifiée d'insertion par Quantification d'Index de Modulation (QIM) avec utilisation d'un quantificateur uniforme de pas  $\Delta$ .

Dans ce manuscrit, nous ne nous intéresserons pas davantage à ce type de méthodes de stéganographie héritées du watermarking. En effet, comme expliqué dans l'introduction générale, chapitre 1, la problématique abordée concerne l'application de la

théorie de la décision statistique pour la détection d'informations cachées dans les LSB d'une image (par substitution voire par correspondance de LSB); par la suite, la méthodologie originale proposée pourra être utilisée pour la détection de schémas de stéganographie plus évolués.

### 2.2.2 Insertion d'informations dans un domaine transformé

Jusqu'à présent, nous avons évoqué les schémas d'insertion dans le domaine spatial, c'est à dire considérant l'image sous sa forme originale d'une matrice d'entiers. Il est courant, dans le domaine du traitement des images (et plus généralement pour le traitement des signaux), d'utiliser un domaine de représentation transformé, notamment pour permettre une représentation parcimonieuse des images. Dans le cadre de la stéganographie, la dissimulation d'informations dans un domaine transformé de l'image présente deux principaux intérêts :

- L'amélioration de la robustesse du stégosystème : la transformée de Fourier-Mellin est par exemple utilisée dans le domaine du tatouage numérique [44, 45] pour sa propriété d'invariance aux rotations, aux translations et aux changements d'échelle (agrandissement ou réduction de la taille de l'image).
- L'amélioration de la sécurité du stégosystème : la modélisation des images peut être simplifiée en utilisant une transformation de cette dernière; il a été ainsi conçu des stégosystèmes permettant que l'image stéganographiée respecte un modèle (empirique) prédéfini [46, 47, 48]. Naturellement ces modèles empiriques des images transformées ont également été exploités pour détecter l'insertion d'informations cachées [48, 49, 50, 51].

En outre, la très grande majorité des méthodes de compression d'images avec perte reposent sur une transformée discrète (tel le JPEG ou le JPEG2000 exploitant respectivement la transformée en cosinus discrète et la transformée en ondelettes discrète). Compte-tenu de la grande proportion de ce type d'images, des stégosystèmes adaptés à ces formats de fichiers ont inévitablement vu le jour. Cependant, compte-tenu de la spécificité de ces formats de fichier, un stégosystème doit être spécifiquement adapté à l'insertion d'informations dans un standard de compression d'images. En effet, dans le cas contraire, une modification inappropriée risquerait, d'une part de rendre l'image illisible et, d'autre part, de mettre en péril la discrétion du stégosystème.

À y regarder de plus près [18, 22, 19], la très grande majorité des stégosystèmes opérant dans un domaine transformé utilisent les mêmes méthodes d'insertion que celles décrites pour le domaine spatial. Les algorithmes de stéganographie de référence pour les images JPEG, tels que F5 [52], Outguess et Jsteg [18], reposent sur la méthode d'insertion par correspondance de LSB.

Les trois principales spécificités des images dont le principe de compression repose sur une transformée discrète des données sont :

- le grand nombre de méta-informations nécessaires,
- la plus grande étendue de valeurs possibles.
- la distribution de probabilité spécifique des valeurs codées.

Le premier point ne nécessite que des changements d'implémentation informatique (identification des métadonnées et des données visuelles dans le fichier). Les derniers points peuvent en revanche amener à des modifications de la méthode d'insertion. Il faut en effet tenir compte de la méthode de compression et des propriétés particulières des valeurs codées sans quoi la présence du message stéganographié peut devenir évidente. On pourra citer à titre d'exemple les schémas d'insertion dans les images codées à l'aide d'une palette de couleurs (par exemple, les images au format GIF ou PNG) procédant préalablement à un tri préalable de la palette. L'analyse de ces schémas de stéganographie a montré une faiblesse vis-à-vis d'une détection par signature [53, 54].

*On distinguera les images compressées avec perte des images compressées sans perte. En effet, pour les images compressées sans perte les méthodes usuelles de stéganographie adaptées au domaine spatial demeurent applicables.*

## 2.3 La stéganalyse, détection de stéganographie : difficultés et solutions

Dans le domaine plus connu de la cryptographie, la cryptanalyse a pour but de décrypter un message, ayant été préalablement crypté, sans connaissance de la clé. En ce qui concerne la stéganalyse, le *Saint-Graal* de la stéganalyse serait, de manière analogue, de tenter d'extraire l'information dissimulée sans connaissance de la clé du stégosystème. Seulement, un rapide calcul de dénombrement permet de comprendre la difficulté : supposons que nous sommes en possession d'une image de  $128 \times 128$  pixels parmi lesquels 32 sont porteurs d'informations cachées. Pour extraire le message de ces 32 pixels, il faut avant toute chose retrouver les bons pixels dans le bon ordre. Or cette configuration offre  $(128^2)! / (128^2 - 32)! \approx 7.051509 \cdot 10^{134}$  combinaisons possibles... L'attaque par force brute pour extraire l'information cachée semble donc vaine. En outre, si l'on suppose que les 32 pixels porteurs d'informations cachées sont identifiés, il est immédiat de noter que  $32! \approx 2.6313 \cdot 10^{35}$  combinaisons de ces LSB dans un ordre différent sont possibles. Enfin, le message peut avoir été préalablement crypté avant son insertion, il faut alors faire appel à des méthodes de cryptanalyse pour finalement retrouver le sens du message caché.

Ainsi, les objectifs de la stéganalyse sont beaucoup plus modestes que l'extraction du message secret. Il existe deux classes de stéganalyse se différenciant par les objectifs recherchés et les moyens utilisés. Dans le cadre de la stéganalyse active, l'objectif est d'empêcher la communication du message secret en modifiant le support (l'image) suffisamment peu pour en conserver le contenu perceptible mais de façon suffisamment importante pour en supprimer le contenu dissimulé. Par exemple, dans le cas de méthodes de stéganographie dans les LSB (par substitution ou par correspondance) écraser tous les bits de poids faible, en les mettant par exemple à 0 ou à 1, efface l'éventuel message dissimulé.

*Comme le montre la figure 2.7, l'apparente indiscernabilité entre images avant et après stéganographie n'est pas révélatrice de la possibilité statistique de détecter la présence du message.*

Les travaux présentés s'intéressent uniquement à la stéganalyse passive qui, comme son nom le laisse entendre, ne s'autorise pas la modification du média lors de l'analyse. Le but premier est alors de détecter la présence d'informations cachées ou mieux, d'estimer la taille du message dissimulé : la terminologie de "stéganalyse quantitative" [55, 56, 57] est alors usuellement utilisée. Détecter la présence d'un message caché dans une image peut-être formalisé de manière générique comme un test d'hypothèses dont le but est de décider entre les deux hypothèses suivantes :

*Des explications approfondies sur les tests statistiques d'hypothèses sont données dans l'annexe A.*

$$\begin{cases} \mathcal{H}_0 : \mathbf{Z} = \mathbf{C} & \text{le média } \mathbf{Z} \text{ est sain} \\ \mathcal{H}_1 : \mathbf{Z} = \mathbf{S} & \text{le média } \mathbf{Z} \text{ est stéganographié.} \end{cases} \quad (2.8)$$

Notons enfin qu'il a été récemment proposé des méthodes de stéganalyse dont le but est d'estimer la clé utilisée pour insérer le message. Ces quelques méthodes ont été appliquées pour la stéganographie des images dans le domaine spatial [58, 59, 19] et dans un domaine transformé [60]. Ces méthodes reposent sur l'hypothèse forte que plusieurs médias -de même taille- stéganographiés avec la même clé sont disponibles lors de la stéganographie. La même supposition a été utilisée dans [61] pour montrer qu'il est -au moins théoriquement- possible de localiser les pixels ayant été modifiés par l'insertion d'un message. Dans ce dernier cas, le principe est simplement de ne pas analyser chacune des images mais chacun des pixels de même localisation dans les images (de taille identique). Bien que ces méthodes soient intéressantes elles ne seront pas étudiées dans ce manuscrit. En effet, il semble d'une part peu probable que de nombreux médias stéganographiés avec la même clé et de taille identique soient disponibles. D'autre part, la détection statistique d'informations cachées est un travail suffisamment difficile pour que ce manuscrit y soit entièrement consacré.

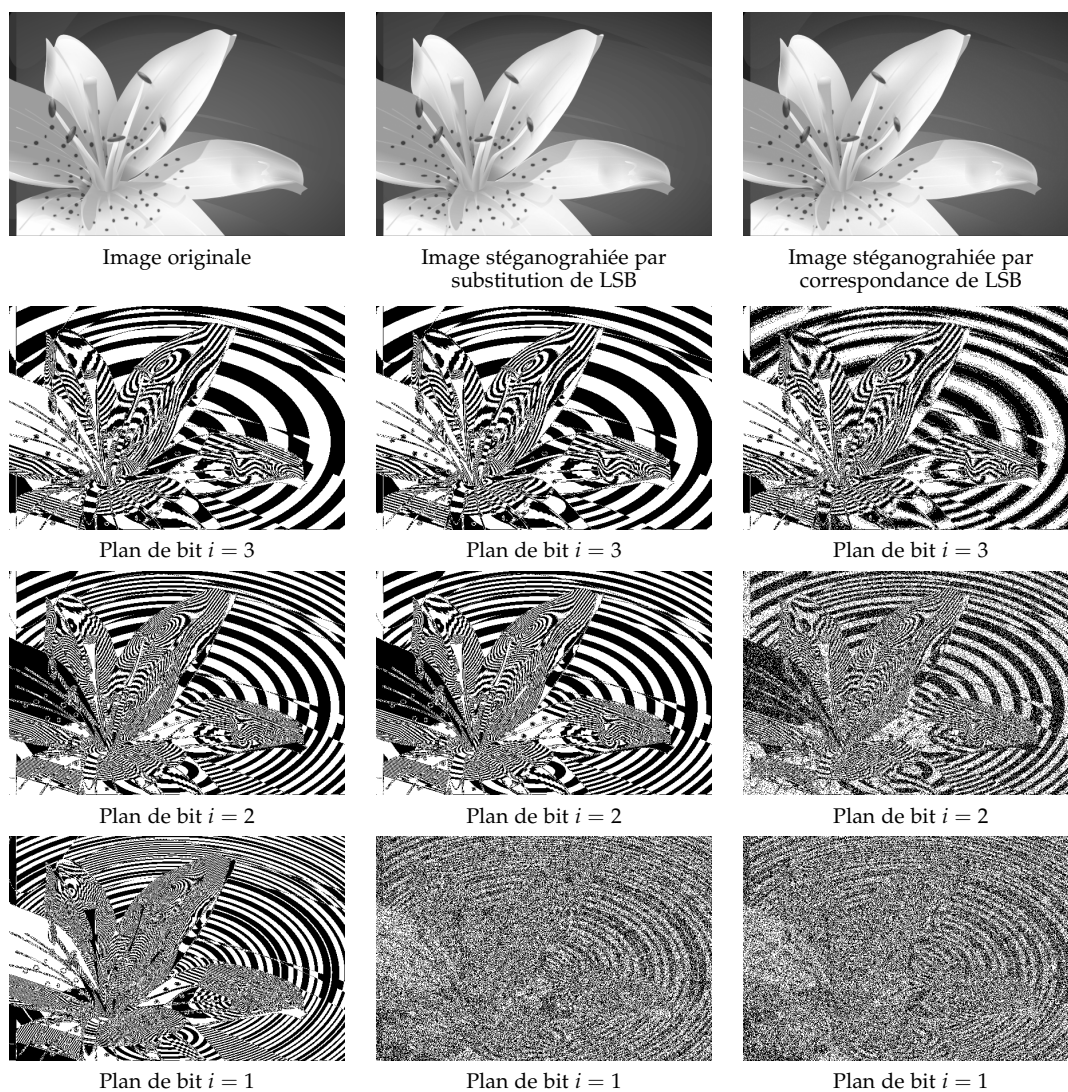
(a) Image originale et stéganographiée avec un taux d'insertion  $R = 0.7$ 

FIGURE 2.7: Décomposition des plans de bits d'une image en niveaux de gris avant et après stéganographie : les changements opérés par l'insertion d'un message ne sont observables que dans le(s) plan(s) des bits de poids faible. L'image utilisée dans cet exemple étant générée par ordinateur, les bits de poids faibles sont structurés mais cela n'est pas valable pour une image naturelle, comme l'illustre la figure 2.3.

Dans la suite de ce manuscrit, le terme de stéganalyse sera utilisé pour désigner uniquement les méthodes statistiques permettant la détection d'informations cachées ; celle-ci peut en effet se faire différemment, notamment par :

- l'analyse sensorielle du média, méthode qui demeure anecdotique ;
- la comparaison avec le fichier original, lorsque ce dernier est disponible ;
- l'analyse de signature dont le but est de détecter les médias stéganographiés par des informations annexes en général liées à l'implémentation informatique. L'image 1.1, issue de [1], a par exemple été détectée par le fait que F5 introduit par défaut un commentaire dans les méta-données JPG ;

Dans cette section 2.3 les méthodes de détection existantes sont synthétiquement présentées. N'ayant pas volonté à l'exhaustivité, cette section pourra être complétée par les ouvrages [18, 19, 22].

### 2.3.1 Stéganalyses universelles ou aveugles

Lors du renouveau de la stéganographie et de la stéganalyse (à la fin des années 90), différentes attaques de stégosystèmes ont été répertoriées dans [53] selon les informations disponibles lors de l'analyse (connaissance du message, de l'image de couverture, de la clé, etc. . .). Les deux méthodes de stéganalyse, qui sont effectivement étudiées, se distinguent par la connaissance ou non du stégosystème. Dans cette section, nous nous intéresserons aux méthodes de stéganalyse universelles ou aveugles pour lesquelles le stégosystème n'est pas connu. Le problème de décision est alors proche de celui défini de façon générique par l'équation (2.8). Comme nous le verrons plus en détails dans la section 2.4, le problème de la stéganalyse universelle tel qu'il est explicité par la relation (2.8) ne peut être résolu formellement en utilisant les résultats classiques issus de la théorie de la décision statistique. Malgré cela, les méthodes de stéganalyse universelle sont sans conteste les plus populaires actuellement.

Ne pouvant exploiter la théorie de la décision statistique, les techniques de stéganalyse universelle s'appuient largement sur les méthodes d'apprentissage supervisé. Ces méthodes reposent sur l'extraction de caractéristiques et fonctionnent en deux temps ; lors de la phase d'apprentissage, une large base de données est utilisée pour extraire les caractéristiques de chacun des médias associés à une classe préalablement connue (stéganographié ou non). Un classifieur est ensuite utilisé pour discriminer le plus précisément possible les deux classes d'objets à partir des caractéristiques extraites. À la suite de cette première phase d'apprentissage, le détecteur est opérationnel. Dans la seconde phase de test, des médias sont fournis au détecteur qui doit décider de la classe à laquelle chacun d'eux appartient. En général, l'utilisateur connaît la classe des images qui ont été fournies au détecteur et peut ensuite juger des performances de ce dernier en comparant les résultats de la classification avec les résultats corrects. Après cette seconde phase de test le détecteur peut être amélioré ou mis en situation opérationnelle de détection.

*L'amélioration ou "réglage" des détecteurs universels peut se faire en redéfinissant les caractéristiques utilisées ou en modifiant les paramètres de la méthode de classification.*

*Mesurer les performances des méthodes de classification par apprentissage supervisé ne peut se faire que par benchmarking, i. e. à travers des résultats empiriques sur de larges bases de données*

L'intérêt pratique et les bonnes performances obtenues empiriquement par les détecteurs universels, par apprentissage, restent sujet à caution compte-tenu des limites théoriques de la méthodologie.

Tout d'abord, dans toutes les applications des méthodes d'apprentissage (supervisé ou non-supervisé), le choix des caractéristiques pertinentes pour la classification est crucial. Pour la stéganalyse, ce choix est davantage guidé par l'intuition ou l'observation de changements induits par la stéganographie que par les propriétés intrinsèques des images naturelles. Ne pouvant donc pas disposer d'une modélisation statistique précise de ces caractéristiques, il n'est pas possible de fixer *a priori* les probabilités d'erreurs de classification. La seule possibilité de connaître les performances des méthodes de stéganalyse universelle et de maîtriser les erreurs de classification est l'expérimentation numérique sur une large base de données (que l'on souhaite la plus représentative possible de l'ensemble des données potentiellement analysables). Il a d'ailleurs été montré expérimentalement dans [62] que les résultats obtenus par les principaux algorithmes de stéganalyse universelle dépendent fortement de la base de données utilisée.

En outre, la méthode d'apprentissage et de classification doit, elle aussi, être choisie empiriquement ; en effet, il n'existe pas de résultat théorique permettant de prédire quelle méthode offrira les meilleures performances pour un problème donné (en terme de minimisation d'erreurs de classification par exemple). Trois méthodes de classification sont usuellement utilisées dans le domaine de la stéganalyse dont les fondements théoriques sont brièvement expliqués dans la section A.5 de l'annexe A. Historiquement, le premier algorithme de stéganalyse universelle proposé par H.Farid [49] est basé sur l'analyse discriminante linéaire de Fisher, ou FLD de l'anglais *Fisher Linear Discriminant analysis*, issue de [63]. Bien que l'approche de H.Farid ait permis d'ou-

vrir la voie de la stéganalyse universelle, l'utilisation de la FLD dans le domaine de la stéganalyse demeure cependant très minoritaire [64, 65]. Les méthodes à base de réseaux de neurones, ou NN de l'anglais "Neural Network", ont été utilisées plus largement [50, 66, 67]. Dans le domaine de la stéganalyse universelle, les méthodes de classification très majoritairement utilisées sont incontestablement les Séparateurs à Vastes Marges, ou SVM de l'anglais *Support Vector Machine*, issus des travaux de V.Vapnik [68, 69]. Devant les bonnes performances des techniques de stéganalyse universelle basées sur les SVM [70, 71, 72, 73, 74, 75, 31], il a également été proposé d'utiliser ces méthodes d'apprentissage pour faire de la régression. Suivant la méthodologie initialement proposée dans [76], le but recherché est de faire de la stéganalyse quantitative [55, 57] *i.e.* d'estimer la taille du message dissimulé.

Selon les caractéristiques utilisées pour l'apprentissage et la classification, les méthodes de stéganalyses universelle sus-décrites ont été utilisées indifféremment pour détecter la présence d'informations dissimulées dans le domaine spatial ou dans un domaine transformé.

Un second type de stéganalyse universelle a été étudié dans une moindre mesure. En reprenant la définition 2.1 d'un stégosystème, il est supposé que l'image stéganographiée  $\mathbf{S} = (s_1, \dots, s_N)^T$  peut-être définie comme une combinaison linéaire de l'image saine  $\mathbf{C} = (c_1, \dots, c_N)^T$  et d'un "bruit"  $\mathbf{W}(\mathbf{M}) = (w(m_1), \dots, w(m_l), 0, \dots)^T$  représentant l'impact d'insertion du message  $\mathbf{M}$ . Pour chaque pixel (ou plus largement pour chaque échantillon) indexé par  $k$ , avec par convention  $\mathbf{p}(\ell) = k$ , on a la relation :

$$s_k = c_k + \Delta w(m_\ell), \quad (2.9)$$

où le paramètre de pondération  $\Delta$  représente la "puissance d'insertion", comme précédemment décrit pour l'insertion par étalement de spectre dans l'équation (2.7). La stéganalyse consiste alors à séparer, à partir de  $\mathbf{S}$ , les composantes  $\mathbf{C}$  et  $\mathbf{M}$ . Supposons que  $i$  images stéganographiées  $(\mathbf{S}^{(1)}, \dots, \mathbf{S}^{(i)})^T$  créées à partir de la même image de couverture  $\mathbf{C}$  et du même message  $\mathbf{M}$  sont disponibles lors de l'analyse. En regroupant le  $k$ -ième pixel de chaque image  $\mathbf{s}_k = (s_k^{(1)}, \dots, s_k^{(i)})^T$ , l'équation (2.9) peut être ré-écrite algébriquement par :

$$\mathbf{s}_k = \mathbf{A} \mathbf{R}_k \quad \text{où} \quad \mathbf{A} = \begin{pmatrix} 1 & \Delta^{(1)} \\ \vdots & \vdots \\ 1 & \Delta^{(i)} \end{pmatrix} \quad \text{et} \quad \mathbf{R}_k = \begin{pmatrix} c_k \\ w(m_\ell) \end{pmatrix}. \quad (2.10)$$

Le problème décrit par (2.10) est alors celui de la séparation de sources en aveugle, le but étant alors d'extraire le message, qui peut être résolu avec les méthodes classiques du domaine [77]. Parmi ces méthodes, l'analyse en composantes indépendantes (ou ICA, *Independent Component Analysis* en anglais) a été utilisée, dans [78] notamment, pour la stéganalyse universelle. Notons néanmoins qu'en pratique, ces stéganalyses demeurent plutôt adaptées à la détection d'informations cachées par étalement de spectre [37, 38], elle-même davantage utilisée pour l'insertion de tatouages numériques.

Par ailleurs l'approche de stéganalyse par séparation aveugle de sources possède plusieurs limites théoriques. D'une part le problème ne peut être résolu que si les distributions statistiques des stégo-objets sont stationnaires (ce qui en pratique est difficilement vérifiable pour des images). D'autre part l'insertion d'informations doit pouvoir être représentée par une matrice de mélange (matrice notée  $\mathbf{A}$  dans le présent cas) ce qui implique une relation linéaire. Si l'insertion d'informations n'est plus linéaire (ce qui est le cas par exemple pour les méthodes par substitution ou correspondance de LSB), la séparation aveugle de sources, et donc l'extraction du message, devient beaucoup plus complexe. Enfin, bien qu'intéressant d'un point de vue théorique, ce



problème ne trouve que peu d'applications pratiques car il semble assez peu réaliste de supposer que  $i$  stégo-médias générés à partir de la même image de couverture et du même message sont disponibles.

Les méthodes de stéganalyse universelle sortent légèrement du cadre d'étude de ce manuscrit, aussi le lecteur intéressé pourra compléter cette présentation succincte par la lecture de [79].

### 2.3.2 Stéganalyses adaptées à un stégosystème

Dans cette section, nous nous intéresserons davantage aux méthodes de stéganalyse spécifiquement dédiées à un schéma d'insertion ; dans la terminologie usuelle ces méthodes sont dites spécifiques (la dénomination initiale de "*Chosen stego attack*" issue de [53] n'est plus guère utilisée). Le problème de détection sous-jacent à la stéganalyse adaptée à un stégosystème est différent de celui défini dans (2.8) ; il s'agit ici de décider entre les deux hypothèses suivantes :

$$\begin{cases} \mathcal{H}_0 : \mathbf{Z} = \mathbf{C} & \text{le média } \mathbf{Z} \text{ est sain} \\ \mathcal{H}_1 : \mathbf{Z} = \mathbf{E}(\mathbf{M}, \mathbf{C}, K) & \text{le média } \mathbf{Z} \text{ est stéganographié avec la méthode } E. \end{cases} \quad (2.11)$$

Une grande partie des méthodes de stéganalyse adaptées ont pour but de détecter la présence d'informations insérées par substitution de LSB. En effet, bien que des méthodes d'insertions plus évoluées existent (voir section 2.2.1), la substitution de LSB demeure certainement la méthode d'insertion la plus populaire.

### 2.3.3 Stéganalyses adaptées à la stéganographie par substitution de LSB

Les méthodes de stéganalyse adaptées à la substitution de LSB exploitent toutes, de différentes façons, la particularité des modifications engendrées par ce type d'insertion d'informations. Comme expliqué dans la section 2.2.1, un pixel dont la valeur est paire ne peut être qu'incrémenté alors qu'à l'inverse un pixel dont la valeur est impaire ne peut être que décrémenté. Cet impact "asymétrique" engendré par l'insertion d'informations par substitutions de LSB a été exploité à de nombreuses reprises dans la littérature, de façons différentes, pour permettre de détecter la présence d'informations cachées dans un média analysé  $\mathbf{Z}$ .

*La classification des méthodes de stéganalyse adaptées à la substitution de LSB est due à A. Ker, voir par exemple [80]*

#### 2.3.3.1 Détecteurs structuraux

Historiquement, le premier détecteur structural de stéganographie par substitution de LSB est l'analyse RS (pour Régulier-Singulier) proposé par J. Fridrich dans [81]. Le fondement de cette analyse intuitive a été justifié et utilisé par Dumitrescu pour proposer, dans [82], l'analyse par paire de pixels ou SPA de l'anglais *Sample Pair Analysis*. D'un point de vue pédagogique, le fonctionnement de l'algorithme SPA est détaillé pour expliquer le principe de fonctionnement des détecteurs structuraux.

Considérons l'image analysée  $\mathbf{Z}$  comme un vecteur  $\mathbf{Z} = (z_1, \dots, z_N)^T$  dont les pixels sont indexés par  $n \in \mathcal{I}$ . En analysant les couples de pixels adjacents d'une image  $\forall (2n, 2n+1) \in \mathcal{I}^2, (z_{2n}, z_{2n+1})$ , il est proposé de définir les sous-ensembles suivants :

$$\begin{aligned} \mathcal{C}_k &: \{(2n, 2n+1) \in \mathcal{I}^2 \mid \lfloor z_{2n}/2 \rfloor = \lfloor z_{2n+1}/2 \rfloor + k\} \\ \mathcal{E}_k &: \{(2n, 2n+1) \in \mathcal{I}^2 \mid z_{2n} = z_{2n+1} + k, z_{2n+1} \text{ est pair}\} \\ \mathcal{O}_k &: \{(2n, 2n+1) \in \mathcal{I}^2 \mid z_{2n} = z_{2n+1} + k, z_{2n+1} \text{ est impair}\} \end{aligned}$$

où  $\lfloor \cdot \rfloor$  représente l'opération d'arrondi par défaut.

Pour chaque élément de ces sous-ensembles, ici indexés par  $k$ , il est proposé de calculer

les probabilités de transition d'un ensemble à un autre lors de l'insertion d'informations cachées. En utilisant, ici encore, l'hypothèse que tous les pixels puissent être utilisés pour cacher un bit d'information avec la même probabilité  $p = L/2N = R/2$ , un simple calcul permet de représenter les probabilités de transitions d'un sous-ensemble à un autre sous la forme du diagramme état/transition illustré par la figure 2.8.

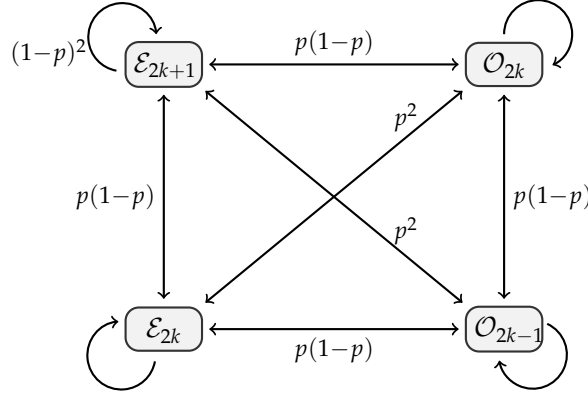


FIGURE 2.8: Diagramme état / transition illustrant les probabilités de passage d'un sous-ensemble à un autre pour chaque paire de valeurs.

De manière analogue, en notant  $N_c(e_k)$  et  $N_c(o_k)$  (respectivement  $N_s(e_k)$  et  $N_s(o_k)$ ) le nombre de pixels appartenant aux ensembles  $\mathcal{E}_k$  et  $\mathcal{O}_k$  dans l'image de couverture (respectivement dans la stégo-image), les probabilités de transition peuvent s'écrire sous la forme algébrique suivante :

$$\begin{pmatrix} N_s(e_{2k}) \\ N_s(o_{2k-1}) \\ N_s(e_{2k+1}) \\ N_s(o_{2k}) \end{pmatrix} = \begin{pmatrix} (1-p)^2 & p(1-p) & p(1-p) & p^2 \\ p(1-p) & (1-p)^2 & p^2 & p(1-p) \\ p(1-p) & p^2 & (1-p)^2 & p(1-p) \\ p^2 & p(1-p) & p(1-p) & (1-p)^2 \end{pmatrix} \begin{pmatrix} N_c(e_{2k}) \\ N_c(o_{2k-1}) \\ N_c(e_{2k+1}) \\ N_c(o_{2k}) \end{pmatrix} \quad (2.12)$$

La matrice constituée des probabilités de transition étant inversible, il découle de la relation (2.12) que :

$$\begin{pmatrix} N_c(e_{2k}) \\ N_c(o_{2k-1}) \\ N_c(e_{2k+1}) \\ N_c(o_{2k}) \end{pmatrix} = \frac{1}{(1-2p)^2} \begin{pmatrix} (1-p)^2 & -p(1-p) & -p(1-p) & p^2 \\ -p(1-p) & (1-p)^2 & p^2 & -p(1-p) \\ -p(1-p) & p^2 & (1-p)^2 & -p(1-p) \\ p^2 & -p(1-p) & -p(1-p) & (1-p)^2 \end{pmatrix} \begin{pmatrix} N_s(e_{2k}) \\ N_s(o_{2k-1}) \\ N_s(e_{2k+1}) \\ N_s(o_{2k}) \end{pmatrix} \quad (2.13)$$

Lors de l'analyse d'une image  $\mathbf{Z}$ , le nombre d'éléments de chacun des sous-ensembles  $N_s(\cdot)$  est accessible. Une résolution de l'équation (2.13) pourrait permettre de calculer  $p$  à condition de pouvoir connaître le nombre d'éléments de chacun des sous-ensembles dans l'image originale. Pour ce faire, l'algorithme du SPA repose sur le "modèle" ou "l'a priori" que  $\forall k, N_c(e_k) = N_s(e_k)$ . L'utilisation de ce modèle dans l'équation (2.13) permet d'obtenir, après quelques calculs, l'équation du second degré suivante :

$$0 = N_s(e_k) - N_s(e_{k+1})p^2 + N_s(e_{2k+1}) - N_s(o_{2k+1}) + [N_s(e_{2k+2}) + N_s(o_{2k+2}) + 2N_s(o_{2k+1}) - 2N_s(e_{2k+1}) - N_s(e_{2k}) + N_s(o_{2k})] p. \quad (2.14)$$

dont la solution  $\hat{p}$  est utilisée comme "estimateur" de la taille du message inséré  $\hat{L} = 2\hat{p}N$ .

Notons que dans la version originale de l'algorithme du SPA [82], l'équation du second degré définie dans (2.14) n'admet souvent pas de racine réelle. Ce problème a donné

lieu à l'amélioration proposée par l'algorithme du SPA-LSM[83] utilisant le même principe que le SPA mais amenant à la résolution d'une équation de degré 3 qui a une racine réelle (le problème est alors de choisir "la bonne" racine).

Le principe de fonctionnement de tous les détecteurs structuraux est similaire à celui décrit ci-dessus pour le SPA. Les algorithmes Triples [80] et Quadruples [84] diffèrent respectivement par l'utilisation de triplets et de quadruplets de pixels consécutifs, là où l'algorithme du SPA utilise des couples de pixels (cela conduit à une modification des équations (2.12) - (2.13) et *in fine* de (2.14)).

### 2.3.3.2 Détecteurs statistiques

Historiquement, la première méthode de stéganalyse qui a vu le jour [85] repose sur la théorie de la décision statistique. Néanmoins, l'application de cette approche statistique dans le domaine de la stéganalyse est restée assez peu étudiée en raison de la complexité de modélisation des images. En effet, les tests statistiques d'hypothèses binaires ont pour vocation de décider si une observation est distribuée suivant une loi  $P_0$  définissant l'hypothèse  $\mathcal{H}_0$  ou bien suivant une loi  $P_1$  définissant l'hypothèse  $\mathcal{H}_1$ . Le problème clé est alors de formuler une définition statistique précise de l'hypothèse  $\mathcal{H}_0$  "l'image ne contient pas d'information cachée".

Pour comprendre les méthodes de stéganalyse exploitant la détection statistique, il nous faut tout d'abord étudier brièvement l'impact de la dissimulation d'informations par substitution de LSB.

On désigne sous le terme de *paire de valeurs* deux entiers relatifs dont le codage binaire ne diffère que par leurs LSB. Mathématiquement parlant, un couple d'entiers relatifs  $(z_1, z_2) \in \mathbb{Z}^2$  forme une paire de valeurs si et seulement si  $\lfloor z_1/2 \rfloor = \lfloor z_2/2 \rfloor$  où  $\lfloor \cdot \rfloor$  représente l'opération d'arrondi par défaut. Par exemple (0;1), (112;113) et (12;13) sont des paires de valeurs.

La définition des paires de valeurs est fondamentale pour comprendre le fonctionnement de certaines méthodes de stéganalyse adaptées à la substitution de LSB. En effet, comme illustré sur la figure 2.9, la substitution de LSB a tendance (sous certaines conditions) à égaliser les fréquences d'occurrences au sein des paires de valeurs.

**Exemple 2.1.** *Avant d'expliquer ce phénomène plus en détails, étudions, à des fins pédagogiques, un exemple trivial. Considérons le schéma de stéganographie par substitution de LSB dans un vecteur de valeurs binaires  $\mathbf{C} \in \mathbb{F}_2^N = \{0;1\}^N$ . Supposons que chacune des valeurs  $c_n$  est la réalisation d'une variable aléatoire binomiale  $\mathcal{B}(1, p_1)$  c'est à dire que pour chaque élément  $c_n$  on a  $\mathbb{P}[c_n=0] = 1-p_1$  et  $\text{Prob}[c_n=1] = p_1$ . Considérons à présent que l'on insère un message  $\mathbf{M} \in \mathbb{F}_2^L$ , avec  $L \leq N$  dont chaque élément est la réalisation d'une variable aléatoire distribuée selon une loi binomiale  $\mathcal{B}(1, p_2)$ ; les bits du message à insérer, 0 ou 1 sont alors équidistribués si  $p_2 = 1/2$ . Le stégo-objet  $\mathbf{S} \in \mathbb{F}_2^N$  sera donc au final constitué des  $L$  bits du message  $\mathbf{M}$  et de  $N-L$  bits de couverture issus de  $\mathbf{C}$ . Supposant que la probabilité d'insertion d'un bit est identique pour tous les échantillons de  $\mathbf{C}$  (i.e. que pour tout  $n \in \mathcal{I}$  on a  $\mathbb{P}[s_n=m_\ell] = L/N$ ) un simple calcul, en utilisant le théorème des probabilités totales, permet alors de montrer que, pour chaque élément du stégo-objet  $\mathbf{S}$ , on a :*

$$\mathbb{P}[s_n=1] = p_1 \left(1 - \frac{L}{N}\right) + p_2 \frac{L}{N}. \quad (2.15)$$

Il est alors aisé de constater que  $\mathbb{P}[s_n=1] \rightarrow p_2$  lorsque  $L \rightarrow N$ .

En pratique, l'exemple 2.1 est applicable aux LSB des pixels d'une image. Le message est généralement considéré comme étant constitué de bits  $\{0;1\}$  uniformément distribués, i.e. pour tout  $\ell \in \{0, \dots, L\}$ ,  $m_\ell \sim \mathcal{B}(1, 1/2)$ , et cela pour deux raisons principales. D'une part, cela correspond au cas le plus difficilement détectable (ou dans un jargon plus proche de la théorie de l'information, ce message binaire correspond au

*Les apports et les difficultés que présente l'approche statistique sont présentés plus en détails dans la section 2.4.*

cas d'entropie maximale). D'autre part, le message est supposé avoir été compressé, afin de réduire l'impact d'insertion, et/ou crypté, afin d'éviter qu'un attaquant puisse en avoir connaissance. Or les méthodes de compression sans perte de données comme les méthodes de chiffrement produisent des séquences binaires uniformément distribuées.

L'exemple 2.1 peut-être généralisé par la proposition suivante :

**Proposition 2.1.** *Sous les deux hypothèses qui sont généralement admises pour les méthodes d'insertion d'informations par substitution ou par correspondance de LSB :*

1. *la probabilité d'insertion est identique pour chacun des pixels (distribution uniforme du message dans les pixels);*
2. *chacun des  $L$  bits du message  $\mathbf{M} \in \mathbb{F}_2^L$  est distribué selon une loi binomiale  $\mathcal{B}(1, 1/2)$ .*

*Il découle alors que pour tout  $n \in \mathcal{I}$ , chacun des pixels  $s_n$  suit une loi de distribution définie par :*

$$\mathbb{P}[s_n = c_n] = 1 - \frac{R}{2} \quad \text{et} \quad \mathbb{P}[s_n = c_n + (-1)^{c_n}] = \frac{R}{2} \quad (2.16)$$

**Corollaire 2.1.** *La distribution de probabilité au sein des paires de valeurs converge vers une loi binomiale  $\mathcal{B}(1, 1/2)$  lorsque  $L \rightarrow N$ .*

Le plus important n'est pas ici la démonstration de la proposition 2.1 ou du corollaire 2.1, mais de prendre conscience de leurs significations. La propriété 2.1 est en effet cruciale pour comprendre le principe de fonctionnement des trois méthodes de stéganalyse adaptées à la substitution de LSB décrites dans ce manuscrit.

### Détecteur basé sur le test du $\chi^2$

La méthode de stéganalyse basée sur le test du  $\chi^2$ , proposée dans [85] par A. Westfeld et A. Pfitzmann, repose sur les fréquences d'occurrences des valeurs de pixels dans une image. En notant  $n_i^c$  le nombre de pixels dont l'intensité est  $i \in \mathcal{I}$  dans l'image de couverture  $C$ , il découle de la proposition 2.1, qu'après insertion d'un bit d'information dans tous les pixels de l'image  $C$ , *i.e.* lorsque  $R = 1$ , le nombre  $n_i^s$  de pixels, dans l'image stéganographiée, dont l'intensité est  $i \in \mathcal{I}$  vérifie au sein de chacune des paires de valeurs :

$$\forall 2i \in \mathcal{I}, n_{2i}^s \rightarrow \hat{n}_{2i}^* = \frac{n_{2i}^c + n_{2i+1}^c}{2} \quad \text{et} \quad n_{2i+1}^s \rightarrow \hat{n}_{2i}^*,$$

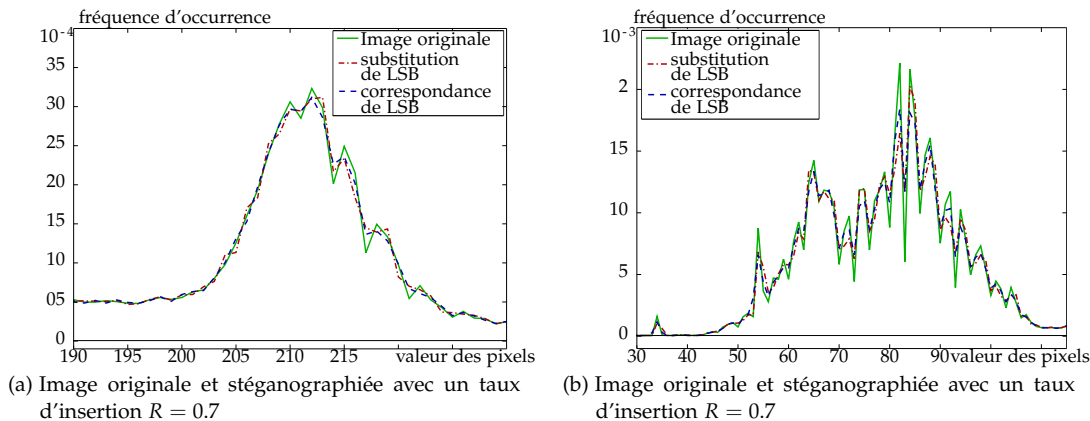


FIGURE 2.9: Fréquence d'occurrence (en ordonnée) de valeurs (abscisse) des pixels avant et après stéganographie.

lorsque le nombre total de pixels  $N \rightarrow \infty$ .

Lors de l'analyse d'une image  $\mathbf{Z}$ , le nombre  $n_i^z$  d'occurrences de pixels d'intensité  $i$  est dénombré pour tout  $i \in \mathcal{I}$  et l'indicateur statistique suivant :

$$X_{i-1}^2 = \sum_{i \in \mathcal{I}} \frac{(n_i^z - \hat{n}_{2i}^*)^2}{n_i^z}$$

est calculé. La statistique  $X_{i-1}^2$  sus-définie correspond à l'indicateur du test statistique cumulatif de Pearson qui doit asymptotiquement, lorsque le nombre de pixels  $M \times N \rightarrow \infty$ , converger en distribution vers une loi du  $\chi_{i-1}^2$  si l'image est stéganographiée. En utilisant les propriétés de la distribution du  $\chi^2$ , les auteurs proposent finalement d'utiliser la valeur :

$$p = 1 - \int_0^{X_{i-1}^2} f_{\chi_{i-1}^2}(u) du = 1 - \frac{1}{2^{i-1/2} \Gamma(i-1/2)} \int_0^{X_{i-1}^2} u^{i-1/2-1} e^{-u/2} du$$

où  $f_{\chi_{i-1}^2}$  est la densité de probabilité de la distribution du  $\chi_{i-1}^2$  définie ci-dessus.

On notera donc que l'une des particularités de ce détecteur du  $\chi^2$  est de caractériser statistiquement l'hypothèse  $\mathcal{H}_1$  "le média est stéganographié" et de tester ensuite les hypothèses suivantes :

$$\delta_{\chi^2}(\mathbf{Z}) = \begin{cases} \mathcal{H}_1 & \text{si } p \geq p_0 \Leftrightarrow X_{i-1}^2 \leq [\chi_{i-1}^2]^{-1}(1-p_0)\tau \\ \mathcal{H}_0 & \text{si } p < p_0 \Leftrightarrow X_{i-1}^2 > [\chi_{i-1}^2]^{-1}(1-p_0)\tau, \end{cases}$$

où  $[\chi_{i-1}^2]^{-1}(\cdot)$  correspond à la fonction réciproque de la densité de probabilité de la distribution du  $\chi_{i-1}^2$ . Autrement dit, c'est une façon de contourner le problème consistant à définir précisément l'hypothèse statistique  $\mathcal{H}_0$  en acceptant cette hypothèse par "défaut", lorsque  $\mathcal{H}_1$  est rejetée.

### Détection basée sur le critère du maximum a posteriori (MAP)

Une approche Bayésienne a également été proposée dans [86] par R. Chandramouli et N. Memon. Dans ce travail, les auteurs présentent deux avancées. Il est proposé, d'une part, un détecteur basé sur le critère du Maximum (de Probabilité) A Posteriori (MAP). D'autre part, une relation explicite est établie (sous certaines hypothèses simplificatrices) entre les caractéristiques du média, le nombre de bits insérés et les probabilités d'erreurs du détecteur proposé. Les auteurs suggèrent donc également d'utiliser cette relation afin de calculer la "capacité stéganographique" d'un média, c'est à dire le nombre de bits qu'il est possible d'insérer tout en assurant que les probabilités d'erreurs de détection demeurent à un certain niveau jugé sécurisé (préalablement défini par l'utilisateur).

Le principe de fonctionnement de ce détecteur repose sur l'observation que les pixels ne peuvent être modifiés que de  $\pm 1$  par substitution de LSB. Il est alors proposé de classer les pixels d'une image stéganographiée en 3 sous-ensembles disjoints définis par :

$$\begin{cases} \Lambda_1 : \{n \in \mathcal{I} | s_n = c_n + 1\} \\ \Lambda_2 : \{n \in \mathcal{I} | s_n = c_n - 1\} \\ \Lambda_3 : \{n \in \mathcal{I} | s_n = c_n\} \end{cases}$$

Les trois hypothèses suivantes sont alors envisageables pour chaque pixel :

$$\mathcal{H}_j : \{s_n \in \Lambda_j\}, \quad i = \{1, 2, 3\}$$

L'utilisation de l'approche Bayésienne (voir section A.2.2 de l'annexe A) proposée par les auteurs consiste alors à supposer que pour chacun des pixels  $s_n$ ,  $n \in \mathcal{I}$ , l'occurrence de l'hypothèse  $\mathcal{H}_i$  est une variable aléatoire dont la distribution  $Q = (q_1, q_2, q_3)$

est définie par  $\mathbb{P}[\mathcal{H}_i] = \mathbb{P}[s_n \in \Lambda_i] = q_i$ , avec bien sûr  $\sum_{i=1}^3 q_i = 1$ . L'approche Bayésienne consiste alors à définir un test  $\delta$  minimisant le critère suivant appelé coût de Bayes (voir section A.2.2) :

$$J(\delta) = \sum_{i=1}^3 q_i \alpha_i = \sum_{i=1}^3 q_i \mathbb{P}[\delta(s_n) \neq \mathcal{H}_i | \mathcal{H}_i]$$

où  $\alpha_i$  est la probabilité d'erreur du test  $\delta$  lorsque l'hypothèse  $\mathcal{H}_i$  est vérifiée :  $\alpha_i = \mathbb{P}[\delta(s_n) \neq \mathcal{H}_i | \mathcal{H}_i]$ . Pour chacun des pixels, la règle de décision de MAP qui minimise le coût de Bayes est donnée [87, 88] par :

$$\delta = \mathcal{H}_i \quad \text{si} \quad \forall i \neq j, q_i \mathbb{P}[s_n | \mathcal{H}_i] > q_j \mathbb{P}[s_n | \mathcal{H}_j]$$

Suivant cette approche Bayésienne, les auteurs supposent, pour rendre ce test exploitable en pratique, que les pixels de couverture  $c_n$  sont indépendants et identiquement distribués selon une loi normale  $\mathcal{N}(0, \sigma^2)$ . L'introduction de cet *a priori* permet d'obtenir une règle de décision assez simple dont les probabilités d'erreurs  $\alpha_i$ ,  $i = \{1, 2, 3\}$  sont établies en fonction du nombre de pixels dans chacun des ensembles  $\Lambda_i$ .

Cette approche est l'une des rares reposant sur la théorie de la décision statistique, ce qui la rend originale. Cela permet, en outre, aux auteurs de pouvoir calculer explicitement les probabilités d'erreurs à partir desquelles une mesure de la capacité stéganographique d'un média est proposée.

Les apports de l'article [86] sont cependant à contrebalancer par rapport au modèle d'image utilisé : chaque pixel est supposé suivre une loi de distribution  $\mathcal{N}(0, \sigma^2)$  qui, certes simplifie grandement les calculs, mais manque beaucoup de réalisme dans la pratique (notamment pour le cas d'une image dont les observations sont supposées être des valeurs positives). De même, la distribution de probabilité  $Q = (q_1, q_2, q_3)$  n'est pas définie ; le soin est laissé à l'utilisateur de déterminer les probabilités d'occurrence de chacune des hypothèses.

### Détection basé sur le test de Hoeffding et sur le rapport de vraisemblance (RV)

Enfin une approche intéressante est celle proposée par O. Dabeer & al. dans [89] reposant sur la théorie de la décision statistique et sur le modèle simpliste selon lequel, pour tout  $n \in \mathcal{I}$ , les pixels  $z_n$  sont indépendants et identiquement distribués (i.i.d) ; en outre, le taux d'insertion  $R$  est inconnu mais supposé plus grand qu'un taux minimal  $R_0$  à partir duquel il est souhaitable de détecter la présence d'un message :  $R_0 < R < 1$ . En notant la loi de distribution (discrète) des pixels du média de couverture sous la forme d'un vecteur  $P_0$  de taille  $2^b \times 1$ , les auteurs montrent qu'en vertu de la proposition 2.1 la distribution  $P_R$  des pixels après insertion avec un taux  $R$ , par substitution de LSB, peut être représentée par l'application linéaire  $P_R = Q_R P_0$  où  $Q_R$  est la matrice de taille  $2^b \times 2^b$  donnée par :

$$\begin{pmatrix} 1-R/2 & R/2 & 0 & 0 & 0 & \dots & \dots \\ R/2 & 1-R/2 & 0 & 0 & 0 & \dots & \dots \\ 0 & 0 & 1-R/2 & R/2 & 0 & \dots & \dots \\ 0 & 0 & R/2 & 1-R/2 & 0 & \dots & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \dots \\ 0 & 0 & \dots & \dots & 0 & 1-R/2 & R/2 \\ 0 & 0 & \dots & \dots & 0 & R/2 & 1-R/2 \end{pmatrix}$$

Le problème de détection de la présence d'un message caché par substitution de LSB est alors formalisé comme un test statistique entre les deux hypothèses composites (voir définition 2.2 dans l'annexe A) suivantes :

$$\begin{cases} \mathcal{H}_0 : \{z_n \sim P_0, \forall n = \{1, \dots, N\}, R = 0\} & \text{image de couverture} \\ \mathcal{H}_1 : \{z_n \sim P_R, \forall n = \{1, \dots, N\}, \forall R > R_0\} & \text{image stéganographiée.} \end{cases}$$

Les auteurs considèrent deux cas particuliers selon que la loi de distribution des pixels de l'image de couverture  $P_0$  est connue ou pas :

- Lorsque  $P_0$  est inconnu, le principe du test proposé est de caractériser  $\mathcal{P}_1$  l'ensemble des distributions de pixels d'une image stéganographiée avec un taux  $R > R_0$ . En notant  $p_Z$  la distribution des pixels empiriquement observée, l'hypothèse  $\mathcal{H}_1$  est acceptée si :

$$\min_{q \in \mathcal{P}_1} D_{KL}(p_Z, q) < \tau$$

où  $\tau$  est un seuil fixé préalablement et  $D_{KL}(p, q)$  est la distance de Kullback-Leibler entre les distributions  $p$  et  $q$ , voir définition C.1. Autrement dit, l'hypothèse  $\mathcal{H}_1$  est rejetée si, dans l'ensemble  $\mathcal{P}_1$  aucune distribution n'est "suffisamment proche" de  $p_Z$ , celle empiriquement observée, au sens de la distance de Kullback-Leibler. Ce test repose sur les résultats de W. Hoeffding [90] sur les tests asymptotiques (lorsque le nombre de pixels  $N \rightarrow \infty$ ) pour les lois multinomiales.

- Lorsque  $P_0$  est connu, les auteurs établissent théoriquement que le test statistique le plus puissant (ou optimal au sens de Neyman-Pearson, voir la définition 2.5) est donné, quel que soit le taux d'insertion, par le test statistique entre les hypothèses simples correspondant au "pire cas", celui où  $R = R_0$ . L'hypothèse  $\mathcal{H}_1$  est alors acceptée si :

$$D_{KL}(p_Z, P_{R_0}) - D_{KL}(p_Z, P_0) < \tau \quad (2.17)$$

où  $\tau$  est, ici encore, un seuil préalablement fixé permettant notamment le respect d'une contrainte sur la probabilité de fausse alarme. Autrement dit, l'hypothèse  $\mathcal{H}_1$  est désormais acceptée si la distance entre les distributions  $p_Z$  et  $P_{R_0}$  est "suffisamment petite" comparativement à la distance entre les distributions  $p_Z$  et  $P_0$ .

Dans l'article [89], ces deux approches sont présentées comme complémentaires. Selon les auteurs, un détecteur pour la stéganalyse ne connaîtra en pratique, ni parfaitement, ni aucunement la distribution  $P_0$  des pixels du média de couverture. Ainsi la performance d'un détecteur ne connaissant que partiellement la distribution  $P_0$  est nécessairement bornée par les performances établies dans les cas extrêmes étudiés.

Enfin, les auteurs se basent sur le test supposant connue la distribution  $P_0$  pour montrer la pertinence de l'approche par des résultats d'expérimentations numériques. La distribution  $P_0$  n'est pas connue *a priori* lors de l'analyse d'une image  $Z$  mais il est proposé par les auteurs 3 méthodes *ad-hoc* d'estimation de cette dernière.

Les résultats présentés dans [89] montrent que les performances du détecteur statistique proposé sont (quasi) systématiquement inférieures aux performances de détection de l'algorithme RS. Ces faibles performances peuvent s'expliquer par le modèle d'image proposé. En effet, l'analyse statistique repose sur des pixels indépendants et identiquement distribués (i.i.d). Si la supposition d'indépendance entre les pixels peut être admise, la modélisation d'une image comme la réalisation d'un champs de variables aléatoires identiquement distribuées est assez peu réaliste en pratique. Autrement dit, considérer que les pixels sont identiques revient à ne tenir compte, ni du contenu structuré de l'image, ni des propriétés de non-stationnarité du bruit d'acquisition.

Nous verrons dans les chapitres suivants que l'un comme l'autre de ces deux aspects est à prendre en compte pour permettre une détection statistique fiable et performante de la substitution de LSB.

### 2.3.3.3 Stéganalyse des images par Weighed-Stego (WS)

La méthode de stéganalyse du WS, de l'anglais *Weighed-Stego image* ou "stégo-image pondérée" en français, a été initialement proposée par J. Fridrich dans [91] pour estimer la taille du message inséré, ou de manière équivalente le taux d'insertion  $R$ . Quelques années plus tard, cette méthode a été largement étudiée et améliorée par A. Ker [92, 93] notamment pour localiser les pixels porteurs d'informations cachées [61] (voir le début de la section 2.3) ou encore pour proposer une méthode séquentielle [94] de détection.

Dans un souci de pédagogie, la présentation proposée de l'algorithme du WS (et de ses multiples variantes) est directement inspirée de [91, 92]. Pour le  $n$ -ième pixel, avec  $n \in \mathcal{I}$ , dont la valeur est  $z_n$ , on définit la valeur  $\bar{z}_n = z_n + 1 - 2\text{lsb}(z_n) = z_n + (-1)^{z_n}$  qui correspond en pratique à inverser le LSB de la valeur  $z_n$  :  $\bar{z}_n = z_n + 1$  si  $z_n$  est pair et  $\bar{z}_n = z_n - 1$  si  $z_n$  est impair. Il découle d'une analyse similaire à la proposition 2.1 que, dans un média stéganographié avec un taux  $R$  par substitution de LSB, pour tout  $n \in \mathcal{I}$  la valeur du pixel  $s_n$  suit une loi de distribution donnée par :

$$\mathbb{P}[s_n = \bar{c}_n] = \frac{R}{2} \text{ et } \mathbb{P}[s_n = c_n] = 1 - \frac{R}{2}$$

dont l'espérance est donnée par  $\mathbb{E}[s_n] = (1 - R/2)\mathbb{E}[c_n] + R/2\mathbb{E}[\bar{c}_n]$ . Ne pouvant connaître le média de couverture  $\mathbf{C}$  (avant insertion) à partir du média étudié  $\mathbf{Z}$ , l'algorithme du WS s'appuie sur un filtrage linéaire  $\hat{\mathbf{C}} = \mathcal{F}(\mathbf{S})$  qui permet d'obtenir une valeur approchée de la valeur de chacun des pixels de couverture  $\hat{c}_n$ . L'estimation du taux d'insertion  $\hat{R}$ , but initial de l'algorithme du WS, est obtenu par minimisation de la distance suivante :

$$D(R) = \sum_{n=1}^N \left( \left(1 - \frac{R}{2}\right)z_n + \frac{R}{2}\bar{z}_n - \hat{c}_n \right)^2. \quad (2.18)$$

Autrement dit on cherche une estimation  $\hat{R}$  qui correspond à celle minimisant l'écart (au sens des moindres carrés) entre la valeur estimée du pixel  $\hat{c}_n$  et l'espérance de cette valeur après insertion avec un taux  $R$ . La minimisation de la distance  $D(R)$ , telle que définie dans (2.18), peut s'obtenir en annulant la dérivée de  $D(R)$  par rapport à la variable  $R$  :

$$\begin{aligned} \frac{\partial D(R)}{\partial R} &= 2 \sum_{n=1}^N \left( \left(1 - \frac{R}{2}\right)z_n + \frac{R}{2}\bar{z}_n - \hat{c}_n \right) \left( \frac{\bar{z}_n}{2} - \frac{z_n}{2} \right) \\ &= \sum_{n=1}^N \frac{R}{2} (\bar{z}_n - z_n)^2 + (z_n - \hat{c}_n)(\bar{z}_n - z_n) = \frac{R}{2} + \sum_{n=1}^N (z_n - \hat{c}_n)(\bar{z}_n - z_n) = 0 \end{aligned} \quad (2.19)$$

En remarquant que  $\forall n \in \mathcal{I}, \forall z_n \in \mathcal{Z}, (\bar{z}_n - z_n)^2 = 1$  il est aisé de résoudre analytiquement l'équation (2.19) pour obtenir l'estimation du taux d'insertion  $\hat{R}$  par :

$$\hat{R} = \frac{1}{N} \arg \min_{R \in [0;1]} D(R) = \frac{2}{N} \sum_{n=1}^N (z_n - \hat{c}_n)(z_n - \bar{z}_n). \quad (2.20)$$

À partir de l'estimation du taux d'insertion  $\hat{R}$  telle que formulée dans (2.20), une amélioration importante était déjà proposée dans [91]. Cette correction consiste à introduire un facteur de pondération dans la distance  $D(R)$  (2.18) afin que, dans le calcul de  $\hat{R}$ , les pixels appartenant à des "zones texturées" de l'image aient une influence moindre que les pixels appartenant à des "zones uniforme" de l'image. La distance considérée devient alors :

$$D(R) = \sum_{n=1}^N w_n \left( \left(1 - \frac{R}{2}\right)z_n + \frac{R}{2}\bar{z}_n - \hat{c}_n \right)^2$$



où, afin de ne pas biaiser l'estimation  $\hat{R}$ , il est nécessaire d'avoir  $\sum_{n=1}^N w_n = 1$ , ce qui dans la pratique s'obtient en divisant le résultat final par  $\sum_{n=1}^N w_n$ . En tenant compte pour tout  $n \in \mathcal{I}$ , des facteurs de pondération  $w_n$ , les calculs précédents (2.19) permettent d'obtenir l'estimation du taux d'insertion  $\hat{R}$  par :

$$\hat{R} = \frac{1}{N} \arg \min_{R \in [0;1]} D(R) = \frac{2}{N} \sum_{n=1}^N w_n (z_n - \hat{c}_n)(z_n - \bar{z}_n). \quad (2.21)$$

Parmi les différentes versions de l'algorithme du WS qui ont été publiées, les améliorations proposées consistent principalement à modifier deux points de l'algorithme. D'une part, l'influence du filtrage linéaire utilisé pour obtenir la valeur approchée  $\hat{\mathbf{C}} = \mathcal{F}(\mathbf{Z})$  a été étudiée. Le filtrage initial proposé dans [91] consiste à calculer, pour chaque pixel  $z_n$ , la moyenne des pixels appartenant au voisinage  $\mathcal{V}_n$  défini initialement par les 4 pixels adjacents à  $z_n$ . Cette opération de filtrage peut donc s'écrire  $\hat{\mathbf{C}} = \mathbf{F} ** \mathbf{Z}$  où  $**$  représente le produit de convolution discret 2D et la matrice  $\mathbf{F}$  est définie par

$$\mathbf{F} = \frac{1}{4} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Des filtrages linéaires avec des matrices différentes (par leurs tailles, comme par leurs éléments) ont été proposés par la suite.

D'autre part, les facteurs de pondération  $w_n$  utilisés dans [91] sont définis en fonction de la variance localement calculée sur le voisinage  $\mathcal{V}_n$ , par :

$$w_n = \frac{1}{1 + \sigma_n} \quad \text{avec :} \quad \sigma_n^2 = \frac{1}{|\mathcal{V}_n|} \sum_{m \in \mathcal{V}_n} (z_m - \hat{c}_n)^2.$$

Là encore, différents voisinages  $\mathcal{V}_n$  ont été utilisés (correspondant en général au voisinage utilisé pour le calcul de  $\hat{c}_n$ ) et différents facteurs de pondération ont été proposés sous la forme  $w_n = \frac{1}{\alpha + \sigma_n}$  avec  $\alpha \geq 0$ .

Dans la littérature, il est notable que quelques publications [92, 93, 95] étudient l'algorithme du WS pour améliorer le filtrage ou les facteurs de pondération, en analysant empiriquement sur une base d'images l'influence de ces paramètres sur la qualité de l'estimation de  $R$  résultante. En revanche, aucun article sur les fondements théoriques de l'algorithme du WS, en lien avec la théorie de la décision statistique, n'a été publié avant la présentation [23, 96] de travaux issus de cette thèse ; le lien entre l'approche proposée et le WS est détaillé dans la section 3.6.

### 2.3.4 Stéganalyses adaptées à d'autres méthodes de stéganographie

La stéganographie par substitution de LSB est certes la méthode d'insertion la plus étudiée, mais elle n'est pas la seule. Aussi, pour de nombreuses méthodes de stéganographie, des détecteurs adaptés ont vu le jour. La méthode d'insertion par correspondance de LSB semble n'être qu'une modification mineure de la méthode de substitution de LSB. Dans les faits, en évitant l'asymétrie de l'insertion créée par la substitution de LSB (voir proposition 2.1), la correspondance de LSB est plus délicate à détecter ; en conséquence quelques travaux seulement ayant pour objet de détecter spécifiquement la correspondance de LSB ont été publiés [97, 98, 99] et les performances de détection obtenues sont nettement moins bonnes que celles des détecteurs spécifiques à la substitution de LSB, détaillés dans la section 2.3.3.

À l'opposé, de nombreuses méthodes de stéganographie proposant d'appliquer la correspondance de LSB aux images JPEG (voir section 2.2.2) ont fait l'objet de travaux de détection adaptés. On citera à titre d'exemple l'algorithme de stéganographie F5

proposé par A. Westfeld [52] pour lequel J. Fridrich [100] a proposée une méthode de détection reposant sur la distribution particulière des coefficients DCT, usuellement modélisée par une distribution Laplacienne [20], utilisée dans la compression JPEG. Ce jeu du chat et de la souris a été prolongé par P. Salle [46, 48] proposant d'utiliser ce modèle de loi de distribution Laplacienne afin qu'après insertion du message, la distribution des coefficients DCT respecte cette loi. Là encore, pour cette méthode de stéganographie "basée sur un modèle", un premier détecteur adapté a été proposé par R. Böhme et A. Westfeld dans [101], d'autres détecteurs adaptés ont ensuite été proposés [102, 103].

Les méthodes d'insertion par étalement de spectre ou par quantification d'index de modulation (QIM), brièvement évoquées dans la section 2.2.1, ont également fait l'objet de quelques travaux de détection [104, 105, 106]. Ces méthodes d'insertion étant davantage appliquées pour le tatouage numérique, elles sont légèrement en dehors du cadre de cette thèse. Notons toutefois, qu'à notre connaissance, aucune méthode reposant sur la théorie de la décision ou de l'estimation statistique n'a été proposée dans ce cas particulier de stéganalyse.

Ce manuscrit n'ayant pas vocation à présenter un état de l'art exhaustif, le lecteur intéressé est invité à compléter cette présentation succincte des méthodes de stéganographie et de stéganalyse par la lecture des ouvrages [18, 19, 22].

## 2.4 Apports de la théorie de la décision statistique à la stéganalyse

Comme nous l'avons vu dans la section 2.3, les méthodes de stéganalyse proposées ne reposent que rarement sur la théorie de la décision et de l'estimation statistique, à l'exception notable des travaux évoqués dans la section 2.3.3.2. Cette approche reste peu mise en avant dans la communauté *Information Hiding* alors que le cadre de la théorie de la décision et de l'estimation statistique s'applique à la stéganalyse (quantitative ou pas). Dans la pratique, l'intérêt de cette approche n'est pas nécessairement l'obtention des meilleurs résultats expérimentaux mais davantage la maîtrise des résultats ou, en d'autres termes, l'étude des statistiques utilisées pour la détection d'informations cachées. En effet, cette thèse s'inscrit dans le cadre du projet ANR RIC (recherche d'informations cachées) pour lequel la principale difficulté est le respect d'un taux préalablement fixé de fausse alarme (très faible).

La théorie de l'estimation et de la détection statistique jouant donc un rôle fondamental dans l'ensemble des travaux présentés, une introduction plus détaillée est proposée dans l'annexe A.

### 2.4.1 Motivations et notations

Le fondement de la théorie de la détection et de l'estimation statistique consiste à admettre que les observations disponibles sont entachées d'erreurs de nature aléatoire liées notamment aux imperfections des capteurs, aux erreurs de modélisation à la quantification etc, ... Pour prendre en compte ces erreurs aléatoires, les observations  $\mathbf{Z} \in \mathbb{R}^N$ ,  $\mathbf{Z} = (z_1 \dots, z_N)^T$  sont supposées, sans perte de généralité aucune, être la réalisation d'un vecteur aléatoire  $Z$  suivant une loi de distribution connue  $Z \sim P_{\theta, R}$  où  $\theta$  est un paramètre vectoriel inconnu. Rappelons que, pour tout  $n \in \mathcal{I}$ , la valeur mesurée  $z_n$  (intensité du pixel dans le cas d'une image) appartient à l'ensemble  $\mathcal{Z}$  qui constitue l'espace des observations  $\mathcal{Z}^N$ . De manière analogue, le paramètre  $\theta$  appartient, en toute généralité, à l'espace des paramètres  $\Theta$  contenant les valeurs possibles  $\theta \in \Theta \subset \mathbb{R}^p$  avec  $p < N$ . L'espace des paramètres permet de définir la famille des

lois envisagées  $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$ . Pour des raisons théoriques évidentes, il est supposé que la famille de distribution  $\mathcal{P}$  est identifiable, *i. e.* que la condition suivante est toujours vérifiée :

$$\theta_1 \neq \theta_2 \Leftrightarrow P_{\theta_1} \neq P_{\theta_2}.$$

Autrement dit, l'application  $\theta \mapsto P_\theta$  est bijective sans quoi le paramètre n'est plus nécessairement "identifiable". Il sera admis dans cette section que les observations sont continues, *i. e.*  $\mathcal{Z} \subset \mathbb{R}$ , et que la distribution  $P_\theta$  admet la densité  $f_\theta : \mathbf{Z} \mapsto f_\theta(\mathbf{Z})$  par rapport à une mesure quelconque (celle de Lebesgue par exemple).

Dans le cas de la décision statistique, l'objectif est de choisir une "classe" à laquelle appartient le paramètre  $\theta$  à partir de la réalisation  $\mathbf{Z}$ ; cela suppose que l'espace des paramètres  $\Theta$  soit partitionné en un nombre fini de sous-espaces. Dans les travaux présentés, le problème de décision statistique considéré est restreint aux cas de deux "classes" distinctes *i. e.* d'une partition binaire de l'espace  $\Theta$  :

$$\Theta = \Theta_0 \cup \Theta_1, \quad \text{avec } \Theta_0 \cap \Theta_1 = \emptyset. \tag{2.22}$$

**Définition 2.2** (hypothèse statistique). Une hypothèse statistique paramétrique désigne un ensemble  $\Theta^*$  de valeurs possibles pour le paramètre  $\theta$ . L'hypothèse est dite simple si l'ensemble  $\Theta^*$  est un singleton, *i. e.*  $\Theta^* = \{\theta^*\}$  et définit donc de façon unique la distribution du vecteur aléatoire  $\mathbf{Z} \sim P_{\theta^*}$ , dans le cas contraire l'hypothèse est dite composite.

Étant donnée la partition finie de l'espace des paramètres  $\Theta$  (2.22), l'hypothèse  $\mathcal{H}_k$  est définie par  $\mathcal{H}_k = \{\theta \in \Theta_k\}$ ; lorsque l'hypothèse  $\mathcal{H}_k$  est vérifiée, les observations  $\mathbf{Z}$  sont donc la réalisation d'un vecteur aléatoire  $\mathbf{Z} \sim P_\theta$  de densité  $f_\theta$  avec  $\theta \in \Theta_k$ .

**Définition 2.3** (test statistique d'hypothèse). On appelle test statistique entre les hypothèses  $\mathcal{H}_0, \mathcal{H}_1$  toute application surjective et mesurable  $\delta : \mathcal{Z}^N \mapsto \{\mathcal{H}_0, \mathcal{H}_1\}$

Autrement dit, la définition d'un test statistique  $\delta$  revient à partitionner l'espace des observations  $\mathcal{Z}^N$  en deux ensembles disjoints :

$$\mathcal{Z}^N = \Omega_0 \cup \Omega_1, \quad \text{avec } \Omega_k \cap \Omega_j = \emptyset$$

tels que l'hypothèse  $\mathcal{H}_k$  est acceptée si  $\mathbf{Z} \in \Omega_k$ . L'ensemble  $\Omega_k$  est appelé *région d'acceptation* de l'hypothèse  $\mathcal{H}_k$ .

Par ailleurs on notera  $\Delta$  l'ensemble des tests statistiques dont le but est de décider entre les mêmes hypothèses,  $\mathcal{H}_0, \mathcal{H}_1$  dans le cas présent.

Une région d'acceptation de l'hypothèse  $\mathcal{H}_k$  peut être une union de sous-ensembles disjoints.

### 2.4.2 Approche proposée et critère d'optimalité utilisé

Le cas particulier des tests (binaires) entre les deux hypothèses simples  $\mathcal{H}_0 = \{\theta = \theta_0\}$  et  $\mathcal{H}_1 = \{\theta = \theta_1\}$  constitue le problème de "base" de la théorie de la décision

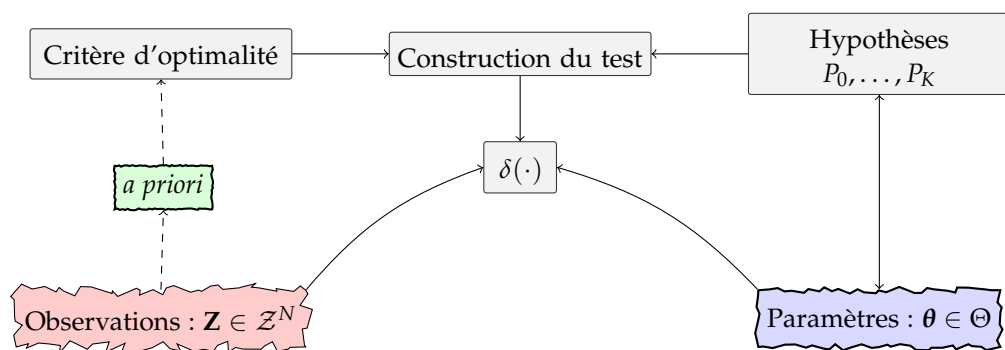


FIGURE 2.10: Illustration d'un test statistique d'hypothèses.

statistique au travers duquel est introduit l'approche proposée dans ce manuscrit. Dans ce cas, une terminologie spécifique est adoptée,  $\mathcal{H}_0$  est appelée *l'hypothèse de base* et  $\mathcal{H}_1$  la *contre-hypothèse* ou *l'hypothèse alternative*. En outre, seuls deux types d'erreurs peuvent être commis, décider  $\mathcal{H}_0$  lorsque  $\mathcal{H}_1$  est vérifiée ou inversement  $\mathcal{H}_0$  lorsque  $\mathcal{H}_1$ .

**Définition 2.4** (erreur de  $k$ -ième espèce). *On appelle erreur de  $k$ -ième espèce du test  $\delta$  le rejet de l'hypothèse  $\mathcal{H}_k$  lorsque cette dernière est vraie. La probabilité d'erreur de  $k$ -ième espèce est donc donnée par :*

$$\mathbb{P}_{\theta_k}[\delta(\mathbf{Z}) \neq \mathcal{H}_k] = \alpha_0(\delta) \quad (2.23)$$

Dans la terminologie usuelle,  $\alpha_0(\delta)$  est appelée *probabilité de fausse alarme*,  $\alpha_1(\delta)$  est appelée *probabilité de non-détection* et  $\beta(\delta) = 1 - \alpha_1(\delta)$  est appelée *puissance du test  $\delta$* .

*Dans le jargon médical notamment les termes faux-positif et faux-négatif remplacent, respectivement, ceux de fausse alarme et non-détection*

Comme illustré dans la figure 2.10, la notion même d'optimalité d'un test dépend de la définition d'un critère qui doit être optimisé; les trois principales approches, sont présentées plus en détails dans l'annexe A. Dans le présent manuscrit, l'approche utilisée est celle "de Neyman-Pearson" qui se fonde sur les critères  $\alpha_0(\delta)$  et  $\beta(\delta)$ , respectivement la probabilité de fausse alarme et la puissance. Le principe de cette approche est de fixer la probabilité maximale de fausse alarme acceptable  $\alpha_0(\delta) \leq \alpha_0$  et de maximiser la probabilité de détection  $\beta(\delta)$ . Plus formellement, la classe  $\mathcal{K}_{\alpha_0} \subset \Delta$  des tests dont la probabilité de fausse alarme est majorée par  $\alpha_0$  est définie par :

$$\mathcal{K}_{\alpha_0} = \{\delta \in \Delta \mid \alpha_0(\delta) \leq \alpha_0\} \quad (2.24)$$

**Définition 2.5.** *Un test  $\delta^* \in \mathcal{K}_{\alpha_0}$  est le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  (ou optimal au sens de Neyman-Pearson) si :*

$$\forall \delta \in \mathcal{K}_{\alpha_0}, \beta(\delta^*) \geq \beta(\delta)$$

Le principal intérêt de cette approche réside dans la maîtrise de la probabilité de fausse alarme. En effet, un utilisateur peut exiger un niveau maximal de probabilité de fausse alarme acceptable, le test le plus puissant permet de garantir simultanément le respect de cette probabilité de fausse alarme et la plus grande puissance de détection. Dans la pratique, la construction d'un test le plus puissant est donnée par le lemme de Neyman-Pearson :

**Lemme 2.1** (de Neyman-Pearson). *Soient données deux distributions  $P_{\theta_0}$  et  $P_{\theta_1}$  admettant respectivement la densité de probabilité  $f_{\theta_0}$  et  $f_{\theta_1}$ . Pour  $\alpha_0 \in ]0; 1[$ , le test le plus puissant de la classe  $\mathcal{K}_{\alpha_0}$  pour décider entre les hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$  est le test  $\delta^*$  du rapport de vraisemblance défini par :*

$$\delta^*(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \Lambda(\mathbf{Z}) = \frac{f_{\theta_1}(\mathbf{Z})}{f_{\theta_0}(\mathbf{Z})} < \tau \\ \mathcal{H}_1 & \text{sinon} \end{cases} \quad (2.25)$$

où  $\tau$  est un seuil fixé afin que  $\alpha_0(\delta^*) = \alpha_0$  i. e. est la solution de l'équation

$$\mathbb{P}_{\theta_1}[\delta^*(\mathbf{Z}) \geq \tau] = \alpha_0.$$

On constatera par ailleurs que la modification de la probabilité maximale de fausse alarme acceptable  $\alpha_0$  ne change en rien la structure du test, seul le seuil  $\tau$  change.

**Remarque 2.3.** *Si les observations sont indépendantes, il est évident que  $f_{\theta_0}(\mathbf{Z}) = \prod_{n=1}^N f_{\theta_0}(z_n)$  et de la même façon  $f_{\theta_1}(\mathbf{Z}) = \prod_{n=1}^N f_{\theta_1}(z_n)$ . Le rapport de vraisemblance qui intervient dans (2.24) s'écrit alors :*

$$\Lambda(\mathbf{Z}) = \frac{f_{\theta_1}(\mathbf{Z})}{f_{\theta_0}(\mathbf{Z})} = \prod_{n=1}^N \frac{f_{\theta_1}(z_n)}{f_{\theta_0}(z_n)} = \prod_{n=1}^N \Lambda(z_n).$$

*L'approche bi-critère "de Neyman-Pearson" est uniquement défini lorsqu'il s'agit de décider entre deux hypothèses simples.*

Dans le cas où au moins une des deux hypothèses  $\mathcal{H}_0, \mathcal{H}_1$  est composite, de nombreuses difficultés apparaissent rendant l'utilisation de l'approche bi-critère de "Neyman-Pearson" plus délicate. Tout d'abord, la probabilité d'erreur de  $k$ -ième espèce doit être redéfinie en conséquence :

**Définition 2.6** (risque de  $k$ -ième espèce). *On appelle risque de  $k$ -ième espèce d'un test  $\delta$  entre des hypothèses composites la quantité :*

$$\alpha_k(\delta) = \sup_{\theta_k \in \Theta_k} \mathbb{P}_{\theta_k}[\delta(\mathbf{Z}) \neq \mathcal{H}_k].$$

*Autrement dit, la notion de probabilité d'erreur de  $k$ -ième espèce (voir définition 2.4) est remplacée dans le cas d'un test entre hypothèses composites par la probabilité maximale de rejeter l'hypothèse  $\mathcal{H}_k$  lorsque cette dernière est vraie.*

Dans le cas d'un test entre deux hypothèses composites, la classe  $\mathcal{K}_{\alpha_0} = \{\delta \in \Delta \mid \alpha_0(\delta) \leq \alpha_0\}$  définit désormais la classe des tests dont la probabilité de fausse alarme est majorée par  $\alpha_0$  quel que soit  $\theta_0 \in \Theta_0$ . La fonction de puissance d'un test entre deux hypothèses composites est notée  $\beta(\theta) = \beta(\theta; \delta)$ , puisque cette dernière dépend désormais du paramètre  $\theta$ , et est définie par  $\beta(\theta) = \mathbb{P}_{\theta \in \Theta_1}[\delta(\mathbf{Z}) = \mathcal{H}_1]$ .

Dans le cas d'hypothèses composites, le pendant du test le plus puissant (ou optimal au sens de Neyman-Pearson, voir définition 2.5), est le test uniformément le plus puissant (UPP).

**Définition 2.7** (Test uniformément le plus puissant (UPP) entre deux hypothèses). *Un test  $\delta^* \in \mathcal{K}_{\alpha_0}$  est uniformément le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  si :*

$$\forall \delta \in \mathcal{K}_{\alpha_0}, \forall \theta_1 \in \Theta_1, \beta(\theta_1; \delta^*) \geq \beta(\theta_1; \delta)$$

*ou de manière équivalente si :*

$$\forall \delta \in \mathcal{K}_{\alpha_0}, \sup_{\theta_1 \in \Theta_1} [\beta(\theta_1; \delta) - \beta(\theta_1; \delta^*)] \leq 0$$

*Autrement dit, un test est uniformément le plus puissant dans  $\mathcal{K}_{\alpha_0}$  s'il permet de garantir un taux de fausse alarme majoré par  $\alpha_0$  (quelque soit  $\theta_0 \in \Theta_0$ ) et assure pour tout  $\theta_1 \in \Theta_1$  la meilleure puissance de détection. L'illustration de la fonction de puissance d'un test UPP est donnée dans la figure A.3.*

L'obtention d'un test uniformément le plus puissant est assez difficile en pratique et n'est pas certaine dans le cadre de la détection d'informations cachées. Une solution alternative satisfaisante, notamment dans le cadre applicatif de la détection d'informations cachées dans les images, est de définir un test qui soit asymptotiquement le plus puissant.

**Définition 2.8** (Test asymptotiquement uniformément le plus puissant (AUPP) entre deux hypothèses [107]). *Soit une séquence de variables aléatoires,  $X_1, \dots, X_n$ . Un test  $\delta_n^* \in \mathcal{K}_{\alpha_0}$  est asymptotiquement uniformément le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  si pour tout  $\delta_n \in \mathcal{K}_{\alpha_0}$  on a :*

$$\limsup_{n \rightarrow \infty} \sup_{\theta_1 \in \Theta_1} [\beta(\theta_1; \delta_n) - \beta(\theta_1; \delta_n^*)] \leq 0$$

Néanmoins, construire un test AUPP ou démontrer qu'un test est AUPP n'est en général pas simple. De manière informelle, il est notable que lorsque les hypothèses sont composites, les performances d'un test sont amoindries lorsque le paramètre est "proche de la frontière séparant les hypothèses". Une approche intéressante consiste alors à rechercher un test qui soit AUPP uniquement dans un voisinage de la frontière séparant les hypothèses.

**Définition 2.9** (Test localement asymptotiquement uniformément le plus puissant (LAUPP) entre deux hypothèses [107]). *Soit une séquence de variables aléatoires,  $X_1, \dots, X_n$  et soit le paramètre de distribution  $\theta \in \mathbb{R}$ . Pour tester les hypothèses  $\mathcal{H}_0 = \{\theta \leq \theta^*\}$  contre  $\mathcal{H}_1 = \{\theta > \theta^*\}$ , le test  $\delta_n^* \in \mathcal{K}_{\alpha_0}$  est localement asymptotiquement uniformément le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  si pour tout  $\delta_n \in \mathcal{K}_{\alpha_0}$  on a :*

$$\limsup_{n \rightarrow \infty} \sup_{\theta_1 \in \Theta_1 | 0 < n^{1/2}(\theta_1 - \theta^*) \leq c} [\beta(\theta_1; \delta_n) - \beta(\theta_1; \delta_n^*)] \leq 0,$$

pour tout  $c > 0$ .

De manière générale, l'état de l'art sur la théorie de la décision est tel qu'il n'est pas toujours possible de définir un test qui soit optimal. Néanmoins, le test du rapport de vraisemblance généralisé est souvent utilisé en théorie de la décision car des résultats établissent, dans de nombreux cas, les propriétés d'optimalité de ce test [87, 108, 109].

**Définition 2.10** (Test du rapport de vraisemblance généralisé (RVG) entre deux hypothèses composites). *Soit deux hypothèses composites  $\mathcal{H}_0$  et  $\mathcal{H}_1$ , le rapport de vraisemblance généralisé entre ces deux hypothèses est défini par :*

$$\Lambda^{\text{rvg}}(\mathbf{Z}) = \frac{\sup_{\theta \in \Theta_1} f_{\theta_1}(\mathbf{Z})}{\sup_{\theta \in \Theta_0} f_{\theta_0}(\mathbf{Z})}$$

On appelle test du rapport vraisemblance généralisé le test  $\delta^{\text{rvg}}$  défini par la règle de décision suivante :

$$\delta^{\text{rvg}}(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \Lambda^{\text{rvg}}(\mathbf{Z}) < \tau \\ \mathcal{H}_1 & \text{sinon,} \end{cases}$$

où le seuil de décision  $\tau$  est fixé de façon à ce que  $\delta^{\text{rvg}} \in \mathcal{K}_{\alpha_0}$ , i. e. est la solution de l'équation  $\sup_{\theta_0 \in \Theta_0} \mathbb{P}[\Lambda^{\text{rvg}}(\mathbf{Z}) \geq \tau] = \alpha_0$

L'idée est alors, en quelque sorte, de choisir pour une observation  $\mathbf{Z}$  le paramètre le plus vraisemblable sous chacune des deux hypothèses. Le test du rapport de vraisemblance généralisé consiste ensuite à tester ces deux hypothèses de façon analogue à ce qui a été décrit pour le cas d'hypothèses simples (voir le lemme 2.1 de Neyman-Pearson).

### 2.4.3 Détection en présence de paramètres de nuisance

Dans [89], les auteurs proposent de modéliser les observations comme  $Z \sim P_R$  où  $R$  est le taux d'insertion, voir également section 2.3.3.2. En pratique, cette approche est rendue délicate, notamment dans le cadre des images, car elle suppose que la distribution statistique de chacun des pixels soit connue, au taux d'insertion près.

Une approche plus réaliste consiste à considérer que les observations  $\mathbf{Z}$  sont la réalisation d'une variable aléatoire  $Z$  distribuée suivant la loi  $P_{\theta, \eta}$  où, comme précédemment,  $\theta \in \Theta \subset \mathbb{R}^p$  sont des *paramètres informatifs* (ou *paramètres d'intérêts* pour la détection d'informations cachées) et  $\eta \in Y \subset \mathbb{R}^m$  sont des *paramètres "de nuisance"*. La prise en compte de ces paramètres est fondamentale puisqu'ils interviennent dans la définition de la famille de distribution paramétrique  $\mathcal{P} = \{P_{\theta, \eta}; \theta \in \Theta, \eta \in Y\}$  mais ne présentent pas d'intérêt pour proposer une estimation  $\hat{\theta}$  de  $\theta$  ou pour décider en faveur d'une des deux hypothèses  $\mathcal{H}_k = \{\theta \in \Theta_k, \eta \in Y\}$ .

De manière générale, la présence d'un paramètre de nuisance est délicate, notamment lorsque la dimension de ce dernier est grande, en raison des inférences éventuelles entre paramètres de nuisance et informatifs. Pour solutionner rigoureusement un tel problème, plusieurs approches peuvent être envisagées. Celle proposée dans le cadre de ces travaux présente une certaine analogie avec la théorie de l'invariance statistique brièvement présentée dans la section 2.4.3 de l'annexe A, voir également [107, chap.6],[87, chap.47] pour une explication approfondie. La méthodologie proposée repose sur la modélisation du paramètre nuisance pour "éliminer" ce dernier.

*En toute rigueur, le domaine de définition du paramètre de nuisance peut varier selon l'hypothèse considérée.*

## 2.4.4 Apprentissage supervisé pour la décision statistique

Un rapide état de l'art des méthodologies récemment publiées dans la littérature montre que les approches issues de l'apprentissage supervisé sont très populaires lorsqu'il s'agit d'appliquer la théorie de la détection statistique aux images. Comme cela a déjà été brièvement évoqué dans la section 2.3.1, l'intérêt principal de ces méthodes réside dans leur caractère très général ; de nombreux problèmes de décision peuvent être correctement résolus par l'apprentissage alors qu'une modélisation statistique précise de ces derniers est parfois difficile.

L'apprentissage supervisé s'oppose à l'apprentissage non-supervisé dans le sens où le classifieur dispose d'une base d'apprentissage pour laquelle les classes de chacun des éléments sont connues ; ; aussi la base d'apprentissage  $\mathcal{B}_a$  pour un problème d'apprentissage supervisé peut s'écrire sous la forme de  $N_e$  couples  $(\mathbf{x}_i; c_i) \in \mathcal{X} \times \mathcal{C}$  où  $i \in \{1, \dots, N_e\}$ ,  $\mathbf{x}_i \in \mathcal{X} \subset \mathbb{R}^Q$  représente les  $Q$  caractéristiques extraites du  $i$ -ième média et  $c_i$  représente la classe à laquelle appartient ce dernier. La sélection des caractéristiques jugées pertinentes pour la discrimination des classes est généralement le problème le plus épineux et sort du cadre des travaux abordés dans ce manuscrit (une discussion à ce sujet pourra être consultée dans [110]).

*Dans le cadre de l'apprentissage non-supervisé seules les caractéristiques  $\mathbf{x}_i$  sont connues et il faut alors en plus identifier une structure de classes dans ces données.*

Le fondement théorique des méthodes de décision statistique basées sur l'apprentissage supervisé est décrit dans la section 2.4.3 de l'annexe A. Dans le cadre des travaux présentés, il a déjà été précisé que l'une des difficultés réside dans la construction d'un test permettant de respecter un niveau  $\alpha_0$  très faible de probabilité de fausse alarme. Or, l'approche par apprentissage supervisé est fondée sur la recherche d'un plan séparant les régions d'acceptation par minimisation du critère d'erreur de classification suivant :  $\alpha_0(\delta) + \alpha_1(\delta)$ . Si les fondements théoriques de cette approche sont intéressants, il semble difficile dans une telle démarche de caractériser statistiquement les performances d'un test fondé sur les méthodes d'apprentissage supervisé. Aussi, bien que quelques résultats préliminaires aient été avancés dans [111, 112] notamment, garantir le respect d'une probabilité de fausse alarme demeure encore hors de portée dans ce contexte.

## 2.4.5 Applications à la stéganographie

Comme cela a été brièvement évoqué dans l'introduction de cette section, il a été supposé que les observations sont décrites par une loi de distribution continue, *i. e.*  $\mathcal{Z}$  est un sous-ensemble compact de  $\mathbb{R}$ , et les principaux résultats issus de la théorie de la décision et de l'estimation statistique qui ont été exposés dans cette section 2.4 reposent sur cette supposition. Or, en pratique, les médias numériques sont sujets à une conversion analogique-numérique (CAN), les échantillons sont donc quantifiés (de façon individuelle) préalablement à l'enregistrement d'un média. Aussi, le média de couverture  $\mathbf{c}$  est modélisé par une loi de distribution discrète, avec pour tout  $n \in \mathcal{I}$  sans perte de généralité  $c_n \in \mathcal{Z} \subset \mathbb{Z}$ . La quantification étant uniforme pour la très grande majorité des médias, il est donc possible de décrire la valeur de chacun des échantillons  $c_n = Q_\Delta(y_n)$  sous la forme :

$$\begin{aligned} Q_\Delta(y_n) = 0 &\Leftrightarrow y_n < \Delta/2 \quad ; \quad Q_\Delta(y_n) = p-1 \Leftrightarrow y_n \geq \Delta(p-3/2) \\ \forall k \in \mathcal{Z}, k \neq \{0; p-1\}, Q_\Delta(y_n) = k &\Leftrightarrow y_n \in [\Delta(k-1/2); \Delta(k+1/2)] \end{aligned} \quad (2.26)$$

où  $y_n \in \mathbb{R}$  est la valeur du  $n$ -ième échantillon avant quantification et  $Q_\Delta(\cdot)$  représente l'opérateur de quantification de pas  $\Delta$ .

Dans bien des cas, l'anomalie est d'intensité bien supérieure au pas de quantification et étendue spatialement sur plusieurs pixels ; la détection de ce type d'anomalie dans un média peut alors se faire en négligeant la quantification. Dans le cas particulier

“d’anomalies” résultant de la dissimulation d’informations dans les LSB, c’est précisément sur la quantification que repose le mécanisme d’insertion. Les modifications que l’on souhaite détecter sont donc d’amplitudes égales au pas de quantification et cette opération ne peut alors plus être négligée. Ainsi, il est tout à fait pertinent de s’intéresser à la stéganalyse en utilisant la théorie de la détection statistique, mais beaucoup des résultats usuels ne pourront être exploités puisque leurs cadres d’application ne sont pas vérifiés.

Avant d’étudier plus en détail le rôle prépondérant de la quantification, il nous faut formaliser le problème de détection d’informations cachées dans le cadre de la théorie de l’estimation et de la détection statistique.

Sans perte aucune de généralité, la valeur de chacun des échantillons avant quantification peut-être être exprimée :

$$y_n = \theta_n + \xi_n$$

où  $\theta = \mathbb{E}[y_n] \in \mathbb{R}$  est une valeur déterministe et  $\xi_n$  est la réalisation de la variable aléatoire continue  $\Xi_n$  de moyenne nulle et distribuée suivant une loi connue  $\Xi_n \sim T$  modélisant les bruits d’acquisition. On peut donc considérer que  $y_n$  est la réalisation d’une variable aléatoire continue  $Y_n \sim T_{\theta_n}$  où la distribution  $T_{\theta_n}$ , supposée dans un premier temps connue pour tout  $n$  de  $\mathcal{I}$ , admet la densité  $t_\theta$ .

La valeur quantifiée selon (2.26) de  $y_n$  est donc elle aussi la réalisation d’une variable aléatoire  $c_n \sim P_{\theta_n;\Delta}$  où la distribution  $P_{\theta_n;\Delta}$  admet la loi de probabilité  $p_\theta$  donnée selon (2.26) par  $\forall k \in \mathcal{Z}, k \neq \{0; p-1\}, P_{\theta_n;\Delta}[k] = \mathbb{P}[c_n=k] = \mathbb{P}[y_n \in [\Delta(k-1/2); \Delta(k-1/2)]]$  soit :

$$\begin{aligned} \forall k \in \mathcal{Z}, k \neq \{0; p-1\}, p_{\theta_n;\Delta}[k] = \mathbb{P}[c_n=k] &= \int_{\Delta(k-\frac{1}{2})}^{\Delta(k+\frac{1}{2})} t_{\theta_n}(u) du & (2.27) \\ p_{\theta_n;\Delta}[0] = \int_{-\infty}^{\frac{\Delta}{2}} t_{\theta_n}(u) du & ; \quad p_{\theta_n;\Delta}[p-1] = \int_{\Delta(p-\frac{3}{2})}^{+\infty} t_{\theta_n}(u) du . \end{aligned}$$

Comme cela a été présenté dans la section 2.3.3.2, après insertion d’information par substitution de LSB avec un taux  $R$ , chacun des échantillons  $s_n, \forall n \in \mathcal{I}$  vérifie selon (2.16) :

$$\forall k \in \mathcal{Z}, \mathbb{P}[s_n = c_n] = 1 - \frac{R}{2} \quad \text{et} \quad \mathbb{P}[s_n = c_n + (-1)^{c_n}] = \frac{R}{2} .$$

Il en découle donc que chacun des échantillons suit la loi de distribution définie par :

$$\forall k \in \mathcal{Z}, Q_{\theta_n;\Delta}^R[k] = 1 - \frac{R}{2} P_{\theta_n;\Delta}[k] + \frac{R}{2} P_{\theta_n;\Delta}[k + (-1)^k] . \quad (2.28)$$

Aussi, le problème de détection statistique d’informations cachées (par substitution LSB) consiste, selon (2.27) - (2.28), à tester les hypothèses statistiques suivantes :

$$\begin{cases} \mathcal{H}_0 = \{z_n = c_n \Rightarrow z_n \sim P_{\theta_n;\Delta}[k]; \forall n \in \mathcal{I}\} \\ \mathcal{H}_1 = \{z_n = s_n \Rightarrow z_n \sim Q_{\theta_n;\Delta}^R[k]; \forall n \in \mathcal{I}, \forall R > 0\} \end{cases} . \quad (2.29)$$

Le taux d’insertion  $R > 0$  dans la définition de l’hypothèse alternative peut être remplacé par  $R > R_0$  où  $R_0$  est un faible taux d’insertion correspondant au taux minimal qu’il est souhaitable de détecter.

Le problème de détection statistique d’informations cachées par substitution de LSB, tel que formulé dans (2.29), présente les trois difficultés majeures suivantes :

1. Il s’agit d’un problème de décision entre hypothèses composites puisque le taux d’insertion ne peut être considéré comme *a priori* connu.



2. Le paramètre  $\theta_n$  agit comme un paramètre de nuisance, ce dernier ne présentant pas d'intérêt pour la détection d'informations cachées mais intervient dans la loi de distribution caractérisant la valeur des pixels. Il est donc primordial de tenir compte du contenu structuré des images afin que la détection soit, autant que possible, indépendante du contenu. Hélas, comme détaillé dans les chapitres 4 et 5, modéliser le contenu d'une image est délicat compte-tenu de sa complexité (notamment à cause de la présence de non-stationnarités et de changements abrupts).
3. Enfin, les méthodes de décision reposant sur des observations quantifiées font l'objet de peu de travaux dans la littérature. De nombreux résultats issus de la théorie de la décision statistique ne sont plus applicables dès lors que les variables aléatoires sont discrètes.

Dans le cas simpliste où il est supposé que le taux d'insertion  $R$ , les paramètres de nuisance,  $\forall n \in \mathcal{I}$ ,  $\theta_n$  et la distribution de probabilité des pixels  $P_{\theta_n, \Delta}$  sont parfaitement connus, le problème de la détection statistique d'informations cachées n'a jamais été résolu. L'expression des probabilités d'erreurs n'ayant alors jamais été explicitée il n'est pas possible de garantir le respect d'une contrainte sur la probabilité de fausse alarme. Ce problème de décision statistique "dans un média parfaitement connu" sera donc le premier abordé dans le chapitre 3.

### Conclusions du chapitre :

Au cours du présent chapitre 2, un état de l'art critique des principales méthodes de détection d'informations cachées permet notamment de souligner que seules quelques méthodes proposées s'appuient sur la théorie de la décision statistique. Les intérêts de l'exploitation de la théorie de la décision statistique sont d'une importance cruciale compte-tenu du contexte dans lequel s'inscrivent ces travaux. Cette approche est en effet la seule permettant la conception de tests dont la probabilité de fausse alarme est théoriquement maîtrisable et dont la puissance peut être comparée avec un niveau optimal. Néanmoins, il ressort d'une rapide analyse que la quantification des observations, la présence de paramètres de nuisance, ou encore les hypothèses statistiques composites testées sont autant de verrous scientifiques empêchant une application directe de la théorie de la décision statistique à la détection d'informations cachées dans les images (ou plus largement dans les médias numériques). La levée de ces verrous pour permettre une application pratique de la théorie de la décision statistique et *in fine* pour concevoir un (ou des) test(s) répondant aux contraintes imposées dans le cadre du projet ANR RIC constitue donc le fil conducteur du reste de ce manuscrit.

---

# Détection statistique optimale de l'insertion d'information dans les LSB

---

# 3

«*Alea Jecta est.*»

— Jules César, franchissant le Rubicon

«*Dieu ne joue pas aux dès.*»

— Albert Einstein, 1879 – 1955

«*Non seulement Dieu n'existe pas,  
mais essayez de trouver un plombier le dimanche !* »

— Woody Allen, 1935 – ...

## Sommaire

---

Introduction au chapitre . . . . .	41
3.1 Décision statistique basée sur des observations quantifiées . . . . .	42
3.2 Test entre deux hypothèses simples : taux d'insertion connu . . . . .	44
3.2.1 Expression du rapport de vraisemblance . . . . .	44
3.2.2 Étude asymptotique du rapport de vraisemblance . . . . .	47
3.2.3 Performances asymptotique du test du RV . . . . .	49
3.3 Test entre deux hypothèses composites : taux d'insertion inconnu . . . . .	53
3.4 Test en présence de paramètres de nuisance : modèle régressif de l'image. . . . .	56
3.5 Détection optimale sous condition d'un faible pas de quantification . . . . .	60
3.5.1 Performance asymptotique du test entre hypothèses composites	61
3.5.2 Test en présence de paramètres de nuisance : modèle régres- sif de l'image . . . . .	65
3.6 Relation entre le test proposé et la méthode du WS . . . . .	65
3.7 Extension de la méthodologie proposée . . . . .	67
Conclusions du chapitre . . . . .	71
3.8 Annexes du chapitre 3 : démonstration du théorème 3.7 . . . . .	73

---

### Introduction au chapitre :

Dans le présent chapitre 3, les difficultés scientifiques que présente l'application de la théorie de la décision statistique à la détection d'informations ( la quantification des observations, la présence de paramètres de nuisance, ou encore des hypothèses statistiques composites) sont étudiées plus en détail. Le problème de la stéganalyse est formalisé dans la section 3.1 comme un problème de décision statistique. Dans la section 3.2, le test le plus puissant PP (ou optimal au sens de Neyman-Pearson) est étudié dans le cas où le taux d'insertion et les paramètres décrivant l'image analysée sont connus. Ce cas, qui peut paraître simpliste, met en évidence la difficulté fondamentale dans l'application des méthodes de décision statistiques lorsque les données disponibles sont quantifiées. Le cas plus complexe d'un taux d'insertion inconnu, donc d'un test statistique entre hypothèses composites, est ensuite étudié dans la section 3.3 et un test localement optimal (LAUPP) est proposé.

Une approche différente, consistant à supposer que le pas de quantification est "suffisamment petit", est étudiée dans la section 3.5, et un test asymptotiquement le plus puissant (AUPP) est proposé.

L'application pratique de ces tests est proposée dans les sections 3.4 et 3.5.2, en utilisant un modèle régressif linéaire du contenu des images. L'expression des tests obtenus présente une grande similarité avec la formulation du test du WS comme cela est mis en exergue dans la section 3.6.

Enfin, les travaux présentés pour le cas d'informations cachées par substitution de LSB sont étendus pour le cas de correspondance de LSB; une brève étude statistique est présentée et permet la conception du test le plus puissant, issu de l'approche de Neyman-Pearson.

Les principales contributions de ce chapitre, comparativement aux travaux publiés précédemment, sont :

1. L'application de la théorie de la décision statistique pour la stéganalyse.
2. La prise en compte de la quantification des données dans la formulation du test statistique optimal au sens de Neyman-Pearson et, plus important, l'étude de l'impact de la quantification sur les performances du test statistique proposé.
3. L'établissement d'un test AUPP pour la stéganographie par substitution de LSB lorsque ni le taux d'insertion ni les paramètres caractérisant l'image ne sont connus.
4. La justification théorique de la méthode du WS ainsi que l'établissement des paramètres du test (jusqu'alors fixés heuristiquement).
5. L'établissement du test PP pour la détection de stéganographie par correspondance de LSB, dont les probabilités d'erreurs sont explicitement données.

### 3.1 Décision statistique basée sur des observations quantifiées : formulation du problème

Dans bon nombre de problèmes de détection, l'application de la théorie de la décision statistique aux médias numériques est réalisée sans tenir compte du fait que les échantillons sont quantifiés [113, 114, 115]. Cette négligence s'explique par le fait que ces travaux étudient des problèmes de détection statistique dans des conditions où la quantification est présente mais négligeable devant l'événement qu'il est proposé de détecter (notamment dans le domaine du médical et/ou de la tomographie).

À l'inverse, dans le cas de l'insertion d'informations cachées dans les LSB d'un média numérique (par substitution ou par correspondance), les modifications qu'il est souhaitable de détecter sont de l'ordre du pas de quantification. Il est donc primordial de tenir compte du fait que les observations sont quantifiées, et donc que le test statistique est nécessairement basé sur des variables aléatoires discrètes. La principale difficulté n'est pas tant de formuler le test le plus puissant, puisque le lemme 2.1 de Neyman-Pearson nous indique comment doit être construit ce test, que de donner une expression analytique de l'impact attendu de la quantification sur les performances du détecteur.

Comme cela a été brièvement évoqué dans la section 2.4.5, pour tout  $n \in \mathcal{I}$  la valeur  $y_n \in \mathbb{R}_+$  du  $n$ -ième pixel du média de couverture peut être décomposée, avant quantification, sous la forme :

$$y_n = \theta_n + \xi_n$$

où  $\theta_n = \mathbb{E}[y_n]$  est une valeur déterministe représentant l'énergie lumineuse mesurée par le photo-détecteur avant enregistrement, voir section 4.1, et  $\xi_n$  est la réalisation d'une variable aléatoire  $\Xi_n$  due aux divers phénomènes stochastiques intervenant durant l'acquisition. Comme cela est mis en évidence dans la section 4.4 (décrivant en détails le processus d'acquisition des images naturelles),  $\Xi_n$  peut-être fidèlement modélisé par une variable aléatoire Gaussienne de moyenne nulle et de variance  $\sigma_n^2$  ; Il est donc considéré dans ce chapitre que  $\Xi_n \sim \mathcal{N}(0; \sigma_n^2)$ . Lors de l'analyse d'un média, les échantillons disponibles (pixels dans le cas d'une image) résultent de la quantification :

$$c_n = Q_\Delta(y_n)$$

où le quantificateur  $Q_\Delta(\cdot)$  de pas  $\Delta$  est décrit par la relation (2.26).

Il est immédiat de constater que pour tout  $y_n \in \mathbb{R}$ ,  $n \in \mathcal{I}$  et pour tout  $\Delta \in \mathbb{R}_*^+$  on a :

$$Q_\Delta(y_n) = Q_1(\Delta y_n)$$

autrement dit, la quantification de la valeur  $y_n$  en utilisant un pas  $\Delta$  est équivalente à l'application du quantificateur  $Q_1$  de pas unitaire à  $\Delta y_n$ . Il est donc considéré par la suite, dans un souci de clarté, que le quantificateur  $Q_1$  est utilisé et que les valeurs  $y_n$  sont préalablement multipliées par  $\Delta$ .

En conservant les notations précédemment introduites, la loi de probabilité du  $n$ -ième pixel est donnée par  $P_{\theta_n} = \{P_{\theta_n}[0], \dots, P_{\theta_n}[2^b - 1]\}$ . En supposant que pour tout  $n \in \mathcal{I}$ , les paramètres  $\theta_n$  et  $\sigma_n$  sont connus et en supposant que les phénomènes de saturation du quantificateur sont négligeables ( $y_n$  est toujours supposée dans la plage du quantificateur :  $\mathbb{P}[y_n < 0] + \mathbb{P}[y_n > 2^b - 1/2] \approx 0$ ), pour tout  $k \in \mathcal{Z}$  la probabilité  $P_{\theta_n}[k] = \mathbb{P}[c_n = k | \theta_n, \sigma_n]$  peut s'écrire :

$$P_{\theta_n}[k] = \frac{1}{\sigma_n} \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} \varphi\left(\frac{y - \theta_n}{\sigma_n}\right) dy = \Phi\left(\frac{k + 1/2 - \theta_n}{\sigma_n}\right) - \Phi\left(\frac{k - 1/2 - \theta_n}{\sigma_n}\right) \quad (3.1)$$

où  $\varphi(x)$  et  $\Phi(x)$  sont respectivement la densité de probabilité et la fonction de répartition d'une variable aléatoire (continue) Gaussienne centrée-réduite définies par :

$$\varphi(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \quad \text{et} \quad \Phi(y) = \int_{-\infty}^y \varphi(x) dx.$$

En vertu du théorème des valeurs intermédiaires, il est notable que pour tout  $k \in \mathcal{Z}$ , il existe  $\epsilon(k)$  tel que :

$$P_{\theta_n}[k] = \frac{1}{\sigma_n} \varphi\left(\frac{k - \theta_n}{\sigma_n} + \epsilon(k)\right). \quad (3.2)$$

Les valeurs de  $\epsilon(k) = \epsilon(k; \theta_n, \sigma_n)$ , pour  $\theta \in [126; 130]$  et pour quelques valeurs distinctes de  $\sigma_n$ , sont illustrées dans la figure 3.1.

**Assertions 3.1.** Comme cela a été évoqué dans la proposition 2.1, il est supposé que les assertions suivantes sont vérifiées :

A-1 La probabilité d'insertion est identique pour tous les pixels et vaut donc  $L/N = R$ . Par ailleurs, si le  $n$ -ième pixel est utilisé pour dissimuler le  $k$ -ième bit d'information, dont la valeur est notée  $m_k = b_s \in \{0; 1\}$ , la valeur  $s_n$  du stégo-pixel est donnée par  $s_n = c_n^{(0)} + b_s$ , où  $c_n^{(0)} = c_n - b_{n,0}$ . Cette situation permet de définir la distribution statistique de la valeur du stégo-pixel  $s_n$  après insertion avec un taux  $R$  par substitution de LSB par :

$$\begin{cases} \mathbb{P}[s_n = c_n] & = (1 - R) \\ \mathbb{P}[s_n = c_n^{(0)} + b_s] & = R. \end{cases}$$

*Les bruits d'acquisition sont alors rigoureusement modélisés comme la réalisation des variables aléatoires  $\Delta \Xi_n \sim \mathcal{N}(0; \Delta^2 \sigma_n^2)$ . La notation de la variance restera néanmoins inchangée afin de simplifier les notations.*

*Rigoureusement on a  $P_{\theta_n}[0] = \Phi\left(\frac{k+1/2-\theta_n}{\sigma_n}\right)$  et  $P_{\theta_n}[2^b-1] = 1 - \Phi\left(\frac{2^b-3/2-\theta_n}{\sigma_n}\right)$ . La relation 3.1 permet par suite de simplifier l'expression des calculs.*

*Il est notable que  $c_n^{(0)} = c_n - b_{n,0}$  correspond, en pratique, à la "mise à zéro systématique" du bit de poids faible.*

A-2 La valeur  $b_s$  d'un bit éventuellement dissimulé dans le  $n$ -pixel  $c_n$  étant inconnu, il est considéré que  $b_s$  est la réalisation d'une variable aléatoire  $B_s \sim \mathcal{B}(1, p_1) : \mathbb{P}[b_s=0] = p_0 = (1 - p_1)$  et  $\mathbb{P}[b_s=1] = p_1$ ; cela suppose en outre que les bits d'informations sont dissimulés indépendamment des pixels de couverture i.e.  $\forall c_n \in \mathcal{Z}, \mathbb{P}[b_s = i] = \mathbb{P}[b_s = i|c_n]$  avec  $i = \{0; 1\}$ . Dans la suite de ce manuscrit il est considéré que  $p_1 = 1/2$ .

Il est ici considéré que  $p_1 = 1/2$  pour les raisons évoquées dans la proposition 2.1 : le message  $\mathbf{M}$  est supposé compressé et/ou chiffré.

Les assertions A-1 et A-2 permettent l'expression, pour tout  $n \in \mathcal{I}$ , de la distribution statistique de la valeur  $s_n$  des stégo-pixels après insertion par substitution de LSB avec un taux  $R$  :

$$Q_{\theta_n}^R[s_n] = (1 - R)P_{\theta_n}[c_n] + RP_{\theta_n}[c_n^{(0)} + b_s]. \quad (3.3)$$

Le problème de détection statistique qu'il est souhaitable de résoudre, lors de l'analyse d'un média  $\mathbf{Z}$ , consiste à décider entre les deux hypothèses :

$$\mathcal{H}_0 = \{z_n \sim P_{\theta_n}, \forall n \in \mathcal{I}\} \quad \text{vs} \quad \mathcal{H}_1 = \{z_n \sim Q_{\theta_n}^R, \forall n \in \mathcal{I}\}. \quad (3.4)$$

## 3.2 Test entre deux hypothèses simples : taux d'insertion connu

### 3.2.1 Expression du rapport de vraisemblance

Dans le cas où le paramètre vectoriel  $\theta = (\theta_1, \dots, \theta_n)^T$  ainsi que le taux d'insertion  $R$  sont connus, le problème de décision statistique (3.4) est un test binaire entre deux hypothèses simples. En vertu du Lemme 2.1, le test optimal au sens de Neyman-Pearson est basé sur le rapport de vraisemblance  $\Lambda_R(\mathbf{Z})$  donné, compte-tenu de l'indépendance statistique entre les pixels, par :

$$\Lambda_R(\mathbf{Z}) = \prod_{n=1}^N \Lambda_R(z_n) = \prod_{n=0}^N \frac{Q_{\theta_n}^R[z_n]}{P_{\theta_n}[z_n]}. \quad (3.5)$$

De façon analogue à ce qui a été proposé précédemment  $z_n^{(0)} = z_n - b_{n,0}$  correspond à la "mise à zéro" du LSB de  $z_n$ . Un rapide calcul permet de montrer que pour tout  $n \in \mathcal{I}$ , le rapport de vraisemblance appliqué au  $n$ -ième pixel  $\Lambda(z_n)$  peut s'écrire :

$$\frac{Q_{\theta_n}^R[z_n]}{P_{\theta_n}[z_n]} = \frac{(1 - R)P_{\theta_n}[z_n] + RP_{\theta_n}[z_n - b_{n,0} + b_s]}{P_{\theta_n}[z_n]} = (1 - R) + R \frac{P_{\theta_n}[z_n^{(0)} + b_s]}{P_{\theta_n}[z_n]}.$$

Il est alors évident que  $\Lambda(z_n)$  ne dépend des observations  $z_n$  qu'à travers le terme :

$$\Lambda_R(z_n) = R \frac{P_{\theta_n}[z_n^{(0)} + b_s]}{P_{\theta_n}[z_n]}. \quad (3.6)$$

Ainsi, le test le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  des tests dont la probabilité de fausse alarme est majorée par  $\alpha_0$  peut s'écrire :

Ici encore, le seuil de détection  $\tau_{\alpha_0}$  est défini comme la solution de :  $\mathbb{P}_0[\Lambda_R(\mathbf{Z}) \geq \tau_{\alpha_0}] = \alpha_0$  afin que  $\delta_R \in \mathcal{K}_{\alpha_0}$ .

$$\delta_R(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \Lambda_R(\mathbf{Z}) < \tau_{\alpha_0} \\ \mathcal{H}_1 & \text{si } \Lambda_R(\mathbf{Z}) \geq \tau_{\alpha_0} \end{cases}. \quad (3.7)$$

Dans un souci de clarté, il est considéré dans un premier temps que  $R = 1$ ; dans ce cas le rapport de vraisemblance  $\Lambda_R(z_n)$  (3.6) peut donc se ré-écrire :

$$\Lambda_1(z_n) = \frac{P_{\theta_n}[z_n - b_{n,0} + b_s]}{P_{\theta_n}[z_n]}. \quad (3.8)$$

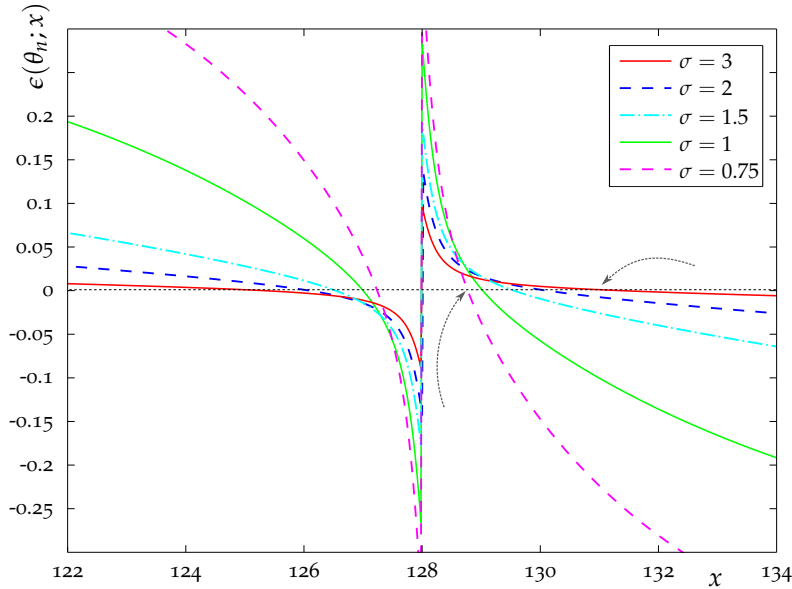


FIGURE 3.1: Graphique représentatif des valeurs du terme correctif  $\epsilon(\theta_n; x)$  pour quelques valeurs de  $\sigma$  avec  $\theta = 128$ .

En supposant que l'assertion A-2 est vérifiée, on peut immédiatement calculer :

$$P_{\theta_n}[z_n^{(0)} + b_s] = \mathbb{P}[b_s=0]P_{\theta_n}[z_n^{(0)}] + \mathbb{P}[b_s=1]P_{\theta_n}[z_n^{(0)} + 1] = \frac{1}{2} \left( P_{\theta_n}[z_n^{(0)}] + P_{\theta_n}[z_n^{(0)} + 1] \right).$$

Il est noté par la suite, pour tout  $2k \in \mathcal{Z}$ ,  $P_{\theta_n}^0[k]$  la probabilité que la variable aléatoire discrète  $z_n^{(0)}$  soit égale à l'entier pair  $k$  :  $P_{\theta_n}^0[k] = P_{\theta_n}[k] + P_{\theta_n}[k + 1]$ . Suivant (3.1) - (3.2), cette probabilité peut être exprimée par :

$$\begin{aligned} P_{\theta_n}^0[k] &= \frac{1}{\sigma_n} \int_{k-\frac{1}{2}}^{k+\frac{3}{2}} \varphi \left( \frac{y - \theta_n}{\sigma_n} \right) dy = \Phi \left( \frac{k + 3/2 - \theta_n}{\sigma_n} \right) - \Phi \left( \frac{k - 1/2 - \theta_n}{\sigma_n} \right) \\ &= \frac{1}{\sigma_n} \varphi \left( \frac{k - \theta_n}{\sigma_n} + \epsilon^{(0)}(k) \right). \end{aligned}$$

En utilisant les assertions A-1 et A-2 et les notations précédemment définies, il découle de la définition de la loi de probabilité  $P_{\theta_n}$  (3.2) que le rapport de vraisemblance  $\Lambda_1(z_n)$ , défini par la relation (3.8), peut être exprimé :

$$\begin{aligned} \Lambda_1(z_n) &= \frac{P_{\theta_n}^0[z_n]}{2P_{\theta_n}[z_n]} = \frac{1}{2} \frac{2 \exp \left( -\frac{1}{2\sigma_n^2} \left( z_n^{(0)} + 1/2 - \theta_n + \epsilon^{(0)}(z_n) \right)^2 \right)}{\exp \left( -\frac{1}{2\sigma_n^2} \left( z_n - \theta_n + \epsilon(z_n) \right)^2 \right)} \\ \Leftrightarrow \log(\Lambda_1(z_n)) &= \frac{-1}{2\sigma_n^2} \left[ \left( z_n^{(0)} + 1/2 - \theta_n + \epsilon^{(0)}(z_n) \right)^2 - \left( z_n - \theta_n + \epsilon(z_n) \right)^2 \right]. \quad (3.9) \end{aligned}$$

Le rapport de vraisemblance exact (ou le logarithme de cette quantité) tel que donné dans (3.9) est complexe à calculer en raison de la présence des termes correctifs  $\epsilon^{(0)}(z_n)$  et  $\epsilon(z_n)$  qui ne sont pas analytiquement calculables. En outre, comme l'illustre la figure 3.1, l'influence de ces termes est négligeable lorsque la variance du bruit  $\sigma_n^2$  est importante. Afin de pouvoir calculer analytiquement le rapport de vraisemblance  $\Lambda_1(\mathbf{Z})$

il est proposé de négliger les termes correctifs ; l'équation (3.9) se simplifie et devient alors :

$$\begin{aligned}\Lambda_1(z_n) &= \frac{\exp\left(-\frac{1}{2\sigma_n^2}\left(z_n^{(0)} + 1/2 - \theta_n\right)^2\right)}{\exp\left(-\frac{1}{2\sigma_n^2}\left(z_n - \theta_n\right)^2\right)} + o_p(\sigma_n^{-2}) \\ \Leftrightarrow \log(\Lambda_1(z_n)) &\approx \frac{-1}{2\sigma_n^2} \left[ \left(z_n^{(0)} + 1/2 - \theta_n\right)^2 - \left(z_n - \theta_n\right)^2 \right].\end{aligned}\quad (3.10)$$

où la notation  $x = o_p(y)$  signifie que  $x/y \rightarrow 0$  en probabilité lorsque  $y \rightarrow 0$ .

Sous l'hypothèse  $\mathcal{H}_0$ , un développement du terme quadratique  $(z_n^{(0)} + 1/2 - \theta_n)^2$  permet (en utilisant le fait que  $z_n^{(0)} = z_n - b_{n,0}$ ) immédiatement d'obtenir :

$$(z_n - \theta_n + 1/2 - b_{n,0})^2 = (z_n - \theta_n)^2 + (1/2 - b_{n,0})^2 + 2(z_n - \theta_n)(1/2 - b_{n,0}).$$

Enfin, en constatant que  $(1/2 - b_{n,0})^2 = 1/4$  quel que soit  $b_{n,0} \in \{0; 1\}$  et en négligeant le terme correctif  $\epsilon(z_n)$ , il découle que :

$$\log(\Lambda_1(\mathbf{Z})) = \sum_{n=1}^N \frac{(z_n - \theta_n)(b_{n,0} - 1/2)}{\sigma_n^2} - \frac{1}{8\sigma_n^2}.\quad (3.11)$$

À l'inverse, sous l'hypothèse  $\mathcal{H}_1$ , un développement du terme quadratique  $(z_n - \theta_n)^2$  permet (en utilisant le fait que  $z_n = z_n^{(0)} + b_s$ ) immédiatement d'obtenir :

$$(z_n^{(0)} + 1/2 - \theta_n + b_s - 1/2)^2 = (z_n^{(0)} + 1/2 - \theta_n)^2 + (b_s - 1/2)^2 + 2(z_n^{(0)} + 1/2 - \theta_n)(b_s - 1/2),$$

dont il découle que :

$$\log(\Lambda_1(\mathbf{Z})) = \sum_{n=1}^N \frac{(z_n^{(0)} + 1/2 - \theta_n)(b_s - 1/2)}{\sigma_n^2} + \frac{1}{8\sigma_n^2}.\quad (3.12)$$

**Remarque 3.1.** L'approche présentée jusqu'à présent est très similaire à celle proposée par O. Dabeer & al. dans [89] : dans les deux cas le lemme 2.1 de Neyman-Pearson est utilisé afin de donner une expression simple du test le plus puissant. Comme cela a été esquissé dans la section 2.3.3.2, les auteurs de [89] supposent que les pixels sont i.i.d. et suivent une loi de distribution  $P_0$ , si l'image ne contient pas d'information cachée, ou  $P_R$ , si l'image est stéganographiée avec un taux  $R$ . Les auteurs montrent que le test du RV peut être écrit sous la forme, voir (2.17) :

$$D_{KL}(p_{\mathbf{Z}}, P_{R_0}) - D_{KL}(p, P_0) < \tau,$$

où  $p_{\mathbf{Z}}$  est la distribution empirique du média analysé et  $D_{KL}(p, q)$  est la distance de Kullback-Leibler entre les distributions  $p$  et  $q$  (voir définition C.1).

En considérant que pour tout  $n \in \mathcal{I}$ ,  $\theta_n = \theta$  et  $\sigma_n^2 = \sigma^2$ , un rapide calcul permet d'obtenir à partir de (3.5) :

$$\log(\Lambda_R(\mathbf{Z})) = \sum_{n=1}^N \log\left(\frac{Q_{\theta_n}^R[z_n]}{P_{\theta_n}[z_n]}\right).$$

En notant  $p_{\mathbf{Z}} = (p_{\mathbf{Z}}[k], \dots, p_{\mathbf{Z}}[2^b - 1])^T$  la loi de probabilité empirique (ou fréquence d'occurrence) de  $\mathbf{Z}$ , on peut immédiatement obtenir :

$$\begin{aligned}\log(\Lambda_R(\mathbf{Z})) &= N \sum_{k \in \mathcal{I}} p_{\mathbf{Z}}[k] \log\left(\frac{Q_{\theta_n}^R[k]}{P_{\theta_n}[k]}\right) \\ &= N \sum_{k \in \mathcal{I}} p_{\mathbf{Z}}[k] \left( \log(p_{\mathbf{Z}}[k]) - \log(P_{\theta_n}[k]) - \left( \log(p_{\mathbf{Z}}[k]) \log(Q_{\theta_n}^R[k]) \right) \right) \\ &= -N \left( D_{KL}(p_{\mathbf{Z}}, Q_{\theta_n}^R[k]) - D_{KL}(p_{\mathbf{Z}}, P_{\theta_n}[k]) \right).\end{aligned}$$

chapitre On constate alors aisément que les deux approches mènent à un résultat identique. Deux différences importantes demeurent cependant. D'une part, dans l'approche présentée dans ce mémoire, les pixels ne sont pas nécessairement considérés comme i.i.d ; cette supposition est la pierre angulaire dans [89] puisque le test ne repose que sur la loi de distribution empirique  $p_{\mathbf{Z}}$ . D'autre part, l'approche des présents travaux est de considérer chaque pixel comme le résultat d'une V.A. continue puis quantifiée ; O. Dabeer & al. considèrent que les pixels sont discrets par nature.

Comme nous le verrons dans la suite, ces différences changent radicalement la nature des tests proposés.

### 3.2.2 Étude asymptotique du rapport de vraisemblance

Compte tenu du grand nombre de pixels présents dans une image à analyser, l'utilisation de résultats asymptotiques (lorsque  $N \rightarrow \infty$ ) semble pertinent ; dans ce contexte, le théorème 3.1 [107, théorème 11.2.5] est incontournable.

**Théorème 3.1** (de la limite centrale de Lindeberg). Soit  $\{X_n\}_{n=1}^N$  une suite de variables aléatoires indépendantes telles que pour tout  $n \in \{1; \dots; N\}$ , l'espérance  $m_{x_n}$  et la variance  $\sigma_{x_n}^2 \neq 0$  de  $X_n$  existent et sont finies. Alors pour tout  $(a, b) \in \mathbb{R}^2$  tels que  $a < b$  :

Les conditions de Lindeberg sont plus restrictives mais aisément vérifiables pour les pixels  $\mathbf{z}$ .

$$\mathbb{P} \left[ a < \frac{\sum_{n=1}^N X_n - m_{x_n}}{\sqrt{\sum_{n=1}^N \sigma_{x_n}^2}} < b \right] \rightarrow \Phi(b) - \Phi(a)$$

lorsque  $N \rightarrow \infty$ . Autrement dit, le théorème de la limite centrale stipule que :

$$\frac{\sum_{n=1}^N [X_n - m_{x_n}]}{\sqrt{\sum_{n=1}^N \sigma_{x_n}^2}} \rightsquigarrow \mathcal{N}(0; 1),$$

où  $\rightsquigarrow$  désigne la convergence en distribution (parfois aussi appelée convergence en loi) lorsque  $N \rightarrow \infty$ .

**Corollaire 3.1.** Dans le cas du rapport de vraisemblance  $\log(\Lambda_1(\mathbf{Z}))$ , en notant, pour  $i = \{0; 1\}$ ,  $\mathbb{E}_i[\cdot]$  et  $\text{Var}_i[\cdot]$  respectivement l'espérance et la variance sous l'hypothèse  $\mathcal{H}_i$ , un corollaire immédiat du théorème 3.1 est :

$$\frac{\sum_{n=1}^N \log(\Lambda_1(z_n)) - \mathbb{E}_i[\log(\Lambda_1(z_n))]}{\sqrt{\sum \text{Var}_i[\log(\Lambda_1(z_n))]} } \rightsquigarrow \mathcal{N}(0, 1). \quad (3.13)$$

Autrement dit, il est possible de caractériser, en vertu du théorème 3.1 de la limite centrale de Lindeberg, la convergence en distribution de  $\log(\Lambda_1(\mathbf{Z}))$ , sous chacune des hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$ . Ce résultat offre donc la possibilité de calculer les probabilités d'erreurs du détecteur pour un seuil de décision  $\tau$  donné et, réciproquement, permet de calculer la valeur du seuil de détection permettant de respecter une contrainte sur la probabilité de fausse alarme. Néanmoins, pour appliquer le corollaire (3.1) il est nécessaire de pouvoir exprimer, pour  $i = \{0; 1\}$ , les espérances mathématiques  $\mathbb{E}_i[\log(\Lambda_1(z_n))]$  et les variances  $\text{Var}_i[\log(\Lambda_1(z_n))]$ . La principale difficulté est alors de prendre en compte l'impact de la quantification des variables  $z_n$  dans le calcul du rapport du vraisemblance  $\Lambda_1(\mathbf{Z})$  pour *in fine* donner une expression analytique de cet impact sur les performances du test proposé.

Sous l'hypothèse  $\mathcal{H}_0$ , un calcul direct montre que l'espérance mathématique de la log vraisemblance (en omettant ici encore les termes correctifs) est donnée pour tout  $n \in \mathcal{I}$  par :

$$\mu_0 = \mathbb{E}_0[\log(\Lambda_1(z_n))] = \frac{-1}{8\sigma_n^2} + \frac{m_0}{\sigma_n^2}, \quad (3.14)$$



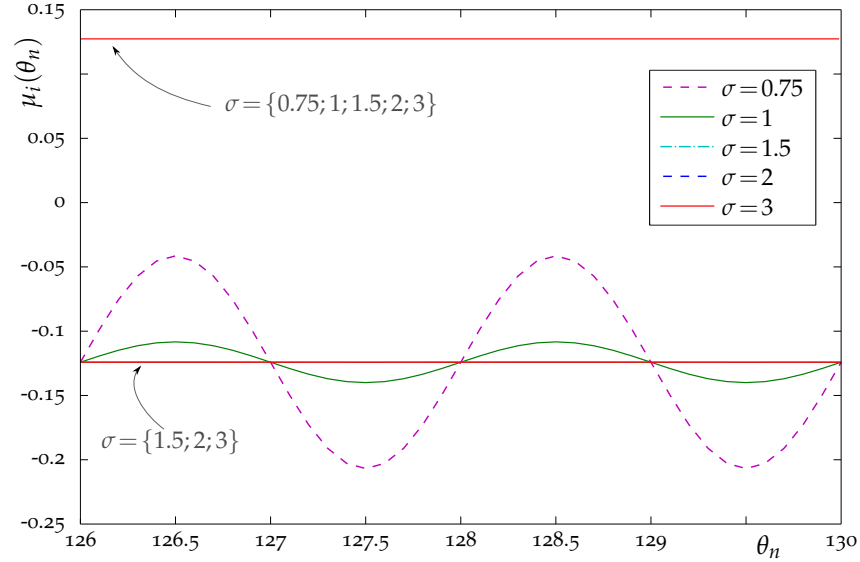


FIGURE 3.2: Représentations graphiques des moyennes  $\mu_0$ ,  $\mu_1$  et de leurs approximations, en négligeant la quantification,  $\widehat{\mu}_0$ ,  $\widehat{\mu}_1$  en fonction de  $\theta_n \in [126; 130]$  et pour  $\sigma_n = \{0.75; 1; 1.5; 2; 3\}$ . Dans un souci de clarté, les valeurs sont multipliées par  $\sigma_n^2$  afin que  $\widehat{\mu}_0 = -1/8$ .

où le terme  $m_0$ , permettant d'évaluer l'impact de la quantification sur le rapport de vraisemblance (voir ci-dessous), peut s'écrire :

$$\begin{aligned} m_0 &= \sum_{k \in \mathcal{Z}} \mathbb{P}[z_n = k] (k - \theta_n) (b_{n,0} - 1/2) \\ &= \sum_{2k \in \mathcal{Z}} \left( \Phi\left(\frac{k + 3/2 - \theta_n}{\sigma}\right) - \Phi\left(\frac{k + 1/2 - \theta_n}{\sigma}\right) \right) \frac{2k + 1 - \theta_n}{2} \\ &\quad - \sum_{2k \in \mathcal{Z}} \left( \Phi\left(\frac{k + 1/2 - \theta_n}{\sigma}\right) - \Phi\left(\frac{k - 1/2 - \theta_n}{\sigma}\right) \right) \frac{2k - \theta_n}{2}. \end{aligned}$$

De même la variance de la log vraisemblance est donnée, sous  $\mathcal{H}_0$ , par :

$$\begin{aligned} \sigma_0^2 &= \mathbb{E}_0 \left[ \log (\Lambda_1(z_n))^2 \right] - \mu_0^2 = \sum_{k \in \mathcal{Z}} \frac{\mathbb{P}[z_n = k]}{\sigma_n^4} \left( (k - \theta_n) (b_{n,0} - 1/2) - \frac{1}{8} \right)^2 - \mu_0^2 \\ &= \sum_{k \in \mathcal{Z}} \frac{\mathbb{P}[z_n = k]}{\sigma_n^4} \left( \frac{(k - \theta)^2}{4} - \frac{1}{4} (k - \theta) (b_{n,0} - 1/2) + \frac{1}{8^2} \right) - \frac{1}{8^2 \sigma_n^4} + \frac{m_0^2}{\sigma_n^4} - \frac{m_0}{4\sigma_n^4} \\ \sigma_0^2 &= \frac{\mathbb{E}_0[(z_n - \theta_n)^2] - 4m_0^2}{4\sigma_n^4}. \end{aligned} \quad (3.15)$$

De façon analogue, sous l'hypothèse  $\mathcal{H}_1$ , en considérant ici encore que les assertions A-1 et A-2 sont vérifiées, un calcul direct permet d'obtenir :

$$\begin{aligned} &\sum_{2k \in \mathcal{Z}} \mathbb{P}[z_n^{(0)} = k] ((k + 1/2 - \theta_n) (b_s - 1/2)) \\ &= \sum_{2k \in \mathcal{Z}} \mathbb{P}[z_n^{(0)} = k] (k + 1/2 - \theta_n) ((-1/2)\mathbb{P}[b_s = 0] + 1/2\mathbb{P}[b_s = 1]) = 0, \end{aligned}$$

dont il découle immédiatement l'espérance mathématique de la log vraisemblance :

$$\mu_1 = \mathbb{E}_1 \left[ \log (\Lambda_1(z_n)) \right] = \frac{1}{8\sigma_n^2}. \quad (3.16)$$

Il est notable que  $\mu_0 = -D_{\text{KL}}(P_{\theta_n}, P_{\theta_n}^0)$  et que  $\mu_1 = D_{\text{KL}}(P_{\theta_n}^0, P_{\theta_n})$  : la non-symétrie de la distance de Kullback-Leibler apparaît ici clairement.

Enfin, la variance de la log vraisemblance est donnée, sous l'hypothèse  $\mathcal{H}_1$  par :

$$\begin{aligned}
\sigma_1^2 &= \mathbb{E}_1 \left[ \log (\Lambda_1(z_n))^2 \right] - \mu_1^2 = \sum_{2k \in \mathcal{Z}} \frac{\mathbb{P}_1[z_n^{(0)}=k]}{\sigma_n^4} \left( (k+1/2-\theta_n)(b_s-1/2) + \frac{1}{8} \right)^2 - \frac{1}{8^2 \sigma_n^4} \\
&= \sum_{2k \in \mathcal{Z}} \frac{\mathbb{P}_1[z_n^{(0)}=k]}{\sigma_n^4} \left( \frac{(k+1/2-\theta_n)^2}{4} + \frac{1}{8^2} + \frac{1}{4}(k+1/2-\theta_n)(b_s-1/2) \right) - \frac{1}{8^2 \sigma_n^4} \\
&= \sum_{2k \in \mathcal{Z}} \left( \Phi \left( \frac{k+3/2-\theta_n}{\sigma} \right) - \Phi \left( \frac{k-1/2-\theta_n}{\sigma} \right) \right) \frac{(k+1/2-\theta_n)^2}{4\sigma_n^4}. \quad (3.17)
\end{aligned}$$

### Négligence de la quantification : impact sur le rapport de vraisemblance

Il est possible de simplifier le calcul des moments (3.14) - (3.17) en considérant que les distributions  $P_{\theta_n}$  et  $P_{\theta_n}^0$  sont symétriques par rapport à la moyenne  $\theta_n$ . De manière plus formelle cette considération revient à faire l'hypothèse forte que pour  $k \in \mathcal{N}$ ,  $\mathbb{P}[z_n = \theta - k + 1] = \mathbb{P}[z_n = \theta + k]$ . Ne pas tenir compte de l'impact de la quantification dans le calcul du rapport de vraisemblance revient à négliger (de façon plus ou moins erronée) le premier terme des équations (3.11) et (3.12). Plus précisément, un rapide calcul permet immédiatement d'obtenir, dans le cas où  $z_n$  est supposé continu, que :

$$\begin{aligned}
\sum_{k \in \mathcal{Z}} \mathbb{P}[z_n = k](k - \theta_n)(b_{n,0} - 1/2) &= 0 = \sum_{2k \in \mathcal{Z}} \mathbb{P}[z_n^{(0)} = k](k + 1/2 - \theta_n)(b_s - 1/2) \\
\sum_{k \in \mathcal{Z}} \mathbb{P}[z_n = k](k - \theta_n)^2 &= \sigma_n^2 = \sum_{2k \in \mathcal{Z}} \mathbb{P}[z_n^{(0)} = k](k + 1/2 - \theta_n)^2 \quad (3.18)
\end{aligned}$$

*Les notations  $b_{n,0}$  et  $z_n^{(0)}$  sont conservées inchangées afin de ne pas alourdir les notations.*

Il découle de l'utilisation des simplifications (3.18) qu'en négligeant la quantification dans les calculs (3.14) - (3.17) les deux premiers moments du rapport de vraisemblance  $\log (\Lambda_1(\mathbf{Z}))$  sont donnés par :

$$\begin{cases} \widehat{\mu}_0 = \frac{-1}{8\sigma_n^2} \text{ et } \widehat{\mu}_1 = \frac{1}{8\sigma_n^2} \\ \widehat{\sigma}_0^2 = \frac{1}{4\sigma_n^2} = \widehat{\sigma}_0^2 \end{cases} \quad (3.19)$$

Les figures 3.2 et 3.3 illustrent l'impact de la prise en compte, ou non, de la quantification sur les valeurs des deux premiers moments du rapport de vraisemblance  $\log (\Lambda_1(\mathbf{Z}))$  (respectivement  $\mu_i$ ,  $\sigma_i$ ,  $\widehat{\mu}_0$  et  $\widehat{\sigma}_0^2$ ) pour  $\theta_n \in [126; 130]$  et pour quelques valeurs de  $\sigma_n$ .

### 3.2.3 Performances asymptotique du test du RV

Les calculs précédents ont été réalisés en considérant que  $R = 1$ , comme le montre les relations (3.5) - (3.8). Avant de pouvoir exprimer analytiquement l'impact de la quantification sur les performances du test du RV proposé, il est nécessaire d'élargir les résultats précédents aux cas d'un taux d'insertion quelconque, *i.e.*  $R \in ]0; 1[$ . L'approche proposée est, dans la continuité des résultats précédents, de "supprimer" le bit de poids faible afin que l'insertion d'informations cachées n'ait aucune inférence. Le calcul du rapport de vraisemblance  $\Lambda_R(z_n)$  (3.6), pour tout taux d'insertion  $R \in ]0; 1[$  est donné par :

$$\Lambda_R(z_n) = R \frac{P_{\theta_n}^0[z_n]}{2P_{\theta_n}[z_n]}.$$

Il découle du théorème des probabilités totales que :

$$\begin{aligned}
\mu_R &= \mathbb{E}_R [\log (\Lambda_R(z_n))] = \mathbb{P}[z_n=c_n] \mathbb{E}_0 [\log (\Lambda_R(z_n))] + \mathbb{P}[z_n=s_n] \mathbb{E}_1 [\log (\Lambda_R(z_n))] \\
&\Rightarrow \mu_R = (1 - R)\mu_0 + R\mu_1 \quad (3.20)
\end{aligned}$$

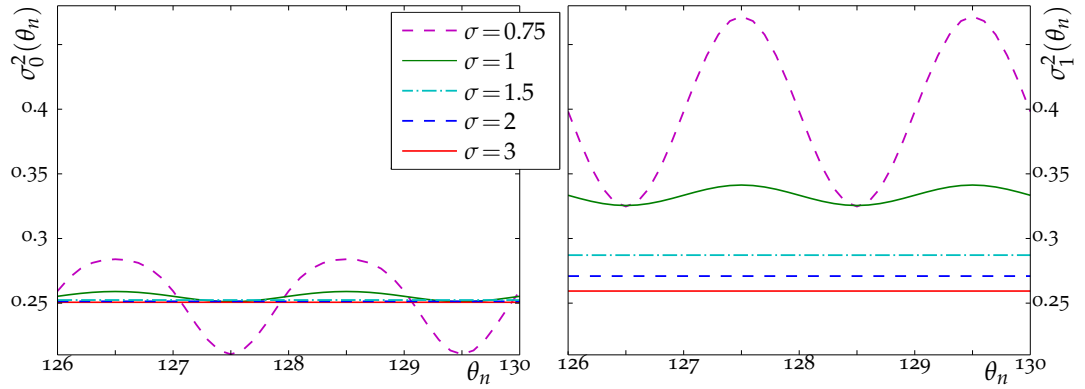


FIGURE 3.3: Représentations graphiques des variances  $\sigma_0^2$ ,  $\sigma_1^2$  et de leurs approximations, en négligeant la quantification,  $\widehat{\sigma}_0^2$ ,  $\widehat{\sigma}_1^2$  en fonction de  $\theta_n \in [126; 130]$  et pour  $\sigma_n = \{0.75; 1; 1.5; 2; 3\}$ . Dans un souci de clarté, les valeurs sont multipliées par  $\sigma_n^2$  afin que  $\widehat{\sigma}_0^2 = \widehat{\sigma}_1^2 = 1/4$ .

où, pour  $i = \{0; 1\}$ , les quantités  $\mu_i$  correspondent à l'espérance du rapport de vraisemblance sous l'hypothèse  $\mathcal{H}_i$  et sont données par (3.14) et (3.16). De manière analogue, il est immédiat, en utilisant la loi des variances totales, de montrer que :

$$\begin{aligned} \sigma_R^2 &= \mathbb{E} [\text{Var}_i [(\Lambda_R(z_n))]] + \text{Var} [\mathbb{E}_i [(\Lambda_R(z_n))]] \\ &= (1-R)\sigma_0^2 + R\sigma_1^2 + (1-R)\mu_0^2 + R\mu_1^2 - ((1-R)\mu_0 + R\mu_1)^2 \\ &= \sigma_R^2 = (1-R)(\sigma_0^2 + \mu_0^2) + R(\sigma_1^2 + \mu_1^2) - ((1-R)\mu_0 + R\mu_1)^2 \end{aligned} \quad (3.21)$$

Enfin, en négligeant la quantification, les expressions données par (3.19) permettent d'obtenir l'expression simplifiée de la variance :

$$\widehat{\sigma}_R^2 = \frac{1}{4\sigma_n^2} \left( 1 + \frac{R-R^2}{4\sigma_n^2} \right). \quad (3.22)$$

Dans un souci de clarté, on adopte la notation :

$$\bar{\mu}_i = \frac{1}{N} \sum_{n=1}^N \mathbb{E}_i [\log(\Lambda_i(z_n))] \quad \text{et} \quad \bar{\sigma}_i = \frac{1}{N} \sum_{n=1}^N \text{Var}_i [\log(\Lambda_i(z_n))]. \quad (3.23)$$

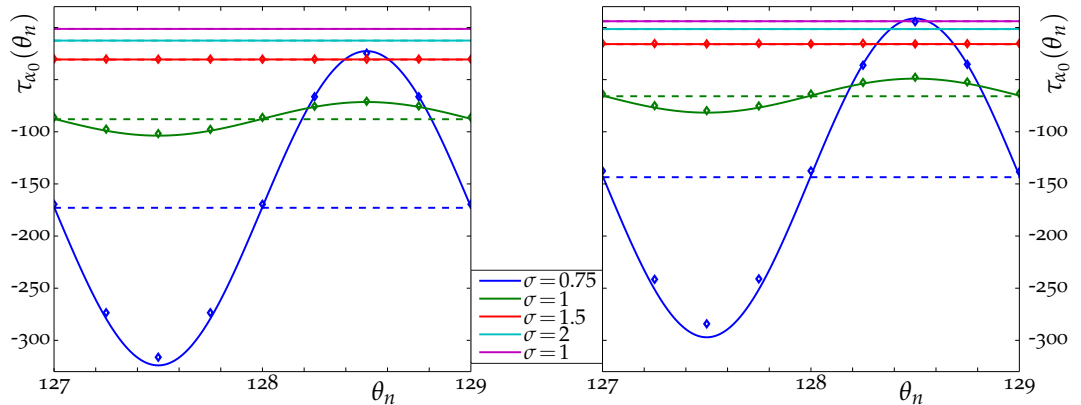


FIGURE 3.4: Représentations graphiques du seuil de décision  $\tau_{\alpha_0}$ , et de la valeur approchée en négligeant la quantification  $\widehat{\tau}_{\alpha_0}$  en fonction de  $\theta_n \in [127; 129]$  pour  $\sigma_n = \{0.75; 1; 1.5; 2; 3\}$ ,  $\alpha_0 = 10^{-2}$  (gauche) et  $\alpha_0 = 10^{-4}$  (droite),  $R = 0.1$  et  $N = 1000$ . Les lignes pointillées (--) correspondent aux valeurs théoriques en négligeant la quantification, les valeurs en traits pleins (—) aux valeurs théoriques avec prise en compte de la quantification (3.26) et les losanges ( $\diamond$ ) sont les résultats de simulation Monte-Carlo avec  $2.10^6$  réalisations.

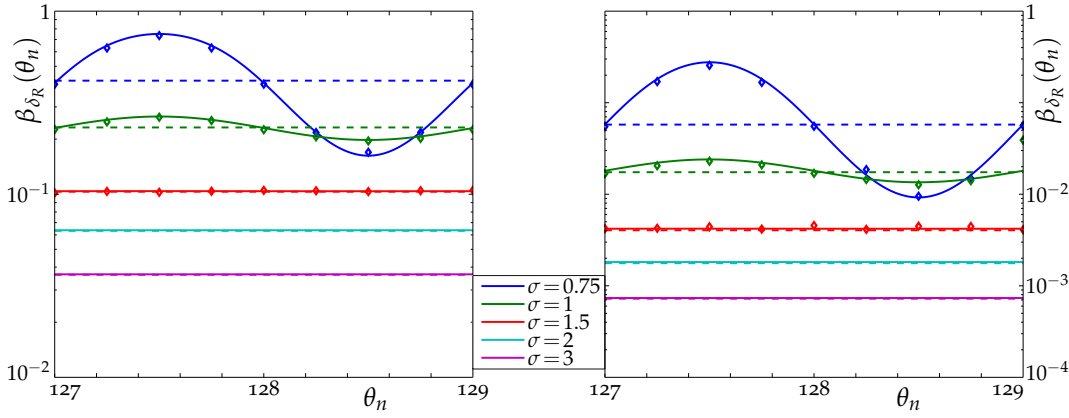


FIGURE 3.5: Représentations graphiques de la fonction de puissance  $\beta(\delta_R)$ , et de son approximation en négligeant la quantification  $\hat{\beta}(\delta_R)$  en fonction de  $\theta_n \in [127; 129]$  pour  $\sigma_n = \{0.75; 1; 1.5; 2; 3\}$ ,  $\alpha_0 = 10^{-2}$  (gauche) et  $\alpha_0 = 10^{-4}$  (droite),  $R = 0.1$  et  $N = 1000$ . Les lignes pointillées (-) correspondent aux valeurs théoriques en négligeant la quantification, les valeurs en traits pleins (—) aux valeurs théoriques avec prise en compte de la quantification (3.26) et les losanges ( $\diamond$ ) sont les résultats de simulation Monte-Carlo avec  $2.10^6$  réalisations.

Par ailleurs, les résultats numériques présentés sont obtenus, sauf lorsque cela est explicitement précisé, avec pour tout  $n \in \mathcal{I}$ ,  $\theta_n = \theta$  et  $\sigma_n^2 = \sigma^2$ . Dans ce cas particulier il est évident que  $\bar{\mu}_i = \mu_i$  et  $\bar{\sigma}_i = \sigma_i$ , où les expressions des moments  $\mu_i$  et  $\sigma_i$  sont définies dans les équations (3.14) - (3.17). En vertu du théorème (3.1) de la limite centrale de Lindeberg, un rapide calcul permet d'obtenir pour tout  $R \in ]0; 1[$  :

$$\begin{cases} \frac{1}{\sqrt{N}} \log(\Lambda_0(\mathbf{Z})) = \frac{1}{\sqrt{N}} \sum_{n=1}^N \log(\Lambda_0(z_n)) \rightsquigarrow \mathcal{N}(\bar{\mu}_0 \sqrt{N}; \bar{\sigma}_0^2) \text{ sous } \mathcal{H}_0 \\ \frac{1}{\sqrt{N}} \log(\Lambda_R(\mathbf{Z})) = \frac{1}{\sqrt{N}} \sum_{n=1}^N \log(\Lambda_R(z_n)) \rightsquigarrow \mathcal{N}(\bar{\mu}_R \sqrt{N}; \bar{\sigma}_R^2) \text{ sous } \mathcal{H}_1. \end{cases} \quad (3.24)$$

Dans l'approche du test le plus puissant de Neyman-Pearson, (voir définition 2.5 voir section 2.4.2), une fois fixée la probabilité  $\alpha_0$  maximale de fausse alarme admise, il est ensuite nécessaire de calculer le seuil de décision  $\tau_{\alpha_0}$  assurant que  $\delta_R \in \mathcal{K}_{\alpha_0}$ . La proposition 3.1 permet de calculer la valeur du seuil en fonction de  $\alpha_0$ , de  $\theta$ .

**Proposition 3.1.** *Un calcul direct permet de montrer que pour tout  $\alpha_0 \in ]0; 1[$ , en fixant*

$$\tau_{\alpha_0} = \left( \bar{\sigma}_0 \sqrt{N} \right) \Phi^{-1}(1 - \alpha_0) + (\bar{\mu}_0 N), \quad (3.25)$$

où  $\Phi^{-1}(\cdot)$  est la fonction réciproque de la fonction de répartition Gaussienne  $\Phi(\cdot)$ , on a asymptotiquement, lorsque  $N \rightarrow \infty$ ,  $\alpha_0(\delta_R) = \alpha_0$  et donc  $\delta_R \in \mathcal{K}_{\alpha_0}$ .

On rappellera que  $\mathcal{K}_{\alpha_0}$  définit la classe des tests dont la probabilité de fausse alarme est majorée par  $\alpha_0$  :  $\mathcal{K}_{\alpha_0} = \{\delta \in \Delta \mid \alpha_0(\delta) \leq \alpha_0\}$ .

**Démonstration 3.1.** *Un calcul direct, en utilisant les résultats (3.24) de convergence asymptotique de  $\log(\Lambda_R(\mathbf{Z}))$ , permet d'obtenir :*

$$\alpha_0 \geq \alpha_0(\delta_R) = \mathbb{P}_0[\log(\Lambda_R(\mathbf{Z})) \geq \tau_{\alpha_0}] = 1 - \Phi\left(\frac{\tau_{\alpha_0} - \bar{\mu}_0 N}{\bar{\sigma}_0 \sqrt{N}}\right).$$

L'application  $\Phi$  étant continue et strictement croissante, il en découle que :

$$\begin{aligned} \Phi^{-1}(1 - \alpha_0) &\leq \Phi^{-1}(1 - \alpha_0(\delta_R)) = \frac{\tau_{\alpha_0} - \bar{\mu}_0 N}{\bar{\sigma}_0 \sqrt{N}} \\ \Leftrightarrow \tau_{\alpha_0} &\leq (\bar{\sigma}_0 \sqrt{N}) \Phi^{-1}(1 - \alpha_0) + (\bar{\mu}_0 N). \end{aligned}$$

Il est immédiat de vérifier que le seuil de décision  $\tau_{\alpha_0}$  permet de maximiser la fonction de puissance tout en ayant  $\alpha_0 = \alpha_0(\delta_R)$  ce qui implique  $\delta_R \in \mathcal{K}_{\alpha_0}$ . La figure 3.4

illustre la valeur du seuil de décision ainsi obtenu pour  $\alpha_0 = \{10^{-2}; 10^{-4}\}$  en fonction de  $\theta_n \in [127; 129]$  et pour quelques valeurs de  $\sigma_n$ .

Enfin, une fois la valeur du seuil  $\tau_{\alpha_0}$  déterminée, le théorème 3.2 fournit une expression de la fonction de puissance  $\beta_{\delta_R}$  associée au test  $\delta_R$  le plus puissant de la classe  $\mathcal{K}_{\alpha_0}$ .

**Théorème 3.2.** *Supposons que le taux d'insertion  $R \in ]0; 1[$  ainsi que, pour tout  $n \in \{0, \dots, N\}$ , les paramètres  $\theta_n$  et  $\sigma_n$  sont connus. La puissance  $\beta_{\delta_R}$  associée au test  $\delta_R$  (3.7) est approximativement donnée par :*

*Cette expression de la fonction de puissance est approchée dans le sens où les termes correctifs ont été négligés, voir (3.9) et (3.10).*

$$\beta_{\delta_R} = 1 - \Phi \left( \frac{\tau_{\alpha_0} - \bar{\mu}_R \sqrt{N}}{\bar{\sigma}_R N} \right) = 1 - \Phi \left( \frac{\bar{\sigma}_0 \Phi^{-1}(1 - \alpha_0) + (\bar{\mu}_0 - \bar{\mu}_1) R \sqrt{N}}{\bar{\sigma}_R} \right) \quad (3.26)$$

où les quantités  $\bar{\mu}_i$  et  $\bar{\sigma}_i^2$  sont définies par les équations (3.14) - (3.17) et (3.21) - (3.21).

**Démonstration 3.2.** *La démonstration du théorème 3.2 est issue d'un calcul direct en utilisant le résultat (3.24) de convergence asymptotique de  $\log(\Lambda_R(\mathbf{Z}))$ . Ainsi, sous l'hypothèse  $\mathcal{H}_1$ , lorsque  $N \rightarrow \infty$ , on a :*

$$\mathbb{P}_R [\log(\Lambda_R(\mathbf{Z})) \geq \tau_{\alpha_0}] = 1 - \Phi \left( \frac{\tau_{\alpha_0} - \bar{\mu}_R N}{\bar{\sigma}_R \sqrt{N}} \right)$$

ce qui, en utilisant l'expression du seuil  $\tau_{\alpha_0}$  (3.25), permet l'obtention de la relation (3.26).

**Remarque 3.2.** *Lorsque la quantification est négligée, les valeurs des moments utilisés sont données par les relations (3.19), (3.20) et (3.22). Dans ces conditions, l'expression de la fonction de puissance peut être simplifiée dans le cas où pour tout  $n \in \mathcal{I}$ ,  $\theta_n = \theta$  et  $\sigma_n = \sigma$  :*

$$\hat{\beta}_{\delta_R} = 1 - \Phi \left( \frac{\Phi^{-1}(1 - \alpha_0) - R \sqrt{N}}{1 + \frac{R - R^2}{4\sigma_n^2}} \right).$$

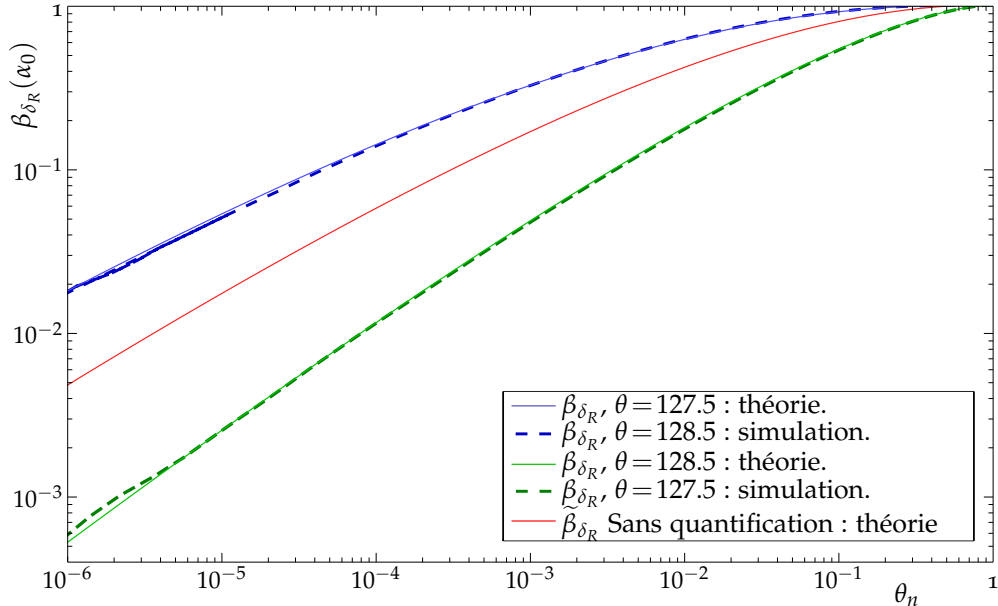


FIGURE 3.6: Courbes COR illustrant les performances du test PP  $\delta_R$  : représentation graphique de la fonction de puissance  $\beta_{\delta_R}$  en fonction de la probabilité de fausse alarme  $\alpha_0(\delta_R)$ . Les résultats présentés correspondent aux cas de  $N = 1000$  pixels de paramètres  $\theta = \{127.5; 128.5\}$   $\sigma_n = 0.75$  et, pour les simulations Monte-Carlo,  $10^8$  répétitions ont été réalisées.

L'impact de la prise en compte, ou de la négligence, de la quantification sur la fonction puissance est illustré dans les figures 3.5 et 3.6. La figure 3.5 illustre les valeurs prises par la fonction de puissance  $\beta_{\delta_R}$  et par son approximation  $\hat{\beta}_{\delta_R}$  en fonction de  $\theta \in [127; 129]$ . De manière analogue, la figure 3.6 offre une représentation graphique de la puissance  $\beta_{\delta_R}(\alpha_0)$  et de son approximation  $\hat{\beta}_{\delta_R}(\alpha_0)$  en fonction de la probabilité de fausse alarme  $\alpha_0$ .

### 3.3 Test entre deux hypothèses composites : taux d'insertion inconnu

Comme cela a été abordé dans la section 3.2, dans le cas où le taux d'insertion est supposé connu, la stéganalyse correspond à un problème de décision statistique entre deux hypothèses simples. En vertu du lemme de Neyman-Pearson 2.1, il est possible de construire le test le plus puissant et d'établir asymptotiquement ses performances statistiques.

Dans cette section, la détection statistique de la stéganographie dans le cas où le taux d'insertion  $R$  n'est pas connu, est étudiée. On suppose néanmoins que pour tout  $n \in \mathcal{I}$  les paramètres  $\theta_n$  et  $\sigma_n$  sont connus. La loi de probabilité  $P_{\theta_n}$ , voir relation (3.1), est donc supposée connue pour chacun des pixels  $c_n$ . À l'inverse, puisque le taux d'insertion  $R$  est inconnu, la loi de probabilité d'un pixel,  $Q_{\theta_n}^R$ , voir relation (3.3), est alors inconnue.

Dans ces conditions il est souhaitable, lors de l'analyse d'un média  $\mathbf{Z}$ , de décider entre les deux hypothèses composites suivantes :

$$\begin{cases} \mathcal{H}_0 = \{z_n \sim Q_{\theta_n}^R, \forall n \in \mathcal{I}, \forall R \leq R^*\} \\ \mathcal{H}_1 = \{z_n \sim Q_{\theta_n}^R, \forall n \in \mathcal{I}, \forall R > R^*\}. \end{cases} \quad (3.27)$$

ou  $R^*$  est un taux d'insertion "critique" au delà duquel le média est considéré porteur d'informations cachées. Le rapport de vraisemblance (3.5) peut alors s'écrire :

$$\Lambda_{R_0, R_1}(\mathbf{Z}) = \prod_{n=1}^N \Lambda_{R_0, R_1}(z_n) = \prod_{n=0}^N \frac{Q_{\theta_n}^{R_1}[z_n]}{Q_{\theta_n}^{R_0}[z_n]} = \prod_{n=0}^N \frac{(1 - R_1)P_{\theta_n}[z_n] + R_1/2P_{\theta_n}^0[z_n^{(0)}]}{(1 - R_0)P_{\theta_n}[z_n] + R_0/2P_{\theta_n}^0[z_n^{(0)}]}. \quad (3.28)$$

où  $R_0 \leq R^* < R_1$ . Dans le présent contexte d'un test entre deux hypothèses composites, la contrainte sur le respect d'une probabilité de fausse alarme doit être vérifiée pour tout  $R \leq R^*$ . Cela impose de chercher un test dans la classe  $\mathcal{K}_{\alpha_0}$  désormais définie par :  $\mathcal{K}_{\alpha_0} = \left\{ \delta : \sup_{R_0 \leq R^*} \mathbb{P}_{R_0}[\delta(\mathbf{Z}) = \mathcal{H}_1] \leq \alpha_0 \right\}$ , voir section 2.4.2. Une solution idéale consisterait à obtenir un test uniformément le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$ . Pour la détection de stéganographie par substitution de LSB, l'obtention d'un test UPP demeure incertaine car les hypothèses n'admettent pas un rapport de vraisemblance monotone, voir définition 2.7.

Dans cette section, une solution satisfaisante, reposant sur la normalité asymptotique locale de L. Le Cam [116], est proposée. Le fondement de cette méthodologie est de considérer deux hypothèses simples en fonction du nombre d'échantillons  $N$  de sorte que les deux hypothèses testées convergent l'une vers l'autre, de façon non dégénérée, lorsque  $N \rightarrow \infty$ . Le principal intérêt de cette approche théorique est que l'utilisation d'un développement limité de la log-vraisemblance  $\log(\Lambda_1(z_n))$  permet de réduire le problème statistique considéré à un test entre deux hypothèses Gaussiennes simples différant par leur moyenne. Dans le problème de décision considéré (3.27),

*La convergence des deux hypothèses est dite dégénérée si les hypothèses sont asymptotiquement parfaitement séparable ou, inversement, inséparable, i.e. s'il existe un test  $\delta^*$  tel que  $\alpha(\delta^*) \rightarrow 0$  et  $\beta(\delta^*) \rightarrow 1$  ou, inversement, si tout test vérifie  $\beta(\alpha_0) \rightarrow \alpha_0$  lorsque  $N \rightarrow \infty$ .*

*Cette approche peut-être liée à la définition 2.9 d'un test LAUPP pour lequel le domaine d'optimalité est proportionnel à  $N^{-1/2}$ .*

l'approche locale asymptotique de Le Cam est appliquée aux hypothèses statistiques définies par :

$$\begin{cases} \mathcal{H}_0 &= \{z_n \sim Q_{\theta_n}^{R^*}, \forall n \in \mathcal{I}, R = R^*\} \\ \mathcal{H}_{1,N} &= \{z_n \sim Q_{\theta_n}^{R_1^*}, \forall n \in \mathcal{I}, R_1^* = R_1^*(N) = R^* + cN^{-1/2}\} \end{cases} \quad (3.29)$$

De manière informelle, on peut considérer que le paramètre  $c$  fixe la séparabilité des hypothèses.

avec  $c > 0$ .

Le test du rapport de vraisemblance local, autour du taux d'insertion  $R = R^*$ , est donc donné par :

$$\Lambda_{R^*}(\mathbf{Z}) = \prod_{n=1}^N \frac{Q_{\theta_n}^{R_1^*}[z_n]}{Q_{\theta_n}^{R^*}[z_n]} \Leftrightarrow \log(\Lambda_{R^*}(\mathbf{Z})) = \sum_{n=1}^N \log(Q_{\theta_n}^{R_1^*}[z_n]) - \log(Q_{\theta_n}^{R^*}[z_n]). \quad (3.30)$$

Il est immédiat de vérifier que pour tout  $n \in \mathcal{I}$ ,  $\log(Q_{\theta_n}^{R_1^*}[z_n]) \rightarrow \log(Q_{\theta_n}^{R^*}[z_n])$  lorsque  $N \rightarrow \infty$ . Or un développement autour du taux d'insertion  $R^*$  permet d'obtenir, en admettant que  $Q^R[z_n]$  est deux fois continûment différentiable par rapport à  $R$  :

$$\log(\Lambda_{R^*}(\mathbf{Z})) = \frac{c}{\sqrt{N}} \sum_{n=1}^N \left. \frac{\partial \log(Q^R[z_n])}{\partial R} \right|_{R=R^*} + \frac{c^2}{2N} \sum_{n=1}^N \left. \frac{\partial^2 \log(Q^R[z_n])}{\partial R^2} \right|_{R=R^*} + o_p\left(\frac{h^2}{N}\right).$$

Dans cette relation, il est notable que  $\left. \frac{\partial \log(Q^R[z_n])}{\partial R} \right|_{R=R^*}$  correspond au score efficace, noté  $\zeta(\mathbf{Z}; R^*)$  dans la suite, dont un rapide calcul permet de donner l'expression :

La relation (3.31) est une valeur approchée du score efficace  $\zeta(\mathbf{Z}; R^*)$  car, ici encore, les termes correctifs  $\epsilon(k)$  (3.2) sont négligés.

$$\zeta(\mathbf{Z}; R^*) = \sum_{n=1}^N \frac{1/2 P_{\theta_n}^0[z_n^{(0)}] - P_{\theta_n}[z_n]}{(1 - R^*) P_{\theta_n}[z_n] + R^*/2 P_{\theta_n}^0[z_n^{(0)}]} \approx \sum_{n=1}^N \frac{\Lambda_1(z_n) - 1}{R^* \Lambda_1(z_n) + (1 - R^*)}. \quad (3.31)$$

De même, le second terme différentiel correspond asymptotiquement lorsque  $N \rightarrow \infty$ , en vertu de la loi des grands nombres, à :

$$\frac{c^2}{2N} \sum_{n=1}^N \left. \frac{\partial^2 \log(Q^R[z_n])}{\partial R^2} \right|_{R=R^*} \xrightarrow{p.s} \mathbb{E}_{R^*} \left[ \frac{\partial^2 \log(Q^R[z_n])}{\partial R^2} \right] = -\mathcal{F}(R^*)$$

où  $\mathcal{F}(R^*)$  représente l'information de Fisher dont une expression alternative est :

$$\mathcal{F}(R^*) = \mathbb{E} \left[ \left( \frac{\partial \log(Q^R[z_n])}{\partial R} \right)^2 \right] = \mathbb{E} \left[ \left( \frac{\Lambda_1(z_n) - 1}{R^* \Lambda_1(z_n) + (1 - R^*)} \right)^2 \right]. \quad (3.32)$$

De manière analogue, en vertu du théorème 3.1 de la limite centrale de Lindeberg, le score efficace admet la convergence asymptotique suivante :

$$\begin{cases} \zeta(\mathbf{Z}; R^*) \rightsquigarrow \mathcal{N}(0, \mathcal{F}(R^*)) & \text{sous } \mathcal{H}_0 \\ \zeta(\mathbf{Z}; R^*) \rightsquigarrow \mathcal{N}(\mathcal{F}(R^*)c, \mathcal{F}(R^*)) & \text{sous } \mathcal{H}_{1,N}, \end{cases} \quad (3.33)$$

lorsque  $N \rightarrow \infty$ .

**Remarque 3.3.** Il est notable que la vitesse de convergence utilisée dans l'approche locale asymptotique,  $N^{-1/2}$ , permet d'assurer que le score efficace et l'information de Fisher (sous  $\mathcal{H}_0$  comme sous  $\mathcal{H}_{1,N}$ ) convergent en loi vers des distributions Gaussiennes de paramètres finis et distincts.

Finalement, les précédents calculs (3.30) - (3.33) permettent d'établir le développement asymptotique, lorsque  $N \rightarrow \infty$ , de la fonction  $\log(\Lambda_{R^*})$  :

$$\log(\Lambda_{R^*}(\mathbf{Z})) \simeq \frac{c}{\sqrt{N}} \zeta(\mathbf{Z}; R^*) - \frac{c^2}{2} \mathcal{F}(R^*). \quad (3.34)$$

En conséquence, pour tester les hypothèses  $\mathcal{H}_0 = \{R=R^*\}$  contre  $\mathcal{H}_1 = \{R=R^* + c/N\}$ , le test  $\delta_{R^*}$  défini par :

$$\delta_{R^*}(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \zeta(\mathbf{Z}; R^*) < \tau_{R^*} \\ \mathcal{H}_{1,N} & \text{si } \zeta(\mathbf{Z}; R^*) \geq \tau_{R^*}, \end{cases} \quad (3.35)$$

est caractérisé par la proposition 3.2 et le théorème 3.3 suivants.

**Proposition 3.2.** *Pour tout  $\alpha_0 \in ]0; 1[$ , en fixant :*

$$\tau_{R^*} = \mathcal{F}(R^*)\Phi^{-1}(1 - \alpha_0). \quad (3.36)$$

*permet asymptotiquement, lorsque  $N \rightarrow \infty$ , de garantir que le test  $\delta_{R^*}$  (3.35) associé au seuil de décision  $\tau_{R^*}$  vérifie  $\delta_{R^*} \in \mathcal{K}_{\alpha_0}$ .*

*Démonstration.* La démonstration de la proposition 3.2, laissée au lecteur, repose sur le résultat de convergence (3.33), et est similaire en tout point à la démonstration 3.1.  $\square$

**Théorème 3.3.** *En outre, un calcul direct permet d'établir le fonction de puissance du test  $\delta_{R^*}$  (3.35) :*

$$\beta_{\delta_{R^*}} = 1 - \Phi\left(\Phi^{-1}(1 - \alpha_0) - c\right). \quad (3.37)$$

*Démonstration.* La démonstration de la proposition 3.3, laissée au lecteur, repose sur le résultat de convergence (3.33) est similaire en tout point à la démonstration 3.2.  $\square$

Enfin, l'intérêt fondamental du test  $\delta_{R^*}$  réside dans le théorème 3.4

**Théorème 3.4.** *Quel que soit  $R^* \in ]0; 1[$ , pour tester les hypothèses  $\mathcal{H}_0 = \{R = R^*\}$  contre  $\mathcal{H}_1 = \{R = R^* + c/N\}$ , le test  $\delta_{R^*}$  associé au seuil de décision  $\tau_{R^*}$  est LAUPP, voir définition 2.9.*

*Démonstration.* La démonstration du théorème 3.4 découle de [107, théorème 13.3.3].  $\square$

L'approche asymptotique locale peut également être appliquée pour résoudre le problème de décision statistique (3.27) au prix de l'abandon de l'optimalité du test, cette dernière n'étant vérifiée que localement (autour de  $R^*$ ) et asymptotiquement (pour  $N \rightarrow \infty$ ).

Cependant, lors de l'analyse du média  $\mathbf{Z}$ , supposer que l'on dispose d'un nombre arbitrairement grand de pixels n'est pas réaliste. Le test proposé en utilisant l'approche locale asymptotique ne permet pas, en toute rigueur, de garantir la plus grande puissance lorsque le nombre de pixels est fixé. Autrement dit, l'utilisation pratique du test (3.35) LAUPP à distance finie est possible mais l'optimalité de ce dernier (au sens de Neyman-Pearson) n'est plus assurée, notamment lorsque le nombre de pixel n'est pas "suffisamment grand" ou que le taux d'insertion  $R$  n'est pas "suffisamment proche" du taux critique testé  $R^*$ . Ce dernier cas est illustré dans la figure 3.7 présentant la fonction de puissance  $\beta_{\delta_{R^*}}$  du test LAUPP pour  $R^* = \{0.1; 1\}$  pour un taux d'insertion  $R = 0.05$ ; il est notable que la perte de puissance du test  $\delta_{R^*}$ , LAUPP pour  $R^* = 1$ , est plus importante que la perte de puissance du même test pour  $R^* = 0.1$ .

**Remarque 3.4.** *La normalité asymptotique du score efficace  $\zeta(\mathbf{Z}; R^*)$ , voir équation (3.33), est assurée par le développement (3.34) de  $L$ . Le Cam fondé sur le théorème 3.1 de la limite centrale. Lors de l'analyse d'un média stéganographié avec un taux d'insertion  $R$  la comparaison de performance du test PP, le plus puissant,  $\delta_1$  défini par (3.7) et conçu pour  $R = 1$  avec celle du test (3.35) LAUPP  $\delta_{R^*}$  conçu pour  $R = R^*$  et celle du test PP  $\delta_R$  conçu pour  $R$  nécessite le calcul des deux premiers moments des statistiques de décision (resp.  $\log(\Lambda_1(\mathbf{Z}))$ ,  $\zeta(\mathbf{Z}; R^*)$  et  $\log(\Lambda_R(\mathbf{Z}))$ ) sous les hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$ . La figure 3.7 présente graphiquement une comparaison de ce type.*



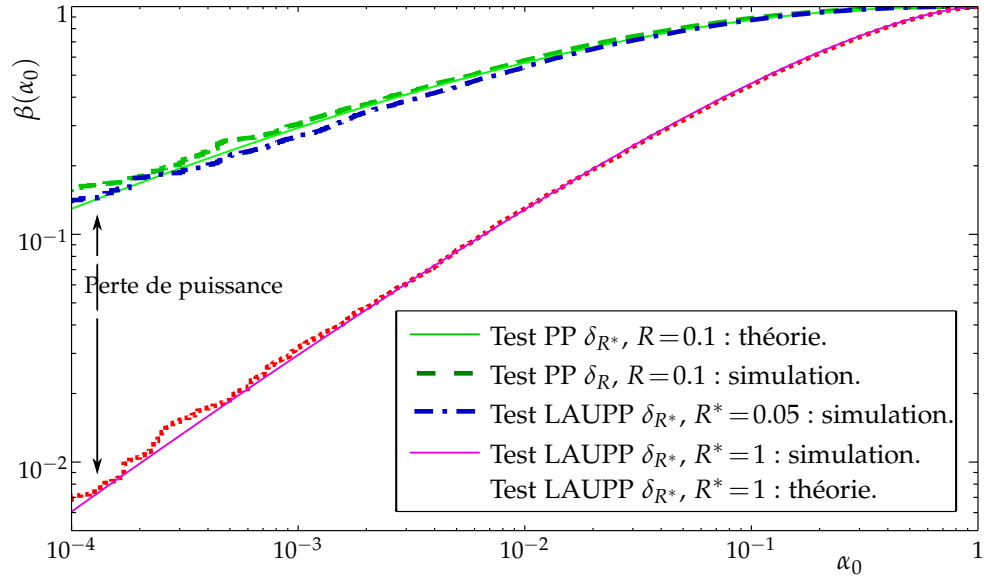


FIGURE 3.7: Comparaison de la puissance entre le test PP  $\delta_R$  (dans le cas d'hypothèses simples) et les tests LAUPP. Le taux d'insertion est ici  $R = 0.05$  et les taux d'insertion considérés par les tests LAUPP sont  $R^* = 0.1$  et  $R^* = 1$ . Les résultats présentés correspondent aux cas de  $N = 1000$  pixels de paramètres  $\theta = 127$   $\sigma_n = 0.75$  et, pour les simulations Monte-Carlo,  $10^5$  répétitions ont été réalisées.

### 3.4 Test en présence de paramètres de nuisance : modèle régressif de l'image.

En pratique, la difficulté majeure provient de la présence des paramètres de nuisance que constituent l'espérance  $\theta = (\theta_1, \dots, \theta_N)^T$  et la variance du bruit  $\sigma^2 = (\sigma_1^2, \dots, \sigma_N^2)^T$ ; ces paramètres n'ont en effet aucun intérêt pour la détection mais interviennent dans l'expression des lois de probabilité,  $P_\theta$  et  $Q_\theta^R$ , et donc dans l'expression du rapport de vraisemblance. Les aspects théoriques des tests d'hypothèses en présence de paramètres de nuisance sont détaillés dans [87, 107]. Une approche performante consiste à utiliser la théorie de l'invariance statistique abordée dans la section 2.4.3. La conception de tests invariants optimaux dans le cadre du traitement des images a déjà été proposée et étudiée dans [117, 115]. L'approche proposée présente une certaine similitude avec l'utilisation du principe d'invariance, mais il faut souligner qu'en tenant compte de la quantification l'invariance du groupe n'est pas vérifiée et la statistique invariant maximale n'a pas été explicitée, voir section 2.4.3 ou, plus d'informations, [107, chap.6] et [87, chap.47].

Pour appliquer la méthodologie de détection décrite dans la section 3.2, il est proposé de découper l'image à analyser en  $K$  segments contigus de  $L$  pixels consécutifs, voir figure 3.8. Il est en outre supposé que pour tout  $k \in \{1, \dots, K\}$  le vecteur  $\mathbf{z}_k \in \mathcal{Z}^L$  peut être représenté de façon linéaire comme suit :

$$\mathbf{z}_k = Q_1[\mathbf{y}_k] , \quad \mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \boldsymbol{\zeta}_k = \boldsymbol{\theta}_k + \boldsymbol{\zeta}_k \quad (3.38)$$

où  $\mathbf{H} \in \mathcal{M}_{L,p}$ ,  $L > p$  est une matrice de rang colonne plein,  $\boldsymbol{\zeta}_k$  est la réalisation d'un vecteur aléatoire Gaussien  $\boldsymbol{\Xi}_k \sim \mathcal{N}(0, \sigma_k^2 \mathbf{I}_L)$  avec  $\mathbf{I}_L$  la matrice identité de taille  $L \times L$  et, pour tout  $k \in \{1, \dots, K\}$ ,  $\mathbf{x}_k \in \mathbb{R}^p$ ,  $\mathbf{y}_k \in \mathbb{R}^L$ ,  $\boldsymbol{\theta}_k \in \mathbb{R}^L$ . Les modèles linéaires tels que (3.38) ont été étudiés largement dans la littérature. Les éléments des vecteurs  $\mathbf{z}_k$ ,  $\mathbf{y}_k \in \mathbb{R}^L$  et  $\boldsymbol{\theta}_k$  sont respectivement notés pour tout  $l \in \{1, \dots, L\}$ ,  $z_{k,l}$ ,  $y_{k,l}$  et  $\theta_{k,l}$ . En outre, en vertu de l'indépendance statistique des pixels, la loi de probabilité conjointe du vecteur  $\mathbf{z}_k$  est définie par le produit des lois de probabilité :  $P_{\theta_n} = P_{\theta_{k,1}} \times \dots \times P_{\theta_{k,L}}$

Rigoureusement l'invariance statistique n'est vérifiée que dans le cas où les données sont à valeurs réelles.

Le modèle linéaire (3.38) ne permet pas une modélisation réaliste du contenu structuré des images mais est utilisé à titre d'exemple. La modélisation du contenu des images est l'objet des chapitres 4 et 5.

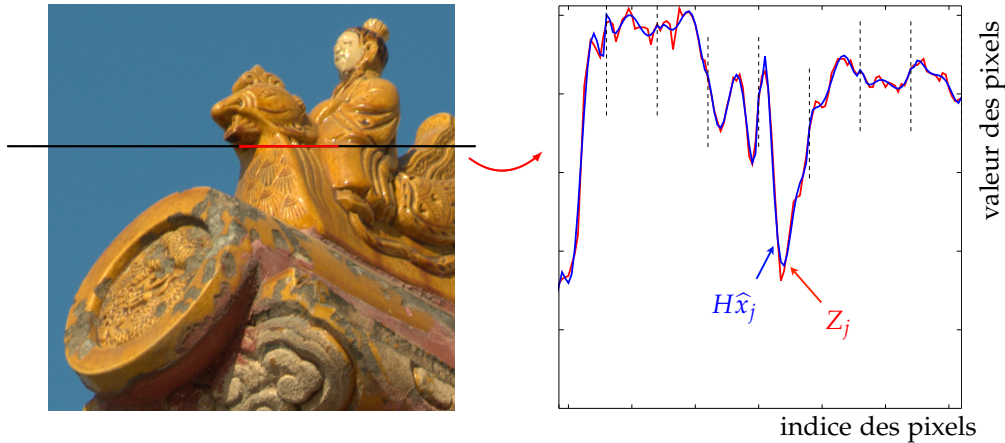


FIGURE 3.8: Illustration du découpage d'une image en segments  $\mathbf{z}_k$  de  $L = 16$  pixels et des approximations par  $\hat{\mathbf{z}}_k = \mathbf{H}\hat{\mathbf{x}}_k$ .

sous  $\mathcal{H}_0$  et  $Q_{\theta_n}^R = Q_{\theta_{k,1}}^R \times \dots \times Q_{\theta_{k,L}}^R$  sous  $\mathcal{H}_1$ .

Comme expliqué dans la section 3.3, lorsque le taux  $R$  est inconnu, le problème statistique qu'il est souhaitable de résoudre consiste à décider entre les deux hypothèses composites (3.27) :

$$\begin{aligned} \mathcal{H}_0 &= \{ \mathbf{z}_k \sim Q_{\theta_n}^{R_0}, \forall k \in \{1, \dots, K\}, \forall R_0 \leq R^* \} \\ \text{vs } \mathcal{H}_1 &= \{ \mathbf{z}_k \sim Q_{\theta_n}^{R_1}, \forall k \in \{1, \dots, K\}, \forall R_1 > R^* \}. \end{aligned}$$

Comme cela a été fait en début de section 3.2, les termes correctifs (3.2) sont considérés comme négligeables. Il découle de la définition de  $P_{\theta_n}$  (3.1) et  $Q_{\theta_n}^R$  et (3.3) que, dans le cas du modèle paramétrique (3.38), le logarithme du rapport de vraisemblance  $\log(\Lambda_1(\mathbf{z}_k))$  du  $k$ -ième segment peut approximativement s'écrire :

$$\begin{aligned} \log(\Lambda_1(\mathbf{z}_k)) &\simeq \frac{-1}{2\sigma_k^2} \left[ \left( z_{k,l}^{(0)} - \theta_l + 1/2 \right)^2 - \left( z_{k,l} - \theta_l \right)^2 \right] \\ &= -\frac{1}{2\sigma^2} \left[ \left\| \mathbf{z}_k^{(0)} - \mathbf{H}\mathbf{x}_k + 0.5 \cdot \mathbf{1}_L \right\|_2^2 - \left\| \mathbf{z}_k - \mathbf{H}\mathbf{x}_k \right\|_2^2 \right] \quad (3.39) \end{aligned}$$

où  $\mathbf{1}_L \in \mathbb{R}^L$  est défini par  $\mathbf{1} = (1, \dots, 1)^T$  et  $\mathbf{z}_k^{(0)}$  (resp.  $z_{k,l}^{(0)}$ ) représente le vecteur  $\mathbf{z}_k$  (resp. l'entier  $z_{k,l}$ ) dont les LSB sont remplacés par des "0".

Il est immédiat de vérifier que l'expression (3.39) correspond rigoureusement à utiliser le modèle paramétrique (3.38) dans les relations (3.6) ou (3.10).

La formulation (3.39) présente, en outre, l'intérêt de permettre une élimination algébrique du paramètre de nuisance que constitue  $\mathbf{H}\mathbf{x}_k$ . Cette méthode n'est pas sans rappeler le principe d'invariance lorsque les données sont réelles, par opposition aux pixels quantifiés lors de l'acquisition.

Enfin, il est proposé d'estimer le paramètre vectoriel  $\mathbf{x}_k$  par maximum de vraisemblance à partir de  $\mathbf{z}_k^{(0)}$  pour supprimer ainsi toute inférence éventuelle avec les bits insérés :

$$\hat{\mathbf{x}}_k = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{z}_k^{(0)}.$$

De façon similaire, l'estimation de  $\sigma_k$  par maximum de vraisemblance à partir de  $\mathbf{z}_k^{(0)}$  est donnée par :

$$\hat{\sigma}_k^2 = \frac{1}{L-p} \left\| \mathbf{P}_H^\perp \mathbf{z}_k \right\|_2^2 \quad \text{où} \quad \mathbf{P}_H^\perp = \mathbf{I}_L - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T,$$

*Les méthodes classiques d'estimation sont présentées dans la section A.6.*

la matrice  $\mathbf{P}_H^\perp$  correspond au projecteur sur l'espace supplémentaire de l'image de  $\mathbf{H}$ .

En utilisant ces deux estimations dans (3.39), le logarithme du rapport de vraisemblance généralisé (RVG) peut alors s'écrire :

$$\log(\Lambda_1^{\text{rvg}}(\mathbf{z}_k)) = -\frac{1}{2\hat{\sigma}_k^2} \left[ \|\mathbf{z}_k^{(0)} - \mathbf{H}\hat{\mathbf{x}}_k + 0.5 \cdot \mathbf{1}_L\|_2^2 - \|\mathbf{z}_k - \mathbf{H}\hat{\mathbf{x}}_k\|_2^2 \right]. \quad (3.40)$$

Un développement des termes quadratiques, voir (3.11) et (3.12), permet d'écrire le terme  $\log(\hat{\Lambda}_1(\mathbf{z}_k))$  (3.40) comme suit :

$$\log(\Lambda_1^{\text{rvg}}(\mathbf{z}_k)) = \frac{1}{\hat{\sigma}_k^2} \left( \mathbf{P}_H^\perp \mathbf{z}_k^{(0)} \right)^T \left( \mathbf{z}_k^{(0)} - 1/2 \cdot \mathbf{1}_L \right) - \frac{1}{8\hat{\sigma}_k^2}. \quad (3.41)$$

Le test du RVG est donc similaire au test du RV (3.7) et est donné par la règle de décision suivante :

$$\delta_R^{\text{rvg}}(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \log(\Lambda_1^{\text{rvg}}(\mathbf{Z})) = \sum_{k=1}^K \log(\Lambda_1^{\text{rvg}}(\mathbf{z}_k)) < \tau_{\alpha_0} \\ \mathcal{H}_1 & \text{si } \log(\Lambda_1^{\text{rvg}}(\mathbf{Z})) = \sum_{k=1}^K \log(\Lambda_1^{\text{rvg}}(\mathbf{z}_k)) \geq \tau_{\alpha_0} \end{cases} \quad (3.42)$$

où le seuil de décision  $\tau_{\alpha_0}$  est la solution de  $\sup_{R_0 \leq R^*; \mathbf{x}_k \in \mathbb{R}^p} \mathbb{P}_{R_0, \mathbf{x}} \left[ \log(\Lambda_1^{\text{rvg}}(\mathbf{Z})) \geq \tau_{\alpha_0} \right] = \alpha_0$ .

**Remarque 3.5.** Il est notable que les appellations de RVG concernant la quantité (3.40) - (3.41), et, de manière analogue, de test du RVG concernant (3.42), sont abusives. En effet le RVG est rigoureusement défini par :

$$\log \Lambda^{\text{rvg}}(\mathbf{Z}) = \frac{\sup_{R_1 > R^*; \mathbf{x}_k \in \mathbb{R}^p} Q_{\mathbf{x}_k}^{R_1}(\mathbf{Z})}{\sup_{R_0 \leq R^*; \mathbf{x}_k \in \mathbb{R}^p} Q_{\mathbf{x}_k}^{R_0}(\mathbf{Z})},$$

voir définition 2.10. En pratique cela revient donc à estimer simultanément par maximum de vraisemblance le paramètre  $\mathbf{x}_k$  et le taux d'insertion  $R$ .

Dans le cas présent, la maximisation des vraisemblances par rapport au taux d'insertion  $R$  n'est pas considérée car, d'une part, l'estimation de ce dernier reste un problème ouvert et, d'autre part, le traitement de la présence des paramètres de nuisance  $\mathbf{x}_k$  demeure la principale difficulté.

**Remarque 3.6.** Le RVG qu'il est proposé d'utiliser correspond donc au rapport :

$$\log \Lambda^{\text{rvg}}(\mathbf{Z}) = \frac{\sup_{\mathbf{x}_k \in \mathbb{R}^p} Q_{\mathbf{x}_k}^1(\mathbf{Z})}{\sup_{\mathbf{x}_k \in \mathbb{R}^p} P_{\mathbf{x}_k}(\mathbf{Z})}.$$

Or, il est notable que le RVG qu'il est proposé d'utiliser ne correspond au supremum que lorsque les données sont à valeurs réelles. Il est donc considéré implicitement que l'impact de la quantification sur la valeur de l'estimation  $\mathbf{x}_k$  est négligeable.

Ici encore, l'établissement des performances du test (3.42) n'est pas évidente à distance finie. L'approche asymptotique est de nouveau utilisée pour caractériser les performances du test lorsque le nombre de segments  $K \rightarrow \infty$ . Comme développé dans la section 3.2, l'idée est alors d'utiliser le théorème 3.1 de la limite centrale de Lindeberg pour expliciter la distribution asymptotique de  $\log(\hat{\Lambda}_1(\mathbf{Z}))$ , sous chacune des hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$ , au moyen du calcul des deux moments.

En supposant que les segments d'une image naturelle peuvent être modélisés par (3.38), un calcul direct similaire à (3.14) permet de montrer [23] que :

$$\hat{\mu}_0 = \mathbb{E}_0 \left[ \sum_{k=1}^K \log(\Lambda_1^{\text{rvg}}(\mathbf{z}_k)) \right] \simeq K \frac{2p - L}{8\hat{\sigma}^2}, \quad (3.43)$$

où  $\frac{1}{\bar{\sigma}^2} = \frac{1}{K} \sum_{k=1}^K \frac{1}{\sigma_k^2}$  et  $\simeq$  représente la valeur approchée considérant que la quantification a une influence négligeable sur l'estimation de  $\mathbf{x}_k$ , voir remarque 3.6. De manière analogue, un calcul direct montre que :

$$\hat{\sigma}_0^2 = \text{Var}_0 \left[ \sum_{k=1}^K \log (\Lambda_1^{\text{rvg}}(\mathbf{z}_k)) \right] \simeq K(L-p) \left( \frac{1}{4\bar{\sigma}^2} + \frac{1}{16\bar{\sigma}^2} \right), \quad (3.44)$$

Enfin, pour un taux d'insertion quelconque  $R \in ]0; 1]$ , les deux premiers moments du RVG approché sont donnés par :

$$\hat{\mu}_R = \mathbb{E}_R \left[ \sum_{k=1}^K \log (\Lambda_1^{\text{rvg}}(\mathbf{z}_k)) \right] \simeq K \frac{2p-L+2R(L-p)}{8\bar{\sigma}^2}, \quad (3.45)$$

$$\text{et } \hat{\sigma}_R^2 = \text{Var}_R \left[ \sum_{k=1}^K \log (\Lambda_1^{\text{rvg}}(\mathbf{z}_k)) \right] \simeq K(L-p) \left( \frac{1}{4\bar{\sigma}^2} + \frac{(1-R)^2}{16\bar{\sigma}^2} \right). \quad (3.46)$$

Dans le cas où la moyenne du signal est connue, la proposition 3.1 et le théorème 3.2 fournissent explicitement la valeur du seuil de décision  $\tau_{\alpha_0}$ , en fonction de la probabilité de fausses alarmes maximale  $\alpha_0$ , et la puissance du test du RV  $\beta_{\delta_R}$  respectivement. Dans le cas où la présence du paramètre de nuisance  $\mathbf{x}_k$  impose l'utilisation du RVG (3.40) - (3.41) et, *in fine*, du test (3.42) dont les performances statistiques sont analytiquement établies par ci-dessous par la proposition 3.3 et le théorème 3.5.

**Proposition 3.3.** *Un calcul direct permet de montrer que pour tout  $\alpha_0 \in ]0; 1[$ , en fixant*

$$\tau_{\alpha_0} = \left( \bar{\sigma}_0 \sqrt{N} \right) \Phi^{-1}(1 - \alpha_0) + (\bar{\mu}_0 N), \quad (3.47)$$

*on a asymptotiquement, lorsque  $N \rightarrow \infty$ ,  $\alpha_0(\delta_R^{\text{rvg}}) = \alpha_0$  et donc  $\delta_R^{\text{rvg}} \in \mathcal{K}_{\alpha_0}$ .*

*Démonstration.* La démonstration de la proposition 3.2, laissée au lecteur, repose sur les moments (3.43) et (3.44) est similaire en tout point à la démonstration 3.1.  $\square$

Enfin, une fois la valeur du seuil  $\tau_{\alpha_0}$  déterminée, le théorème 3.5 fournit une expression de la fonction de puissance  $\beta_{\delta_R^{\text{rvg}}}$  associée au test  $\delta_R^{\text{rvg}}$ .

**Théorème 3.5.** *Supposons que le taux d'insertion  $R \in ]0; 1]$  ainsi que, pour tout  $n \in \{0, \dots, N\}$ , les paramètres  $\theta_n$  et  $\sigma_n$  sont connus. La puissance  $\beta_{\delta_R}$  associée au test  $\delta_R$  (3.42) est donnée par :*

$$\beta_{\delta_R} = 1 - \Phi \left( \frac{\tau_{\alpha_0} - \bar{\mu}_R \sqrt{N}}{\bar{\sigma}_R N} \right) = 1 - \Phi \left( \frac{\bar{\sigma}_0 \Phi^{-1}(1 - \alpha_0) + (\bar{\mu}_0 - \bar{\mu}_1) R \sqrt{N}}{\bar{\sigma}_R} \right), \quad (3.48)$$

où les quantités  $\bar{\mu}_i$  et  $\bar{\sigma}_i^2$  sont définies par les équations (3.43) - (3.46).

*Démonstration.* La démonstration de la proposition 3.3, laissée au lecteur, repose sur les moments (3.43) - (3.46) est similaire en tout point à la démonstration 3.2.  $\square$

**Remarque 3.7.** *Il est notable que l'optimalité du test du RVG n'est plus assurée. La principale source de perte de puissance, en comparaison avec le test  $\delta_R$  du RV, est due à l'utilisation du modèle provoquant une réduction du nombre de degré de liberté (passant de KL à  $K(L-p)$ ); cependant le taux d'insertion influe également puisque le test  $\delta_R^{\text{rvg}}$  du RVG découle du test  $\delta_R$  du RV conçu pour  $R = 1$ , voir figure 3.7.*

*On rappellera que  $\mathcal{K}_{\alpha_0}$  définit la classe des tests dont la probabilité de fausse alarme est majorée par  $\alpha_0$  :  $\mathcal{K}_{\alpha_0} = \{\delta \in \Delta \mid \alpha_0(\delta) \leq \alpha_0\}$ .*

*Il est rappelé que cette expression de la fonction de puissance est approchée dans le sens où les moments ont été calculés en négligeant la quantification.*

### 3.5 Détection optimale sous condition d'un faible pas de quantification

Les résultats présentés dans la section 3.1 permettent de mettre en évidence deux phénomènes importants :

- D'une part, l'impact de la quantification sur les performances des tests présentés peut être important, notamment lorsque  $\sigma_n$  est faible et lorsque  $\theta \approx k + 1/2$ ,  $k \in \mathcal{Z}$ , voir figure 3.6.
- D'autre part, le cas d'hypothèses composites, lorsque le taux d'insertion  $R$  est inconnu, ne peut être résolu de façon optimale que dans un cadre local asymptotique. Si le taux d'insertion testé  $R^*$  diffère grandement du taux d'insertion réel  $R$ , la puissance du test  $\delta_{R^*}$  peut-être grandement diminuée, voir figure 3.7.

Ces résultats, bien qu'intéressants, sont contrebalancés par les deux inconvénients majeurs que présente l'utilisation pratique des tests PP  $\delta_R$  et LAUPP  $\delta_{R^*}$  :

1. D'une part le calcul du seuil de décision, reposant sur la proposition 3.1 et sur la proposition 3.2, peut être coûteux numériquement. De façon analogue, la puissance des tests, bien que calculable numériquement, ne peut être formulée analytiquement de façon simple.
2. D'autre part, il est immédiat de constater que l'information utile fournie par les pixels, pour le tests PP comme pour le test LAUPP, est contenue pour tout  $n \in \mathcal{I}$  dans le terme

$$\log(\Lambda_1(z_n)) = \frac{P_{\theta_n}^0[z_n^{(0)}]}{P_{\theta_n}[z_n]}.$$

La suppression du bit de poids faible (dans  $z_n^{(0)}$ ) présente l'intérêt d'éviter toute inférence statistique avec les éventuels bits d'informations cachées. Cependant, dans la pratique il est précisément souhaitable de pouvoir détecter la dissimulation d'informations avec un faible taux d'insertion. On peut alors se questionner sur la pertinence de supprimer tous les LSB si, typiquement, une proportion de  $R = 0.05$  d'entre eux sont utilisés pour dissimuler de l'information.

Dans cette section, une méthode alternative de stéganalyse basée sur la théorie de la décision statistique est proposée. L'idée est de considérer le "pire cas" de détection, *i.e.* lorsque le taux d'insertion  $R$  est faible et que la variance du bruit est importante, typiquement  $\sigma \gg 1$ . Dans ces conditions, négliger la quantification au sens qui a été expliqué plus en détails dans la section 3.2 semble pertinent. En outre, les paramètres de distributions  $\theta_n$  et  $\sigma_n$  n'étant pas constants sur l'ensemble de l'image, l'impact de la négligence de la quantification en est réduit davantage.

Le problème de décision statique considéré est, dans un premier temps, celui du test entre les deux hypothèses simples (3.4) :

$$\mathcal{H}_0 = \{z_n \sim P_{\theta_n}, \forall n \in \mathcal{I}\} \quad \text{vs} \quad \mathcal{H}_1 = \{z_n \sim Q_{\theta_n}^R, \forall n \in \mathcal{I}\}.$$

où  $R \in ]0;1[$  est supposé connu et les distributions  $P_{\theta_n}$  et  $Q_{\theta_n}^R$  sont respectivement définies par les équations (3.1) et (3.3). Comme dans la section 3.1, il est proposé d'utiliser le logarithme du rapport de vraisemblance  $\log(\Lambda_R(\mathbf{Z}))$  défini par (3.5) :

$$\log(\Lambda_R(\mathbf{Z})) = \sum_{n=1}^N \log(\Lambda_R(z_n)) = \sum_{n=0}^N \log\left(\frac{Q_{\theta_n}^R[z_n]}{P_{\theta_n}[z_n]}\right).$$

*Les causes de la non-stationnarité de  $\theta_n$  et  $\sigma_n$  sont décrites dans le chapitre 4.*

*Si par exemple  $\theta_n$  est uniformément distribuée sur  $[127.5;128.5]$  les figures 3.2-3.6 montrent un amoindrissement de l'impact de la quantification.*

Sous les assertions A-1 et A-2, définies dans la section 3.1, la log-vraisemblance  $\log(\Lambda_R(\mathbf{Z}))$  peut être écrite :

$$\log(\Lambda_R(z_n)) = \log\left(1 - R + \frac{R}{2} \frac{P_{\theta_n}[z_n] + P_{\theta_n}[\bar{z}_n]}{P_{\theta_n}[z_n]}\right) = \log\left(1 + R \left[\frac{P_{\theta_n}[z_n] + P_{\theta_n}[\bar{z}_n]}{2P_{\theta_n}[z_n]} - 1\right]\right), \quad (3.49)$$

où  $\bar{z}_n = z_n + (-1)^{z_n}$  correspond à l'entier  $z_n$  dont la valeur du LSB a été inversée, voir section 2.3.3.3. Ici encore il est proposé de négliger les termes correctifs définis dans (3.2) pour écrire, en utilisant la méthode du point médian :

$$\begin{aligned} P_{\theta_n}[z_n] &= \phi\left(\frac{z_n - \theta_n}{\sigma_n} + \epsilon(z_n)\right) = \phi\left(\frac{z_n - \theta_n}{\sigma_n}\right) + o_p(\sigma_n^{-2}) \\ P_{\theta_n}[z_n] + P_{\theta_n}[\bar{z}_n] &= 2\phi\left(\frac{z_n - \theta_n - 1/2(z_n - \bar{z}_n)}{\sigma_n} + \epsilon^{(0)}(z_n)\right) = 2\phi\left(\frac{z_n - \theta_n - 1/2(z_n - \bar{z}_n)}{\sigma_n}\right) + o_p(\sigma_n^{-2}), \end{aligned} \quad (3.50)$$

où, comme précédemment, la notation  $x = o_p(y)$  signifie que  $x/y \rightarrow 0$  en probabilité lorsque  $y \rightarrow 0$ .

De manière analogue aux calculs (3.11) et (3.12), le développement du terme quadratique  $(z_n - \theta_n + 1/2(z_n - \bar{z}_n))^2$  permet d'obtenir à partir de (3.50) :

$$\frac{P_{\theta_n}[z_n] + P_{\theta_n}[\bar{z}_n]}{2P_{\theta_n}[z_n]} = \exp\left(\frac{-1}{8\sigma_n^2}\right) \exp\left(\frac{1}{2\sigma_n^2}(z_n - \theta_n)(z_n - \bar{z}_n)\right) + o_p(\sigma_n^{-2}). \quad (3.51)$$

Le résultat (3.51) avec (3.49) permet d'écrire le log-RV comme suit :

$$\log(\Lambda_R(z_n)) = \log\left(1 + R \left[\exp\left(\frac{1}{2\sigma_n^2}(z_n - \theta_n)(z_n - \bar{z}_n)\right) \exp\left(\frac{-1}{8\sigma_n^2}\right) - 1\right]\right) + o_p(\sigma_n^{-2}).$$

Il est alors immédiat de constater que lorsque  $R \rightarrow 0$  ou lorsque  $\sigma_n \rightarrow +\infty$  on a :

$$\begin{aligned} \exp(-1/8\sigma_n^2) &\rightarrow 1 \\ R \left[ \exp\left(\frac{1}{2\sigma_n^2}(z_n - \theta_n)(z_n - \bar{z}_n)\right) \exp\left(\frac{-1}{8\sigma_n^2}\right) - 1 \right] &\rightarrow 0. \end{aligned}$$

Un développement de Taylor-Maclaurin des fonctions  $\log(1+x)$  et  $\exp(x)$ , au voisinage de  $x=0$ , permet alors d'établir :

$$\log(\Lambda_R(z_n)) \xrightarrow{p} R \underbrace{\frac{1}{2\sigma_n^2}(z_n - \theta_n)(z_n - \bar{z}_n)}_{\log(\bar{\Lambda}(z_n))} + o_p(\sigma_n^{-2}). \quad (3.52)$$

**Remarque 3.8.** S'il a été considéré que le taux d'insertion  $R$  est connu, il est notable que les calculs précédents (3.49) - (3.52) sont valables pour tout  $R \in ]0, 1[$ .

### 3.5.1 Performance asymptotique du test entre hypothèses composites

Ici encore, en vertu du théorème 3.1 de la limite centrale de Lindeberg, il est possible de fournir une expression simple de la distribution de probabilité du rapport de vraisemblance  $\log(\bar{\Lambda}(z_n))$  et donc *in fine* de calculer simplement la performance asymptotique du test du RV simplifié basé sur la valeur approchée de la log-RV  $\log(\bar{\Lambda}(z_n))$  définie dans (3.52) par :

$$\log(\bar{\Lambda}(z_n)) = \frac{1}{2\sigma_n^2}(z_n - \theta_n)(z_n - \bar{z}_n). \quad (3.53)$$

Ici encore il est donc nécessaire de calculer les premiers moments de la valeur approchée du log-RV  $\log(\Lambda(z_n))$  (3.52) dans les conditions d'utilisation de cette dernière, *i.e.* lorsque  $R \rightarrow 0$  et pour tout  $n \in \mathcal{I}$ ,  $\sigma_n \rightarrow \infty$ . En utilisant d'une part le fait que la

Ici encore (3.50) revient à considérer que sur un intervalle de  $\Delta$  aussi bien que sur  $2\Delta$  on a :  $\int_{z-t}^{z+t} \phi(y) dy \approx 2t\phi(z)$ .

Le cadre d'étude  $R \rightarrow 0$  ou  $\sigma_n \rightarrow +\infty$  correspondent au "pire cas" décrit en début de section.

Par souci de clarté il a été considéré que le signal est quantifié avec un pas unitaire  $\Delta = 1$  et que le signal est pré-amplifié par  $1/\Delta$ . Aussi dans la pratique supposé que  $\sigma_n \rightarrow \infty$  revient à supposer que  $\Delta \rightarrow 0$ .

distribution  $P_{\theta_n}$  est statistiquement indépendante du LSB de  $z_n$  (ou de manière équivalente de  $z_n - \bar{z}_n$ ) et d'autre part la relation (3.50), il est immédiat de constater que :

$$\begin{aligned} \mathbb{P}[(z_n - \bar{z}_n) = 1] &= 1/2 + o(\sigma_n^{-2}) = \mathbb{P}[(z_n - \bar{z}_n) = -1] \\ \mathbb{E}[z_n - \theta_n | z_n - \bar{z}_n = 1] &= \mathbb{E}[z_n - \theta_n] = \mathbb{E}[z_n - \theta_n | z_n - \bar{z}_n = -1]. \end{aligned} \quad (3.54)$$

Il découle alors des relations (3.54) que sous l'hypothèse nulle  $\mathcal{H}_0$  on a :

$$\mu_0 = \mathbb{E}_0[(z_n - \theta_n)(z_n - \bar{z}_n)] = \frac{1}{2}\mathbb{E}_0[z_n - \theta_n] - \frac{1}{2}\mathbb{E}_0[z_n - \theta_n] + o_p(\sigma_n^{-2}) = 0 + o_p(\sigma_n^{-2}). \quad (3.55)$$

Pour calculer rigoureusement la variance  $\sigma_0^2 = \text{Var}_0[(z_n - \theta_n)(z_n - \bar{z}_n)]$ , il est notable que  $z_n = \mathcal{Q}_1[\theta_n + \xi_n] = \theta_n + \xi_n + q_n$  où  $q_n$  est une variable aléatoire représentant le bruit de quantification qui, sous la condition  $\sigma_n \gg 1$ , peut être considérée comme la réalisation d'une variable aléatoire uniformément distribuée sur l'intervalle  $[-1/2; 1/2]$ , voir [118, 119]. Cette formalisation permet d'écrire :

$$\begin{aligned} \sigma_0^2 &= \mathbb{E}_0[(z_n - \theta_n)^2(z_n - \bar{z}_n)^2] - \mu_0^2 = \mathbb{E}_0[(z_n - \theta_n)^2] + o_p(\sigma_n^{-4}) \\ &= \mathbb{E}_0[\xi_n^2] + \mathbb{E}_0[q_n^2] + 2\mathbb{E}_0[q_n \xi_n] + o_p(\sigma_n^{-4}) = \sigma_n^2 + \frac{1}{12} + o_p(\sigma_n^{-4}). \end{aligned} \quad (3.56)$$

De manière équivalente, sous l'hypothèse alternative  $\mathcal{H}_1$  on a :

$$\begin{aligned} \mu_1 &= \mathbb{E}_1[(z_n - \theta_n)(z_n - \bar{z}_n)] = \frac{1}{2}\mathbb{E}_1[z_n - \theta_n] - \frac{1}{2}\mathbb{E}_1[z_n - \theta_n] + o_p(\sigma_n^{-2}) \\ &= \left(1 - \frac{R}{2}\right)\mu_0 + \frac{R}{4}(\mathbb{E}_0[z_n - \theta_n + 1] - \mathbb{E}_0[z_n - \theta_n - 1]) + o_p(\sigma_n^{-2}) = \frac{R}{2} + o_p(\sigma_n^{-2}). \end{aligned} \quad (3.57)$$

Enfin, la variance est donnée sous l'hypothèse  $\mathcal{H}_1$  par :

$$\begin{aligned} \sigma_1^2 &= \mathbb{E}_1[(z_n - \theta_n)^2(z_n - \bar{z}_n)^2] - \mu_1^2 = \mathbb{E}_1[(z_n - \theta_n)^2] - \frac{R^2}{4} + o_p(\sigma_n^{-2}) \\ &= \left(1 - \frac{R}{2}\right)\sigma_0^2 + \frac{R}{4}(\mathbb{E}_0[(\theta_n + q_n + 1)^2] - \mathbb{E}_0[(\theta_n + q_n - 1)^2]) - \frac{R^2}{4} + o_p(\sigma_n^{-2}) \\ &= \sigma_0^2 + 2\frac{R}{4} - \frac{R^2}{4} + o_p(\sigma_n^{-2}) = \sigma_0^2 + \frac{R}{2}\left(1 - \frac{R}{2}\right) + o_p(\sigma_n^{-4}). \end{aligned} \quad (3.58)$$

Pour des raisons de clarté, le rapport insertion-sur-bruit (ISB)  $\varrho$  est défini par la quantité :

$$\varrho = \frac{R\Delta}{2\bar{\sigma}} \quad \text{avec} \quad \frac{1}{\bar{\sigma}^2} = \frac{1}{N} \sum_{n=1}^N \frac{1}{\sigma_n^2}. \quad (3.59)$$

Finalement, les relations (3.55) - (3.58) permettent de déduire de l'application du théorème 3.1 de la limite centrale de Lindeberg que :

$$\log(\bar{\Lambda}(\mathbf{Z})) = \frac{\bar{\sigma}}{\sqrt{N}} \sum_{n=1}^N \frac{(z_n - \theta_n)(z_n - \bar{z}_n)}{\sigma_n^2} \rightsquigarrow \mathcal{N}\left(\sqrt{N}\varrho; 1 + \gamma\right). \quad (3.60)$$

où  $\gamma = 1/\bar{\sigma}^2 R/2(1 - R/2)$ . Cette relation est illustrée par la figure 3.9 présentant graphiquement une comparaison des distributions  $\log(\bar{\Lambda}(\mathbf{Z}))$  et  $\log(\Lambda_R(\mathbf{Z}))$  empiriquement obtenues avec les distributions Gaussienne théoriques.

**Remarque 3.9.** Il est immédiat de constater que le RV  $\log(\Lambda_R(\mathbf{Z}))$  (3.6) et (3.11) sur le lequel repose le test PP  $\delta_R(\mathbf{Z})$  et la quantité  $\log(\bar{\Lambda}(\mathbf{Z}))$  (3.60) se déduisent l'une de l'autre par la relation :

$$\frac{R}{2} \frac{\sqrt{N}}{\bar{\sigma}} \log(\bar{\Lambda}(\mathbf{Z})) = \log(\Lambda_R(\mathbf{Z})) - \mu_0. \quad (3.61)$$

En considérant que pour tout  $n \in \mathcal{I}$  les paramètres de distribution  $\sigma_n$  et  $\theta_n$  sont connus, le test proposé est donc donné par la fonction de décision suivante :

$$\bar{\delta}(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \log(\bar{\Lambda}(\mathbf{Z})) = \sum_{n=1}^N \log(\bar{\Lambda}(z_n)) < \tau_{\alpha_0} \\ \mathcal{H}_1 & \text{si } \log(\bar{\Lambda}(\mathbf{Z})) = \sum_{n=1}^N \log(\bar{\Lambda}(z_n)) \geq \tau_{\alpha_0} \end{cases} \quad (3.62)$$

où le seuil de décision est la solution de l'équation  $\mathbb{P}_0[\log(\bar{\Lambda}(\mathbf{Z})) \geq \tau_{\alpha_0}] = \alpha_0$  explicitée dans la proposition 3.4.

**Proposition 3.4.** *Un calcul direct permet de montrer que pour tout  $\alpha \in ]0; 1[$ , en fixant*

$$\tau_{\alpha_0} = \Phi^{-1}(1 - \alpha_0), \quad (3.63)$$

où  $\Phi^{-1}(\cdot)$  est la fonction réciproque de la fonction de répartition Gaussienne  $\Phi(\cdot)$ , on a asymptotiquement, lorsque  $N \rightarrow \infty$ ,  $\alpha_0(\bar{\delta}) = \alpha_0$  et donc  $\bar{\delta} \in \mathcal{K}_{\alpha_0}$ .

*Démonstration.* La démonstration de la proposition 3.4, similaire en tout point à la démonstration 3.1, est donnée dans l'annexe 3.8.  $\square$

**Théorème 3.6.** *Supposons que le taux d'insertion  $R \in ]0; 1[$  ainsi que, pour tout  $n \in \{0, \dots, N\}$ , les paramètres  $\theta_n$  et  $\sigma_n$  sont connus. La puissance  $\beta_{\bar{\delta}}$  associée au test  $\bar{\delta}$  du rapport de vraisemblance approché (3.62) est donnée par :*

$$\beta_{\bar{\delta}} = 1 - \Phi\left(\frac{\tau_{\alpha_0} - \sqrt{N}\varrho}{1 + \gamma}\right) = 1 - \Phi\left(\frac{\Phi^{-1}(1 - \alpha_0) - \sqrt{N}\varrho}{1 + \gamma}\right). \quad (3.64)$$

*Démonstration.* La démonstration de la proposition 3.6, est similaire à la démonstration 3.2, est donnée dans l'annexe 3.8.  $\square$

La proposition 3.4 et le théorème 3.6 mettent en évidence deux propriétés intéressantes du test  $\bar{\delta}(\mathbf{Z})$ . D'une part, ni le seuil de décision ni le log-RV  $\log(\bar{\Lambda}(\mathbf{Z}))$  ne dépendent du taux d'insertion  $R$ ; cela permet une extension immédiate aux cas d'hypothèses composites (*i.e.* lorsque le taux d'insertion  $R$  n'est pas connu). D'autre part, comparativement aux tests présentés dans les sections 3.3-3.2, les expressions du seuil de décision  $\tau_{\alpha_0}$  et de la fonction de puissance  $\beta_{\bar{\delta}}$  sont grandement simplifiées.

Enfin, l'intérêt fondamental du test  $\bar{\delta}(\mathbf{Z})$  réside dans le théorème 3.7 précisant le caractère optimal de ce dernier.

**Théorème 3.7.** *Supposons que le taux d'insertion  $R \in ]0; 1[$  ainsi que, pour tout  $n \in \{0, \dots, N\}$ , les paramètres  $\theta_n$  et  $\sigma_n$  sont connus. Si  $R = R_N$ ,  $\Delta = \Delta_N$  et  $\bar{\sigma} = \bar{\sigma}_N$  évoluent en fonction de  $N$  de sorte que :*

$$\varrho_N \sqrt{N} = \frac{R_N \Delta_N}{2\bar{\sigma}_N} \sqrt{N} \rightarrow C \in \mathbb{R}_+^* \text{ lorsque } N \rightarrow \infty.$$

Alors, le test  $\bar{\delta}(\mathbf{Z})$  est AUPP au sens de la définition 2.8 dans la classe  $\mathcal{K}_{\alpha_0}$ .

*Démonstration.* La démonstration du théorème 3.7 est donnée dans l'annexe 3.8. Cette annexe permet en outre de comprendre les conditions d'application du théorème.  $\square$

La figure 3.10 illustre graphiquement la validité de l'utilisation du log-RV approché  $\log(\bar{\Lambda}(\mathbf{Z}))$ . Pour  $\sigma_n = 2.25$  on constate que la performance des tests  $\delta_R$  et  $\bar{\delta}$  sont (quasi) identiques et, plus important, correspondent parfaitement avec celles attendues. À l'inverse pour  $\sigma = 0.75$ , la figure 3.10 met en évidence une différence de puissance entre les tests (voir également la figure 3.6) et, plus important, que la fonction de puissance  $\beta_{\bar{\delta}}$  ne correspond plus à celle donnée par (3.64).



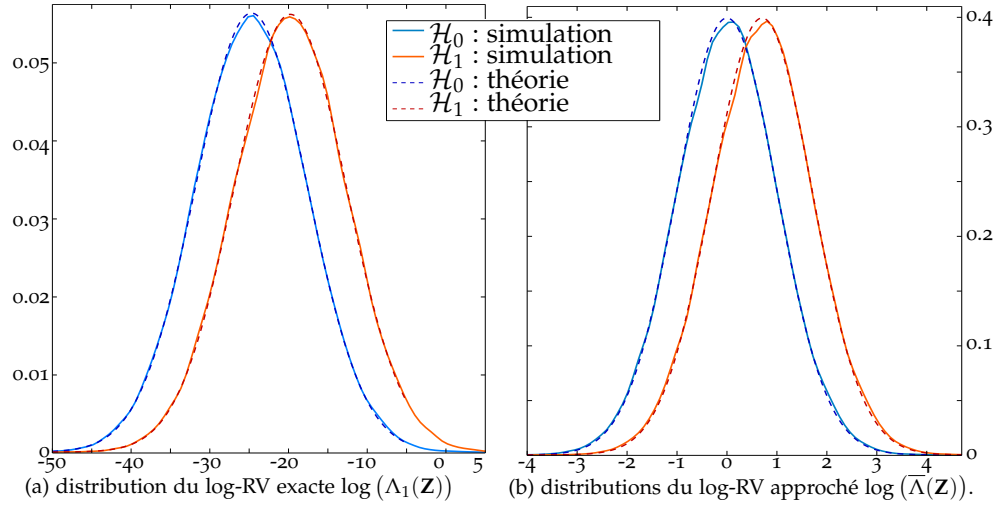


FIGURE 3.9: Distribution des statistiques de décisions  $\log(\Lambda_1(\mathbf{Z}))$  et  $\log(\bar{\Lambda}(\mathbf{Z}))$  sous  $\mathcal{H}_0$  et  $\mathcal{H}_1$  avec  $N = 1000$ ,  $R = 0.1$ ,  $\sigma_n = 2.25$  et  $\theta_n = 127.5$ .

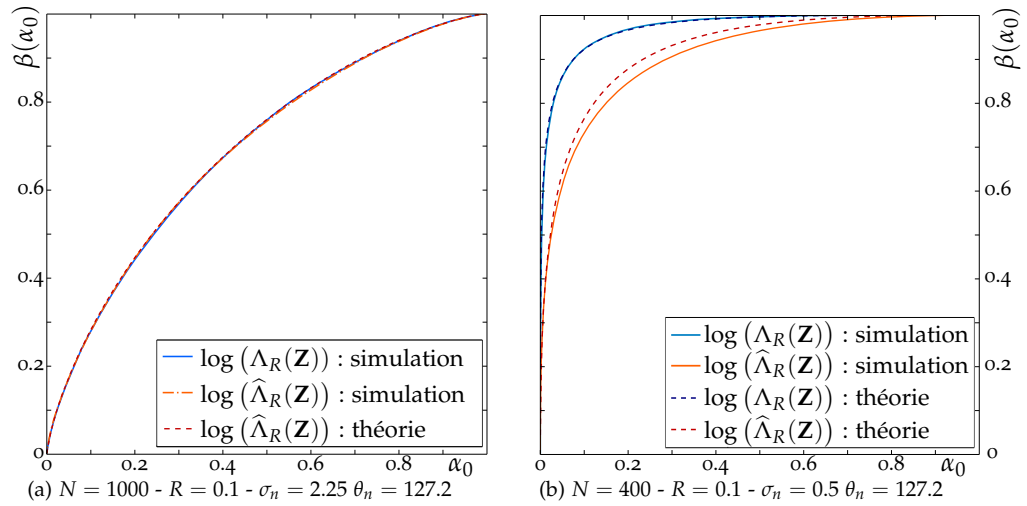


FIGURE 3.10: Courbes COR,  $\beta(\alpha_0)$ , illustrant la puissance des tests  $\delta_R$  et  $\bar{\delta}$  selon la variance du bruit.

**Remarque 3.10.** Les approches présentées dans les sections 3.5 et 3.1 sont complémentaires dans le sens où elles sont chacune spécifiquement dédiées à un cas différent. Dans des conditions de fort rapport signal-sur-bruit (typiquement  $\sigma_n < 1$ ), l'approche proposée dans les sections 3.1 et 3.4 permet d'établir le test NP et de calculer ses performances. À l'inverse dans le cas où la détection est rendue plus difficile par un faible rapport signal-sur-bruit, l'approche décrite dans la section 3.5 présente quelques simplifications permettant de proposer un test asymptotiquement optimal dont le seuil et la puissance sont plus aisément calculables. La figure 3.10 illustre notamment cette différence au travers d'une courbe COR pour différentes valeurs de paramètres.

### 3.5.2 Test en présence de paramètres de nuisance : modèle régressif de l'image

Comme cela a été fait dans la section 3.4 pour le cas où l'image  $\mathbf{Z}$  à analyser est inconnue, il est proposé de découper cette dernière en  $K$  segments contigus de  $L$  pixels consécutifs. Il est en outre supposé que pour tout  $k \in \{1, \dots, K\}$  le vecteur  $\mathbf{z}_k \in \mathcal{Z}^L$  peut être représenté par le modèle linéaire (3.38) :

$$\mathbf{z}_k = Q_1[\mathbf{y}_k] \quad , \quad \mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \boldsymbol{\zeta}_k = \boldsymbol{\theta}_k + \boldsymbol{\zeta}_k.$$

Le problème qu'il est alors souhaitable de résoudre consiste à décider entre les deux hypothèses statistiques composites définies par :

$$\begin{aligned} \mathcal{H}_0 &= \{ \mathbf{z}_k \sim Q_{\boldsymbol{\theta}_k}^R, \forall k \in \{1, \dots, K\}, \forall R \leq R^* \} \\ \text{vs } \mathcal{H}_1 &= \{ \mathbf{z}_k \sim Q_{\boldsymbol{\theta}_k}^R, \forall k \in \{1, \dots, K\}, \forall R > R^* \}, \end{aligned}$$

où la distribution de probabilités  $Q_{\boldsymbol{\theta}_k}^R$  est le produit des distributions des pixels  $\prod_{l=1}^L Q_{\theta_{k,l}}^R$  et  $R^*$  est un taux d'insertion critique préalablement fixé.

Il est alors proposé de remplacer les paramètres  $\boldsymbol{\theta}_k$  et  $\sigma_k$  par l'estimation au sens du maximum de vraisemblance de ces derniers, respectivement  $\hat{\boldsymbol{\theta}}_k$  et  $\hat{\sigma}_k$ . La fonction de décision du test  $\bar{\delta}(\mathbf{Z})$  est alors remplacée par :

$$\bar{\delta}^{\text{rvg}}(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \log(\bar{\Lambda}^{\text{rvg}}(\mathbf{Z})) = \sum_{k=1}^K \log(\bar{\Lambda}(\mathbf{z}_k; \hat{\mathbf{z}}_k)) < \tau_{\alpha_0} \\ \mathcal{H}_1 & \text{si } \log(\bar{\Lambda}^{\text{rvg}}(\mathbf{Z})) = \sum_{k=1}^K \log(\bar{\Lambda}(\mathbf{z}_k; \hat{\mathbf{z}}_k)) \geq \tau_{\alpha_0}, \end{cases} \quad (3.65)$$

où le log-RV approché  $\log(\bar{\Lambda}^{\text{rvg}}(\mathbf{Z}))$  basé sur les estimations  $\hat{\boldsymbol{\theta}}_k$  et  $\hat{\sigma}_k$  est donné par :

$$\log(\bar{\Lambda}^{\text{rvg}}(\mathbf{Z})) = \frac{\bar{\sigma}}{\sqrt{K(L-p)}} \sum_{k=1}^K \frac{(\mathbf{P}_H^\perp \mathbf{z}_k)^T (\mathbf{z}_k - \bar{\mathbf{z}}_k)}{\hat{\sigma}_k^2}, \quad (3.66)$$

et le seuil de décision est la solution de l'équation  $\sup_{R < R^*, \mathbf{x} \in R^p} \mathbb{P}_0[\log(\bar{\Lambda}(\mathbf{Z}; \hat{\mathbf{Z}})) \geq \tau_{\alpha_0}] = \alpha_0$ .

**Remarque 3.11.** Dans un souci de synthèse, les test précédemment étudiés sont résumés dans la tableau 3.3 ci dessous.

## 3.6 Relation entre le test proposé et la méthode du WS

Le lien entre l'approche proposée dans la section 3.2 et le test de O. Dabeer & al. publié dans [89] a déjà été décrit dans la remarque 3.1.

Les travaux présentés jusqu' alors ont notamment permis la conception de deux tests complémentaires  $\delta^{\text{rvg}}(\mathbf{Z})$  (3.39) et  $\bar{\delta}(\mathbf{Z})$  (3.62) dont l'utilisation pratique à l'aide d'un modèle linéaire simpliste du contenu d'une image a été décrite dans les sections 3.4 et 3.5.2 respectivement. L'expression des statistiques de décision de ces tests, resp.  $\log(\hat{\Lambda}_1(\mathbf{Z}))$  et  $\log(\bar{\Lambda}(\mathbf{Z}))$ , présente une similarité singulière avec la méthode du WS, décrite dans la section 2.3.3.3, initialement proposée dans [91] pour estimer le taux d'insertion  $R$  d'un média analysé.

notation	description
$\delta_R$	Test PP du RV entre les hypothèses simples $\mathcal{H}_0 = \{z_n \sim P_{\theta_n}\}$ et $\mathcal{H}_1 = \{z_n \sim Q_{\theta_n}^R\}$ , le taux d'insertion $R$ étant connu, ainsi que, pour tout $n \in \mathcal{I}$ , $\theta_n$ et $\sigma_n$ , voir relation (3.7).
$\delta_R^{\text{rvg}}$	Test du RVG entre les hypothèses composites $\mathcal{H}_0 = \{z_n \sim Q_{\theta_n}^{R_0}, R_0 \leq R^*\}$ et $\mathcal{H}_1 = \{z_n \sim Q_{\theta_n}^{R_1}, R_1 > R^*\}$ en présence du paramètre de nuisance linéaire $\theta$ inconnu.
$\delta_{R^*}$	Test LAUPP entre les hypothèses simples $\mathcal{H}_0 = \{z_n \sim Q_{\theta_n}^{R^*}\}$ et $\mathcal{H}_1 = \{z_n \sim Q_{\theta_n}^{R_1^*}, R_1^*(N) = R^* + cN^{-1/2}\}$ nécessitant pour tout $n \in \mathcal{I}$ , la connaissance de $\theta_n$ et $\sigma_n$ , voir relation (3.35).
$\bar{\delta}$	Test AUPP entre les hypothèses composites $\mathcal{H}_0 = \{z_n \sim P_{\theta_n}\}$ et $\mathcal{H}_1 = \{z_n \sim Q_{\theta_n}^R, R > 0\}$ coïncidant quelque soit $R$ avec le test du RV sous condition de faible pas de quantification.
$\bar{\delta}^{\text{rvg}}$	Test entre les hypothèses composites $\mathcal{H}_0 = \{z_n \sim P_{\theta_n}\}$ et $\mathcal{H}_1 = \{z_n \sim Q_{\theta_n}^R, R > 0\}$ en présence du paramètre de nuisance linéaire $\theta$ coïncidant quelque soit $R$ avec le test du RVG sous condition de faible pas de quantification, voir relation (3.62).

TABLE 3.3: Rappels des différents tests proposés dans cette section ainsi que leur notation et leur caractère d'optimalité éventuel.

Afin de mettre en évidence la similarité entre les tests proposés et le WS, le log-RV  $\log(\hat{\Lambda}_1(\mathbf{Z}))$  utilisé par le test du RVG  $\delta^{\text{rvg}}(\mathbf{Z})$  est formulé, de manière strictement équivalente à (3.41) de la façon suivante :

$$\log(\Lambda_1^{\text{rvg}}(\mathbf{Z})) = \underbrace{\sum_{k=1}^K \hat{\sigma}_k^{-2}}_{\text{poids}} \underbrace{\sum_{l=1}^L (z_{k,l}^{(0)} - \hat{\theta}_{k,l})}_{\text{résidus}} \underbrace{(z_{k,l}^{(0)} - 1/2)}_{\text{lsb}(z_{k,l}) - 1/2} - \underbrace{\sum_{k=1}^K \frac{1}{8\hat{\sigma}_k^2}}_{\text{terme correctif}}. \quad (3.67)$$

De façon analogue, le log-RV approché  $\log(\bar{\Lambda}_1(\mathbf{Z}))$  utilisé dans le test  $\bar{\delta}(\mathbf{Z})$  peut être donné, de manière strictement équivalente à (3.66) par :

$$\log(\bar{\Lambda}(\mathbf{Z})) = \underbrace{\frac{\bar{\sigma}}{\sqrt{K(L-p)}}}_{\text{normalisation}} \sum_{k=1}^K \underbrace{\hat{\sigma}_k^{-2}}_{\text{poids}} \sum_{l=1}^L \underbrace{(z_{k,l} - \hat{\theta}_{k,l})}_{\text{résidus}} \underbrace{(z_{k,l} - \bar{z}_{k,l})}_{2 \cdot \text{lsb}(z_{k,l}) - 1}. \quad (3.68)$$

Il est immédiat de constater que les log-RV explicités ci-dessus présentent une forte similarité : dans les deux cas, l'information issue de l'image se trouve concentrée dans un produit scalaire entre un vecteur de résidus  $\mathbf{z}_k - \hat{\mathbf{z}}_k$  et un facteur de signe  $\text{lsb}(z_{k,l}) - 1/2$ , le résultat étant pondéré par  $\sigma^{-2}$  de sorte que les observations les moins bruitées aient une influence plus importante. Il est cependant notable que les expressions (3.67) et (3.68) diffèrent par la présence du terme correctif  $1/8\hat{\sigma}_k^2$  dans la première, qui ne dépend pas du taux d'insertion mais doit être pris en compte pour le calcul du seuil de décision, et par la présence du facteur de normalisation  $\frac{\bar{\sigma}}{\sqrt{K(L-p)}}$  dans la seconde expression.

En ré-écrivant l'équation (2.21), la comparaison avec la méthode du WS est immédiate :

$$\hat{R} = \underbrace{\frac{2}{N}}_{\text{normalisation}} \sum_{m=1}^M \sum_{n=1}^N \underbrace{w_{m,n}}_{\text{poids}} \underbrace{(z_{m,n} - \hat{c}_{m,n})}_{\text{résidus}} \underbrace{(z_{m,n} - \bar{z}_{m,n})}_{2 \cdot \text{lsb}(z_{k,l}) - 1}, \quad (3.69)$$

où  $\hat{c}_{m,n} = \mathcal{F}(\mathbf{Z})_{m,n}$  représente une opération de filtrage discret s'appuyant sur un voisinage  $\mathcal{V}_{m,n}$  du pixel  $z_{m,n}$ , et où les facteurs de pondérations  $w_{m,n}$  sont définis de sorte que  $\sum_{m,n} w_{m,n} = 1$  par la relation  $w_{m,n} \propto (\alpha + \hat{\sigma}_{k,n}^2)^{-1}$ ,  $\alpha > 0$  où  $\hat{\sigma}_{k,n}^2$  représente la variance "locale" des pixels de  $\mathcal{V}_{m,n}$ .

La structure des log-RV  $\log(\hat{\Lambda}_1(\mathbf{z}_k))$  et  $\log(\bar{\Lambda}(\mathbf{z}_k))$  sus-décrite se retrouve dans l'expression de la statistique utilisée par le WS (3.69). Cependant, contrairement à l'approche conduisant à la méthode du WS (décrite dans [91, 92] et dans la section 2.3.3.3 du présent manuscrit) trois différences fondamentales sont à souligner :

1. D'une part, l'utilisation des méthodes de décision statistique permet d'établir théoriquement les facteurs de "pondération"  $\hat{\sigma}_k^{-2}$ , par opposition à la méthode du WS pour laquelle les termes  $w_{m,n}$  sont choisis empiriquement (la valeur du terme  $\alpha$  est déterminée dans [91, 92] à partir de résultats numériques).
2. D'autre part, l'estimation de la valeur du pixel avant insertion  $\hat{c}_{m,n}$  est remplacée par une estimation de l'intensité lumineuse  $\hat{\theta}_{m,n}$  (à valeur réelle) mesurée par le photo-détecteur dépendant des paramètres physiques décrivant l'acquisition de l'image.
3. Enfin, la méthodologie proposée reposant sur la théorie de la décision statistique. une justification de la statistique de décision du WS.

### 3.7 Extension de la méthodologie proposée

Il est rappelé en préambule de cette section 3.7 que la détection d'informations cachées par correspondance de LSB n'est pas exactement au centre des travaux présentés. En effet, la méthodologie proposée est en premier lieu appliquée à la substitution de LSB. Néanmoins cette méthodologie est assez aisément transposable à la correspondance de LSB. Les quelques résultats présentés dans cette section constituent davantage un élargissement des travaux proposés. Le lecteur comprendra donc que les résultats de l'étude statistique présentée ne sont, pas encore, aussi aboutis que pour le cas de la substitution de LSB ; le caractère optimal du test proposé demeure notamment incertain.

Il est ici encore considéré un média de couverture  $\mathbf{C}$  dont les échantillons sont le résultat, pour tout  $n \in \mathcal{I}$  de la quantification :  $c_n = Q[\theta_n + \xi_n]$ . La loi de distribution des pixels avant insertion est donnée par  $P_{\theta_n} = \{P_{\theta_n}[0], \dots, P_{\theta_n}[2^b - 1]\}$  où pour tout  $k \in \mathcal{Z}$  la probabilité  $P_{\theta_n}[k]$  est explicitée dans (3.1) et (3.2). Enfin, comme cela a été évoqué dans le proposition 2.1, il est admis que :

1. La valeur des bits insérés suit une loi de distribution uniforme :  $b_s \sim \mathcal{B}(1, 1/2)$
2. La probabilité d'insertion est identique pour tous les pixels, on a alors  $\forall n \in \mathcal{I}$  :

$$\begin{cases} \mathbb{P}[s_n = c_n] & = (1 - R) \\ \mathbb{P}[s_n = c_n + \text{ins}(b_s, c_n)] & = R, \end{cases}$$

où  $\text{ins}(b_s, c_n)$  représente la valeur ajoutée en fonction de la valeur du pixel de couverture  $c_n$  et du bit d'information à insérer  $b_s$ .

La spécificité de l'insertion par correspondance de LSB, par rapport à la substitution de LSB, réside dans la fonction d'insertion  $\text{ins} : \{0, 1\} \mapsto \{-1, 0, 1\}$ . Lorsque la valeur du bit à insérer  $b_s$  est celle du LSB du pixel  $c_n$  avant insertion, aucune modification n'est nécessaire, *i.e.* si  $b_s = c_{n,0}$ , alors  $\text{ins}(b_s, c_n) = 0$ . À l'inverse si la valeur du bit d'information à insérer est différente de celle du LSB du pixel  $c_n$  avant insertion alors :

$$\begin{cases} \mathbb{P}[\text{ins}(b_s, c_n) = 1] & = 1/2 \\ \mathbb{P}[\text{ins}(b_s, c_n) = -1] & = 1/2. \end{cases}$$

Il est ainsi possible de modéliser statistiquement la valeur des pixels d'un média stéganographié par correspondance avec un taux d'insertion  $R$  par :

$$\begin{cases} \mathbb{P}[s_n=c_n] = (1 - \frac{R}{2}) \\ \mathbb{P}[s_n=c_n + 1] = \mathbb{P}[s_n=c_n - 1] = R/4. \end{cases}$$

Il en découle alors que la loi de probabilité des pixels après insertion est donnée par  $T_{\theta_n}^R = (T_{\theta_n}^R[0], \dots, T_{\theta_n}^R[2^b - 1])$  avec pour tout  $k \in \mathcal{Z}$  :

$$T_{\theta_n}^R[k] = \frac{R}{4} (P_{\theta_n}[k-1] + P_{\theta_n}[k+1]) + \left(1 - \frac{R}{2}\right) P_{\theta_n}[k]. \quad (3.70)$$

Il est en conséquence souhaitable, lors de l'analyse d'un média  $\mathbf{Z}$ , de décider entre les deux hypothèses statistiques suivantes :

$$\mathcal{H}_0 = \{z_n \sim P_{\theta_n}, \forall n \in \mathcal{I}\} \quad \text{vs} \quad \mathcal{H}_1 = \{z_n \sim T_{\theta_n}^R, \forall n \in \mathcal{I}\}. \quad (3.71)$$

En considérant, dans un premier temps, le cas le plus simple où le taux d'insertion  $R$  ainsi que, pour tout  $n \in \mathcal{I}$ , les paramètres  $\theta_n$  et  $\sigma_n$  sont connus, il est immédiat de constater que le problème de décision statistique (3.71) consiste à tester deux hypothèses simples. En vertu du Lemme 2.1 de Neyman-Pearson, le test le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  existe et est donné par :

$$\delta_R^\pm(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \Lambda_R^\pm(\mathbf{Z}) < \tau \\ \mathcal{H}_1 & \text{si } \Lambda_R^\pm(\mathbf{Z}) \geq \tau, \end{cases} \quad (3.72)$$

où le rapport de vraisemblance  $\Lambda_R^\pm(\mathbf{Z})$  peut s'écrire, compte-tenu de l'indépendance statistique des pixels :

$$\Lambda_R^\pm(\mathbf{Z}) = \prod_{n=1}^N \Lambda_R^\pm(z_n) = \prod_{n=1}^N \frac{R}{4} \frac{P_{\theta_n}[z_n - 1] + P_{\theta_n}[z_n + 1]}{P_{\theta_n}[z_n]} + \left(1 - \frac{R}{2}\right). \quad (3.73)$$

Il est aisé de constater que  $\Lambda_R^\pm(z_n)$  ne dépend des échantillons  $z_n$  qu'au travers du rapport de vraisemblance :

$$\Lambda_R^\pm(z_n) = \frac{R}{4} \frac{P_{\theta_n}[z_n - 1] + P_{\theta_n}[z_n + 1]}{P_{\theta_n}[z_n]}.$$

En utilisant la définition de la distribution  $P_{\theta_n}$  (3.2), un calcul élémentaire permet d'obtenir la relation :

$$\begin{aligned} \frac{P_{\theta_n}[z_n-1]}{P_{\theta_n}[z_n]} &= \exp\left(-\frac{1}{2\sigma_n^2}\right) \exp\left(\frac{\theta_n-k}{\sigma_n^2}\right) \\ \frac{P_{\theta_n}[z_n+1]}{P_{\theta_n}[z_n]} &= \exp\left(-\frac{1}{2\sigma_n^2}\right) \exp\left(\frac{k-\theta_n}{\sigma_n^2}\right). \end{aligned} \quad (3.74)$$

En utilisant la relation (3.74), le rapport de vraisemblance  $\Lambda_R^\pm(z_n)$  (3.73) peut s'écrire :

$$\Lambda_R^\pm(z_n) = \frac{R}{4} \exp\left(-\frac{1}{2\sigma_n^2}\right) \left[ \exp\left(\frac{(k-\theta_n)}{\sigma_n^2}\right) + \exp\left(\frac{-(k-\theta_n)}{\sigma_n^2}\right) \right]. \quad (3.75)$$

Ici encore, le nombre élevé de pixels présents dans une image analysée rend pertinent l'utilisation du théorème 3.1 de la limite centrale de Lindeberg. Le problème de détection d'informations cachées par correspondance de LSB est alors asymptotiquement assimilé à un problème de décision statistique entre deux hypothèses Gaussiennes de paramètres de distributions distincts. L'expression de ces paramètres nécessite de calculer les deux premiers moments du rapport de vraisemblance  $\Lambda_R^\pm(z_n)$  (3.75). En considérant, dans un premier temps, le problème de décision statistique (3.70) dans

le cas où  $R = 1$ , un calcul permet immédiatement de ré-écrire le rapport de vraisemblance  $\Lambda_1^\pm(z_n)$  comme suit :

$$\begin{aligned} \Lambda_1^\pm(z_n) &= \frac{1}{4} \exp\left(-\frac{1}{2\sigma_n^2}\right) \left[ \exp\left(\frac{k-\theta_n}{\sigma_n^2}\right) + \exp\left(\frac{\theta_n-k}{\sigma_n^2}\right) \right] \\ \Leftrightarrow \log(\Lambda_1^\pm(z_n)) &= -\frac{1}{2\sigma_n^2} - \log(4) + \log\left(\left[ \exp\left(\frac{k-\theta_n}{\sigma_n^2}\right) + \exp\left(\frac{\theta_n-k}{\sigma_n^2}\right) \right]\right). \end{aligned} \quad (3.76)$$

Sous l'hypothèse  $\mathcal{H}_0$ , un calcul rigoureux montre que l'espérance mathématique de la log-vraisemblance  $\log(\Lambda_1^\pm(z_n))$  est donnée, en omettant les termes correctifs (3.2), pour tout  $n \in \mathcal{I}$  par :

$$\begin{aligned} \mu_0 &= \mathbb{E}_0[\log(\Lambda_1^\pm(z_n))] = -\frac{1}{2\sigma_n^2} - \log(4) \\ &+ \sum_{k \in \mathcal{Z}} \left[ \Phi\left(\frac{k+1/2-\theta_n}{\sigma}\right) - \Phi\left(\frac{k-1/2-\theta_n}{\sigma}\right) \right] \log\left(\exp\left(\frac{k-\theta_n}{\sigma_n^2}\right) + \exp\left(\frac{\theta_n-k}{\sigma_n^2}\right)\right). \end{aligned} \quad (3.77)$$

Sous l'hypothèse  $\mathcal{H}_1$ , la loi de probabilité des pixels après insertion (3.71) permet d'obtenir :

$$\begin{aligned} \mathbb{P}_1[z_n = k] &= \frac{1}{2} \left[ \Phi\left(\frac{k+1/2-\theta_n}{\sigma}\right) - \Phi\left(\frac{k-1/2-\theta_n}{\sigma}\right) \right] \\ &+ \frac{1}{4} \left[ \Phi\left(\frac{k+3/2-\theta_n}{\sigma}\right) - \Phi\left(\frac{k+1/2-\theta_n}{\sigma}\right) + \Phi\left(\frac{k-1/2-\theta_n}{\sigma}\right) - \Phi\left(\frac{k-3/2-\theta_n}{\sigma}\right) \right] \\ &= \frac{1}{4} \left[ \Phi\left(\frac{k+3/2-\theta_n}{\sigma}\right) + \Phi\left(\frac{k+1/2-\theta_n}{\sigma}\right) - \Phi\left(\frac{k-1/2-\theta_n}{\sigma}\right) - \Phi\left(\frac{k-3/2-\theta_n}{\sigma}\right) \right]. \end{aligned}$$

Il en découle alors que de manière équivalente, l'espérance mathématique de la log-vraisemblance  $\log(\Lambda_1^\pm(z_n))$  est donnée, sous l'hypothèse  $\mathcal{H}_1$ , pour tout  $n \in \mathcal{I}$  par :

$$\begin{aligned} \mu_1 &= \mathbb{E}_1[\log(\Lambda_1^\pm(z_n))] = -\frac{1}{2\sigma_n^2} - \log(4) \sum_{k \in \mathcal{Z}} \frac{1}{4} \left[ \Phi\left(\frac{k+3/2-\theta_n}{\sigma}\right) + \Phi\left(\frac{k+1/2-\theta_n}{\sigma}\right) \right. \\ &\left. - \Phi\left(\frac{k-1/2-\theta_n}{\sigma}\right) - \Phi\left(\frac{k-3/2-\theta_n}{\sigma}\right) \right] \log\left(\exp\left(\frac{k-\theta_n}{\sigma_n^2}\right) + \exp\left(\frac{\theta_n-k}{\sigma_n^2}\right)\right). \end{aligned} \quad (3.78)$$

Les expressions (3.77) - (3.78) des espérances mathématiques ne sont pas aisément calculables analytiquement pour  $\theta_n$  et  $\sigma_n$  donnés. Cependant le calcul numérique de ces quantités ne pose pas de problème majeur.

Les relations (3.77) et (3.78) sont graphiquement illustrées dans la figure 3.11a. Les valeurs approchées des deux premiers moments sont calculées en supposant que  $\forall n \in \mathcal{I}$ ,  $\theta_n = k \in \mathcal{Z}$  i.e. en considérant que les lois de distributions des pixels, sous  $\mathcal{H}_0$  et sous  $\mathcal{H}_1$ , sont "symétriques" par rapport à  $\theta_n$ . cela correspond à négliger, dans le calcul du rapport de vraisemblance, la quantification au sens étudié dans la section 3.2.

De façon analogue aux calculs (3.77) et (3.78), les variances du RV  $\log(\Lambda_1^\pm(z_n))$  sous les hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$  sont données par :

$$\sigma_0^2 = \sum_{k \in \mathcal{Z}} P_{\theta_n}[k] \log\left(\exp\left(\frac{k-\theta_n}{\sigma_n^2}\right) + \exp\left(\frac{\theta_n-k}{\sigma_n^2}\right)\right)^2 - m_0^2 \quad (3.79)$$

$$\sigma_1^2 = \sum_{k \in \mathcal{Z}} Q_{\theta_n}^R[k] \log\left(\exp\left(\frac{k-\theta_n}{\sigma_n^2}\right) + \exp\left(\frac{\theta_n-k}{\sigma_n^2}\right)\right)^2 - m_1^2, \quad (3.80)$$

Les relations (3.79) et (3.80) sont graphiquement illustrées dans la figure 3.11b. Les figures 3.11 montrent que comme pour le cas de la substitution de LSB, l'impact de la quantification sur la valeur du RV  $\log(\Lambda_1^\pm(z_n))$  peut être important, notamment lorsque  $\forall n \in \mathcal{I}$ ,  $\theta_n = \theta$  et lorsque  $\forall n \in \mathcal{I}$ ,  $\sigma_n \lesssim 1$ .

**Taux d'insertion  $R \in ]0;1[$  quelconque**

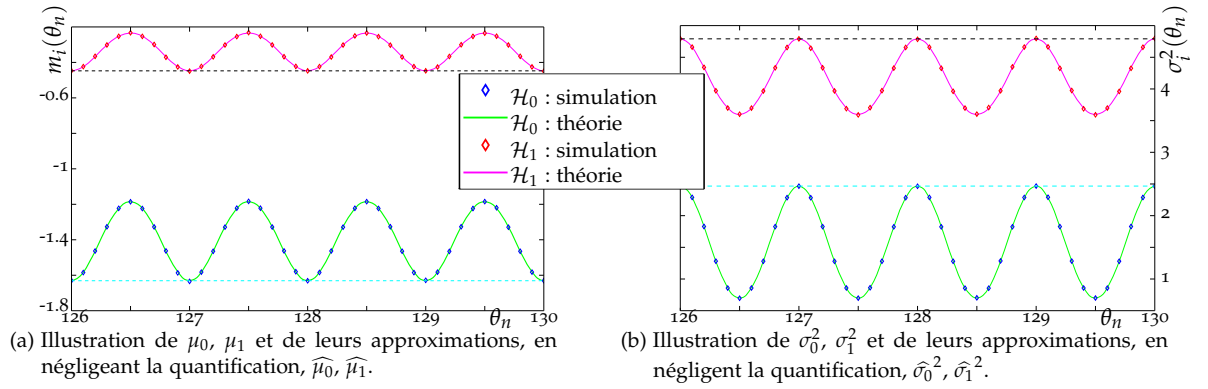


FIGURE 3.11: Représentations graphiques des moments (3.77) - (3.80) et de leurs approximations, en négligeant la quantification, en fonction de  $\theta_n$  et pour  $\sigma_n = 0.5$ . Les valeurs obtenues par simulation Monte Carlo ont été calculées avec  $10^6$  répétitions.

Pour la correspondance de LSB, il est également proposé d'étudier le cas plus général d'un taux d'insertion  $R \in [0; 1]$  quelconque et connu, en utilisant le RV  $\log(\Lambda_1^\pm(z_n))$  initialement étudié pour le cas  $R = 1$ . Le principe demeure toujours de calculer les deux premiers moments de ce RV sous les deux hypothèses simples :

$$\mathcal{H}_0 = \{z_n \sim P_{\theta_n}, \forall n \in \mathcal{I}\} \text{ vs } \mathcal{H}_1 = \{z_n \sim T_{\theta_n}^R, \forall n \in \mathcal{I}\}.$$

Il est immédiat de constater que l'hypothèse nulle, d'un média sain, n'est pas modifiée ; les deux moments du RV  $\log(\Lambda_1^\pm(z_n))$  sont donc inchangés sous l'hypothèse  $\mathcal{H}_0$ . Sous l'hypothèse  $\mathcal{H}_1$ , il découle du théorème des probabilités totales et des relations (3.77) et (3.78) :

$$\mu_R = \mathbb{E}_R [\log(\Lambda_1^\pm(z_n))] = R\mu_1 + (1 - R)\mu_0. \quad (3.81)$$

où les valeurs  $\mu_0$  et  $\mu_1$  sont explicitées par les relations (3.77) et (3.78).

De manière analogue, l'application de la loi des variances totales aux résultats (3.77) - (3.80) permet d'obtenir par un calcul direct :

$$\sigma_R^2 = R(\sigma_1^2 + \mu_1^2) + (1 - R)(\sigma_0^2 + \mu_0^2) - (R\mu_1 + (1 - R)\mu_0)^2. \quad (3.82)$$

Pour clore la présentation de l'application de la méthodologie proposée au cas de correspondance de LSB, la proposition 3.5 et le théorème 3.8 fournissent respectivement la valeur du seuil de décision selon la probabilité de fausse alarme acceptable et l'expression qui en découle de la puissance du test  $\delta_R^\pm$ .

**Proposition 3.5.** *Un calcul direct permet de montrer que pour tout  $\alpha_0 \in ]0; 1[$ , en fixant :*

$$\tau_{\alpha_0} = (\sqrt{N}\sigma_0)\Phi^{-1}(1 - \alpha_0) + N\mu_0, \quad (3.83)$$

*on a asymptotiquement, lorsque  $N \rightarrow \infty$ ,  $\alpha_0(\delta_R^\pm) = \alpha_0$  et donc  $\delta_R^\pm \in \mathcal{K}_{\alpha_0}$ .*

Enfin, une fois la valeur du seuil  $\tau_{\alpha_0}$  déterminée, le théorème 3.8 fournit une expression de la fonction de puissance  $\beta_{\delta_R^\pm}$  associée au test  $\delta_R^\pm$  le plus puissant de la classe  $\mathcal{K}_{\alpha_0}$ .

**Théorème 3.8.** *Supposons que le taux d'insertion  $R \in ]0; 1[$  ainsi que, pour tout  $n \in \{0, \dots, N\}$ , les paramètres  $\theta_n$  et  $\sigma_n$  sont connus. La puissance  $\beta_{\delta_R^\pm}$  associée au test  $\delta_R^\pm$  (3.42) est approximativement donnée par :*

$$\beta_{\delta_R^\pm} = 1 - \Phi\left(\frac{\tau_{\alpha_0} - N\mu_R}{\sigma_R}\right) = 1 - \Phi\left(\frac{\sigma_0}{\sigma_R}\Phi^{-1}(1 - \alpha_0) + \frac{\sqrt{N}(\mu_0 - \mu_1)}{\sigma_R}\right), \quad (3.84)$$

*où les quantités  $\mu_i$  et  $\sigma_i^2$  sont définies par les équations (3.77) - (3.82).*

Il est rappelé que cette expression de la fonction de puissance est approchée dans le sens où les moments ont été calculés en négligeant la quantification.

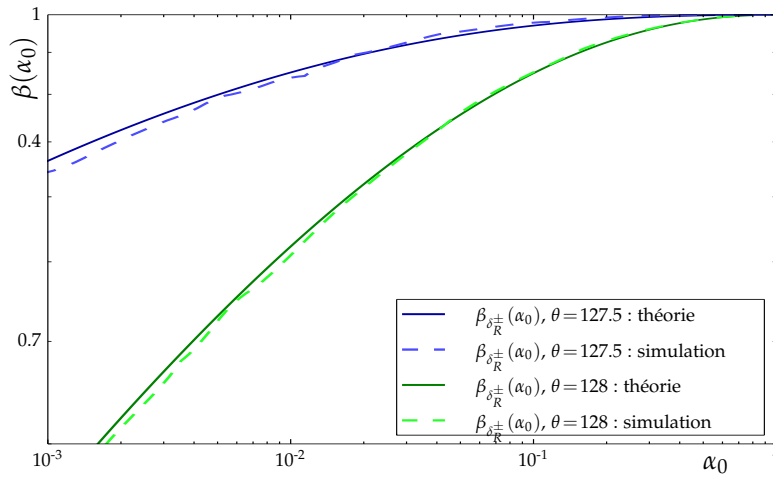


FIGURE 3.12: Représentation graphique de la fonction de puissance du détecteur  $\beta_{\delta_R^\pm}$  en fonction de la probabilité de fausse alarme  $\alpha_0$  (courbes COR) pour  $N = 1000$  pixels,  $R = 0.1$ ,  $\sigma_n = 0.5$  et  $\theta = \{127.5; 128\}$ . Les valeurs théoriques sont données par (3.84) et les valeurs empirique sont obtenues par simulation Monte Carlo avec  $10^6$  répétitions.

La fonction de puissance  $\beta_{\delta_R^\pm}$  est graphiquement illustrée dans la figure 3.12 par la comparaison des puissances empiriquement obtenues et de celles théoriquement établies par le théorème 3.8 pour deux valeurs de  $\theta = \{127.5; 128\}$ . Il est immédiat de constater que d'une part, la puissance de détection est supérieure pour le cas  $\theta = 127.5$  que pour  $\theta = 128$ ; le premier cas étant plus aisément "détectable", par exemple, la distance  $D_{KL}(P_\theta, T_\theta)$  est maximale lorsque  $\theta_n = k + 1/2$ ,  $k \in \mathbb{Z}$ . Par ailleurs, il est notable que la puissance établie théoriquement dans le théorème 3.8 est confirmée par les résultats numériquement obtenus.

## Conclusions du chapitre :

Dans le présent chapitre 3, le problème de détection d'informations cachées par substitution de LSB est formalisé comme un test statistique entre deux hypothèses. Dans le cas où les paramètres de distribution de tous les pixels  $\theta = (\theta_1, \dots, \theta_N)^T$  et  $\sigma = (\sigma_1, \dots, \sigma_N)^T$  ainsi que le taux d'insertion  $R$  sont connus, l'expression de la fonction de puissance du test le plus puissant (PP) permet de mettre en évidence l'impact important que peut avoir la quantification des données sur les performances du test (et sur la valeur du seuil permettant le respect d'une contrainte sur la probabilité de fausse alarme). Dans le cas où le taux d'insertion n'est pas connu, les hypothèses sont donc composites, la conception d'un test localement optimal (test LAUPP) permet de montrer que la perte de puissance, lorsque le taux d'insertion testé ne correspond pas au taux d'insertion effectif, peut être importante.

À l'inverse, une approche différente, consistant à négliger l'impact de la quantification dans le calcul du rapport de vraisemblance, permet la conception d'un test asymptotiquement le plus puissant uniformément par rapport au taux d'insertion  $R \in ]0; 1]$  (test AUPP) lorsque  $\varrho_N \sqrt{N} = \frac{R_N \Delta_N}{2\sigma_N} \sqrt{N} \rightarrow C \in \mathbb{R}_+^*$ . En outre, l'expression analytique de la valeur du seuil, selon le niveau de fausse alarme souhaité, et de la fonction puissance de ce test sont explicitées de façon simple.



Enfin, une brève étude statistique ouvrant la voie à l'extension de la méthodologie proposée pour la détection d'informations cachées par correspondance de LSB conclut le chapitre. L'expression du rapport de vraisemblance (RV) sur lequel repose ce test permet de comprendre les difficultés supplémentaires que présente la substitution de LSB ; les paramètres de distribution des pixels  $\theta$  et  $\sigma$  interviennent de façon complexe. Comparativement aux méthodes de détection d'informations cachées précédemment publiées dans la littérature, l'originalité de la méthodologie proposée réside dans :

1. la formalisation du problème de la stéganalyse comme un test statistique entre deux hypothèses (simples ou composites) en présence de paramètres de nuisance représentant le contenu de l'image ;
2. l'utilisation de la théorie de la décision statistique permet la conception de tests garantissant le respect d'une contrainte sur la probabilité de fausse alarme et dont le caractère optimal est démontré ;
3. la prise en compte de la quantification dans la caractérisation statistique de la performances des tests.

D'un point de vue applicatif, cette approche ouvre la voie à la conception de tests répondant aux contraintes exigées dans le cadre du projet ANR RIC.

### 3.8 Annexes du chapitre 3 : démonstration du théorème 3.7

Le but principal de cette annexe, outre la démonstration du théorème 3.7, est de décrire l'approche proposée et les conditions d'applications du théorème.

L'approche proposée consiste à admettre que le pas de quantification  $\Delta_N = \Delta(N)$  dépend du nombre d'observations de sorte que, lorsque  $N \rightarrow \infty$ , les deux hypothèses testées convergent l'une vers l'autre, de façon non dégénérée

**Assertions 3.2.** Plus formellement, les hypothèses permettant l'application du théorème 3.7 sont détaillées ci-dessous

A-1 le paramètre vectoriel  $\theta = (\theta_1, \dots, \theta_N)^T$  est borné, i. e.  $\theta \in \Theta \subset \mathbb{R}^N$  ;

A-2 puisque le pas de quantification  $\Delta_N$  dépend de  $N$ , le nombre de seuil de quantification  $p_N$ , tel que  $\mathcal{Z} = \mathbb{F}_{p_N} = \{0, \dots, p_N - 1\}$  doit permettre de vérifier :

$$\forall y_n \in \mathbb{R}, \lim_{N \rightarrow \infty} \tilde{y}_n = \Delta_N Q_{\Delta_N}(y_n) = y_n.$$

Notons qu'ici  $Q_{\Delta_N}(y_n)$  représente l'indice de quantification et  $\tilde{y}_n = \Delta_N Q_{\Delta_N}(y_n)$  la valeur quantifiée de  $y_n$  ; cette distinction n'a pas été considérée jusqu'à présent car il a été considéré que  $\Delta = 1$ .

A-3 enfin, il est supposé que le pas de quantification est choisi de sorte que :

$$q_N \sqrt{N} = \frac{R \Delta_N}{2 \bar{\sigma}} \sqrt{N} \rightarrow C \in \mathbb{R}_+^*$$

La première assertion A-1 semble assez naturelle mais permet d'éviter des difficultés théoriques n'ayant pas un grand intérêt en pratique. De manière analogue, la seconde assertion A-2 stipule, moins formellement, que malgré la diminution du pas de quantification  $\Delta_N$ , les phénomènes de saturation du quantificateur demeurent négligeables ; en vertu de l'assertion A-1, l'assertion A-2 peut être vérifiée si  $p_N \Delta_N \rightarrow \infty$  lorsque  $\Delta_N \rightarrow 0$ . Enfin, similairement à l'approche présentée dans la section 3.3, le but de l'assertion A-3 est d'éviter qu'asymptotiquement le problème de décision statistique ne soit trivial ou insoluble.

Les calculs (3.55) - (3.58) demeurent valable pour le cas d'un quantificateur de pas  $\Delta_N$  quelconque, en substituant  $\sigma_n^2/\Delta_N$  à  $\sigma_n^2$ . Les premiers moments sont donc données par :

$$\begin{aligned} \mu_0 &= 0 + o_p\left(\frac{\Delta_N^2}{\sigma_n^2}\right) & \sigma_0^2 &= \sigma_n^2 + \frac{\Delta_N^2}{12} + o_p\left(\frac{\Delta_N^2}{\sigma_n^4}\right) = \sigma_n^2 + O_p\left(\frac{\Delta_N^2}{\sigma_n^4}\right) \\ \mu_1 &= \frac{R \Delta_N}{2} + o_p\left(\frac{\Delta_N^2}{\sigma_n^2}\right) & \sigma_1^2 &= \sigma_0^2 + \Delta_N^2 \frac{R}{2} \left(1 - \frac{R}{2}\right) + o_p\left(\frac{\Delta_N^2}{\sigma_n^4}\right) = \sigma_0^2 + O_p\left(\frac{\Delta_N^2}{\sigma_n^4}\right). \end{aligned}$$

De même la convergence asymptotique de la distribution du RV  $\log(\bar{\Lambda}(\mathbf{Z}))$  (3.60), découlant du théorème (3.1) de la limite centrale de Lindeberg, devient, en tenant compte du pas de quantification  $\Delta_N$  :

$$\log(\bar{\Lambda}(\mathbf{Z})) = \frac{\bar{\sigma}}{\sqrt{N}} \sum_{n=1}^N \frac{(z_n - \theta_n)(z_n - \bar{z}_n)}{\sigma_n^2} \rightsquigarrow \mathcal{N}\left(q\sqrt{N}; 1\right),$$

où, compte-tenu du fait que  $\Delta_N \rightarrow 0$  lorsque  $N \rightarrow \infty$  le terme correctif (3.60)  $\gamma = \Delta_N^2/\bar{\sigma}^2 R/2(1 - R/2) = O_p(\Delta_N^2/\sigma_n^4)$  est négligé.

Enfin, la démonstration du caractère AUPP découle de la relation (3.61), qui en tenant compte de la quantification devient :

$$\frac{R \Delta_N}{2} \frac{\sqrt{N}}{\bar{\sigma}} \log(\bar{\Lambda}(\mathbf{Z})) = \log(\Lambda_R(\mathbf{Z})) - \mu_0.$$

En vertu du théorème de Slutsky [107, théorème 11.2.11] le test  $\bar{\delta}(\mathbf{Z})$  est asymptotiquement équivalent au test le plus puissant PP  $\delta_R(\mathbf{Z})$  et cela quel que soit le taux d'insertion  $R$ . Le test  $\bar{\delta}(\mathbf{Z})$  est donc AUPP, ce qui clôt la démonstration.



Deuxième partie

ACQUISITION DES IMAGES NATURELLES ET  
MODÉLISATION LOCALE DU CONTENU



---

# Processus de formation des images : Description de la chaîne d'acquisition des images

---

# 4

*« Le savant n'étudie pas la nature parce que cela est utile ;  
il l'étudie parce qu'il y prend plaisir,  
et il y prend plaisir parce qu'elle est belle. »*

— Henri Poincaré, 1854 – 1912

## Sommaire

---

Introduction au chapitre . . . . .	77
4.1 Modélisation du processus de formation d'une image optique . . . . .	79
4.1.1 De la scène à l'image optique - Parcours de l'énergie lumineuse	79
4.1.2 Éléments stationnaires du système optique : introduction de la fonction d'étalement du point (PSF) . . . . .	81
4.1.3 Non-stationnarité de la fonction d'étalement du point (PSF) . . . . .	83
4.2 Modèle simplifié du processus optique de formation d'une image . . . . .	86
4.3 Modélisation du processus d'acquisition de l'image numérique . . . . .	88
4.3.1 Modélisation stochastique détaillée . . . . .	89
4.3.2 Modèle simplifié du bruit . . . . .	92
Conclusions du chapitre . . . . .	94
4.4 Annexes du chapitre 4 : De l'image "RAW" à l'image "TIFF" . . . . .	95
4.4.1 Dématricage d'une image "RAW" . . . . .	96
4.4.2 Balance des blancs . . . . .	98
4.4.3 Correction gamma . . . . .	99

---

### Introduction au chapitre :

Afin de permettre une application pratique des tests étudiés dans le chapitre 3, il est nécessaire d'estimer les paramètres de distribution de tous les pixels,  $\theta = (\theta_1, \dots, \theta_N)^T$  et  $\sigma = (\sigma_1, \dots, \sigma_N)^T$ . Le nombre d'observations étant inférieur au nombre de paramètres, l'estimation de ces derniers n'est possible qu'en introduisant un a priori physique, au travers d'une modélisation des images, afin de diminuer le nombre de paramètres à estimer. Puisque les informations cachées qu'il est souhaitable de détecter sont insérées dans les LSB de la valeur des pixels, la modélisation doit être la plus précise possible. Ce modèle des images doit en outre pouvoir être exploité pour permettre la détection d'informations cachées tout en respectant une contrainte sur la probabilité de fausse alarme ; cela suppose que les paramètres décrivant une image doivent être maîtrisés. Pour ces raisons, il est nécessaire d'analyser l'ensemble du processus physique permettant l'acquisition d'une image numérique, à partir de la scène imagée, afin de proposer une modélisation tenant compte du contenu structuré des images pour permettre l'estimation des paramètres  $\theta$  et  $\sigma$ . Les propriétés physiques de la scène imagée sont décrites et exploitées dans le chapitre 5 suivant.

Dans le présent chapitre 4, les phénomènes physiques permettant l'acquisition d'une image sont présentés. Cette chaîne d'acquisition est considérée sous deux angles distincts :

1. L'image optique formée sur le plan du photo-détecteur est étudiée, dans la section 4.1. Les dégradations introduites par le système optique sont modélisées afin d'établir une relation analytique entre la scène imagée et l'intensité de l'image optique.

Le modèle exact du système optique étant difficilement exploitable, quelques propositions sont présentées dans la section 4.2 afin de simplifier la modélisation des dégradations introduites par le système optique.

2. Dans la section 4.3, l'acquisition numérique de l'image optique est brièvement décrite et les divers phénomènes stochastiques de dégradation sont modélisés. Là encore, une attention particulière est portée à la simplification du modèle du bruit d'acquisition présent dans l'image numérique.

Lors de l'acquisition d'une image, la lumière provenant de la scène observée parcourt l'atmosphère, puis pénètre dans le système optique où un jeu de lentilles permet une focalisation des rayons qui donne ainsi naissance à l'image optique sur le photo-détecteur. L'image optique subit alors un échantillonnage par intégration sur la surface des éléments photo-sensibles (pixels). Au cours de cette première étape de l'acquisition, l'image subit des dégradations déterministes d'origine physique dont l'analyse fait l'objet de la section 4.1. L'intensité lumineuse est ensuite convertie au sein des pixels en charges électriques qui sont successivement amplifiées et transférées avant que leurs potentiels ne soient mesurés et quantifiés. C'est schématiquement au cours de ces opérations que le signal électrique est contaminé par des bruits d'origine diverse; ces facteurs stochastiques de dégradation d'une image sont décrits et modélisés dans la section 4.3.1. Enfin, le travail ne s'arrête pas après l'obtention de cette image numérique "brute". Cette dernière subit finalement des traitements numériques post-acquisition ayant pour but d'améliorer le rendu visuel de l'image. Cette troisième et dernière étape de l'acquisition d'une image est décrite dans l'annexe 4.4.

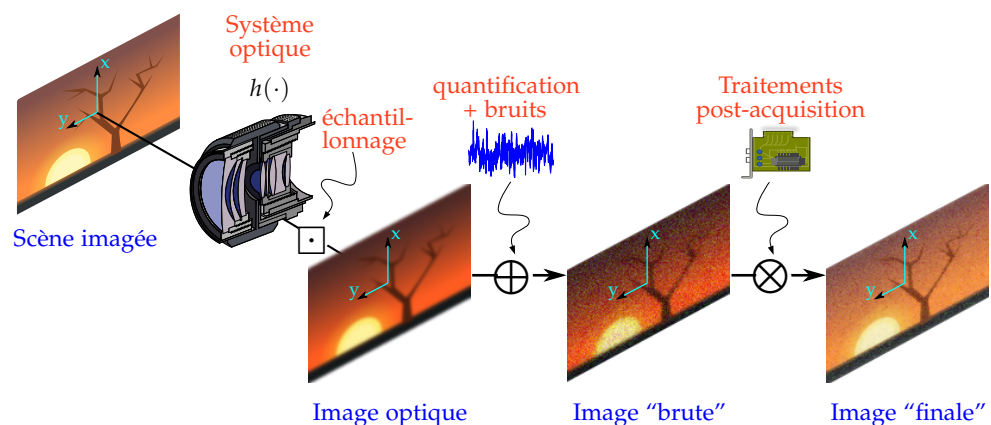


FIGURE 4.1: De la scène à l'image numérique - illustration de la chaîne d'acquisition des images naturelles.

Cette étape d'analyse de la formation des images permet de proposer un modèle générique des images utilisable pour l'estimation des paramètres  $\theta_n$  et  $\sigma_n$ , voir chapitre 5. L'exploitation pratique de ce modèle des images numériques pour la détection d'informations cachées est abordée dans la troisième partie de ce manuscrit, voir chapitres 6 et 7.

## 4.1 Modélisation du processus de formation d'une image optique

Un appareil photographique ou notre œil peuvent être vus comme des instruments de radiométrie dans le sens où ces derniers sont constitués d'un nombre important d'éléments photo-sensibles mesurant l'énergie lumineuse incidente. Pour modéliser une image numérique naturelle, il est donc nécessaire d'analyser le processus de formation de cette image optique.

Dans l'ordre de lecture de la figure 4.1, il est proposé de commencer par décrire, dans la section 4.1.1, la relation liant l'énergie lumineuse de la scène et celle de l'image optique du point de vue de l'optique géométrique. L'influence du système optique sur la répartition de l'énergie lumineuse incidente sur le détecteur est ensuite analysée, dans des conditions d'acquisition simples, au cours de la section 4.1.2. Les phénomènes physiques complexes non pris en compte dans la modélisation du système d'acquisition sont succinctement décrits dans la section 4.1.3.

Dans la section 4.1.1,  $(x, y)$  désigne un point de la scène (dans le plan focal objet  $\Pi$ ) et  $(x', y')$  l'image de ce point sur le photo-détecteur (dans le plan focal image  $\Pi'$ ) ; ces grandeurs étant reliées par le grandissement  $\gamma$  du système optique  $\gamma \stackrel{\text{def.}}{=} x'/x = y'/y$ .

### 4.1.1 De la scène à l'image optique - Parcours de l'énergie lumineuse

Le but de cette sous-section est de fournir une expression analytique de l'énergie reçue en un point  $(x', y')$  du photo-détecteur en fonction de l'énergie du point  $(x, y)$  de la scène imagée. Aussi, il n'est pas nécessaire de tenir compte du fait que la scène est généralement une source secondaire ; elle n'émet pas de lumière par elle-même mais réfléchit l'énergie lumineuse incidente. En outre, nous avons précisé que le système optique est, dans un premier temps, supposé être stationnaire (invariant par rotation et par translation) et n'induire aucune aberration. De même la scène imagée est d'abord supposée être dans le plan de mise au point (ou plan focal objet)  $\Pi$ . Enfin, le système optique peut toujours être assimilé à deux lentilles convergentes (définissant les pupilles d'entrée et de sortie) séparées par un diaphragme circulaire et un filtre spectral (également appelé filtre colorimétrique). Pour des raisons de simplicité évidentes, c'est un système optique de ce type qui sera considéré dans ce chapitre.

Les quelques définitions nécessaires afin d'établir une relation entre l'énergie émise par la scène et celle reçue au niveau du photo-détecteur sont données ci-dessous.

**Définition 4.1.** L'énergie rayonnée  $Q$  par une source étendue  $S$  entre les instants  $t_0$  et  $t_1$  dans le spectre de longueur d'onde compris entre  $\lambda_0$  et  $\lambda_1$  vaut, d'un point de vue corpusculaire :

$$\mathcal{Q}(S, t_0, t_1, \lambda_0, \lambda_1) = \int_{\lambda_0}^{\lambda_1} N_p(S, t_0, t_1, \lambda) \frac{h.c}{\lambda} d\lambda. \quad (4.1)$$

où  $h \approx 6.626 \cdot 10^{-34}$  est la constante de Planck,  $c \approx 2.998 \cdot 10^8$  la célérité de la lumière dans le vide et  $N_p(S, t_0, t_1, \lambda)$  le nombre de photons de longueur d'onde  $\lambda$  émis par la source  $S$  entre les instants  $t_0$  et  $t_1$ .

*Les éléments photo-sensibles de l'œil sont les cônes et les bâtonnets ; ce sont des macromolécules de bromure d'argent dans une pellicule photographique.*

*Ces conditions d'acquisition consistent à supposer que le système optique est stigmatique, i.e. tous les rayons issus d'un même point de la scène convergent, selon les lois de l'optique géométrique, vers un même point de l'image.*

*L'énergie rayonnée s'exprime, comme toute énergie, en  $W = J.s^{-1}$ .*

*La flux -ou puissance- spectral est donné en  $W.nm^{-1}$  dans les unités S.I.*



**Définition 4.2.** Le flux spectral d'une source lumineuse étendue  $S$ , noté  $\Phi(S, \lambda; t)$ , correspond à la puissance lumineuse -rayonnée ou reçue- par unité élémentaire de longueur d'onde à l'instant  $t$  :

$$\Phi(S, \lambda, t) = \frac{d^2 Q(S)}{d\lambda dt} = \frac{1}{\delta_\lambda \delta_t} \int_\lambda^{\lambda+\delta_\lambda} N_p(S, t_1, t + \delta_t, \lambda) \frac{h.c}{\lambda} d\lambda \quad (4.2)$$

Pour caractériser la puissance émise par un point de la scène imagée en direction de la pupille d'entrée du système optique, il est nécessaire de définir la puissance ponctuellement émise, par une source étendue, de façon directionnelle.

**Définition 4.3 (Luminance spectrale).** La luminance spectrale d'un point  $(x, y)$  d'une source étendue correspond au flux spectral émis par unité élémentaire d'angle solide (dans la direction donnée par les angles  $\phi$  et  $\omega$ ) et par unité élémentaire de surface autour du point  $(x, y)$ . La luminance est physiquement définie par :

La luminance s'exprime en  $Wm^{-2}.sr^{-1}.nm^{-1}$  dans les unités S.I.

$$S(x, y, \lambda, \omega, \phi, t) = \frac{d^2 \Phi(S, \lambda)}{d\Omega_{\omega, \phi} d\Sigma_{x, y}}$$

où  $d\Sigma_{x, y}$  représente la surface élémentaire autour du point  $(x, y)$  et  $d\Omega_{\omega, \phi}$  représente l'angle solide élémentaire dans la direction donnée par les angles  $\phi$  et  $\omega$ .

Enfin, la puissance rayonnante sur un point du photo-détecteur est mesurée indifféremment de l'angle d'incidence.

**Définition 4.4 (Éclairement spectral).** L'éclairement spectral d'un point  $(x', y')$  correspond au flux spectral reçu par unité élémentaire de surface, notée  $d\Sigma_{x', y'}$  autour du point  $(x', y')$  :

L'unité de l'éclairement spectral est le  $Wm^{-2}.nm^{-1}$  dans les unités S.I.

$$E(x', y', \lambda) = \frac{d\Phi(\Sigma_{x', y'}, \lambda; t)}{d\Sigma_{x', y'}}$$

Pour plus de détails sur la radiométrie et son utilisation en imagerie, le lecteur intéressé pourra en outre consulter [120, 121].

La scène est définie, d'un point de vue radiométrique, par sa *luminance spectrale*  $S(x, y, \lambda; \omega, \phi, t)$  définissant pour tout  $(x, y) \in \Pi = [0; x_{\max}] \times [0; y_{\max}]$  le flux spectral émis par le point  $(x, y)$  en direction de la pupille d'entrée. En toute rigueur, la luminance d'un point dépend de la direction sous laquelle ce dernier est observé (notée  $(\omega, \phi)$  en coordonnées sphériques) et de l'instant d'observation  $t$ . Dans le cas de la photographie, la durée d'acquisition (ou temps de pose) est suffisamment courte pour qu'il soit considéré que la luminance est constante dans le temps. En outre, l'angle solide sous-tendu par la pupille d'entrée du système optique étant négligeable et les objets observés étant en général Lambertiens, l'influence des angles d'observation sera réduite à un facteur  $\cos(\omega)$  où  $\omega$  est par convention l'angle entre la normale locale à la surface de la scène en  $(x, y)$  et la direction du système optique.

Ces considérations permettent de décrire la scène par sa luminance spectrale  $S : (x, y, \lambda; \omega, \tau) \mapsto \cos(\omega)S_p(x, y, \lambda)$  où sous les hypothèses précédentes, la fonction  $S_p$  représente le flux spectral émis par le point  $(x, y)$ .

L'énergie émise par un point de la scène traverse l'atmosphère pour parvenir à la pupille d'entrée du système optique qui, dans les conditions considérées, permet une focalisation des rayons dans le plan focal objet. L'atmosphère traversée modifie l'énergie incidente sur la pupille d'entrée de deux manières distinctes. D'une part, les éléments constituant l'atmosphère absorbent et/ou diffusent la lumière agissant ainsi comme un filtre spectral dont la transmittance est notée  $q_{atm} : \lambda \mapsto q_{atm}(\lambda)$  avec, pour tout  $\lambda \in \mathbb{R}_+$ ,  $0 < q_{atm}(\lambda) < 1$ . D'autre part l'atmosphère traversée est un milieu turbulent : ses propriétés physiques (température, pression, densité, etc. . .) sont en évolution permanente ; cette particularité est décrite plus en détails dans la section 4.1.3.

La transmittance d'un élément est, par définition, le rapport en l'énergie en sortie de cet élément et l'énergie incidente.

Les turbulences atmosphériques étant détaillées dans la section 4.1.3, elles sont également négligées présentement.

Compte-tenu du fait que la distance  $D$  séparant le plan focal objet de la pupille d'entrée du système optique est largement plus grande que le diamètre du diaphragme, il est notable que :

- l'angle solide  $\Omega_d$  sous-tendu par le diaphragme depuis un point  $(x, y)$  de la scène peut être approximé par :  $\frac{\cos(\omega)\Sigma_{\text{dia}}}{D^2} = \frac{\cos(\omega)\pi R_{\text{dia}}^2}{D^2}$  où  $\Sigma_{\text{dia}}$  et  $R_{\text{dia}}$  sont respectivement la surface et le rayon du diaphragme.
- la luminance spectrale émise par un point  $(x, y)$  de la source en direction d'un point du diaphragme peut être considérée constante quel que soit le point du diaphragme et le point  $(x, y)$  de la scène considérée.

En utilisant ces approximations, un calcul direct [120] permet de montrer que l'éclairement du point  $(x', y')$  est donné par :

$$E(x', y', \lambda) \approx \gamma^{-2} S_p(x, y, \lambda) q_{\text{atm}}(\lambda) \cos(\omega) \cos^3(\omega') \left( \frac{\pi R_{\text{dia}}^2}{D^2} \right)^2. \quad (4.3)$$

où le grandissement  $\gamma$  peut être défini par  $\gamma = \frac{D}{f'}$  avec  $f'$  la distance focale du système optique.

**Remarque 4.1.** Dans le cas où la scène imagée est orthogonale à l'axe optique, on a  $\omega = \omega'$  dont il découle  $E(x', y', \lambda) \propto \cos^4(\omega)$ . Ce phénomène de lobe est illustré sur la figure 4.3b.

Enfin, l'éclairement spectral est modifié par les filtres colorimétriques présents dans le système optique dont la transmittance est notée  $q_c$ . Il découle alors de (4.3) que l'on peut approximer l'éclairement d'un point  $(x', y')$  du photo-détecteur par :

$$E(x', y', \lambda) \approx \frac{1}{\gamma^2} S_p(x, y, \lambda) q_{\text{atm}}(\lambda) q_c(\lambda) \cos(\omega) \cos^3(\omega') \left( \frac{\pi R_{\text{dia}}^2}{D^2} \right)^2. \quad (4.4)$$

Dans la suite de ce manuscrit on notera, dans un souci de clarté, la relation (4.4) de la façon suivante :

$$E(x', y', \lambda) \approx S_0(x, y, \lambda) q_{\text{opt}}(\lambda) \quad (4.5)$$

avec  $S_0(x, y, \lambda) = S_p(x, y, \lambda) \frac{1}{\gamma^2} \cos(\omega) \cos^3(\omega') \left( \frac{\pi R_{\text{dia}}^2}{D^2} \right)^2$

et  $q_{\text{opt}}(\lambda) = q_{\text{atm}}(\lambda) \cdot q_c(\lambda)$ .

Le filtre spectral  $q_{\text{opt}}$  représente la transmittance globale du canal optique et  $S_0$  représente la scène pondérée par un facteur d'atténuation périphérique (phénomène appelé vignettage en photographie).

**Remarque 4.2.** La principale fonction des filtres colorimétriques présents dans un appareil photographique est précisée dans l'annexe 4.4.

**Remarque 4.3.** La différentiation entre coordonnées d'un point de la scène imagée  $(x, y)$  et les coordonnées du même point sur le plan focal objet  $(x', y')$  n'est plus nécessaire dans la suite de ce chapitre, aussi les coordonnées seront notées  $(x, y)$  afin d'alléger les expressions, le contexte permettant de comprendre s'il s'agit de coordonnées dans la scène ou dans l'image.

#### 4.1.2 Éléments stationnaires du système optique : introduction de la fonction d'étalement du point (PSF)

Le système optique considéré dans la section 4.1.1 est supposé parfaitement stigmatique d'un point de vue géométrique. Dans les faits, cette hypothèse de travail ne peut être rigoureusement vérifiée. En outre, les calculs de la section 4.1.1 reposent sur les

*Dans la pratique, plusieurs éléments filtrent la lumière dans un appareil photographique ; ces derniers sont ici assimilés à un unique filtre par souci de clarté.*

lois élémentaires de l'optique géométrique. Or il se trouve que l'influence ondulatoire de la lumière (ignorée jusqu'à présent) est fondamentale dans le processus de formation des images.

Il est donc proposé dans cette section et dans la suivante 4.1.3 d'étudier les principaux phénomènes de dégradations qui surviennent de façon déterministe au cours de la formation de "l'image optique"  $E(x, y, \lambda)$ . Ces dégradations sont modélisées par la fonction d'étalement du point, ou "PSF" de l'anglais *Point Spread Function*, qui décrit le canal optique comme la réponse impulsionnelle décrit un filtre.

Le diaphragme de l'appareil photographique modifie la propagation des ondes électromagnétiques. En effet, en application du principe de Huygens-Fresnel, le diaphragme agit comme une source secondaire susceptible de créer des interférences sur le plan image ; c'est le phénomène de diffraction de l'onde.

Dans le cadre de la photographie, la diffraction sous éclairage incohérent dans les conditions dites de Fraunhofer (lorsque la distance focale  $f' \gg \lambda \|(x', y')\|_2^2$ ) sera considérée.

Il est possible de montrer [121, 122] que la diffraction par un diaphragme agit comme un filtre passe-bas :

$$E(x, y; \lambda) = [S_0(x, y, \lambda) ** h_d(x, y; \lambda)] q_{\text{opt}}(\lambda) \quad (4.6)$$

où  $h_d$  est la PSF, ou noyau de convolution, d'un système optique limité par la diffraction et  $**$  représente le produit de convolution bidimensionnel défini par :

$$f(x, y) ** g(x, y) = \int_{\mathbb{R}^2} f(u, v) g(x - u, y - v) du dv = \int_{\mathbb{R}^2} f(x - u, y - v) g(u, v) du dv .$$

Afin de donner une expression analytique du noyau de convolution  $h_d$ , il est nécessaire de définir la fonction pupillaire de la surface diffractante. Cette fonction, notée  $p : (x, y) \mapsto p(x, y)$ , est définie pour tout point  $(x, y) \in \mathbb{R}^2$  comme la transmittance de la surface diffractante en ce point. Dans le cas de la photographie, la seule surface diffractante considérée est un diaphragme circulaire de rayon  $R_{\text{dia}}$  et de transmittance unitaire constante ; la fonction pupillaire s'exprime donc :

$$p(x, y) = \begin{cases} 1 & \text{si } x^2 + y^2 \leq R_{\text{dia}}^2 \\ 0 & \text{sinon} \end{cases} \Leftrightarrow p(\rho) = \begin{cases} 1 & \text{si } \rho^2 \leq R_{\text{dia}}^2 \\ 0 & \text{sinon} . \end{cases} \quad (4.7)$$

avec,  $\rho = (x^2 + y^2)^{1/2}$ .

La PSF  $h_d$  d'un système optique limité par la diffraction est rigoureusement donnée par :

$$h_d(x, y, \lambda) = \frac{2\pi R_{\text{dia}}^2}{\lambda^2 f'^2} \frac{|P(\frac{x}{\lambda f'}, \frac{y}{\lambda f'})|^2}{|P(0, 0)|^2}, \quad (4.8)$$

où la fonction  $P : \mathbb{R}^2 \rightarrow \mathbb{R}$  représente la transformée de Fourier de la fonction pupillaire et, par abus de notation,  $P(0, 0) = P(x, y)|_{x=0, y=0}$ .

En utilisant l'expression en coordonnées polaires de la fonction pupillaire un (fastidieux) calcul permet de montrer que [121, 122] pour tout  $\rho \in \mathbb{R}_+$  on a :

$$P(x, y) = P(\rho) = 2\pi R_{\text{dia}}^2 \frac{J_1(2\pi R_{\text{dia}} \rho)}{2\pi R_{\text{dia}} \rho}$$

où  $\rho = \rho(x, y) = \|(x, y)\|_2$  et  $J_n$  est la fonction de Bessel de 1ère espèce et d'ordre  $n$  (ici  $n = 1$ ).

**Remarque 4.4.** La présence du terme  $\frac{2\pi R_{\text{dia}}^2}{\lambda^2 f'^2}$  dans l'équation 4.8 permet de vérifier que :

$$\iint_{\mathbb{R}^2} |h_d(x, y, \lambda)|_2 = 1,$$

*Les conditions de Fraunhofer sont vérifiées dans la pratique en raison de la présence de lentilles devant et derrière la surface diffractante, le diaphragme.*

ce qui, physiquement, peut s'interpréter comme le fait que la surface diffractante ne provoque aucun gain ou perte d'énergie lumineuse, dès lors que sa transmittance est unitaire. Par souci de clarté, les phénomènes physiques distincts (ici diffraction et absorption de la lumière) sont considérés séparément. L'absorption du système optique est donc incluse dans la transmittance  $q_{\text{opt}}(\lambda)$ , voir relation (4.5).

Enfin, l'éclairement spectral incident sur le photo-détecteur est échantillonné par intégration spatiale sur la surface des cellules photo-sensibles, les pixels. L'énergie  $Q[m, n]$  lumineuse incidente sur la surface du pixel situé sur la  $m$ -ième ligne et sur la  $n$ -ième colonne, notée  $\Sigma_{(m,n)}$ , est donc donnée par :

$$Q[m, n] = \tau \int_{R_+} \left( \iint_{\Sigma_{(m,n)}} E(x, y, \lambda) dx dy \right) d\lambda \quad (4.9)$$

$Q[m, n]$  s'exprime en J puisque  $E$  est donné en  $W.m^{-2}.nm^{-1}$ .

où, l'émission lumineuse étant considérée constante dans le temps, l'intégration sur la durée d'acquisition  $\tau$  est réduite dans (4.9) à un facteur multiplicatif.

En supposant que la surface photo-sensible de chacun des pixels est identique, l'intégration spatiale dans la relation (4.9) peut être représentée par une opération de convolution suivi d'un échantillonnage par multiplication par un peigne de Dirac (bi-dimensionnel).

Dans le cas usuel où la surface photo-sensible d'un pixel correspond à un carré de largeur  $\Delta_x$ , la PSF  $h_{\text{int}}$  liée à l'intégration spatiale sur la surface photo-sensible est donnée par :

$$h_{\text{int}}(x, y) = \begin{cases} \frac{1}{\Delta_x^2} & \text{si } (x, y) \in \left[-\frac{\Delta_x}{2}; \frac{\Delta_x}{2}\right]^2 \\ 0 & \text{sinon.} \end{cases}$$

L'énergie  $Q[m, n]$  lumineuse incidente sur la surface du pixel  $(m, n)$  et sur la durée d'acquisition  $\tau$  (4.9) peut alors s'écrire :

$$Q[m, n] = \tau \left( \int_{R_+} E(x, y, \lambda) ** h_{\text{int}}(x, y) d\lambda \right) \Big|_{x=x_n, y=y_m} \quad (4.10)$$

avec  $(x_n, y_m)$  les coordonnées du centre du pixel  $(m, n)$ .

### 4.1.3 Non-stationnarité de la fonction d'étalement du point (PSF)

Le canal optique d'acquisition d'une image naturelle considéré dans la section 4.1.2 est supposé limité par la diffraction ; *i.e.* le seul élément de dégradation est la diffraction des ondes lumineuses par le diaphragme. Un tel canal optique permet de modéliser avec une grande précision, et de façon assez simple, la relation entre la luminance d'une scène et l'énergie reçue par chaque pixel, voir (4.9) - (4.10). En pratique, d'autres éléments plus complexes à modéliser, notamment en raison de leur non-stationnarité spatiale, modifient l'éclairement incident sur la surface du photo-détecteur. Cette sous-section 4.1.3 décrit brièvement et modélise ces principaux éléments de dégradation.

En préambule, il est à souligner que ces opérateurs de dégradation, introduits par le système optique, impactent l'image de manière analogue à un filtre. La seule, mais ô combien importante, différence réside dans le fait que le système optique n'est pas invariant par translation. L'invariance du système n'est alors pas vérifiée, l'expression formelle des dégradations s'en trouve rendue immédiatement plus complexe.

#### Turbulences atmosphériques [120]

L'atmosphère traversée par les rayons lumineux, entre la source et la pupille d'entrée, est un milieu dit *turbulent* dans le sens où ses caractéristiques physiques sont en constantes variations. A l'échelle de temps d'acquisition d'une photographie, les turbu-

Le temps d'acquisition d'une image est typiquement  $\tau \leq 1/60s$ .

lences atmosphériques ont une influence limitée correctement modélisée en première approximation par les lois de Kolmogorov (ces dernières faisant appel à l'optique statistique, leur étude approfondie sort du cadre de ce manuscrit). Ces lois décrivent la position d'incidence d'un rayon lumineux comme une variable aléatoire suivant une loi de distribution Gaussienne de moyenne nulle. L'écart-type  $\sigma_{\text{atm}}$  de cette variable aléatoire, représentant physiquement l'impact des turbulences, est fonction de la distance  $D$  et de la "constante de structure de l'indice de réfraction", voir [120] pour plus de détails.

Selon les lois de Kolmogorov les turbulences atmosphériques sont modélisées par une PSF donnée par :

$$h_{\text{atm}}(\rho) = \frac{1}{\sigma_{\text{atm}}^2 2\pi} . e^{-\frac{\rho^2}{2\sigma_{\text{atm}}^2}} . \quad (4.11)$$

### Défaut de mise au point [120, 121]

Le défaut de mise au point intervient lorsque la scène imagée (ou une partie de celle-ci) est située en dehors du plan focal objet  $\Pi$ . En effet, pour un système rigoureusement stigmatique, les rayons lumineux convergent devant (resp. derrière) le plan focal image  $\Pi'$  lorsqu'ils sont issus d'un point situé respectivement derrière (resp. devant) le plan  $\Pi$ . Dans les deux cas, une simple construction reposant sur les lois de l'optique géométrique permet de montrer que : 1) l'image d'un point en dehors du plan  $\Pi$  est un disque au sein duquel l'éclairement spectral incident est constant (et nul en dehors de ce disque) et 2) le rayon de ce disque est donné par :

$$R_{\text{focus}} = R_{\text{dia}} \frac{f'}{D - f'} \frac{|D - D_{x,y}|}{D_{x,y}}$$

avec  $R_{\text{dia}}$  le rayon du diaphragme et  $D_{x,y}$  la distance entre le centre de la pupille et le point  $(x, y)$  considéré.

Le défaut de mise au point peut donc être modélisé par une PSF dont l'expression est :

$$h_{\text{focus}}(\rho) = \begin{cases} \frac{1}{\pi R_{\text{focus}}^2} & \text{si } \rho \leq R_{\text{focus}} \\ 0 & \text{sinon} . \end{cases} \quad (4.12)$$

### Mouvement de la scène et du système imageur

Lorsque la scène est en mouvement par rapport à l'axe optique du système imageur durant la durée d'acquisition  $\tau$  d'une photographie, la position de l'image d'un point de la scène est alors soumise au même mouvement sur la surface du photo-détecteur. Aussi, l'énergie lumineuse issue du point  $(x, y)$  est distribuée sur une courbe représentant le mouvement de ce dernier. Pour modéliser simplement ce phénomène on peut considérer [121], compte-tenu du fait que la durée d'acquisition est très courte, que le déplacement de la scène par rapport au système imageur est rectiligne uniforme, *i. e.* le vecteur de vitesse du déplacement  $\vec{v}_0$  est constant.

La PSF de ce phénomène, appelé le "bougé", peut donc être modélisé par une "lame" de Dirac [121] définie par  $h_{\text{mvt}}(\rho, \omega) = \lim_{\delta_\omega \rightarrow 0} L_{\text{mvt}}(\rho, \omega)$  avec :

$$L_{\text{mvt}}(\rho, \omega) = \begin{cases} \frac{2}{\tau^2 \|\vec{v}_0\|_2^2 \delta_\omega} & \text{si } \rho \leq \tau \|\vec{v}_0\| \text{ et } \omega \in [\omega_0; \omega_0 + \delta_\omega] \\ 0 & \text{sinon,} \end{cases}$$

$\|\vec{v}_0\|_2$  et  $\omega_0$  sont respectivement la norme Euclidienne usuelle et l'angle (en coordonnées polaires) du vecteur vitesse  $\vec{v}_0$  caractérisant le mouvement relatif de la scène par rapport au système imageur.

### Aberrations du système optique

Le système optique limité par la diffraction, présenté dans la section 4.1.2, suppose que les conditions paraxiales sont vérifiées, *i. e.* l'angle d'incidence  $\omega$  est supposé très faible et la position des points de la scène sont considérés proches de l'axe optique. Les pupilles d'entrée et de sortie étant sphériques, les approximations paraxiales de Gauss utilisées consistent à approximer les fonctions trigonométriques par leurs développements de Taylor-Maclaurin à l'ordre 1.

On désigne de façon générique par le terme d'*aberrations*, toutes les déformations du système optique par rapport aux approximations paraxiales de Gauss [122, 123]. Ces déformations peuvent être calculées par un développement à un ordre supérieur des fonctions trigonométriques. Les aberrations de Seidel (la coma, l'astigmatisme, l'aberration sphérique, la distorsion et la courbure de champ) sont par exemple calculées en utilisant des développements à l'ordre 3.

L'étude précise des aberrations sort légèrement du cadre de ce manuscrit, aussi une description volontairement laconique est présentée. Le lecteur intéressé pourra se référer aux ouvrages [122, 124].

Les déformations du système optique par rapport aux approximations paraxiales de Gauss introduisent un défaut de phase  $\delta_\phi$  du front d'onde incident sur la pupille d'entrée [122]. L'étude des principales aberrations peut se faire en introduisant la fonction complexe  $p_{\text{abe}}$ , appelée fonction pupillaire généralisée, définie par :

$$p_{\text{abe}}(\rho; \omega) = p(\rho; \omega) \exp(i\delta_\phi),$$

où  $p(\rho; \omega)$  est l'expression, en coordonnées polaires, de la fonction pupillaire (4.7) et  $\delta_\phi = \delta_\phi(\rho, \omega)$  correspond au défaut de phase du front d'onde.

La PSF  $h_{\text{abe}}$  d'un système optique présentant des aberrations est donnée en substituant la transformée de Fourier de la fonction pupillaire généralisée  $P_{\text{abe}}$  à  $P$  dans la relation (4.8) :

$$h_{\text{abe}}(x, y, \lambda) = \frac{2\pi R_{\text{dia}}^2}{\lambda^2 f^2} \frac{|P_{\text{abe}}(\frac{x}{\lambda f}, \frac{y}{\lambda f})|^2}{|P_{\text{abe}}(0, 0)|^2}. \quad (4.13)$$

### Modélisation de l'éclairement spectral incident

La prise en compte des éléments non-stationnaires décrits dans la présente section permet de modéliser l'éclairement spectral de façon plus précise que (4.6) par :

$$E(x, y; \lambda) = \left[ \int_{R^2} S_0(x-u, y-v; \lambda) h_{\text{opt}}^*(u, v, x, y; \lambda) du dv \right] q_{\text{opt}}(\lambda), \quad (4.14)$$

où  $h_{\text{opt}}^*$  représente la PSF globale du canal optique d'acquisition; l'expression de  $h_{\text{opt}}^*$  est donnée par l'application successive des PSF de chacun des éléments non-stationnaires (4.11) - (4.13).

La relation (4.14) permet finalement d'exprimer l'énergie incidente sur l'ensemble de la surface du pixel  $(m, n)$  (4.10) comme suit :

$$\begin{aligned} Q[m, n] &= \tau \left( \int_{R_+} E(x, y, \lambda) d\lambda ** h_{\text{int}}(x, y) \right) \Big|_{x=x_n; y=y_m} \\ &= \tau \left( \int_{R_+} \left[ \int_{R^2} S_0(x-u, y-v; \lambda) h_{\text{opt}}^*(u, v, x, y; \lambda) du dv \right] ** h_{\text{int}}(x, y) \right) q_{\text{opt}}(\lambda) d\lambda \Big|_{x=x_n; y=y_m} \end{aligned} \quad (4.15)$$

avec  $(x_n, y_m)$  les coordonnées du centre du pixel  $(m, n)$ .

*Les aberrations de Seidel sont également appelées aberrations du troisième ordre en raison de l'ordre du développement utilisé.*

## 4.2 Modèle simplifié du processus optique de formation d'une image

Les dégradations déterministes décrites dans la section 4.1 sont entièrement caractérisées par la PSF  $h_{\text{opt}}^*$ . Or comme cela a été détaillé, la PSF globale du canal optique dépend de nombreux éléments, voir [122, 121, 123] pour plus d'informations. Les relations (4.14) - (4.15) laissent entendre qu'une modélisation rigoureuse et simple de la PSF d'un système imageur sur l'ensemble de l'image n'est pas réaliste.

Pour permettre une modélisation exploitable de la PSF  $h_{\text{opt}}^*$ , il est impératif de simplifier le modèle exacte de la PSF tout en conservant les propriétés pertinentes dans le cadre des travaux présentés. Le point crucial dans la modélisation de la PSF  $h_{\text{opt}}^*$  est sa non-invariance, ou non-stationnarité, spatiale sur l'ensemble de l'image : les éléments de dégradations décrits dans la section 4.1.3 engendrent une réponse dépendante des coordonnées du point considéré.

**Propriété 4.1** (Propriétés de la PSF usuellement admises). *Malgré cette non-stationnarité, il peut être considéré que la PSF  $h_{\text{opt}}^*$  vérifie, sur l'ensemble du domaine de l'image, les propriétés suivantes :*

P-1 : *La PSF  $h_{\text{opt}}^*$  est isotrope ; en effet seule l'aberration de coma et le "bougé" (mouvement relatif de la scène par rapport au système optique) sont anisotropes et, dans la très grande majorité des cas, l'influence des ces phénomènes peut être négligée.*

P-2 : *La PSF  $h_{\text{opt}}^*$  ne dépend pas de la longueur d'onde  $\lambda$  du rayonnement incident. En effet, seule la PSF due à la diffraction par le diaphragme dépend de la longueur d'onde et dans les faits, cette dépendance est limitée par la présence de filtres couleurs (réduisant fortement la plage de longueur d'onde incidente sur le photo-détecteur).*

P-3 : *Enfin, la PSF  $h_{\text{opt}}^*$  est considéré localement invariante par translation ; En effet, le formalisme de la relation (4.14) permet de montrer qu'en tout point  $(x_0, y_0)$  de  $\mathcal{X}$  l'image optique est donnée par :*

$$\begin{aligned} & \int_{\mathbb{R}^2} S_0(x_0 - u, y_0 - v; \lambda) h_{\text{opt}}^*(u, v, x_0, y_0; \lambda) du dv \\ &= \int_{\mathbb{R}^2} S_0(x_0 - u, y_0 - v; \lambda) h_0^*(u, v; \lambda) du dv = [S_0(x, y; \lambda) ** h_0^*(y, x; \lambda)] \Big|_{x=x_0, y=y_0} \end{aligned}$$

*correspond à la convolution de la scène  $S_0(x, y; \lambda)$  avec la PSF localement définie par :*

$$h_0^*(u, v; \lambda) = h_{\text{opt}}^*(u, v, x, y; \lambda) \Big|_{x=x_0, y=y_0}$$

L'intérêt de la propriété P1 est mis en évidence dans la section 5.1. À l'inverse, l'utilité de la propriété P-3 a déjà été mis en avant dans la formulation de la propriété elle-même. Enfin, l'intérêt de la propriété P-2 est immédiat en définissant :

$$S(x, y) = \int_{\mathbb{R}_+} S_0(x, y; \lambda) q_{\text{opt}}(\lambda) d\lambda.$$

Il est alors immédiat d'écrire, dans la relation (4.15) :

$$\int_{\mathbb{R}_+} E(x, y, \lambda) d\lambda = \left[ \int_{\mathbb{R}^2} S(x-u, y-v) h_{\text{opt}}^*(u, v, x, y) du dv \right],$$

où, puisque la PSF  $h_{\text{opt}}^*$  ne dépend pas de  $\lambda$ , il est noté  $\forall \lambda \in \mathbb{R}_+, h_{\text{opt}}^*(u, v, x, y) = h_{\text{opt}}^*(u, v, x, y, \lambda)$ .

Ainsi, il découle finalement de l'utilisation des propriétés P-2 et P-3 dans (4.15) l'expression simplifiée suivante de l'énergie incidente sur le pixel  $(m, n)$  :

$$Q[m, n] = \tau \left[ \int_{\mathbb{R}^2} S(x-u, y-v) h_{\text{opt}}^*(u, v, x, y) du dv \right] ** h_{\text{int}}(x, y) \Big|_{x=x_n, y=y_m} \quad (4.16)$$

où, le formalisme de la relation (4.16) ne doit pas faire oublier que l'intégrale, entre [], peut être localement représentée par un produit de convolution.

Il est important de souligner que l'application de la méthodologie détaillée dans les chapitres 5 et 6 repose sur les propriétés 4.1 ; ces propriétés, qui sont motivées physiquement, constituent des postulats incontournables à l'application de la méthodologie proposée. Toute modélisation de la PSF respectant les propriétés 4.1 peut être potentiellement utilisée dans le cadre méthodologique proposé. À l'inverse, il a été choisi d'étudier un modèle particulier de la PSF, défini par les assertions 4.1, dont le principal intérêt est de simplifier les calculs et d'en limiter le nombre. Il ne s'agit cependant que d'un modèle particulier parmi d'autres potentiellement utilisables.

**Assertions 4.1.** *Le modèle du système optique retenu suppose que les assertions suivantes sont vérifiées :*

A-1 : *La PSF  $h_{\text{opt}}^*$  est non-stationnaire par changement d'échelle : en tout point  $(x, y)$  de l'image, il est possible d'approcher la valeur de la PSF par :*

$$h_{\text{opt}}^*(u, v, x, y) \approx \frac{1}{\zeta_0(x, y)^2} h_{\text{opt}}^* \left( \frac{u}{\zeta_0(x, y)}, \frac{v}{\zeta_0(x, y)}; x_0, y_0 \right) \stackrel{\text{def.}}{=} h_{\text{opt}}(u, v; \zeta_0(x, y))$$

où  $\zeta_0(x, y)$  représente le paramètre d'échelle de la PSF et est noté  $\zeta_0$  dans la suite pour alléger les expressions.

A-2 : *La PSF  $h_{\text{opt}}$  (ou de manière analogue  $h_{\text{opt}}^*$ ) peut être correctement modélisée en première approximation par une fonction Gaussienne :*

$$h_{\text{opt}}(u, v, \zeta_0) \approx \frac{1}{\zeta_0^2} \phi \left( \frac{\rho}{\zeta_0} \right) \quad (4.17)$$

$$\text{avec } \rho = \sqrt{u^2 + v^2} \text{ et } \phi(x) = \frac{1}{2\pi} \exp \left( -\frac{x^2}{2} \right).$$

Les assertions A-1 et A-2 sont très difficiles, voire impossibles, à justifier formellement. En effet, l'assertion A-1 est davantage une modélisation mathématique de la PSF ayant pour but de réduire la non-stationnarité du canal optique à un unique paramètre  $\zeta_0 = \zeta_0(x, y)$  dépendant de la position du point  $(x, y)$  considéré.

De manière analogue, le modèle Gaussien imposé par l'assertion A-2 est couramment utilisé dans la littérature pour modéliser la PSF de l'ensemble de canal optique [125, 126], à l'exception du "bougé", en raison de son aspect générique. Dans la pratique, la figure 4.2 montre que le modèle Gaussien reflète assez fidèlement la nature réelle de la PSF, mais une démonstration formelle n'a pas, à notre connaissance, été proposée.

Il découle des assertions A-1 et A-2 (et des propriétés 4.1) que l'éclairement spectral incident sur le photo-détecteur (4.14) est donné par :

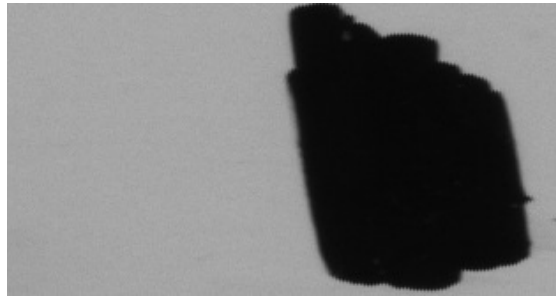
$$E(x, y; \lambda) = S(x, y) ** h_{\text{opt}}(x, y; \zeta_0). \quad (4.18)$$

Il est alors immédiat de montrer que, de manière analogue, les assertions A-1 et A-2 permettent d'exprimer l'énergie incidente (4.15) sur le pixel  $(m, n)$  par la relation :

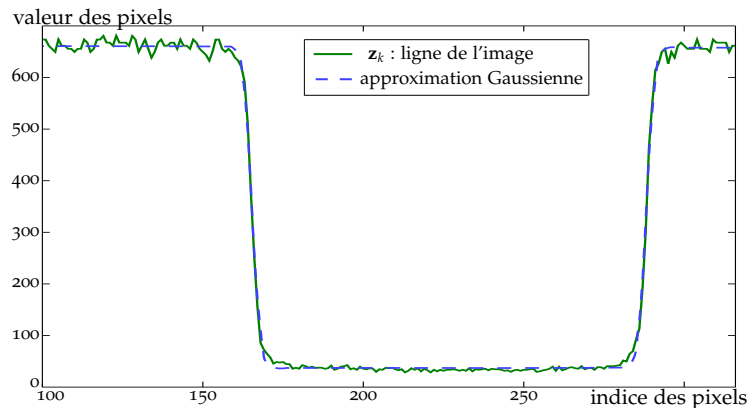
$$Q[m, n] = \tau S(x, y) ** h_{\text{opt}}(x, y; \zeta_0) ** h_{\text{int}}(x, y) \Big|_{x=x_n, y=y_m} \quad (4.19)$$

Ce sont ces deux expressions qui seront utilisées dans la suite. Il est cependant important de souligner que les assertions A-1 et A-2 sont un cas particulier qu'il a été choisi de développer dans les travaux présentés ; la méthodologie proposée dans les chapitres 5 et 6 peut être appliquée en utilisant uniquement les propriétés 4.1 au prix d'un plus grand nombre de calculs.





(a) Image numérique constituée de deux domaines  $\mathcal{X}_1$  (en blanc) et  $\mathcal{X}_2$  (en noir).



(b) Illustration de l'impact de la PSF sur l'image numérique et de la pertinence de l'approximation Gaussienne.

FIGURE 4.2: Illustration de l'effet de la PSF sur une image numérique naturelle simple ; l'approximation locale de la PSF par une fonction Gaussienne semble pertinente.

### 4.3 Modélisation du processus d'acquisition de l'image numérique

Les dispositifs électroniques d'acquisition d'une image (type Charge-coupled device *CCD*, Complementary metal oxide semi-conductor *CMOS*, ou autre) fonctionnent sur le même principe. Un substrat semi-conducteur (type jonction P-N) permet d'exploiter l'effet photovoltaïque et de convertir les photons incidents en électrons. Les charges créées sont ensuite stockées, au sein de chaque photosite, dans un condensateur déchargé avant acquisition de l'image (création d'un puits de potentiel). Enfin, les charges stockées par chaque photosite sont transférées dans des registres verticaux puis dans un registre horizontal pour être finalement mesurées séquentiellement, ligne après ligne, par une unité dédiée [127, 128].

Les tensions électriques  $V[m, n]$  ainsi mesurées sont ensuite quantifiées ; le résultat de cette opération correspond à la valeur  $z[m, n]$  de chacun des pixels constituant l'image numérique. Au cours de cette étape de numérisation de l'image optique formée sur le photo-détecteur, plusieurs processus de natures aléatoires viennent corrompre la valeur  $z[m, n]$  des pixels. Dans la présente section il est notamment montré que la valeur  $z[m, n]$  d'un pixel est donnée par :

$$z[m, n] = \theta[m, n] + \zeta[m, n]$$

où  $\theta[m, n] = \mathbb{E}[z[m, n]]$  est une valeur déterministe et  $\zeta[m, n]$  est la réalisation de la variable aléatoire  $\Xi[m, n]$  représentant l'ensemble de phénomènes aléatoires corrompant l'image lors de sa numérisation. Le but de cette section est double :

1. D'une part, il s'agit d'établir une relation entre l'énergie incidente  $Q[m, n]$  sur le pixel  $(m, n)$ , décrite par les relations (4.15) et (4.19), et l'espérance mathématique  $\mathbb{E}[z[m, n]] = \theta[m, n]$  de la valeur du même pixel après numérisation.
2. D'autre part, il s'agit de caractériser la distribution de la variable aléatoire  $\Xi[m, n]$ .

### 4.3.1 Modélisation stochastique détaillée

Les relations radiométriques données dans la section 4.1 permettent d'exprimer l'éclairement spectral incident. Or cette grandeur représente une énergie qu'il convient de convertir en un nombre de photons pouvant interagir avec le photo-détecteur. L'énergie  $Q[m, n]$  incidente sur la surface d'un pixel (4.15) est converti en un nombre aléatoire de paires électrons-trous  $N_{ep}[m, n]$  dont la loi de distribution est donnée par [129, 130] :

$$N_{ep}[m, n] \sim \mathcal{P}(\mu_{ep}[m, n]) \quad \text{avec} \quad \mu_{ep}[m, n] = \frac{\lambda}{hc} Q[m, n] q_{ep}, \quad (4.20)$$

$\mathcal{P}(\mu)$  représente la loi de distribution Poissonienne de paramètre  $\lambda$ ,  $h \approx 6.626 \cdot 10^{-34}$  la constante de Planck,  $c \approx 2.998 \cdot 10^8$  la célérité de la lumière dans le vide et  $q_{ep}$  l'efficacité quantique du semi-conducteur exprimée en fonction de la longueur d'onde, *i. e.* le nombre moyen d'électrons collectés par photons incidents.

Dans ce manuscrit, il est admis que la totalité des charges collectées sont transférées sans aucune perte. Cette supposition est une approximation raisonnable étant donné que le taux de transfert des charges entre registres est de  $1 - \epsilon$  avec  $\epsilon \lesssim 10^{-6}$  [127, 129]. Ainsi, la diffusion et la rémanence de charges, engendrées par le "transfert" indésirable de charges d'un pixel à un autre, sont implicitement négligées dans le présent manuscrit ; le lecteur intéressé pourra consulter [127, 128] pour une description approfondie.

Dans ce manuscrit, seules les cinq principales sources de bruits sont considérées [129, 130].

#### Non-uniformité des pixels

En raison d'imperfections lors de la fabrication du photo-détecteur, l'efficacité quantique n'est pas constante sur l'ensemble de la surface photo-sensible. D'un point de vue macroscopique, ce phénomène peut se modéliser en considérant que l'efficacité quantique varie entre chacun des pixels au sein desquels elle peut être considérée comme constante [129]. Ce phénomène de non-uniformité de l'efficacité quantique des pixels (abrégé par NUEQ) peut être modélisé, en apportant à la relation (4.20) la "correction" suivante [129, 130, 131] :

$$N_{ep}[m, n] \sim \mathcal{P}(\mu_{ep}[m, n]) \quad \text{avec} \quad \mu_{ep}[m, n] = \frac{\lambda}{hc} Q[m, n] q_{ep} K[m, n], \quad (4.21)$$

où  $K[m, n]$  représente la sensibilité du photo-site à la position  $(m, n)$ . Pour un appareil photographique donné, le champ scalaire discret  $K : (m, n) \mapsto K[m, n]$  est déterministe (constant dans le temps) mais n'est pas connu *a priori*. Aussi il est proposé de modéliser ce terme comme la réalisation d'une variable aléatoire Gaussienne de moyenne unitaire et de variance  $\sigma_K^2 \ll 1$ , ce qui est noté par abus de notations :  $K[m, n] \sim \mathcal{N}(1, \sigma_K^2)$ .

#### Courant d'obscurité

Indépendamment du nombre de photons incidents sur le photo-détecteur, des électrons se déplacent en permanence sous l'effet de l'agitation thermique, générant ainsi des paires électrons-trous dans l'ensemble du semi-conducteur. Les électrons libres ainsi générés se déplacent, comme les électrons d'origine photonique, vers le puits de potentiel créés par le condensateur déchargé. Ce phénomène de courant obscur,

*La loi distribution de Poisson est ici liée au fait que la génération de paires électrons-trous est assimilable à un processus de comptage d'un (grand) nombre d'occurrence.*

*Afin de donner un ordre de grandeur, une valeur approchée de l'efficacité quantique d'un capteur CCD performant est  $q_{ep} \approx 0.7$ .*

*À titre d'exemple, la diffusion des charges est limitée par l'utilisation de CCD à "canaux enfouis" (buried channel) [127].*

indépendant de l'énergie lumineuse incidente, génère un nombre de charges dépendant linéairement du temps d'acquisition  $\tau$  et exponentiellement de la température ambiante (doublant tous les 6°C environ [127, 128]).

Le nombre d'électrons d'origine thermique  $N_{co}[m, n]$  collecté dans le pixel  $(m, n)$  pouvant être assimilé à un processus stochastique de décompte d'un (grand) nombre d'occurrences, ce dernier est modélisé comme la réalisation d'un variable aléatoire suivant une loi de Poisson :

$$N_{co}[m, n] \sim \mathcal{P}(\mu_{co})$$

où le paramètre de distribution  $\mu_{co}$  n'est pas connu en pratique.

**Remarque 4.5.** *Dans la pratique, les pixels ne sont pas tous identiquement sensibles à l'agitation thermique, en raison de la présence d'impuretés dans le substrat semi-conducteur. Le terme de pixel "chaud" est souvent utilisé pour qualifier les pixels les plus sujets à l'agitation thermique.*

*La terminologie de bruit de structure fixe (ou fixed pattern noise, FPN) est usuellement utilisée en référence à la non-uniformité de l'efficacité quantique des pixels et à la présence de pixels chauds, ces phénomènes se répétant "de façon fixe" sur chaque image.*

### Bruit quantique de comptage (Shot noise)

Ce bruit est directement lié à la nature corpusculaire de la lumière et des charges électriques. Les électrons d'origine thermique et photonique générés dans un pixel sont stockés indifféremment dans le condensateur de ce pixel. Comme précédemment évoqué, le nombre d'électrons stockés dans le condensateur est assimilable à un processus stochastique de décompte d'un grand nombre d'occurrences aléatoires. En vertu de la propriété de stabilité de la loi de Poisson par la somme, le nombre total d'électrons collectés dans le pixel à la position  $(m, n)$ , noté  $N_{e^-}[m, n]$ , est également modélisé comme la réalisation d'une variable aléatoire Poissonnienne :

$$\begin{aligned} N_{e^-}[m, n] &= N_{co}[m, n] + N_{ep}[m, n] \\ \Rightarrow N_{e^-}[m, n] &\sim \mathcal{P}(\mu_{ep}[m, n] + \mu_{co}). \end{aligned} \quad (4.22)$$

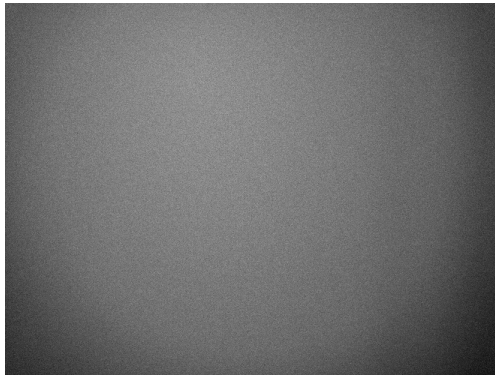
où selon (4.21)  $\mu_{ep}[m, n] = \frac{\lambda}{hc} Q[m, n] q_{ep} K[m, n]$ .

### Bruit de lecture et d'amplification

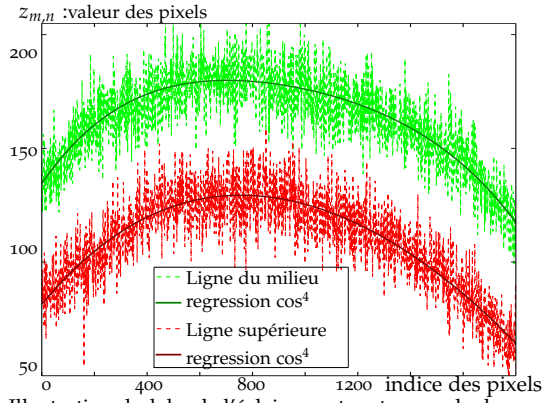
L'amplificateur transforme le nombre d'électrons collectés en une tension électrique mesurable. Les opérations d'amplification et de lecture étant (presque) réalisées simultanément, *i.e.* en sortie du registre horizontal, les bruits intervenant au cours de ces deux opérations sont considérés conjointement. Physiquement, ce bruit est dû aux imperfections des semi-conducteurs de l'unité de lecture et d'amplification ; ce bruit est donc indépendant de la position du pixel et de la charge collectée au sein de ce dernier. Le bruit de lecture, noté  $N_{lec}[m, n]$ , peut être modélisé avec une bonne précision [129, 130] comme la réalisation d'une variable aléatoire Gaussienne de moyenne nulle et de variance  $\sigma_{lec}^2$  (constante sur l'ensemble des pixels). L'amplification du signal, est caractérisée par le gain, noté  $A$ , représentant le facteur multiplicatif de la conversion charge-tension. La tension de sortie  $V[m, n]$  correspondant au pixel  $(m, n)$  est donc finalement donnée par :

$$V[m, n] = A(N_{ep}[m, n] + N_{co}[m, n] + N_{lec}[m, n]) \quad (4.23)$$

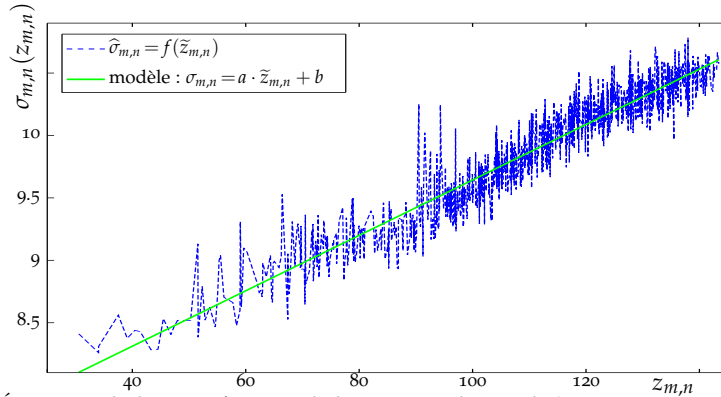
### Bruit de quantification



(a) Image "RAW" d'une surface de luminance constante.



(b) Illustration du lobe de l'éclairage au travers de deux lignes.



(c) Écart-type du bruit en fonction de la moyenne des pixels (régression en utilisant un modèle quadrique du lobe (4.3) - (4.5)).

 FIGURE 4.3: Illustration du lobe de l'éclairage incident pour mettre en évidence le bruit quantique de comptage des photons : l'écart-type du bruit de comptage dépend linéairement de l'énergie incidente. La moyenne et la variance ont été calculées sur des blocs de  $50 \times 50$  pixels modélisés par une quadrique :  $\sum_{i,j \in \{0,2\}} a_{i,j} x^i y^j$ .

Enfin, le signal analogique  $V[m, n]$  est numérisé par un convertisseur analogique-numérique (CAN). À ce stade de l'acquisition d'une image, la valeur des pixels est quantifiée avec un pas  $\Delta_{can}$  très faible (typiquement 12,14 voire 16 bits sont utilisées pour représenter ces valeurs); cela notamment pour permettre une application des traitements post-acquisition décrits dans l'annexe 4.4. En outre, une quantification linéaire uniforme est usuellement utilisée pour réaliser cette conversion analogique-numérique. La valeur  $z[m, n]$  du pixel  $(m, n)$  "pré-quantifiée" est donc rigoureusement donnée d'après (4.23) par :

$$z[m, n] = Q_{\Delta_{can}}(V[m, n]) = Q_{\Delta_{can}}(A(N_{ep}[m, n] + N_{co}[m, n] + N_{lec}[m, n])).$$

où  $Q_{\Delta_{can}}$  représente la fonction de quantification définie par la relation (2.26).

**Remarque 4.6.** Ici encore, le phénomène de "clipping" ou de "censoring" (écrêtage du signal) n'est pas considéré pour des raisons de clarté et de simplicité. Il est donc implicitement supposé que le pas de quantification  $\Delta_{can}$  et le facteur d'amplification  $A$  sont choisis de telle sorte que  $A \cdot V[m, n]$  n'excède jamais la quantité  $(2^b - 1/2)\Delta_{can}$ .

En considérant que le pas de quantification  $\Delta_{can}$  est "suffisamment petit", une valeur approchée de la tension quantifiée  $z[m, n]$  est donnée par :

$$z[m, n] = A(N_{ep}[m, n] + N_{co}[m, n] + N_{lec}[m, n]) + N_{can}[m, n] \quad (4.24)$$

où  $N_{can}[m, n]$  est la réalisation d'une variable aléatoire représentant le bruit dû à la conversion analogique-numérique modélisée [118, 129] par :

$$N_{can}[m, n] \sim \mathcal{U} \left[ -\frac{1}{2}\Delta_{can}; \frac{1}{2}\Delta_{can} \right].$$

où  $\mathcal{U}[a; b]$  représente la loi de distribution uniforme (continue) sur l'intervalle  $[a; b]$ .

Le but de cette section étant de fournir un modèle statistique de l'image enregistrée par un appareil avant application des traitements post-acquisition, la quantification finale (sur 8 bits) et la compression de l'image ne sont pas considérées ici. En outre, la quantification de la valeur des pixels d'une image et son impact sur la détection statistique ont déjà été longuement étudiés dans les sections 3.1 - 3.4.

### 4.3.2 Modèle simplifié du bruit

Les caractéristiques des phénomènes stochastiques décrits dans la section 4.3.1 sont résumées dans le tableau 4.2 ; il est toutefois rappelé que cette liste n'a pas vocation à l'exhaustivité mais correspond aux principaux phénomènes [129, 130].

Bruits dépendants de l'intensité du signal.		
Bruit	Modélisation	Origines / Propriétés
NUEQ $K[m, n] \sim \mathcal{N}(1, \sigma_K^2)$	$N_{ep} \propto K[m, n] \cdot Q[m, n]$ variant spatialement	$\propto Q[m, n]$ (4.21),
Comptage des photons	$N_{ep} \sim \mathcal{P}(\mu_{ep}[m, n])$	hétéroscédastique variant spatialement
Bruits stationnaires.		
Bruit	Modélisation	Origines / Propriétés
Courant obscur	$N_{co} \sim \mathcal{P}(\mu_{co})$	d'origine thermique $\propto$ durée d'exposition
Lecture Amplification	$N_{lec} \sim \mathcal{N}(0, \sigma_{lec}^2)$	bruit additif, blanc et Gaussien
Conversion analog.-num.	$N_{can} \sim \mathcal{U} \left[ \frac{-\Delta_{can}}{2}; \frac{\Delta_{can}}{2} \right]$	bruit de quantification additif de variance $\Delta_{can}^2/12$ .

TABLE 4.2: Bruits d'acquisition pris en compte dans le modèle statistique d'une image.

La distribution statistique rigoureusement exacte de la valeurs de chacun des pixels est difficilement exploitable en pratique ; en particulier, le paramètre de la distribution de Poisson est lui-même la réalisation d'une variable aléatoire puisque selon (4.21)  $\frac{\lambda}{hc} Q[m, n] q_{ep} K[m, n]$ , avec  $K[m, n] \sim \mathcal{N}(1, \sigma_K^2)$ . Il est donc proposé quelques approximations permettant une simplification de la loi de distribution de la valeur des pixels :

- Le nombre d'électrons collectés dans les pixels est très grand. Il est donc possible d'approcher avec une grande précision la distribution Poissonienne de  $N_{e^-}[m, n]$  par une distribution Gaussienne ; on a alors :

$$N_{e^-}[m, n] \sim \mathcal{N}(\mu_{ep}[m, n] + \mu_{co}, \mu_{ep}[m, n] + \mu_{co}). \quad (4.25)$$

- D'autre part, compte-tenu du fait que  $\sigma_K^2 \ll 1$ , il peut être proposé de négliger la NUEQ en considérant que [129]  $\forall (m, n) \in \mathcal{I}, K[m, n] = 1$ . Une approche alternative, retenue dans le présent mémoire, consiste à modéliser le bruit de NUEQ

La charge maximal du condensateur d'un pixel est usuellement comprise entre  $60.000e^-$  et  $300.000e^-$

comme un phénomène intervenant de façon additive dans le nombre d'électrons d'origine photonique [130, 131] :

$$K[m, n] \simeq \mathcal{N}(1, \sigma_K^2) \Rightarrow N_{ep}[m, n] \simeq \mathcal{N}\left(\mu_{ep}[m, n], \mu_{ep}[m, n] + \mu_{ep}[m, n]^2 \sigma_K^2\right). \quad (4.26)$$

En utilisant les simplifications (4.25) et (4.26) dans les relations (4.21) et (4.22), il est immédiat de décrire statistiquement le nombre d'électrons collectés par :

$$\begin{aligned} N_{e^-}[m, n] &= N_{ep}[m, n] + N_{co}[m, n] \\ \Rightarrow N_{e^-}[m, n] &\sim \mathcal{N}\left(A(\mu_{ep}[m, n] + \mu_{co}), A^2(\mu_{co}\mu_{ep}[m, n] + \mu_{ep}[m, n]^2 \sigma_K^2)\right). \end{aligned}$$

Les simplifications (4.25) et (4.26) permettent de décomposer le bruit d'acquisition sous la "forme canonique" suivante

$$\underbrace{z[m, n]}_{\text{signal mesuré}} = \underbrace{\theta[m, n]}_{\text{partie déterministe}} + \underbrace{N_{img}[m, n]}_{\text{bruit hétéroscédastique}} + \underbrace{N_{ind}}_{\text{bruit stationnaire}} \quad (4.27)$$

avec  $N_{img}[m, n] \sim \mathcal{N}(0, A^2\mu_{ep}[m, n] + A^2\sigma_K^2\mu_{ep}[m, n]^2)$

et  $N_{ind} \sim \mathcal{N}(0, A^2(\sigma_{lec}^2 + \mu_{co}))$ .

Dans la relation (4.27), le terme  $N_{img}[m, n]$  est la réalisation d'une variable aléatoire représentant les bruits d'acquisition dont la distribution dépend de l'intensité moyenne du pixel,  $\theta[m, n]$ , et, à l'inverse,  $N_{ind}$  est une variable aléatoire représentant les bruits d'acquisition indépendant de l'énergie reçue par le pixel  $(m, n)$ .

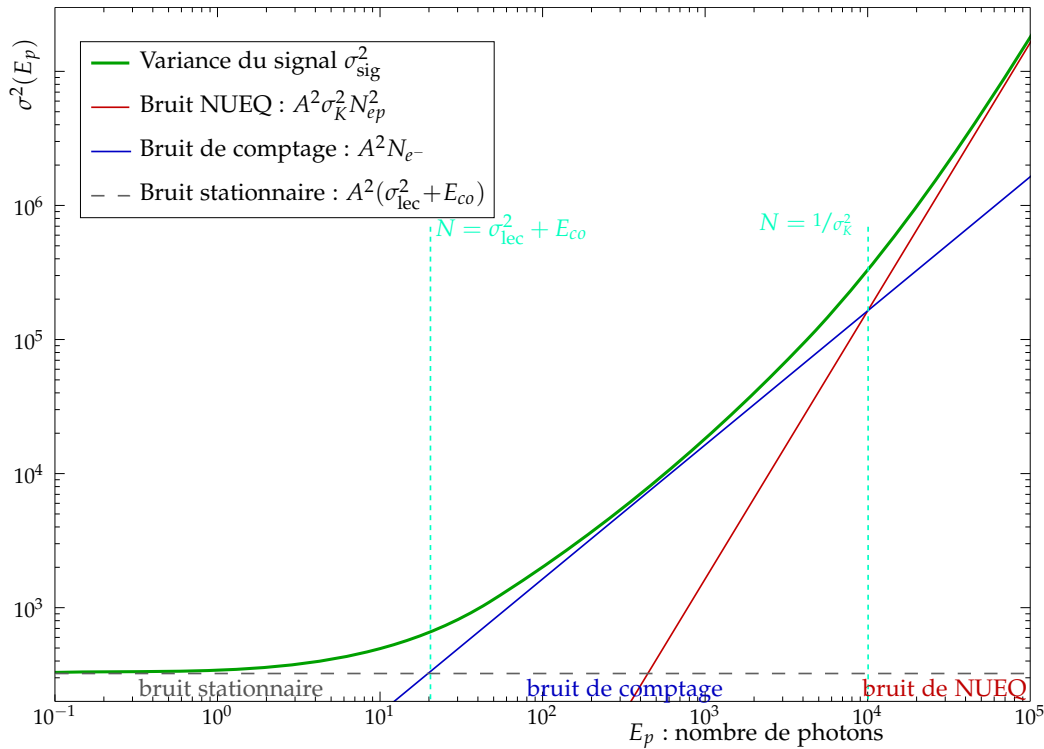


FIGURE 4.4: Importance relative de chaque type de bruit en fonction du nombre de photo-électrons collectés. Dans cet exemple, les paramètres d'acquisition sont caractérisés par un facteur d'amplification  $A = 4$ , un bruit de lecture avec  $\sigma_{lec} = 4$ , un bruit de NUEQ avec  $\sigma_K = 0.01$  et un bruit électronique de courant obscur générant en moyenne  $\mu_{co} = 4e^-$  électrons.

Enfin, l'espérance de chacun des pixels  $\theta[m, n] = \mathbb{E}[z[m, n]]$ , représentant le contenu de l'image numérique, peut également être décomposée de la façon suivante :

$$\underbrace{\theta[m, n]}_{\text{partie déterministe}} = \underbrace{A\mu_{ep}[m, n]}_{\text{signal lumineux}} + \underbrace{A\mu_{co}}_{\text{"biais"}} \quad (4.28)$$

où, selon (4.21) le contenu de l'image dû à l'énergie lumineuse incidente est donnée par :

$$\mu_{ep}[m, n] = \mathbb{E}[N_{ep}[m, n]] = \frac{\lambda}{hc} Q[m, n] q_{ep} K[m, n].$$

L'importance relative de chaque type de bruit en fonction du nombre de photo-électrons collectés est illustrée dans la figure 4.4. Il est notable qu'à très faible flux,  $N_{ep} \lesssim \sigma_{lec}^2 + \mu_{co}$ , ce sont les bruits électronique et de courant obscur dont l'importance est prépondérante. À l'inverse lorsque le nombre de photons incidents par pixel devient très important,  $N_{ep} \gtrsim 1/\sigma_k^2$ , le "bruit" de NUEQ devient prépondérant. Enfin, dans une région médiane  $\sigma_{lec}^2 + \mu_{co} \lesssim N_{ep} \lesssim 1/\sigma_k^2$ , qui correspond assez raisonnablement au nombre de photons collectés par les pixels d'un appareil photographique le bruit quantique de comptage des photons a une importance prépondérante.

*Pour la radiographie, le flux de rayons X étant faible les détecteurs sont refroidis à environ  $-40^\circ\text{C}$  afin d'accroître le rapport signal-sur-bruit.*

## Conclusions du chapitre :

Au cours de ce chapitre, une description et une modélisation du système d'acquisition d'une image naturelle ont été présentées. L'analyse des phénomènes physiques donnant naissance à l'image numérique permet de mettre en exergue deux particularités des images numériques :

- Le bruit présent dans une image est non-stationnaire : la variance du bruit d'acquisition d'un pixel dépendant, notamment, de l'énergie reçue par ce pixel et donc n'est pas identique sur l'ensemble d'une image numérique.
- Le système optique est également non-stationnaire, il ne peut rigoureusement être modélisé comme un système invariant par translation : divers phénomènes optiques engendrent une réponse dépendante des coordonnées du point  $(x, y)$  de l'image considéré.

Ces deux phénomènes ont des impacts distincts. La non-stationnarité du bruit doit être prise en compte dans le test statistique visant à détecter d'éventuelles informations cachées, comme par exemple dans les RV (3.11) et (3.53).

La non-stationnarité du système optique d'acquisition doit être prise en compte dans la modélisation du contenu,  $\theta$ , des images. Toutefois, il apparaît clairement qu'une modélisation exploitable des images n'est possible que si l'on peut modéliser la scène  $S(x, y)$ . À la suite du présent chapitre 4, il demeure nécessaire d'étudier les propriétés physiques de la scène imagée afin de proposer une modélisation des images naturelles acquises, ce point est abordé dans le chapitre suivant 5.

## 4.4 Annexes du chapitre 4 : De l'image "RAW" à l'image "TIFF"

Il est à noter en préambule que dans le cadre des travaux présentés, le modèle fondamental d'acquisition des images est donné dans les sections 4.1-4.3. Cependant, si la méthodologie proposée ne demeure applicable que pour des images "raw", l'intérêt pratique en sera très limité. Le but de cette section est donc de présenter brièvement les principaux éléments de la chaîne des traitements post-acquisition afin que le modèle des images proposé dans le chapitre 5 suivant puisse être utilisé pour analyser des images "tiff" au prix de légères modifications.

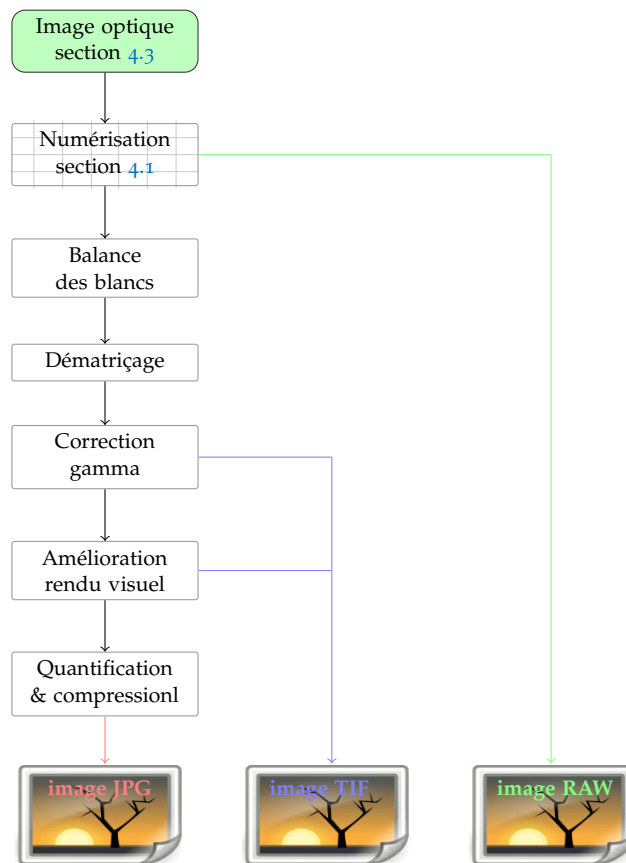


FIGURE 4.5: Chaîne des traitements post-acquisition appliquée à une image numérique préalablement à son enregistrement par l'appareil.

De nombreux traitements post-acquisition peuvent potentiellement être appliqués par l'appareil de photographie numérique après acquisition de l'image brute (4.24). Dans le cadre des travaux présentés, les traitements post-acquisition ne sont pas considérés dans la modélisation des images numérique en raison des deux problèmes pratiques majeurs suivants :

- Les constructeurs n'intègrent pas tous les mêmes méthodes numériques permettant une amélioration du rendu visuel des images.
- En outre, il est très difficile pour un utilisateur de connaître précisément les différents traitements et leurs implémentations, les constructeurs préférant ne pas rendre ces informations publiques pour des raisons commerciales évidentes.

Dans la suite de cette annexe, la chaîne des traitements post-acquisition est réduite aux trois éléments indispensables qui la composent [132] : le dématriçage, la correction gamma et la balance des blancs, voir figure 4.5. De ces trois traitements post-



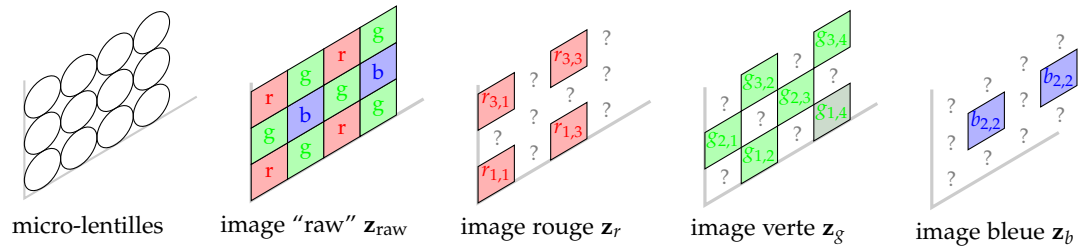


FIGURE 4.6: Illustration de la colorisation des pixels d'une image et de la problématique du dématricage.

acquisition, le dématricage est largement le plus complexe (autant d'un point de vue conceptuel que calculatoire) ; ce traitement est donc le plus détaillé dans la présente section.

#### 4.4.1 Dématricage d'une image "RAW"

Le lecteur attentif aura remarqué que la notion de couleur est absente des sections 4.1 et 4.3. En effet, seule des relations radiométriques et statistiques entre l'énergie rayonnée par la scène et la valeur des pixels ont été établies au travers de (4.15) - (4.24). Schématiquement, les pixels accumulent l'éclairement incident sur l'ensemble des longueurs d'onde. La couleur est introduite artificiellement dans les images par l'ajout d'un filtre devant chacun des pixels. Sous l'effet de ce filtre, seule une plage de longueur d'onde limitée atteint la surface photosensible de chacun des pixels ; ces derniers agissent alors comme les cônes de la rétine en étant sensibles à un intervalle déterminé de longueur d'onde correspondant à un canal de couleur.

*L'ajout de ce filtre ne change guère le modèle proposé de la valeur des pixels (4.24) ; le filtre couleur étant simplement ajouté dans l'expression de  $q_{acq}(\lambda)$ .*

Différentes matrices de filtres couleurs, également appelé *Color Filter Array* ou CFA, ont été proposées. Ce manuscrit n'ayant pas vocation à présenter de façon exhaustive les traitements post-acquisition, seul le CFA défini par Bayer [133] est considéré dans la suite. La structure de base du CFA de Bayer est une matrice de taille  $2 \times 2$  composées de trois filtres couleurs, correspondant respectivement aux canaux rouge, vert et bleu, dont la répartition est représentée dans la figure 4.6. La sensibilité de la vision humaine étant plus importante dans l'intervalle de longueur d'onde correspondant aux couleurs vertes, deux filtres verts sont présents dans la matrice de base du CFA de Bayer.

Quel que soit le CFA utilisé, l'image "raw" obtenue est composée de différents canaux de couleurs indexés par  $c \in \mathcal{C}$  (usuellement  $\mathcal{C} = \{r, g, b\}$ ), mais la valeur mesurée pour chaque pixel ne correspond qu'à un seul de ces canaux. Or pour obtenir une image couleur, les pixels doivent posséder une valeur d'intensité pour chacun des canaux. Le problème du dématricage consiste donc à retrouver, pour chaque pixel, l'intensité des canaux de couleurs correspondant aux filtres différents de celui situé devant le pixel considéré.

De façon générale, on peut distinguer deux méthodologies différentes pour aborder le problème du dématricage. Les premiers algorithmes proposés abordent le dématricage comme un problème d'interpolation [134]. Historiquement ces méthodes présentent l'avantage d'avoir un coût calculatoire très faible et de pouvoir être implémentées efficacement sur des DSP (processeurs de signaux numériques). En utilisant la figure 4.6, il est par exemple immédiat de constater qu'une interpolation linéaire bidimensionnelle (appelée aussi interpolation bilinéaire) permet de calculer les pixels manquants du canal vert comme la somme des quatre pixels adjacents, pondérée par un facteur  $1/4$ . La même interpolation bilinéaire sur les canaux rouge et bleu permet d'écrire cette méthode de dématricage sous la forme de convolutions discrètes :

$$\hat{z}_c = z_c ** H_c,$$

où les noyaux de convolutions  $H_g$  et  $H_{r,b}$ , respectivement pour le canal vert et pour les canaux rouge et bleu sont, données par :

$$H_g = \frac{1}{4} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad H_{r,b} = \frac{1}{4} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

Les différentes méthodes de dématricage par interpolation engendrent notamment des artefacts chromatiques dans les régions où l'intensité varie de manière importante, notamment au niveau des contours des objets [134, 135]. Pour amoindrir ces défauts des méthodes adaptatives de dématricage par interpolation ont été proposées. Le principe est de classer chaque pixel dont la valeur doit être calculée selon que ce dernier appartienne à un contour vertical ou horizontal. L'interpolation est ensuite réalisée dans la direction orthogonale au contour supposé. Il serait trop long de dresser un état détaillé de l'art de ces méthodes heuristiques proposées, le lecteur intéressé pourra néanmoins consulter les articles [134, 135, 136, 137].

La seconde approche souvent rencontrée dans le dématricage est de considérer qu'il s'agit d'un problème inverse mal posé. Formellement, l'ensemble des pixels  $\Omega$  est partitionné en  $|\mathcal{C}|$  sous ensemble  $\Omega = \bigcup_{c \in \mathcal{C}} \Omega_c$  où  $\Omega_c$  représente l'ensemble des pixels appartenant au canal de couleur  $c \in \{r, g, b\}$ . Le phénomène de dématricage peut être représenté algébriquement en notant  $\mathbf{z}_{\text{raw}} \in \mathcal{M}_{MN,1}$  le vecteur des pixels raw rangés dans l'ordre lexicographique  $\mathbf{z}_{\text{raw}} = (r_{1,1}, g_{1,2}, \dots, b_{M,N-1}, g_{M,N})^T$ . De la même façon en notant  $\mathbf{z}_c$  le vecteur des pixels raw du canal de couleur  $c \in \{r, g, b\}$ , on peut représenter la relation entre  $\mathbf{z}_{\text{raw}}$  et  $\mathbf{z}_c$  sous la forme de la projection associée à la matrice diagonale creuse  $\mathbf{M}_c \in \mathcal{M}_{MN,MN}$  de matricage du canal de couleur  $c$  qui, dans le cas du filtre de Bayer, est donnée par :

$$\mathbf{M}_c \mathbf{z}_{\text{raw}} = \mathbf{z}_c \quad \text{avec} \quad \mathbf{M}_c = \begin{pmatrix} \delta_{1,c} & 0 & \cdots & 0 & 0 \\ 0 & \delta_{2,c} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \delta_{MN-1,c} & 0 \\ 0 & 0 & \cdots & 0 & \delta_{MN,c} \end{pmatrix}$$

où le symbole de Kroenecker  $\delta_{i,c}$  est défini par  $\delta_{i,c} = 1$  si  $i \in \Omega_c$  et  $\delta_{i,c} = 0$  si  $i \notin \Omega_c$ . Il est immédiat de constater que la propriété de partitionnement de  $\Omega$  se traduit algébriquement par  $\sum_{c \in \mathcal{C}} \mathbf{M}_c = \mathbf{I}_{MN}$ . De manière analogue, cette propriété se traduit par le fait que pour tout  $c \in \{r, g, b\}$  on a  $\text{im}(\mathbf{M}_c) = \Omega_c$  et  $\text{ker}(\mathbf{M}_c) = \Omega \setminus \Omega_c$ . En notant  $\bar{\mathbf{z}}_c$  la valeur (théorique) de tous les pixels de  $\Omega$  pour le canal  $c$ , le problème du dématricage consiste donc à proposer une estimation  $\hat{\mathbf{z}}_c$  de  $\bar{\mathbf{z}}_c$  à partir des valeurs connues  $\mathbf{z}_c = \mathbf{M}_c \bar{\mathbf{z}}_c$ .

**Définition 4.5** (Problème inverse). Soit  $s \in \Omega_s$  un signal quelconque défini par  $s = f(u)$  où  $u \in \Omega_u$  est un paramètre d'observation. Un problème inverse consiste à déterminer la valeur du paramètre d'observation  $u$  à partir du signal mesuré  $s$ .

typiquement  $\Omega_s \subseteq \mathbb{R}^N$  si le signal est échantillonné et  $\Omega_s \subset \mathcal{L}^2(\mathbb{R})$  si le signal est à temps continu

**Définition 4.6** (Problème mal posé). Soit  $s \in \Omega_s$  un signal quelconque défini par  $s = f(u)$  et soit le problème inverse consistant à déterminer la valeur du paramètre d'observation  $u \in \Omega_u$  à partir du signal  $s$ . Selon J.Hadamard [138], un problème inverse est bien posé si et seulement si les trois conditions sont vérifiées :

1. existence d'une solution,  $f$  est surjective :  $\forall s \in \Omega_s, \exists u \in \Omega_u / s = f(u)$ ;
2. unicité de la solution,  $f$  est injective :  $\forall (u_0, u_1) \in \Omega_u^2, u_0 \neq u_1 \Rightarrow f(u_0) \neq f(u_1)$ ;
3. stabilité de la solution : la solution unique  $u$  dépend continûment de  $s$ .

Par contraposée, le problème est mal posé si et seulement si l'une des trois conditions n'est pas respectée.

Dans le cas du dématricage, pour tout  $c \in \mathcal{C}$  on a  $\ker(\mathbf{M}_c) \neq \emptyset$ ; il en découle immédiatement que l'endomorphisme associé à la matrice  $\mathbf{M}_c$  n'est pas injectif,  $\forall \xi_c \in \ker(\mathbf{M}_c)$ ,  $\mathbf{M}_c \bar{z}_c = \mathbf{M}_c (\bar{z}_c + \xi_c)$ . En outre, compte-tenu de la structure diagonale creuse des matrices de projections  $\mathbf{M}_d$ , il n'est pas possible d'inverser ou même de pseudo-inverser (au sens de Penrose ou de Moore)  $\mathbf{M}_d$ .

Néanmoins, certaines méthodes de résolution des problèmes inverses demeurent exploitables dans le cadre du dématricage [135, 134], notamment les méthodes issues de la régularisation, les méthodes Bayésiennes (avec modèle *a priori* Markovien), ou encore les méthodes issues de l'apprentissage supervisé et de la reconnaissance des formes [139, 140, 141].

De manière générale, les méthodes de dématricage exploitant les algorithmes de résolution numérique des problèmes inverses ne sont pas utilisées en pratique dans les appareils de photographie en raison de leur complexité calculatoire prohibitive.

Les méthodes de dématricage fondées sur une approche de régression, les plus couramment implémentés dans les appareils photographiques, modifient peu (voire pas du tout) le modèle de l'intensité des pixels proposé.

#### 4.4.2 Balance des blancs

Le système visuel humain possède la capacité de percevoir la couleur des objets sous des conditions d'éclairage très différentes : notre perception des couleurs ne change par exemple pas (ou très peu) au fil de la journée alors que le spectre de la lumière issu du soleil change (par effet de filtration de l'atmosphère). Cette capacité d'adaptation au spectre de la lumière ambiante résulte d'une interprétation des images observées et, de façon générale, d'un mécanisme de rétro-action permettant une adaptation de la sensibilité à un stimulus par le stimulus même [142].

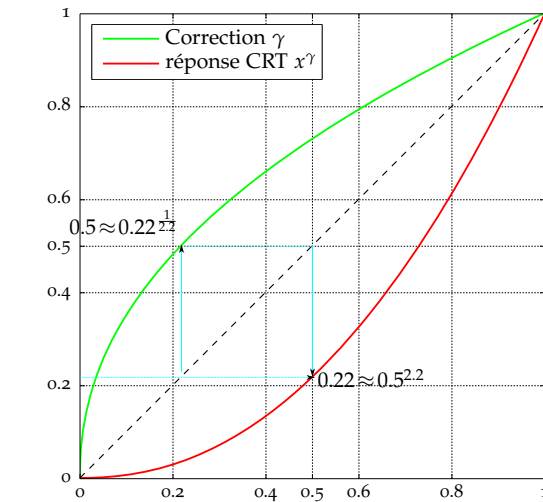
Comparativement à notre oeil, un photo-détecteur est un appareil de mesure et ne possède donc pas cette capacité d'adaptation au spectre de la lumière incidente. Notre vision des couleurs d'une scène peut donc sensiblement varier des couleurs restituées lors de l'acquisition d'une image. Il est donc nécessaire d'intégrer dans les appareils photographiques un système d'auto-adaptation au spectre de la lumière ambiante pour obtenir un rendu des couleurs proche de celui des utilisateurs. La "balance des blancs" est à l'origine de cette adaptation, elle est principalement réalisée par amplification de chaque canal de couleur (rouge, vert et bleu). En réalité seuls 2 coefficients multiplicatifs sont nécessaires, en considérant par exemple que l'intensité du canal vert demeure inchangée. En notant  $\bar{z}_c$  l'intensité des pixels du canal  $c \in \mathcal{C}$  après dématricage (sous la forme d'un vecteur), l'opération de balance des blancs peut être représentée sous la forme :

$$\bar{z}_c^b = w_c \cdot \bar{z}_c, \quad (4.29)$$

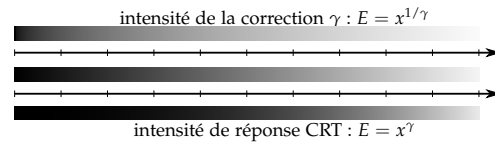
où  $\bar{z}_c^b$  représente l'intensité des pixels du canal  $c$  après balance des blancs et  $w_c$  est le facteur d'amplification. La problématique de la balance des blancs se résume donc à fixer les facteurs d'amplification  $w_c$ .

Il existe trois grandes méthodes d'algorithmes de balance des blancs, ou plus précisément de détermination des coefficients  $w_c$ , qui sont les algorithmes automatiques, semi-automatiques, et manuels. Sans rentrer dans le détail de ces différents algorithmes, il est important de souligner que quelle que soit la méthode de balance des blancs utilisée, la distribution statistique des pixels n'est modifiée qu'au travers de la multiplication par un scalaire. Le modèle statistique proposé pour décrire la valeurs des pixels (4.27) - (4.24) n'est donc *mutatis mutandis* pas modifié.

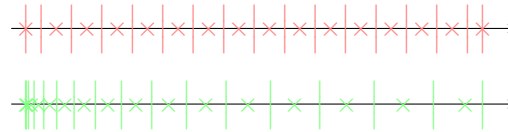
Le système visuel humain s'adapte également à l'intensité de l'éclairage reçu. Cette adaptation se fait de façon plus mécanique (par dilatation/rétraction de la pupille d'entrée) et est généralement présente dans les appareils photographiques sous le nom de "auto-exposure".



(a) Réponse d'un appareil type CRT et de la correction  $\gamma$  d'un appareil photo en fonction de la tension du signal d'entrée.



(b) Réponse d'un appareil type CRT et de la correction  $\gamma$  d'un appareil photo en fonction de la tension du signal d'entrée.



(c) Illustration de l'impact de la correction  $\gamma$  sur la quantification des données.

FIGURE 4.7: Illustration de la réponse non-linéaire des tubes cathodiques, de la fonction de correction  $\gamma$ .

### 4.4.3 Correction gamma

Selon le processus d'acquisition d'une image décrit dans les sections 4.1 et 4.3, la valeur d'un pixel dépend linéairement de l'énergie incidente à la surface de ce dernier elle-même linéairement dépendante de l'énergie émise par la scène. Historiquement, la correction gamma proviendrait de la non-linéarité de la réponse du tube cathodique[121]. Soumis à une tension  $U$  la luminance d'une tube cathodique est schématiquement donnée par  $S = f_\gamma(U) = U^\gamma$  avec  $0.39 \lesssim \gamma \lesssim 0.46$ . Or idéalement, la luminance d'une scène observée sur un écran doit être proportionnelle à la luminance de la scène imagée.

Pour rétablir une relation linéaire entre ces deux quantités, il aurait été proposé d'introduire une relation non linéaire entre énergie  $Q_{m,n}$  reçue et valeur  $z_{m,n}$  d'un pixel de la forme  $z_{m,n} = Q_{m,n}^{1/\gamma}$  de sorte que si le signal  $\mathbf{z}$  est ensuite utilisé pour un affichage sur un écran cathodique on ait  $S = z_{m,n}^\gamma = (Q_{m,n}^{1/\gamma})^\gamma = Q_{m,n}$ .

La très grande majorité des appareils photographiques utilisent comme valeur  $\gamma$  de référence pour le standard NTSC, stipulant que  $\gamma = 2,2$ .

*Il est en outre notable que d'un point de vue signal, la correction gamma a un avantage certain pour la compression. La correction gamma permet de coder des signaux dont l'étendue des valeurs est large et permet une quantification adaptée à la représentation des faibles valeurs.*

Les fonctions non-linéaires caractérisant la réponse des écrans type cathodique  $f[0,1] \rightarrow [0,1] : x \mapsto x^\gamma$  et la correction gamma associée  $g : x \mapsto x^{1/\gamma}$  sont représentées sur la figure 4.7 ci-contre.

---

# Vers une modélisation des images numériques

---

# 5

«*La science ne cherche pas à énoncer des vérités éternelles, [...], elle cherche à cerner la vérité par approximations successives.*»

— Bertrand Russell, 1872 – 1970

«*Modelling is an essential and inseparable part of all scientific.*»

— William Silvert, 1953 – ...

## Sommaire

---

Introduction au chapitre . . . . .	101
5.1 Modèle du contenu structuré des images naturelles . . . . .	102
5.1.1 Modélisation de la scène imagée . . . . .	102
5.1.2 De la scène à un modèle de l'image optique formée sur le capteur . . . . .	103
5.1.3 Modèle continu du contenu des images en 1D . . . . .	104
5.2 Nécessité de la modélisation non-linéaire des discontinuités . . . . .	109
5.2.1 Modélisation locale retenue du contenu des images . . . . .	110
5.3 Discussions sur quelques modèles alternatifs des images . . . . .	112
5.3.1 Approximation par transformation en ondelettes . . . . .	112
5.3.2 Apprentissage de dictionnaire parcimonieux . . . . .	114
Conclusions du chapitre . . . . .	115
5.4 Annexes du chapitre 5 : démonstrations des théorèmes 5.1 et 5.2 . . . . .	116

---

### Introduction au chapitre :

À partir de la modélisation du système d'acquisition présentée dans le chapitre (4), le présent chapitre est dévolu à la construction d'un modèle physique de l'image optique incidente sur le photo-détecteur. Pour cela, il est nécessaire d'analyser les propriétés physiques de la scène  $S(x, y)$  imagée, et d'étudier l'énergie incidente sur le plan du photo-détecteur après passage par le système optique d'acquisition.

La section 5.1.1 présente un modèle générique décrivant la scène comme une juxtaposition d'objets solides distincts au sein desquels les propriétés lumineuses varient de façon continue. Ce modèle de la scène est ensuite exploité, dans la section 5.1.3, conjointement avec la description de la chaîne d'acquisition présentée dans le chapitre 4 pour obtenir une modélisation de l'image numérique acquise. Afin de proposer une modélisation des images tenant compte des propriétés physiques non-stationnaires du système d'acquisition, le modèle est simplifié en analysant l'image suivant une dimension. Enfin, la section 5.2 présente formellement quelques approximations permettant de proposer un modèle local paramétrique d'une image analysée.

L'intérêt de ce modèle dans le cadre des travaux présentés est souligné dans la section 5.3 au travers d'une discussion autour d'autres modèles d'images usuellement rencontrés dans la littérature.

Le chapitre 4 est dévolu à la description et à la et modélisation des différents phénomènes physiques et électroniques intervenant au cours de processus d'acquisition d'une image naturelle. Néanmoins, cette analyse du processus d'acquisition d'une image ne permet une modélisation de l'image acquise que si la scène, décrite par sa luminance  $S(x, y)$ , est connue. Il est donc proposé dans le présent chapitre de modéliser la scène  $S(x, y)$  imagée puis, d'exploiter l'analyse de la chaîne d'acquisition décrite dans le chapitre 4, afin obtenir une modélisation de l'image numérique acquise fondé sur les propriétés physiques de la scène et du système d'acquisition. Dans la section 5.1, une modélisation simple de la scène est d'abord présentée puis analysée au cours des différentes étapes d'acquisition d'une image. En utilisant quelques résultats de la théorie de l'approximation, cette description de la scène imagée permet l'obtention d'un modèle local paramétrique de l'image acquise décrit en détails en dans la section 5.2.

## 5.1 Modèle du contenu structuré des images naturelles

### 5.1.1 Modélisation de la scène imagée

Afin de modéliser l'image numérique produite à partir d'une scène naturelle, il est nécessaire d'analyser le canal optique donnant naissance à l'image et de décrire de façon générique la scène imagée. Le canal optique d'acquisition a été décrit dans le chapitre 4; le modèle simplifié des dégradations déterministes proposé dans la section 4.2 est notamment largement utilisé tout au long de cette section 5.1.

Modéliser la scène (ou plus précisément sa luminance spectrale) est l'objectif poursuivi dans cette section 5.1. La PSF ne dépendant pas de la longueur d'onde de la lumière (voir propriétés 4.1 de la section 4.2), la luminance de la scène est caractérisée, pour le canal de couleur  $c \in \mathcal{C} = \{r; g; b\}$  :

$$S^c(x, y) = \int_{\mathbb{R}^+} S_0(x, y; \lambda) q_{\text{opt}}^c(\lambda) d\lambda,$$

représentant la fraction de puissance rayonnée par le point  $(x, y)$ , dans le spectre du canal  $c$ , incidente sur le capteur optique.

**Remarque 5.1.** Dans la suite de ce manuscrit, les canaux de couleurs sont considérés individuellement aussi l'indice  $c$  est omis.

Le modèle de la luminance de la scène  $S(x, y)$  utilisé est directement inspiré de [125, 143, 144] et illustré par la figure 5.1. La scène est supposée constituée de  $N_O$  différents objets  $O_i$  possédant chacun des propriétés de réflectance spécifiques. Du point de vue du système optique, chaque objet  $O_i$  est donc associé à une fonction de luminance  $S_i(x, y)$  et à un compact ouvert  $\mathcal{X}_i \subset \mathcal{X}$  de sorte que  $S_i(x, y) \neq 0 \Leftrightarrow (x, y) \in \mathcal{X}_i$ . Ainsi, une modélisation générique de la scène peut être la suivante :

$$S(x, y) = \sum_{i=1}^{N_O} S_i(x, y) \mathbf{1}_{\mathcal{X}_i}(x, y), \quad (5.1)$$

où  $\mathbf{1}_E$  symbolise la fonction indicatrice de l'ensemble  $E$  définie par :

$$\mathbf{1}_E(x) = \begin{cases} 1 & \text{si } x \in E \\ 0 & \text{sinon.} \end{cases}$$

Les frontières des sous-ensembles  $\mathcal{X}_i$ , notées  $\partial\mathcal{X}_i$ , représentent physiquement les arêtes des différents objets  $O_i$  qui forment les discontinuités de la luminance  $S$  de la scène. Il est supposé que chacune des frontières peut être représentée sous la forme d'une

*Le modèle de la scène présenté est proche du modèle des feuilles mortes [145] aux différences près que la densité spectrale de puissance d'une image n'est pas considéré ici dans le modèle et que les objets ne sont pas supposés avoir une luminance constante.*

*En pratique les arêtes présentes dans la scène peuvent avoir diverses origines physiques, présence d'ombres, discontinuité de la réflectance d'un objet, etc... Ces différents types d'arêtes ne sont pas distinguées dans ce manuscrit.*

courbe paramétrique notée  $\gamma_i : J_i \mapsto \mathcal{X}$  avec  $J_i \subset \mathbb{R}$ . Aussi, l'ensemble des frontières,  $\Gamma = \bigcup_{i=1}^{N_O} \partial \mathcal{X}_i$ , peut également être représenté sous la forme d'une courbe paramétrique  $\Gamma : J \mapsto \mathcal{X}$  qui admet la représentation suivante :

$$\Gamma(t) = \sum_{i=1}^{N_C} \gamma_i(t) \mathbf{1}_{J_i}(t), \quad (5.2)$$

où  $J$  est un intervalle compact dans  $\mathbb{R}$  et  $\{J_i\}$  forme une partition finie de  $J$ . Pour permettre une formulation mathématique simple de la courbe paramétrique  $\Gamma$ , il est supposé sans perte de généralité, que les ensembles  $\{J_i\}_{i=1}^{N_O}$  ainsi que les ensembles  $\{\gamma_i(t); t \in J_i\}$ , représentant une arête, sont disjoints.

**Propriété 5.1.** *Comme souligné dans [143], compte-tenu de la nature de la luminance, issue du rayonnement des objets solides, les propriétés suivantes sont supposées vérifiées.*

1. Les fonctions  $S_i : (x, y) \rightarrow S_i(x, y)$  sont continues sur  $\mathcal{X}_i$ .
2. La luminance de la scène  $S$  est discontinue sur l'ensemble des frontières  $\{\Gamma(t), t \in J\}$ .
3. Les objets sont géométriquement réguliers par morceaux ; chacune des courbes  $\gamma_i$  représentant la position de la frontière  $\partial \mathcal{X}_i$  vérifie  $\gamma_i(t) \in C^2$  pour presque tout  $t \in J_i$ .

L'exploitation du modèle de la scène fourni par (5.1) - (5.2), est en pratique considérablement limitée par les deux difficultés majeures suivantes.

1. D'une part, la courbe paramétrique  $\Gamma$  doit être connue, ou de manière équivalente l'ensemble des courbes paramétriques  $\{\gamma_i\}$  représentant les frontières  $\partial \mathcal{X}_i$ .
2. D'autre part, la luminance des objets n'étant pas supposée constante mais continue, l'amplitude des discontinuités de la fonction  $S$  varie le long de la courbe  $\{\Gamma(t), t \in J\}$ .

Ces deux paramètres interviennent de façon non-linéaire dans le modèle de la luminance (5.1) et sont difficilement représentables avec un nombre limité (idéalement faible) de paramètres.

**Remarque 5.2.** *Il est en outre notable que les ensembles  $\mathcal{X}_i$  sont définis comme des compacts ouverts, correspondant à la surface intérieure des objets constituant la scène. Ainsi, selon la relation (5.1), la luminance spectrale de la scène n'est pas définie sur les frontières  $\partial \mathcal{X}_i$ . Cette considération théorique n'a guère d'importance en pratique puisque l'ensemble des frontières  $\{\Gamma(t), t \in J\}$  est de mesure négligeable au sens de Lebesgue.*

Le modèle de la luminance proposé dans (5.1) - (5.2) est suffisamment général pour permettre de décrire une large classe de scènes naturelles [143, 125]. Il est proposé d'étudier les propriétés de l'éclairement incident sur le plan du photo-détecteur à partir du modèle de la luminance proposé. Une attention toute particulière est portée à la présence des "courbes  $\gamma_i$  de discontinuités" inconnues qui, comme cela a été évoqué, rendent difficile l'exploitation pratique de ce modèle.

### 5.1.2 De la scène à un modèle de l'image optique formée sur le capteur

En utilisant le modèle de la scène (5.1) et le modèle de la PSF décrit dans la section 4.2, voir relation (4.17), l'éclairement  $E_i$  issu de l'objet  $O_i$  incident sur le photo-détecteur est défini par (4.18) :

$$\begin{aligned} E_i(x, y) &\stackrel{\text{def.}}{=} \int_{\mathbb{R}^2} \left( S_i(x - u, y - v) \mathbf{1}_{\mathcal{X}_i}(x - u, y - v) \right) h_{opt}(u, v, \zeta_0) du dv \\ &= \left[ S_i(x, y) \mathbf{1}_{\mathcal{X}_i}(x, y) \right] ** h_{opt}(u, v, \zeta_0). \end{aligned}$$



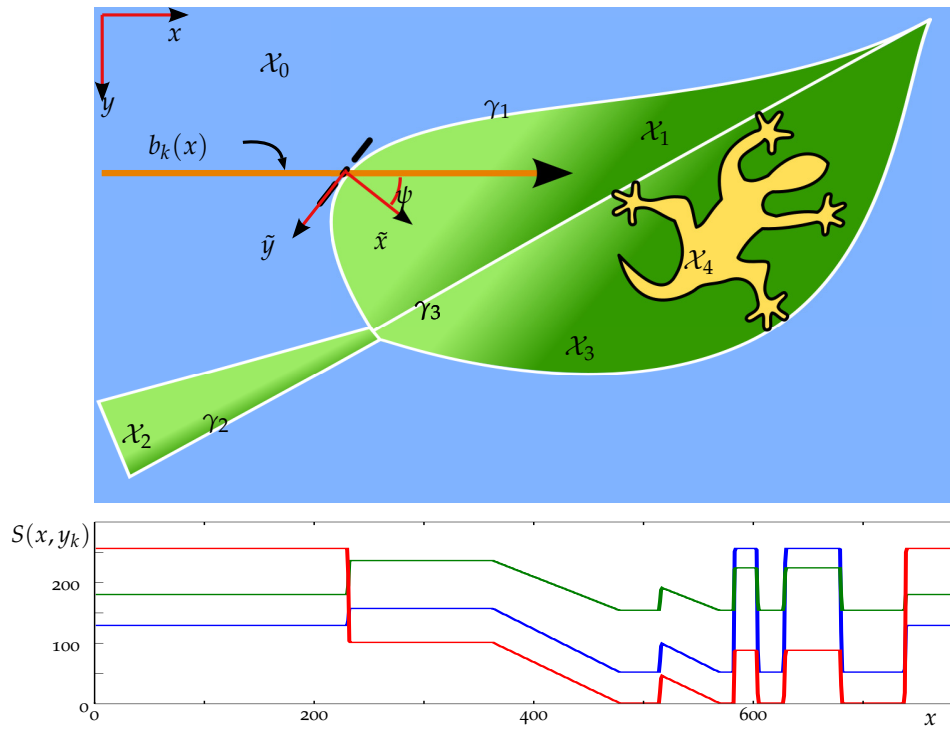


FIGURE 5.1: Illustration du modèle de scène local 1D pour une scène simple 2D (avant dégradation déterministes).

L'éclairement peut donc être modélisé de façon analogue à la scène, en vertu de la linéarité de l'intégration, comme la somme des éclairements de chacun des objets  $E_i$  :

$$E(x, y) = \sum_{i=1}^{N_0} E_i(x, y) = \sum_{i=1}^{N_0} [S_i(x, y) \mathbf{1}_{\mathcal{X}_i}(x, y)] ** h_{opt}(x, y; \zeta_0). \quad (5.3)$$

Enfin, la valeur théorique du signal échantillonné est obtenue par intégration sur la surface des pixels, voir (4.10), (4.15) et (4.19), et est donc finalement donnée, en omettant volontairement la corruption par les divers phénomènes stochastiques de bruit, par :

$$z_{m,n} = \theta(x, y) \Big|_{x=x_n; y=y_m} = \tilde{A} E(x, y) ** h_{int}(x, y) \Big|_{x=x_n; y=y_m} \quad (5.4)$$

où  $h_{int}$  représente la PSF liée à l'intégration spatiale et  $(x_n, y_m)$  les coordonnées du centre du pixel  $(m, n)$  (4.9) - (4.10). La constante  $\tilde{A} = A \tau \frac{\lambda}{hc} q_{ep}$  représente le facteur global de conversion de l'éclairement en une tension ; ce facteur est négligé dans la suite car ce dernier est supposé constant sur l'ensemble de l'image et n'a donc qu'une importance mineure dans le modèle proposé.

Ici encore, l'éclairement incident (5.3) et donc la valeur du signal échantillonné (5.4) peuvent difficilement être exprimés sous une forme analytique simple en raison de la PSF spatialement variante et en raison de la présence des arêtes des objets (ou plus formellement, en raison de la discontinuité de la luminance sur l'ensemble  $\{\Gamma(t), t \in J\}$ ).

### 5.1.3 Modèle continu du contenu des images en 1D

Afin d'obtenir un modèle de l'éclairement incident représentant fidèlement le contenu de la scène (5.1) - (5.3), tout en étant suffisamment simple pour permettre une utilisation pratique, il est proposé d'analyser l'image suivant une dimension (typiquement une ligne ou une colonne).

**Assertions 5.1.** Afin de donner une expression analytique de l'énergie  $\theta$  incidente, il est en outre considéré que les deux assertions suivantes sont vérifiées :

A-1 : La courbe  $\Gamma$  est régulière par morceaux,  $\Gamma \in \mathcal{C}^2(J, \mathbb{R}^2)$ , excepté en un nombre fini de points, voir propriétés 5.1. Il est supposée que la courbe paramétrique  $\Gamma : t \rightarrow \Gamma(t) = \begin{pmatrix} \Gamma_x(t) \\ \Gamma_y(t) \end{pmatrix}$  peut être fidèlement approximée, dans un voisinage de  $t_0 \in J$ , par la droite  $\bar{D}$  d'équation paramétrique :

$$\bar{D}(t_0 + \delta_t) = \Gamma(t_0) + \begin{pmatrix} \delta_t d\Gamma_x(t)|_{t=t_0} \\ \delta_t d\Gamma_y(t)|_{t=t_0} \end{pmatrix}$$

A-2 : L'éclairement incident est constant sur un intervalle de la largeur d'un pixel. Autrement dit, l'éclairement de la  $k$ -ième ligne de l'image noté  $\theta(x, y_k)$ , (resp.  $\theta(x_i, y)$  pour la  $k$ -ième colonne) vérifie :  $\forall y \in [y_k - \Delta_x/2; y_k + \Delta_x/2]$ ,  $\theta(x, y) = \theta(x, y_k)$ .

**Remarque 5.3.** Dans la suite de ce chapitre, la luminance de la scène  $S$ , l'éclairement du photo-détecteur  $\theta$  et le signal intégré spatialement  $\theta$  seront analysés en 1D par restriction de ces fonctions sur le segment  $D_{y_k} : \{(x, y) \in \mathcal{X} / y = y_k\}$ . Il est immédiat de généraliser les résultats présentés à toutes restrictions de ces fonctions sur un segment quelconque au prix d'une plus grande complexité des notations.

Avant de modéliser la valeur du signal continu enregistré  $\theta_k(x)$ , il est notable que selon la définition de la luminance de la scène (5.1), la fonction  $S_k(x) = S(x, y_k)$  n'est pas définie en un nombre fini de points, qui constituent l'ensemble noté  $\mathcal{D}_k$ . La nature de la fonction  $S_k$  est formalisée, pour tout  $y_k$ , par les définitions suivantes :

**Définition 5.1** (Singularité du premier ordre d'une fonction). Soit  $f : \mathcal{X} \mapsto \mathbb{R}$  une fonction continue définie sur le compact  $\mathcal{X}$ . La fonction  $f$  admet une singularité du premier ordre en  $x_1 \in \mathcal{X}$  si et seulement si les trois conditions suivantes sont vérifiées :

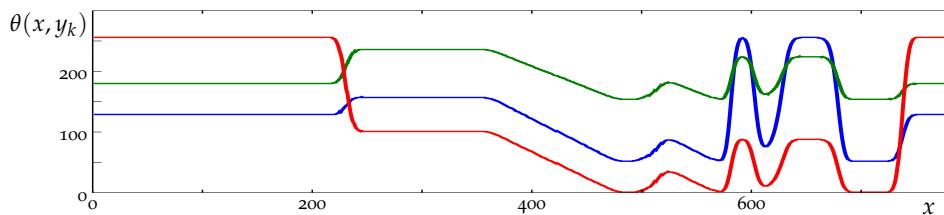
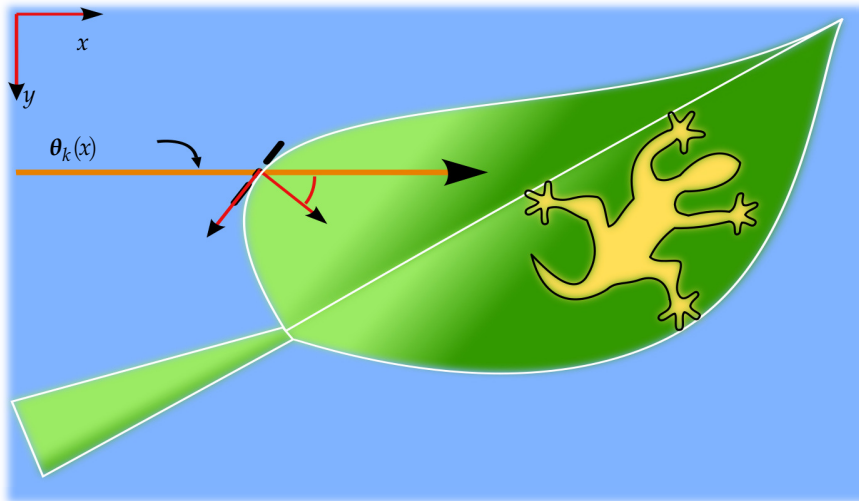


FIGURE 5.2: Illustration du modèle d'image local 1D pour une scène simple 2D (après dégradations déterministes).

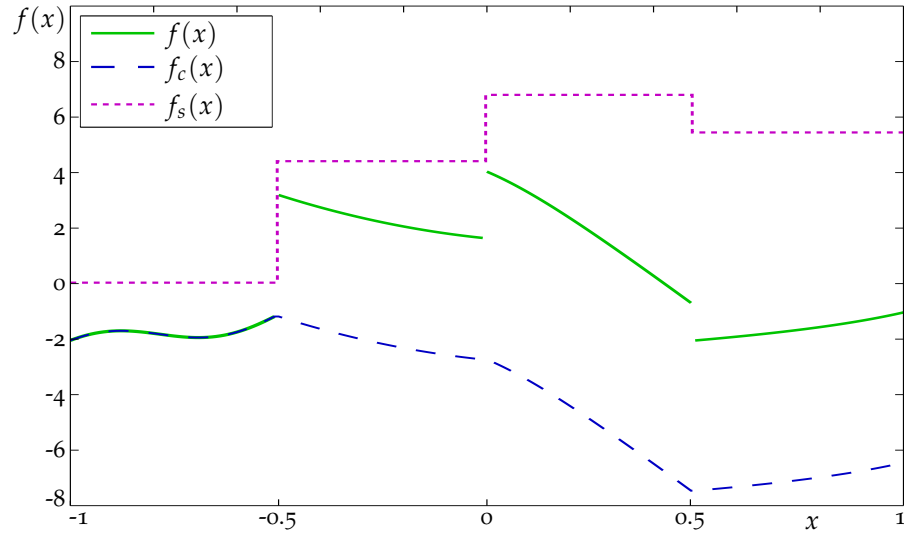


FIGURE 5.3: Illustration du théorème 5.1 sur la décomposition d'une fonction  $f$  continue par morceaux en une somme d'une fonction  $f_c$  continue et d'une fonction  $f_s$  constante par morceaux.

1.  $\lim_{\delta_x \rightarrow 0; \delta_x < 0} f(x_1 + \delta_x) = L^- \in \mathbb{R};$
2.  $\lim_{\delta_x \rightarrow 0; \delta_x > 0} f(x_1 + \delta_x) = L^+ \in \mathbb{R};$
3.  $L^+ \neq L^-$

**Définition 5.2** (Application continue par morceaux). Soit  $f : \mathcal{X} \mapsto \mathbb{R}$  une fonction continue sur le compact  $\mathcal{X} \setminus \{\mathcal{D}\}$ . La fonction  $f$  est dite continue par morceaux sur  $\mathcal{X}$  si et seulement si les trois conditions suivantes sont vérifiées :

1. l'ensemble  $\mathcal{D}$  est dénombrable :  $\mathcal{D} = \{x_1, x_2, \dots\}$ ;
2. pour tout  $x_i \in \mathcal{D}$ , la fonction  $f$  admet une singularité du premier ordre en  $x_i$ ;
3. pour tout  $x \notin \mathcal{D}$ , la fonction  $f$  est continue en  $x$ .

La définition 5.2 peut aisément être étendue à des champs scalaires  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

En vertu de la définition de la luminance de la scène (5.1) et des propriétés 5.1, il est immédiat de constater que toute restriction de la fonction  $S$  à une droite, par exemple  $S_k$ , est continue par morceaux. Le théorème 5.1 permet finalement d'exprimer simplement la fonction  $S_k$  de luminance en 1D.

**Théorème 5.1** (décomposition d'une fonction continue par morceaux). Soit  $f : \mathcal{X} \setminus \{\mathcal{D}\} \mapsto \mathbb{R}$  une fonction continue par morceaux sur un compact  $\mathcal{X}$ . Il existe deux fonctions,  $f_c$  et  $f_d$ , telles que pour tout  $x \in \mathcal{X} \setminus \{\mathcal{D}\}$ ,  $f(x) = f_s(x) + f_c(x)$ , vérifiant  $f_c \in C^0(\mathcal{X}, \mathbb{R})$  et

$$f_s = \sum_{i=1}^{|\mathcal{D}|} u_i \mathbf{1}_{x > x_i}(x).$$

où pour tout  $i \in I$ ,  $u_i \in \mathbb{R}$ .

Il est en outre aisé de montrer l'unicité de cette décomposition (à l'ajout d'une constante près), mais cela ne présente pas d'intérêt dans le cadre des travaux présentés.

*Démonstration.* La démonstration du théorème 5.1 est donnée dans l'annexe 5.4, démonstration 5.1.  $\square$

De manière plus informelle, le théorème 5.1 stipule que toute fonction univariée continue par morceaux, peut se décomposer comme la somme d'une fonction continue (ou prolongeable par continuité) et d'une fonction constante par morceaux.

La fonction  $S_k$  (ou plus largement toute restriction de la fonction  $S$  à une droite) est continue par morceaux et, en vertu du théorème 5.1, admet donc la décomposition suivante :

$$\forall (x; y_k) \in \mathcal{X}, S_k = S_{k,c}(x) + S_{k,s}(x) = S_{k,s}(x) + \sum_{i=1}^{r_k} u_i \mathbf{1}_{x>x_i}(x) \quad (5.5)$$

où le nombre de discontinuités  $r_k = |\mathcal{D}_k|$ , comme la position  $x_i$  de ces dernières, dépend de l'ordonnée  $y_k$  considérée. La construction de cette décomposition, et notamment l'obtention de  $S_c \in \mathcal{C}^0$  sont données dans la démonstration 5.1.

La relation (5.5) permet de mesurer l'intérêt d'analyser les images par ligne (ou par colonne). La localisation des discontinuités définit, en 2D, la courbe paramétrique  $\Gamma$  difficile à modéliser avec un nombre réduit de paramètres ; à l'inverse, l'intersection de la courbe paramétrique  $\Gamma$  avec la ligne d'analyse forme un ensemble de points. Plus largement, l'analyse en 1D permet de s'affranchir de la géométrie, éventuellement complexe, des objets de la scène.

L'application du théorème 5.1 est immédiate si l'on considère un système imageur mono-dimensionnel (constitué d'un photo-détecteur linéique) de PSF spatialement invariante  $h_{\text{lin}}$  :

**Proposition 5.1.** *Soit  $f : \mathcal{X} \rightarrow \mathbb{R}$  une fonction continue par morceaux et  $\mathcal{H}$  un opérateur linéaire et invariant pas translation (LIT) de réponse impulsionnelle  $h_{\text{lin}}$  et soit  $g = \mathcal{H}[f] = f ** h_{\text{lin}}$ . En vertu du théorème 5.1 la fonction  $f$  admet la décomposition suivante :  $\forall x \in \mathcal{X}, f(x) = f_c(x) + \sum_{i=1}^l u_i \mathbf{1}_{x>x_i}(x)$ . Il en découle alors, en vertu de la linéarité de l'intégration, la décomposition suivante de la fonction  $g$  :*

$$\begin{aligned} g(x) &= f(x) ** h_{\text{lin}}(x) = f_c(x) ** h_{\text{lin}}(x) + \sum_{i=1}^l u_i \int_{-\infty}^{x-x_i} h(t) dt \\ &= g_c(x) + \sum_{i=1}^l u_i H(x - x_i), \end{aligned} \quad (5.6)$$

avec  $g_c = f_c ** h_{\text{lin}}$  et  $H(x) = \int_{-\infty}^x h(t) dt$

Dans le cas plus classique d'un système imageur bi-dimensionnel, l'application du théorème 5.1 n'est pas immédiate. En effet, l'intensité incidente sur le photo-détecteur  $\theta(x, y)$  peut-être vue, voir relation (5.4) comme le résultat de la convolution de la luminance de la scène  $S(x, y)$  en 2D par la PSF spatialement variante  $h_{\text{opt}}(u, v; \zeta_0)$  du système optique. Aussi, bien que la restriction de la luminance suivant la droite  $D_{y_k} : \{(x, y) \in \mathcal{X} / y = y_k\}$  soit une fonction univariée continue par morceaux, et puisse être décomposée suivant le théorème 5.1, deux difficultés empêchent une extension de la proposition 5.1 au cas d'un système imageur 2D :

- d'une part, la convolution en 2D engendrée par la PSF, ne permet pas de considérer la luminance d'une ligne,  $S_k(x)$ , indépendamment du reste de la scène ;
- d'autre part, la PSF  $h_{\text{opt}}$  du système imageur n'est pas spatialement invariante.

Le théorème 5.2 permet, sous certaines conditions, d'exprimer simplement le "profil", suivant une dimension, d'une discontinuité convoluée en 2D. Il est rappelé que le modèle proposé dans la section 4.2 suppose que la PSF vérifie les propriétés 4.1 d'isotropie et stationnarité locale. Par ailleurs, pour les besoins du théorème 5.2, le champ scalaire de discontinuité  $U : \partial\mathcal{X} \rightarrow \mathbb{R}$  est défini par :

$$\forall (x, y) \in \partial\mathcal{X}, U(x, y) = \lim_{\substack{(x_i, y_i) \rightarrow (x, y) \\ (x_i, y_i) \in \mathcal{X}_2}} f(x_j, y_j) - \lim_{\substack{(x_j, y_j) \rightarrow (x, y) \\ (x_j, y_j) \in \mathcal{X}_1}} f(x_j, y_j).$$

**Théorème 5.2** (Restriction en 1D d'une discontinuité convoluée en 2D). *Avec les notations précédemment introduites, soit un champ scalaire  $S : \mathcal{X} \rightarrow \mathbb{R}$  défini par  $S(x, y) = \sum_{i=1}^{N_0} S_i(x, y) \mathbf{1}_{\mathcal{X}_i}(x, y)$  et  $S_k$  la restriction de  $S$  à la droite  $D = \{(x, y) \in \mathcal{X} / y = y_k\}$  avec  $D \cup \partial\mathcal{X} = \{(x_i, y_k) \mid i = 1^{r_k}\}$ . La fonction  $S_k$  est continue par morceaux et admet donc, en vertu du théorème 5.1, la décomposition suivante :  $S_k(x) = S_{0,c}(x) + \sum_{k=1}^{r_k} u_i \mathbf{1}_{x > x_i}(x)$ . Soit, enfin, la PSF  $h_{\text{opt}} : \mathbb{R}^2 \times \mathbb{R}^+ \rightarrow \mathbb{R}$  et  $\theta : \mathbb{R}^2 \rightarrow \mathbb{R}$  défini pour tout  $(x, y) \in \mathcal{X}$  par  $\theta(x, y) = S(x, y) ** h_{\text{opt}}(x, y; \zeta_0)$ .*

*Si dans un voisinage  $\mathcal{V}_k$  de  $(x_k, y_0) \in \partial\mathcal{X}$  "suffisamment large" les conditions suivantes sont vérifiées :*

C-1 *la réponse impulsionnelle  $h_{\text{opt}}$  est isotrope et spatialement invariante dans  $\mathcal{V}_k$ , i. e.  $h_{\text{opt}}(x_k - x, y_k - y; \zeta_0) = h(\rho)$  avec  $\rho^2 = x^2 + y^2$  ;*

C-2  *$\partial\mathcal{X} \cap \mathcal{V}_k$  peut être représentée par la droite  $D_k$  formant un angle  $\psi_k$  avec  $D$  ;*

C-3 *pour tout  $(x, y) \in \mathcal{V}_k$  on a  $U(x, y) = u_k \mathbf{1}_{\partial\mathcal{X}}(x, y)$ , avec  $u \in \mathbb{R}$ .*

*alors  $\theta_k$ , la restriction de  $\theta$  à la droite  $D = \{(x, y) \in \mathcal{X} / y = y_k\}$ , admet la décomposition suivante :*

$$\theta_k(x) = \theta_{k,c}(x) + \sum_{k=1}^{r_k} u_k H\left(\frac{x - x_k}{\cos(\psi_k)}\right),$$

*avec  $\theta_{k,c} \in \mathcal{C}^0$  et  $H(x) = \int_{\mathbb{R}} \int_{-\infty}^x h(u, y) du dy$*

*Démonstration.* La démonstration du théorème 5.2, donnée dans l'annexe 5.4, est issue d'un calcul direct en utilisant un changement de variable adéquate.  $\square$

Il est notable qu'ici encore le théorème 5.2 peut-être immédiatement élargi pour le cas d'une restriction à une droite  $D$  quelconque par un changement de repère adéquate, voir la démonstration 5.2.

Le théorème 5.2 fournit de façon générale un modèle de l'intensité d'une ligne de l'image optique incidente sur le photo-détecteur  $\theta_k$  en fonction de la luminance de la scène  $S$  et de la PSF du système optique  $h_{\text{opt}}$ . La force de ce théorème réside dans le fait qu'il permet de s'affranchir d'un certain nombre de difficultés. D'une part, l'analyse de la scène en 1D permet de réduire la courbe  $\Gamma$ , représentant les arêtes des objets, à un ensemble de points. En outre, l'amplitude des discontinuités qui varie le long de la courbe  $\Gamma$  n'a alors plus de sens. D'autre part, la modélisation de la PSF spatialement variante sur le domaine  $\mathcal{X}$  est rendue simple en utilisant le fait que cette fonction peut-être localement supposée invariante. En outre, l'influence de la géométrie des discontinuités  $\psi_k$  et de la non-stationnarité de la PSF globale du système optique  $\zeta_0$  peuvent être modélisées de façon identique, par un "changement d'échelle" de la PSF.

**Corollaire 5.1** (-du théorème 5.2 : Application au cas d'un noyau Gaussien dépend d'un paramètre d'échelle). *Dans les conditions C-1, C-2 et C-3, d'application du théorème 5.2, si en outre la PSF peut-être modélisée sur le domaine  $\mathcal{X}$  par l'expression 4.17, alors  $g_0$ , la restriction de  $g$  à la droite  $D = \{(x, y) \in \mathcal{X} / y = y_0\}$ , admet la décomposition suivante :*

$$g_0(x) = g_{0,c}(x) + \sum_{k=1}^{K_0} u_k \Phi\left(\frac{x - x_k}{\zeta_0 \cos(\psi_k)}\right),$$

*avec  $g_{0,c} \in \mathcal{C}^0$ ,  $\eta_k = \zeta_{x_k, y_k} \psi_k$  et  $\Phi$  représente la fonction de distribution Gaussienne (unidimensionnelle) dont l'expression est donnée par  $\Phi(x) = \int_{-\infty}^x \phi(-u^2/2) du$ .*

Il découle donc finalement du théorème (5.2) et de son corollaire (5.1) que l'énergie  $\theta_k(x)$  intégrée sur la  $k$ -ième ligne des pixels du photo-détecteur peut s'exprimer :

$$\forall (x, y_k) \in \mathcal{X}, \theta_k(x) = \theta_{k,c}(x) + \theta_{k,s}(x) = \theta_{k,c}(x) + \sum_{i=1}^{r_k} u_i \Phi\left(\frac{x - x_{k,i}}{\zeta_{k,i}}\right) \quad (5.7)$$

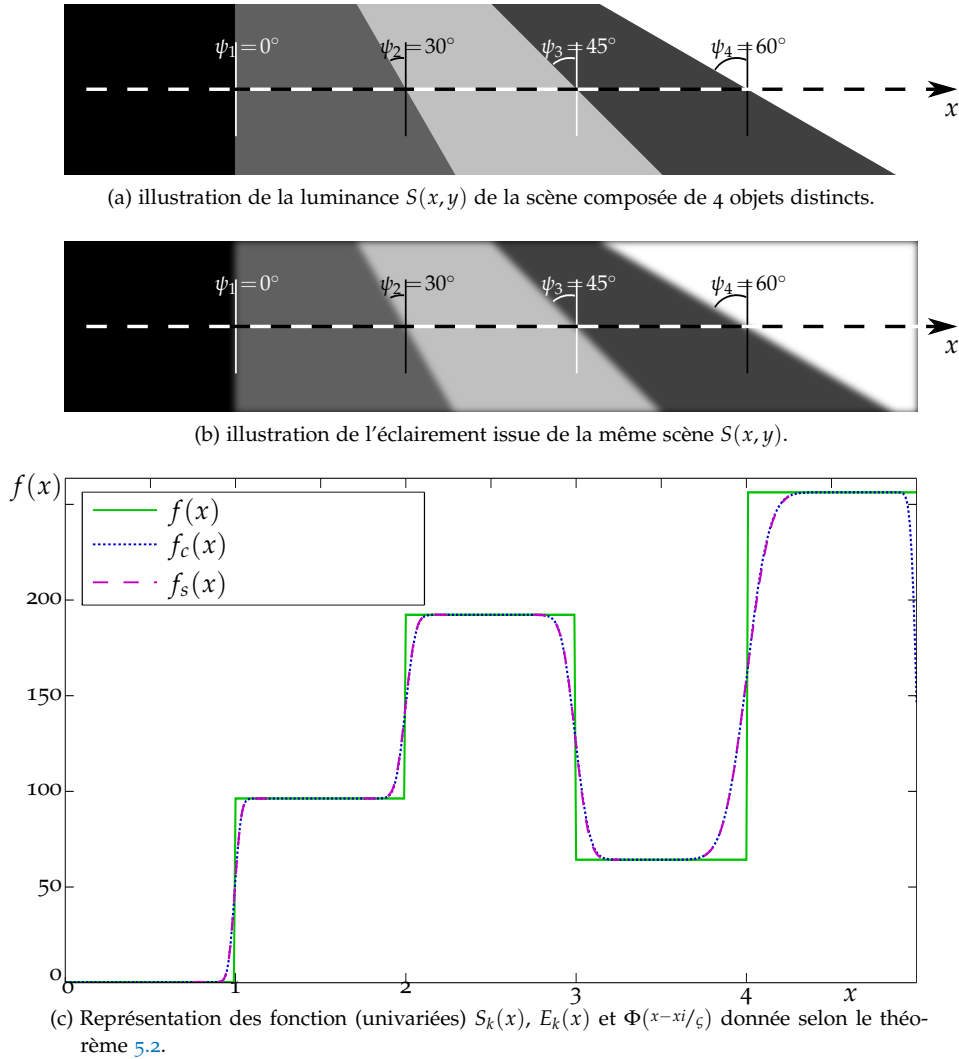


FIGURE 5.4: Illustration du théorème 5.1 :

Dans la suite de ce manuscrit, l'image n'est pas analysée ligne par ligne mais segment par segment, car la dimension des lignes est trop importante pour permettre une modélisation précise de façon simple. En revanche, chaque segment est extrait d'une ligne afin que le modèle 5.7 ne soit en rien modifié. Une image  $\mathbf{Z}$  analysée est donc typiquement représentée par  $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_N)$  où pour tout  $k \in \{1, \dots, N\}$ ,  $\mathbf{z}_k = (z_{k,1}, z_{k,L})^T$ .

Il est en outre considéré, sans perte de généralité, que l'ensemble  $\mathcal{X}$  de définition de l'image (et de la scène) est choisi de sorte que les pixels soient des carrés de côté unitaire. De la même façon, la fonction  $\theta_k$  représentant l'énergie lumineuse incidente sur chacun des segments est supposée définie sur le domaine  $\mathcal{X}_k = [-L/2; L/2]$ .

## 5.2 Nécessité de la modélisation non-linéaire des discontinuités

La modélisation qui a été proposée de l'énergie incidente sur une ligne des pixels du photo-détecteur n'est pas, telle que formulée dans (5.7), directement exploitable. En effet, la décomposition  $\theta_k(x) = \theta_{k,c}(x) + \theta_{k,s}(x)$  n'est qu'une étape préalable et il demeure nécessaire de modéliser les fonctions continues et singulières (resp.  $\theta_{k,c}$  et

$\theta_{k,s}$ ).

L'objectif de cette section 5.2 est de proposer un modèle paramétrique de la fonction  $\theta_k$  représentant l'énergie incidente sur une ligne de l'image. Pour cela, les fonctions  $\theta_{k,c}$  et  $\theta_{k,s}$  sont considérées distinctement :

- La fonction  $\theta_{k,c}$  représente la partie continue de la scène convoluée par la PSF ; la régularité supposée de cette fonction laisse penser qu'un modélisation paramétrique sous une forme linéaire est pertinente.
- À l'inverse, la fonction  $\theta_{k,s}$  représentant les frontières entre les différents objets de la scène est constituée de discontinuités convoluées ; proposer une modélisation générale de cette fonction  $\theta_{k,s}$  sous une forme linéaire quelle que soit la position des discontinuités ou le paramètre local de la PSF  $\zeta_0$  semble difficile. En outre, ces discontinuités sont d'une importance prépondérante dans la modélisation d'une image.

Dans le cadre des travaux présentés, la partie continue  $\theta_{k,c}$  de l'image, est localement modélisée par un polynôme ; ce choix d'un modèle polynomial est fondé sur la théorie "classique" de l'approximation, voir annexe B et [146, 147] pour plus d'informations. En effet, certains résultats permettent de conjecturer que l'erreur due à la modélisation de  $\theta_{k,c}$  par un polynôme est efficace au sens où avec un polynôme de degré faible (relativement au nombre de pixels) l'erreur d'approximation peut être uniformément bornée par une constante (très petite).

Plus formellement, en vertu du théorème B.3 de Jackson et de son corollaire B.2 l'erreur d'approximation de la fonction  $\theta_{k,c}$  par un polynôme de degré  $n$  vérifie :

$$\mathcal{E}_n^*(\theta_{k,c}) = \min_{p_n \in \mathbb{R}_n[X]} \sup_{x \in \mathcal{X}} \|\theta_{k,c}(x) - p_n(x)\| \leq \frac{6^2}{n(n-1)} \omega \left( (n-1)^{-1}; \frac{d}{dx} h_{\zeta_{k,i}}(x) * S_{k,c}(x) \right), \quad (5.8)$$

où  $\omega(\delta; f)$  représente le module de continuité de la fonction  $f$  défini par :

$$\omega(\delta; f) = \sup_{(x,y) \in \mathcal{X}^2 / |x-y| \leq \delta} \|f(x) - f(y)\|.$$

Il ne semble pas possible de fournir une expression analytique simple de la borne donnée dans la relation (5.8). En revanche, il peut être conjecturé que le module de continuité de  $\omega \left( (n-1)^{-1}; \frac{d}{dx} h_{\zeta_{k,i}}(x) * S_{k,c}(x) \right)$  est faible compte-tenu des propriétés de régularité de la fonction  $S_{k,c}(x)$  et compte-tenu du fait que  $\frac{d}{dx} h_{\zeta_{k,i}}(x)$  agit comme un filtre passe-haut.

À l'inverse cette l'approche consistant à utiliser la théorie "classique" de l'approximation fournit quelques éléments permettant de conjecturer qu'une modélisation de la fonction singulière  $\theta_{k,s}$  par un polynôme de degré faible (relativement au nombre de pixels) et avec une erreur d'approximation uniforme faible n'est pas envisageable. La figure 5.5 illustre ce fait en présentant l'erreur d'approximation définie par :

$$\mathcal{E}_n^*(\theta_{k,s}) = \min_{p_n \in \mathbb{R}_n[X]} \sup_{x \in \mathcal{X}} \|\theta_{k,s}(x) - p_n(x)\|$$

obtenue en approchant la fonction  $\theta_{k,s}(x) = \Phi(x)$  sur un intervalle  $[-a; a]$  par des polynômes de degré impairs  $n = \{1; 3; 5; 7; 9\}$ . Ces résultats sont obtenus numériquement en utilisant l'algorithme de Remez [148].

### 5.2.1 Modélisation locale retenue du contenu des images

Compte-tenu du rôle prépondérant des discontinuités dans une image, et de la difficulté soulevée par la modélisation de  $\theta_{k,s}$  (quelle que soit la position des discontinuités et le paramètre local de la PSF  $\zeta_{k,s}$ ), il est proposé de conserver le modèle de la par-

*Dans théorie "classique" de l'approximation les polynômes utilisés sont des polynômes algébriques ou trigonométrique; les résultats obtenus sont similaires dans tous les cas.*

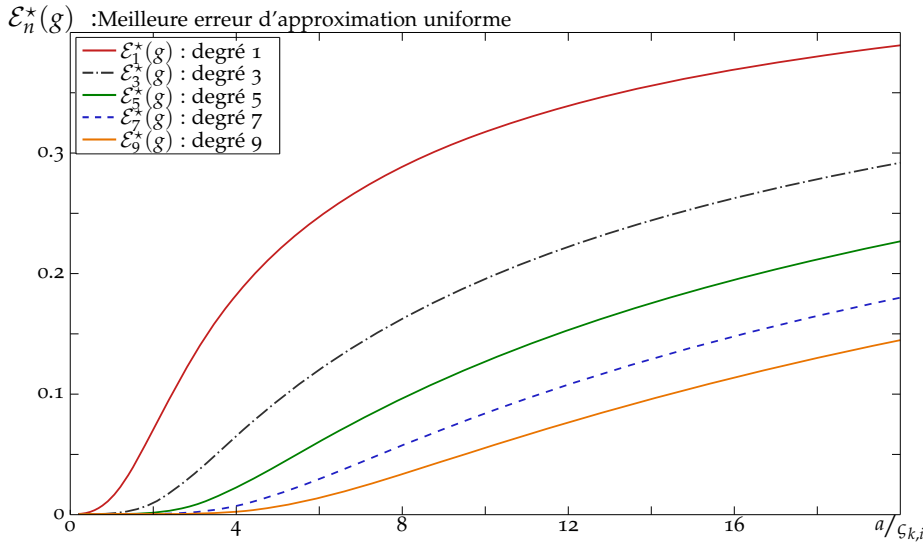


FIGURE 5.5: Valeur approchée de la meilleure erreur d'approximation  $\mathcal{E}_n^*(\theta_{k,s})$  obtenue numériquement par l'algorithme de Remez avec  $\theta_{k,s}$  définie sur l'intervalle  $[-a; a]$  par  $\theta_{k,s} = \Phi(x/\zeta_{k,i})$ ; les résultats sont ici présentés en fonction du rapport  $a/\zeta_{k,i}$ .

tie discontinue présentée dans (5.7) La modélisation retenue de l'énergie lumineuse incidente sur le domaine  $\mathcal{X}_k$  du  $k$ -ième "segment" est donnée par :

$$\forall x \in \mathcal{X}_k, \theta_k(x) = \theta_{k,c}(x) + \theta_{k,s}(x) = \sum_{j=0}^n c_{k,j} x^j + \sum_{i=1}^{r_k} u_{k,i} \Phi\left(\frac{x - x_{k,i}}{\zeta_{k,i}}\right) \quad (5.9)$$

Rappelons que dans le cas d'un système imageur, le signal discret est obtenu par intégration spatiale de l'énergie lumineuse incidente sur la surface des éléments photosensibles (les pixels). Aussi, le signal photo-électrique enregistré correspondant au  $k$ -ième segment,  $\theta_k = (\theta_{k,1}, \dots, \theta_{k,L})^T$  est donné par :

$$\theta_{k,l} = \theta_k(x) * h_{\text{int}}(x) \Big|_{x=x_l} dx,$$

où le noyau de convolution  $h_{\text{int}}$  est donné en 1D pour des pixels carrés de côtés unitaires, voir relation (4.10), par :  $h_{\text{int}}(x) = \mathbf{1}_{[-1/2; 1/2]}(x)$ . Il est immédiat de montrer que la convolution d'une fonction polynomiale demeure un polynôme. Ainsi, en notant le paramètre (vectoriel) de non-linéarité par  $\eta_k = (x_{k,1}, \zeta_{k,1}, \dots, x_{k,r_k}, \zeta_{k,r_k})$ , il en découle que le signal photo-électrique  $\mathbf{z}_k$  peut être modélisé par :

$$\theta_k = \mathbf{H}\mathbf{c}_k + \mathbf{F}(\eta_k)\mathbf{u}_k, \quad (5.10)$$

où

$$\mathbf{H} = \begin{pmatrix} x_1^0 & x_1^1 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \vdots \\ x_l^0 & x_l^1 & \dots & x_l^n \\ \vdots & \vdots & \vdots & \vdots \\ x_L^0 & x_L^1 & \dots & x_L^n \end{pmatrix}, \quad \mathbf{F}(\eta_k) = \begin{pmatrix} f(x_1; x_{k,1}, \zeta_{k,1}) & \dots & f(x_1; x_{k,i}, \zeta_{k,i}) & \dots & f(x_1; x_{k,r_k}, \zeta_{k,r_k}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f(x_l; x_{k,1}, \zeta_{k,1}) & \dots & f(x_l; x_{k,i}, \zeta_{k,i}) & \dots & f(x_l; x_{k,r_k}, \zeta_{k,r_k}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f(x_L; x_{k,1}, \zeta_{k,1}) & \dots & f(x_L; x_{k,i}, \zeta_{k,i}) & \dots & f(x_L; x_{k,r_k}, \zeta_{k,r_k}) \end{pmatrix}$$

et les vecteurs inconnus  $\mathbf{c}_k$  et  $\mathbf{u}_k$  sont les coefficients  $\mathbf{c}_k = (c_{k,1}, \dots, c_{k,n})^T$  et  $\mathbf{u}_k = (u_{k,1}, \dots, u_{k,r_k})^T$ .

Chacun des vecteurs  $\mathbf{f}$ , colonne de la matrice  $\mathbf{F}(\eta_k)$ , représente une des  $r_k$  discontinuités et a pour expression :

$$f(x_l; x_{k,i}, \zeta_{k,i}) = \int_{x_k-1/2}^{x_k+1/2} \Phi\left(\frac{x - x_{k,i}}{\zeta_{k,i}}\right) dx.$$



Le modèle du signal photo-électrique (5.10) peut alternativement être représenté sous la forme :

$$\begin{aligned} \boldsymbol{\theta}_k &= \mathbf{G}(\boldsymbol{\eta}_k) \mathbf{v}_k, \\ \text{où } \mathbf{G}(\boldsymbol{\eta}_k) &= \left( \mathbf{H} \mid \mathbf{F}(\boldsymbol{\eta}_k) \right) \text{ et } \mathbf{v}_k = \left( \mathbf{c}_k, \mathbf{u}_k \right)^T. \end{aligned} \quad (5.11)$$

Néanmoins, la formulation (5.10) permet plus aisément que (5.11) de mettre en évidence que le modèle de  $\boldsymbol{\theta}_k$  est composé d'une partie linéaire  $\mathbf{H}\mathbf{c}_k$ , représentant la partie continue de l'énergie incidente sur les segments, et une partie non-linéaire  $\mathbf{F}(\boldsymbol{\eta}_k)\mathbf{u}_k$  représentant les éventuelles discontinuités de la luminance de la scène sur le domaine  $\mathcal{X}_k$ . La présence dans le modèle (5.10) - (5.11) du paramètre  $\boldsymbol{\eta}_k$  intervenant de façon non-linéaire est un problème délicat. Il faut toutefois souligner que cette non-linéarité résulte des propriétés physiques de la scène imagée et demeure incontournable pour permettre l'obtention d'un modèle paramétrique représentant fidèlement les images. Les difficultés soulevées par la présence des paramètres non-linéaires  $x_{k,i}$  et  $\zeta_{k,i}$  dans le cadre de la théorie de l'estimation statistique sont abordés dans la section 6.1. Enfin, une méthodologie permettant d'estimer simplement et précisément  $\theta(x)$  lorsque les paramètres  $x_{k,i}$  et  $\zeta_{k,i}$  ne sont pas connus est proposée dans la section 6.2.

### 5.3 Discussions sur quelques modèles alternatifs des images

L'originalité de l'approche présentée pour modéliser les images naturelles réside dans la prise en compte des phénomènes physiques permettant la formation d'une image optique incidente sur le photo-détecteur. Cette approche des images permet de proposer un modèle local paramétrique et non-linéaire des images naturelles (5.10) - (5.11). Le lecteur attentif comprendra que le modèle proposé n'a pas vocation à l'optimalité, en un quelconque sens, mais de permettre la conception d'un test pour détecter des informations cachées tout en respectant la contrainte sur la probabilité de fausse alarme. Dans ce cadre applicatif, la présente section 5.3 souligne l'intérêt du modèle proposé au travers d'une brève description des inconvénients que présente l'utilisation de quelques modèles usuels du contenu des images ; compte-tenu du nombre de modèles d'images rencontrés dans la littérature, cette section 5.3 n'a pas vocation à présenter un état de l'art exhaustif.

#### 5.3.1 Approximation par transformation en ondelettes

Le problème de l'approximation par reconstruction partielle de transformée en ondelettes, proche de la méthodologie proposée, peut se résumer ainsi ; soit une fonction  $f \in \mathcal{L}^2[0; 1]$  inconnue dont les échantillons  $s(t_i)$ , du signal  $\mathbf{s} = \{s_i\}_{i=0}^{N-1}$ , sont donnés par la relation :

$$s_i = f(t_i) + \sigma \xi_i \quad (5.12)$$

où, pour tout  $i \in \{0, \dots, N-1\}$ ,  $t_i = i/N$  et les  $\{\xi_i\}_{i=0}^{N-1}$ , représentant le bruit présent dans le signal  $s$ , sont les réalisations de variables aléatoires indépendantes et identiquement distribuées suivant une loi de probabilité Gaussienne, ce qui est noté par abus de langage,  $\boldsymbol{\xi} \sim \mathcal{N}(0, \mathbf{I}_N)$ .

L'approche proposée dans [149, 150, 151] consiste à obtenir une estimation  $\hat{f}$  de  $f$  en projetant  $\mathbf{s}$  sur une base orthonormale d'ondelettes discrétisées à la résolution  $2^j$ , représentée par une matrice unitaire  $\mathbf{W}_{2^j}$ , voir [152, 153] pour une présentation formelle de la transformation en ondelettes continues. Le résultat de cette projection est un vecteur noté  $\mathbf{c} = \{c_i\}_{i=0}^{N-1}$ , de coefficients de transformée en ondelettes, qui peut être représenté par :

$$\mathbf{c} = \mathbf{W}_{2^j} \mathbf{s} = \mathbf{d} + \sigma \boldsymbol{\epsilon}$$

où

$$\mathbf{d} = \mathbf{W}_{2j} \mathbf{f} \quad \text{et} \quad \sigma \boldsymbol{\epsilon} = \sigma \mathbf{W}_{2j} \boldsymbol{\zeta}.$$

Comme tenu du fait que la matrice  $\mathbf{W}_{2j}$  est orthonormale, le vecteur  $\boldsymbol{\epsilon}$  est, comme  $\boldsymbol{\zeta}$ , un vecteur dont les composantes sont des variables aléatoires indépendantes et identiquement distribuées suivant une loi de probabilité Gaussienne [149, 150], ce qui est noté par abus de langage,  $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I}_N)$ .

En notant  $\hat{\mathbf{d}} = \{\delta_\lambda(c_i)\}_{i=0}^{N-1}$  une approximation quelconque du vecteur de coefficients  $\mathbf{c}$ , il est immédiat de montrer qu'en vertu de l'orthonormalité de  $\mathbf{W}_{2j}$ , l'approximation  $\hat{\mathbf{f}}$  de  $\mathbf{f}$  qui découle de l'utilisation de  $\hat{\mathbf{d}}$  est donnée par :

$$\hat{\mathbf{f}} = \mathbf{W}_{2j}^T \hat{\mathbf{d}}.$$

L'erreur quadratique moyenne (EQM), également appelée risque, issue de l'approximation de  $\mathbf{f}$  par  $\hat{\mathbf{f}}$  est définie par :

$$R_{N,\sigma}(\hat{\mathbf{f}}, \mathbf{f}) = N^{-1} \|\mathbf{f} - \hat{\mathbf{f}}\|_2^2 = N^{-1} \|\mathbf{d} - \hat{\mathbf{d}}\|_2^2. \quad (5.13)$$

La seconde égalité de (5.13) est immédiate à démontrer, en vertu du théorème de Plancherel, en utilisant ici encore l'orthonormalité de la matrice  $\mathbf{W}_{2j}$ .

Dans la suite de cette section, la fonction d'approximation des coefficients  $\delta_\lambda$  considérée est la fonction de *shrinkage* par seuillage doux, très appréciée pour ses propriétés de lissage et d'adaptativité [150, 151], formellement définie par :

$$\delta_\lambda(c_i) = \begin{cases} x - \text{signe}(x)\lambda & \text{si } |x| \geq \lambda \\ 0 & \text{si } |x| < \lambda. \end{cases} \quad (5.14)$$

L'intérêt de cette fonction de seuillage doux (5.14) réside dans l'obtention de résultats sur l'erreur d'approximation en terme de majoration du risque (5.13). Le seuil universel est par exemple issu d'une analyse de la contribution du bruit  $\boldsymbol{\epsilon}$  dans les coefficients de la transformée en ondelettes discrètes  $\mathbf{c}$ . Dans le cadre du signal défini par 5.12, la contribution du bruit est  $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \mathbf{I}_N)$  ce qui permet de montrer (voir [153, chap.9]) que :

$$\lim_{N \rightarrow \infty} \mathbb{P} \left[ \lambda_u(N) - \sigma \frac{\ln \ln N}{\ln N} \leq \max \left( \{|\epsilon_n|\}_{n=0}^{N-1} \right) \leq \lambda_u(N) \right] = 1.$$

où  $\lambda_u(N) = \sigma \sqrt{2 \ln N}$  est le seuil dit universel. Autrement dit, le maximum en valeur absolue de la contribution du bruit dans les coefficients de transformée en ondelettes discrètes  $\max \left( \{|\epsilon_n|\}_{n=0}^{N-1} \right)$  a une grande probabilité d'être proche, mais inférieur, au seuil universel  $\lambda_u(N)$  lorsque  $N \rightarrow \infty$ . Le théorème 5.3, d'après [149, théorème 1], offre en conséquence une majoration du risque (ou de l'EQM) d'approximation de  $\mathbf{f}$  par  $\hat{\mathbf{f}}$  par seuillage doux associé au seuil universel :

**Théorème 5.3** (Majoration du risque  $R_{\lambda_u(N)}$  [149]). *Suivant le modèle (5.12), soit donné l'approximation  $\hat{\mathbf{f}}$  de  $\mathbf{f}$  donné par  $\mathbf{W}_{2j} \delta_{\lambda_u(N)}(\mathbf{c})$  où  $\delta_{\lambda_u(N)}$  représente la fonction de seuillage doux associée au seuil universel  $\lambda_u(N)$ , alors :*

$$R_{\lambda_u(N)}(\hat{\mathbf{f}}, \mathbf{f}) \leq (1 + 2 \ln N) (N^{-1} \sigma^2 + R_0(\mathbf{f})),$$

où  $R_0(\mathbf{f}) = \sum_{n=0}^{N-1} \min(d_i^2; \epsilon^2)$ .

Deux autres résultats intéressants ont été obtenus pour le problème dit de débruitage de la fonction  $f$  ou du signal  $\mathbf{f}$ , formellement défini dans [150]; ce problème consiste à obtenir une valeur approchée  $\hat{f}$  de  $f$  (et donc une approximation  $\hat{\mathbf{f}}$  de  $\mathbf{f}$ ) sous les deux contraintes suivantes :

*La fonction  $\delta$  appliquée indépendamment aux composantes du vecteur  $\mathbf{c}$  de coefficients est volontairement pas précisé présentement.*

*Un résultat similaire au théorème 5.3 existe dans le cas du seuil dit minimax. Le lecteur intéressé peut consulter [149].*

- 1- l'erreur quadratique moyenne issue de l'approximation de  $\mathbf{f}$  par  $\hat{\mathbf{f}}$  est optimale (en un certain sens qui reste à définir);
- 2- l'approximation  $\hat{f}$  de la fonction  $f$  est, avec une grande probabilité, au moins aussi régulière que  $f$ .

Plus formellement, considérant que la fonction  $f \in \mathcal{F} \in \mathcal{L}^2(\mathbb{R})$  où  $\mathcal{F}$  est un espace de fonctions régulières (espaces de Besov par exemple), la première contrainte sur l'optimalité de l'approximation est vérifiée par le théorème 5.4, d'après [150, théorème 1.2].

**Théorème 5.4** (approximation quasi minimax de  $\mathbf{f}$  [150]). *Suivant le modèle (5.12), soit donné l'approximation  $\hat{\mathbf{f}}$  de  $\mathbf{f}$  donné par  $\mathbf{W}_{2^j} \delta_{\lambda_u(N)}(\mathbf{c})$  où  $\delta_{\lambda_u(N)}$  représente la fonction de seuillage doux associée au seuil universel  $\lambda_u(N)$ , alors il existe une constante  $C_1 = C_1(\mathcal{F}, \phi) > 0$  telle que :*

$$\sup_{f \in \mathcal{F}} R_{\lambda_u(N)}(\hat{\mathbf{f}}, \mathbf{f}) \leq C_1 \ln N \inf_{\hat{\mathbf{f}}} \sup_{f \in \mathcal{F}} R_{\lambda_u(N)}(\hat{\mathbf{f}}, \mathbf{f})$$

La notation  $C_1 = C_1(\mathcal{F}, \phi)$  indique ici le fait que cette constante dépendant uniquement de l'ondelette  $\phi$  et de l'espace de fonctions considérée  $\mathcal{F}$  mais ne dépend pas de  $f$  ni de  $N = 2^j$ .

De manière analogue, la seconde contrainte sur la régularité de l'approximation obtenue est formellement définie et vérifiée par le théorème 5.5, d'après [150, théorème 1.1].

**Théorème 5.5** (Régularité de l'approximation  $\hat{f}$  [150]). *Suivant le modèle (5.12), soit l'approximation  $\hat{f}$  de la fonction  $f$  basée sur le seuillage doux associé au seuil universel :  $\mathbf{W}_{2^j} \delta_{\lambda_u(N)}(\mathbf{c})$ . Alors, il existe une suite  $\pi_n \rightarrow 1$  lorsque  $N = 2^j \rightarrow \infty$  et une constante  $C_2 = C_2(\mathcal{F}, \phi) > 0$  telles que :*

$$\mathbb{P} \left[ \|\hat{f}_n\|_{\mathcal{F}} \leq C_1 \|f\|_{\mathcal{F}} \right] \geq \pi_n \stackrel{\text{def.}}{=} \mathbb{P} \left[ \max \left( \{|\epsilon_n|\}_{n=0}^{N-1} \right) \leq \sqrt{2 \ln N} \right].$$

La méthode d'interpolation permettant l'obtention de  $\hat{f}$  découle de la définition même de la définition de transformation en ondelettes continues et est détaillées dans [150, sec.V].

Ici encore  $C_2 = C_2(\mathcal{F}, \phi) > 0$  indique que  $C_2$  dépend de  $\phi$  et de  $\mathcal{F}$  mais ne dépend pas de  $f$  ni de  $N = 2^j$ .

Si ces résultats sont incontestablement intéressants, il faut noter toutefois que par rapport au modèle proposé, l'approximation par transformation en ondelettes ne tient nullement compte de la structure particulière des images naturelles (le modèle des ondelettes se voulant le plus générique possible). En outre, le problème de la détermination du seuil, longuement discuté dans [149, 150, 151], semble d'autant plus difficile dans le cadre des travaux présentés puisque le but est *in fine* de pouvoir concevoir un test dont les performances statistiques sont connues. Dans ce cadre l'utilisation de fonction de seuillage complexifie grandement le problème.

### 5.3.2 Apprentissage de dictionnaire parcimonieux

Comme cela a été évoqué dans la section 5.3.1, la décomposition en ondelettes est parcimonieuse au sens de la norme  $\ell^2$ , *i.e.* l'énergie du signal est concentrée dans un nombre limité de coefficients issus de la transformée en ondelettes. La transformée en ondelettes est rendue parcimonieuse au sens de la norme  $\ell^0$  par l'application d'une fonction de seuillage (5.14), voir également [154] pour une discussion sur les fonctions de seuillage.

L'apprentissage de dictionnaire est une approche non-paramétrique alternative pour l'approximation d'un signal inconnu. L'originalité de cette méthodologie réside dans le fait que c'est la base de représentation elle-même qui est adaptée au signal en vue d'obtenir une représentation parcimonieuse au sens de la norme  $\ell^0$ . Plus précisément, l'obtention du dictionnaire  $\mathbf{D}_0$  permettant une représentation exacte la plus parcimonieuse du signal  $\boldsymbol{\theta} \in \mathbb{R}^N$  (au sens de la norme  $\ell^0$ ) est formellement donnée, de manière générale, par la résolution de l'équation (5.15) :

$$\inf_{\mathbf{x} \in \mathbb{R}^N} \{ \|\mathbf{x}\|_0 / \boldsymbol{\theta} = \mathbf{D}_0 \mathbf{x} \}, \quad (5.15)$$

Les termes de dictionnaire est d'atome utilisés dans le domaine peuvent être resp. interprétés, du point de vue algébrique, comme une famille génératrice et un vecteur de cette famille.

De manière informelle la norme  $\ell_0$  consiste à compter le nombre de composantes non-nulles ce qui justifie sont intérêt dans la recherche de représentation parcimonieuse.

où  $\|\cdot\|_0$  représente la norme  $\ell^0$  définie par :  $\|\mathbf{x}\|_0 = \sum_{n=1}^N x_n^0 = \sum_{n=1}^N \mathbf{1}_{\mathbb{R}^*}(x_n)$  avec la convention  $\forall x \in \mathbb{R}, x^0 = \mathbf{1}_{\mathbb{R}^*}(x_n)$ .

De manière analogue, la contrainte de représentation exacte peut être relaxée; l'obtention de la base  $\mathbf{D}_{0,\epsilon} \in \mathcal{M}_{N,N}$  de représentation à  $\epsilon$ -près la plus parcimonieuse du signal  $\boldsymbol{\theta} \in \mathbb{R}^N$  est donnée par la résolution de l'équation (5.16) :

$$\inf_{\mathbf{x} \in \mathbb{R}^N} \left\{ \|\mathbf{x}\|_0 / \|\boldsymbol{\theta} - \mathbf{D}_{0,\epsilon} \mathbf{x}\|_2 \leq \epsilon \right\}. \quad (5.16)$$

La principale difficulté à la détermination de la base de représentation la plus parcimonieuse d'un signal, est de nature calculatoire; il est en effet possible de démontrer [155] que la résolution de l'une des équations (5.15) et (5.16) est un problème NP-difficile.

De nombreuses méthodes disponibles dans la littérature permettent une résolution approchée des équations (5.15) et (5.16). Ces algorithmes utilisent un dictionnaire  $\mathbf{D}$  prédéfini pour construire le dictionnaire  $\hat{\mathbf{D}}_0$  en sélectionnant séquentiellement les atomes parmi ceux de  $\mathbf{D}$ . Les algorithmes de *Matching Pursuit* (MP) [156] et *Orthogonal Matching Pursuit* (OMP) [157] sont de cette catégorie, la sélection des atomes implique de manière itérative un produit scalaire entre le signal et les atomes et, usuellement, la résolution de problème de moindres carrés.

La méthode de *Basis pursuit* (BP) [158] propose de rendre le problème convexe en substituant la norme  $\ell^1$  à la norme  $\ell^0$ ; le problème peut alors être considéré comme une minimisation au sens des moindres carrés avec un terme de régularisation  $\ell^1$ .

Certains algorithmes disponibles dans la littérature, notamment le K-SVD [159], autorisent une modification des atomes du dictionnaire fonctionnant, de façon générale, en deux étapes répétées itérativement : sélection puis mise-à-jour des atomes du dictionnaire. Par ailleurs, la sélection des atomes du dictionnaire permet d'envisager la représentation par un dictionnaire sur-complet *overcomplete dictionary*. Le principe consiste alors à utiliser un dictionnaire contenant un nombre d'atomes supérieur au nombre d'échantillons du signal analysé; la sélection des atomes permettant dans tous les cas l'obtention *in fine* d'une représentation parcimonieuse.

## Conclusions du chapitre :

Dans la lignée directe du chapitre 4 précédent, ce chapitre est dévolu à la modélisation des images naturelles en utilisant les propriétés physiques du système d'acquisition et de la scène imagée. Contrairement aux modèles d'images "adaptatifs" usuellement rencontrés dans la littérature, l'originalité du modèle proposé réside dans son fondement physique qui le rend, par nature, adapté aux propriétés que présente les images naturelles; ces dernières présentent notamment un contenu structuré caractérisé par la présence de discontinuités convoluées par le système optique d'acquisition qui agit localement comme un filtre.

Sans perdre de vue le cadre applicatif de la détection d'informations cachées, il est proposé de modéliser les images suivant une dimension. Cette approche permet de décrire localement le contenu d'une image en utilisant un nombre limité de paramètres. Compte tenu de la nature continue par morceaux de la scène imagée, ce modèle local paramétrique est non-linéaire vis-à-vis des paramètres décrivant physiquement les discontinuités. Enfin, l'utilisation de ce modèle ouvre la voie à la conception d'un test dont les performances statistiques sont explicitement données dans le chapitre 6 suivant, ce qui permet *in fine* de respecter une contrainte sur la probabilité de fausses alarmes.

## 5.4 Annexes du chapitre 5 : démonstrations des théorèmes 5.1 et 5.2

La démonstration des théorèmes 5.1 et 5.2 est réalisée, dans un but pédagogique, en trois étapes. Dans un premier temps, le théorème 5.1, sur la décomposition d'une fonction (univariée) continue par morceaux, est démontré. La proposition 5.2 permet une extension du théorème aux fonctions de deux variables. Une attention particulière est portée aux conditions sous lesquelles cette extension est possible. Enfin, sous des conditions similaires à celles exigées dans la proposition 5.2, le théorème 5.2 est démontré.

**Démonstration 5.1.** Soit l'ensemble  $\mathcal{D} = \bigcup_{i \in I \subseteq \mathbb{N}} x_i$  avec pour tout  $i \in I$ ,  $x_i < x_{i+1}$  et soit  $f : \mathcal{X} \setminus \{\mathcal{D}\} \mapsto \mathbb{R}$  une fonction continue par morceaux sur un compact  $\mathcal{X} = [x_{\min}; x_{\max}]$ . Pour tout  $i \in I$  le compact ouvert  $\mathcal{Y}_i$  est défini par  $\mathcal{Y}_i = [x_{\min}; x_{i+1}[$  et la fonction  $f_i : \mathcal{Y}_i \rightarrow \mathbb{R}$  est définie pour tout  $x \in \mathcal{Y}_i$  par  $f_i(x) = f(x)$ .

Il est notable que :  
 $\mathcal{Y}_i = \mathcal{Y}_{i-1} \cup [x_i; x_{i+1}[$ .

Il est proposé de démontrer le théorème 5.1 par récurrence. Pour cela le cas de la fonction  $f_1$ , n'admettant donc qu'une unique discontinuité de premier ordre, est considéré dans un premier temps. En utilisant la définition 5.1 il est possible de définir deux fonctions  $f_{1,c} : \mathcal{Y}_1 \mapsto \mathbb{R}$  et  $f_{1,s} : \mathcal{Y}_1 \mapsto \mathbb{R}$  comme suit :

$$f_{1,c}(x) = \begin{cases} f(x) & \text{si } x < x_1 \\ f(x) + \left( \lim_{x \rightarrow x_1^-} f(x) - \lim_{x \rightarrow x_1^+} f(x) \right) & \text{si } x > x_1 \end{cases}$$

$$\text{et } f_{1,s}(x) = \begin{cases} 0 & \text{si } x < x_1 \\ \lim_{x \rightarrow x_1^+} f(x) - \lim_{x \rightarrow x_1^-} f(x) & \text{si } x > x_1 \end{cases}$$

La fonction  $f_1$  admet, pour tout  $x \in \mathcal{Y}_1$ , la décomposition  $f_1(x) = f_{1,c}(x) + f_{1,s}(x)$  avec  $f_{1,c} \in \mathcal{C}^0(\mathcal{Y}_1, \mathbb{R})$  et, en posant  $u_1 = \lim_{x \rightarrow x_1^+} f(x) - \lim_{x \rightarrow x_1^-} f(x)$  on a  $f_{1,s} = u_1 \mathbf{1}_{x > x_1}(x)$ . La démonstration par récurrence se poursuit en considérant que  $i - 1 \in I$  la fonction  $f_{i-1}$  admet, pour tout  $x \in \mathcal{Y}_{i-1}$ , la décomposition  $f_{i-1}(x) = f_{i-1,c}(x) + f_{i-1,s}(x)$  avec  $f_{i-1,c} \in \mathcal{C}^0(\mathcal{Y}_{i-1}, \mathbb{R})$  et  $f_{i-1,s} = \sum_{k=1}^{i-1} u_k \mathbf{1}_{x > x_k}(x)$ . Le cas de la fonction  $f_i$  défini sur l'ensemble  $\mathcal{Y}_i$  est traité en posant :

$$f_{i,c}(x) = \begin{cases} f_{i-1}(x) & \text{si } x < x_i \\ f(x) - \sum_{k=0}^{i-1} u_k + \left( \lim_{x \rightarrow x_i^-} f(x) - \lim_{x \rightarrow x_i^+} f(x) \right) & \text{si } x > x_i \end{cases}$$

$$\text{et } f_{i,s}(x) = \begin{cases} f_{i-1,s} & \text{si } x < x_i \\ f_{i-1,s} + \lim_{x \rightarrow x_i^+} f(x) - \lim_{x \rightarrow x_i^-} f(x) & \text{si } x > x_i \end{cases}$$

La décomposition étant admise au rang  $i - 1$  on a, pour tout  $x < x_i$  on a  $f(x) = f_{i,c}(x) + f_{i,s}(x)$ . En posant  $u_i = \lim_{x \rightarrow x_i^+} f(x) - \lim_{x \rightarrow x_i^-} f(x)$  il est immédiat de vérifier  $f_{i,c} \in \mathcal{C}^0(\mathcal{Y}_i, \mathbb{R})$  et que  $f_{i,s}(x) = f_{i-1,s}(x) + u_i \mathbf{1}_{x > x_i}(x) = \sum_{k=1}^i u_k \mathbf{1}_{x > x_k}(x)$ . En outre, un rapide calcul permet de montrer que pour tout  $x \in \mathcal{Y}_i \mid x > x_i$  on a  $f_{i,c}(x) + f_{i,s}(x) = f(x) - \sum_{k=0}^i u_k + f_{i,s}(x) = f(x)$ . La propriété est donc vérifiée au rang  $i$  ce qui clôt la démonstration. ■

Comme cela est détaillé dans la section 5.1.3, le résultat de la convolution d'une fonction continue par morceaux par un noyau spatialement invariant est aisément calculable en vertu de la linéarité de l'intégration et du théorème 5.1 sur la décomposition d'une telle fonction, voir proposition 5.1. En revanche étendre au cas bi-dimensionnel

afin d'obtenir une relation entre intensité d'une ligne de la scène et éclairage incident au niveau de la même ligne du photo-détecteur n'est pas évident en raison du caractère 2D de la convolution (et de la scène) et la PSF spatialement variante.

L'introduction de la notion d'amplitude d'une discontinuité est nécessaire pour formaliser la relation entre  $S_k$  et  $\theta_k$ , resp. les fonctions univariées représentant l'intensité rayonnée par la  $k$ -ième ligne de la scène et l'intensité incidente sur la  $k$ -ième ligne du photo-détecteur.

Pour comprendre les difficultés que pose l'extension de la proposition 5.1 au cas de fonctions à deux variables, il est proposé de brièvement étudier le cas simple de deux ensembles ouverts  $\mathcal{X}_1$  et  $\mathcal{X}_2$ , séparés par une frontière  $\partial\mathcal{X}$ , et d'un champ scalaire  $f : \mathcal{X} \setminus \partial\mathcal{X}$  défini par  $f(x, y) = \sum_{i=1}^2 f_i(x, y) \mathbf{1}_{\mathcal{X}_i}(x, y)$ . Il est immédiat de constater qu'il existe deux fonctions  $f_c$  et  $f_s$  telles que  $f_c \in \mathcal{C}^0(\mathcal{X}, \mathbb{R})$  et  $f_d = u_2 \mathbf{1}_{\mathcal{X}_2}(x, y)$  vérifiant :

$$\forall (x, y) \in \mathcal{X}, f(x, y) = f_c(x, y) + f_d(x, y) = f_c(x, y) + u_2 \mathbf{1}_{\mathcal{X}_2}(x, y), \quad (5.17)$$

si et seulement si l'amplitude des discontinuités est constante, i.e.  $\forall (x, y) \in \partial\mathcal{X}, U(x, y) = u \in \mathbb{R}$ . Ce cas, très restrictif, ne permet pas une modélisation réaliste des discontinuités de la fonction  $S$  représentant l'intensité rayonnée par la scène imagée.

En outre, si l'on suppose que cette condition est vérifiée, le calcul du résultat de la convolution du champ scalaire  $f$  par un noyau  $h$  est délicat en raison de la géométrie de la discontinuité :

$$\begin{aligned} g(x, y) &= [f**h](x, y) = [f_c**h](x, y) + [f_d**h](x, y) \\ &= [f_c**h](x, y)(x, y) + u_2 \int_{\mathcal{X}_2} h(x - u, y - v) du dv. \end{aligned} \quad (5.18)$$

Le dernier terme de cette somme peut avoir une expression difficilement exploitable suivant la géométrie de l'ensemble  $\mathcal{X}_2$  (ou de la courbe de discontinuité  $\partial\mathcal{X}$ ).

Pour résoudre cette difficulté et prendre en compte la PSF spatialement variante, il est proposé d'utiliser l'étendue du support de la PSF autour d'un point  $(x_k, y_k) = \Gamma(t_k)$ ,  $t_k \in J$  pour définir un voisinage  $\mathcal{V}_k^\epsilon$  autour de ce point. Il est rappelé en préambule que la PSF  $h_{\text{opt}}$  vérifie, par définition, la propriété :  $\int_{\mathbb{R}^2} |h_{\text{opt}}(x, y; \zeta_0)| dx dy = 1$ .

*Le fait que la PSF soit une fonction positive, i.e.  $h_{\text{opt}}(x, y; \zeta_0) \geq 0$  n'est pas utilisé ici.*

**Définition 5.3.** Soit un point  $(x_k, y_k) = \Gamma(t_k)$ ,  $t_k \in J$ , un noyau de convolution isotrope (la PSF optique dans notre cas)  $h_{\text{opt}}$  et le disque de centre  $(x_k, y_k)$  et de rayon  $\rho$  i.e.  $\mathcal{D}_k(\rho) = \{(x, y) \in \mathbb{R}^2 / (x - x_k)^2 + (y - y_k)^2 \leq \rho\}$ . Le voisinage  $\mathcal{V}_k^\epsilon$  du point  $(x_k, y_k)$  est défini comme le plus petit disque  $\mathcal{D}_k(\rho_{\min})$  permettant de vérifier  $\iint_{\mathcal{D}_k(\rho_{\min})} h_{\text{opt}}(x - x_k, y - y_k) dx dy \geq 1 - \epsilon$ , ou plus formellement, en posant  $\rho_{\min} = \arg \min_{\rho \in \mathbb{R}^+} 1 - \iint_{\mathcal{D}_k(\rho)} h_{\text{opt}}(x, y) dx dy \leq \epsilon$ .

La définition du voisinage  $\mathcal{V}_k^\epsilon$  est difficilement applicable en pratique mais permet de formaliser l'analyse locale d'une discontinuité. L'idée de cette approche est de simplifier l'expression analytique de l'éclairage incident autour d'un point de discontinuités,  $(x_k, y_k) = \Gamma(t_k)$ ,  $t_k \in J$  en limitant le domaine d'intégration à un voisinage  $\mathcal{V}_k^\epsilon$ .

Un dernier exemple permet alors de mettre en évidence comment l'expression de l'éclairage peut être modélisée simplement à l'aide des notions précédemment définies et, surtout, de formaliser les conditions d'applications.

**Exemple 5.1.** Soit un compact fermé  $\mathcal{X}$  et deux sous ensembles ouverts  $\mathcal{X}_1$  et  $\mathcal{X}_2$  séparés par une frontière  $\partial\mathcal{X}$  tels que  $\{\mathcal{X}_1, \mathcal{X}_2, \partial\mathcal{X}\}$  forment une partition de  $\mathcal{X}$ . Soit un champ scalaire  $f : \mathcal{X} \rightarrow \mathbb{R}$  défini pour tout  $(x, y) \in \mathcal{X} \setminus \partial\mathcal{X}$  par  $f(x, y) = \sum_{i=1}^2 f_i(x, y) \mathbf{1}_{\mathcal{X}_i}(x, y)$ . Supposons qu'il existe un point  $(x_k, y_k) \in \partial\mathcal{X}$  et  $\epsilon \gtrsim 0$  tel

1. l'ensemble  $\partial\mathcal{X} \cap \mathcal{V}_k$  peut être représenté par le segment  $\mathcal{S}_k = \{(x, y) \in \mathcal{V}_k / y = y_k\}$ ;
2. pour tout  $(x, y) \in \partial\mathcal{X} \cap \mathcal{V}_k$  on a  $U(x, y) = u_k \mathbf{1}_{\partial\mathcal{X}}(x, y)$ , avec  $u_k \in \mathbb{R}$ .

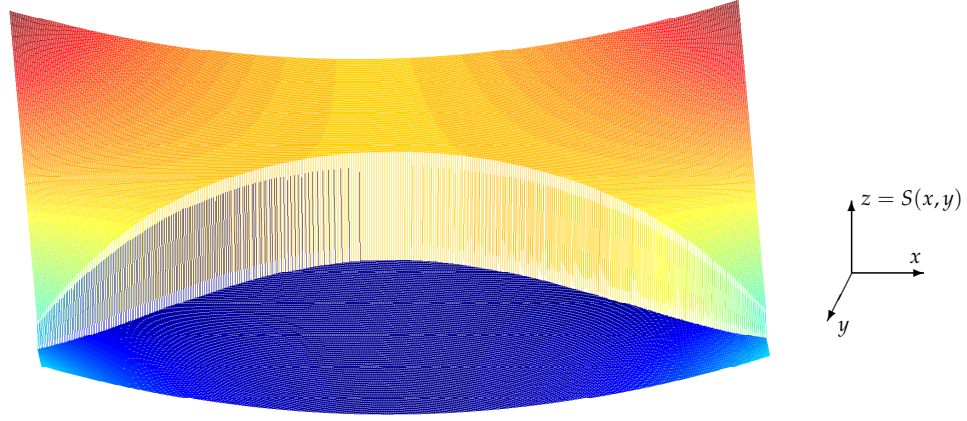


FIGURE 5.6: Illustration d'un champ scalaire simple,  $S(x, y)$ , continu par morceaux mais ne pouvant être décomposé tel (5.17).

Alors, un calcul direct permet de montrer que pour tout  $(x, y) \in \mathcal{V}_k$ , le résultat de la convolution du champ scalaire  $f$  par un noyau  $h$  isotrope est donné par :

$$\begin{aligned} g(x, y) &= [f_c ** h](x, y) + u_k \iint_{\mathcal{V}_k} \mathbf{1}_{u > x_k}(u) h(x - u, y - v) du dv \\ &= g_c(x, y) + u_k \iint_{\mathbb{R}^2} \mathbf{1}_{u > x_k}(u) h(x - u, y - v) du dv - \epsilon \\ &= g_c(x, y) + u_2 H(x - x_k) - \epsilon \end{aligned} \quad (5.19)$$

où, en vertu de l'isotropie de  $h$ ,  $H(x - x_k) = \int_{\mathbb{R}} \int_{-\infty}^{x - x_k} h(u, v) du dv = \int_{\mathbb{R}} \int_{x_k - x}^{\infty} h(u, v) du dv$ .

La relation (5.19) permet une modélisation simple de l'éclairement incident dans un voisinage du point de discontinuité  $(x_k, y_k)$ .

Les conditions d'applications de l'exemple précédent peuvent être étendues au cas où, dans le voisinage  $\mathcal{V}_k$ , l'amplitude des discontinuités dépend linéairement de  $y$ , *i. e.* s'il existe deux réels  $u_{k,0}$  et  $u_{k,1}$  tels que  $U(x, y) = (u_{k,0} + u_{k,1}(y - y_k)) \mathbf{1}_{\partial \mathcal{X}}(x, y)$ . En effet, le calcul (5.19) devient alors :

$$\begin{aligned} g(x, y) &= g_c(x, y) + u_{k,0} \iint_{\mathbb{R}^2} \mathbf{1}_{u > x_k}(u) h(x - u, y - v) du dv \\ &\quad + u_{k,1} \iint_{\mathbb{R}^2} \mathbf{1}_{u > x_k}(u) (y - y_k) h(x - u, y - v) du dv - \epsilon. \end{aligned}$$

Or en vertu de l'isotropie de  $h$ , cette fonction est paire alors que  $(y - y_k)$  est impaire ; le dernier terme de la somme est donc nul.

En outre, il est immédiat d'étendre, en utilisant un changement de repère adéquat, les deux exemples précédents au cas où l'ensemble  $\partial \mathcal{X} \cap \mathcal{V}_k$  peut être représenté par un segment quelconque  $\mathcal{S}_k$ .

Similairement, le problème de la PSF spatialement variant peut-être résolu de façon analogue ; cette dernière pouvant en effet être considérée localement comme spatialement invariante. Ainsi, la proposition 5.2 résume ci-dessous le modèle de l'éclairement incident sur une ligne de l'image et les conditions permettant l'obtention de ce résultat.

**Proposition 5.2** (Décomposition d'un champ scalaire continu par morceaux). *Soit un compact  $\mathcal{X} \subseteq \mathbb{R}^2$  et  $N$  sous-ensembles compacts ouverts  $\mathcal{X}_i$ ,  $i = \{1, \dots, N\}$  dont l'ensemble des frontières est noté  $\partial \mathcal{X} = \{\Gamma(t), t \in J\}$ . Soit un champ scalaire  $f : \mathcal{X} \rightarrow \mathbb{R}$  défini par :  $f(x, y) = \sum_{i=1}^N f_i(x, y) \mathbf{1}_{(x, y) \in \mathcal{X}_i}(x, y)$  définissant le champ scalaire d'amplitude des discontinuités  $U$  et  $f_0$  la restriction de  $f$  à la droite  $D = \{(x, y) \in \mathcal{X} / y = y_0\}$  avec*

$D \cup \partial\mathcal{X} = \{(x_k, y_0)\}_{k=1}^{K_0}$ . Soit enfin un noyau de convolution  $h : \mathbb{R}^2 \rightarrow \mathbb{R}$  et  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  défini par  $g = f**h$ .

Si dans un voisinage  $\mathcal{V}_k$  de  $(x_k, y_0) \in \partial\mathcal{X}$  "suffisamment large" les conditions suivantes sont vérifiées :

C-1 la réponse impulsionnelle  $h$  est isotrope et spatialement invariante dans  $\mathcal{V}_k$ , i. e. si  $h(u, v, x, y) = h(\rho_k)$  avec  $\rho_k^2 = (x_k - u)^2 + (y_0 - v)^2$ ;

C-2  $\partial\mathcal{X} \cap \mathcal{V}_k$  peut être représentée par la droite  $D_k$  formant un angle  $\psi_k$  avec la droite  $D$ ;

C-3 pour tout  $(x, y) \in \mathcal{V}_k$  on a  $U(x, y) = u_k \mathbf{1}_{\partial\mathcal{X}}(x, y)$ , avec  $u \in \mathbb{R}$ .

alors la fonction  $g_0$ , la restriction du champ scalaire  $g$  à la droite  $D = \{(x, y) \in \mathcal{X} / y = y_0\}$ , admet la décomposition suivante :

$$g_0(x) = g_{0,c}(x) + \sum_{k=1}^{K_0} u_k H\left(\frac{x - x_k}{\psi_k}\right)$$

avec  $g_{0,c} \in \mathcal{C}^0$  et  $H(x) = \int_{\mathbb{R}} \int_{-\infty}^x h(u, y) du dy$

**Démonstration 5.2.** Pour démontrer le théorème, il est proposée d'analyser  $g_0(x)$  dans un voisinage  $\mathcal{V}_k$  de  $(x_k, y_0) \in \partial\mathcal{X}_{i,j}$ . En utilisant le changement de repère  $\Psi : (\tilde{x}, \tilde{y}) \rightarrow (x, y)$

défini par  $\Psi : (\tilde{x}, \tilde{y}) = \begin{pmatrix} x_k + \tilde{x} \cos(\psi_k) + \tilde{y} \sin(\psi_k) \\ y_k + \tilde{x} \sin(\psi_k) - \tilde{y} \cos(\psi_k) \end{pmatrix}$  de sorte que le nouveau repère soit centré sur le point de discontinuité et l'axe  $\tilde{x}$  orthogonal à l'axe optique. Pour tout  $(\tilde{x}, \tilde{y}) \in \mathcal{V}_k$  on a :

$$g(\tilde{x}, \tilde{y}) = [f**h](\tilde{x}, \tilde{y}) = \int_{\mathcal{V}_k} f(\tilde{x} - u, \tilde{y} - v) h(u, v) du dv.$$

Or en vertu des conditions C-1, C-2 et C-3, la fonction  $f$  admet, pour tout  $(\tilde{x}, \tilde{y}) \in \mathcal{V}_k$ , la décomposition  $f(\tilde{x}, \tilde{y}) = f_c(\tilde{x}, \tilde{y}) + U(\tilde{x}, \tilde{y})$ , où  $f_c \in \mathcal{C}^0(\mathcal{V}_k, \mathbb{R})$  et le champ scalaire est défini dans la proposition 5.2.

décomposition à mettre en proposition

Il est alors immédiat de vérifier que :

$$\begin{aligned} g(\tilde{x}, \tilde{y}) &= [f_c**h](\tilde{x}, \tilde{y}) + \int_{\mathcal{V}_k} u_k \mathbf{1}_{\mathcal{X}_j}(\tilde{x} - u, \tilde{y} - v) h(u, v) du dv \\ &= [f_c**h](\tilde{x}, \tilde{y}) + u_k \int_{\mathbb{R}} \int_{-\infty}^{\tilde{x}} h(u, v) du dv \\ &= g_c(\tilde{x}, \tilde{y}) + u_k H(\tilde{x}) \end{aligned}$$

il en découle alors, en utilisant le changement de repère  $\Psi^{-1}$ , que :

$$g(x, y_0) = [f**h](x, y) \Big|_{y=y_0} = g_c(x, y_0) + u_k H\left(\frac{x - x_k}{\cos(\psi_k)}\right)$$

■





Troisième partie

EXPLOITATION DU MODÈLE DES IMAGES POUR LA  
DÉTECTION STATISTIQUE D'INFORMATIONS  
CACHÉES



---

# Exploitation du modèle non-linéaire pour la détection statistique d'informations cachées

---

# 6

*«Les choses devraient être faites aussi simples que possible,  
mais pas plus simples.»*

— Albert Einstein, 1879 – 1955

*«For every problem there is a solution which is simple, clean and wrong.»*

— Henry Louis Mencken, 1880 – 1956

## Sommaire

---

Introduction au chapitre . . . . .	123
6.1 Introduction à la détection statistique en présence de paramètres non-linéaires . . . . .	124
6.1.1 Motivations et difficultés . . . . .	125
6.1.2 Méthodes numériques d'estimation du paramètre non-linéaire de nuisance . . . . .	126
6.2 Principe de l'approche par linéarisation . . . . .	127
6.3 Estimation du point de fonctionnement et biais des résidus . . . . .	130
6.3.1 Algorithmes utilisés pour l'estimation des discontinuités . . . . .	132
6.3.2 Étude qualitative de la linéarisation du modèle local . . . . .	135
6.4 Détection statistique d'informations cachées : impact de la linéarisa- tion sur les performances du test statistique . . . . .	137
6.4.1 Contenu de l'image inconnu : test statistique proposé . . . . .	137
6.4.2 Calcul des performances du test quasi-optimal proposé . . . . .	138
Conclusions du chapitre . . . . .	140
6.5 Annexes du chapitre 5 . . . . .	141
6.5.1 Borne de Rao-Cramer sur l'estimation des paramètres de nui- sance non-linéaires . . . . .	141
6.5.2 Démonstration des théorèmes 6.1 et 6.2. . . . .	142

---

### Introduction au chapitre :

L'objectif du chapitre 6 est de présenter la méthodologie permettant la conception d'un test statistique pour la détection d'informations cachées en exploitant le modèle réaliste des images naturelles, proposé dans le chapitre 5 précédent. Compte-tenu de la finalité applicative visée, une attention particulière est portée à la simplicité calculatoire de la méthodologie dans son ensemble et à la possibilité d'explicitier analytiquement les propriétés du test statistique, en terme de probabilités d'erreurs notamment.

La première difficulté, abordée dans la section 6.1, est la détection et l'estimation statistique en présence de paramètres de nuisance intervenant de façon non-linéaire. Pour surmonter ce problème de façon satisfaisante, une approche par linéarisation du modèle autour des discontinuités est présentée dans la section 6.2. Cette méthodologie de linéarisation nécessite d'estimer les paramètres non-linéaires décrivant une discontinuité. Ainsi, l'approche utilisée en pratique, issue de la littérature, est décrite dans la section 6.3 et l'impact sur l'estimation du contenu est étudié. L'intérêt fondamental de cette approche est exposé au travers du test statistique présenté dans la section 6.4. L'impact, sur les performances du test, de la présence de discontinuités est maîtrisé dans le sens où une contrainte sur la probabilité de fausse alarme peut être garantie et la perte de puissance est bornée (comparativement au test le plus puissant).

En résumant de façon lacunaire les chapitres 3-5, la stéganalyse des images est, de façon générale, un problème de décision statistique entre hypothèses composites (le taux d'insertion  $R$  étant inconnu), à partir de variables aléatoires quantifiées et en présence de paramètres de nuisance intervenant de façon non-linéaire (contenu de l'image  $\theta_k$ ).

Le problème de décision d'informations cachées a été abordé sous un aspect purement statistique dans le chapitre 3; différentes approches ont conduit à proposer différents tests optimaux. L'application pratique de ces tests est toutefois sujette à l'estimation des paramètres de distribution de chacun des pixels, notamment du paramètre de nuisance  $\theta_k$ . Or, pour conserver une modélisation parcimonieuse du modèle malgré le caractère continu par morceau de la scène, une paramétrisation non-linéaire du paramètre de nuisance  $\theta_k$  est nécessaire ce qui rend délicat le problème d'estimation du contenu, et donc *in fine* de détection statistique.

Le présent chapitre est consacré à la description de la méthodologie proposée pour la détection statistique d'informations cachées dans une image dont les paramètres sont inconnus. Dans la section 6.1 sont présentées les difficultés engendrées par la présence du paramètre de nuisance  $\theta_k$ , intervenant de façon non-linéaire dans le modèle du contenu des images (5.10) - (5.11). La section 6.2 présente l'approche par linéarisation proposée pour surmonter les difficultés induites par le modèle non-linéaire. Cette méthodologie repose sur une estimation du paramètre  $\eta_k$ ; aussi, bien que cela ne soit pas le centre d'intérêt des travaux menés, la section 6.3 présente brièvement quelques méthodes d'estimation de  $\eta_k$  et l'impact de la linéarisation sur l'estimation de  $\theta_k$ . Enfin, les principaux résultats obtenus dans le cadre de cette thèse sont présentés dans la section 6.4 : le test statistique pour la détection d'informations cachées  $y$  est défini et ses performances  $y$  sont analytiquement établies. Il est en particulier souligné que la perte de puissance (ou d'optimalité), comparativement au test asymptotiquement le plus puissant proposé dans la section 3.5, est bornée.

## 6.1 Introduction à la détection statistique en présence de paramètres non-linéaires

Il est proposé en préambule de "prendre un peu de recul" en mettant en exergue la difficulté d'application de la théorie de la détection statistiques présentée dans la section 2.4 et dans l'annexe A, lorsque le modèle de représentation des données est non-linéaire.

### 6.1.1 Motivations et difficultés

Dans le chapitre 3, une modélisation *réaliste* du contenu structuré des images n'a pas été évoquée. Les tests proposés sont intéressants d'un point de vue théorique, néanmoins pour une application pratique, une estimation  $\hat{\theta}_k$  de  $\theta_k$  s'avère nécessaire. Une application pratique a été proposée dans la section 3.4 en utilisant une modélisation linéaire (3.38) des segments issus des images et les résultats obtenus dans [23, 96, 160] avec cette approche sont encourageants.

En pratique, l'utilisation du modèle non-linéaire du contenu des images naturelles (5.10) - (5.11) est nécessaire afin de conserver un paramétrage parcimonieux et donc *in fine* de préserver la puissance du test qui en découle. Il est rappelé que les segments  $\mathbf{z}_k$ ,  $k \in \{1, \dots, K\}$  sont modélisés par :

$$\mathbf{z}_k = Q_1(y_k) = Q_1(\theta_k + \boldsymbol{\zeta}_k) \quad \text{avec : } \boldsymbol{\zeta}_k \sim \mathcal{N}(0, \Sigma_k)$$

avec (5.9) - (5.10) :

$$\begin{aligned} \theta_k &= \mathbf{H}\mathbf{c}_k + \mathbf{F}(\boldsymbol{\eta}_k)\mathbf{u}_k = \mathbf{G}(\boldsymbol{\eta}_k)\mathbf{v}_k, \\ \text{et } \mathbf{G}(\boldsymbol{\eta}_k) &= \left( \mathbf{H} \mid \mathbf{F}(\boldsymbol{\eta}_k) \right) \quad \text{et } \mathbf{v}_k = \left( \mathbf{c}_k, \mathbf{u}_k \right)^T. \end{aligned} \quad (6.1)$$

Le modèle de  $\theta_k$  est donc composé d'une partie linéaire  $\mathbf{H}\mathbf{c}_k$ , représentant la partie continue de la scène et d'une partie non-linéaire  $\mathbf{F}(\boldsymbol{\eta}_k)\mathbf{u}$  représentant les éventuelles discontinuités présentes dans le  $k$ -ième segment.

La présence, dans le modèle des observations, de paramètres de nuisance intervenant de façon non-linéaire peut *a priori* être traitée par une application directe de la théorie de la décision statistique décrite dans l'annexe A, au coût d'éventuelles adaptations. Le test du RVG demeure notamment utilisable. Dans le cadre du problème de détection d'informations cachées considéré, le RVG est rigoureusement donné, voir définition 2.10, par :

$$\log \Lambda^{\text{rvg}}(\mathbf{Z}) = \frac{\sup_{R_1 > R^*; \theta_k \in \Theta_k} Q_{\theta_k}^{R_1}(\mathbf{Z})}{\sup_{R_0 \leq R^*; \theta_k \in \Theta_k} Q_{\theta_k}^{R_0}(\mathbf{Z})},$$

où  $\Theta_k = \text{Im}(\mathbf{G}(\boldsymbol{\eta}_k)) \subset \mathbb{R}^L$  représente l'espace du paramètre  $\theta_k \in \Theta_k$ ,  $Q_{\theta_k}^R(\mathbf{Z})$  représente la loi de distribution des pixels du  $k$ -ième segment après insertion avec un taux  $R$  et  $R^*$  correspond au taux à partir duquel il est souhaitable de détecter les informations cachées.

Similairement le test "type RVG" proposé dans (3.40) - (3.41) peut être aisément adapté au modèle non-linéaire du contenu :

$$\log \Lambda^{\text{rvg}}(\mathbf{Z}) = \frac{\sup_{\theta_k \in \Theta_k} Q_{\theta_k}^{R_1}(\mathbf{Z})}{\sup_{\theta_k \in \Theta_k} Q_{\theta_k}^{R_0}(\mathbf{Z})}.$$

En pratique, le RVG "exact" revient à estimer simultanément par maximum de vraisemblance le paramètre  $\theta_k$  et le taux d'insertion  $R$ , alors que, dans le RVG proposé, la maximisation par rapport au taux d'insertion  $R$  n'est pas considérée (voir remarques 3.5 et 3.6).

Cependant, l'application du test du RVG, et plus généralement de la théorie de la détection et de l'estimation statistique, peut être rendue complexe en pratique par la présence de paramètres non-linéaires. De manière générale les difficultés suivantes apparaissent lorsque le modèle des observations est non-linéaire :

1. D'un point de vue théorique, l'inférence de l'anomalie qu'il est souhaitable de détecter sur l'estimation du paramètre de nuisance  $\theta_k$  doit être prise en compte [115, 161].

*Il est à souligner que l'application  $\mathbf{F}(\boldsymbol{\eta}_k)$  n'est pas linéaire par rapport au paramètre  $\boldsymbol{\eta}_k$  (sans quoi le formalisme matricielle est préféré).*

2. Le calcul des performances du test statistique qui en découle est rendu complexe ; dans le cas du RVG par exemple l'impact sur les propriétés statistiques du test de la procédure d'estimation par maximum de vraisemblance n'est pas toujours évident à "mesurer".
3. D'un point de vue pratique, le problème d'optimisation sous-jacent au calcul du maximum de vraisemblance est beaucoup plus complexe dans le cas d'un modèle non-linéaire que dans le cas d'un modèle linéaire.

Dans le cadre des travaux présentés, les "anomalies" qu'il est souhaitable de détecter sont les modifications engendrées par l'insertion d'informations dans les LSB des pixels d'une image ; en outre le cas d'intérêt suppose que le taux d'insertion  $R$  est faible. L'inférence de la modification de  $\pm 1$  d'une proportion réduite de pixels d'une image avec le contenu  $\theta_k$  n'est donc pas considéré. Le problème consistant à expliciter les propriétés statistiques du détecteur proposé, fil conducteur des présents travaux, est le centre de gravité des sections 6.3.2 et 6.4.

Enfin, la résolution numérique du problème d'optimisation sous-jacent à l'estimation du paramètre de nuisance  $\theta_k$ , ou de manière analogue du paramètre non-linéaire  $\eta_k$ , est brièvement abordé ci-dessous dans la section 6.1.2.

### 6.1.2 Méthodes numériques d'estimation du paramètre non-linéaire de nuisance

Dans un souci de clarté, il est supposé que les observations  $\mathbf{z}_k$  sont des valeurs réelles, décrites par (6.1) :

$$\mathbf{z}_k = \mathbf{G}(\eta_k) + \xi_k,$$

où l'application  $\mathbf{G} : \mathbb{R}^p \rightarrow \mathbb{R}^L$  est connue,  $\eta \in \Omega \subset \mathbb{R}^p$  et  $\xi \in \mathbb{R}^N$  est un vecteur aléatoire dont la distribution statistique est supposée connue. Le problème qu'il est souhaitable de résoudre est celui de l'estimation numérique de  $\eta_k$  à partir de  $\mathbf{z}_k$ . Sans perte de généralité, il est possible de considérer la famille de distribution paramétrique  $\mathcal{P} = \{P_\eta \mid \eta \in \Omega\}$  et de décrire  $\mathbf{z}$  comme la réalisation d'un vecteur aléatoire de distribution  $P_\eta$ , admettant la densité  $f_\eta$ , ce qui est noté de manière abusive  $\mathbf{z} \sim P_\eta$ . Dans un cadre non-Bayésien (lorsque la distribution *a priori* du paramètre  $\eta_k$  n'est pas connue) l'estimation  $\hat{\eta}_k$  de  $\eta_k$  est usuellement obtenue par maximum de vraisemblance :

$$\hat{\eta} = \arg \max_{\eta \in \Omega} f_\eta(\mathbf{z}_N). \quad (6.2)$$

Dans le cadre des travaux présentés dans ce manuscrit, il est considéré (voir section 4.3) que les variables aléatoires  $\xi_n$ ,  $n \in \{1, \dots, L\}$  sont statistiquement indépendantes. L'estimation de  $\eta$  par maximum de vraisemblance est alors donnée, voir définition A.22, par :

$$\hat{\eta} = \arg \max_{\eta \in \Omega} \prod_{n=1}^N f_\eta(z_n). \quad (6.3)$$

L'estimation par maximum de vraisemblance est donc, formellement, un problème d'optimisation (recherche d'un maximum global) sous la contrainte  $\eta \in \Omega$  qui peut être traité avec les méthodes usuelles d'optimisation. La résolution des équations (6.2) et (6.3) peut s'avérer délicate selon la nature de la fonction  $\mathbf{G}$ , selon l'espace des contraintes et selon la densité de probabilité  $f_{\text{veta}}$ .

Dans le cas où le bruit  $\xi$  est i.i.d. et est distribué suivant une loi Gaussienne, le vecteur des observations  $\mathbf{z}_N$  est alors décrit statistiquement par  $\mathbf{z}_N \sim \mathcal{N}(\mathbf{G}(\eta), \sigma^2 \mathbf{I}_L)$ . L'estimation par maximum de vraisemblance est alors donné par résolution du problème de moindres carrés non-linéaires suivant :

$$\hat{\eta} = \arg \min_{\eta \in \Omega} \|\mathbf{z}_N - \mathbf{h}(\eta)\|_2^2. \quad (6.4)$$

Ce résultat est analogue à la remarque A.3.

La minimisation du critère quadratique (6.4) a longuement été étudiée dans la littérature et de nombreuses méthodes d'optimisation convexe ont été proposées.

**Remarque 6.1.** *De manière générale, l'existence et l'unicité de solutions pour les problèmes (6.2) - (6.4) ne sont pas assurées. Il est néanmoins possible de montrer que sous certaines conditions de régularité, aisément vérifiables dans le cas du modèle non-linéaire étudié, que l'estimation  $\hat{\eta}$  de  $\eta$  existe et est unique.*

Pour procéder à la résolution des problèmes d'optimisation (6.2) - (6.4), de nombreux algorithmes existent [162]. De manière générale, il est possible d'utiliser des méthodes déterministes, tels que la descente de gradient, le gradient conjugué, le GNC ou le Quasi-Newton, ou des méthodes stochastiques tels que le recuit simulé [163] ou la descente stochastique de gradient. Qu'ils soient déterministes ou stochastiques, ces algorithmes font appel à des procédés répétés itérativement jusqu'à la vérification d'un critère d'arrêt et, en général, demandent une valeur initiale  $\eta_0$  du paramètre à estimer. La résolution d'un problème de moindres carrés non-linéaires est donc d'un coût calculatoire nettement plus important qu'une simple minimisation au sens des moindres carrés linéaires. En outre, la convergence vers un extremum global des méthodes déterministes n'est, dans le cas général, pas garantie, notamment en présence d'extrema locaux, comparativement aux méthodes stochastiques qui, en revanche, présentent un coût calculatoire bien plus important.

Dans le cadre des travaux présentés dans ce manuscrit, l'estimation des paramètres non-linéaires de nuisance décrivant les discontinuités dans chacune des lignes d'une image ne peut se faire en utilisant les méthodes d'optimisation brièvement décrites ci-dessus. En effet, compte-tenu des contraintes de temps d'analyse d'une image (qui doit être réduit au maximum) et du nombre très élevé de discontinuités présentes dans l'ensemble d'une image, il ne peut être envisagé une résolution directe des équations (6.2) - (6.4). En outre, le principal problème abordé dans ce manuscrit concerne la description précise des performances du test statistique en vue de la détection d'informations cachées dans une image. Les propriétés statistiques du test dépendent de l'estimation du paramètre de nuisance que représente le contenu structuré d'une image. Or, il n'est pas possible d'explicitement les performances statistiques de l'estimation du contenu au moyen d'algorithmes d'optimisation, ce qui compromet la possibilité de calculer analytiquement les performances du test.

Pour ces raisons, l'estimation directe des paramètres de nuisance  $\eta_k$  intervenant de façon non-linéaire est à proscrire dans le cadre des travaux présentés dans ce manuscrit.

## 6.2 Principe de l'approche par linéarisation

Si l'on suppose, pour l'exemple, que le paramètre  $\eta_k$  peut être déterminé avec exactitude le modèle (6.1) devient linéaire par rapport aux variables inconnues  $\mathbf{c}_k$  et  $\mathbf{u}_k$ . De manière analogue, l'approche par linéarisation consiste à admettre qu'une valeur approchée  $\hat{\eta}_k$  du paramètre  $\eta_k$  peut être déterminée. Il est alors possible de recourir à un développement de Taylor au premier-ordre de l'application  $\mathbf{F}$  autour de la valeur approchée  $\hat{\eta}_k$  de  $\eta_k$ , ce qui présente comme intérêt immédiat de rendre le modèle linéaire par rapport à la quantité  $\eta_k - \hat{\eta}_k$ . Par analogie à l'étude des systèmes dynamiques le terme de *point de fonctionnement* est utilisé pour qualifier  $\hat{\eta}_k$ .

De manière générale, si l'application  $\mathbf{F}$  est deux fois continûment différentiable elle admet le développement au second-ordre suivant :

$$\mathbf{F}(\eta_k) = \mathbf{F}(\hat{\eta}_k) + (\eta_k - \hat{\eta}_k) \dot{\mathbf{F}}_{\hat{\eta}_k} + \frac{1}{2} (\eta_k - \hat{\eta}_k)^T \ddot{\mathbf{F}}_{\hat{\eta}_k} (\eta_k - \hat{\eta}_k) + o(\|\eta_k - \hat{\eta}_k\|_1^2), \quad (6.5)$$



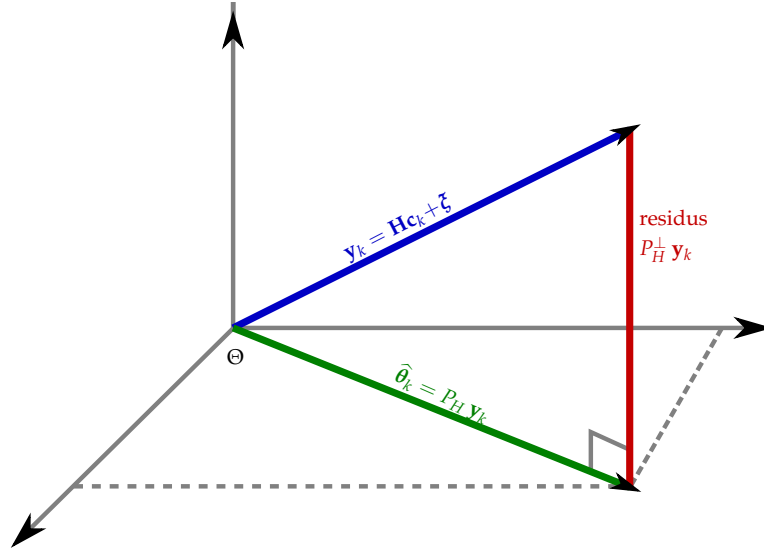


FIGURE 6.1: Illustration en 3D de l'estimation d'un paramètre linéaire ; dans le cas d'un bruit additif Gaussien l'estimation  $\theta$  est obtenue par projection orthogonale de  $z$  sur l'espace des paramètres  $\Theta$ .

où  $\dot{\mathbf{F}}_{\hat{\eta}_k}$  représente la matrice Jacobienne de  $\mathbf{F}$ , de taille  $N \times p$ , dont les éléments sont donnés, pour tout  $(n, i)$ , par :

$$\dot{\mathbf{F}}_{\hat{\eta}_k}[n, i] = \left. \frac{\partial f}{\partial \eta_{k,i}}(x; \boldsymbol{\eta}_k) \right|_{x=n, \boldsymbol{\eta}_k=\hat{\boldsymbol{\eta}}_k}$$

et, de manière analogue,  $\ddot{\mathbf{F}}_{\hat{\eta}_k}$  représente le Hessian de  $\mathbf{F}$  par un tableau tri-dimensionnel de taille  $N \times p \times p$  dont les éléments sont donnés, pour tout  $(n, i, j)$ , par :

$$\ddot{\mathbf{F}}_{\hat{\eta}_k}[n, i, j] = \left. \frac{\partial^2 f}{\partial \eta_{k,i} \partial \eta_{k,j}}(x; \boldsymbol{\eta}_k) \right|_{x=n, \boldsymbol{\eta}_k=\hat{\boldsymbol{\eta}}_k}.$$

**Remarque 6.2.** Le Hessian  $\ddot{\mathbf{F}}_{\hat{\eta}_k}$  représenté par un tableau tri-dimensionnel peut être considéré comme une juxtaposition de  $N$  matrices de taille  $p \times p$ . Le produit  $(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)^T \ddot{\mathbf{F}}_{\hat{\eta}_k}$ , ou de manière commutative le produit  $\ddot{\mathbf{F}}_{\hat{\eta}_k}(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)$ , doit donc être considéré comme  $N$  produits distincts. Ainsi, le résultat du produit  $(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)^T \ddot{\mathbf{F}}_{\hat{\eta}_k}(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)$  est un vecteur dont les  $N$  composantes sont données par  $(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)^T \ddot{\mathbf{F}}_{\hat{\eta}_k}[n, :, :](\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)$ , où  $\ddot{\mathbf{F}}_{\hat{\eta}_k}[n, :, :]$  représente la  $n$ -ième matrice de dimension  $p \times p$  extraite du Hessian  $\ddot{\mathbf{F}}_{\hat{\eta}_k}$ .

De manière alternative, le vecteur  $(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)^T \ddot{\mathbf{F}}_{\hat{\eta}_k}(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)$  peut également être exprimé par  $\sum_i \sum_j (\eta_{k,i} - \hat{\eta}_{k,i})^T \ddot{\mathbf{F}}_{\hat{\eta}_k}[:, i, j](\eta_{k,j} - \hat{\eta}_{k,j})$  où  $\ddot{\mathbf{F}}_{\hat{\eta}_k}[:, i, j]$  est un vecteur de  $N$  composantes représentant la dérivée de seconde de  $\mathbf{F}$  par rapport à  $\eta_{k,i}$  et  $\eta_{k,j}$ ; plus de détails sur ce point précis sont donnés dans [162, Annexe B4].

En supposant qu'une unique discontinuité est présente, l'omission de l'influence du terme  $(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)^T \ddot{\mathbf{F}}_{\hat{\eta}_k}(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k)$  d'ordre 2 dans la relation (6.5) permet d'obtenir la modélisation "linéarisée" approximative du vecteur  $\boldsymbol{\theta}_k$  :

$$\boldsymbol{\theta}_k = \mathbf{H}\mathbf{c}_k + u_k \mathbf{F}(\hat{\boldsymbol{\eta}}_k) + u_k (\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k) \dot{\mathbf{F}}_{\hat{\eta}_k} + \mathbf{b}_k = \dot{\mathbf{G}}_{\hat{\eta}_k} \mathbf{v}_k + \mathbf{b}_k \quad (6.6)$$

où, avec les notations précédentes, la matrice  $\dot{\mathbf{G}}_{\hat{\eta}_k}$  est donnée par  $\dot{\mathbf{G}}_{\hat{\eta}_k} = (\mathbf{H} | \mathbf{F}(\hat{\boldsymbol{\eta}}_k) | \dot{\mathbf{F}}_{\hat{\eta}_k})$ , le paramètre inconnu  $\mathbf{v}_k$  est désormais donné par  $\mathbf{v}_k = (\mathbf{c}_k, u_k, u_k(\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k))^T$  et  $\mathbf{b}_k$  re-

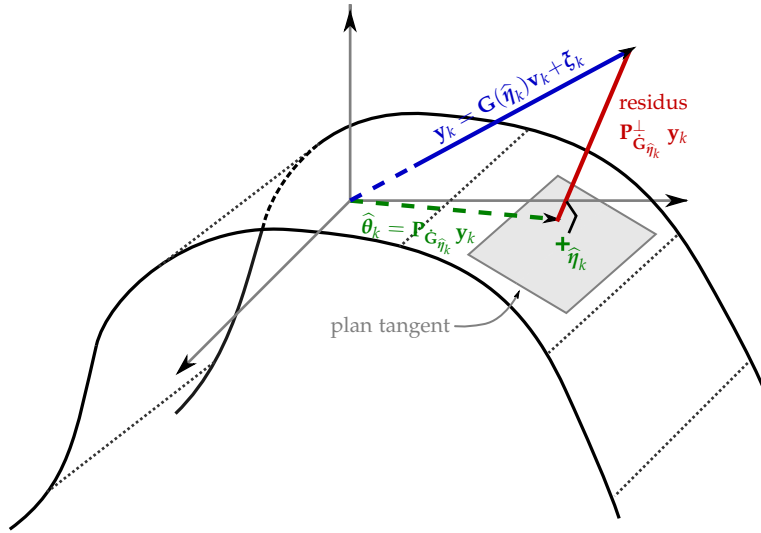


FIGURE 6.2: Illustration en 3D de la méthode proposée pour l'estimation d'un paramètre non-linéaire par linéarisation autour du point  $\hat{\eta}_k$ ; l'estimation  $\hat{\theta}$  est obtenue projection orthogonale sur la plan tangent à  $\Theta$  en  $\hat{\eta}_k$  :  $\hat{\mathbf{z}} = \mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}} \mathbf{z}$ .

présente l'erreur de modélisation induite par la troncature du modèle (6.5) au premier ordre dans (6.6).

### Biais d'estimation engendré par la linéarisation du modèle

L'approximation (6.6) du modèle (6.5) étant linéaire, il est possible d'appliquer les méthodes de détection et d'estimation statistique évoquées dans le chapitre 3. En particulier, l'estimation  $\hat{\theta}_k$  de  $\theta_k$  au sens du maximum de vraisemblance est donnée par :

$$\hat{\theta}_k = \mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}} \mathbf{y}_k = \dot{\mathbf{G}}_{\hat{\eta}_k} \left( \dot{\mathbf{G}}_{\hat{\eta}_k}^T \Sigma_k^{-1} \dot{\mathbf{G}}_{\hat{\eta}_k} \right)^{-1} \dot{\mathbf{G}}_{\hat{\eta}_k}^T \Sigma_k^{-1} \mathbf{y}_k. \quad (6.7)$$

Il est rappelé que le bruit additif est modélisé par  $\xi_k \sim \mathcal{N}(0, \Sigma_k)$ .

Dans le cas où le bruit additif Gaussien est i.i.d., *i.e.*  $\xi_k$  vérifie  $\xi_k \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_N)$ , il est immédiat de vérifier que  $\mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}}$  est alors donnée par  $\mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}} = \dot{\mathbf{G}}_{\hat{\eta}_k} \left( \dot{\mathbf{G}}_{\hat{\eta}_k}^T \dot{\mathbf{G}}_{\hat{\eta}_k} \right)^{-1} \dot{\mathbf{G}}_{\hat{\eta}_k}^T$ . Dans ce cas particulier,  $\mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}}$  représente la projection orthogonale sur l'espace tangent à la surface  $\mathbf{G}(\eta_k)$  au point  $\hat{\eta}_k$ , voir illustration par la figure 6.2; autrement dit, comme rappelé dans l'annexe A.6, l'estimateur du maximum de vraisemblance coïncide alors avec celui des moindres carrés.

De même, la rejection du paramètre de nuisance "linéarisé", voir section 2.4.3, permet de montrer que l'erreur résiduelle (ou résidus) peut être décomposée de la manière suivante :

$$\begin{aligned} \mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}}^\perp \mathbf{y}_k &= \mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}}^\perp \theta_k + \mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}}^\perp \xi_k \\ &= \underbrace{\mathbf{b}_k}_{\text{biais déterministe}} + \underbrace{\tilde{\xi}_k}_{\text{erreur stochastique}} \end{aligned} \quad (6.8)$$

où la matrice de rejection  $\mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}}^\perp$  est donnée par  $\mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}}^\perp = \mathbf{I}_N - \mathbf{P}_{\dot{\mathbf{G}}_{\hat{\eta}_k}}$ .

Les termes non-explicites qui interviennent dans le développement (6.5) sont usuellement négligés. En utilisant les relations (6.5) et (6.6), il est alors proposé dans [162] de

décomposer le terme de second-ordre  $\ddot{\mathbf{F}}_{\hat{\eta}_k}$  en deux composantes orthogonales définies de la façon suivante :

$$\begin{aligned}\ddot{\mathbf{F}}_{\hat{\eta}_k} &= \ddot{\mathbf{F}}_{\hat{\eta}_k}^T + \ddot{\mathbf{F}}_{\hat{\eta}_k}^N \\ \text{où} &= \ddot{\mathbf{f}}_{\hat{\eta}_k}^T[:, i, j] = \mathbf{P}_{\hat{\mathbf{F}}_{\hat{\eta}_k}} \ddot{\mathbf{f}}_{\hat{\eta}_k}[:, i, j] \\ \text{et} &= \ddot{\mathbf{f}}_{\hat{\eta}_k}^N[:, i, j] = \mathbf{P}_{\hat{\mathbf{F}}_{\hat{\eta}_k}}^\perp \ddot{\mathbf{f}}_{\hat{\eta}_k}[:, i, j] = \ddot{\mathbf{f}}_{\hat{\eta}_k}[:, i, j] - \ddot{\mathbf{f}}_{\hat{\eta}_k}^T[:, i, j]\end{aligned}\quad (6.9)$$

avec  $\mathbf{P}_{\hat{\mathbf{F}}_{\hat{\eta}_k}} = \hat{\mathbf{F}}_{\hat{\eta}_k} (\hat{\mathbf{F}}_{\hat{\eta}_k}^T \hat{\mathbf{F}}_{\hat{\eta}_k})^{-1} \hat{\mathbf{F}}_{\hat{\eta}_k}^T$  et  $\mathbf{P}_{\hat{\mathbf{F}}_{\hat{\eta}_k}}^\perp$  la matrice de projection orthogonale sur l'image de l'application  $\hat{\mathbf{F}}_{\hat{\eta}_k}$  et  $\mathbf{P}_{\hat{\mathbf{F}}_{\hat{\eta}_k}}^\perp = \mathbf{I}_N - \mathbf{P}_{\hat{\mathbf{F}}_{\hat{\eta}_k}}$  est la matrice de projection orthogonale sur l'espace supplémentaire.

Cette décomposition permet de définir respectivement la courbure paramétrique et la courbure intrinsèque du modèle à partir desquelles il est usuel d'étudier la légitimité de l'approximation linéaire d'un modèle. De manière informelle, la courbure paramétrique  $\ddot{\mathbf{F}}_{\hat{\eta}_k}^T$ , représente la composante de  $\ddot{\mathbf{F}}_{\hat{\eta}_k}$  contenue dans le plan tangent à  $\mathbf{F}(\eta_k)$  au point  $\hat{\eta}_k$  et n'intervient donc pas dans le biais déterministe  $\mathbf{b}_k$ . À l'inverse, la courbure intrinsèque  $\ddot{\mathbf{F}}_{\hat{\eta}_k}^N$ , représente la composante de  $\ddot{\mathbf{F}}_{\hat{\eta}_k}$  normale au plan d'approximation  $\hat{\mathbf{F}}_{\hat{\eta}_k}$ .

La méthodologie de linéarisation brièvement présentée ci-dessus soulève deux questions qui sont intrinsèquement liées :

1. Quel est l'impact du biais  $\mathbf{b}_k$  sur les performances statistiques des tests proposés pour la détection d'informations cachées ?
2. Comment choisir le point de fonctionnement  $\hat{\eta}_k$  pour minimiser le biais d'approximation  $\mathbf{b}_k$  ? D'ailleurs en quel sens faut-il minimiser le biais d'approximation ?

Ces deux questions sont abordées dans les sections 6.3 et 6.3.2 suivantes.

**Remarque 6.3.** Une méthode alternative de linéarisation est proposée dans [164, chap.3]. L'approche proposée consiste à rechercher le meilleur plan affine qui approche la variété non-linéaire engendrée par  $\mathbf{G}(\eta_k)$ . Bien que cette approche soit intéressante, elle n'a pas été utilisée dans ces travaux en raison de la complexité calculatoire du problème d'optimisation sous-jacent.

## 6.3 Estimation du point de fonctionnement et biais des résidus

Au regard de ce qui a été présenté dans la section précédente 6.2, l'application du principe de linéarisation est sujette à l'estimation des paramètres non-linéaires,  $\eta_k$ , caractérisant une discontinuité dans la  $k$ -ième ligne d'une image. La théorie de l'estimation statistique n'étant pas le centre des travaux présentés, un rappel succinct est donné dans l'annexe 6.2, le lecteur intéressé pourra en outre consulter les ouvrages [87, 107, 165].

Un rapide état de l'art permet de constater que le problème d'estimation de la position d'une discontinuité abrupte (ou de l'instant d'occurrence d'un changement) a été longuement étudié dans la littérature [166, 167, 168]. En revanche, le problème d'estimation statistique d'une discontinuité convoluée n'a donné lieu qu'à quelques publications dont notamment [169, Chap.3]. De même, l'estimation de la position d'une discontinuité lorsque les paramètres du signal avant et après la discontinuité ne sont pas connus, ou partiellement connus, n'a été que peu étudiée dans la littérature [170]. Or dans le cadre des travaux présentés, quatre difficultés majeures sont à considérer :

- les moyennes des signaux (avant et après la discontinuité) ne peuvent être considérées comme constantes, la modélisation de ces signaux est, en outre, délicate et fait l'objet du chapitre 5, voir plus précisément section 5.2 ;

- la discontinuité qu’il est souhaitable de détecter n’est pas abrupte mais convoluée car comme détaillé dans la section 4.1, le canal d’acquisition agit comme un filtre (passe-bas) ;
- compte-tenu du modèle décrit dans la section 4.3, le bruit d’acquisition est non-stationnaire : la variance du bruit Gaussien n’est pas constante mais dépendante de l’intensité du signal ;
- enfin, les paramètres des signaux avant et après acquisition ne sont pas connus et l’estimation de ces derniers n’est pas évidente dans ce contexte.

Le cumul de ces difficultés théoriques ne permet pas de proposer une méthode d’estimation statistique, ni des paramètres non linéaires  $\eta_k$  ni du contenu  $\theta_k$ , dont les critères qualitatifs sont établis. En outre, les méthodes d’estimation statistique des paramètres d’une discontinuité convoluée [169, 171] nécessitent la résolution d’un problème d’optimisation ; cette approche est à proscrire dans le cadre des travaux présentés en raison de son calculatoire important.

Aussi, bien que les approches proposées dans la littérature issues de l’estimation de changement abrupte dans le paramètre de distribution sont très intéressantes, elles ne sont guères exploitables dans le cadre des travaux présentés dans ce manuscrit.

Dans le domaine du traitement des images, l’estimation de la position des contours des objets a été longuement étudiée dans la littérature, voir notamment [172, chap.5] pour une introduction détaillée. Les méthodes issues du traitement d’images sont bien différentes des méthodes de détection / localisation de changement de paramètre de distribution issues du domaine statistique. En effet, les problèmes considérés, les finalités et les moyens utilisés sont différents dans les deux cas : dans le cas du traitement des images, il est souhaité détecter des changements importants dans la moyenne (contenu des l’image).

La grande majorité des méthodes de détection de contours issues du traitement des images reposent sur un filtrage bi-dimensionnel passe-haut de l’image analysée (et, le cas échéant par un filtrage du résultat) ce qui présente l’intérêt d’être d’un coût calculatoire très faible ; le lecteur intéressé par des informations plus détaillées pourra consulter [172, chap.5][173, chap.5][174]. Ce type d’algorithme de détection de contours n’a pas été considéré dans le cadre des présents travaux en raison de l’utilisation de l’aspect 2D des images, alors que la modélisation retenue décrit les images en 1D, par ligne ou par colonne.

Les méthodes variationnelles pour le traitement des images ont également été appliquées à la détection de contours [175, 176]. L’approche variationnelle est très intéressante notamment en raison de son fondement assez proche du modèle de la scène proposée dans [143] ; ce modèle étant lui-même utilisé dans la section 5.1.1. Les méthodes itératives d’approximation du contenu des images et de détection de contours qui en découlent sont généralement calculatoirement coûteuses et sont donc, au même titre que les approches purement statistiques, à proscrire.

Enfin, dans la continuité de l’approche proposée dans [177], les méthodes de décomposition multi-résolution des signaux (et notamment la transformation en ondelettes) ont également été utilisées pour permettre la détection de changement dans un signal. Cette approche est particulièrement intéressante dans le cadre des travaux présentés car le signal considéré est analysé en 1D (bien que des extensions pour le cas 2D existent [178]) leur coût calculatoire faible et l’extension au cas de discontinuités convoluées a été proposée [144, 179].

Pour ces raisons cette approche a été choisie pour l’estimation des paramètres  $\eta_k$  caractérisant une discontinuité.

*On trouvera d’ailleurs dans l’introduction de [172, chap.5] la définition d’un contour comme «un changement local significatif de l’intensité d’une image usuellement associée à une discontinuité».*

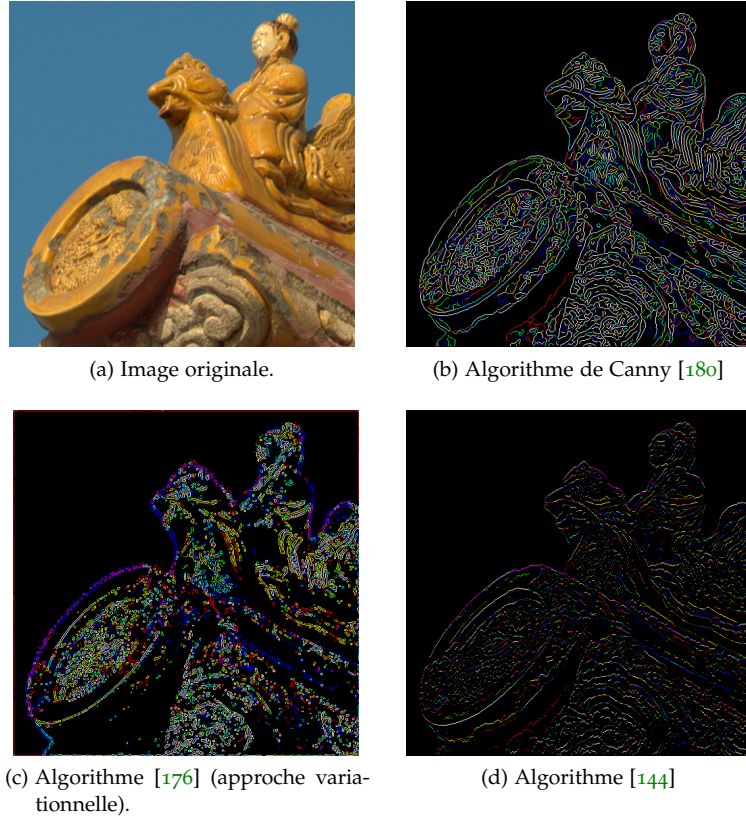


FIGURE 6.3: Illustration graphique de la détection de contours.

### 6.3.1 Algorithmes utilisés pour l'estimation des discontinuités

La méthode d'estimation des paramètres  $(v_{k,i}, \zeta_{k,i})$ , caractérisant une discontinuité, implémentés dans le cadre des travaux présentés s'appuie sur [144]. Le principe de cette approche repose sur l'analyse multi-résolution d'un signal par application d'un banc de filtre. Généralement le filtrage passe-haut d'un signal en vue de détecter les éventuelles discontinuités est soumis au compromis suivant ; un large noyau offre une meilleure robustesse vis-à-vis du bruit et un noyau dont le support est très restreint permet de détecter de faibles changements dans la valeur du signal. L'application d'un banc de filtre est une solution usuelle, pour détecter les discontinuités dans une image, permettant d'utiliser les informations après filtrage avec différentes échelles.

Compte tenu de l'expression de la modélisation du  $\theta_k$  proposée par (5.7) :

$$\theta_k(x) = \theta_{k,c}(x) + \theta_{k,s}(x) = \sum_{j=0}^p c_j x^j + \sum_{i=1}^{r_k} u_i \Phi \left( \frac{x - v_{k,i}}{\zeta_{k,i}} \right), \quad (6.10)$$

les auteurs de [144] suggèrent de filtrer le signal analysée par convolution avec un noyau défini par la  $j$ -ième dérivée de la fonction Gaussienne, notée dans la suite :

$$\varphi_{\zeta_0}^{(j)}(x) = \frac{d^j}{dx^j} \frac{1}{\zeta_0} \varphi \left( \frac{x}{\zeta_0} \right) = \left( \frac{1}{\zeta_0} P_j^h(x) \right) \left( \frac{1}{\zeta_0} \varphi \left( \frac{x}{\zeta_0} \right) \right) \quad (6.11)$$

où  $\zeta_0$  représente le paramètre d'échelle utilisé pour la détection des discontinuités (donc connu) et la fonction  $P_k^h$  est définie à partir des polynômes de Hermite [181, chap.22] par :  $P_j^h(x) = \frac{(-1)^j}{2^{j/2}} H_j \left( \frac{x}{\sqrt{2}} \right)$ .

Ce choix du noyau de convolution est guidé par l'expression simple du résultat  $f(x) = \theta_k(x) * \varphi_{\zeta_0}^{(j)}(x)$  qu'un rapide calcul permet d'obtenir :

$$f(x) = \left[ \theta_k * \varphi_{\zeta_0}^{(j)} \right] (x) = \left[ \theta_{k,c} * \varphi_{\zeta_0}^{(j)} \right] (x) + \sum_{i=1}^{r_k} u_{k,i} \varphi_{\zeta_{k,i}}^{(j-1)}(x - v_{k,i}). \quad (6.12)$$

où  $\zeta_{k,i} = (\zeta_0^2 + \zeta_{k,i}^2)^{1/2}$ .

Le premier terme de (6.12) ne peut être explicité de manière générale, néanmoins :

- comme décrit dans la section 5.2, compte-tenu que de la régularité de la fonction  $\theta_{k,c}$  et du fait la convolution par  $\varphi_{\zeta_0}^{(j)}$  agit comme un filtre passe-haut, il semble raisonnable de conjecturer que la contribution du premier terme de (6.12) est négligeable, comparativement au second ;
- en particulier, dans le cas où  $\theta_{k,c}$  est un polynôme de degré  $p$ , en vertu de la règle d'intégration de Leibniz on a  $\theta_{k,c} * \varphi_{\zeta_0}^{(j)} = \theta_{k,c}^{(j)} * \varphi_{\zeta_0}$ . Il en découle que pour  $p < j$  le premier de (6.12)  $\theta_{k,c} * \varphi_{\zeta_0}^{(j)}$  est nul pour tout  $x$ .

A partir de ces résultats, les auteurs de [144] proposent d'étudier le cas pratique d'un signal discret  $\mathbf{z}_k = (z_{k,1}, \dots, z_{k,N})^T$  (correspondant à une ligne ou un segment de l'image, voir section 5.1.3) défini par :

$$\mathbf{z}_k = \boldsymbol{\theta}_k + \boldsymbol{\zeta}_k \quad \text{avec} \quad \boldsymbol{\theta}_k = \boldsymbol{\theta}_{k,c} + \boldsymbol{\theta}_{k,s}, \quad \boldsymbol{\zeta}_k \sim \mathcal{N}(0, \Sigma_k), \quad (6.13)$$

où, compte-tenu de l'hétéroscédasticité du bruit d'acquisition,  $\Sigma_k = \text{diag}(\sigma_{k,1}^2, \dots, \sigma_{k,L}^2)$  et  $\boldsymbol{\theta}$  correspond au signal  $\theta$  discrétisé  $\boldsymbol{\theta} = (\theta_{k,1}, \dots, \theta_{k,L})^T$ , avec dans le cas considéré  $\forall l \in \{1, \dots, L\}$ ,  $\theta_n = \int_{x_n - \Delta/2}^{x_n + \Delta/2} \theta(x) dx$  et  $\Delta = x_n - x_{n-1}$ . De manière analogue au cas continu, en notant  $\varphi_{\zeta_0}^{(j)}$  le noyau de convolution discrétisé il est considéré dans [144] que l'on a :

$$\begin{aligned} \mathbf{f}_k &= \mathbf{z}_k * \varphi_{\zeta_0}^{(j)} = \boldsymbol{\theta}_{k,c} * \varphi_{\zeta_0}^{(j)} + \boldsymbol{\theta}_{k,s} * \varphi_{\zeta_0}^{(j)} + \boldsymbol{\zeta}_k * \varphi_{\zeta_0}^{(j)} \\ &\approx \left[ \boldsymbol{\theta}_{k,c} * \varphi_{\zeta_0}^{(j)} \right] + \sum_{i=1}^{r_k} u_{k,i} \varphi_{\zeta_{k,i}}^{(j-1)}(x - v_{k,i}) + \left[ \boldsymbol{\zeta}_k * \varphi_{\zeta_0}^{(j)} \right], \end{aligned}$$

La figure 6.4 propose une illustration de quelques résultats de convolution tel que fourni par (6.14). Le premier terme de (6.14) est inconnu mais peut être négligé sous les conditions précédemment précisés. Dans [144], il est proposé d'estimer la localisation  $v_k$  et de l'intensité  $\mathbf{u}_k$  des  $r_k$  discontinuités par maximum de vraisemblance. Dans la cadre des travaux présentés l'estimation du paramètre  $\mathbf{u}_k$  ne présente pas d'intérêt ; à l'inverse l'application de l'approche par linéarisation nécessaire l'estimation du paramètre  $\boldsymbol{\zeta}_k = (\zeta_{k,1}, \dots, \zeta_{k,r_k})$ . En outre, l'estimation par maximum de vraisemblance nécessite de connaître le nombre  $r_k$  de discontinuités et demande, ici encore, l'utilisation d'algorithmes d'optimisation, décrits dans la section 6.1.2, dont le temps de calcul est prohibitif dans le cadre des travaux présentés.

L'approche proposée inspirée des méthodes d'analyse multi-résolution pour la détection des discontinuités [178, 182] est de simplement seuiliser les signaux résultant de la convolution  $\mathbf{z}_k * \varphi_{\zeta_0}^{(j)}$ . L'estimation (dans un sens qui n'est pas statistique) de la localisation des discontinuités est ensuite basée sur la localisation des maxima locaux et, pour chacune des discontinuités, l'estimation du paramètre  $\zeta_{k,i}$  est obtenu par recherche du premier passage à zéro du résultat de la convolution  $\mathbf{z}_k * \varphi_{\zeta_0}^{(j)}$ . Plus précisément, l'implémentation algorithmique a été réalisée en utilisant la 3-ième dérivée  $\varphi_{\zeta_0}^{(3)}$  en raison du fait que  $u_{k,i} \Phi\left(\frac{x - v_{k,i}}{\zeta_{k,i}}\right) * \varphi_{\zeta_0}^{(3)}(x) = u_{k,i} \varphi_{\zeta_{k,i}}^{(2)}(x - v_{k,i})$ , or un rapide calcul montre que les racines de l'équation  $\varphi_{\zeta_{k,i}}^{(2)}(x - v_{k,i}) = 0$  sont données par  $x = v_{k,i} \pm \widetilde{\zeta}_{k,i} = v_{k,i} \pm (\zeta_{k,i}^2 + \zeta_0^2)^{1/2}$ .

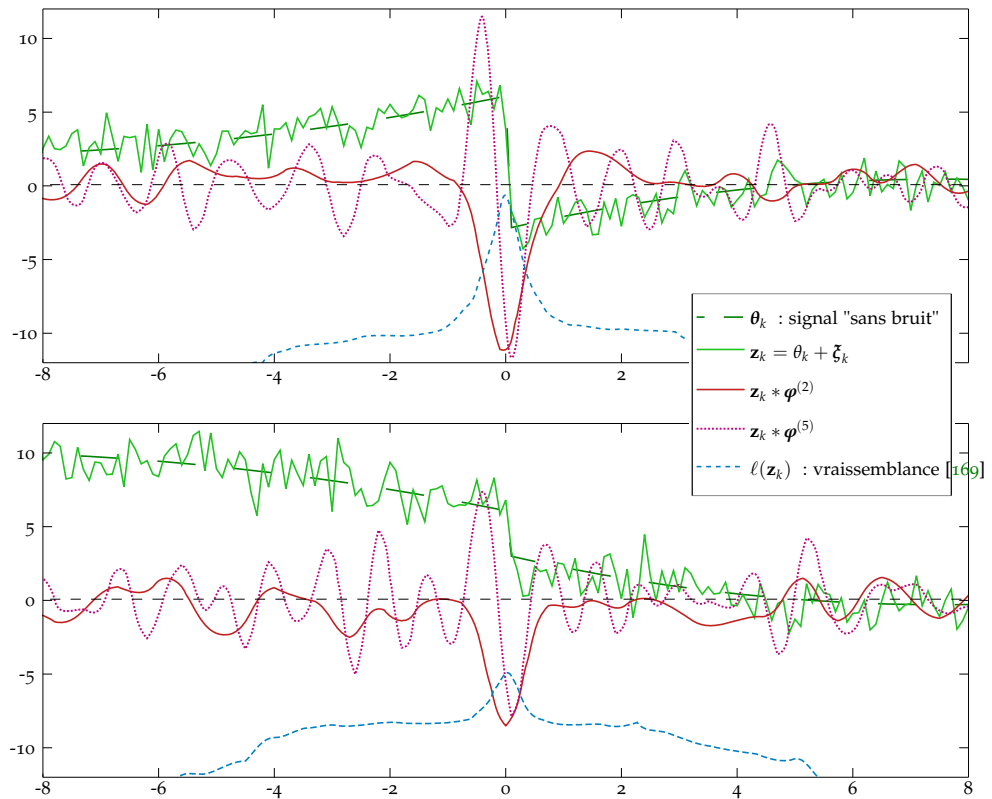


FIGURE 6.4: Illustration graphique d'un exemple de signal à analyser (simulation), localisation des discontinuités convolution des données par  $\varphi_{\zeta_0}^{(j)}$  et par maximum de vraisemblance.

En pratique, les trois "échelles" suivantes ont été arbitrairement choisies  $\zeta_0 = \{1; 2; 4\}$ , et les estimations  $\hat{x}_{k,i}$  et  $\hat{\zeta}_{k,i}$  sont obtenues par moyennes des résultats obtenues avec les trois échelles différentes.

Il est important de noter que cette estimation du paramètre  $\eta_k$  n'a aucune prétention qualitative et a simplement pour but de permettre l'application de l'approche par linéarisation proposée. Le but est donc davantage de proposer une première ébauche de ce qui est davantage une "astuce" pour permettre l'estimation  $\hat{\eta}_k$  de  $\eta_k$  de façon simple, les convolutions étant les opérations les plus coûteuses calculatoirement. Tout autre "technique" d'estimation de  $\eta_k$  peut être substituée à celle proposée avec, potentiellement, des meilleurs résultats *in fine*.

La principale difficulté ici est, qu'à notre connaissance, lorsque les discontinuités sont convoluées et leur nombre inconnue, aucune méthode existante permet d'estimer le paramètre  $\eta_k$  de façon simple ou avec des résultats qualitatifs sur la qualité de l'estimation. Dans notre cas, il faudrait idéalement les deux, l'estimateur idéal doit être simple et avec des résultats qualitatifs. Cela explique le recours, pour l'application numérique, à cette technique *ad hoc* d'estimation.

Notons enfin que dans l'annexe 6.5.1, la borne de Rao-Cramer, voir théorème A.4, sur l'estimation des paramètres de discontinuités  $\eta_k$  est calculée pour le cas d'une unique discontinuité présente dans un segment. En outre, les tableaux 6.4 et 6.5, de la même annexe 6.5.1, fournissent quelques résultats sur la qualité de l'estimation  $\hat{\eta}_k$  de  $\eta_k$  avec l'algorithme utilisé.

### 6.3.2 Étude qualitative de la linéarisation du modèle local

Afin d'étudier la légitimité de l'approche choisie de linéarisation du modèle (5.10) - (5.11) du contenu des images, l'analyse des courbures paramétrique et intrinsèque décrites dans la section 6.2 est incontournable. En outre l'impact de cette procédure de linéarisation du modèle requiert l'étude des courbures du modèle non-linéaire considéré.

Pour des raisons de clarté, il est proposé d'étudier le cas d'un segment  $\mathbf{z}_k = Q(\boldsymbol{\theta}_k + \boldsymbol{\zeta}_k)$  issu de l'image présentant une unique discontinuité. L'occurrence du bruit  $\boldsymbol{\zeta}_k$  et de la quantification  $Q(\cdot)$  n'intervenant pas dans l'analyse de la linéarisation du modèle, le signal considéré est exempt de ces phénomènes :

$$\boldsymbol{\theta}_k = \mathbf{H}\mathbf{c}_k + \mathbf{f}(\boldsymbol{\eta}_k)\mathbf{u}_k. \quad (6.14)$$

où l'expression des composantes du vecteur  $\mathbf{f}(\boldsymbol{\eta}_k)$  est, consécutivement à la discrétisation par l'intégration, donnée par  $f(x_l; v_{k,i}, \zeta_{k,i}) = \int_{x_l-1/2}^{x_l+1/2} \Phi\left(\frac{x-v_{k,i}}{\zeta_{k,i}}\right) dx$ , voir section 5.2.1 pour plus de détails. .

Le calcul des courbures du modèle non-linéaire représentant le signal discret  $\boldsymbol{\theta}_k$  est théoriquement possible en application directe de la méthodologie présentée dans la section 6.2. Cependant, ces calculs algébriques rendent l'expression de ces courbures difficilement exploitable en pratique. Pour contourner cette difficulté, il est proposé de considérer le signal  $\theta_k(x)$  avant discrétisation par intégration spatiale :

$$\begin{aligned} \theta_k(x) &= \theta_{k,c}(x) + \mathbf{u}_k f(x; \boldsymbol{\eta}_k) \\ &= \underbrace{\sum_{j=0}^p c_{k,j} x^j}_{\text{partie linéaire} \simeq \mathbf{H}\mathbf{c}_k} + \underbrace{\sum_{i=1}^{r_k} u_{k,i} \Phi\left(\frac{x-v_{k,i}}{\zeta_{k,i}}\right)}_{\text{partie non-linéaire} \simeq \mathbf{f}(\boldsymbol{\eta}_k)\mathbf{u}_k}. \end{aligned} \quad (6.15)$$

Dans le cas considéré où une unique discontinuité (au plus) est présente dans un segment, il découle immédiatement que  $r_k = 1$ ,  $\boldsymbol{\eta}_k = (v_{k,1}, \zeta_{k,1})$  et la non-linéarité du modèle réside dans la fonction  $f(x; \boldsymbol{\eta}_k) = \Phi\left(\frac{x-v_{k,1}}{\zeta_{k,1}}\right)$ .

**Remarque 6.4.** Dans un souci évident de clarté, l'indice  $i$  des discontinuités est omis dans la suite de cette section puisqu'une unique discontinuité est supposée éventuellement présente dans chacun des segments.

Le calcul des courbures, paramétrique et intrinsèque, nécessitant le calcul des dérivées partielles de la fonction  $f(x; \boldsymbol{\eta}_k)$  jusqu'au second-ordre, par analogie avec la section 6.2, les notations suivantes seront utilisées pour alléger les expressions :

$$\begin{aligned} f_{\boldsymbol{\eta}_k} &= f(x; \boldsymbol{\eta}_k) & \dot{f}_{v_k} &= \frac{\partial}{\partial v_k} \Phi\left(\frac{x-v_k}{\zeta_k}\right) & \dot{f}_{\zeta_k} &= \frac{\partial}{\partial \zeta_k} \Phi\left(\frac{x-v_k}{\zeta_k}\right) \\ \ddot{f}_{v_k^2} &= \frac{\partial^2}{\partial v_k^2} \Phi\left(\frac{x-v_k}{\zeta_k}\right) & \ddot{f}_{\zeta_k^2} &= \frac{\partial^2}{\partial \zeta_k^2} \Phi\left(\frac{x-v_k}{\zeta_k}\right) & \ddot{f}_{v_k \zeta_k} &= \frac{\partial^2}{\partial v_k \partial \zeta_k} \Phi\left(\frac{x-v_k}{\zeta_k}\right) \end{aligned} \quad (6.16)$$

**Remarque 6.5.** Le calcul des courbures est en outre simplifié en considérant que la fonction  $f_{\boldsymbol{\eta}_k}$  est définie sur  $\mathbb{R}$ , afin d'éviter l'occurrence de phénomènes de bords qui, s'ils ne modifient pas considérablement les résultats présentés, complexifient grandement l'expression analytique des calculs. Bien que cette condition, ainsi que la présence supposée d'une unique discontinuité, sont assez irréalistes en pratique, ces contraintes peuvent être relaxées en considérant que la fonction  $f_{\boldsymbol{\eta}_k}$  est constante en dehors d'un voisinage  $K$  autour de  $v_k$ , défini par exemple par  $K = \{x \in \mathbb{R} \mid |x-v_k/\zeta_k| < x_{max}\}$ , i. e. les dérivées partielles  $\dot{f}_{v_k}$  et  $\dot{f}_{\zeta_k}$  sont non-nulles uniquement sur  $K$ .

*Le cas d'un segment présentant plusieurs discontinuités peut être analysé de manière analogue au prix de notations moins simples.*

*$x_l$  vs  $v_{k,i}$  attention confusions : noter  $v_{k,i}$  la localisation des discontinuités !*



	$\dot{f}_{v_k}$	$\dot{f}_{\zeta_k}$
$\ddot{f}_{v_k^2}$	$\langle \dot{f}_{v_k}, \ddot{f}_{v_k^2} \rangle = 0$	$\langle \dot{f}_{\zeta_k}, \ddot{f}_{v_k^2} \rangle = \frac{1}{\theta} \ \dot{f}_{\zeta_k}\ _{\mathcal{L}^2}^2$
$\ddot{f}_{v_k \zeta_k}$	$\langle \dot{f}_{v_k}, \ddot{f}_{v_k \zeta_k} \rangle = \frac{-1}{\theta} \ \dot{f}_{v_k}\ _{\mathcal{L}^2}^2 - \frac{1}{4\sqrt{\pi\theta}}$	$\langle \dot{f}_{\zeta_k}, \ddot{f}_{v_k \zeta_k} \rangle = 0$
$\ddot{f}_{\zeta_k^2}$	$\langle \dot{f}_{v_k}, \ddot{f}_{\zeta_k^2} \rangle = 0$	$\langle \dot{f}_{\zeta_k}, \ddot{f}_{\zeta_k^2} \rangle = \frac{3}{8\sqrt{\pi\theta^2}}$

TABLE 6.2: Normes des composantes tangentielles des dérivées partielles de second-ordre.

La notation  $\|f\|_{\mathcal{L}^p}$  représente ici le produit scalaire usuel de  $\mathcal{L}^p$  sur  $R : (\int_R |f(x)|^p)^{1/p}$ .

Les produits scalaires des dérivées partielles de premier et de second-ordre, issus d'un calcul direct, sont résumés dans le tableau 6.2.

Enfin, les contributions normales et tangentielles de chacune des dérivées seconde est immédiat en calculant la norme  $\|\cdot\|_{\mathcal{L}^p}$  des dérivées de premier et de second-ordre. Le cas de  $\ddot{f}_{v_k^2}$  est immédiat puisque l'on a  $\ddot{f}_{v_k^2} = \dot{f}_{\zeta_k} / \zeta_k$ ; pour les autres dérivées partielles de second-ordre,  $\ddot{f}_{v_k \zeta_k}$  et  $\ddot{f}_{\zeta_k^2}$  on a :

$$\begin{aligned} \ddot{f}_{\zeta_k^2}^T &= \frac{\langle \dot{f}_{\zeta_k}, \ddot{f}_{\zeta_k^2} \rangle}{\|\dot{f}_{\zeta_k}\|_{\mathcal{L}^2}^2} \dot{f}_{\zeta_k} = \frac{1}{16\sqrt{\pi\zeta_k^3}}, \quad \|\ddot{f}_{\zeta_k^2}^N\|_{\mathcal{L}^2}^2 = \|\ddot{f}_{\zeta_k^2}\|_{\mathcal{L}^2}^2 - \|\ddot{f}_{\zeta_k^2}^T\|_{\mathcal{L}^2}^2 = \frac{6}{16\sqrt{\pi\zeta_k^3}}, \\ \ddot{f}_{v_k \zeta_k}^T &= \frac{\langle \dot{f}_{v_k}, \ddot{f}_{v_k \zeta_k} \rangle}{\|\dot{f}_{v_k}\|_{\mathcal{L}^2}^2} \dot{f}_{v_k} = \frac{1}{8\sqrt{\pi\zeta_k^3}}, \quad \|\ddot{f}_{v_k \zeta_k}^N\|_{\mathcal{L}^2}^2 = \|\ddot{f}_{v_k \zeta_k}\|_{\mathcal{L}^2}^2 - \|\ddot{f}_{v_k \zeta_k}^T\|_{\mathcal{L}^2}^2 = \frac{2}{8\sqrt{\pi\zeta_k^3}}. \end{aligned} \quad (6.17)$$

En notant, de manière analogue à ce qui proposé dans [162],  $(\eta_k - \hat{\eta}_k) = (\delta_{v_k}, \delta_{\zeta_k})$ , le terme de second-ordre du développement (6.5), base de linéarisation du modèle, peut être ré-écrit :

$$\frac{1}{2}(\eta_k - \hat{\eta}_k)^T \ddot{\mathbf{f}}_{\hat{\eta}_k} (\eta_k - \hat{\eta}_k) = \frac{1}{2} \delta_{v_k}^2 \ddot{f}_{v_k^2} + \frac{1}{2} \delta_{\zeta_k}^2 \ddot{f}_{\zeta_k} + \delta_{v_k} \delta_{\zeta_k} \ddot{f}_{v_k \zeta_k}. \quad (6.18)$$

La composante tangentielle  $\ddot{\mathbf{f}}_{\hat{\eta}_k}^T[:, i, j] = \mathbf{P}_{\hat{\eta}_k} \ddot{\mathbf{f}}_{\hat{\eta}_k}[:, i, j]$  du terme de second-ordre, voir relations 6.9 est immédiatement donnée par :

$$\ddot{\mathbf{f}}_{\hat{\eta}_k}^T = \frac{1}{2} (4\sqrt{\pi\zeta_k^3})^{-\frac{1}{2}} \left( \delta_{v_k}^2 + \frac{\delta_{\zeta_k}^2}{4} \right) \dot{f}_{\zeta_k} + \delta_{\zeta_k} \delta_{v_k} (8\sqrt{\pi\zeta_k^3})^{-\frac{1}{2}} \dot{f}_{v_k}.$$

L'expression de la composante normale  $\ddot{\mathbf{f}}_{\hat{\eta}_k}^N$  est plus délicate, cette dernière étant en effet définie indirectement comme la soustraction de la composante tangentielle  $\ddot{\mathbf{f}}_{\hat{\eta}_k}^T$  au terme de second-ordre 6.18. Il peut être pertinent de caractériser la composante normale afin de fournir une expression analytique du biais des résidus obtenu par l'estimateur découlant de la linéarisation "approximative" du modèle autour du point de fonctionnement  $\hat{\eta}_k \neq \eta_k$ . Cela n'a qu'un intérêt limité dans le cadre des travaux présentés compte-tenu du nombre important de discontinuités et de la grande variabilité des paramètres  $\eta_k$  pour  $k = \{1, \dots, K\}$ .

En revanche, les résultats présentés dans la section 6.4 reposent sur la norme de l'erreur résiduelle, liée à la linéarisation du modèle; or les segments de l'image analysées sont des vecteurs, alors que les calculs présentés sont réalisées en considérant modélisant la discontinuité comme une fonction. En rappelant que l'erreur due à la linéarisation du modèle est donnée par :  $\mathbf{b}_k = \mathbf{P}_{\hat{\eta}_k}^\perp \boldsymbol{\theta}_k$  (6.8), la proposition 6.1 permet d'explicitement l'erreur résiduelle à partir des résultats précédents.

**Proposition 6.1.** *En utilisant les notations précédemment introduites et en considérant que les effets de bords sont absents, i.e. que  $\forall x_l \notin [0, L], \dot{f}_{v_k}(x_l) = 0$  et  $\dot{f}_{c_k}(x_l) = 0$ , alors, il découle de la définition (6.8) de  $\mathbf{b}_k$  et des relations (6.17) et d'un calcul direct, que :*

$$\sum_{l=1}^L \mathbf{b}_{k,l} = \int_0^L \dot{\mathbf{f}}_{\hat{\eta}_k}^N(x) dx = 0 \quad (6.19)$$

$$\|\mathbf{b}_k\|_2^2 = \sum_{l=1}^L \mathbf{b}_{k,l}^2 \leq \left( \frac{1}{8\sqrt{\pi c_k^3}} \right) \left( 2|\delta_{v_k} \delta_{c_k}| + 3/2\delta_{c_k}^2 \right) \quad (6.20)$$

$$\leq \left( \frac{2}{8\sqrt{\pi c_k^3}} \right) \|\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k\|_1^2, \quad (6.21)$$

Il faut souligner que la linéarisation du modèle autour de l'estimation  $\hat{\boldsymbol{\eta}}_k$  de  $\boldsymbol{\eta}_k$  introduit une erreur déterministe et bornée, mais cette dernière inconnue. L'intérêt de cette borne est de permettre de linéariser le modèle (6.1) pour obtenir une l'estimation  $\hat{\boldsymbol{\theta}}_k$  de  $\boldsymbol{\theta}_k$ , voir relation (6.7), ensuite utilisée dans un test statistique dont les performances sont alors bornées.

## 6.4 Détection statistique d'informations cachées : impact de la linéarisation sur les performances du test statistique

### 6.4.1 Contenu de l'image inconnu : test statistique proposé

Dans les travaux présentés dans ce manuscrit il est proposé d'utiliser le test statistique avec conservation du LSB, approximation du test du RV, présenté dans la section 3.5. Le contenu de l'image  $\boldsymbol{\theta}$  n'étant pas connu, l'approche avancée dès la section 3.4 consiste à estimer ce paramètre inconnu ce qui conduit à un test similaire dans son approche au test du RVG.

Plus précisément, le test pratique considéré est décrit par la fonction de décision, voir relations (3.42), (3.62) et (3.65), suivante :

$$\hat{\delta}(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \log(\hat{\Lambda}(\mathbf{Z})) = \sum_{n=1}^N \log(\hat{\Lambda}(\mathbf{z}_k)) < \tau_{\alpha_0} \\ \mathcal{H}_1 & \text{si } \log(\hat{\Lambda}(\mathbf{Z})) = \sum_{n=1}^N \log(\hat{\Lambda}(\mathbf{z}_k)) \geq \tau_{\alpha_0} \end{cases} \quad (6.22)$$

où la valeur approchée du RV  $\log(\hat{\Lambda}(\mathbf{Z}))$  est donnée, selon (3.53) et (3.66), par :

$$\log(\hat{\Lambda}(\mathbf{Z})) = \frac{\bar{\sigma}}{\sqrt{K}} \sum_{k=1}^K \frac{(\mathbf{P}_{\hat{\boldsymbol{\eta}}_k}^{\perp} \mathbf{z}_k)^T (\mathbf{z}_k - \bar{\mathbf{z}}_k)}{\sigma_k^2 \sqrt{(L-p-3r_k)}} = \frac{\bar{\sigma}}{\sqrt{K}} \sum_{k=1}^K \sum_{l=1}^L \frac{(z_{k,l} - \hat{\theta}_{k,l})(z_{k,l} - \bar{z}_{k,l})}{\sigma_k^2 \sqrt{(L-p-3r_k)}}, \quad (6.23)$$

où ici encore,  $\frac{1}{\bar{\sigma}^2} = \frac{1}{N} \sum_{n=1}^N \frac{1}{\sigma_n^2}$  et  $r_k$  représente le nombre de discontinuités dans le  $k$ -ième segment.

Comme cela a été évoqué dans les chapitres 4 et 5, la principale difficulté dans l'application du test  $\hat{\delta}(\mathbf{Z})$  (6.22), ou plus précisément dans le calcul du log-RV  $\log(\hat{\Lambda}(\mathbf{Z}))$  (6.23), réside dans l'estimation  $\hat{\boldsymbol{\theta}}_k$  du contenu du  $k$ -ième segment  $\boldsymbol{\theta}_k$ . Au cours de la section 6.2, une méthode d'estimation fondée sur la linéarisation du modèle du contenu  $\boldsymbol{\theta}_k$  (5.10) - (5.11) a été proposée. Si l'impact de la linéarisation de ce modèle sur le biais des résidus a été étudié dans la section 6.3.2, le plus important demeure, dans le cadre des travaux présentés, de caractériser les performances statistiques du test (6.22) reposant sur la procédure d'estimation proposée afin :

1. de quantifier la perte de puissance du test  $\hat{\delta}(\mathbf{Z})$  liée à l'estimation  $\hat{\boldsymbol{\theta}}_k$ , comparativement au test AUPP  $\bar{\delta}(\mathbf{Z})$  (3.65) dans le cas où le contenu  $\boldsymbol{\theta}_k$  est connu ;

- de fournir une expression analytique du seuil de décision  $\tau_{\alpha_0}$  permettant le respect d'une probabilité de fausse alarme  $\alpha_0$  préalablement fixée et de caractériser la puissance du test (6.22).

**Remarque 6.6.** Il est notable que l'expression du log-RV  $\log(\hat{\Lambda}(\mathbf{Z}))$  (6.23) est très similaire au log-RV  $\log(\hat{\Lambda}(\mathbf{Z}))$  (3.66) (3.68) détaillé dans la section 3.5. La différence fondamentale entre les deux tests réside dans l'exploitation du modèle ad-hoc des images naturelles présenté dans les chapitres 4 et 5. L'exploitation de ce modèle permet, outre une augmentation possible de la taille des segments utilisés, une estimation  $\hat{\theta}_k$  réaliste du contenu alors qu'en pratique la paramétrisation linéaire des images (3.38) est assez irréaliste.

La figure 6.5 résume le principe du test statistique proposé pour la détection d'informations cachées dans les LSB d'une image naturelle. D'un point de vue algorithmique,

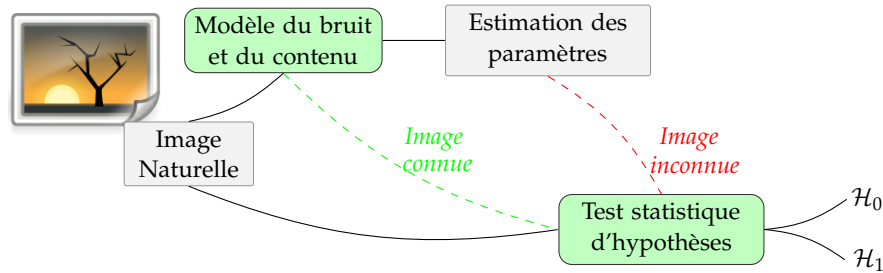


FIGURE 6.5: Diagramme représentant l'application pratique du test proposé.

les différentes étapes du test, illustrées dans la figure 6.5, peuvent être résumées de la façon suivante :

- l'image naturelle inconnue est analysée ligne par ligne (ou de manière équivalente colonne par colonne) ;
- dans un premier temps, une estimation "approximative" des paramètres des discontinuités,  $\eta_{k,i}$ , est réalisée : dans les résultats numériques présentés dans le chapitre 7, la méthodologie issue de [144, 179] détaillée dans la section 6.3.1 est utilisée ;
- les estimations  $\hat{\eta}_{k,i}$  des paramètres  $\eta_{k,i}$  sont ensuite utilisées comme "point de fonctionnement" autour duquel le modèle (5.10) - (5.11) est linéarisé selon la méthodologie décrite dans la section 6.2, voir également section 6.3.2 pour une étude du biais d'estimation engendré par la linéarisation du modèle ;
- l'estimation  $\hat{\theta}_k$  du contenu de chacun des segments  $\theta_k$  obtenue en utilisant le modèle linéarisé (6.7) est utilisée dans le calcul du log-RV (6.23), et donc *in fine* dans le test statistique (6.22) ;

## 6.4.2 Calcul des performances du test quasi-optimal proposé

L'approche asymptotique proposée dans les sections 3.2,3.4 et 3.5.1 est ici encore utilisée pour calculer les performances du test. En vertu du théorème 3.1 de la limite centrale de Lindeberg il est proposé de fournir une expression simple de la distribution de probabilité du rapport de vraisemblance  $\log(\hat{\Lambda}(\mathbf{Z}))$  (6.23), et donc *in fine* de calculer simplement la performance asymptotique du test  $\hat{\delta}(\mathbf{Z})$  (6.22). Cette approche nécessite le calcul des premiers moments  $\mathbb{E}_R [\log(\hat{\Lambda}(z_n))]$  et  $\text{Var}_R [\log(\hat{\Lambda}(z_n))]$ .

En supposant qu'il existe, pour chacune des discontinuités, un estimateur  $\hat{\eta}_k$  de  $\eta_k$  tel que  $\|\eta_k - \hat{\eta}_k\|_1 \leq \vartheta$  où  $\vartheta$  est une constante positive ; *i.e.* statistiquement cela revient à supposer que cet éventuel estimateur vérifie  $\mathbb{P}[\|\eta_k - \hat{\eta}_k\|_1 > \vartheta] = 0$ . Alors,

en utilisant l'estimateur  $\hat{\theta}_k$  (6.7) de  $\theta_k$ , la distribution du log-RV  $\log(\hat{\Lambda}(\mathbf{Z}))$  (6.23) converge asymptotiquement, quel que soit  $R \in [0; 1]$  vers :

$$\log(\hat{\Lambda}(\mathbf{Z})) \rightsquigarrow \mathcal{N}(\sqrt{\kappa}\varrho; 1 + \gamma + b). \quad (6.24)$$

où  $\gamma = 1/\sigma^2 R/2(1 - R/2)$ ,  $\kappa = K(L - p) - \sum_{k=1}^K r_k$  représente le nombre de degré de liberté de la modélisation de l'ensemble de l'image et  $b$  est un accroissement de la variance due à la linéarisation du modèle vérifiant, voir annexe 6.5.2 pour plus de détails :

$$0 \leq b \leq \frac{o(\vartheta^2)}{\sigma^2(L - p - 3)} = b_{\max}.$$

Le résultat (6.24) permet immédiatement, au travers de la proposition 6.2 et du théorème 6.1, un encadrement de la puissance du test  $\hat{\delta}$  (6.22) sous la contrainte de respect d'une probabilité de fausse alarme  $\alpha_0$  préalablement fixée.

**Proposition 6.2.** *Un calcul direct permet de montrer que pour tout  $\alpha_0 \in ]0; 0.5]$ , en fixant*

$$\hat{\tau}_{\alpha_0} = \Phi^{-1}(1 - \alpha_0)\sqrt{1 + \gamma + b_{\max}} \geq \Phi^{-1}(1 - \alpha_0)\sqrt{1 + \gamma + b}, \quad (6.25)$$

où  $\Phi^{-1}(\cdot)$  est la fonction réciproque de la fonction de répartition Gaussienne  $\Phi(\cdot)$ , on a asymptotiquement, lorsque  $K \rightarrow \infty$ ,  $\alpha_0(\hat{\delta}) \leq \alpha_0$  et donc  $\hat{\delta} \in \mathcal{K}_{\alpha_0}$ .

*Démonstration.* La démonstration de la proposition 6.2 est donnée dans l'annexe 6.5.2.  $\square$

**Théorème 6.1** (Encadrement de la puissance du test, approche géométrique). *Supposons que le taux d'insertion  $R : 0 < R \leq 1$ , qu'il existe un estimateur  $\hat{\eta}_k$  de  $\eta_k$  vérifiant  $\|\eta_k - \hat{\eta}_k\|_1 \leq \vartheta$  et que pour tout  $k \in \{1, \dots, K\}$ , la covariance  $\Sigma_k$  est connue. La puissance  $\beta_{\hat{\delta}}$  associée au test  $\hat{\delta}$  du rapport de vraisemblance approché (6.22) est donnée par :*

$$1 - \Phi\left(\frac{\hat{\tau}_{\alpha_0} - \sqrt{\kappa}\varrho}{1 + \gamma + b_{\max}}\right) \leq \beta_{\hat{\delta}} \leq 1 - \Phi\left(\frac{\hat{\tau}_{\alpha_0} - \sqrt{\kappa}\varrho}{1 + \gamma}\right) \quad (6.26)$$

*Démonstration.* La démonstration du théorème 6.1, et accessoirement des relations (6.24) - (6.28), est donnée dans l'annexe 6.5.2.  $\square$

Les résultats (6.24) - (6.28) présentés sont fort intéressants ; cependant, il convient de se questionner sur le fondement de ces relations. En effet, supposer qu'il existe un estimateur  $\hat{\eta}_k$  de  $\eta_k$  vérifiant  $\|\eta_k - \hat{\eta}_k\|_1 \leq \vartheta$  ou, plus statistiquement,  $\mathbb{P}[\|\eta_k - \hat{\eta}_k\|_1 > \vartheta] = 0$  permet de borner "géométriquement" la perte de puissance du test  $\hat{\delta}$  (6.22), mais l'obtention, en pratique, d'un tel estimateur est douteux.

Aussi, une approche alternative plus satisfaisante d'un point de vue statistique est étudiée en supposant qu'il existe, pour chacune des discontinuités, un estimateur  $\hat{\eta}_k$  de  $\eta_k$  tel que  $\mathbb{E}[\|\eta_k - \hat{\eta}_k\|_2^2] \leq \vartheta_2$  ; autrement dit que l'erreur quadratique moyenne (EQM) d'estimation est bornée.

**Remarque 6.7.** *Il est notable que la présente approche est moins restrictive (ou plus large) que l'approche précédente qualifiée de "géométrique" dans le sens où un calcul direct permet de vérifier que :*

$$\begin{aligned} \|\eta_k - \hat{\eta}_k\|_1 \leq \vartheta &\Rightarrow \mathbb{E}[(\eta_k - \hat{\eta}_k)^T(\eta_k - \hat{\eta}_k)] \leq \vartheta_2 \\ \mathbb{E}[(\eta_k - \hat{\eta}_k)^T(\eta_k - \hat{\eta}_k)] \leq \vartheta_2 &\nRightarrow \|\eta_k - \hat{\eta}_k\|_1 \leq \vartheta \end{aligned}$$

**Proposition 6.3.** *Un calcul direct permet de montrer que pour tout  $\alpha \in ]0; 0.5]$ , en fixant*

$$\hat{\tau}_{\alpha_0} = \Phi^{-1}(1 - \alpha_0)\sqrt{1 + b_{\max}}, \quad (6.27)$$

où  $\Phi^{-1}(\cdot)$  est la fonction réciproque de la fonction de répartition Gaussienne  $\Phi(\cdot)$ , on a asymptotiquement, lorsque  $K \rightarrow \infty$ ,  $\alpha_0(\hat{\delta}) = \alpha_0$  et donc  $\hat{\delta} \in \mathcal{K}_{\alpha_0}$ .

*La contrainte  $\alpha_0 \geq 0.5$  est uniquement nécessaire pour vérifier l'inégalité (6.25).*

*Démonstration.* La démonstration de la proposition 6.3 est donnée dans l'annexe 6.5.2.  $\square$

**Théorème 6.2** (Encadrement de la puissance du test (approche statistique)). *Supposons que le taux d'insertion  $R \in ]0; 1]$ , qu'il existe un estimateur  $\hat{\boldsymbol{\eta}}_k$  de  $\boldsymbol{\eta}_k$  vérifiant  $\mathbb{E}[\|\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k\|_2^2] \leq \vartheta_2$  et que pour tout  $k \in \{1, \dots, K\}$ , la covariance  $\Sigma_k$  est connue. La puissance  $\beta_{\hat{\delta}}$  associée au test  $\hat{\delta}$  du rapport de vraisemblance approché (6.22) est donnée par :*

$$1 - \Phi\left(\frac{\hat{\tau}_{\alpha_0} - \sqrt{\kappa}\varrho}{1 + \gamma + b_{\max}}\right) \leq \beta_{\hat{\delta}} \leq 1 - \Phi\left(\frac{\hat{\tau}_{\alpha_0} - \sqrt{\kappa}\varrho}{1 + \gamma}\right) \quad (6.28)$$

*Démonstration.* La démonstration du théorème 6.2, est donnée dans l'annexe 6.5.2.  $\square$

## Conclusions du chapitre :

Dans ce chapitre 6, le modèle local non-linéaire du contenu des images naturelles est utilisé dans un test statistique pour la détection d'informations cachées. L'étude des courbures paramétrique et intrinsèque du modèle permet d'envisager une linéarisation afin de surmonter les difficultés dues à la non-linéarité du modèle. Les paramètres non-linéaires décrivant les discontinuités doivent être estimés ; pour cela la méthode utilisée, issue de la littérature, est décrite. L'intérêt de cette approche est alors, d'une part, de limiter la complexité calculatoire de l'ensemble de la procédure de test et, d'autre part, de pouvoir analytiquement expliciter les performances statistiques du test proposé.

L'expression de la fonction de puissance du test finalement proposé montre que la perte de puissance (comparativement au test le plus puissant) est bornée. Par ailleurs, les deux sources de perte de puissance sont clairement identifiées. D'une part, la linéarisation du modèle, vis-à-vis des paramètres décrivant les discontinuités, engendre une incertitude qui est bornée. D'autre part, l'estimation des paramètres inconnus (linéaires) restant diminue le nombre de "degré de liberté" et donc la puissance du test.

Ces résultats justifient donc les efforts faits pour proposer un modèle d'image avec peu de paramètres.

## 6.5 Annexes du chapitre 5

### 6.5.1 Borne de Rao-Cramer sur l'estimation des paramètres de nuisance non-linéaires

Le théorème 6.1 de la section 6.4.2 reposent sur le fait que la performance de l'erreur d'estimation  $\eta_k - \hat{\eta}_k$  est bornée, voir également la démonstration de l'annexe 6.5.2. Afin de vérifier la pertinence de cette l'hypothèse et afin de comparer les performances de l'estimateur utilisé avec une borne d'optimalité, il est proposé dans cette annexe de calculer la "borne de Rao-Cramer", voir théorème A.4, pour l'estimation du paramètre de nuisance non-linéaire  $\eta_k$ . L'idée n'est pas ici de fournir une expression analytique de  $\vartheta$  pour un estimateur donné, mais davantage de vérifier qu'admettre l'existence d'une telle (petite) borne est une hypothèse plausible (ou au moins possible).

Il est donc proposé d'étudier le cas simpliste d'un signal  $\mathbf{z}$  de  $N$  observations contenant une unique discontinuité;  $\mathbf{z}$  est donc modélisé par  $\mathbf{z} \sim \mathcal{N}(\mu_0 + u(\boldsymbol{\eta}), \Sigma)$  les composantes du vecteur sont données, pour tout  $n \in \{1, \dots, N\}$ , par  $\mathbf{f}(\boldsymbol{\eta}) = \int_{n-\frac{1}{2}}^{n-\frac{1}{2}} f(x; \boldsymbol{\eta}_k) dx = \int_{n-\frac{1}{2}}^{n-\frac{1}{2}} \Phi\left(\frac{x-v_0}{\zeta_0}\right) dx$  et  $\boldsymbol{\eta} = (v_0, \zeta_0)$ . Il est immédiat de montrer que, en vertu de la distribution Gaussienne du bruit additif, la matrice d'information de Fisher, définie dans (A.15), est donnée par :

$$\mathcal{F}_{i,j}(\boldsymbol{\eta}) = \dot{\mathbf{f}}_{\eta_i}^T \Sigma^{-1} \dot{\mathbf{f}}_{\eta_j} \quad (6.29)$$

où, comme déjà détaillé dans la section 6.2, les vecteurs  $\dot{\mathbf{f}}_{\eta_1} = \dot{\mathbf{f}}_{v_k}$  et  $\dot{\mathbf{f}}_{\eta_2} = \dot{\mathbf{f}}_{\zeta_0}$  représentent les dérivées partielles du vecteur  $\mathbf{f}(\boldsymbol{\eta})$  considéré comme un champ scalaire  $\mathbf{f}(\boldsymbol{\eta}) : \mathbb{R}^2 \rightarrow \mathbb{R}^N$ . Dans le cas général, l'évaluation de la matrice d'information de Fisher (6.29), est assez complexe d'un point de vue algébrique. Il est donc proposé, à l'instar de la méthodologie présentée dans la section 6.3.2, de simplifier ces calculs en utilisant les expressions analytiques de l'application  $f(\boldsymbol{\eta}) = \Phi(\boldsymbol{\eta})$  et de ses dérivées partielles. Dans le cas où le bruit additif Gaussien est i.i.d *i.e.* lorsque  $\Sigma = \sigma^2 \mathbf{I}_N$ , il découle d'un rapide calcul l'expression suivante de la matrice d'information de Fisher (6.29) :

$$\mathcal{F}(\boldsymbol{\eta}) \simeq \frac{u^2}{\sigma^2} \begin{pmatrix} \|\dot{f}_{v_0}\|_{\mathcal{L}^2} & \\ 0 & \|\dot{f}_{\zeta_0}\|_{\mathcal{L}^2} \end{pmatrix} = \frac{u^2}{\sigma^2} \frac{1}{2\sqrt{\pi}\zeta_0} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}. \quad (6.30)$$

La borne de Rao-Cramer d'un estimateur, non-biaisé, du paramètre non-linéaire  $\boldsymbol{\eta}$  est alors immédiatement donnée par :

$$\text{Cov}(\hat{\boldsymbol{\eta}}) \geq \frac{\mathcal{F}(\boldsymbol{\eta})^{-1}}{N} = \frac{2\sqrt{\pi}\zeta_0}{u^2} \frac{\sigma^2}{N} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

où le signe  $A \geq B \Leftrightarrow A - B \geq 0$  signifie que la matrice  $A - B$  est semi-définie positive.

Les tableaux 6.4 et 6.5 présentent le biais et la covariance de l'estimateur du maximum de vraisemblance implémenté numériquement au moyen de la méthode de Newton pour le problème de recherche du maximum.

Les résultats numériques, données dans le tableau 6.4, concernant le biais numériquement obtenus montrent que le l'estimateur est approximativement non-biaisé.

Les résultats obtenus pour la variance des estimateurs, données dans le tableaux 6.5, montrent en revanche qu'en pratique, la variance sur l'estimation des paramètres non-linéaire par maximum de vraisemblance sont loin d'atteindre la borne de Rao-Cramer. On notera cependant que le tableau 6.5 confirme la relation 6.30 dans le sens où il semble que  $\text{Cov}(\hat{\boldsymbol{\eta}}) \propto \zeta_0/u^2$ . La covariance  $\mathbb{E}[(v_0 - \hat{v}_0)(\sigma_0 - \hat{\sigma}_0)]$  n'est pas présentée afin d'alléger le tableau et car la valeur de cette dernière est très faible (en comparaison aux variances).

	Estimateur MV ( $\times 10^{-3}$ )	Algorithme utilisé
paramètres	$\nu_0 = 0.25, \zeta_0 = 1.25, u = 12.5, N = 33$	
$\mathbb{E}[\nu_0 - \hat{\nu}_0]$	3.53	$2.03 \cdot 10^{-3}$
$\mathbb{E}[\zeta_0 - \hat{\zeta}_0]$	-3.73	0.156
paramètres	$\nu_0 = 0.25, \zeta_0 = 2.75, u = 12.5, N = 33$	
$\mathbb{E}[\nu_0 - \hat{\nu}_0]$	8.20	$1.10 \cdot 10^{-3}$
$\mathbb{E}[\zeta_0 - \hat{\zeta}_0]$	-0.67	0.146
paramètres	$\nu_0 = 0.25, \zeta_0 = 1.25, u = 47.5, N = 33$	
$\mathbb{E}[\nu_0 - \hat{\nu}_0]$	-2.73	$-7.19 \cdot 10^{-4}$
$\mathbb{E}[\zeta_0 - \hat{\zeta}_0]$	1.13	0.126

TABLE 6.4: Biais des estimateurs du paramètre non linéaire  $\eta$  par maximum de vraisemblance (calculé par la méthode de Newton) et par l'algorithme utilisé.

	Borne de Rao-Cramer ( $\times 10^{-3}$ )	Estimateur MV ( $\times 10^{-3}$ )	Algorithme utilisé
paramètres	$\nu_0 = 0.25, \zeta_0 = 1.25, u = 12.5, N = 33$		
$\text{Var}[\hat{\nu}_0]$	0.86	33.36	0.121
$\text{Var}[\hat{\zeta}_0]$	1.72	66.99	0.233
paramètres	$\nu_0 = 0.25, \zeta_0 = 2.75, u = 12.5, N = 33$		
$\text{Var}[\hat{\nu}_0]$	1.89	75.13	0.307
$\text{Var}[\hat{\zeta}_0]$	3.78	145.77	0.461
paramètres	$\nu_0 = 0.25, \zeta_0 = 1.25, u = 47.5, N = 33$		
$\text{Var}[\hat{\nu}_0]$	0.06	2.23	0.033
$\text{Var}[\hat{\zeta}_0]$	0.12	4.25	0.073

TABLE 6.5: Variance des estimateurs du paramètre non linéaire  $\eta$  par maximum de vraisemblance (calculé par la méthode de Newton) et par l'algorithme utilisé.

**Remarque 6.8.** Les conditions d'applications des précédents calculs, présence d'une unique discontinuité et absence de phénomène de bord (puisque les intégrales sont calculées sur  $\mathbb{R}$ ), peuvent sembler peu réaliste. En pratique cependant, le support sur lequel les fonctions  $\dot{f}_{\nu_0}$  et  $\dot{f}_{\zeta_0}$  sont nulles peut être réduits à  $x \in K = \{x \in \mathbb{R} \mid |x - \nu_0/\zeta_0| < x_{max}\}$ . En admettant l'approximation que pour tout  $x \notin K$ , on a  $\dot{f}_{\nu_0}(x) = 0$  et  $\dot{f}_{\zeta_0}(x) = 0$ , les conditions d'applications des précédents sont relaxer au sens il est seulement supposé qu'une unique discontinuité est présenté sur  $K \subset \mathbb{R}$ .

### 6.5.2 Démonstration des théorèmes 6.1 et 6.2.

En préambule, il est importante de noter que les démonstrations présentées sont restreintes au cas où une seule discontinuité est éventuellement présente dans chacun des segments et aux cas où les effets de bords sont, ici encore, omis. Supposer le présence, au plus, d'une discontinuité présente un intérêt de simplification des notations ; les démonstration peuvent être aisément étendue aux cas de présence de multiples discontinuités au prix de notations beaucoup plus lourdes. Le second point, omission des effets de bords, consiste à supposer que sur chacun des segments analysés, la fonction  $f_{\eta_k}$ ,  $k \in \{1, \dots, K\}$  est constante en dehors de son domaine d'étude  $\mathcal{X}_k$  ; ce point,

brèvement abordé dans la section 6.3.2, permet de simplifier grandement les calculs et peut être vérifié en pratique par concaténation de deux segments si nécessaire, voir plus de détails dans [166, 167, 168, 170].

Pour le besoin des démonstrations, il est rappelé qu'en vertu du corollaire 6.1 le biais d'estimation lié à la linéarisation du modèle vérifie, voir relations (6.3.2) :

$$\sum_{l=1}^L \mathbf{b}_{k,l} = \int_0^L \ddot{\mathbf{f}}_{\hat{\boldsymbol{\eta}}_k}^{\mathbf{N}}(x) dx = 0$$

$$\|\mathbf{b}_k\|_2^2 = \sum_{l=1}^L \mathbf{b}_{k,l}^2 \leq \left( \frac{2}{8\sqrt{\pi}\sigma_k^3} \right) \|\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k\|_1^2,$$

où  $\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k = (\delta_{v_k}, \delta_{\zeta_k})^T$ .

Les deux démonstration proposées sont structurées de façon analogue. Dans un premier les deux moment du RV  $\log(\hat{\Lambda}(\mathbf{z}_k))$  sont caractérisés, l'espérance mathématique  $\mathbb{E}_{\mathcal{R}}[\log(\hat{\Lambda}(\mathbf{z}_k))]$  est calculée en fonction de  $R$  et la variance  $\text{Var}_{\mathcal{R}}[\log(\hat{\Lambda}(\mathbf{z}_k))]$  est bornée. Ce calcul permet ensuite de calculer, en vertu du théorème (3.1) de la limite centrale de Lindeberg, la distribution asymptotique (lorsque  $K \rightarrow \infty$ ) du RV  $\log(\hat{\Lambda}(\mathbf{Z}))$  sur l'ensemble de l'image analysée. Enfin, la distribution asymptotique du RV  $\log(\hat{\Lambda}(\mathbf{Z}))$  permet de calculer directement le seuil de décision  $\tau_{\alpha_0}$ , garantissant le respect de la contrainte sur la probabilité de fausse alarme, et de borner la puissance du test.

*S'il est possible de calculer l'espérance mathématique  $\mathbb{E}_{\mathcal{R}}[\log(\hat{\Lambda}(\mathbf{z}_k))]$ , la variance ne peut être que bornée car  $\boldsymbol{\eta}_k$  demeure inconnu la quantité  $\|\mathbf{b}_k\|_2^2$  ne peut être que bornée.*

### Calculs des moments du RV $\log(\hat{\Lambda}(\mathbf{z}_k))$

Pour des raisons de clarté évidentes, les calculs des moments du RV sont calculés en utilisant les conventions de notations suivantes :

$$\zeta_{k,l} = z_{k,l} - \hat{\theta}_{k,l}, w_{k,l} = z_{k,l} - \bar{z}_{k,l}, v_{k,l} = (z_{k,l} - \theta_{k,l})(z_{k,l} - \bar{z}_{k,l}) = \zeta_{k,l}w_{k,l}, \quad (6.31)$$

de sorte que :

$$\log(\hat{\Lambda}(\mathbf{Z})) = \frac{\bar{\sigma}}{\sqrt{K}} \sum_{k=1}^K \frac{1}{\sigma_k^2 \sqrt{L-p-3r_k}} \sum_{l=1}^L v_{k,l}.$$

Il est, par ailleurs, rappelé que le segment analysé  $\mathbf{z}_k$  est donné par  $\mathbf{z}_k = Q(\boldsymbol{\theta}_k + \boldsymbol{\zeta}_k) = \boldsymbol{\theta}_k + \boldsymbol{\zeta}_k + \mathbf{q}_k$ , où  $\mathbf{q}_k$  représente l'erreur due à la quantification modélisée par une des variables aléatoires i.i.d de moyennes nulle et de variances  $\text{Var}[\mathbf{q}_{k,l}] = 1/12$ . En outre, le bruit d'acquisition  $\boldsymbol{\zeta}_k$  est supposé être la réalisation de variables aléatoire Gausienne i.i.d, de moyenne nulle et de variance  $\sigma_k^2$ . Enfin, l'estimation  $\hat{\boldsymbol{\theta}}_k$  de  $\boldsymbol{\theta}_k$ , donnée par (6.7)  $\hat{\boldsymbol{\theta}}_k = \mathbf{P}_{\hat{\mathbf{G}}_{\hat{\boldsymbol{\eta}}_k}} \mathbf{z}_k$  vérifie  $\hat{\boldsymbol{\theta}}_k = \boldsymbol{\theta}_k + \mathbf{b} + o(\|\boldsymbol{\eta}_k\|_1^2)$ , ce qui permet d'écrire :

$$\mathbf{z}_k - \hat{\boldsymbol{\theta}}_k = \mathbf{P}_{\hat{\mathbf{G}}_{\hat{\boldsymbol{\eta}}_k}}^\perp \mathbf{z}_k = \boldsymbol{\zeta}_k = \tilde{\boldsymbol{\zeta}}_k + \tilde{\mathbf{q}}_k + \mathbf{b}_k + o(\|\boldsymbol{\eta}_k\|_1^2), \quad (6.32)$$

où  $\tilde{\boldsymbol{\zeta}}_k = \mathbf{P}_{\hat{\mathbf{G}}_{\hat{\boldsymbol{\eta}}_k}}^\perp \boldsymbol{\zeta}_k$ ,  $\tilde{\mathbf{q}}_k = \mathbf{P}_{\hat{\mathbf{G}}_{\hat{\boldsymbol{\eta}}_k}}^\perp \mathbf{q}_k$  et  $\mathbf{b}_k = \mathbf{P}_{\hat{\mathbf{G}}_{\hat{\boldsymbol{\eta}}_k}}^\perp \boldsymbol{\theta}_k$ .

En utilisant la relation (6.19) de la proposition 6.1 on montre que, sous l'hypothèse  $\mathcal{H}_0$ , l'espérance mathématique  $\mathbb{E}_{\mathcal{H}_0}[v_{k,l}]$  est donnée, de manière analogue à la relation (3.55), par :

$$\begin{aligned} \mathbb{E}_{\mathcal{H}_0} \left[ \sum_{l=1}^L v_{k,l} \right] &= \sum_{l=1}^L \mathbb{E}_{\mathcal{H}_0}[v_{k,l}] \\ &= \sum_{l=1}^L \mathbb{P}_{\mathcal{H}_0}[w_{k,l}=1] \mathbb{E}_{\mathcal{H}_0}[\zeta_{k,l}|w_{k,l}=1] - \mathbb{P}_{\mathcal{H}_0}[w_{k,l}=-1] \mathbb{E}_{\mathcal{H}_0}[\zeta_{k,l}|w_{k,l}=-1] \\ &= 0 + \text{trace}(\mathbf{P}_{\hat{\mathbf{G}}_{\hat{\boldsymbol{\eta}}_k}}^\perp) o_p(\sigma_k^{-2}) \end{aligned} \quad (6.33)$$



où  $\text{trace}(\mathbf{P}_{\mathbf{G}_{\hat{\eta}_k}}^\perp) = L - p - 3r_k$  représente la trace de la matrice  $\mathbf{P}_{\mathbf{G}_{\hat{\eta}_k}}^\perp$ .

De manière analogue, l'espérance mathématique  $\mathbb{E}_{\mathcal{H}_1}[v_{k,l}]$  est donnée sous l'hypothèse  $\mathcal{H}_1$  par :

$$\begin{aligned} \mathbb{E}_{\mathcal{H}_1} \left[ \sum_{l=1}^L v_{k,l} \right] &= \sum_{l=1}^L \mathbb{E}_{\mathcal{H}_1}[v_{k,l}] \\ &= \sum_{l=1}^L \mathbb{P}_{\mathcal{H}_1}[w_{k,l}=1] \mathbb{E}_{\mathcal{H}_1}[\zeta_{k,l}|w_{k,l}=1] - \mathbb{P}_{\mathcal{H}_1}[w_{k,l}=-1] \mathbb{E}_{\mathcal{H}_1}[\zeta_{k,l}|w_{k,l}=-1] \\ &= \left(1 - \frac{R}{2}\right) \mathbb{E}_{\mathcal{H}_0} \left[ \sum_{l=1}^L v_{k,l} \right] + \frac{R}{4} \sum_{l=1}^L \mathbb{E}_{\mathcal{H}_0}[\zeta_{k,l} + 1] - \mathbb{E}_{\mathcal{H}_0}[\zeta_{k,l} - 1] \\ &= \text{trace}(\mathbf{P}_{\mathbf{G}_{\hat{\eta}_k}}^\perp) \left( \frac{R}{2} + o_p(\sigma_k^{-2}) \right). \end{aligned} \quad (6.34)$$

Le calcul de la variance est réalisé, sous chacune des deux hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$ , en utilisant le fait que :

$$\text{Var}_{\mathcal{H}_i} \left[ \sum_{l=1}^L v_{k,l} \right] = \sum_{l=1}^L \sum_{n=1}^L \text{Cov}_{\mathcal{H}_i}[v_{k,l}; v_{k,n}], \quad i = \{0, 1\}.$$

En considérant que  $\tilde{\zeta}_k$ ,  $\tilde{\mathbf{q}}_k$  et  $\mathbf{b}_k$  sont la réalisation de vecteurs mutuellement indépendants et qu'en outre, le bruit d'acquisition  $\zeta_k$  et de quantification  $\mathbf{q}_k$  sont la réalisation de variables aléatoires i.i.d, un rapide calcul permet de montrer que sous l'hypothèse  $\mathcal{H}_0$  on a :

$$\begin{aligned} \text{Var}_{\mathcal{H}_0} \left[ \sum_{l=1}^L v_{k,l} \right] &= \sum_{l=1}^L \mathbb{E}_{\mathcal{H}_0} [\tilde{\zeta}_{k,l}^2 + \tilde{q}_{k,l}^2] + \sum_{m=1}^L \sum_{n=1}^L \text{Cov}_{\mathcal{H}_0}[w_{k,l}b_{k,l}; w_{k,n}b_{k,n}], \quad (6.35) \\ \text{avec } \sum_{l=1}^L \mathbb{E}_{\mathcal{H}_0} [\tilde{\zeta}_{k,l}^2 + \tilde{q}_{k,l}^2] &= \text{trace}(\mathbf{P}_{\mathbf{G}_{\hat{\eta}_k}}^\perp) (\sigma_k^2 + 1/12) + o_p(\sigma_k^{-2}). \end{aligned}$$

Enfin, en utilisant le fait que :

$$\begin{aligned} \mathbb{E}_{\mathcal{H}_0}[w_{k,l}b_{k,l}] &= \mathbb{P}_{\mathcal{H}_0}[w_{k,l}=1] \mathbb{E}_{\mathcal{H}_0}[b_{k,l}|w_{k,l}=1] - \mathbb{P}_{\mathcal{H}_0}[w_{k,l}=-1] \mathbb{E}_{\mathcal{H}_0}[b_{k,l}|w_{k,l}=-1] \\ &= 0 + o_p(\sigma_k^{-2}), \end{aligned}$$

il en découle que :

$$\text{Cov}_{\mathcal{H}_0}[w_{k,l}b_{k,l}; w_{k,n}b_{k,n}] = \mathbb{E}[(w_{k,l}b_{k,l})(w_{k,n}b_{k,n})] + 0 + o_p(\sigma_k^{-2})$$

Pour tout  $l = n$  il est immédiat de constater que  $w_{k,l}w_{k,n} = 1$ . À l'inverse, pour tout  $l \neq n$  il découle de la relation (3.54) que  $\mathbb{P}[w_{k,l}w_{k,n} = 1] = 1/2 + o(\sigma_k^{-1}) = \mathbb{P}[w_{k,l}w_{k,n} = -1]$ . Il en découle donc que :

$$\sum_{l=1}^L \sum_{n=1}^L \text{Cov}_{\mathcal{H}_0}[w_{k,l}b_{k,l}; w_{k,n}b_{k,n}] = \sum_{l=1}^L b_{k,l}^2 + o_p(L\sigma_k^{-2}). \quad (6.36)$$

L'utilisation de l'assertion  $\|\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k\|_1^2 \leq \vartheta$  et de la relation (6.19) de la proposition 6.1, permet d'obtenir :

$$\sum_{l=1}^L b_{k,l}^2 = \|\mathbf{b}_k\|_2^2 \leq \left( \frac{2}{8\sqrt{\pi}\zeta_k^3} \right) \|\boldsymbol{\eta}_k - \hat{\boldsymbol{\eta}}_k\|_1^2 \leq \left( \frac{2}{8\sqrt{\pi}\zeta_k^3} \right) \vartheta^2 = b_{\max}.$$

Il découle finalement des relations (6.35), et (6.36) que :

$$\begin{aligned} \text{Var}_{\mathcal{H}_0} \left[ \sum_{l=1}^L v_{k,l} \right] &= \text{trace}(\mathbf{P}_{\mathbf{G}_{\hat{\eta}_k}}^\perp) (\sigma_k^2 + 1/12) + \|\mathbf{b}_k\|_2^2 + o_p(L\sigma_k^{-2}) \\ &\leq \text{trace}(\mathbf{P}_{\mathbf{G}_{\hat{\eta}_k}}^\perp) (\sigma_k^2 + 1/12) + b_{\max} + o_p(L\sigma_k^{-2}) \end{aligned} \quad (6.37)$$

Enfin, des calculs similaires aux précédents permettent d'obtenir la relation suivante, analogue à (3.58) :

$$\text{Var}_{\mathcal{H}_1} \left[ \sum_{l=1}^L v_{k,l} \right] = \text{Var}_{\mathcal{H}_0} \left[ \sum_{l=1}^L v_{k,l} \right] + \text{trace}(\mathbf{P}_{\hat{\mathbf{G}}_{\eta_k}}^\perp) \frac{R}{2} \left( 1 - \frac{R}{2} \right). \quad (6.38)$$

**Démonstration 6.1** (de la proposition 6.2). En application du théorème 3.1 de la limite centrale de Lindenberg, il découle des relations (6.33) et (6.37) que sous l'hypothèse  $\mathcal{H}_0$  on a  $\forall k \in \{1, \dots, K\}$  :

$$\log(\hat{\Lambda}(\mathbf{Z})) = \frac{\bar{\sigma}}{\sqrt{K}} \sum_{k=1}^K \frac{1}{\sigma_k^2 \sqrt{L-p-3r_k}} \sum_{l=1}^L v_{k,l} \rightsquigarrow \mathcal{N}(\sqrt{\kappa} \varrho; 1+b),$$

$$\text{avec } 0 \leq b \leq \frac{b_{\max}}{\bar{\sigma}^2(L-p-3)} + o_p(\bar{\sigma}^{-2}) = \frac{o(\vartheta)}{\bar{\sigma}^2(L-p-3)}.$$

Ainsi pour tout  $\alpha_0 \in ]0; 0.5]$ , en fixant

$$\hat{\tau}_{\alpha_0} = \Phi^{-1}(1 - \alpha_0) \sqrt{1 + b_{\max}},$$

un calcul direct permet de montrer que :

$$\mathbb{P}_{\mathcal{H}_0}[\log(\hat{\Lambda}(\mathbf{Z})) > \hat{\tau}_{\alpha_0}] \leq \alpha_0.$$

**Démonstration 6.2** (du théorème 6.1). En application du théorème 3.1 de la limite centrale de Lindenberg, il découle des relations (6.34) et (6.38) que sous l'hypothèse  $\mathcal{H}_1$  on a  $\forall k \in \{1, \dots, K\}$  :

$$\log(\hat{\Lambda}(\mathbf{Z})) = \frac{\bar{\sigma}}{\sqrt{K}} \sum_{k=1}^K \frac{1}{\sigma_k^2 \sqrt{L-p-3r_k}} \sum_{l=1}^L v_{k,l} \rightsquigarrow \mathcal{N}(\sqrt{\kappa} \varrho; 1+\gamma+b),$$

$$\text{avec } \gamma = \frac{1}{\bar{\sigma}^2} \frac{R}{2} \left( 1 - \frac{R}{2} \right).$$

Ainsi pour tout  $\hat{\tau}_{\alpha_0} \in \mathbb{R}$ , en vertu du fait  $b \leq b_{\max}$ , un rapide permet de montrer que :

$$\beta_{\hat{\delta}} \geq 1 - \Phi \left( \frac{\hat{\tau}_{\alpha_0} - \sqrt{\kappa} \varrho}{1 + \gamma + b_{\max}} \right). \quad (6.39)$$

De manière analogue, en vertu du fait que  $b \geq 0$ , un calcul direct permet de montrer que :

$$\beta_{\hat{\delta}} \leq 1 - \Phi \left( \frac{\hat{\tau}_{\alpha_0} - \sqrt{\kappa} \varrho}{1 + \gamma} \right). \quad (6.40)$$

### Démonstration du théorème 6.2

Les calculs des espérances mathématiques, détaillées dans les relations (6.33) et (6.34), reposent sur le fait que  $\sum_{l=1}^L b_{k,l} = 0$ . Ainsi, en considérant que l'estimation  $\hat{\eta}_k$  vérifie  $\mathbb{E}[\|\eta_k - \hat{\eta}_k\|_2^2] \leq \vartheta_2$  et en admettant, ici encore, que les phénomènes de bord sont exclus et qu'au plus une unique discontinuité est présente dans le segment  $z_k$  l'espérance  $\mathbb{E}_{\mathcal{H}_i} \left[ \sum_{l=1}^L v_{k,l} \right]$  est donnée sous chacune des hypothèses  $i = \{0; 1\}$  par les relations suivantes :

$$\mathbb{E}_{\mathcal{H}_0} \left[ \sum_{l=1}^L v_{k,l} \right] \text{ et } \mathbb{E}_{\mathcal{H}_1} \left[ \sum_{l=1}^L v_{k,l} \right]. \quad (6.41)$$

Afin de calculer la variance  $\text{Var}_{\mathcal{H}_i} \left[ \sum_{l=1}^L v_{k,l} \right]$  sous les hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$ , il est nécessaire de noter en préambule que pour tout  $x \in \mathbb{R}^n$ ,  $\|x\|_1 \leq \sqrt{n}\|x\|_2$ ; dans le présent cas, cette relation permet d'obtenir, en utilisation les équations (6.3.2), l'inégalité suivante :

$$\|b_k\|_2^2 \leq \left( \frac{2}{8\sqrt{\pi}\zeta_k^3} \right) \|\eta_k - \hat{\eta}_k\|_1^2 \leq \left( \frac{4}{8\sqrt{\pi}\zeta_k^3} \right) \|\eta_k - \hat{\eta}_k\|_2^2. \quad (6.42)$$

En outre, les calculs permettant d'aboutir à la relation (6.36) demeurent inchangés; sous l'hypothèse  $\mathcal{H}_0$  on a donc toujours :

$$\text{Var}_{\mathcal{H}_0} \left[ \sum_{l=1}^L v_{k,l} \right] = \text{trace}(\mathbf{P}_{\hat{\mathbf{G}}_{\eta_k}}^\perp) (\sigma_k^2 + 1/12) + \|b_k\|_2^2 + o_p(L\sigma_k^{-2}). \quad (6.43)$$

Comparativement aux calculs réalisés dans le cas où l'on suppose que  $\|\eta_k - \hat{\eta}_k\|_1^2 \leq \vartheta$ , considérer que l'erreur quadratique moyenne (EQM) de l'estimation  $\hat{\eta}_k$  est bornée  $\mathbb{E} [\|\eta_k - \hat{\eta}_k\|_2^2] \leq \vartheta_2$  ne permet pas de borner le terme  $b_{k,l}^2$  de la relation (6.43). Il est cependant possible, en vertu de la relation (6.42), de borner la variance moyenne :

$$\mathbb{E} [\|\eta_k - \hat{\eta}_k\|_2^2] \leq \vartheta_2 \Rightarrow \mathbb{E} [\|b_k\|_2^2] \leq \left( \frac{4}{8\sqrt{\pi}\zeta_k^3} \right) \vartheta_2 = b_{\max}. \quad (6.44)$$

En vertu de la loi faible des grands nombres pour des variables aléatoires indépendantes, pas nécessairement identiquement distribuées voir [183, p.254-260], il découle des relations (6.43) - (6.44) que :

$$\lim_{K \rightarrow \infty} \mathbb{P} \left[ \frac{1}{K} \sum_{k=1}^K \text{Var}_{\mathcal{H}_0} \left[ \sum_{l=1}^L v_{k,l} \right] \geq \sigma_{\max}^2 \right] = 0 \quad (6.45)$$

avec

$$\sigma_{\max}^2 = \text{trace}(\mathbf{P}_{\hat{\mathbf{G}}_{\eta_k}}^\perp) (\sigma_k^2 + 1/12) + b_{\max} + o_p(L\sigma_k^{-2}).$$

**Remarque 6.9.** Formellement, la relation (6.44) ne permet à elle seule de vérifier les conditions d'application du théorème 3.1 de la limite centrale de Lindeberg. Ce problème d'avantage théorique n'a guère d'intérêt en pratique car la valeur des pixels est bornée, ce qui permet de garantir les condition d'application du théorème 3.1.

Enfin, des calculs similaires aux précédents (6.43) - (6.44) permettent d'obtenir, sous l'hypothèse alternative  $\mathcal{H}_1$ , la relation suivante, similaire en tout point à (6.38) :

$$\text{Var}_{\mathcal{H}_1} \left[ \sum_{l=1}^L v_{k,l} \right] = \text{Var}_{\mathcal{H}_0} \left[ \sum_{l=1}^L v_{k,l} \right] + \text{trace}(\mathbf{P}_{\hat{\mathbf{G}}_{\eta_k}}^\perp) \frac{R}{2} \left( 1 - \frac{R}{2} \right). \quad (6.46)$$

**Démonstration 6.3** (de la proposition 6.3). En application du théorème 3.1 de la limite centrale de Lindenberg, il découle des relations (6.41) et (6.43) - (6.44) que sous l'hypothèse  $\mathcal{H}_0$  on a, lorsque  $K \rightarrow \infty$  :

$$\log(\hat{\Lambda}(\mathbf{Z})) = \frac{1}{\sqrt{K}} \sum_{k=1}^K \log(\hat{\Lambda}(\mathbf{z}_k)) = \frac{\bar{\sigma}}{\sqrt{K}} \sum_{k=1}^K \frac{1}{\sigma_k^2 \sqrt{L-p-3r_k}} \sum_{l=1}^L v_{k,l} \rightsquigarrow \mathcal{N}(\sqrt{\kappa} \varrho; 1+b),$$

$$\text{avec } b = \frac{1}{K} \sum_{k=1}^K \frac{\|b_k\|_2^2}{\bar{\sigma}^2 L - p - 3r_k} + o_p(\bar{\sigma}^{-2}) \text{ et } \lim_{K \rightarrow \infty} \mathbb{P} \left[ b \geq \frac{b_{\max}}{\bar{\sigma}^2(L-p-3)} \right] = 0$$

Ainsi pour tout  $\alpha_0 \in ]0; 0.5]$ , en fixant

$$\hat{\tau}_{\alpha_0} = \Phi^{-1}(1 - \alpha_0) \sqrt{1 + b_{\max}},$$

un calcul direct permet de montrer que lorsque  $N \rightarrow \infty$  :

$$\lim_{K \rightarrow \infty} \mathbb{P}_{\mathcal{H}_0} \left[ \frac{1}{\sqrt{K}} \sum_{k=1}^K \log(\hat{\Lambda}(\mathbf{z}_k)) > \hat{\tau}_{\alpha_0} \right] \leq \alpha_0.$$

**Démonstration 6.4** (du théorème 6.2). En application du théorème 3.1 de la limite centrale de Lindenberg, il découle des relations (6.41) et (6.46) que sous l'hypothèse  $\mathcal{H}_1$  on a, lorsque  $K \rightarrow \infty$  :

$$\log(\widehat{\Lambda}(\mathbf{Z})) = \frac{1}{\sqrt{K}} \sum_{k=1}^K \log(\widehat{\Lambda}(\mathbf{z}_k)) = \frac{\bar{\sigma}}{\sqrt{K}} \sum_{k=1}^K \frac{1}{\sigma_k^2 \sqrt{L-p-3r_k}} \sum_{l=1}^L v_{k,l} \rightsquigarrow \mathcal{N}(\sqrt{\kappa} \varrho; 1 + \gamma + b),$$

$$\text{avec } \gamma = \frac{1}{\bar{\sigma}^2} \frac{R}{2} \left(1 - \frac{R}{2}\right).$$

Ainsi pour tout  $\widehat{\tau}_{\alpha_0} \in \mathbb{R}$ , en vertu de la relation (6.45), un rapide permet de montrer que :

$$\lim \mathbb{P}_{\mathcal{H}_1} \left[ \frac{1}{\sqrt{K}} \sum_{k=1}^K \log(\widehat{\Lambda}(\mathbf{z}_k)) > \widehat{\tau}_{\alpha_0} \right] \geq 1 - \Phi \left( \frac{\widehat{\tau}_{\alpha_0} - \sqrt{\kappa} \varrho}{1 + \gamma + b_{\max}} \right),$$

ce qui sera abusivement notée :

$$\beta_{\widehat{\delta}} \geq 1 - \Phi \left( \frac{\widehat{\tau}_{\alpha_0} - \sqrt{\kappa} \varrho}{1 + \gamma + b_{\max}} \right). \quad (6.47)$$

De manière analogue, en vertu du fait que  $b \geq 0$ , un calcul direct permet de montrer que :

$$\beta_{\widehat{\delta}} \leq 1 - \Phi \left( \frac{\widehat{\tau}_{\alpha_0} - \sqrt{\kappa} \varrho}{1 + \gamma} \right). \quad (6.48)$$



«Le fondement de la théorie, c'est la pratique.»  
— Mao Tsé-Toung, 1893 – 1976

«Une théorie peut être vérifiée par l'expérience,  
mais aucun chemin ne mène de l'expérience à la création d'une théorie. ».

— Albert Einstein, 1879 – 1955

«Entre une pratique sans tête et une théorie sans jambes, il n'y aura jamais à choisir.»  
— Régis Debray, 1940 – ...

## Sommaire

---

Introduction au chapitre . . . . .	149
7.1 Discussion autour des hypothèses de travail . . . . .	149
7.2 Résultats numériques sur des images naturelles : comparaisons avec l'état de l'art . . . . .	151
7.3 Vérifications numériques des résultats théoriques . . . . .	156
Conclusions du chapitre . . . . .	161

---

### Introduction au chapitre :

En préambule de ce chapitre 7, les principales hypothèses de travail utilisées dans ce manuscrit sont discutées dans la section 7.1. La suite du présent chapitre, est dévolue à la présentation de résultats numériques. La section 7.2 présente les principaux résultats numériques obtenus sur des images naturelles ainsi que des comparaisons avec certains détecteurs publiés dans la littérature. Ces résultats mettent en évidence la pertinence de l'approche proposée pour offrir une solution satisfaisant les contraintes fixées dans le cadre du projet ANR RIC, en terme de respect d'une probabilité de fausse alarme faible et de temps de calcul minimal. Enfin, des résultats issus de simulations numériques, permettant de vérifier expérimentalement les principaux résultats théoriques obtenus, sont présentés dans la section 7.3.

## 7.1 Discussion autour des hypothèses de travail

Avant de présenter quelques résultats numériques pour confirmer la pertinence de la méthodologie proposée, il est intéressant de discuter les hypothèses de travail permettant l'application pratique de cette méthodologie. Cette discussion n'a pas vocation à l'exhaustivité, seuls les principales hypothèses de travail sont donc discutées.

- Le théorème 6.1 suppose implicitement que 1) la variance  $\sigma_k^2$  du bruit  $\xi_k$  est constante sur chacun des segments et 2) pour tous les segments  $k = \{1, \dots, K\}$  la

variance  $\sigma_k^2$  est connue. Le premier de ces deux points n'est pas fondamental car une simple modification du log-RV (6.23) permet de prendre en compte le cas où la variance de chacun des pixels d'un segment diffère, pour peu que la variance de chacun des pixels soit connue. En revanche, le second point est plus problématique car la variance  $\sigma_k^2$  est, dans les faits, inconnue ; cette dernière doit donc être estimée ce qu'il est proposé de faire en utilisant l'estimateur :

$$\hat{\sigma}_k^2 = \frac{1}{L - p - 3r_k} \left\| \mathbf{P}_{\hat{\mathbf{G}}_{\hat{\eta}_k}}^\perp \mathbf{z}_k \right\|_2^2 \quad \text{où} \quad \mathbf{P}_{\hat{\mathbf{G}}_{\hat{\eta}_k}}^\perp = \mathbf{I}_L - \hat{\mathbf{G}}_{\hat{\eta}_k} \left( \hat{\mathbf{G}}_{\hat{\eta}_k}^T \hat{\mathbf{G}}_{\hat{\eta}_k} \right)^{-1} \hat{\mathbf{G}}_{\hat{\eta}_k}^T,$$

où la matrice  $\mathbf{P}_{\hat{\mathbf{G}}_{\hat{\eta}_k}}^\perp = \mathbf{I}_L - \hat{\mathbf{G}}_{\hat{\eta}_k} \left( \hat{\mathbf{G}}_{\hat{\eta}_k}^T \hat{\mathbf{G}}_{\hat{\eta}_k} \right)^{-1} \hat{\mathbf{G}}_{\hat{\eta}_k}^T$  correspond au projecteur sur le sous-espace supplémentaire de l'image de  $\hat{\mathbf{G}}_{\hat{\eta}_k}$ .

Il faut toutefois souligner que la prise en compte du fait que la variance des pixels d'un segments (qu'elle soit constante ou pas) n'est pas connue et doit donc être estimée, entraîne une modification de la valeurs des deux premiers moments sous chacune des hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$ . En conséquence, la valeur du seuil de décision donnée par la proposition 6.2 s'en trouve modifiée tout comme la puissance du test explicitée dans le théorème 6.1.

- La principale difficulté en pratique lorsque la variance n'est pas connue est la gestion de certains segments pour lesquelles la valeur de la variance estimée est "anormalement" élevée ou "anormalement" faible. L'exemple d'un segment constitué uniquement de pixels surexposés est typiquement problématique car alors  $\hat{\boldsymbol{\theta}}_k = \mathbf{z}_k$  et la valeur estimée de la variance est  $\sigma_k = 0$ . Numériquement la prise en compte de ce segment n'est pas souhaitable. En pratique la règle *ad-hoc* arbitrairement adopté est la suivante : tout segment pour lequel  $\sigma_k \notin [1/6\bar{\sigma}; 6\bar{\sigma}]$  n'est pas pris en compte dans le calcul du log-RV. Sur l'ensemble des 9 076 images de la base BOSS [184], cette règle a conduit à ne pas considérer moins de 1% des segments
- De manière analogue, le phénomène de saturation du quantificateur, ou d'écrêtage du signal, est négligé dans l'ensemble de ce manuscrit, ce qui revient concrètement à supposer que le pas de quantification  $\Delta_{can}$  est fixé de sorte que  $\forall n \in \mathcal{I}, \mathbb{P}[y_n > (2^b - 1/2)\Delta_{can}] = 0$ , avec  $b$  le nombre de bits utilisé pour coder la valeur (entière) des pixels. Cette hypothèse n'a qu'une raison pédagogique de clarté, évitant de systématiquement discerner les cas  $z_n = 0$ ,  $z_n = 2^b - 1$  et  $z_n \in \{2, \dots, 2^b - 2\}$  dans les expressions. En pratique il est simple de modifier les expressions des distributions de probabilité (3.1) et (3.3) afin de tenir compte du phénomène de clipping ; l'expression analytique de propriétés statistiques des test proposés s'en trouve en revanche largement compliquées.
- Il est également notable que la proposition 6.2 ainsi que le théorème 6.1 repose sur l'hypothèse que chacun des segments ne contient pas plus d'une unique discontinuité. Néanmoins l'extension de cette proposition et de ce théorème au cas de segment pouvant contenir plusieurs discontinuités est immédiat, au prix de notations moins claires, en modifiant en conséquence les paramètres  $b_{\max}$  et  $\kappa$ , voir démonstration 6.1 de l'annexe 6.5.2 pour plus de détails.
- Enfin, les hypothèses de travail supposées vérifiées par le système d'acquisition sont plus difficile à discuter. Comme détaillé dans la section 4.2, l'approximation de la PSF par un noyau Gaussien, non-stationnaire par changement d'échelle, n'est pas justifiable formellement. On notera cependant que dès lors qu'un segment peut être modélisé par (6.1) :

$$\boldsymbol{\theta}_k = \mathbf{H}\mathbf{c}_k + \mathbf{F}(\boldsymbol{\eta}_k)\mathbf{u}_k = \mathbf{G}(\boldsymbol{\eta}_k)\mathbf{v}_k,$$

où l'application  $\mathbf{F}(\boldsymbol{\eta}_k)$  est deux fois continûment différentiable par rapport au paramètres non-linéaire  $\boldsymbol{\eta}_k$ , l'approche par linéarisation demeure applicable. On comprendra également que supposer que la PSF est différente sur chacun des

segments, outre la modélisation *a priori* de la PSF sur chaque segment, nécessite d'appliquer l'approche par linéarisation avec une variété  $\mathbf{F}(\eta_k)$  différente pour chaque discontinuité ce qui est calculatoirement hors de portée.

## 7.2 Résultats numériques sur des images naturelles : comparaisons avec l'état de l'art

Dans la directe lignée du chapitre 6 précédent, il est proposé de présenter dans cette première section des résultats numériques sur le test statistique  $\hat{\delta}(\mathbf{Z})$  (6.22), proposé pour le cas pratique d'une image inconnue. Il est crucial de comparer les performances du test proposé avec les algorithmes de stéganalyse proposés dans la littérature. D'une part, la majorité des algorithmes proposés récemment reposent sur l'apprentissage supervisé et notamment sur les SVM, voir section 2.3; compte-tenu des enjeux des travaux présentés, il a été choisi de ne pas inclure ce type de classifieur dans les résultats présentés. En effet la problématique de cette thèse est de proposer une méthodologie exploitant un modèle *ad-hoc* des images naturelles pour permettre *in fine* le respect d'une probabilité de fausse alarme  $\alpha_0$  préalablement fixée. En outre, dans le contexte opérationnel envisagé, l'utilisation d'une base d'image en vue de l'apprentissage statistique n'est pas nécessairement disponible.

Parmi les méthodes de détection d'informations cachées dans les bits de poids faibles d'une image, il semble assez naturel de comparer le test statistique proposé avec le WS [91, 92, 93] en raison de leurs formalismes similaires décrits dans la section 3.6. De même, parmi les quelques tests statistiques proposés dans la littérature, voir section 2.3.3.2, celui proposé par Dabeer & al. dans [89] est le compétiteur le plus intéressant en raison de ses relatives bonnes performances et de son utilisation de la théorie de la décision statistique; la comparaison formelle entre l'approche de [89] et celle proposée dans ce manuscrit fait l'objet de la remarque 3.1. Le test du  $\chi^2$  [85] a été également inclus dans les comparaisons numériques car il s'agit du premier test proposé en 1999 et il demeure intéressant de constater les améliorations de performance réalisées depuis lors. Enfin, malgré leur manque de fondement statistique, les détecteurs structuraux ont une popularité telle dans la communauté "stéganographie/stéganalyse" que la comparaison avec ces derniers semble incontournable. Aussi, le RS [81] et le SPA/LSM [82] ont été inclus dans les comparaisons numériques.

Les figures 7.1 illustrent graphiquement, sous la forme de courbe COR (avec une échelle logarithmique), les résultats obtenus sur deux bases d'images usuellement utilisées dans le domaine de la criminalistique des images (base BOSS [184] constituée de 9 076 images pour la figure 7.1b et base UCID [185] constituée de 1338 images pour la figure 7.1a). Bien que le taux d'insertion utilisé  $R = 0.05$  est identique pour l'ensemble des images, l'écart-type du bruit n'est pas identique pour ces images; cet aspect est discuté ci-dessous et illustré dans la figure 7.3.

De manière analogue, les figures 7.2 illustrent graphiquement, sous la forme de courbes COR (avec une échelle linéaire), les résultats obtenus sur les deux mêmes bases d'images en utilisant un taux d'insertion  $R = 0.1$ .

Les figures 7.1 et 7.2 montrent que le WS offre une puissance de détection légèrement supérieure à celle du test proposé, pour une probabilité de fausse alarme "intermédiaire", pour  $0.2 \lesssim \alpha_0 \lesssim 0.4$  selon les courbes présentées dans les figures 7.2 avec  $R = 0.1$ . Néanmoins, il convient de s'interroger sur la pertinence d'une telle probabilité de fausse alarme dans le contexte opérationnel des travaux présentés. Par exemple, lors de l'analyse d'un disque dur il ne semble pas opportun de fixer une telle probabilité de fausse alarme; en pratique il est suggéré de respecter une contrainte de l'ordre de  $\alpha_0 = 10^{-3}$  voire moins... Or les résultats présentés dans les figures 7.1 et 7.2 montrent très clairement que le test proposé possède une puissance largement meilleure que ses principaux compétiteurs pour des probabilités de fausses alarmes

Dans les figures 7.1 et 7.2, les messages insérés, par un programme ad-hoc, sont des vecteurs binaires i.i.d. distribués suivants une lois uniforme  $\mathcal{U}(\{0;1\})$ .



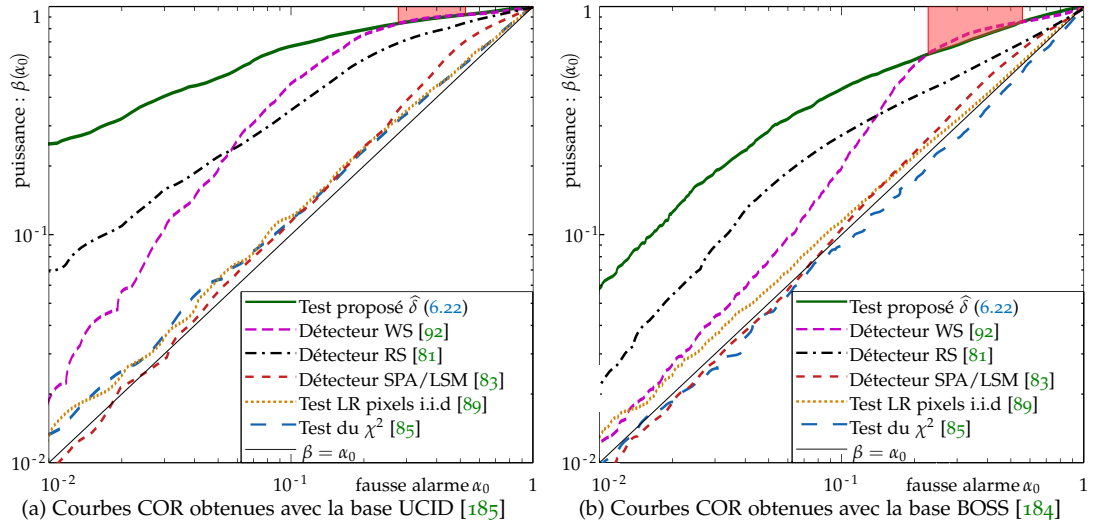


FIGURE 7.1: Comparaison numérique de la performance des détecteurs : courbes COR,  $\beta(\alpha_0)$ , en échelle logarithmique obtenues pour différents algorithmes avec  $R = 0.05$ .

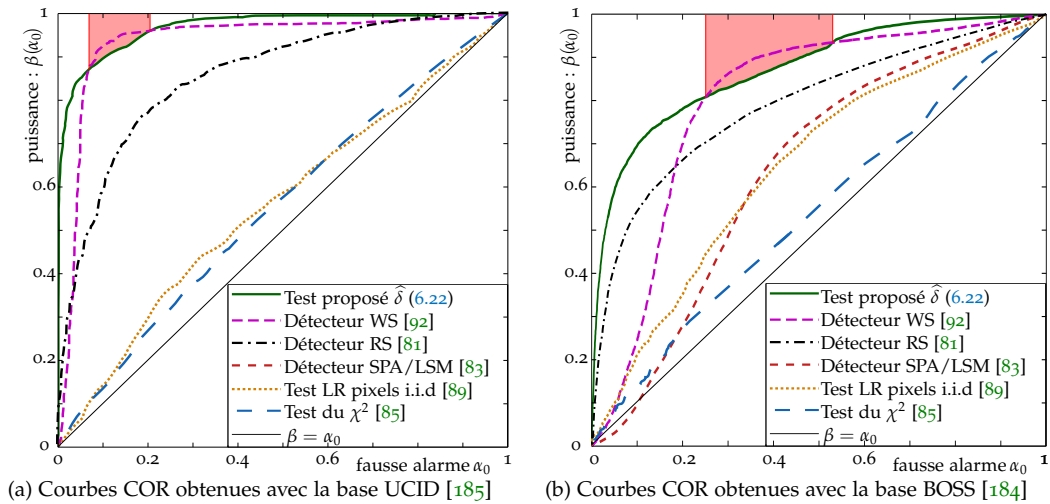


FIGURE 7.2: Comparaison numérique de la performance des détecteurs : courbes COR,  $\beta(\alpha_0)$ , obtenues pour différents algorithmes avec  $R = 0.1$ .

faibles.

Il est par ailleurs notable que pour des faibles taux de fausses-alarmes, la puissance de détection du WS devient inférieure à celle du détecteur du R-S, bien que ne reposant sur aucune exploitation de la théorie de la décision statistique. Cette constatation, déjà été soulignée dans [91], demeure cependant complexe à expliquer formellement compte-tenu du fonctionnement.

Enfin, le lecteur attentif constatera que la puissance de tous les détecteurs est plus importante sur la base d'image UCID, figures 7.1a et 7.2a, en raison du nombre de pixels,  $512 \times 384$  plus importants que sur les images BOSS, figures 7.1b et 7.2b, qui ont été préalablement réduites à une taille de  $256 \times 256$  pixels par "rognage".

Afin de mieux comprendre et d'interpréter plus finement les résultats présentés dans les figures 7.1 et 7.2, il est nécessaire de comparer la distributions empiriquement observée des statistiques de décision utilisées par différents tests. Plus précisément, les figures 7.3 présentent les distributions empiriquement obtenues de la statistique de décision du WS, du RS et du test proposé, sous  $\mathcal{H}_0$  sur la base d'image UCID (les images

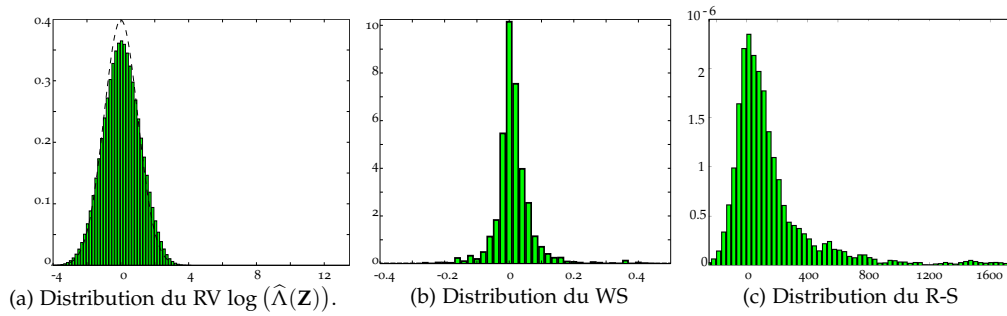


FIGURE 7.3: Distributions empiriques de la statistique de décision du test  $\hat{\delta}$  proposé, du WS et du RS sous l'hypothèse  $\mathcal{H}_0$ .

sont donc analysées sans modifications aucunes). Bien que ces images présentent des caractéristiques différentes (écart-type du bruit notamment) il est notable que conformément à ce qui est établi dans la section 6.4, le RV  $\hat{\delta}(Z)$  semble empiriquement distribué suivant une loi Gaussienne  $\mathcal{N}(0, 1 + \zeta)$  (représentée par des pointillés noirs sur la figure 7.3a). À l'inverse, les lois de distribution empirique du WS, figure 7.3b, et du RS, figure 7.3c, présentent des "queues lourdes". Autrement dit, pour respecter une probabilité de fausse alarme très faible, il est nécessaire de fixer un seuil de décision très élevé, ce qui pénalise en conséquence la puissance découlant du test. En outre, les distributions statistiques du WS et du RS ne sont pas clairement établies, bien qu'une étude empirique soit présentée dans [95]; la seule façon de respecter une contrainte sur la probabilité de fausse alarme est donc de recourir à une simulation numérique puis de fixer le seuil de décision en conséquence, ce qui ne garantit en rien le respect de la contrainte pour d'autres images (de tailles différentes, avec un bruit plus important, etc ...).

Les deux principales différences dans l'expression mathématique du WS et du test statistique proposé résident dans le facteur de pondération, théoriquement établi pour le test proposé, et l'exploitation d'un modèle d'images *ad hoc* pour l'estimation du contenu  $\hat{\theta}_k$ , là où WS exploite un modèle auto-régressif assez simple. Pour mettre en évidence l'intérêt de l'utilisation du modèle des images naturelles décrit dans les chapitres 4 - 5, une simulation Monte-Carlo a été menée sur une trentaine d'images naturelles "fortement texturées". Ces images ont donc subies un ajout de bruit (Gaussien i.i.d avec  $\sigma = 0.5$ ) et ont été testées par le WS et le test statistique

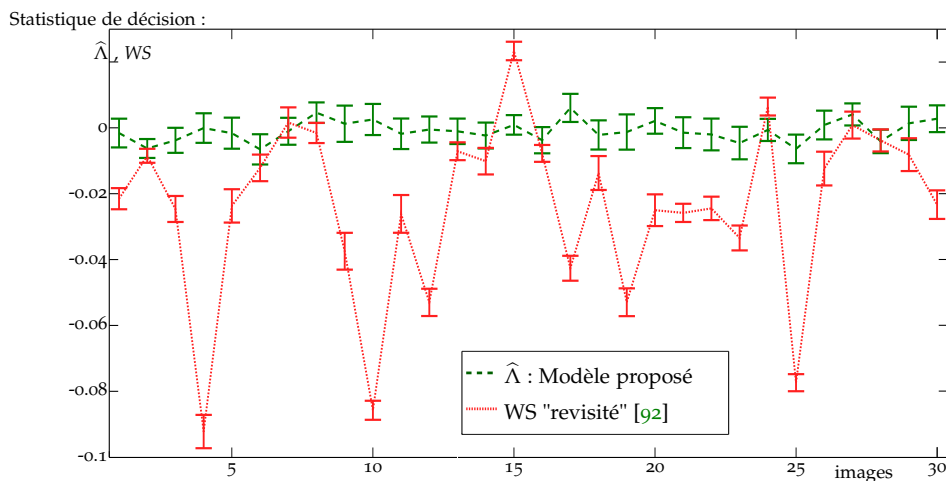


FIGURE 7.4: Espérance et écart-type du WS et du RV  $\log(\hat{\Lambda})$  obtenus par simulation Monte-Carlo en utilisant des images texturées.

	image 7.10a texturée	image 2.3a "classique"
WS	$\approx 0.198$	$\approx 0.197$
Algorithme proposé ( $L = 32, p = 3$ )	$\approx 0.989$	$\approx 0.506$
( $L = 16, p = 3$ )	$\approx 0.535$	$\approx 0.353$
Modèle linéaire ( $L = 32, p = 3$ )	$\approx 0.108$	$\approx 0.105$
( $L = 16, p = 3$ )	$\approx 0.103$	$\approx 0.101$

TABLE 7.2: Temps de calcul nécessaire à l'exécution de l'algorithme proposé en comparaison avec le WS sur des images de  $1024 \times 1024$  pixels. L'algorithme proposé nécessitant l'inversion d'une matrice de taille  $L \times L$ , où  $L$  est la taille des segments analysés, pour le calcul de  $\mathbf{P}_{\mathbf{G}_{\eta_k}}^\perp$ , les deux cas  $L = 16$  et  $32$  sont reportés dans le tableau. Pour information, les calculs ont été réalisés avec un processeur Intel®Core(TM)2 Duo CPU T8300, 2.40GHz et le système d'exploitation Unix / Ubuntu 10.04.

$\hat{\delta}$  proposé. Pour chacune des images, la figure 7.4 représente la moyenne et l'écart-type empiriques des statistiques de décision du WS (courbe rouge continue) et du test  $\hat{\delta}$  (courbe verte pointillée), *i.e.* le RV  $\log(\hat{\Lambda}(\mathbf{Z}))$ . Il est important de noter que les résultats ont été normalisés car le WS propose une estimation de  $R$  et fournit donc un résultat supposé être dans  $[0; 1]$  alors que  $\log(\hat{\Lambda}(\mathbf{Z}))$  peut tout à fait prendre des valeurs bien plus grandes que 1. Le RV  $\log(\hat{\Lambda}(\mathbf{Z}))$  est donc multiplié par  $2\bar{\sigma}/\sqrt{\bar{\kappa}}$  afin que l'espérance de ces deux statistiques convergent vers  $R$ , qui ici vaut 0, voir relation (6.24).

La figure 7.4 montre clairement que les variances des deux statistiques de décision sont très similaires sur l'ensemble des trente images testées. En revanche, il est notable que l'espérance du WS est assez sporadique. En pratique le modèle d'image auto-régressif sur lequel repose le WS n'est pas très réaliste et l'inadéquation de ce modèle avec les données peut biaiser de façon importante les résultats. À l'inverse, le modèle d'image proposé représente fidèlement le contenu des images; ce biais d'estimation devient alors négligeable (notamment en comparaison avec l'écart-type).

Il a été souligné que, dans le cadre applicatif du projet ANR RIC, les partenaires souhaitent que le coût calculatoire de l'algorithme soit aussi faible que possible. Aussi, il est présenté dans le tableau 7.2 quelques résultats sur le temps d'analyse de l'algorithme proposé, implémenté en langage C, en comparaison avec l'algorithme du WS, également implémenté en C en utilisant les mêmes bibliothèques. À titre de comparaison, le temps de calcul du test  $\delta^{\text{TVG}}$  est également présenté. Ce test est similaire au test proposé mais repose sur un modèle paramétrique linéaire de la scène. Le tableau 7.2 montre que, sans surprise, le coût calculatoire du test proposé dépend de la structure de l'image puisque le modèle est linéarisé autour de chacune des discontinuités. En outre, il est remarquable que si le temps de calcul du test proposé est supérieur au temps de calcul du WS, l'ordre de grandeur reste sensiblement identique.

Par ailleurs, il a été présenté dans la section 3.7 une extension de l'approche au cas de la détection d'informations cachées par correspondance de LSB. En effet, l'un des principaux apports des travaux présentés réside dans le fait que l'utilisation de la théorie de la décision statistique permet la stéganalyse quel que soit le schéma d'insertion d'informations supposé utilisé. Aussi, bien que le centre d'intérêt des travaux présentés soit la détection d'informations cachées par substitution des LSB d'une image, il est intéressant d'illustrer graphiquement la puissance du test  $\delta_R^\pm$  proposé dans le cas d'informations cachées par correspondance de LSB.

Il est notable qu'aucune adaptation des détecteurs précédemment présentés et utilisés pour la détection de stéganographie par substitution de LSB n'a été proposée pour la détection de stéganographie par correspondance de LSB. Il est donc nécessaire, à des

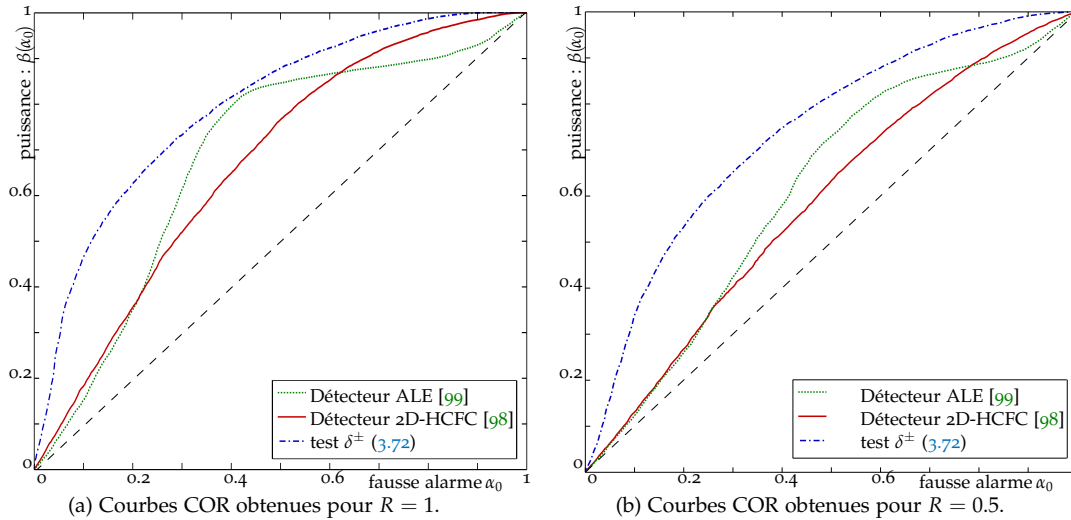


FIGURE 7.5: Comparaison numérique de la performance des détecteurs adaptés à la correspondance de LSB : courbes COR,  $\beta(\alpha_0)$  pour différents algorithmes avec  $R = 1$  (a) et  $R = 0.5$  (b).

fins de comparaison, d'utiliser d'autres détecteurs. Comme cela a été brièvement évoqué dans la section 2.3.4, peu de travaux ayant pour objet la détection d'informations cachées par correspondance de LSB ont été publiés. Les deux détecteurs les plus récemment publiés [98, 99] ont été choisis ; ces derniers sont basés sur la modification de l'histogramme engendrée par l'insertion d'informations dans les bits de poids faibles, et diffèrent donc fondamentalement de la méthodologie proposée dans ce manuscrit.

Il est ici uniquement présenté les résultats empiriquement obtenus pour la base d'images BOSS [184], en raison notamment du grand nombre d'images qu'elle contient. Les performances des détecteurs sont illustrées graphiquement, sous la forme de courbes COR, par la figure 7.5a pour le cas  $R = 1$ , et par la figure 7.5b pour le cas  $R = 0.5$ . Les figures 7.5 montrent que le test statistique proposé  $\delta^\pm$  offre une puissance de détection supérieure aux deux détecteurs, pour le cas d'un taux d'insertion  $R = 0.5$ . La différence de puissance est légèrement moindre pour  $R = 1$ . En outre, comme détaillé dans la section 3.7, les performances du test  $\delta^\pm$  ont été établies statistiquement ce qui permet de fixer le seuil de décision en fonction de la probabilité de fausse alarme  $\alpha_0$  souhaitée et permet de calculer la puissance de détection attendue en fonction de l'écart-type du bruit et du taux d'insertion ; cela n'est pas possible pour les deux autres détecteurs [98, 99] car leurs performances ne sont connues qu'empiriquement.

Enfin, l'une des principales contributions des travaux présentés réside dans le développement d'un modèle des images naturelles adapté à la détection d'informations cachées. Pour mettre en évidence l'intérêt de ce modèle il est proposé de comparer les performances de détection empiriquement obtenues par le test  $\delta^\pm$  en utilisant deux autres modèles d'images pour estimer le contenu  $\theta$ .

Plus précisément, il est rappelé que le test statistique  $\delta^\pm$  (3.72) est basé par le RVG donné par (3.76) :

$$\Lambda_R^\pm(\mathbf{Z}) = \sum_{k=1}^K \frac{1}{4} \exp\left(-\frac{1}{2\sigma_n^2}\right) \sum_{l=1}^L \left[ \exp\left(\frac{(k - \theta_{k,l})}{\sigma_k^2}\right) + \exp\left(\frac{-(k - \theta_{k,l})}{\sigma_k^2}\right) \right],$$

où, le paramètre  $\theta_k = (\theta_{k,1}, \dots, \theta_{k,L})^T$  est en pratique inconnu et est donc être substitué par l'estimation  $\hat{\theta}_k$ , qui, en utilisant le modèle d'image proposé, est donnée par la

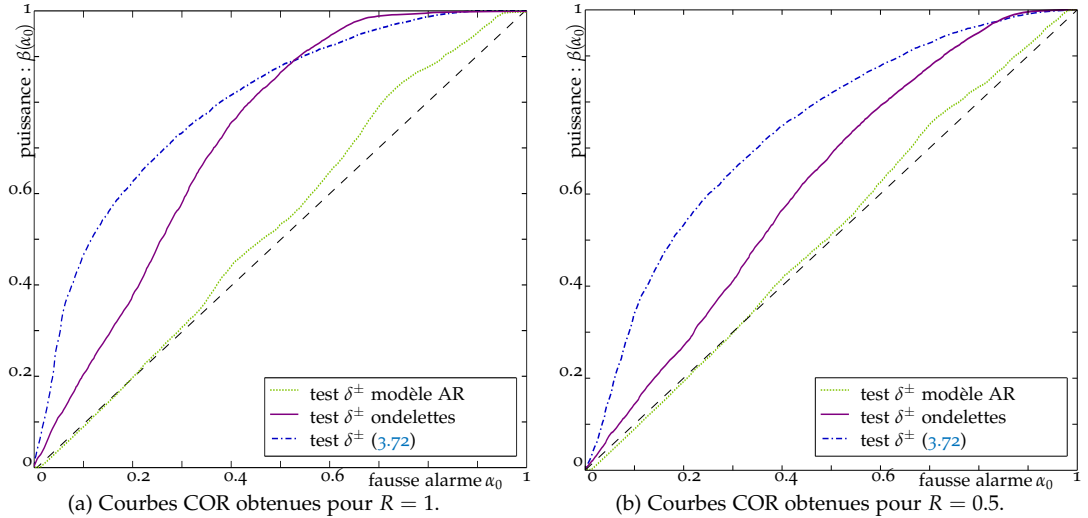


FIGURE 7.6: Comparaison numérique de la puissance des tests reposant sur différents estimateurs des paramètres  $\theta$  et  $\sigma$  : courbes COR obtenues avec  $R = 1$  (a) et  $R = 0.5$  (b).

projection  $\mathbf{P}_{\hat{\mathbf{G}}_{\hat{\eta}_k}} \mathbf{z}_k$ .

En omettant le problème consistant à respecter la contrainte sur la probabilité de fausse alarme (et accessoirement à expliciter la puissance du test), il est tout à fait envisageable d'utiliser en pratique une méthode d'estimation alternative du contenu  $\theta_k$ . Les résultats présentés dans la figure 7.6 offrent une comparaison graphique des résultats empiriquement obtenus en utilisant 1) une estimation de  $\theta$  par reconstruction partielle de transformée en ondelettes (voir section 5.3.1) et 2) une estimation reposant sur le modèle auto-régressif utilisé par le WS. La figure 7.6a présente la puissance des tests, sous la forme de courbes COR, empiriquement obtenues en utilisant la base d'images BOSS [184] avec un taux d'insertion  $R = 1$ ; de manière analogue, la figure 7.6b illustre la puissance des tests pour un taux d'insertion de  $R = 0.5$ .

### 7.3 Vérifications numériques des résultats théoriques

Le lecteur attentif aura noté que l'aboutissement des travaux présentés pour la détection d'informations cachées dans les bits de poids faibles d'une image naturelle réside dans les résultats présentés dans la section 6.4. Il semble donc primordial d'illustrer graphiquement ces résultats.

Les figures 7.7 et 7.8 offrent une représentation graphique des résultats (6.24) - (6.28) au travers d'une simulation Monte-Carlo avec  $2.5 \cdot 10^4$  répétitions. A chaque itération, une "image" constituée de 400 segments de 32 pixels est analysée. Chacun des segments est décrit par le modèle (5.10) - (5.11) et comporte une unique discontinuité, dont la localisation est uniformément distribuée, de paramètres  $u_k=96$ ,  $\zeta_k = 1.75$ ,  $x_k = 0.25$  et la valeur  $\hat{\eta}_k = (x_k, \zeta_k)$  utilisée pour la linéarisation vérifie  $|\eta_k - \hat{\eta}_k| < \vartheta = 1$ . Un polynôme algébrique de degré 3 (avec des coefficients fixes) est également utilisé pour la partie linéaire du modèle. Le taux d'insertion utilisé est de  $\simeq 0.47$  et un bruit i.i.d Gaussien d'écart-type  $\sigma = 5,43$  est ajouté avant quantification avec un pas  $\Delta = 1$ .

La figure 7.7 présente une comparaison des distributions théoriques et empiriques des RV  $\log(\hat{\Lambda}(\mathbf{Z}))$  (cas d'une image inconnue) et  $\log(\bar{\Lambda}(\mathbf{Z}))$  (cas du contenu  $\theta_k$  connu). Il apparaît clairement dans la figure 7.7 que les distributions empiriquement observées sont sensiblement identiques aux distributions théoriquement établies, à l'exception notable du facteur  $b$  d'accroissement de variance qui demeure inconnu, voir (6.24).

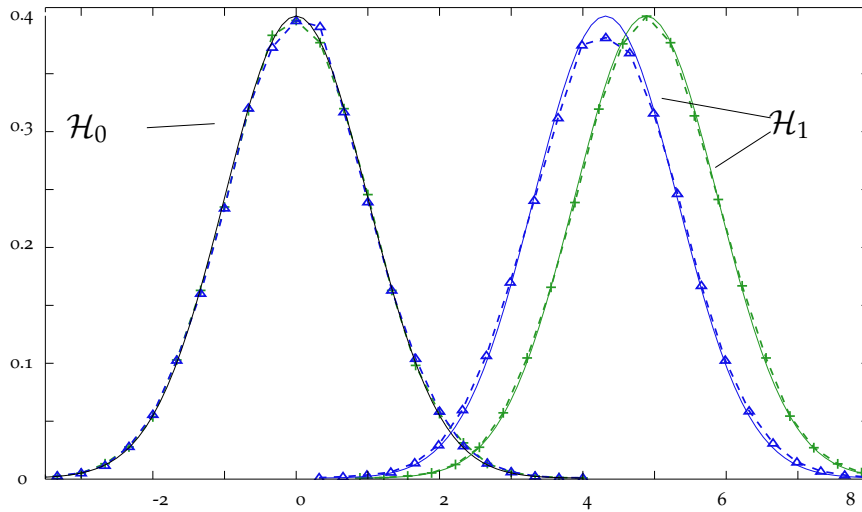


FIGURE 7.7: Distributions théoriques et empiriques des LR  $\bar{\delta}$  (+) et du GLR  $\hat{\delta}$  ( $\Delta$ ); le LR nécessite la connaissance de  $\theta_k$ , alors que le GLR exploite le modèle d'image proposé.

La figure 7.8 offre une comparaison graphique des fonctions de puissance, représentées sous la forme de courbes COR, de trois différents tests afin de mettre en évidence les facteurs de perte de puissance du test pratique  $\hat{\delta}(\mathbf{Z})$ . Les trois tests graphiquement comparés sont le test AUPP  $\bar{\delta}(\mathbf{Z})$  (3.62), applicable lorsque le contenu  $\theta_k$  est supposé connu, le test  $\bar{\delta}^{\text{rvg}}$  (3.65), reposant sur une modélisation linéaire de l'image et nécessitant donc la connaissance des paramètres  $\eta_k$ , et le test proposé en pratique  $\hat{\delta}(\mathbf{Z})$  (6.22), reposant uniquement sur la linéarisation du modèle (5.10) - (5.11).

La perte de puissance du test  $\bar{\delta}^{\text{rvg}}$  (représenté en bleu sur la figure 7.8) est uniquement due à la diminution du nombre de "degré de liberté" introduit par le modèle linéaire. En outre, la perte de puissance du test  $\hat{\delta}(\mathbf{Z})$  (représenté en rouge sur la figure 7.8) est due à une diminution accrue du nombre de "degré de liberté", puisque la linéarisation

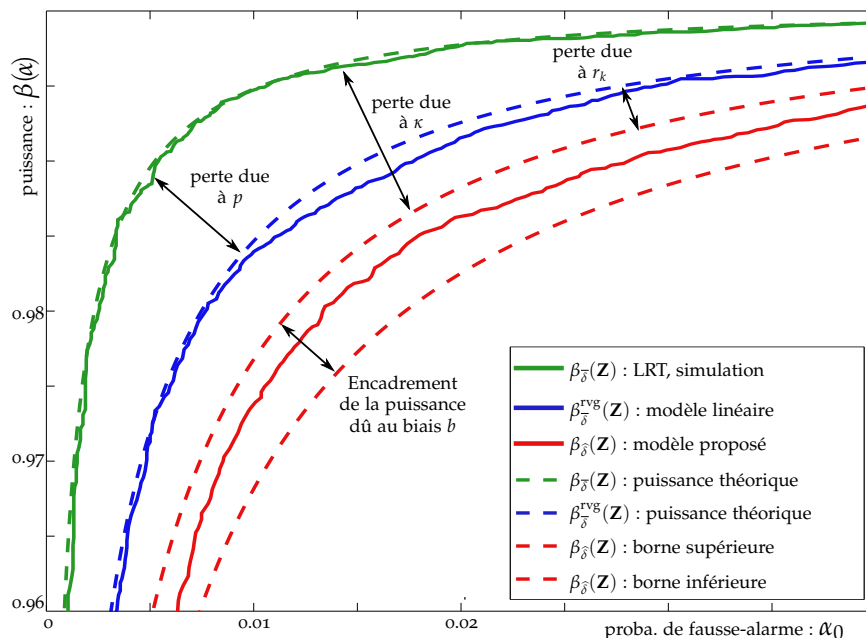


FIGURE 7.8: Représentation graphique, sous la forme d'une courbe COR de la puissance des tests  $\bar{\delta}(\mathbf{Z})$  (vert),  $\bar{\delta}^{\text{rvg}}(\mathbf{Z})$  (bleu) et  $\hat{\delta}(\mathbf{Z})$  (rouge); le but est ici de mettre en évidence les différents facteurs de perte de puissance du test proposé.

du modèle introduit de nouveaux vecteurs dans le modèle, et au biais d'estimation du contenu  $\hat{\theta}_k$  qui demeure inconnu, d'où la présence de bornes sur la fonction de puissance.

Toujours dans un but pédagogique, la figure 7.9 présente les RV  $\hat{\Lambda}(z_k)$  calculés par colonnes pour le cas de deux images contenant des informations cachées par substitution de LSB avec un taux  $R = 0.28$ . Dans l'image 7.9a, les informations sont insérées dans un ordre lexicographique, suivant les colonnes et les lignes; les 732 premières colonnes contiennent un taux  $R=1$  d'informations cachées et les 1881 dernière colonnes ne contiennent pas d'informations cachées. À l'inverse, dans l'image 7.9b chacune des 1536 colonnes contient un taux  $R = 0.28$  d'informations cachées. Il est important de noter que les colonnes de l'image 7.9a sont constituées de 3900 pixels alors que les colonnes de l'image 7.9b ne sont constituées que de 1000 pixels (l'image a été réduite par rognage); pour cette raison les RV présentées dans la figure 7.9d sont calculées sur des blocs de 9 colonnes, afin de permettre une meilleure distinction visuelle des deux hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$ .

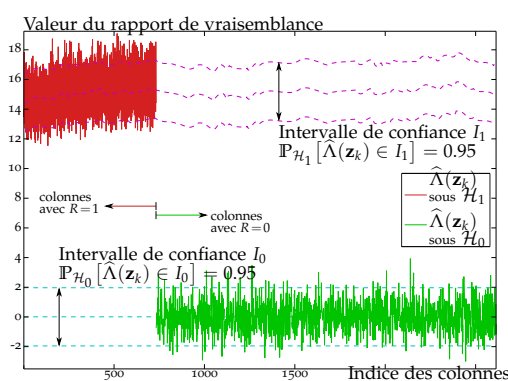
Ce formalisme de présentation des résultats, utilisé dans de articles précédemment pu-



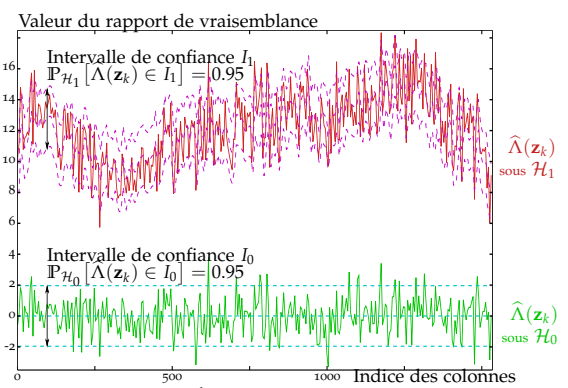
(a) Photographie stéganographiée lexicographiquement au taux  $R = 0.28$ .



(b) Photographie stéganographiée au taux  $R = 0.28$  pour chaque colonne.



(c) Valeurs du RV  $\hat{\Lambda}(z_k)$  obtenues par colonne.



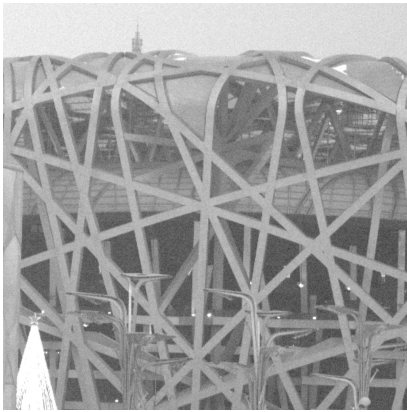
(d) Valeurs du RV  $\hat{\Lambda}(z_k)$  obtenues par colonne.

FIGURE 7.9: Représentation graphique du RV  $\hat{\Lambda}(z_k)$  illustrant les résultats (6.24) - (6.28). Les deux photographies analysées, présentées dans les figures 7.9a et 7.9b, contiennent un taux  $R=0.28$  d'informations cachées. Les résultats présentés dans les figures 7.9c et 7.9d ont été obtenus en analysant les images par colonnes.

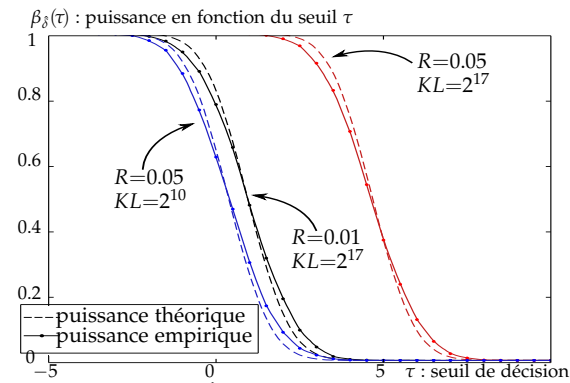
bliés tels que [85], a pour principal intérêt de représenter graphiquement la maîtrise des résultats obtenue par la méthodologie proposée. Le plus important dans les figures 7.9c et 7.9d réside dans la cohérence entre les résultats empiriquement obtenus et l'espérance mathématique du RV  $\hat{\Lambda}$  théoriquement calculée par  $\varrho_k \sqrt{\kappa_k} = \frac{R}{2\bar{\sigma}_k} \sqrt{\kappa_k}$  pour chacune des colonnes indexée par  $k$ . Afin de mettre en lumière cette maîtrise des résultats, les enveloppes correspondant à l'intervalle dans lequel le RV  $\hat{\Lambda}$  a une probabilité de 0.95 d'être présent sont également présentés.

Enfin, il est notable, particulièrement sur la figure 7.9d, que la non-stationnarité de l'image a une importance qui est loin d'être négligeable dans l'analyse d'une image; il est immédiat de constater que sous l'hypothèse  $\mathcal{H}_1$ , l'espérance mathématique du RV  $\hat{\Lambda}(z_k)$  varie grandement selon les colonnes utilisées en raison de changement dans l'écart-type du bruit  $\bar{\sigma}_k$  et dans le contenu analysé (modifiant le nombre de paramètres libres  $\kappa_k$ ).

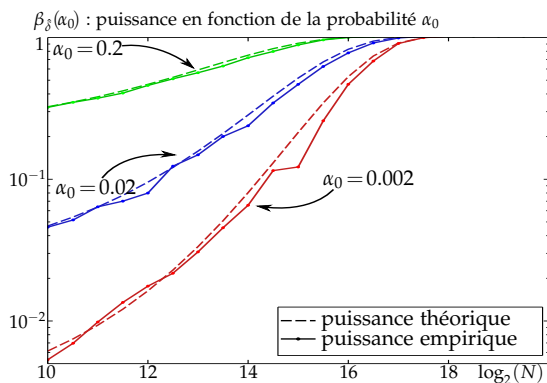
Lors de la soumission de l'article [186] pour la conférence *Information Hiding*, l'un des "reviewers" a suggéré le recours à des simulations Monte-Carlo sur des images naturelles dans le but de vérifier si la puissance du test empiriquement obtenue, en fonction du seuil de décision, était en adéquation avec la puissance théoriquement calculée. De manière légèrement plus exhaustive, il est proposé de confronter les résultats présentés dans la section 6.4, sur les performances du test statistique proposé, au cas pratique d'analyse d'une image naturelle acquise par un appareil photographique. À cette fin, les figures 7.10b, 7.10c et 7.10d présentent sous différentes formes la com-



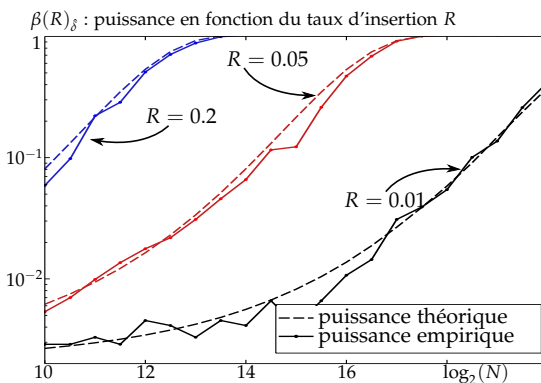
(a) Photographie numérique du nid d'oiseau utilisée pour les simulations



(b) Puissance du test  $\hat{\delta}(\mathbf{Z})$  (6.22) en fonction du seuil de décision : théorie et simulation



(c) Puissance du test  $\hat{\delta}(\mathbf{Z})$  (6.22) en fonction du nombre de pixels pour  $\alpha_0 = \{2 \cdot 10^{-3}; 2 \cdot 10^{-2}; 2 \cdot 10^{-1}\}$  : théorie et simulation



(d) Puissance du test  $\hat{\delta}(\mathbf{Z})$  (6.22) en fonction du nombre de pixels pour  $R = \{0.01; 0.05; 0.2\}$  : théorie et simulation

FIGURE 7.10: Vérification des résultats de la section 6.4 au travers de simulations numériques basées sur une image naturelle.



paraison entre les résultats théoriques et ceux obtenus empiriquement pour l'image naturelle, constituée de  $1024 \times 1024$  pixels, présentée dans la figure 7.10a. Cette image n'a pas été choisie de façon arbitraire, mais en raison de sa structure rendue complexe par la présence de nombreuses discontinuités, de façon à mettre en évidence l'intérêt du modèle proposé. Dans les figures présentées, l'analyse de cette image a été répétée à  $2.5 \cdot 10^3$  reprises après ajout d'un bruit Gaussien i.i.d, de moyenne nulle et de variance  $\sigma^2 = 1$ , et ajout d'informations cachées, par substitution de LSB, avec différents taux d'insertion.

La figure 7.10b présente la puissance du test proposé en fonction du seuil de décision  $\tau_{\alpha_0}$ . De manière analogue, la figure 7.10c illustre, sous la forme d'une courbe COR, la puissance de détection en fonction du taux de fausse alarme préalablement fixé. Enfin la figure 7.10d représente graphiquement la puissance du test proposée en fonction du taux d'insertion, pour une probabilité de fausse alarme constante  $\alpha_0 = 0.02$ .

Il est notable que pour les trois figures 7.10b, 7.10c et 7.10d, les résultats empiriquement obtenus sont en adéquation avec les résultats théoriquement calculés.

### Conclusions du chapitre :

Les résultats numériques présentés dans ce chapitre 7 permettent, dans leur ensemble, de confirmer la pertinence de la méthodologie proposée. La bonne puissance de détection empiriquement observée est toutefois à relativiser ; il a été souligné que pour un niveau fausse alarme important (typiquement  $\alpha_0 \gtrsim 0.15$ ), le détecteur du WS peut fournir une puissance légèrement supérieure. Néanmoins, le cadre applicatif du projet ANR RIC dans lequel s'inscrit ces travaux impose que la probabilité alarme soit très faible. Dans ces conditions, le test proposé offre une puissance de détection bien supérieure à celle des autres tests précédemment publiés ; en outre, et ce point est probablement le plus important, il a été vérifié empiriquement, aussi bien sur des données simulées qu'avec des images réelles, que la valeur du seuil théoriquement établie permet effectivement de garantir une probabilité de fausse alarme prescrite.



«*Une suite de petites volontés fait un gros résultat.*»  
— Charles Baudelaire, 1821 – 1867

«*Yes, there are two paths you can go by, but in the long run,  
There's still time to change the road you're on.* ».  
*And it makes me wonder.* ».

— Robert Plant, 1948 – ...

### 8.1 Synthèse des travaux présentés

Il ressort de l'état de l'art présenté dans le chapitre 2 que dans le domaine de la détection d'informations cachées, les approches issues de la théorie de la décision statistique demeurent très largement minoritaires. De même les seuls modèles d'image qui ont été utilisés dans ce cadre reposent davantage sur des considérations empiriques dont l'exactitude n'est pas vérifiable, telle la distribution Laplacienne des coefficients DCT, et qui ne tiennent ni compte du contenu structuré des images ni de considérations physiques quelconques. Les travaux présentés dans ce manuscrit ont été motivés par ce vide laissé, de manière plus générale, entre l'analyse criminalistique des images (ou *digital forensics*), la modélisation des images naturelles et la théorie de la décision statistique.

Un premier pas vers l'application des méthodes de décision statistique pour la détection d'informations cachées est présenté dans le chapitre 3. Le problème de la stéganalyse est abordé, avec un point de vue purement statistique, comme un problème de décision entre deux hypothèses, éventuellement composites, à base d'observations quantifiées. Le principal apport de l'approche proposée réside dans l'expression analytique des probabilités d'erreur ce qui permet en pratique le respect d'une contrainte sur la probabilité de fausse alarme. Les principaux résultats obtenus dans ce cadre sont :

1. Dans la lignée des travaux de J. Neyman et E. Pearson, le test le plus puissant (PP) est établi lorsque les hypothèses testées sont simples, *i. e.* dans le cas où le contenu de l'image et le taux d'insertion sont connus. L'expression de la fonction de puissance du test PP permet de mettre en évidence l'impact de la quantification sur les performances du test.
2. Lorsque les hypothèses testées sont composites, ou que le taux d'insertion n'est pas connu, l'approche de L. Le Cam permet de proposer un test localement asymptotiquement le plus puissant qui montre combien la perte de puissance peut être importante si le taux d'insertion testé diffère du taux d'insertion effectif.
3. Enfin, pour le cas où le pas de quantification est faible, une approche complémentaire est présentée; cette dernière permet de proposer un test statistique asymptotiquement le plus puissant uniformément par rapport au taux d'insertion.

Avec cette approche purement statistique, le problème le plus délicat est l'intervention du contenu de l'image comme un paramètre de nuisance. Une première ébauche

d'approche reposant sur une modélisation paramétrique locale et linéaire est présentée. Cette application pratique des tests proposés met en évidence une certaine coïncidence avec le détecteur du WS; ce dernier ayant historiquement été proposé sans exploitation aucune des méthodes de décision statistique, certains paramètres ne peuvent être fixés qu'empiriquement dans le WS alors que l'approche proposée permet de calculer théoriquement ces paramètres.

Pour une application pratique des tests proposés, il est nécessaire d'estimer la valeur des paramètres inconnus (pour chacun des pixels), ce qui ne peut se faire qu'à travers d'une modélisation de l'image analysée. Les images numériques considérées étant des images naturelles, acquises par un appareil photographique, il est proposé de construire un modèle des images numériques reposant sur l'analyse de la chaîne d'acquisition et sur les propriétés physiques de la scène. Le chapitre 4 est donc dévolu à la description et la modélisation de la chaîne d'acquisition d'une image naturelle. Cette chaîne d'acquisition est considérée sous deux angles distincts :

- Le système optique permettant la formation d'une image optique sur le plan du photo-détecteur est étudié. Les dégradations introduites à ce stade sont déterministes, une relation analytique entre la scène imagée et l'intensité de l'image optique est donc établie.
- L'acquisition numérique de l'image optique est brièvement décrite et les divers phénomènes de dégradation de nature aléatoire sont modélisés statistiquement.

La modélisation rigoureusement exacte de ces deux parties de la chaîne d'acquisition d'une image est très complexe; en conséquence l'exploitation directe dans un test statistique tout en explicitant analytiquement les performances de ce dernier est hors d'atteinte. Aussi, quelques suppositions sur le système d'acquisition permettent de simplifier le modèle de l'image acquise tout en conservant les propriétés physiques fondamentales de non-stationnaires du système d'acquisition.

Dans le chapitre 5, la scène est décrite comme la juxtaposition de différents objets solides au sein desquels les propriétés lumineuses varient peu et de façon continue. L'image optique que restitue le système optique d'une telle scène est modélisée. Là encore, un compromis entre complexité et exploitation pour la détection d'informations cachées est proposé en analysant une image suivant une dimension (typiquement une ligne). Il est notamment montré qu'en vertu du caractère continu par morceaux de la scène imagée, les lignes extraites de l'image peuvent être modélisées comme la somme d'une composante continue et d'une composante représentant les discontinuités entre les objets. Des éléments montrant que la partie continue peut être localement approximée de façon linéaire avec une bonne précision et un nombre réduit de paramètres sont présentés. À l'inverse une modélisation parcimonieuse de la partie représentant les discontinuités de la scène est difficile, cette dernière étant notamment non-linéaire par rapport à des paramètres décrivant physiquement les frontières entre les objets.

L'exploitation de ce modèle d'image pour la détection d'informations cachées est détaillée dans le chapitre 6. La principale difficulté que présente le modèle proposé réside dans la présence de paramètres de nuisance intervenant de façon non-linéaire. Compte-tenu de la finalité applicative envisagée, l'estimation statistique de l'image au moyen des méthodes d'optimisation usuelles dans le cadre de modèles non-linéaires est à proscrire en raison d'un coût calculatoire élevé. Pour surmonter ce problème de façon satisfaisante, avec une faible complexité calculatoire et une bonne précision, une approche par linéarisation du modèle est présentée. L'intérêt de cette approche est qu'une analyse du modèle permet de quantifier les erreurs de modélisation dues à cette procédure de linéarisation. Au final l'exploitation du modèle pour la détection d'informations cachées permet de proposer un test statistique dont les propriétés statistiques sont analytiquement établies ce qui permet :

- de respecter une contrainte sur la probabilité de fausse alarme;
- de garantir une perte d'optimalité bornée (par rapport au test le plus puissant).

En pratique l'application de la méthodologie de linéarisation repose sur l'estimation des paramètres décrivant les discontinuités ; pour obtenir cette estimation, une méthode issue de la littérature est implémentée.

Enfin, les résultats numériques présentés dans le chapitre 7 mettent en évidence la pertinence de l'approche proposée. D'une part, quelques résultats issus de simulation numériques confirment les résultats théoriques précédemment exposés. D'autre part, les résultats numériques obtenus sur de larges bases d'images mettent en perspective la bonne puissance de détection du test proposé, pour des faibles taux de fausse alarme, par rapport à l'état de l'art dans le domaine de la stéganalyse. Ces résultats expérimentaux soulignent l'intérêt des travaux présentés pour apporter une solution satisfaisant les contraintes fixées dans le cadre du projet ANR RIC.

## 8.2 Perspectives ouvertes à la suite des travaux menés

À la suite de cette brève synthèse des travaux présentés, il semble raisonnable de se questionner sur les problèmes laissés ouverts et qu'il serait souhaitable d'aborder dans un avenir plus ou moins proche. D'autres perspectives n'ayant pas été abordées dans ce manuscrit sont également présentées. Les pistes de travail ouvertes ont été schématiquement classées suivant les trois domaines de recherche abordés dans le présent mémoire :

- la falsification des images et la criminalistique numérique ;
- la modélisation des images naturelles ;
- la théorie de la décision et de l'estimation statistique.

### **Perspectives ouvertes dans le domaine de la criminalistique des images**

L'approche statistique proposée pour la détection d'informations cachées ouvre plusieurs perspectives, aussi bien dans le domaine restreint de la stéganalyse que dans le domaine plus vaste de la criminalistique des images.

- Tout d'abord, dans le prolongement direct des travaux proposés, il est envisageable d'exploiter l'aspect tri-chromatique des images naturelles usuelles pour la détection d'informations cachées. Dans le cadre des présents travaux, les images couleurs sont analysées comme autant d'images en niveaux de gris et la corrélation forte existante entre les différents canaux n'a pas été utilisée.
- Dans le même esprit, l'application de la théorie de l'estimation statistique pour la stéganalyse quantitative (en vue de l'estimation du taux d'insertion) est une piste envisageable. Ces travaux ne sont toutefois pas dans un prolongement directe car les critères de performance d'un estimateur sont différents (biais et variance notamment) et la comparaison avec un estimateur optimal atteignable n'est pas simple ; borner le plus précisément possible la variance ou l'EQM d'un estimateur demeure un problème ouvert.
- Toujours dans la suite des travaux de détection d'informations cachées, il pourrait être envisagé d'étendre l'approche proposée aux images naturelles compressées, au format JPEG notamment. Il est par exemple imaginable de proposer le détecteur optimal au sens de Neyman-Pearson lorsque les paramètres de distribution statistique des coefficients DCT sont connus, et, lorsque ces paramètres ne sont pas connus, de proposer un détecteur dont la perte de puissance est bornée.
- Dans le domaine plus vaste de la criminalistique des images, il est envisageable à court terme d'appliquer la méthodologie proposée (pour la détection d'informations cachées dans les LSB) à la détection de différentes manipulations des images. La détection de photomontage [187], par exemple, n'a pas encore été abordée dans la littérature sous la forme d'un test entre deux hypothèses statistiques ou en considérant le respect d'une contrainte sur le taux de fausse alarme.

- Enfin, dans le domaine de la criminalistique des images, des travaux de recherche ont été proposés [188], visant à identifier l'appareil utilisé pour capturer une photo donnée ou à déceler si deux images proviennent d'un même appareil. Là encore, une approche statistique permettant d'établir le caractère optimal d'un test ou le respect d'une contrainte sur la probabilité de fausse alarme serait très originale.

### Perspectives ouvertes dans le domaine de la modélisation des images

L'exploitation d'un modèle des images naturelles reposant sur la physique de la scène imagée et du système d'acquisition est une approche originale dans le domaine du traitement des images. Dans ce domaine, il serait intéressant de justifier physiquement certaines propriétés des images empiriquement observées.

- Le modèle statistique proposé pourrait par exemple permettre de justifier formellement la distribution des coefficients DCT (usuellement modélisée par une loi Laplacienne [20]) et la distribution des coefficients de transformée en ondelettes discrètes (usuellement modélisée par une loi Normale généralisée [189]).
- Le modèle déterministe des images proposé pourrait être étendu pour permettre de prendre en compte l'aspect bi-dimensionnel des images. Quelques difficultés fondamentales, esquissées dans la section 5.1, demeurent à surmonter. Néanmoins, la parcimonie du modèle serait accrue en conséquence. On peut en outre imaginer, par exemple, un grand pratique pour la description de contenu des images.

### Perspectives ouvertes dans la théorie de l'estimation et de la décision statistiques

Le problème de détection d'informations cachées à partir d'observations quantifiées est abordé d'un point de vue purement statistique dans le chapitre 3 et dans [23]. Cette approche a permis de mettre en évidence plusieurs difficultés fondamentales, liées à la quantification des données, dans l'application de la théorie de la décision statistique. Ces problèmes soulevés sont autant de perspectives de futurs travaux potentiels.

- Fondamentalement, l'impact de la quantification des observations sur les performances d'un test statistique entre hypothèses simples n'a pas été formellement établi. Or, comme le montrent les résultats présentés dans le chapitre 3, les probabilités d'erreurs d'un test peuvent être impactées de manière importante par la quantification.
- De manière analogue, l'impact de la quantification sur les performances d'un estimateur statistique n'a pas été formellement établi.
- Le théorème de Gauss-Markov n'est par exemple pas applicable à des observations quantifiées. Il serait intéressant de "mesurer" l'impact de la quantification sur l'estimation de paramètres intervenant de façon linéaire. Cela aurait notamment une application dans le cadre d'un test statistique en présence d'un paramètre de nuisance à partir d'observations quantifiées ; dans la continuité de [190], par exemple.

Enfin, suite aux travaux présentés, mon *Saint-Graal* serait d'exploiter la modélisation proposée des images naturelles dans le but de détecter toute image qui aurait subi une manipulation quelconque (ou, en étant plus réaliste, une manipulation d'une classe donnée). L'idée serait alors de concevoir un test visant à décider entre les deux hypothèses suivantes  $\mathcal{H}_0 = \{\mathbf{Z} \text{ est une image naturelle non-modifiée}\}$  vs  $\mathcal{H}_1 = \{\mathbf{Z} \text{ est une image naturelle modifiée}\}$ . Le formalisme statistique de ces hypothèses n'a pas encore été clarifié et, d'un point de vue statistique, ce test s'apparenterait davantage à la détection d'une anomalie quelconque abordée dans le cadre du traitement d'image dans [161, 115] notamment.

## 8.3 Conclusion

Les travaux menés durant ses trois années de doctorat ont permis de faire un premier pas vers l'application de la théorie de la décision statistique pour la détection d'informations cachées, et plus largement pour la criminalistique numérique. Les difficultés soulevées sont nombreuses : d'une part, le problème statistique est complexe (en raison de la quantification, d'hypothèses composites et de la présence de paramètres de nuisance, etc. . . .) et, d'autre part, la modélisation des images est complexe compte tenu de la structure particulière des ces objets (présences de courbes de discontinuités, système optique non-stationnaire, etc. . . .).

La pertinence de méthodologie proposée, consistant à étudier le problème sous un angle purement statistique et à développer un modèle des images naturelles adapté à une exploitation dans ce cadre, est confirmée par des résultats numériques encourageants. Ces résultats confirment en outre que la méthodologie proposée permet finalement de répondre de manière satisfaisante aux contraintes exigée par le projet ANR RIC dans lequel s'inscrivent ces travaux.

Il nous faut enfin souligner que les travaux présentés ouvrent également de nombreuses perspectives aussi bien dans le domaine de la théorie de la décision statistique, dans le domaine de la modélisation des images et, entre les deux, dans le domaine de la criminalistique numérique :

**Cette thèse n'est pas une fin, c'est un commencement.**





---

# Éléments de la théorie de la décision et de l'estimation statistique

---



Les méthodes statistique se déclinent en deux classes (selon la problématique envisagée) : les méthodes d'estimation et les méthodes de détection. Cette section s'intéresse davantage au problème de détection statistique qui est le principal intérêt pour la stéganalyse (non-quantitative) ; l'estimation statistique est abordée dans la section A.6.

## A.1 Motivations et notations

Le fondement de la théorie de la détection et de l'estimation statistique consiste à admettre que les observations dont nous disposons sont entachées d'erreurs de nature aléatoire liées notamment aux imperfections des capteurs, aux erreurs de modélisation et/ou de quantification etc. . . . La théorie de la décision statistique trouve sa place dans la stéganalyse car elle permet de :

- prendre en compte explicitement les erreurs (bruits d'acquisition, de quantification, erreurs de modélisation, etc. . . ) ;
- caractériser l'incertitude des résultats de la solution proposée ;
- tenir compte de paramètres de nuisance ou d'informations a priori ;
- maîtriser la probabilité de fausses alarmes ;
- définir l'optimalité d'un test statistique d'hypothèses.

Pour ces raisons notamment, la théorie de la décision et de l'estimation statistique offre une méthodologie permettant de répondre à la problématique qui est abordée dans cette thèse, voir section 1.2.

La présentation succincte, exposée dans le présent manuscrit, de la théorie de la décision statistique est inspirée de [165, 161, 191] et pourra donc être complétée par la lecture de ces ouvrages.

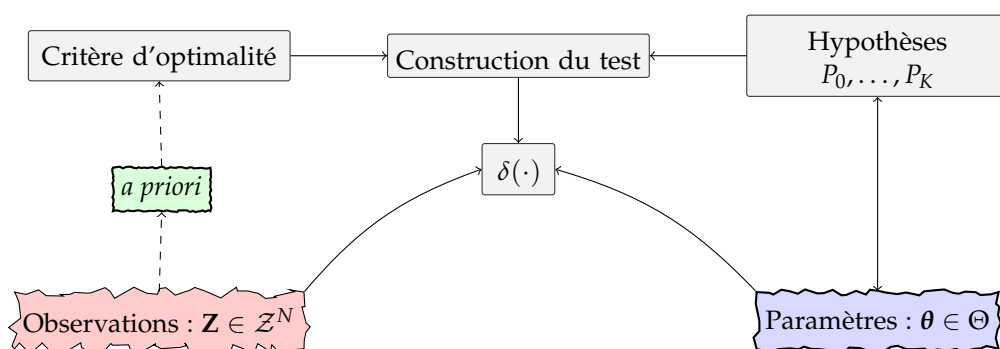


FIGURE A.1: Illustration d'un test statistique d'hypothèses.

Pour prendre en compte les multiples phénomènes physiques sources d'erreurs stochastiques, les observations  $\mathbf{Z}$  sont supposées, sans perte de généralité aucune, être la réalisation d'un vecteur aléatoire  $\mathbf{Z}$  suivant une loi de distribution connue  $\mathbf{Z} \sim P_\theta$  où  $\theta$  est un paramètre vectoriel de distribution inconnu. Rappelons que pour tout  $n \in \mathcal{I}$  la

valeur mesurée  $z_n$  (intensité du pixel dans le cas d'une image) appartient à l'ensemble  $\mathcal{Z}$  qui constitue l'espace des observations  $\mathcal{Z}^N$ . De manière analogue, le paramètre  $\theta$  appartient, en toute généralité, à l'espace des paramètres  $\Theta$  contenant les valeurs possibles  $\theta \in \Theta \subset \mathbb{R}^p$  avec  $p < N$ . L'espace des paramètres permet de définir la famille des lois envisagées  $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$ . Pour des raisons théoriques évidentes il est supposé que la famille de distribution  $\mathcal{P}$  est identifiable, *i. e.* que la condition suivante est toujours vérifiée :

$$\theta_1 \neq \theta_2 \Leftrightarrow P_{\theta_1} \neq P_{\theta_2}.$$

Autrement dit, l'application  $\theta \mapsto P_\theta$  est bijective sans quoi le paramètre n'est plus nécessairement "identifiable". Il sera admis dans cette section que les observations sont continues, *i. e.*  $\mathcal{Z} \subset \mathbb{R}$ , et que la distribution  $P_\theta$  admet la densité  $f_\theta : \mathbf{Z} \mapsto f_\theta(\mathbf{Z})$  par rapport à une mesure quelconque (celle de Lebesgue par exemple).

Dans le cas de la décision statistique, l'objectif est de choisir une "classe" à laquelle appartient le paramètre  $\theta$  à partir de la réalisation  $\mathbf{Z}$  du vecteur aléatoire  $Z \sim P_\theta$ . Il est donc nécessaire que l'espace de paramètre soit fini ou, si ce n'est pas le cas, soit partitionné en un nombre fini de sous-espaces :

$$\Theta = \bigcup_{k=0}^K \Theta_k, \quad k \neq j \Leftrightarrow \Theta_k \cap \Theta_j = \emptyset \quad (\text{A.1})$$

**Définition A.1** (hypothèse statistique). Une hypothèse statistique paramétrique désigne un ensemble  $\Theta^*$  de valeurs possibles pour le paramètre  $\theta$ . L'hypothèse est dite simple si l'ensemble  $\Theta^*$  est un singleton, *i. e.*  $\Theta^* = \{\theta^*\}$  et définit donc de façon unique la distribution du vecteur aléatoire  $Z \sim P_{\theta^*}$ , dans le cas contraire l'hypothèse est dite composite.

Étant donné la partition finie de l'espace des paramètres  $\Theta$  (A.1), notons  $\mathcal{H}_k$  l'hypothèse définie par  $\mathcal{H}_k = \{\theta \in \Theta_k\}$ . Si l'hypothèse  $\mathcal{H}_k$  est vérifiée alors les observations  $\mathbf{Z}$  sont la réalisation d'un vecteur aléatoire suivant une loi  $Z \sim P_\theta$  de densité  $f_\theta$  avec  $\theta \in \Theta_k$ .

**Définition A.2** (test statistique d'hypothèse). On appelle test statistique entre les hypothèses  $\mathcal{H}_0, \dots, \mathcal{H}_K$  toute application surjective et mesurable  $\delta : \mathcal{Z}^N \mapsto \{\mathcal{H}_0, \dots, \mathcal{H}_K\}$

Autrement dit, la définition d'un test statistique  $\delta$  revient à partitionner l'espace des observations  $\mathcal{Z}^N$  en  $K+1$  ensembles disjoints :

$$\mathcal{Z}^N = \bigcup_{k=0}^K \Omega_k, \quad k \neq j \Leftrightarrow \Omega_k \cap \Omega_j = \emptyset$$

Une région d'acceptation de l'hypothèse  $\mathcal{H}_k$  peut être une union de sous-ensembles ou l'ensemble vide (l'hypothèse  $\mathcal{H}_k$  n'est alors jamais acceptée).

tels que l'hypothèse  $\mathcal{H}_k$  est acceptée si  $\mathbf{Z} \in \Omega_k$ . L'ensemble  $\Omega_k$  est appelé région d'acceptation de l'hypothèse  $\mathcal{H}_k$ .

Par ailleurs on notera  $\Delta$  l'ensemble des tests statistiques dont le but est de décider entre les mêmes hypothèses,  $\mathcal{H}_0, \dots, \mathcal{H}_K$  dans le cas présent.

**Définition A.3** (fonction de décision). La fonction de décision associée au test  $\delta$  est une application vectorielle  $\psi = (\phi_0, \dots, \phi_K)^T : \mathcal{Z}^N \mapsto \{0;1\}^{K+1}$  telle que pour tout  $\mathbf{Z} \in \mathcal{Z}^N$  et pour tout  $k \in \{0, \dots, K\}$  :

$$\psi_k(\mathbf{Z}) = \begin{cases} 1 & \text{si } \mathbf{Z} \in \Omega_k \\ 0 & \text{sinon} \end{cases},$$

La connaissance d'un test  $\delta$  est équivalente à celle de la fonction de décision  $\psi$  associée à  $\delta$ . La définition A.3 de la fonction de décision est formellement utile pour caractériser les probabilités d'erreurs d'un test. Par convention, notons  $P_{\theta_k}$  une distribution possible lorsque l'hypothèse  $\mathcal{H}_k$  est vérifiée, *i. e.* lorsque le paramètre  $\theta_k \in \Theta_k$ ,

de façon analogue notons  $\mathbb{P}_{\theta_k}[E]$  la probabilité conditionnelle d'occurrence de l'événement  $E$  et  $\mathbb{E}_{\theta_k}$  l'opérateur d'espérance mathématique conditionné par la vérification de l'hypothèse  $\mathcal{H}_k$ .

Pour présenter la théorie de la décision statistique et quelques résultats importants qui en sont issus, commençons par aborder le problème de décision entre des hypothèses simples.

## A.2 Test entre des hypothèses simples

**Définition A.4** (erreur de  $k$ -ième espèce). *On appelle erreur de  $k$ -ième espèce du test  $\delta$  le rejet de l'hypothèse  $\mathcal{H}_k$  lorsque cette dernière est vraie :  $\mathbf{Z} \sim P_{\theta_k} \mid \theta_k \in \Theta_k, \delta(\mathbf{Z}) \neq \mathcal{H}_k$ .*

*Dans le cas d'un test entre hypothèses simples, le risque (ou probabilité d'erreur) de  $k$ -ième espèce correspond à la probabilité de rejeter l'hypothèse  $\mathcal{H}_k$  lorsque cette dernière est vérifiée :*

$$\mathbb{P}_{\theta_k}[\delta(\mathbf{Z}) \neq \mathcal{H}_k] = 1 - \mathbb{E}_{\theta_k}[\psi_k(\mathbf{Z})] \quad (\text{A.2})$$

Nous avons vu que le problème de détection statistique consiste, à choisir une hypothèse  $\mathcal{H}_k = \{\theta_k \in \Theta_k\}$  à partir de la réalisation  $\mathbf{Z}$  du vecteur aléatoire  $\mathbf{Z} \sim P_{\theta_k}$ . Dans le cas d'un test entre plusieurs hypothèses simples l'espace des paramètres  $\Theta$  est une union de singletons  $\Theta = \bigcup_{k=0}^K \{\theta_k\}$  définissant chacun de façon unique la distribution  $P_{\theta_k}$ .

Le cas particulier des tests (binaires) entre les deux hypothèses simples  $\mathcal{H}_0 = \{\theta = \theta_0\}$  et  $\mathcal{H}_1 = \{\theta = \theta_1\}$  constitue le problème de "base" de la théorie de la décision statistique. Lorsque seulement deux hypothèses sont testées des termes spécifiques sont adoptés,  $\mathcal{H}_0$  est appelée l'hypothèse de base et  $\mathcal{H}_1$  la contre-hypothèse ou l'hypothèse alternative. De même pour un test binaire  $\delta$ ,  $\alpha_0(\delta)$  est appelée probabilité de fausses-alarmes (ou de faux-positif),  $\alpha_1(\delta)$  est appelée probabilité de non-détection et enfin  $\beta(\delta) = 1 - \alpha_1(\delta) = \mathbb{E}_{\theta_1}[\psi_1(\mathbf{Z})]$  est appelée probabilité de détection ou puissance du test  $\delta$ .

La notion même d'optimalité d'un test dépend de la définition d'un critère qui doit être optimisé (minimisé ou maximisé). Certaines approches, dites mono-critère, cherchent à minimiser un unique critère de performances (test bayésien et test min-max notamment) alors que d'autres approches cherchent, dites bi-critères, à optimiser deux critères simultanément (test le plus puissant par exemple).

**Définition A.5** (Caractéristique opérationnelle du récepteur). *Dans le cas d'un test dont le but est de décider entre deux hypothèses simples  $\mathcal{H}_0$  et  $\mathcal{H}_1$ , il est souvent pratique d'illustrer les performances d'un détecteur  $\delta$  au travers de la courbe paramétrique  $\beta(\alpha_0) = \beta(\alpha_0(\delta))$ . Dans la terminologie usuelle, issue du radar, la représentation graphique de cette courbe paramétrique  $\beta(\alpha_0) : [0; 1] \mapsto [0; 1]$  est appelée caractéristique opérationnelle du récepteur ou COR.*

### a.2.1 Test le plus puissant entre deux hypothèses simples (approche bi-critère)

Le premier test étudié est basé sur une approche bi-critère uniquement définie lorsqu'il s'agit de décider entre deux hypothèses simples. Les deux critères sur lesquels repose l'approche proposée par Neyman et Pearson sont  $\alpha_0(\delta)$  et  $\beta(\delta)$ , respectivement la probabilité de fausse-alarme et la puissance. L'idée est de fixer la probabilité maximale de fausse-alarme acceptable  $\alpha_0(\delta) \leq \alpha_0$  et de maximiser la probabilité de détection  $\beta(\delta)$ . Plus formellement définissons  $\mathcal{K}_{\alpha_0} \subset \Delta$  la classe des tests dont la probabilité de fausse-alarme est majorée par  $\alpha_0$  :

$$\mathcal{K}_{\alpha_0} = \{\delta \in \Delta \mid \alpha_0(\delta) \leq \alpha_0\} \quad (\text{A.3})$$

*Dans ce mémoire on ne s'intéressera pas au problème de détection séquentielle pour lequel le nombre d'observations  $N$  n'est pas fixé à l'avance. Les observations sont utilisées séquentiellement (une à une) jusqu'à ce qu'une décision puisse être prise selon un critère de performance déterminé. D'autres critères sont alors à prendre en compte tel que la taille moyenne de l'échantillon nécessaire.*

*Dans le jargon médical les termes faux-positif et faux-négatif remplacent, respectivement, ceux de fausses-alarmes et non-détection*

**Définition A.6.** Un test  $\delta^* \in \mathcal{K}_{\alpha_0}$  est le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  (ou optimal au sens de Neyman-Pearson) si :

$$\forall \delta \in \mathcal{K}_{\alpha_0}, \beta(\delta^*) \geq \beta(\delta)$$

Le principal intérêt de cette approche réside dans la maîtrise de la probabilité de fausse-alarmer. En effet, un utilisateur peut exiger un niveau maximal de probabilité de fausse-alarmer acceptable, le test le plus puissant permet de garantir simultanément le respect de cette probabilité de fausse-alarmer et la plus grande puissance de détection. Dans la pratique, la construction d'un test le plus puissant est donné par le lemme de Neyman-Pearson :

**Lemme A.1** (de Neyman-Pearson). Soient données deux distributions  $P_{\theta_0}$  et  $P_{\theta_1}$  admettant respectivement la densité de probabilité  $f_{\theta_0}$  et  $f_{\theta_1}$ . Pour  $\alpha_0 \in ]0; 1[$ , le test le plus puissant de la classe  $\mathcal{K}_{\alpha_0}$  pour décider entre les hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$  est le test  $\delta^*$  du rapport de vraisemblance défini par :

$$\delta^*(\mathbf{Z}) = \begin{cases} \mathcal{H}_0 & \text{si } \Lambda(\mathbf{Z}) = \frac{f_{\theta_1}(\mathbf{Z})}{f_{\theta_0}(\mathbf{Z})} < \tau \\ \mathcal{H}_1 & \text{sinon} \end{cases}, \tag{A.4}$$

où  $\tau$  est un seuil fixé afin que  $\alpha_0(\delta^*) = \alpha_0$  i.e. est la solution de l'équation

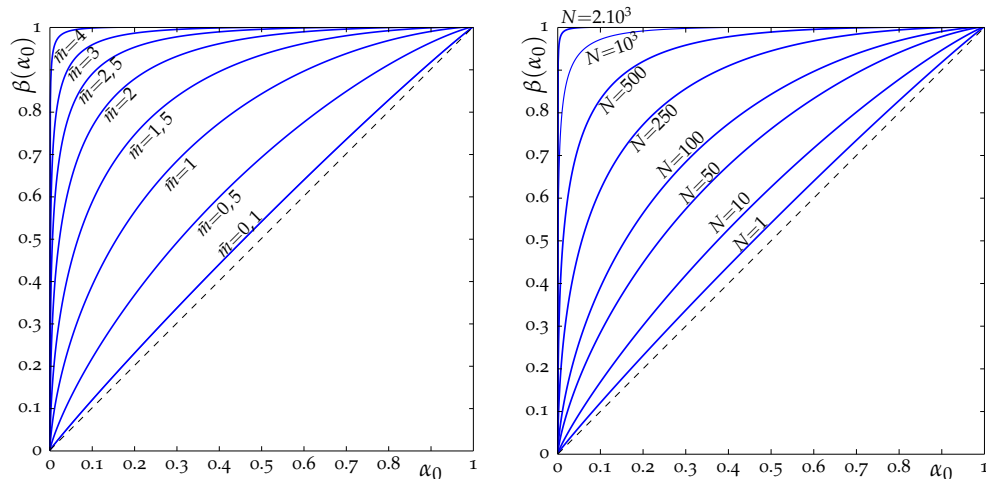
$$\mathbb{P}_{\theta_1}[\delta^*(\mathbf{Z}) \geq \tau] = \alpha_0.$$

On constatera par ailleurs que la modification de la probabilité maximale de fausse-alarmer acceptable ne change en rien la structure du test, seul le seuil  $\tau$  change.

**Remarque A.1.** Le terme de rapport de vraisemblance est dû à l'application  $\theta \mapsto P_\theta$  appelée fonction de vraisemblance.

**Remarque A.2.** Si les observations sont indépendantes il est évident que  $f_{\theta_0}(\mathbf{Z}) = \prod_{n=1}^N f_{\theta_0}(z_n)$  et de la même façon  $f_{\theta_1}(\mathbf{Z}) = \prod_{n=1}^N f_{\theta_1}(z_n)$ . Le rapport de vraisemblance qui intervient dans (A.3) s'écrit alors :

$$\Lambda(\mathbf{Z}) = \frac{f_{\theta_1}(\mathbf{Z})}{f_{\theta_0}(\mathbf{Z})} = \prod_{n=1}^N \frac{f_{\theta_1}(z_n)}{f_{\theta_0}(z_n)} = \prod_{n=1}^N \Lambda(z_n).$$



(a) Courbes COR du test le plus puissant à partir d'un seul échantillon  $\bar{m} = \{0.1; 0.5; 1; 1.5; 2; 2.5; 3; 4\}$  (b) Courbes COR du test le plus puissant à partir de  $N = \{1; 10; 50; 100; 250; 500; 1000; 2000\}$  et  $\bar{m} = 0.1$

FIGURE A.2: Courbes COR du test le plus puissant entre les deux hypothèses simples  $\mathcal{H}_0 = \{\mathbf{Z} \sim \mathcal{N}(0, 1)\}$  et  $\mathcal{H}_1 = \{\mathbf{Z} \sim \mathcal{N}(\bar{m}, 1)\}$ , en fonction de  $\bar{m}$  (figure A.2a) et en fonction de  $N$  (figure A.2b).

Pour cette raison il est dans bien des cas plus pratique d'utiliser le logarithme (Népérien) du rapport de vraisemblance qui s'écrit alors :

$$\ln(\Lambda(\mathbf{Z})) = \sum_{n=1}^N \ln(\Lambda(z_n)) = \sum_{n=1}^N \ln\left(\frac{f_{\theta_1}(z_n)}{f_{\theta_0}(z_n)}\right).$$

La fonction logarithme étant monotone (croissante) au sens strict, le seuil est simplement remplacé par  $\ln(\tau)$  ce qui ne change en rien le détecteur.

### a.2.2 Test Bayésien et minimax entre des hypothèses simples (approches mono-critère)

Dans le cadre de l'approche Bayésienne, l'occurrence des hypothèses  $\mathcal{H}_0, \dots, \mathcal{H}_K$  est elle aussi vue comme une variable aléatoire de distribution supposée connue  $Q = (q_0, \dots, q_K)^T$ , avec  $\sum_{k=0}^K q_k = 1$ . La distribution  $Q$  est un *a priori* utilisé dans un critère appelé *coût de Bayes*, somme pondérée des erreurs de  $k$ -ième espèces, définit par :

$$J_Q(\delta) = q_0\alpha_0 + \dots + q_K\alpha_K = \sum_{k=0}^K q_k\alpha_k$$

**Définition A.7** (Test Bayésien). Le test  $\delta^Q$  qui minimise le coût de Bayes  $J_Q(\delta)$  est appelé *test Bayésien associé à la distribution  $Q$* .

**Théorème A.1** (Ferguson). Le test  $\delta^Q$  qui minimise le coût de Bayes  $J_Q$  est donné [87, 88] par la règle de décision :

$$\psi_k^Q(\mathbf{Z}) = \begin{cases} 1 & \text{si } \forall j \neq k, q_k f_k(\mathbf{Z}) > q_j f_j(\mathbf{Z}) \\ 1 & \text{sinon} \end{cases}.$$

Autrement dit l'approche Bayésienne permet de prendre en compte les probabilités *a priori* d'occurrence de chacune des hypothèses. Cette approche n'est bien sûr possible que si l'on peut disposer d'informations *a priori* (la distribution  $Q$ ), sans ces dernières le calcul du test Bayésien n'est pas possible et il faut opter pour un autre critère.

Un autre critère qui peut être alors utilisé est celui dit du minimax, on s'intéresse alors à la fonction de coût définie par :

$$J(\delta) = \max_{k=0, \dots, K} \alpha_k(\delta)$$

Autrement dit, en l'absence d'*a priori* sur la probabilité d'occurrence des hypothèses, on cherche désormais à minimiser la plus grande probabilité d'erreur.

**Définition A.8** (Test minimax). Le test  $\bar{\delta}$  qui minimise le coût  $\bar{J}(\delta)$  est appelé *test minimax*.

**Théorème A.2.** Supposons qu'il existe un test Bayésien  $\bar{\delta}$  associé à une distribution *a priori*  $Q$  tel que  $\alpha_0(\bar{\delta}) = \dots = \alpha_K(\bar{\delta})$ , alors [87]  $\bar{\delta}$  est un test minimax.

Le théorème A.2 permet la construction d'un test minimax à partir d'un test Bayésien ; pour cela la distribution *a priori* associée au test Bayésien est la distribution  $Q$  appelée la *plus défavorable*.

Intuitivement, il est possible de voir l'approche minimax comme celle du "pire cas" ; en l'absence d'informations *a priori* l'objectif est alors de minimiser l'occurrence de l'erreur d'un test de la plus probable ("situation pessimiste").

### A.3 Test entre des hypothèses composites

Dans le cas où au moins une des hypothèses parmi  $\mathcal{H}_0, \dots, \mathcal{H}_K$  est composite de nombreuses difficultés apparaissent et les résultats de la section A.2 se trouvent profondément modifiés. De façon générale, l'espace des paramètres est supposé partitionné selon (A.1) et les hypothèses statistiques sont alors définies par :  $\mathcal{H}_k = \{\theta \in \Theta_k\}$ . Tout d'abord la probabilité d'erreur de  $k$ -ième espèce doit être redéfinie en conséquence :

**Définition A.9** (risque de  $k$ -ième espèce). *On appelle risque de  $k$ -ième espèce d'un test  $\delta$  entre des hypothèses composites la quantité :*

$$\alpha_k(\delta) = 1 - \inf_{\theta_k \in \Theta_k} \mathbb{E}_{\theta_k}[\psi_k(\mathbf{Z})] = \sup_{\theta_k \in \Theta_k} 1 - \mathbb{E}_{\theta_k}[\psi_k(\mathbf{Z})]$$

*Autrement dit la notion de probabilité d'erreur de  $k$ -ième espèce (voir définition A.4) est remplacée dans le cas d'un test entre hypothèses composites par la probabilité maximale de rejeter l'hypothèse  $\mathcal{H}_k$  lorsque cette dernière est vraie.*

Dans le cas d'un test entre deux hypothèses composites, on note toujours  $\mathcal{K}_{\alpha_0} = \{\delta \in \Delta \mid \alpha_0(\delta) \leq \alpha_0\}$  la classe des tests dont la probabilité de fausse-alarme est majorée par  $\alpha_0$ . Moins formellement, avec la définition précédente A.9 du risque de  $k$ -ième espèce on comprendra la classe  $\mathcal{K}_{\alpha_0}$  comme celle des tests dont la probabilité de fausse-alarme est majorée par  $\alpha_0$  quelle que soit la valeur du paramètre  $\theta_0 \in \Theta_0$ , c'est à dire :  $\sup_{\theta_0 \in \Theta_0} \mathbb{P}[\delta(\mathbf{Z}) = \mathcal{H}_1]$ .

La fonction de puissance d'un test entre deux hypothèses composites sera notée  $\beta(\theta) = \beta(\theta; \delta) = \mathbb{E}_{\theta \in \Theta_1}[\psi_1(\mathbf{Z})]$  puisque cette dernière dépend désormais du paramètre  $\theta$ .

Toujours dans le cas d'un test entre deux hypothèses composites, le pendant du test le plus puissant (ou optimal au sens de Neyman-Pearson, voir définition A.6), est le test uniformément le plus puissant (UPP).

**Définition A.10** (Test uniformément le plus puissant (UPP) entre deux hypothèses). *Un test  $\delta^*$   $\in \mathcal{K}_{\alpha_0}$  est uniformément le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  si :*

$$\forall \delta \in \mathcal{K}_{\alpha_0}, \forall \theta_1 \in \Theta_1, \beta(\theta_1; \delta^*) \geq \beta(\theta_1; \delta)$$

*ou de manière équivalente si :*

$$\forall \delta \in \mathcal{K}_{\alpha_0}, \sup_{\theta_1 \in \Theta_1} [\beta(\theta_1; \delta) - \beta(\theta_1; \delta^*)] \leq 0$$

*Autrement dit, un test est uniformément le plus puissant dans  $\mathcal{K}_{\alpha_0}$  s'il permet de garantir une probabilité de fausse-alarme majorée par  $\alpha_0$  (quel que soit  $\theta_0 \in \Theta_0$ ) et assure pour tout  $\theta_1 \in \Theta_1$  la meilleure puissance de détection. L'illustration de la fonction de puissance d'un test UPP est donnée dans la figure A.3.*

L'obtention d'un test uniformément le plus puissant est assez difficile en pratique. Il est généralement possible d'obtenir un test UPP [87, 107] que lorsque le paramètre  $\theta$  est scalaire et que les distributions associées aux deux hypothèses admettent un rapport de vraisemblance monotone relativement à une fonction  $T(x)$  c'est à dire lorsqu'il existe une fonction  $T(x)$  telle que pour tout  $\theta_0 < \theta_1$  le rapport de vraisemblance :

$$\Lambda(x) = \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)} = g(T(x))$$

*On notera qu'une condition suffisante pour qu'une densité de distribution  $p_\theta(x)$  admette un rapport de vraisemblance monotone par rapport à  $x$  est que la dérivée seconde mixte, si elle existe, vérifie  $\forall x \in \mathcal{X}, \forall \theta \in \Theta, \partial^2 p_\theta(x) / \partial \theta \partial x \geq 0$ .*

est une fonction monotone (croissante ou décroissante) au sens large de  $T(x)$ . Dans tous les cas contraires il n'est généralement pas possible de trouver un test UPP.

Une solution alternative satisfaisante, notamment dans le cadre applicatif de la détection d'informations cachées dans les images, est de définir un test qui soit asymptotiquement le plus puissant.

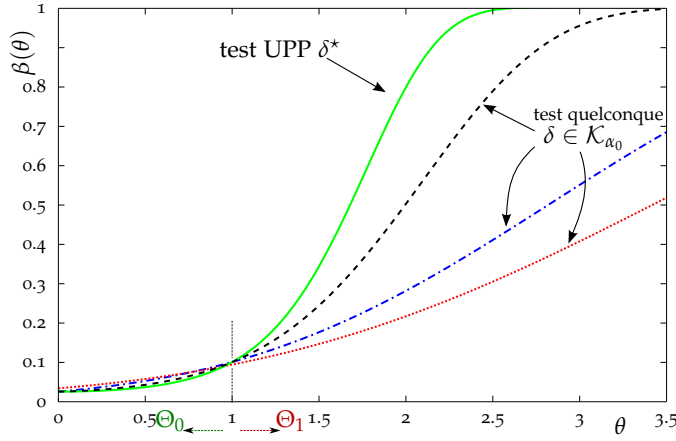


FIGURE A.3: Illustration de la fonction de puissance d'un test uniformément le plus puissant (UPP) entre les deux hypothèses composites  $\mathcal{H}_0 = \{\theta < 1\}$  et  $\mathcal{H}_1 = \{\theta > 1\}$ , avec évidemment  $\theta \in \mathbb{R}$ .

**Définition A.11** (Test asymptotiquement uniformément le plus puissant (AUPP) entre deux hypothèses [107]). Soit une séquence de variables aléatoires,  $X_1, \dots, X_n$ . Un test  $\delta_n^* \in \mathcal{K}_{\alpha_0}$  est asymptotiquement uniformément le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  si pour tout  $\delta_n \in \mathcal{K}_{\alpha_0}$  on a :

$$\limsup_{n \rightarrow \infty} \sup_{\theta_1 \in \Theta_1} [\beta(\theta_1; \delta) - \beta(\theta_1; \delta^*)] \leq 0$$

Moins formellement, le test  $\delta_n$  est AUPP dans la classe  $\mathcal{K}_{\alpha_0}$  si lorsque le nombre d'échantillons  $n$  est arbitrairement élevé le test  $\delta_n$  est le plus puissant quelque soit la valeur de  $\theta_1$ .

Néanmoins, construire un test AUPP ou démontrer qu'un test est AUPP n'est en général pas simple. Une approche

**Définition A.12** (Test localement asymptotiquement uniformément le plus puissant (LAUPP) entre deux hypothèses [107]). Soit une séquence de variables aléatoires,  $X_1, \dots, X_n$  et soit le paramètre de distribution  $\theta \in \mathbb{R}$ . Pour tester les hypothèses  $\mathcal{H}_0 = \{\theta \leq \theta^*\}$  contre  $\mathcal{H}_1 = \{\theta > \theta^*\}$ ,  $\theta \in \mathbb{R}$   $\delta_n^* \in \mathcal{K}_{\alpha_0}$  est localement asymptotiquement uniformément le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  si pour tout  $\delta_n \in \mathcal{K}_{\alpha_0}$  on a :

$$\limsup_{n \rightarrow \infty} \sup_{\theta_1 \in \Theta_1 | 0 < \theta_1 - \theta^* \leq c} [\beta(\theta_1; \delta) - \beta(\theta_1; \delta^*)] \leq 0,$$

pour tout  $c > 0$ .

L'état de l'art sur la théorie de la décision est tel qu'il n'est pas possible de définir un test qui soit optimal, en un certain sens, dans n'importe quelle situation. Néanmoins le test du rapport de vraisemblance généralisé est souvent utilisé en théorie de la décision car des résultats établissent, dans de nombreux cas, les propriétés d'optimalité de ce test [87, 108].

L'introduction de ce test sera faite au travers d'un parallèle avec le lemme A.1 de Neyman-Pearson permettant la construction du test le plus puissant dans la classe  $\mathcal{K}_{\alpha_0}$  pour un test entre deux hypothèses simples. Lorsque les hypothèses  $\mathcal{H}_0$  et  $\mathcal{H}_1$  ne sont plus connues (puisque les hypothèses sont composites) il est proposée d'estimer les paramètres  $\theta_0$  et  $\theta_1$  par maximum de vraisemblance, voir définition A.22.

**Définition A.13** (Test du rapport de vraisemblance généralisé (RVG) entre deux hypothèses composites). Soit deux hypothèses composites  $\mathcal{H}_0$  et  $\mathcal{H}_1$ , le rapport de vraisemblance généralisé entre ces deux hypothèses est défini par :

$$\Lambda^{\text{rvg}}(\mathbf{Z}) = \frac{\sup_{\theta \in \Theta_1} f_{\theta_1}(\mathbf{Z})}{\sup_{\theta \in \Theta_0} f_{\theta_0}(\mathbf{Z})}$$



On appelle test du rapport de vraisemblance généralisé le test  $\delta^{\text{rvG}}$  défini par la règle de décision suivante :

$$\psi_0^{\text{rvG}}(\mathbf{Z}) = \begin{cases} 1 & \text{si } \Lambda^{\text{rvG}}(\mathbf{Z}) < \tau, \\ 0 & \text{sinon} \end{cases},$$

où le seuil de décision  $\tau$  est fixé de façon à ce que  $\delta^{\text{rvG}} \in \mathcal{K}_{\alpha_0}$ , i. e. est la solution de l'équation  $\sup_{\theta_0 \in \Theta_0} \mathbb{P}[\delta^{\text{rvG}}(\mathbf{Z}) \geq \tau] = \alpha_0$

L'idée est alors, en quelque sorte, de choisir pour une observation  $\mathbf{Z}$  le paramètre le plus vraisemblable sous chacune des deux hypothèses. Les tests du rapport de vraisemblance généralisé consistent ensuite à tester ces deux hypothèses de façon analogue à ce que nous avons vu pour le cas d'hypothèses simple (voir le lemme A.1 de Neyman-Pearson). Il est en revanche notable que le test du RVG n'est pas nécessairement optimal en un quelconque sens, il garantit par contre le respect d'une contrainte sur la probabilité de fausse-alarme, le risque  $\alpha_0(\delta^{\text{rvG}})$  étant majoré par la constante  $\alpha_0$ . L'application de cette méthodologie peut tout à fait être étendue aux cas de test entre multiples hypothèses composites. La seule précaution à prendre est d'assurer

**Définition A.14** (Test du rapport de vraisemblance généralisé (RVG) entre plusieurs hypothèses composites). *De manière analogue, soit données  $K$  hypothèses composites  $\mathcal{H}_0, \dots, \mathcal{H}_K$ , le test du rapport de vraisemblance généralisé est défini par la règle de décision suivante :*

$$\psi_0^{\text{rvG}}(\mathbf{Z}) = \begin{cases} 1 & \text{si } \forall k \in \{0, \dots, K\}, \frac{\sup_{\theta \in \Theta_k} f_{\theta_k}(\mathbf{Z})}{\sup_{\theta \in \Theta_0} f_{\theta_0}(\mathbf{Z})} < \tau, \\ 0 & \text{sinon} \end{cases},$$

$$\psi_k^{\text{rvG}}(\mathbf{Z}) = \begin{cases} 1 & \text{si } \frac{\sup_{\theta \in \Theta_k} f_{\theta_k}(\mathbf{Z})}{\sup_{\theta \in \Theta_0} f_{\theta_0}(\mathbf{Z})} \geq \tau \text{ et } \forall j \neq k, \sup_{\theta \in \Theta_k} f_{\theta_k}(\mathbf{Z}) > \sup_{\theta \in \Theta_j} f_{\theta_j}(\mathbf{Z}), \\ 0 & \text{sinon} \end{cases},$$

## A.4 Difficultés soulevées par la présence de paramètres de nuisance

Dans le cadre des travaux présentés dans ce mémoire, le lecteur averti constatera que l'une des principales difficultés pour l'application de décision statistique à la stéganalyse est liée à la définition de la famille de distribution paramétrée. L'une des difficultés est notamment liée à la présence de paramètres de nuisance. Plus précisément, les observations  $\mathbf{Z}$  sont alors la réalisation d'une variable aléatoire  $Z$  distribuée suivant la loi  $P_{\theta, \eta}$  où, comme précédemment,  $\theta \in \Theta \subset \mathbb{R}^p$  sont des *paramètres informatifs* (ou *paramètres d'intérêts*) et  $\eta \in Y \subset \mathbb{R}^m$  sont des *paramètres "de nuisance"*. La prise en compte de ces paramètres est fondamental puisqu'ils interviennent dans la définition de la famille de distribution paramétrique  $\mathcal{P} = \{P_{\theta, \eta}; \theta \in \Theta, \eta \in Y\}$  mais ne présentent pas d'intérêt pour proposer une estimation  $\hat{\theta}$  de  $\theta$  ou pour décider en faveur d'une des hypothèses  $\mathcal{H}_k = \{\theta \in \Theta_k, \eta \in Y\}$ .

Il est par ailleurs notable que la distinction entre paramètres de nuisance et informatifs dépend de l'application envisagée, des paramètres de nuisance peuvent être jugés informatifs pour un autre problème de décision statistique et inversement. À titre d'exemple dans ce manuscrit le but est de détecter la présence d'informations cachées dans un signal. Le contenu du signal en lui-même n'est pas un paramètre informatif, puisqu'il ne renseignera pas sur la présence d'éventuelles informations cachées, mais doit être scrupuleusement pris en compte dans la définition de la distribution des échantillons. On pourrait à l'inverse imaginer utiliser le contenu de ce signal, qui interviendrait alors comme paramètre informatif, pour contrôler le bon fonctionnement de l'appareil d'enregistrement.

De manière générale, la présence de paramètre de nuisance est délicate notamment

En toute généralité, le domaine de définition du paramètre de nuisance peut varier selon l'hypothèse :  $\eta_k \in Y_k$  avec  $\exists j \neq k \mid Y_k \cap Y_j \neq \emptyset$  mais le problème couramment envisagé est que  $\forall k = \{1, \dots, K\}, Y_k = Y$ .

si leur dimension est importante : il peut exister des inférences entre paramètres de nuisance et informatifs. Pour solutionner rigoureusement un tel problème plusieurs approches peuvent être envisagées.

### Détection statistique en présence de paramètres de nuisance

Pour la détection statistique en présence de paramètres de nuisance, l'approche la plus pessimiste (ou "rigoureuse") consisterait à rechercher un test qui minimise l'erreur maximale de non-détection lorsque les paramètres informatifs et les paramètres de nuisance sont simultanément les plus défavorables. Les paramètres de nuisance dissimulent alors le mieux possible les anomalies recherchées : cette recherche du "cas le plus défavorable" conduit naturellement à envisager une approche de type minimax. Toutefois, il faut nécessairement que les hypothèses envisagées soit "suffisamment discernables" car en présence de paramètres de nuisance, si les interférences entre ces derniers et les paramètres informatifs sont trop importantes, il est fort probable que ce type d'approche ne soit plus applicable en pratique du fait des rapprochements multiples entre hypothèses.

Une autre approche séduisante consiste à utiliser la théorie de l'invariance statistique [107, chap.6],[87, chap.47]. Rappelons qu'un ensemble  $G$  muni d'une loi de composition interne possède une structure de groupe algébrique s'il vérifie les trois propriétés suivantes :

1.  $\forall (g_1, g_2, g_3) \in G^3, (g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3)$ .
2.  $\exists e \in G \mid \forall g \in G, g \cdot e = e \cdot g = g$ .
3.  $\forall g \in G, \exists g^{-1} \in G \mid g \cdot g^{-1} = g^{-1} \cdot g = e$ .

Le principe de l'invariance statistique repose sur l'utilisation du groupe algébrique (noté  $G$ ) des transformations mesurables de  $\mathcal{Z}^N$  sur lui-même muni de la loi interne qui est la composition de transformations..

**Définition A.15.** La famille de distribution  $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$  est invariante par le groupe de transformation  $G$  si :

$$\forall g \in G, \forall \theta \in \Theta, \exists \bar{\theta} \in \Theta \mid P_\theta(g(\mathbf{Z})) = P_{\bar{\theta}}(\mathbf{Z})$$

Il est notable que  $\bar{\theta}$  dépend uniquement de la transformation  $g$  et du paramètre  $\theta$ , aussi il est courant de noter  $\bar{\theta} = \bar{g}(\theta)$  où  $\bar{g} : \Theta \mapsto \Theta$  est une transformation bijective.

**Définition A.16.** Le problème consistant à décider entre les hypothèses statistiques  $\mathcal{H}_0, \dots, \mathcal{H}_K$  avec  $\mathcal{H}_k : \{\theta \in \Theta_k\}$  où  $\{\Theta_0, \dots, \Theta_K\}$  forment une partition de  $\Theta$ , voir (A.1), est invariant par le groupe de transformation  $G$  si :

- La famille de distribution paramétrique  $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$  est invariante par le groupe de transformation  $G$ .
- Les ensembles  $\Theta_k$  sont invariants par la transformation  $\bar{g}$  quelque soit  $g \in G$ , i.e. :

$$\forall \theta_k \in \Theta_k, \bar{\theta}_k = \bar{g}(\theta_k) \in \Theta_k$$

Lorsque le problème de décision statistique est invariant par le groupe de transformation  $G$ , la recherche de test préservant cette invariance semble assez naturelle.

**Définition A.17.** Le test statistique  $\delta(\mathbf{Z}) : \mathcal{Z}^N \mapsto \{\mathcal{H}_0, \dots, \mathcal{H}_K\}$  est invariant par le groupe de transformation  $G$  si :

$$\forall g \in G, \forall \mathbf{Z} \in \mathcal{Z}^N, \delta(g(\mathbf{Z})) = \delta(\mathbf{Z})$$

*Il est notable que toute transformation  $g \in G$  est bijective puisqu'il existe une transformation réciproque  $g^{-1}$*

L'utilisation du principe d'invariance statistique est particulièrement adapté à l'élimination des paramètres de nuisance linéaire. Considérons par exemple que le problème de décision statistique entre les hypothèses :

$$\begin{cases} \mathcal{H}_0 = \{ \mathbf{Z} \sim \mathcal{N}(\mathbf{H}\boldsymbol{\eta}, \sigma^2 \mathbf{I}_N); \boldsymbol{\eta} \in \mathbb{R}^p \} \\ \mathcal{H}_1 = \{ \mathbf{Z} \sim \mathcal{N}(\mathbf{H}\boldsymbol{\eta} + \boldsymbol{\theta}, \sigma^2 \mathbf{I}_N); \boldsymbol{\eta} \in \mathbb{R}^p, \boldsymbol{\theta} \neq \mathbf{0}_N \} \end{cases}, \quad (\text{A.5})$$

où  $\mathbf{H} \in \mathcal{M}_{N,p}(\mathbb{R})$  avec  $N \gg p$ ,  $\mathbf{I}_N$  est la matrice d'identité de taille  $N$  et  $\mathbf{0}_N$  est le vecteur nul de  $\mathbb{R}^N$ . À la vue du problème, il semble assez naturel d'utiliser la théorie de l'invariance statistique pour éliminer le paramètre de nuisance que constitue  $\mathbf{H}\boldsymbol{\eta}$ . En admettant que  $\mathbf{H}$  est une matrice rang plein, *i.e.*  $\text{rang}(\mathbf{H}) = p$ , la projection sur le noyau de l'application linéaire défini par  $\mathbf{H}$  permet d'éliminer algébriquement. En notant  $\mathbf{W}$  la matrice orthonormale de projection sur le noyau de  $\mathbf{H}$  vérifiant :

$$\mathbf{H}^T \mathbf{W} = \mathbf{0}_{p,N-p}, \quad \mathbf{W}^T \mathbf{W} = \mathbf{I}_{N-p} \quad \text{et} \quad \mathbf{W} \mathbf{W}^T = \mathbf{P}_H^\perp = \mathbf{I}_N - \mathbf{H} (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$$

et en notant  $\bar{\mathbf{Z}} = \mathbf{W}^T \mathbf{Z}$ , un rapide calcul montre que le problème (A.5) est équivalent au problème de décision suivant :

$$\begin{cases} \mathcal{H}_0 = \{ \bar{\mathbf{Z}} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_{N-p}) \} \\ \mathcal{H}_1 = \{ \bar{\mathbf{Z}} \sim \mathcal{N}(\mathbf{W}^T \boldsymbol{\theta}, \sigma^2 \mathbf{I}_{N-p}); \boldsymbol{\theta} \neq \mathbf{0}_N \}. \end{cases}'$$

pour lequel une solution satisfaisante est donnée par le test [161, 192] :

$$\delta(\bar{\mathbf{Z}}) = \begin{cases} \mathcal{H}_0 & \text{si } \Lambda(\bar{\mathbf{Z}}) < \tau_{\alpha_0} \\ \mathcal{H}_1 & \text{sinon} \end{cases}, \quad \text{avec } \Lambda(\mathbf{Z}) = \|\bar{\mathbf{Z}}\|_2^2 = \mathbf{Z}^T \mathbf{P}_H^\perp \mathbf{Z}.$$

Dans cet exemple, la théorie de l'invariance est utilisée en exploitant le groupe de transformation  $G : \{g : g(\mathbf{Z}) = \mathbf{Z} + \mathbf{H}\mathbf{x}, \mathbf{x} \in \mathbb{R}^p\}$  pour éliminer algébriquement le paramètre de nuisance par  $\mathbf{Z} - \mathbf{H}\boldsymbol{\eta}$ . Dans la pratique  $\boldsymbol{\eta}$  étant inconnu, cette opération s'effectue, dans le cas d'un bruit additif blanc Gaussien, par le projecteur linéaire  $\mathbf{P}_H^\perp = \mathbf{I}_N - \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T = \mathbf{W} \mathbf{W}^T$ .

Ce cas d'utilisation de la théorie de l'invariance a notamment servi pour détecter les anomalies sur lesquelles peu d'information sont *a priori* connues :  $\boldsymbol{\theta} \neq \mathbf{0}_N$ . Si en outre l'anomalie appartient à un sous-espace vectoriel de  $\mathbb{R}^N$  de dimension  $q$ , *i.e.* si  $\boldsymbol{\theta} = \mathbf{M}\boldsymbol{\mu}$ , il est intéressant de prendre en compte cette information *a priori*. En admettant que  $\mathbf{W}^T \mathbf{M}$  est une matrice de rang plein colonne le problème (A.5) devient alors :

$$\begin{cases} \mathcal{H}_0 = \{ \bar{\mathbf{Z}} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_{N-p}) \} \\ \mathcal{H}_1 = \{ \bar{\mathbf{Z}} \sim \mathcal{N}(\mathbf{W}^T \mathbf{M}\boldsymbol{\mu}, \sigma^2 \mathbf{I}_{N-p}); \boldsymbol{\mu} \neq \mathbf{0}_q \}. \end{cases}'$$

Là encore la théorie de l'invariance peut-être utilisée en exploitant, comme précédemment, le groupe de transformation  $G : \{g : g(\mathbf{Z}) = \mathbf{Z} + \mathbf{H}\mathbf{x}, \mathbf{x} \in \mathbb{R}^p\}$ . La connaissance de l'information *a priori* sur  $\boldsymbol{\theta} = \mathbf{M}\boldsymbol{\mu}$  est utilisée en projetant  $\mathbf{Z} - \mathbf{H}\boldsymbol{\eta}$  sur  $\mathbf{M}$ . Un rapide calcul montre que algébriquement la projection de  $\mathbf{Z}$  sur le sous-espace orthogonal à  $\mathbf{H}$  engendré par  $\mathbf{M}$  est donnée par (voir [161] plus de détail) :

$$\mathbf{P}_H^\perp \mathbf{M} (\mathbf{M}^T \mathbf{P}_H^\perp \mathbf{M})^{-1} \mathbf{M}^T \mathbf{P}_H^\perp \mathbf{Z}$$

Hélas, cet exemple de détection d'anomalie connue ou pas n'est pas directement transposable pour la stéganalyse. En effet, d'une part les modifications engendrées par l'insertion d'informations sont de nature stochastique; le détecteur ne peut que décrire statistiquement ces modifications. D'autre part, les modification engendrées par la stéganographie sont de puissance très faible, chacune des modification étant de l'ordre du pas de quantification pour les schémas d'insertion dans les LSB, aussi un

Si  $\mathbf{W}^T \mathbf{M}$  n'est pas une matrice de rang plein colonne cela signifie que certaines anomalies ne sont pas détectables car elles appartiennent au sous-espace vectoriel des paramètres de nuisance engendré par la matrice  $\mathbf{H}$ .

la propriété  $(\mathbf{W}^T \mathbf{M})^T \mathbf{W}^T \mathbf{M} = \mathbf{M}^T \mathbf{P}_H^\perp \mathbf{M}$  est utilisé pour obtenir ce résultat.

test ne tenant pas compte des informations *a priori* sur la nature des modifications ne pourra apporter une réponse satisfaisante en terme de puissance de détection.

Il nous faudra donc envisager une approche permettant d'éliminer les paramètres de nuisance et tenant compte des informations *a priori* sur les propriétés stochastiques des modifications engendrées par l'insertion d'informations.

## A.5 Test entre deux hypothèses et classification par apprentissage

Dans le vaste domaine du traitement et de l'analyse des images, de nombreux articles traitent de l'utilisation de la décision statistique. Lorsque l'environnement imagé est (au moins partiellement) connu ainsi que les hypothèses statistiques le test le plus puissant pourrait être utilisé. Il est cependant usuellement considéré que les hypothèses sont composites auquel cas les méthodes de détection issues de la théorie de la décision statistique (notamment le test du rapport de vraisemblance généralisé) peuvent être utilisées [193, 194]; cette méthodologie a été beaucoup utilisée pour la segmentation des images [114, 195], la détection de changement de scène dans un séquence vidéo [114, 113] ou plus couramment dans le domaine de l'imagerie médicale [196, 197].

Lorsque des informations *a priori* sur les objets sont bien établies les approches Bayésienne donnent de bons résultats. Les travaux dans ce domaine sont beaucoup plus prolifiques puisque de nombreux modèles d'images *a priori* ont été proposés. L'intérêt de l'approche réside dans la prise en compte d'informations *a priori*, mais, c'est aussi là que le bât blesse, modéliser précisément et simplement d'un point de vue statistique un support aussi complexe que l'est une image n'est pas une chose aisée. Ainsi de nombreux modèles probabilistes ont empiriquement été proposés comme *a priori* pour traduire au mieux les propriétés constatées ou attendues des images mais rarement en se fondant sur les propriétés physiques des médias. À titre d'exemple, citons l'utilisation des champs de Gibbs-Markov qui a connu un développement spectaculaire en imagerie depuis 1984 [163, 198, 199] et qui permettent d'incorporer dans une distribution *a priori* des propriétés locales essentielles que doit posséder l'objet. Il s'agit de méthodes complexes qui demandent beaucoup de savoir-faire pour construire un modèle *a priori* pertinent. Plus récemment des modèles probabilistes des coefficients issus de différentes transformation des images (notamment en utilisant les bases d'ondelettes, la DCT, la FFT, etc. . .) ont été utilisées comme *a priori* [200, 201, 202] mais leur utilisation dans le domaine de la stéganalyse n'a, pour le moment, pu être exploitée dans le cadre de tests Bayésiens. Deux principales raisons principales peuvent expliquer cette absence d'exploitation. D'une part selon les modèles probabilistes utilisés les calculs permettant la conception d'un test Bayésien peuvent s'avérer très complexes. D'autre part, la précision des modèles proposés n'est pas suffisante pour détecter une anomalie aussi infime que celle engendrée par l'insertion d'informations cachées. Enfin, pour ce type de modèles probabilistes la description statistique de la contre-hypothèse -d'une image naturelle contenant des informations cachées- n'a jamais été précisément établie à l'exception de certains algorithmes exploitant les images JPEG [52, 100, 48].

Un rapide état de l'art des méthodologies récemment publiées dans la littérature montre que les approches les plus utilisées actuellement pour la détection statistique appliquée aux images sont issues des méthodes d'apprentissage supervisées. Comme cela a déjà été brièvement évoqué dans la section 2.3.1, l'intérêt principal de ces méthodes réside dans leur caractère très général, de nombreux problèmes de classification peut-être résolu de manière satisfaisante par les méthodes d'apprentissage supervisé alors qu'il est au contraire parfois très difficile de proposer une modélisation statistique précise. L'apprentissage supervisé s'oppose à l'apprentissage non-supervisé dans

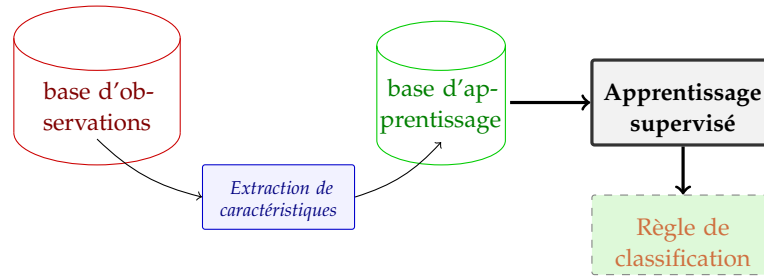


FIGURE A.4: Illustration du principe de fonctionnement des méthodes de classification.

Dans le cadre de l'apprentissage non-supervisé seul les caractéristique  $\mathbf{x}_i$  sont connus et il faut alors en plus identifier une structure de classes dans ces données.

le sens où le classifieur dispose d'une base d'apprentissage pour laquelle les classes de chacun des éléments est connu ; aussi la base d'apprentissage  $\mathcal{B}_a$  pour un problème d'apprentissage supervisé peut s'écrire sous la forme de  $N_e$  couples  $(\mathbf{x}_i; c_i) \in \mathcal{X} \times \mathcal{C}$  où  $i \in \{1, \dots, N_e\}$ ,  $\mathbf{x}_i \in \mathcal{X} \subset \mathbb{R}^Q$  représente les  $Q$  caractéristiques extraites du  $i$ -ième média et  $c_i$  représente la classe à laquelle appartient ce dernier. La sélection des caractéristiques jugées pertinentes pour la discrimination des classes est généralement le problème le plus épineux et sort du cadre des travaux abordés dans ce manuscrit (une discussion à ce sujet pourra être consultée dans [110]). Dans le cadre de la stéganalyse universelle, le problème de décision binaire posé tel que formulé dans (2.8) sera étudié dans un premier temps, l'ensemble des classes est donc représenté par  $\mathcal{C} = \{-1; 1\}$  où par exemple  $c_i=1$  si le  $i$ -ième média contient des informations cachées et inversement  $c_i = -1$  si c'est un média de couverture. Cette base étant ensuite utilisée pour apprendre une règle de décision, sa construction nécessite quelques précautions :

1. la base doit être la plus exhaustive possible.
2. elle doit, autant que possible, refléter les distributions de probabilité conditionnelles pour chacune des classes à discriminer.
3. enfin, elle doit si possible être représentative des probabilités *a priori* d'occurrence de chacune des classes.

Une discussion sur la construction de la base d'apprentissage est disponible dans [203, 204].

Une fois la base d'apprentissage  $\mathcal{B}_a$  construite, la méthode de classification  $\mathcal{C}$  reste à définir, dans ce manuscrit, nous décrirons brièvement le fonctionnement de deux d'entre elles, l'analyse discriminante linéaire et les machines à vecteurs de support.

### Fondement de l'apprentissage statistique

La classe  $\mathcal{L}$  des méthodes de classifications linéaire (ou analyse discriminante linéaire) repose sur une règle de décision qui est définie par la somme pondérée des éléments d'un vecteur de caractéristique  $\mathbf{x}_i$  :

$$\mathcal{L} = \left\{ \delta : \begin{cases} \hat{c}_i = -1 & \text{si } \Lambda(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x}_i \rangle - \lambda_0 \leq 0 \\ \hat{c}_i = 1 & \text{si } \Lambda(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x}_i \rangle - \lambda_0 > 0 \end{cases}, \forall (\mathbf{w}, \lambda_0) \in \mathbb{R}^{Q+1} \right\}, \quad (\text{A.6})$$

Il est notable que la relation  $\mathbf{w}^T \cdot \mathbf{x}_i - \lambda_0 = 0$  définit un hyperplan de dimension  $q-1$ . Un classifieur linéaire a pour but de trouver un hyperplan séparant les deux classes.

où

$$\langle \mathbf{w}, \mathbf{x}_i \rangle = \mathbf{w}^T \cdot \mathbf{x}_i = \sum_{q=1}^Q w_q x_{i,q}$$

est le produit scalaire usuel de  $\mathbb{R}^Q$ . Ce type de détecteur présente l'intérêt d'être à la fois directement lié à la notion de filtrage et de rester simple d'utilisation. Néanmoins il n'offre des solutions d'une bonne qualité que pour une classe réduite de problèmes.

Ici encore, la conception du détecteur (en établissant le vecteur de pondération  $\mathbf{w}$ ) est assujéti à un critère d'optimalité. Parmi les nombreux critères qui ont été proposés, cette présentation se limite au plus utilisé, la probabilité d'erreur  $P_e$  définie par :

$$P_e(\delta) = \sum_{c \in \mathcal{C}} q_c \int_{\mathbb{R}^Q} \mathbf{1}_{\hat{c}_i \neq c} f_c(\mathbf{x}) d\mathbf{x} \quad (\text{A.7})$$

où  $\mathbf{1}_{\hat{c}_i \neq c}$  représente la fonction indicatrice de l'événement  $\hat{c}_i \neq c$  :  $\mathbf{1}_{\hat{c}_i \neq c} = 1$  si  $\hat{c}_i \neq c$  et  $\mathbf{1}_{\hat{c}_i \neq c} = 0$  si  $\hat{c}_i = c$ ,  $q_c$  est la probabilité *a priori* d'occurrence de la classe  $c$  et  $f_c(\mathbf{x})$  est la densité de probabilité conjointe du vecteur  $\mathbf{x}$  conditionnellement à la classe  $c \in \mathcal{C} = \{0; 1\}$ . Un calcul assez simple montre que lorsque tous les éléments de l'équation (A.7) sont connus, le test de Bayes permet de minimiser le critère de probabilité d'erreur  $P_e$ . Les densités de probabilités  $f_i(\mathbf{x})$  comme les probabilité d'occurrence des classes  $q_i$  ne sont usuellement pas connues en pratique, le critère de la probabilité d'erreur empirique  $P_{emp}$  est donc à la probabilité d'erreur (théorique) :

$$P_{emp} = \frac{1}{N_e} \sum_{i=1}^{N_e} \mathbf{1}_{\hat{c}_i \neq c_i} \quad (\text{A.8})$$

Sous certaines conditions, il est établi [203] que la probabilité d'erreur empirique  $P_{emp}$  est un bon estimateur de la probabilité d'erreur  $P_e$ . Cependant, la minimisation de la probabilité d'erreur empirique  $P_{emp}$  est un problème NP-complet qu'il n'est pas envisageable de résoudre directement lorsque la dimension des vecteurs de caractéristique et/ou le nombre de ces derniers dans la base d'apprentissage devient suffisamment grand (alors qu'au contraire il a été précise que cette base se doit d'être la plus exhaustive possible). Pour contourner ce problème, de nombreuses approches ont été proposées. Dans le domaine de la stéganalyse, l'analyse de Fisher a d'abord été utilisée. Cette dernière consiste à supposer (au moins implicitement) que les vecteurs sont i.i.d. au sein de chacune des classes pour réduire les informations utilisées aux moments conditionnels de second ordre définis par :

$$m_c = \mathbb{E}_c[\Lambda(\mathbf{x})] \quad \text{et} \quad \sigma_c^2 = \text{Var}_c[\Lambda(\mathbf{x})]$$

où  $\mathbb{E}_c[\Lambda(\mathbf{x})]$  et  $\text{Var}_c[\Lambda(\mathbf{x})]$  correspondent respectivement à l'espérance mathématique et à la variance conditionnellement à  $c_i = c \in \{-1; 1\}$ .

En supposant que ces premiers moments sont connus (dans le cas contraire ils pourront être estimés à partir de la base d'apprentissage), Fisher propose d'utiliser un critère de séparation entre les deux classes défini par :

$$J_F = \frac{(m_1 - m_{-1})^2}{q_1 \sigma_1^2 + q_{-1} \sigma_{-1}^2} \quad (\text{A.9})$$

avec ici encore  $q_1$  et  $q_{-1}$  les probabilité d'occurrence respectives des classes  $c = 1$  et  $c = -1$ . Dans la classe  $\mathcal{C}$  (A.6) des tests linéaires, le test optimal au sens de Fisher est celui maximisant le critère (A.9). Un rapide calcul montre que les moment conditionnels peuvent s'écrire :

$$m_c = \mathbf{w} \cdot \bar{\mathbf{x}}_c - \lambda_0 ; \quad \sigma_c^2 = \mathbf{w}^T \Sigma_c \mathbf{w}$$

où  $\bar{\mathbf{x}}_c = \mathbb{E}_c[\mathbf{x}]$  et  $\Sigma_c = \text{Cov}_c[\mathbf{x}]$  désignent respectivement l'espérance mathématique et la covariance du vecteur de caractéristique conditionnellement à la classe  $c = \{-1; 1\}$ . La maximisation du critère de Fisher (A.9) conduit alors au résultat classique dans la littérature :

$$\mathbf{w} = \left( q_1 \Sigma_1^2 + q_{-1} \Sigma_{-1}^2 \right)^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_{-1})$$

Une fois vecteur  $\mathbf{w}$  calculé, il reste alors à choisir  $\lambda_0$  pour minimiser la probabilité d'erreur empirique  $P_{emp}$  (A.8). Dans [203], il est montré que les critères du second ordre

peuvent être pertinents et conduire à la conception d'un test minimisant l'erreur de probabilité  $P_e$ . Une procédure est proposée pour parcourir tous les critères pertinents et sélectionner celui qui minimise cette erreur.

### De la classification linéaire aux machines à vecteur de supports

Cependant, les méthodes de classifications linéaires n'offrent une solution optimale que pour une classe réduite de problèmes. La résolution de nombreux problèmes de classification peut se faire en considérant une classe plus vaste de classifieur. Dans le domaine de la stéganographie la classe des méthodes de classifications linéaires généralisées est souvent utilisée au travers des machines à vecteurs de support pouvant également être illustrées par la figure A.4. La description succincte qui en sera faite dans ce manuscrit pourra être complétée par [68, 69].

Pour établir un parallèle avec les méthodes de classifications linéaires, il est notable à partir des équations (A.8) qu'un classifieur linéaire a une probabilité d'erreur nulle si :

$$\forall i \in \{1, \dots, N_e\}, \hat{c}_i(\mathbf{w}^T \cdot \mathbf{x}_i - \lambda_0) \geq 0 \quad (\text{A.10})$$

Deux modifications fondamentales sont à la base des machines à vecteurs de supports, appelées aussi Séparateur à Vaste Marge, SVM, de l'anglais *Support Vector Machine*. Le terme de vaste marge reflète le concept de maximisation de la marge, plus précisément le nouveau problème consiste à résoudre une version légèrement modifiée de (A.10) donnée par :

$$\forall i \in \{1, \dots, N_e\}, c_i(\mathbf{w}^T \cdot \mathbf{x}_i - \lambda_0) \geq 1. \quad (\text{A.11})$$

Les classes  $c_{-1}$  et  $c_1$  sont dites linéairement séparables s'il existe  $\mathbf{w}$  et  $\lambda_0$  permettant de résoudre (A.11).

un calcul élémentaire de géométrie montre que la distance séparant les deux hyperplans  $\mathbf{w}^T \cdot \mathbf{w}_i - \lambda_0 = 1$  et  $\mathbf{w}^T \cdot \mathbf{w}_i - \lambda_0 = -1$  selon que  $c_i = \{-1; 1\}$  est donné par  $2/\|\mathbf{w}\|_2^2$ . Cette distance, la "marge", devant être maximisée, le problème consiste alors à minimiser  $\frac{1}{2}\|\mathbf{w}\|_2^2$  sous contrainte de résolution de (A.11). La minimisation d'une fonction convexe  $f(\mathbf{x})$ , i.e. dans le présent cas  $\frac{1}{2}\|\mathbf{w}\|_2^2$ , sous les contraintes  $g_i(\mathbf{x}) \leq 0, i \in \{1, \dots, N_e\}$  est équivalente à la recherche du point selle du Lagrangien :

$$L(\mathbf{x}, \boldsymbol{\alpha}) = f(\mathbf{x}) + \sum_{i=1}^{N_e} \alpha_i g_i(\mathbf{x}).$$

Le minimum est recherché par rapport à  $\mathbf{x}$  et le maximum est recherché par rapport aux  $N_e$  facteur de Lagrange  $\alpha_i \geq 0$ . Les conditions de Karush-Kuhn-Tucker sont satisfaites à l'optimum :  $\forall i \in \{1, \dots, N_e\}, \alpha_i g_i(\mathbf{x}) = 0$ .

Dans le cas qui nous intéresse il est immédiat que la fonction convexe est  $\frac{1}{2}\|\mathbf{w}\|_2^2$  et que les contraintes sont données par (A.11) ce qui conduit, en conservant les notations précédentes, à résoudre ce problème à l'aide du Lagrangien :

$$L(\mathbf{w}, \lambda_0, \boldsymbol{\alpha}) = \frac{1}{2}\|\mathbf{w}\|^2 - \sum_{i=1}^{N_e} \alpha_i (c_i(\mathbf{w}^T \cdot \mathbf{x}_i - \lambda_0) - 1). \quad (\text{A.12})$$

Un rapide calcul mène aux solutions classiques  $\boldsymbol{\alpha}^* = (\alpha_1^*, \dots, \alpha_{N_e}^*)^T$  et  $\mathbf{w}^*$  suivantes :

$$\sum_{i=1}^{N_e} \alpha_i^* \mathbf{x}_i = 0 \quad \text{et} \quad \mathbf{w}^* = \sum_{i=1}^{N_e} \alpha_i^* c_i \mathbf{x}_i \quad (\text{A.13})$$

avec, d'après les conditions de Karush-Kuhn-Tucker  $\forall i \in \{1, \dots, N_e\}, \alpha_i (c_i(\mathbf{w}^T \cdot \mathbf{w}_i - \lambda_0) - 1) = 0$ . Il en découle immédiatement que  $c_i(\mathbf{w}^T \cdot \mathbf{w}_i - \lambda_0) > 1$  si et seulement si  $\alpha_i = 0$ . Réciproquement  $\alpha_i \neq 0$  si  $\mathbf{w}^T \cdot \mathbf{w}_i - \lambda_0 = 1$ ; ces vecteur définissent seuls la solution  $\mathbf{w}^*$ .

Enfin, les résultats (A.13) peuvent être injectés dans le Lagrangien (A.12) pour finalement aboutir au problème d'optimisation dual cherchant à minimiser :

$$W(\alpha) = \sum_{i=1}^{N_e} \alpha_i - \frac{1}{2} \sum_{i=1}^{N_e} \sum_{j=1}^{N_e} \alpha_i \alpha_j c_i c_j (\mathbf{x}_i \cdot \mathbf{x}_j)$$

par rapport aux seules variables  $\alpha_i$  sous les contraintes, issues de (A.13)  $\sum_{i=1}^{N_e} \alpha_i y_i = 0$ ,  $\alpha_i > 0$ .

### Du linéaire au non-linéaire : l'astuce du noyau

Lorsque les données du problème ne sont plus linéairement séparables, *i. e.* lorsqu'il n'existe pas  $\mathbf{w}$  et  $\lambda_0$  permettant de résoudre (A.13), deux approches qui ne sont pas exclusives sont couramment utilisées. D'une part il peut être intéressant d'introduire une fonction de pénalisation dans la recherche d'optimum. D'autre part, il est courant d'utiliser une application non-linéaire des données  $\phi : \mathcal{X} \rightarrow \mathcal{H}$ ,  $\phi : \mathbf{x} \mapsto \phi(\mathbf{x}) = (\phi_1(\mathbf{x}), \dots)^T$  et d'utiliser ensuite un classifieur linéaire en  $\{\phi(\mathbf{x})\}$  est qui donc (de manière générale) non-linéaire en  $\mathbf{x}$ . Le problème d'optimisation dual linéaire en  $\{\phi(\mathbf{x})\}$  est donc désormais de minimiser

$$W(\alpha) = \sum_{i=1}^{N_e} \alpha_i - \frac{1}{2} \sum_{i=1}^{N_e} \sum_{j=1}^{N_e} \alpha_i \alpha_j c_i c_j \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$$

*D'une manière générale, l'utilisation d'une application non linéaire  $\phi : \mathbf{x} \mapsto \phi(\mathbf{x})$  n'est pas restreinte au cas non-séparable linéairement, mais plus largement a pour vocation d'améliorer les capacités (limitées) des classifieurs linéaires [203].*

par rapport aux seules variables  $\alpha_i$  sous les contraintes, issues de (A.13)  $\sum_{i=1}^{N_e} \alpha_i y_i = 0$ ,  $\alpha_i > 0$ . L'astuce du noyau (en anglais *the kernel trick*) permet de ne jamais explicitement calculer  $\phi(x)$  ce qui pourrait être coûteux calculatoirement, la dimension  $\{\phi(\mathbf{x})\}$  est en général bien plus grande que celle de  $\mathbf{x} \in \mathcal{X}$ . L'idée est donc de parvenir à construire une fonction noyau  $k : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}$  de sorte le produit scalaire  $\phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)$  soit l'image de la fonction noyau  $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$   $k(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)$  et que ce calcul soit peu coûteux calculatoirement même pour des vecteurs d'apprentissage de grande dimension. Il n'est alors aucunement nécessaire de calculer l'image des vecteurs d'apprentissage par la fonction  $\phi$ ; de même à la suite de la phase d'apprentissage, l'obtention de l'optimum  $\mathbf{w}$  permet de classifier un nouveau vecteur  $\tilde{\mathbf{x}}$  (de classe inconnue) par le calcul de  $k(\mathbf{w}, \tilde{\mathbf{x}})$ — En pratique, "l'astuce du noyau" s'appuie sur le théorème de Mercer permettant de trouver des fonctions noyaux ayant les propriétés requises [68, 205].

La théorie de l'apprentissage statistique est certainement bien établie dans la littérature et des méthodes algorithmiques permettent de réduire considérablement le coût calculatoire de leur mise en œuvre. Néanmoins, certaines limites apparaissent assez clairement :

- Les probabilités d'erreurs de décision (fausse-alarme et non-détection) ne sont connues qu'empiriquement pour les données de la base d'apprentissage  $\mathcal{B}_a$ ; ces probabilités ne peuvent donc être déterminées analytiquement durant une seconde phase de classification de données de nature inconnue. Aussi, bien qu'il soit possible de modifier le critère de probabilité d'erreur empirique  $P_{emp}$  défini dans (A.8), le respect d'une contrainte sur la probabilité de fausse-alarme est difficilement envisageable.
- Comme cela a déjà évoqué, l'une des difficultés majeure réside dans la sélection des caractéristiques utilisées. Pour certains problèmes de décision, dont la stéganalyse fait partie, les classes ne sont pas clairement définies d'un point de vue statistique, il est alors difficile de choisir quels sont les critères les plus adéquates pour faire de l'apprentissage statistique.
- Pour assurer une bonne performance des techniques de stéganalyse universelle reposant sur l'apprentissage supervisé il est en pratique souvent proposé d'utiliser



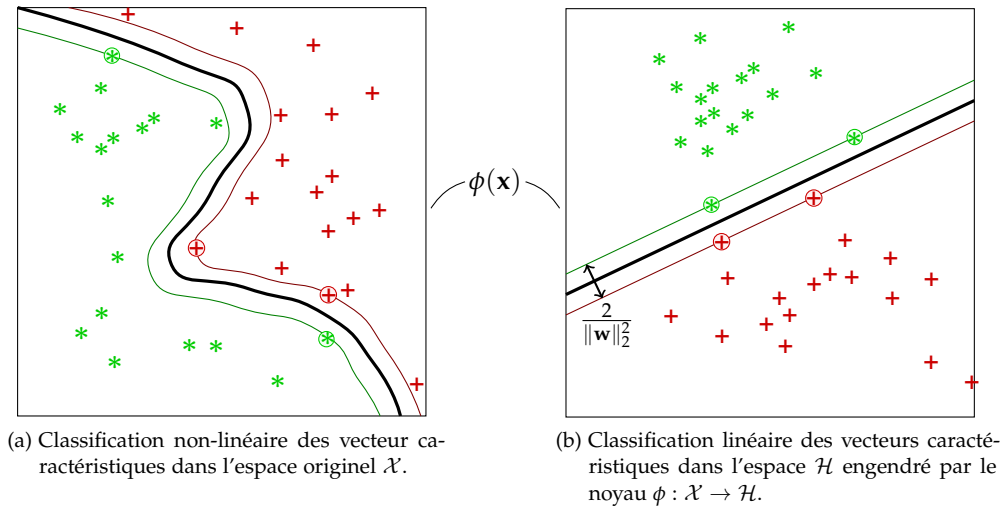


FIGURE A.5: Illustration d'un classifieur à vecteurs de support utilisant l'astuce du noyau : la classification linéaire dans  $\mathcal{H}$  est non-linéaire dans  $\mathcal{X}$ .

un très grand nombre de caractéristiques (1 234 dans [75]). Dans ce cas la base d'apprentissage doit nécessairement grossir, ce qui peut engendrer bon nombre de difficultés appelées malédiction de la dimensionalité dans le domaine [203].

- Pour des signaux aussi complexes que le sont les images, la construction pose quelques problèmes. D'une part il est difficile de refléter les distributions de probabilité conditionnelles pour chacune des classes, et *a fortiori* lorsque ni les classes, ni les caractéristiques ne sont clairement définies statistiquement. D'autre part, il n'est pas envisageable de construire une base d'apprentissage qui soit exhaustive ; en effet, physiquement le nombre de paramètres décrivant physiquement les conditions d'acquisition des images est très grand et une base exhaustive devrait contenir des exemples d'images acquises dans toutes les conditions possibles. Cela peut conduire à édifier des règles de décision spécifique à certains types d'images comme présenté expérimentalement dans [62].

À la vue des contraintes qui sont envisagées dans ce mémoire, il apparaît clairement que les approches de détection d'information cachées reposant sur les méthodes d'apprentissage supervisé ne sont pas les plus adaptées. Il semble donc nécessaire de proposer une nouvelle approche et une méthodologie associée permettant de détecter la présence d'informations cachées dans un média inconnu, sous contrainte de respecter une probabilité de fausse-alarme préalablement fixée.

## A.6 Éléments de la théorie de l'estimation statistique

De manière analogue à ce qui été présenté dans les sections A.2 - A.4 pour le cas de la détection statistique, il est considéré dans cette section que le signal enregistré  $\mathbf{Z}$  est la réalisation d'un vecteur aléatoire  $\mathbf{Z}$  suivant une loi de distribution connue  $\mathbf{Z} \sim P_\theta$ . Le paramètre vectoriel de distribution  $\theta$  (inconnu) appartient, en toute généralité, à l'espace des paramètres  $\Theta \subset \mathbb{R}^p$  et, de manière analogue,  $\mathbf{Z}$  appartient à l'espace des observations  $\mathcal{Z}^N$  avec  $N > p$ .

Dans le cas de la détection statistique, l'espace des paramètres  $\Theta$  est supposé partitionné en un nombre fini de sous-ensembles  $\Theta$ . Le but des méthodes de détection statistique consiste alors à "choisir la classe" à laquelle appartient le paramètre  $\theta$  de la loi de distribution  $P_\theta$ , à partir des observations  $\mathbf{Z}$ .

Dans le cadre de l'estimation statistique, l'objectif est de déterminer une valeur  $\hat{\theta}$  dans

l'ensemble  $\Theta$  qui soit "la plus proche possible" (en un sens qui est précisé dans la suite) de la vraie valeur du paramètre inconnu  $\theta$  à partir de la réalisation  $\mathbf{Z}$  du vecteur aléatoire  $Z \sim P_\theta$ . Le partitionnement (A.1) de l'espace des paramètres n'a alors plus de sens ; il faut au contraire que  $\Theta$  soit un compact. Aussi, les méthodes d'estimation statistique diffèrent grandement des méthodes de détection dans leurs approches comme dans les critères de performances et d'optimalité utilisés.

**Définition A.18** (Estimateur statistique). *On appelle estimateur statistique du paramètre  $\theta$  toute fonction mesurable  $\hat{\theta} : \mathcal{Z}^N \mapsto \Theta$*

Ici encore, comme cela a été précisé dans la section A.1 et implicitement admis dans l'ensemble du manuscrit, il est considéré que la famille de distribution  $\mathcal{P} = \{P_\theta ; \theta \in \Theta\}$  est identifiable :

$$\theta_1 \neq \theta_2 \Leftrightarrow P_{\theta_1} \neq P_{\theta_2}.$$

Avant de présenter quelques méthodes classiques d'estimation, il est nécessaire de présenter les critères de performances d'un estimateur les plus étudiés. **Estimation**

### non Bayésienne

**Définition A.19** (Biais d'un estimateur). *On appelle biais d'un estimateur  $\hat{\theta}$  la quantité  $b_{\hat{\theta}}(\theta)$  définie par :  $\mathbb{E}_\theta[\hat{\theta} - \theta] = \int_{\mathcal{Z}} \hat{\theta}(\mathbf{Z}) - \theta P_\theta(\mathbf{Z}) d\mathbf{Z}$ .*

*Un estimateur  $\hat{\theta}$  est dit sans biais ou non-biaisé lorsque  $\forall \theta \in \Theta, b_{\hat{\theta}}(\theta) = 0$ .*

Il est souvent souhaitable que le biais d'un estimateur soit nul, ou autrement dit, que l'espérance mathématique de l'estimateur soit égal à la valeur du paramètre à estimer quelle que soit cette dernière :  $\forall \theta \in \Theta, \mathbb{E}_\theta[\hat{\theta}] = \theta$ . Néanmoins le biais n'est pas le seul critère de performance d'un estimateur, on comprendra aisément que la dispersion autour de la moyenne de l'estimation joue notamment un rôle important.

**Définition A.20** (Erreur quadratique moyenne d'un estimateur). *On appelle erreur quadratique moyenne (EQM, en anglais MSE, "Mean Square Error") d'un estimateur  $\hat{\theta}$  la quantité  $\Sigma_{\hat{\theta}}(\theta)$  définie par :  $\Sigma_{\hat{\theta}}(\theta) = \mathbb{E}_\theta[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T]$ .*

*Il est notable que l'EQM d'un estimateur peut être formulée :  $\Sigma_{\hat{\theta}}(\theta) = \text{Var}_{\hat{\theta}}[\theta] + b_{\hat{\theta}}(\theta)b_{\hat{\theta}}(\theta)^T$  où  $\text{Var}_{\hat{\theta}}[\theta] = \mathbb{E}_\theta[(\hat{\theta} - \mathbb{E}_\theta[\hat{\theta}])(\hat{\theta} - \mathbb{E}_\theta[\hat{\theta}])^T]$  est la variance de l'estimateur  $\hat{\theta}$ .*

Dans le domaine de l'estimation, l'EQM est souvent préférée à la variance car la première tient compte de la "dispersion de l'estimateur" autour de la valeur exacte du paramètre à estimer  $\theta$ . C'est d'ailleurs en se basant sur l'EQM qu'il a été montré dans [206] qu'admettre un biais d'estimation peut, dans certains cas, permettre d'obtenir des meilleures performances en termes d'EQM d'un estimateur.

**Définition A.21** (Estimateur efficace). *Soit  $\mathcal{K}_b$  la classe des estimateur de biais majoré par la constante  $b$  :*

$$\mathcal{K}_b = \{\hat{\theta} : \mathcal{Z}^N \mapsto \Theta \mid \forall \theta \in \Theta, b_{\hat{\theta}}(\theta) = \mathbb{E}_\theta[\hat{\theta} - \theta] \leq b\}.$$

*Un estimateur  $\hat{\theta}_0$  est efficace dans la classe  $\mathcal{K}_b$  si :*

$$\forall \hat{\theta} \in \mathcal{K}_b, \forall \theta \in \Theta, \Sigma_{\hat{\theta}_0}(\theta) < \Sigma_{\hat{\theta}}(\theta)$$

Autrement dit, un estimateur  $\hat{\theta}_0$  est efficace s'il possède pour toute valeur de  $\theta$  l'erreur quadratique minimale parmi tous les estimateurs dont le biais est majoré par  $b$ . La construction d'un estimateur efficace n'est pas chose aisée en pratique. Un résultat remarquable dans ce domaine est le théorème de Gauss-Markov, sur l'estimation des paramètres d'un modèle linéaire.

**Théorème A.3** (de Gauss-Markov). *S'il existe un modèle linéaire tel que les observations  $\mathbf{Z} = \mathbf{H}\boldsymbol{\theta} + \boldsymbol{\xi}$  où  $\mathbf{H} \in \mathcal{M}_{N,p}$ ,  $p < N$  et que  $\boldsymbol{\xi}$  est la réalisation d'un vecteur aléatoire  $\Xi$  tel que  $\mathbb{E}[\Xi] = (0, \dots, 0)^T = \mathbf{0}_{N,1}$  et  $\text{Cov}[\Xi] = \sigma^2 \mathbf{I}_N$  où  $\sigma^2 \in \mathbb{R}(< \infty)$  et  $\mathbf{I}_N$  est la matrice identité de dimension  $N \times N$  alors l'estimateur  $\hat{\boldsymbol{\theta}}$  du paramètre  $\boldsymbol{\theta}$  donné par :*

$$\hat{\boldsymbol{\theta}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{Z} \quad (\text{A.14})$$

*est l'estimateur linéaire non-biaisé de moindre variance.*

*En outre, si  $\Xi \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_N)$  alors l'estimateur  $\hat{\boldsymbol{\theta}}$  du paramètre  $\boldsymbol{\theta}$  donné par l'équation (A.14) est efficace dans la classe des estimateur  $\mathcal{K}_0$ .*

Il découle donc du théorème A.3 de Gauss-Markov que si le terme stochastique  $\boldsymbol{\xi}$  n'est pas la réalisation d'un vecteur aléatoire suivant une distribution Gaussienne il est tout à fait possible qu'un estimateur (non-linéaire) non-biaisé de moindre EQM existe.

Ainsi la méthode d'estimation du maximum de vraisemblance est couramment utilisée en vertu de ses bonnes propriétés en termes d'efficacité. En effet, sous certaines conditions (voir [87]) s'il existe un estimateur efficace alors il est confondu avec celui du maximum de vraisemblance. Par ailleurs, les propriétés asymptotiques (lorsque le nombre d'échantillons observés  $N$  est arbitrairement grand) ont été largement étudiées [207], ce qui justifie un peu plus l'intérêt porté à cet estimateur.

**Définition A.22** (Estimateur du maximum de vraisemblance (MV)). *On appelle  $\hat{\boldsymbol{\theta}}$  estimateur du maximum de vraisemblance du paramètre  $\boldsymbol{\theta}$  la quantité :*

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta} \in \Theta} f_{\boldsymbol{\theta}}(\mathbf{Z})$$

*Comme précédemment évoqué pour le test du rapport de vraisemblance (A.1), si les échantillons sont indépendants on a alors  $f_{\boldsymbol{\theta}}(\mathbf{Z}) = \prod_{n=1}^N f_{\boldsymbol{\theta}}(z_n)$ .*

*Puisque le logarithme est une fonction monotone croissante  $\arg \max_{\boldsymbol{\theta} \in \Theta} f_{\boldsymbol{\theta}}(\mathbf{Z}) = \arg \max_{\boldsymbol{\theta} \in \Theta} \log(f_{\boldsymbol{\theta}}(\mathbf{Z}))$ .*

On constatera que pour fournir une estimation par maximum de vraisemblance, la densité de distribution  $f_{\boldsymbol{\theta}}$  doit être connue (ce que nous avons supposé). À l'inverse l'intérêt du théorème de Gauss-Markov est qu'il n'est nul besoin de connaître la famille de distribution paramétrique pour faire l'estimation dont l'efficacité n'est alors plus assurée que dans un cadre bien précis.

**Remarque A.3.** *En considérant le modèle linéaire précédemment évoqué dans le cadre du théorème de Gauss-Markov A.3 :  $\mathbf{Z} = \mathbf{H}\boldsymbol{\theta} + \boldsymbol{\xi}$ , avec  $\Xi \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_N)$  un rapide calcul montre que l'estimation par maximum de vraisemblance coïncide avec l'estimation par moindre carrés :*

$$\hat{\boldsymbol{\theta}} = \arg \min_{\boldsymbol{\theta} \in \Theta} \|\mathbf{Z} - \mathbf{H}\boldsymbol{\theta}\|_2^2$$

Étant admis que la construction d'un estimateur efficace n'est pas évidente, il peut être fort utile de savoir si l'estimateur étudié est "proche de l'optimalité" ou pas. Pour ce faire, il nous faut un résultat général sur la meilleure performance qu'il est possible d'attendre d'un estimateur. C'est donc par un résultat général de ce type, l'inégalité ou borne de Rao-Cramer, qu'il est proposé de clore cette section sur les méthodes d'estimations statistiques qui seront utilisées dans ce manuscrit.

Dans le cas général d'estimation d'un paramètre vectoriel, la matrice de Fisher de dimension  $p \times p$  est définie pour chacun de ses éléments par le champ scalaire  $\Theta \mapsto \mathbb{R}$  suivant :

$$\mathbf{F}_{i,j}(\boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}} \left[ \frac{\partial \ln f_{\boldsymbol{\theta}}(\mathbf{Z})}{\partial \theta_i} \frac{\partial \ln f_{\boldsymbol{\theta}}(\mathbf{Z})}{\partial \theta_j} \right] = \int_{\mathcal{Z}^N} \frac{\partial \ln f_{\boldsymbol{\theta}}(\mathbf{Z})}{\partial \theta_i} \frac{\partial \ln f_{\boldsymbol{\theta}}(\mathbf{Z})}{\partial \theta_j} f_{\boldsymbol{\theta}}(\mathbf{Z}) d\mathbf{Z} \quad (\text{A.15})$$

*On rappellera qu'une propriété est vraie presque partout si elle n'est pas vérifiée uniquement sur un domaine de mesure négligeable au sens de Lebesgue.*

**Théorème A.4** (Inégalité de Rao-Cramer). *Soit  $\mathcal{K}_b$  la classe des estimateurs d'un paramètre vectoriel  $\boldsymbol{\theta}$  de biais majoré par  $b$ . Si la densité de probabilité  $f_{\boldsymbol{\theta}}(\mathbf{Z})$  satisfait les deux conditions :*

- La fonction  $\sqrt{f_{\theta}(\mathbf{Z})}$  est dérivable “presque partout” par rapport à  $\theta_i$ ,  $\forall i \in \{1, \dots, p\}$ .
- Les champs scalaires  $F_{i,j}(\theta)$  (A.15) définissant chacun des éléments de la matrice de Fisher sont définis et continus par rapport à  $\theta$ , pour tout  $\theta \in \Theta$ .

Alors, un estimateur  $\hat{\theta} \in \mathcal{K}_b$  admet une variance minimale donnée par l'inégalité suivante :

$$\text{Cov}_{\hat{\theta}}[\theta] \geq \frac{1}{N} (\mathbf{I}_p + b'_{\hat{\theta}}(\theta)) \mathbf{F}(\theta)^{-1} (\mathbf{I}_p + b'_{\hat{\theta}}(\theta))^T \quad (\text{A.16})$$

où  $b'_{\hat{\theta}}(\theta)$  représente la matrice Jacobienne de  $b_{\hat{\theta}}(\theta)$  dont l'élément  $(i, j)$  est donné par  $\partial b_{\hat{\theta}}(\theta_i)/\theta_j$ .

**Remarque A.4.** Il découle de l'inégalité de Rao-Cramer les quelques cas particuliers suivants :

- Si la classe  $\mathcal{K}_0$  des estimateurs d'un paramètre vectoriel est considérée alors :

$$\forall \hat{\theta} \in \mathcal{K}_0, \text{Cov}_{\hat{\theta}}[\theta] > \frac{1}{N} \mathbf{F}(\theta)^{-1}.$$

- Si le paramètre  $\theta$  considéré est un scalaire alors l'inégalité de Rao-Cramer devient :

$$\forall \hat{\theta} \in \mathcal{K}_b, \text{Var}_{\hat{\theta}}[\theta] \geq \frac{(1 + b'_{\hat{\theta}}(\theta))^2}{N \mathbf{F}(\theta)}.$$

La dite borne de Rao-Cramer présente l'intérêt de permettre un calcul simple d'un minorant de la variance d'un estimateur. Par ailleurs, sous des conditions peu restrictives, la méthode d'estimation par maximum de vraisemblance atteint cette borne asymptotiquement (lorsque le nombre d'observations  $N \rightarrow \infty$  ou lorsque pour un bruit Gaussien le rapport signal sur bruit tend vers 0).

la notation  $A \geq B$  où  $A$  et  $B$  sont des matrices (de même dimension) est à comprendre  $A - B$  est semi-définie positive i.e.  $\forall \mathbf{x} \in \mathbb{R}^p, \mathbf{x}^T (A - B) \mathbf{x} \geq 0$ .



---

# Éléments de la théorie “classique” de l’approximation d’une fonction à une variable

---

# B

Le but de cette annexe est d’introduire brièvement la théorie “classique” de l’approximation d’une fonction à une variable. Les principaux résultats issus de cette théorie sont utilisés, de manière sous-jacente, dans la section ?? afin de modéliser les parties continue et discontinue d’une image en une dimension.

En analyse, la théorie de l’approximation consiste à étudier la façon dont une fonction donnée peut être approchée par des fonctions plus simples en fournissant une caractérisation quantitative des erreurs engendrées. De manière générale, les deux principaux problèmes d’intérêt de la théorie de l’approximation consistent à déterminer des bornes aussi précises que possible sur l’erreur d’approximation et à fournir une expression de la meilleure approximation, *i. e.* de la valeur approchée permettant d’atteindre cette borne. L’étude de ces problèmes dépend de la définition de l’erreur d’approximation et du type de fonction utilisée pour obtenir une approximation ; aussi la théorie de l’approximation constitue un vaste domaine de l’analyse.

Dans cette section, seule sera considérée la théorie dite “classique” de l’approximation faisant référence, dans ce manuscrit, aux travaux du début du XX<sup>ème</sup> siècle initiés notamment par S.N. Bernstein, P. Chebyshev, D. Jackson et K. Weierstrass. Ces travaux partagent les mêmes définitions de l’erreur d’approximation et utilisent des fonctions polynomiales pour approcher une fonction donnée. Dans ces travaux, il est en outre considéré, pour la définition, que les fonctions qu’il est souhaitable d’approcher sont continues sur leur domaine de définition, qui est un compact fermé  $\mathcal{X}$ , typiquement  $[0; 1]$  ou  $[-1; 1]$ , et l’erreur d’approximation considérée est formellement définie dans la définition B.1.

*Les fonctions polynomiales considérées peuvent être des polynômes algébrique ou trigonométrique.*

**Définition B.1** (Erreur uniforme d’approximation). *Soit une fonction  $f \in C^0(\mathcal{X}; \mathbb{R})$  et soit un polynôme d’approximation d’ordre  $n$   $p_n \in \mathbb{R}_n[X]$ . L’erreur notée  $\mathcal{E}(f, p)$  induite par l’approximation de la fonction  $f$  par le polynôme  $p_n$  est la quantité :*

$$\mathcal{E}(f, p_n) = \sup_{x \in \mathcal{X}} \|f(x) - p_n(x)\|. \quad (\text{B.1})$$

Une fois l’erreur d’approximation définie, il est possible de définir le polynôme  $p_n^* \in \mathbb{R}_n[X]$  de meilleure approximation pour  $n \in \mathbb{N}$  donné.

**Définition B.2** (Polynôme de meilleure approximation). *Soit une fonction  $f \in C^0(\mathcal{X}; \mathbb{R})$ . On appelle polynôme de meilleure approximation (d’ordre  $n$ ) de la fonction  $f$ , tout polynôme  $p_n^* \in \mathbb{R}_n[X]$  vérifiant :*

$$\forall p_n \in \mathbb{R}_n[X], \mathcal{E}(f, p_n) \geq \mathcal{E}(f, p_n^*).$$

Le polynôme de meilleure approximation est, en outre, associé à la meilleure erreur d’approximation.

**Définition B.3** (Meilleure erreur uniforme d’approximation). *Soit une fonction  $f \in C^0(\mathcal{X}; \mathbb{R})$ , on appelle meilleure erreur uniforme d’approximation par un polynôme de degré  $n$  la quantité :*

$$\mathcal{E}_n^*(f) = \min_{p_n^* \in \mathbb{R}_n[X]} \sup_{x \in \mathcal{X}} \|f(x) - p_n^*(x)\|$$

Ces premières définitions étant établies, quelques théorèmes permettent de caractériser le polynôme de meilleur approximation au sens de la définition B.2.

**Théorème B.1** (Théorème d'équioscillation). *Soit une fonction  $f \in C^0(\mathcal{X}; \mathbb{R})$ . Alors, le polynôme  $p_n^*$  est un polynôme de meilleure approximation de la fonction  $f$  si et seulement si :*

$$\exists \{x_k\}_{k=0}^{n+1} \in \mathcal{X}^{n+2} / \forall k = \{0, \dots, n+1\}, [f - p_n^*](x_k) = \pm (-1)^k \mathcal{E}_n^*(f)$$

Autrement dit, si et seulement si  $p_n^*$  est un polynôme de meilleure approximation de la fonction  $f$ . Alors il existe un ensemble  $\{x_k\}_{k=0}^{n+1}$  de  $n+2$  points de  $\mathcal{X}$ , pour lesquels la différence  $f - p_n^* : \mathcal{X} \rightarrow \mathbb{R}$  atteint l'erreur de meilleure approximation  $\mathcal{E}_n^*(f)$  et cela avec un signe alterné.

**Corollaire B.1** (-du théorème B.1). *Soit une fonction  $f : \mathcal{X} \rightarrow \mathbb{R}$ , et un polynôme de meilleure approximation  $p_n^* \in \mathbb{R}_n[X]$ . Il découle du théorème d'équioscillation B.1 et du théorème des valeurs intermédiaires que la différence  $[f - p_n^*](x_k)$  admet entre chaque extrema (au moins) une racine :  $\exists x_k^* \in [x_k; x_{k+1}] / \mathcal{E}_n(f, x_k^*) = 0$*

**Théorème B.2** (Théorème d'existence et d'unicité du polynôme de meilleure approximation). *Soit une fonction  $f \in C^0(\mathcal{X}; \mathbb{R})$ . Il existe un unique polynôme de meilleure approximation  $p_n^* \in \mathbb{R}_n[X]$  donné par :*

$$p_n^* = \arg \min_{\mathbb{R}_n[X]} \sup_{x \in \mathcal{X}} |f(x) - p_n^*(x)|$$

La démonstration du théorème B.2, bien que ne présentant pas d'intérêt dans la cadre de ce manuscrit, peut être faite en utilisant le théorème B.1.

Pour une fonction  $f \in C^0(\mathcal{X}; \mathbb{R})$  donnée, il est possible de caractériser plus précisément la meilleure erreur d'approximation  $\mathcal{E}_n^*(f)$ . Pour cela, il est préalablement nécessaire de caractériser la fonction  $f$  elle-même au travers, notamment, des deux définitions B.4 et B.2 suivantes.

**Définition B.4** (Module de continuité uniforme). *Soit une fonction  $f \in C^0(\mathcal{X}; \mathbb{R})$ , on appelle module de continuité uniforme [146, 147] la fonction notée  $\omega(\delta; f)$  définie sur  $\mathbb{R}^+$  par :*

$$\omega(\delta; f) = \sup_{(x,y) \in \mathcal{X}^2 / |x-y| \leq \delta} |f(x) - f(y)|. \quad (\text{B.2})$$

Le module de continuité, dont l'intérêt dans la cadre de la théorie "classique" de l'approximation est donné par la théorème B.3, vérifie parmi d'autres les propriétés B.1 suivantes.

**Propriété B.1** (-du modulé de continuité uniforme). *Pour une fonction  $f \in C^0(\mathcal{X}; \mathbb{R})$  quelconque, on peut vérifier les propriétés suivantes :*

P-1  $\omega(\delta; f)$  est une fonction uniformément continue (par rapport à la variable  $\delta$ );

P-2  $\omega(\delta; f)$  est une fonction croissante au sens large (par rapport à la variable  $\delta$ );

P-3 si en outre  $\forall f : C^0(\mathcal{X})$  alors  $\omega(0; f) = 0$ .

P-4  $\omega$  est invariant par changement d'échelle, dans le sens où si  $\tilde{f}$  est une fonction définie pour tout  $x \in \mathcal{X}$  par  $\tilde{f}(x) = f(x/s)$  alors  $\omega(\delta; f) = \omega(s\delta; \tilde{f})$ .

P-5  $\omega$  est une fonction sous-additive, i.e. pour tout  $\delta_1, \delta_2$ ,  $\omega(\delta_1 + \delta_2; f) \leq \omega(\delta_1; f) + \omega(\delta_2; f)$ . Il est immédiat d'en déduire que, de façon analogue, pour tout  $n \in \mathbb{N}$   $\omega(n\delta; f) \leq n\omega(\delta; f)$ .

Un des théorème fondamental de la théorie classique de l'approximation est l'inégalité de Jackson.

**Théorème B.3** (-de D. Jackson : majoration de la meilleure erreur d'approximation). *Soit une fonction  $f \in C^0([-1; 1]; \mathbb{R})$  alors la meilleure erreur uniforme d'approximation par un polynôme algébrique de degré  $n$ ,  $\mathcal{E}_n^*(f)$ , vérifie :*

$$\mathcal{E}_n^*(f) \leq 6\omega\left(n^{-1}; f\right) \quad (\text{B.3})$$

Le cadre d'application du théorème de Jackson est volontairement restreint au cas étudié dans ce manuscrit.

La valeur de la constante 6 a formellement été déterminée par J. Favard dans le cas où  $\mathcal{X} = [-1; 1]$ .

**Corollaire B.2** (-du théorème B.3). *Soit une fonction  $f \in C^k([-1;1];\mathbb{R})$ ; par récurrence, il est également possible de montrer que la meilleure erreur d'approximation par un polynôme de degré  $n$ ,  $\mathcal{E}^*(f)$ , vérifie :*

$$\mathcal{E}_n^*(f) \leq \frac{6^{k-1}}{n(n-1)\cdots(n-k)} \omega\left((n-k)^{-1}; f^{(k)}\right) \quad (\text{B.4})$$

Autrement dit, le théorème de Jackson B.3 permet de donner explicitement une borne supérieure de l'erreur de la meilleure approximation. Il est en outre notable que compte-tenu des propriétés B.1 du module de continuité il est aisé d'étendre le théorème de Jackson B.3 au cas plus général d'une fonction  $f$  définie sur le compact  $\mathcal{X} = [-a;a]$  par un changement d'échelle en définissant  $\tilde{f} : [-1;1] \rightarrow \mathbb{R}$  par  $\tilde{f}(x) = f(x/a)$ ; l'intérêt de la propriété (B.1) P-4 est alors immédiat pour étendre le théorème B.3 à une fonction définie sur un compact  $\mathcal{X}$  quelconque.

**Théorème B.4** (-de C.J. De la Vallée Poussin [146] : obtention d'un minorant de la meilleure erreur d'approximation). *Soit une fonction  $f : \mathcal{X} \mapsto \mathbb{R}$ , avec  $\mathcal{X} \subset [a;b]$ , et soit un polynôme  $p_n$  défini pour tout  $x \in \mathcal{X}$  par  $p_n(x) = \sum_{j=0}^n c_j \phi_j(x)$  où l'ensemble  $\{\phi_j\}_{j=0}^n$  est une famille libre de  $[a;b]$ . S'il existe dans  $\mathcal{X}$  un ensemble de  $n+2$  points  $\{x_0, \dots, x_{n+1}\}$  pour lesquels la différence  $f(x) - p_n(x)$  admet des valeurs  $\{a_0, \dots, a_{n+1}\}$  non-nulles et dont le signe change par alternance, i.e.  $\text{sign}(a_j) = \pm(-1)^j$ , alors l'erreur de la meilleure approximation de  $f$  par un polynôme  $\mathcal{E}_n^*(f)$  vérifie l'inégalité suivante :*

$$\mathcal{E}_n^*(f) = \min_{c_j} \sup_{x \in \mathcal{X}} \left| f(x) - \sum_{j=0}^n c_j \phi_j(x) \right| \geq \mathcal{E}_n^{\min}(f) = \min \left( \{ |a_0|, \dots, |a_{n+1}| \} \right). \quad (\text{B.5})$$

Le théorème B.4 de De la Vallée Poussin est étroitement lié au théorème B.1 d'équioscillation de la meilleure approximation. En effet, de manière moins formelle, le théorème B.4 permet de calculer un minorant de  $\mathcal{E}_n^*(f)$  en utilisant une approximation qui oscille autour de la fonction  $f$  à approcher. Il découle par ailleurs du théorème B.1, que la relation (B.5) est une égalité si les oscillations sont "d'égaux amplitudes"; dans tout autre cas contraire, l'inégalité de (B.5) est stricte. Par ailleurs, dans la formulation du théorème B.4, les  $n+2$  points de signes alternés sont choisis de manière quelconque; il est alors évident que le minorant obtenu est d'autant plus pertinent que l'ensemble  $\{x_0, \dots, x_{n+1}\}$  est choisi de sorte que les erreurs  $\{|a_0|, \dots, |a_{n+1}|\}$  sont importantes.

L'algorithme de Remez permet de calculer numériquement de façon simple une valeur approchée de la meilleure erreur uniforme d'approximation. Cet algorithme repose sur le théorème B.1 d'équioscillation qui fournit une condition nécessaire et suffisante pour l'obtention du polynôme de meilleure approximation. Le principe de l'algorithme de Remez est de résoudre le système d'équations linéaires suivant :

$$\begin{cases} c_0 + c_1 x_0 + c_2 x_0^2 + \dots + c_n x_0^n + E = f(x_0) \\ \vdots + \quad \vdots + \quad \vdots + \dots + \quad \vdots + \quad \vdots = \quad \vdots \\ c_0 + c_1 x_k + c_2 x_k^2 + \dots + c_n x_k^n + (-1)^k E = f(x_k) \Leftrightarrow \mathbf{Xc} = \mathbf{f} \quad (\text{B.6}) \\ \vdots + \quad \vdots + \quad \vdots + \dots + \quad \vdots + \quad \vdots = \quad \vdots \\ c_0 + c_1 x_{n+1} + c_2 x_{n+1}^2 + \dots + c_n x_{n+1}^n + (-1)^{n+1} E = f(x_{n+1}) \end{cases}$$

Ce qui permet d'obtenir une approximation polynomiale de la fonction  $f$ , notée  $p_n$ , dont les coefficients sont déterminés par le vecteur  $\mathbf{c}$ . La seconde étape de l'algorithme de Remez consiste à déterminer l'ensemble des  $n+2$  points pour lesquels la différence  $\|f(x) - p_n(x)\|$  admet un extremum local :  $\mathbf{x} = \arg \max_x |f(x) - p_n(x)|$ . L'ensemble des points  $\{x_0, \dots, x_{n+1}\}$  utilisés dans le système linéaire (B.6) est ensuite remplacé par l'ensemble  $\mathbf{x}$  de manière itérative jusqu'à atteindre un critère de convergence. L'algorithme de Remez est ici volontairement présenté de façon laconique, il ne s'agit



dans le cadre de ce manuscrit que d'une méthode permettant d'approcher la valeur de la meilleure erreur uniforme d'approximation  $\mathcal{E}_n^*(f)$ . Le lecteur intéressé pourra consulter [148].

## c.1 Critères qualitatifs de sécurité d'un stégosystème

Comme nous l'avons vu dans la section 2.1.1 et dans la figure 2.2, le principal enjeu de la stéganographie est d'insérer un message secret de façon à ce que ce dernier soit indétectable. Définir formellement les notions de détectabilité et d'indétectabilité demeurent encore un problème ouvert compte tenu de la méconnaissance de l'impact précis de la quantification et de la complexité intrinsèque des images. Dans la littérature, il existe deux approches complémentaires pour définir la sécurité d'un stégosystème. La première, assez générale, permet de définir la sécurité intrinsèque d'un stégosystème au sens statistique. La seconde approche entend définir la sécurité d'un stégosystème vis-à-vis d'un détecteur spécifique.

### c.1.1 Sécurité au sens probabiliste (face à tout adversaire)

Adoptant une approche exclusivement probabiliste, la détectabilité d'un stégosystème dépend essentiellement des modifications de la distribution de probabilité induite par l'insertion d'informations cachées. Cette approche est la base de la définition de la sécurité d'un stégosystème proposée dans [11]. En considérant qu'Alice peut uniquement envoyer à Bob, soit des médias de couverture  $C$  engendrés par une distribution de probabilité  $P_C$ , soit des stégo-médias  $S$  générés par une distribution de probabilité  $P_S$ . Le gardien, Wendy, doit décider si le média intercepté est sain  $C$  ou est un stégo-média  $S$ . Pour prendre cette décision, il est admis que Wendy possède une connaissance parfaite du stégosystème et a donc accès aux distributions  $P_C$  et  $P_S$  sur lesquelles s'appuie son analyse. D'un point de vue pratique, les distributions de probabilité  $P_C$  et  $P_S$  peuvent représenter :

- La distribution de probabilité de l'ensemble des pixels d'une image (histogramme de la répartition des valeurs des pixels) selon que l'image soit stéganographiée ou non, comme dans [89, 208].
- La distribution de probabilité de la différence entre pixels consécutifs (probabilité de co-occurrence) [72, 31].
- La distribution de probabilité conjointe d'un groupe de pixels selon que l'image soit stéganographiée ou non [209, 186].

**Définition C.1.** On appelle entropie relative ou distance de Kullback-Leibler, entre les deux distributions de probabilité  $P$  et  $Q$  la quantité définie, dans le cas continu, par :

$$D_{KL}(P, Q) = \int_{\mathbb{R}} P(x) \log \left( \frac{P(x)}{Q(x)} \right) dx$$

et de manière analogue, cette quantité est définie dans le cas discret par :

$$D_{KL}(P, Q) = \sum_{k=-\infty}^{+\infty} P[k] \log \left( \frac{P[k]}{Q[k]} \right).$$

On remarquera que cette distance correspond, dans le cas continu ou discret, à l'espérance de la log-vraisemblance  $\log(P/Q)$  en considérant que la variable aléatoire étudiée  $x$  suit la distribution  $x \sim P$ .

On parlera parfois de distance pour désigner des applications qui ne sont pas des distances au sens topologique du terme. Ainsi les distances du  $\chi^2$  ou de Kullback-Leibler ne sont pas symétriques, on a en effet en général,  $D(P, Q) \neq D(Q, P)$ . Le terme distance est ici utilisé car ces quantités quantifient la notion de séparation entre deux éléments

**Définition C.2** (sécurité d'un stégosystème au sens probabiliste). Soit  $(E, D)$  un stégosystème utilisant un objet de couverture  $\mathbf{C}$  engendré à partir d'une distribution de probabilité  $P_C$  et un stégo-objet  $\mathbf{S} = E(\mathbf{M}, \mathbf{C}, K)$  engendré à partir d'une distribution de probabilité  $P_S$ .

- Le stégosystème  $(E, D)$  est parfaitement sécurisé (face à un adversaire passif ayant connaissance des distributions  $P_C$  et  $P_S$ ) si :

$$D_{KL}(P_C, P_S) = 0.$$

- De façon analogue, le stégosystème  $(E, D)$  est  $\epsilon$ -sécurisé (face à un adversaire passif ayant connaissance des distributions  $P_C$  et  $P_S$ ) si :

$$D_{KL}(P_C, P_S) \leq \epsilon,$$

où  $D_{KL}(P_C, P_S)$  désigne la divergence de Kullback-Leibler, définie ci-dessus, entre les deux distributions  $P_C$  et  $P_S$ .

La principale limite de cette définition provient de la définition des distributions  $P_C$  et  $P_S$ , notamment en ce qui concerne les images. Il est en effet très difficile (voire impossible) d'imaginer pouvoir caractériser l'ensemble des images de couverture par une distribution de probabilité  $P_C$ . Par ailleurs, l'adversaire Wendy, cherchant à détecter un message stéganographié dans une image peut exploiter de nombreuses différentes distributions de probabilité; en ce sens cette définition s'appliquera en pratique à un schéma de détection particulier, connu, comme le montre l'exemple C.1

**Exemple C.1** (permutation de pixels). Considérons le schéma de stéganographie de correspondance de LSB par permutation de pixels. Le stéganographe insère un message dans les LSB des pixels d'une image en permutant, si nécessaire, un pixel avec un pixel de son voisinage dont le LSB correspond au bit d'information à insérer. Considérons la distribution de probabilité  $P_C$  comme la fréquence d'occurrence des valeurs des pixels dans l'ensemble de l'image. Puisque la modification générée par l'insertion du message ne modifie en rien cette distribution de probabilité on a  $P_C = P_S$  et ce stégosystème est parfaitement sécurisé, relativement à cette distribution de probabilité. En revanche, si l'on considère la distribution de probabilité conjointe de pixels consécutifs, rien ne garantit plus alors que ce stégosystème est parfaitement sécurisé.

Cette définition de la sécurité d'un stégosystème motive de nombreux travaux [210, 104] dont le but est d'accroître la sécurité d'un stégosystème au sens de la définition C.2. Ces derniers reposent sur l'utilisation de codes issus de la théorie de l'information visant, par exemple, à réduire la distance de Kullback-Leibler entre média de couverture et stéganographié ou à insérer le message sans augmenter la distance de Kullback-Leibler par rapport à un modèle de distribution des pixels [48].

### c.1.2 Sécurité au sens de la décision statistique (face à un adversaire)

Cette approche de la sécurité d'un stégosystème repose davantage sur la théorie de la décision statistique. Comme cela est détaillé dans la section 2.4 de l'annexe A, deux erreurs sont possibles dans le cadre d'un test d'hypothèses binaire, fausse alarme et non-détection ou, respectivement, erreur de première et de seconde espèces. Pour le besoin de la définition C.3, rappelons que  $\delta$  est un test statistique visant à décider entre les deux hypothèses (supposées simples) définies de façon analogue à (2.8)-(2.11) par :

$$\begin{cases} \mathcal{H}_0 : \mathbf{Z} = \mathbf{C} & \text{Le média analysé est sain} \\ \mathcal{H}_1 : \mathbf{Z} = \mathbf{S} & \text{Le média analysé est stéganographié} \end{cases}.$$

dont les probabilités d'erreurs de première et seconde espèces sont respectivement  $\alpha_0(\delta) = \mathbb{P}[\delta(\mathbf{Z}) = \mathcal{H}_1 | \mathcal{H}_0]$  et  $\alpha_1(\delta) = \mathbb{P}[\delta(\mathbf{Z}) = \mathcal{H}_0 | \mathcal{H}_1]$  et la puissance  $\beta(\delta)$  est donnée par  $1 - \alpha_1(\delta)$ .

**Définition C.3** (sécurité d'un stégosystème au sens de la décision statistique). Soit  $(E, D)$  un stégosystème produisant les stego-objets  $\mathbf{S} = E(\mathbf{M}, \mathbf{C}, K)$  à partir du média de couverture  $\mathbf{C}$ .

- Le stégosystème  $(E, D)$  est parfaitement sécurisé par rapport au détecteur  $\delta$  si :  $\forall \alpha_0 \in [0; 1], \beta(\delta) \leq \alpha_0(\delta)$
- De manière analogue, le stégosystème  $(E, D)$  est  $\epsilon$ -sécurisé par rapport au détecteur  $\delta$  si :  $\forall \alpha_0 \in [0; 1], \beta(\delta) \leq \alpha_1(\delta) + \epsilon$

De manière plus informelle, le stégosystème est parfaitement sécurisé par rapport au détecteur  $\delta$  si la probabilité de décider en faveur de l'hypothèse  $\mathcal{H}_1$  est plus importante lorsque l'hypothèse  $\mathcal{H}_0$  est vraie que lorsque l'hypothèse  $\mathcal{H}_1$  est vraie (ou inversement  $\mathbb{P}[\delta(\mathbf{Z}) = \mathcal{H}_0 | \mathcal{H}_1] > \mathbb{P}[\delta(\mathbf{Z}) = \mathcal{H}_0 | \mathcal{H}_0]$ ).

### La loi de la “racine carrée de la capacité stéganographique”

La loi de la “racine carrée de la capacité stéganographique” est un rare résultat théorique sur la sécurité d'un stégosystème. Pour comprendre et interpréter ce résultat, il a été proposé dans [211] puis dans [212, 213] de définir la notion de risque pour un stéganographe. Alice et Bob, les stéganographes, définissent le risque acceptable par un couple de probabilités  $(\alpha_0^*, \beta^*)$  avec  $0 < \alpha_0^* < \beta^* < 1$ , correspondant respectivement aux probabilités d'erreur de première espèce et à la puissance acceptable. Les stéganographes sont en situation de risque si Wendy dispose d'un test  $\delta$  dont la probabilité d'erreur de première espèce et la puissance vérifient respectivement  $\alpha_0 \leq \alpha_0^*$  et  $\beta > \beta^*$ .

**Théorème C.1** (Loi de la racine carrée de la capacité stéganographique). Soit  $N$  la taille de l'image de couverture et  $L$  la longueur du message à insérer, alors quelles que soient les probabilités  $(\alpha_0^*, \beta^*)$  définissant le risque acceptable pour le stéganographe :

- si  $\sqrt{N}/L \rightarrow \infty$  lorsque  $N \rightarrow \infty$  alors il existe une taille d'image  $N$  à partir de laquelle le stéganographe n'est pas en situation de risque ;
- si  $\sqrt{N}/L \rightarrow 0$  lorsque  $N \rightarrow \infty$  alors il existe une taille d'image  $N$  à partir de laquelle le stéganographe est en situation de risque.

De manière plus informelle, ce théorème indique que si l'on fait croître ensemble la taille de l'image et celle du message inséré, il est possible d'envoyer  $L$  bits (au moins asymptotiquement) sans risque si  $L$  ne croît pas plus rapidement que  $\sqrt{N}$ , et cela quel que soit le risque qu'il se fixe comme étant acceptable. Inversement, si le stéganographe (alors imprudent) fait croître la taille  $L$  du message plus rapidement que  $\sqrt{N}$ , la taille de l'image, alors le stéganographe sera en risque (au moins asymptotiquement).

Ce théorème a initialement été proposé dans [211] pour le cas particulier de la “stéganographie commune et/ou pour la stéganalyse par lot” (*batch steganography and pooled steganalysis* en anglais) ; de façon complémentaire, l'application du théorème C.1 dans d'autres cas a été proposée dans [33], pour le cas d'un média de couverture dont la valeur des échantillons est modélisée par un champ de Markov, et dans [214] pour le cas où la distribution du média de couverture n'est que partiellement connue. Cette notion de risque du stéganographe est liée à la définition C.3 de la sécurité d'un stégosystème. Les théorèmes 3.2 et 3.6, énoncés dans les sections 3.1 et 3.5, permettent d'affiner la loi de la racine carrée en explicitant analytiquement les probabilités d'erreurs  $(\alpha_0^*, \beta^*)$  en fonction du nombre de pixels  $N$ .

*La validité expérimentale du théorème C.1 est vérifiée dans [212].*



## Références bibliographiques

---

- [1] N. PROVOS et P. HONEYMAN, « Detecting steganographic content on the internet », in *Proc. Network and Distributed System Security Symposium*, 2002. (Cité dans les pages [xi](#), [2](#), et [21](#).)
- [2] J. BARBIER, « La stéganographie moderne : d'hérodote à nos jours », in *Proc. Débats autour de la science*, (Universités de Samara, Russie), octobre 2006. (Cité dans la page [1](#).)
- [3] S. SINGH, *The Code Book : The Secret History of Codes and Code-breaking*. Fourth Estate, (reissue) éd., 2002. (Cité dans la page [1](#).)
- [4] J. A. REEDS, « Solved : The Ciphers in Book III of Trithemius's Steganographia », *Cryptologia*, vol. 22, p. 291–319, Oct 1996. (Cité dans la page [1](#).)
- [5] T. ERNST, « the numerical-astrological ciphers in the third book of trithemius's steganographia », *Cryptologia*, vol. 22, p. 318–341, Apr 1996. (Cité dans la page [1](#).)
- [6] J. BARBIER, *Analyse de canaux de communication dans un contexte non coopératif : Application aux codes correcteurs d'erreurs et à la stéganalyse*. Thèse doctorat, École Supérieure et d'Application des Transmissions, novembre 2007. (Cité dans les pages [1](#) et [12](#).)
- [7] R. ANDERSON, « Stretching the limits of steganography », *IEEE Journal of Selected Areas in Communications*, vol. 16, p. 474–481, 1996. (Cité dans la page [1](#).)
- [8] M. CHAPMAN, *Hiding the Hidden : A Software System for Concealing Ciphertext in Innocuous Text*. Thèse doctorat, University of Wisconsin-Milwaukee, May 1997. (Cité dans la page [2](#).)
- [9] G. SIMMONS, « The prisoners problem and the subliminal channel », *CRYPTO*, p. 51–67, 1983. (Cité dans la page [10](#).)
- [10] T. MORTEL, J. ELOFF et M. OLIVIER, « An overview of image steganography », in *Proceedings of the Fifth Annual Information Security South Africa Conference (ISSA'2005)*, 2005. (Cité dans la page [10](#).)
- [11] C. CACHIN, « An information-theoretic model for steganography », *Information and Computation*, vol. 192, no. 1, p. 41–56, 2004. (Cité dans les pages [11](#) et [193](#).)
- [12] A. KERCKHOFFS, « La cryptographie militaire », *Journal des sciences militaires*, vol. IX, p. 5–83, jan. 1883. (Cité dans la page [12](#).)
- [13] R. ANDERSON et F. PETITCOLAS, « On the limits of steganography », *Selected Areas in Communications, IEEE Journal on*, vol. 16, p. 474 –481, may 1998. (Cité dans la page [12](#).)
- [14] P. GUILLON, T. FURON et P. DUHAMEL, « Applied public-key steganography », in *Security, Steganography, and Watermarking of Multimedia Contents IV*, SPIE Electronic Imaging, 2002. (Cité dans la page [12](#).)
- [15] J. EGGERS et B. GIROD, « Blind watermarking applied to image authentication », in *Proc. of Acoustics, Speech, and Signal Processing (ICASSP '01)*, vol. 3, p. 1977 –1980 vol.3, 2001. (Cité dans la page [12](#).)
- [16] I. COX, M. L. MILLER et J. A. BLOOM, *Digital watermarking*. San Francisco, CA, USA : Morgan Kaufmann Publishers Inc., 2002. (Cité dans les pages [12](#) et [13](#).)
- [17] A. BARG, G. BLAKLEY et G. KABATIANSKY, « Digital fingerprinting codes : problem statements, constructions, identification of traitors », *Information Theory, IEEE Transactions on*, vol. 49, p. 852 – 865, avril 2003. (Cité dans la page [12](#).)
- [18] R. BÖHME, *Advanced Statistical Steganalysis*. Springer Publishing Company, Incorporated, 1st éd., 2010. (Cité dans les pages [12](#), [15](#), [16](#), [19](#), [21](#), et [33](#).)

- [19] I. COX, M. MILLER, J. BLOOM, J. FRIDRICH et T. KALKER, *Digital Watermarking and Steganography*. Morgan Kaufmann, 2nd edition éd., 2007. (Cité dans les pages 13, 19, 20, 21, et 33.)
- [20] E. LAM et J. GOODMAN, « A mathematical analysis of the DCT coefficient distributions for images », *Image Processing, IEEE Transactions on*, vol. 9, p. 1661–1666, oct 2000. (Cité dans les pages 13, 33, et 166.)
- [21] C. ZITZMANN, R. COGRANNE, F. RETRAINT, L. FILLATRE, I. NIKIFOROV et P. CORNU, « Détection optimale à base de laplacienne quantifiée », in *XXIII<sup>ème</sup> colloque GRETSI*, p. xx – xx, 2011. (Cité dans la page 13.)
- [22] J. FRIDRICH, *Steganography in Digital Media : Principles, Algorithms, and Applications*. Cambridge University Press, 1st edition éd., 2009. (Cité dans les pages 15, 16, 19, 21, et 33.)
- [23] R. COGRANNE, C. ZITZMANN, L. FILLATRE, F. RETRAINT, I. NIKIFOROV et P. CORNU, « Statistical decision by using quantized observations », in *IEEE International Symposium on Information Theory*, p. 1135 – 1139, 2011. (Cité dans les pages 15, 32, 58, 125, et 166.)
- [24] A. KER, « Improved detection of lsb steganography in grayscale images », in *Information Hiding*, vol. 3200, p. 97–115, 2004. (Cité dans la page 15.)
- [25] J. FRIDRICH, M. GOLJAN et D. SOUKAL, « Wet paper codes with improved embedding efficiency », *Information Forensics and Security, IEEE Transactions on*, vol. 1, p. 102 – 110, march 2006. (Cité dans la page 16.)
- [26] W. ZHANG, X. ZHANG et S. WANG, « A double layered “plus-minus one” data embedding scheme », *Signal Processing Letters, IEEE*, vol. 14, p. 848–851, nov. 2007. (Cité dans la page 16.)
- [27] J. BIERBRAUER et J. FRIDRICH, « Constructing good covering codes for applications in steganography », in *Transactions on data hiding and multimedia security III*, p. 1–22, Berlin, Heidelberg : Springer-Verlag, 2008. (Cité dans la page 16.)
- [28] M. KHATIRINEJAD et P. LISONEK, « Linear codes for high payload steganography », *Discrete Applied Mathematics*, vol. 157, no. 5, p. 971 – 981, 2009. (Cité dans la page 16.)
- [29] D.-C. WU et W.-H. TSAI, « A steganographic method for images by pixel-value differencing », *Pattern Recognition Letters*, vol. 24, no. 9-10, p. 1613 – 1626, 2003. (Cité dans la page 16.)
- [30] C.-H. YANG, C.-Y. WENG, S.-J. WANG et H.-M. SUN, « Adaptive data hiding in edge areas of images with spatial lsb domain systems », *Information Forensics and Security, IEEE Transactions on*, vol. 3, p. 488–497, sept. 2008. (Cité dans la page 16.)
- [31] T. PEVNY, P. BAS et J. FRIDRICH, « Steganalysis by subtractive pixel adjacency matrix », *IEEE Trans. Inform. Forensics and Security*, vol. 5, no. 2, p. 215–224, 2010. (Cité dans les pages 16, 23, et 193.)
- [32] W. LUO, F. HUANG et J. HUANG, « Edge adaptive image steganography based on lsb matching revisited », *Information Forensics and Security, IEEE Transactions on*, vol. 5, p. 201–214, june 2010. (Cité dans la page 16.)
- [33] T. FILLER, A. KER et J. FRIDRICH, « The square root law of steganographic capacity for markov covers », in *Media Forensics and Security, Proc. SPIE*, vol. 7254, (San Jose, CA, USA), 2009. (Cité dans les pages 16 et 195.)
- [34] T. SHARP, « An implementation of key-based digital signal steganography », in *Information Hiding*, vol. 2137 in *Lecture Notes in Computer Science*, p. 13–26, Springer Berlin / Heidelberg, 2001. (Cité dans la page 16.)
- [35] W. BENDER, D. GRUHL, N. MORIMOTO et A. LU, « Techniques for data hiding », *IBM Systems Journal*, vol. 35, no. 3.4, p. 313–336, 1996. (Cité dans la page 17.)

- [36] I. COX, J. KILIAN, F. LEIGHTON et T. SHAMOON, « Secure spread spectrum watermarking for multimedia », *Image Processing, IEEE Transactions on*, vol. 6, p. 1673 –1687, dec. 1997. (Cité dans la page 17.)
- [37] L. MARVEL, J. BONCELET, C.G. et C. RETTER, « Spread spectrum image steganography », *Image Processing, IEEE Transactions on*, vol. 8, p. 1075 –1083, aug 1999. (Cité dans les pages 17 et 23.)
- [38] H. MALVAR et D. FLORENCIO, « Improved spread spectrum : a new modulation technique for robust watermarking », *Signal Processing, IEEE Transactions on*, vol. 51, p. 898 – 905, apr 2003. (Cité dans les pages 17, 18, et 23.)
- [39] B. CHEN et G. WORNELL, « Digital watermarking and information embedding using dither modulation », in *Multimedia Signal Processing, IEEE 2nd workshop on*, p. 273 – 278, 1998. (Cité dans la page 18.)
- [40] B. CHEN et G. WORNELL, « Quantization index modulation : a class of provably good methods for digital watermarking and information embedding », *Information Theory, IEEE Transactions on*, vol. 47, p. 1423 – 1443, May 2001. (Cité dans la page 18.)
- [41] P. MOULIN et R. KOETTER, « Data-hiding codes », *Proceedings of the IEEE*, vol. 93, p. 2083 –2126, dec. 2005. (Cité dans la page 18.)
- [42] P. BAS et J. HURRI, « Security of dm quantization watermarking schemes : A practical study for digital images », in *Digital Watermarking*, vol. 3710in *Lecture Notes in Computer Science*, p. 186–200, Springer Berlin, 2005. (Cité dans la page 18.)
- [43] F. CAYRE et P. BAS, « Kerckhoffs-based embedding security classes for woa data hiding », *Information Forensics and Security, IEEE Transactions on*, vol. 3, p. 1 –15, march 2008. (Cité dans la page 18.)
- [44] J. K. JOSEPH, Ö. RUANAIDH et T. PUN, « Rotation, scale and translation invariant spread spectrum digital image watermarking », *Signal Processing*, vol. 66, no. 3, p. 303 – 317, 1998. (Cité dans la page 19.)
- [45] C.-Y. LIN, M. WU, J. BLOOM, I. COX, M. MILLER et Y. LUI, « Rotation, scale, and translation resilient watermarking for images », *Image Processing, IEEE Transactions on*, vol. 10, p. 767 –782, may 2001. (Cité dans la page 19.)
- [46] P. SALLEE, « Model-based steganography », in *Digital Watermarking*, vol. 2939in *Lecture Notes in Computer Science*, p. 254–260, Springer Berlin / Heidelberg, 2004. (Cité dans les pages 19 et 33.)
- [47] M. TOLBA, M.-S. GHONEMY, I.-H. TAHA et A. KHALIFA, « High capacity image steganography using wavelet-based fusion », in *Computers and Communications, 2004. Proceedings. ISCC 2004. Ninth International Symposium on*, vol. 1, p. 430 – 435 Vol.1, june-1 july 2004. (Cité dans la page 19.)
- [48] P. SALLEE, « Model-based methods for steganography and steganalysis », *International Journal of Image and Graphics*, vol. 5, p. 167–189, jan. 2005. (Cité dans les pages 19, 33, 179, et 194.)
- [49] H. FARID, « Detecting hidden messages using higher-order statistical models », in *IEEE International Conference on Image Processing*, vol. 2, p. II–905 – II–908, 2002. (Cité dans les pages 19 et 22.)
- [50] Y. SHI, G. XUAN, D. ZOU, J. GAO, C. YANG, Z. ZHANG, P. CHAI, W. CHEN et C. CHEN, « Image steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image, and neural network », in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, p. 4 pp., july 2005. (Cité dans les pages 19 et 23.)
- [51] Y. WANG et P. MOULIN, « Statistical modelling and steganalysis of DFT-based image steganography », in *Security, Steganography, and Watermarking of Multimedia Contents VIII, Proc. SPIE 6072*, SPIE, 2006. (Cité dans la page 19.)



- [52] A. WESTFELD, « F5—a steganographic algorithm », in *Information Hiding*, vol. 2137 in *Lecture Notes in Computer Science*, p. 289–302, Springer Berlin / Heidelberg, 2001. (Cité dans les pages 19, 33, et 179.)
- [53] N. JOHNSON et S. JAJODIA, « Steganalysis of images created using current steganography software », in *Proceedings of the 2nd Information Hiding Workshop*, 1998. (Cité dans les pages 19, 22, et 24.)
- [54] N. JOHNSON et S. JAJODIA, « Exploring steganography : Seeing the unseen », *Computer*, vol. 31, p. 26–34, feb. 1998. (Cité dans la page 19.)
- [55] T. PEVNÝ, J. FRIDRICH et A. D. KER, « From blind to quantitative steganalysis », in *Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents XI, San Jose, CA, January 19–January 22*, January 2009. (Cité dans les pages 20 et 23.)
- [56] Q. GUAN, J. DONG et T. TAN, « Blind quantitative steganalysis based on feature fusion and gradient boosting », in *Proceedings of the 9th international conference on Digital watermarking, IWDW'10, (Berlin, Heidelberg)*, p. 266–279, Springer-Verlag, 2011. (Cité dans la page 20.)
- [57] Y. MICHE, P. BAS et A. LENDASSE, « Using multiple re-embeddings for quantitative steganalysis and image reliability estimation », Rap. tech. TKK-ICS-R34, Aalto University School of Science and Technology, Aalto, Finland, June 2010. (Cité dans les pages 20 et 23.)
- [58] J. FRIDRICH, M. GOLJAN, D. SOUKAL et H. TARAS, « Forensic steganalysis : Determining the stego key in spatial domain steganography », in *In Security, Steganography, and Watermarking of Multimedia Contents VII, volume 5681 of Proc. SPIE*, p. 631–642, 2005. (Cité dans la page 20.)
- [59] S. TRIVEDI, R. CHANDRAMOULI et R. CH, « Secret key estimation in sequential steganography », *Signal Processing, IEEE Transactions on*, vol. 53, p. 746–757, 2005. (Cité dans la page 20.)
- [60] X. LUO, D. WANG, P. WANG et F. LIU, « Secret key estimation for image sequential steganography in transform domain », in *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, p. 538–542, nov. 2007. (Cité dans la page 20.)
- [61] A. D. KER, « Locating steganographic payload via WS residuals », in *ACM Proceedings of 10th Multimedia and Security Workshop*, p. 27–31, 2008. (Cité dans les pages 20 et 31.)
- [62] G. CANCELLI, G. DOERR, M. BARNI et I. COX, « A comparative study of  $\pm 1$  steganalyzers », in *Multimedia Signal Processing, 2008 IEEE 10th Workshop on*, p. 791–796, oct. 2008. (Cité dans les pages 22 et 184.)
- [63] R. A. FISHER, « The use of multiple measurements in taxonomic problems », *Annals Eugen.*, vol. 7, p. 179–188, 1936. (Cité dans la page 22.)
- [64] J. J. HARMSSEN et W. A. PEARLMAN, « Kernel fisher discriminant for steganalysis of jpeg hiding methods », in *In Security, Steganography, and Watermarking of Multimedia Contents VI, volume 5306 of Proc. SPIE*, p. 13–22, 2004. (Cité dans la page 23.)
- [65] J. FRIDRICH, « Feature-based steganalysis for jpeg images and its implications for future design of steganographic schemes », in *Information Hiding*, vol. 3200 in *Lecture Notes in Computer Science*, p. 67–81, Springer Berlin / Heidelberg, 2005. (Cité dans la page 23.)
- [66] L. SHAOHUI, Y. HONGXUN et G. WEN, « Neural network based steganalysis in still images », in *Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference on*, vol. 2, p. II – 509–12 vol.2, july 2003. (Cité dans la page 23.)
- [67] W.-N. LIE et G.-S. LIN, « A feature-based classification technique for blind image steganalysis », *Multimedia, IEEE Transactions on*, vol. 7, p. 1007–1020, dec. 2005. (Cité dans la page 23.)
- [68] V. N. VAPNIK, *The nature of statistical learning theory*. New York, NY, USA : Springer-Verlag New York, Inc., 1995. (Cité dans les pages 23, 182, et 183.)

- [69] V. N. VAPNIK, *Statistical Learning Theory*. Wiley-Interscience, sept. 1998. (Cité dans les pages 23 et 182.)
- [70] S. LYU et H. FARID, « Steganalysis using higher-order image statistics », *Information Forensics and Security, IEEE Transactions on*, vol. 1, p. 111 – 119, march 2006. (Cité dans la page 23.)
- [71] D. ZOU, Y. SHI, W. SU et G. XUAN, « Steganalysis based on markov model of thresholded prediction-error image », in *Multimedia and Expo, 2006 IEEE International Conference on*, p. 1365 –1368, july 2006. (Cité dans la page 23.)
- [72] K. SULLIVAN, U. MADHOW, S. CH et B. S. MANJUNATH, « Steganalysis for markov cover data with applications to images », *IEEE Trans. Inform. Forensics and Security*, vol. 1, no. 2, p. 275–287, 2006. (Cité dans les pages 23 et 193.)
- [73] H. GOU, A. SWAMINATHAN et M. WU, « Noise features for image tampering detection and steganalysis », in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, vol. 6, p. VI –97 –VI –100, oct 2007. (Cité dans la page 23.)
- [74] Q. LIU, A. SUNG, B. RIBEIRO et R. FERREIRA, « Steganalysis of multi-class jpeg images based on expanded markov features and polynomial fitting », in *IEEE International Joint Conference on Neural Networks*, p. 3352 –3357, june 2008. (Cité dans la page 23.)
- [75] J. KODOVSKÝ, T. PEVNÝ et J. FRIDRICH, « Modern steganalysis can detect YASS », in *Proceedings SPIE, Electronic Imaging, Media Forensics and Security XII*, (San Jose, CA), p. 2 1–2 11, January 17–21, 2010. (Cité dans les pages 23 et 184.)
- [76] H. DRUCKER, C. J. C. BURGESS, L. KAUFMAN, A. J. SMOLA et V. VAPNIK, « Support vector regression machines. », in *NIPS'96*, p. 155–161, 1996. (Cité dans la page 23.)
- [77] A. HYVÄRINEN et E. OJA, « Independent component analysis : algorithms and applications », *Neural Networks*, vol. 13, no. 4-5, p. 411 – 430, 2000. (Cité dans la page 23.)
- [78] R. CHANDRAMOULI, « A mathematical framework for active steganalysis », *Multimedia Systems*, vol. 9, p. 303–311, 2003. 10.1007/s00530-003-0101-8. (Cité dans la page 23.)
- [79] X.-Y. LUO, D.-S. WANG, P. WANG et F.-L. LIU, « A review on blind detection for image steganography », *Signal Processing*, vol. 88, no. 9, p. 2138 – 2157, 2008. (Cité dans la page 24.)
- [80] A. KER, « A general framework for structural steganalysis of lsb replacement », in *Information Hiding*, vol. 3727 in *Lecture Notes in Computer Science*, p. 296–311, Springer Berlin / Heidelberg, 2005. (Cité dans les pages 24 et 26.)
- [81] J. FRIDRICH, M. GOLJAN et R. DU, « Reliable detection of LSB steganography in color and grayscale images », *IEEE Multimedia*, vol. 8, p. 22–28, 2001. (Cité dans les pages 24, 151, et 152.)
- [82] S. DUMITRESCU, X. WU et W. ZHE, « Detection of LSB steganography via sample pair analysis », *Signal Processing, IEEE Transactions on*, vol. 51, no. 6, p. 1995–2007, 2003. (Cité dans les pages 24, 25, et 151.)
- [83] P. LU, X. LUO, Q. TANG et L. SHEN, « An improved sample pairs method for detection of lsb embedding », in *Information Hiding*, vol. 3200 in *Lecture Notes in Computer Science*, p. 215–216, Springer Berlin / Heidelberg, 2005. (Cité dans les pages 26 et 152.)
- [84] A. D. KER, « Fourth-order structural steganalysis and analysis of cover assumptions », in *Security, Forensics, Steganography, and Watermarking of Multimedia Contents VIII, Proc. SPIE 6072*, p. 25–38, 2006. (Cité dans la page 26.)
- [85] A. WESTFELD et P. ANDREAS, « Attacks on steganographic systems », in *Information Hiding, 3rd International Workshop*, p. 61–76, 1999. (Cité dans les pages 26, 27, 151, 152, et 159.)
- [86] R. CHANDRAMOULI et N. MEMON, « Analysis of lsb based image steganography techniques », in *Image Processing, 2001. Proceedings. 2001 International Conference on*, vol. 3, p. 1019 –1022 vol.3, 2001. (Cité dans les pages 28 et 29.)

- [87] A. A. BOROVKOV, *Mathematical Statistics*. Gordon and Breach Sciences Publishers, Amsterdam, 1998. (Cité dans les pages 29, 37, 56, 130, 173, 174, 175, 177, et 186.)
- [88] T. FERGUSON, *Mathematical Statistics : A Decision Theoretic Approach*. Academic Press, 1967. (Cité dans les pages 29 et 173.)
- [89] O. DABEER, K. SULLIVAN, U. MADHOW, S. CHANDRASEKARAN et B. MANJUNATH, « Detection of hiding in the least significant bit », *Signal Processing, IEEE Transactions on*, vol. 52, p. 3046 – 3058, oct. 2004. (Cité dans les pages 29, 30, 37, 46, 47, 65, 151, 152, et 193.)
- [90] W. HOEFFDING, « Asymptotically optimal tests for multinomial distributions », *Annals Math. Statist*, vol. 36, p. 369–400, 1965. (Cité dans la page 30.)
- [91] J. FRIDRICH et M. GOLJAN, « On estimation of secret message length in LSB steganography in spatial domain », in *Security, Steganography, and Watermarking of Multimedia Contents VI*, vol. 5306 in *Proc. SPIE*, p. 23–34, 2004. (Cité dans les pages 31, 32, 65, 67, 151, et 152.)
- [92] A. D. KER et R. BËHME, « Revisiting weighted stego-image steganalysis », in *Security, Forensics, Steganography, and Watermarking of Multimedia Contents X, Proc. SPIE 6819*, p. 501–517, 2008. (Cité dans les pages 31, 32, 67, 151, 152, et 153.)
- [93] A. D. KER, « Optimally weighted least-squares steganalysis », in *In Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505 of *Proc. SPIE*, 2007. (Cité dans les pages 31, 32, et 151.)
- [94] A. KER, « A weighted stego image detector for sequential lsb replacement », in *Information Assurance and Security, 2007. IAS 2007. Third International Symposium on*, p. 453 –456, aug. 2007. (Cité dans la page 31.)
- [95] R. BËHME, « Assessment of steganalytic methods using multiple regression models », in *Information Hiding*, vol. 3727 in *Lecture Notes in Computer Science*, p. 278–295, Springer Berlin / Heidelberg, 2005. (Cité dans les pages 32 et 153.)
- [96] C. ZITZMANN, R. COGRANNE, F. RETRAINT, I. NIKIFOROV, L. FILLATRE et P. CORNU, « Hypothesis testing by using quantized observations », in *Statistical Signal Processing, Proc. of IEEE Workshop on*, p. 501 – 504, 2011. (Cité dans les pages 32 et 125.)
- [97] J. J. FRIDRICH, D. SOUKAL et M. GOLJAN, « Maximum likelihood estimation of length of secret message embedded using  $\pm k$  steganography in spatial domain », in *Security, Steganography, and Watermarking of Multimedia Contents V, volume 5020 of Proc. SPIE*, p. 595–606, 2005. (Cité dans la page 32.)
- [98] A. KER, « Steganalysis of lsb matching in grayscale images », *Signal Processing Letters, IEEE*, vol. 12, p. 441 – 444, june 2005. (Cité dans les pages 32 et 155.)
- [99] J. ZHANG, I. COX et G. DOERR, « Steganalysis for lsb matching in images with high-frequency noise », in *Multimedia Signal Processing, 2007. MMSP 2007. IEEE 9th Workshop on*, p. 385 –388, oct. 2007. (Cité dans les pages 32 et 155.)
- [100] J. FRIDRICH, M. GOLJAN et D. HOGEA, « Steganalysis of jpeg images : Breaking the f5 algorithm », in *Information Hiding*, vol. 2578 in *Lecture Notes in Computer Science*, p. 310–323, Springer Berlin / Heidelberg, 2003. (Cité dans les pages 33 et 179.)
- [101] R. BËHME et A. WESTFELD, « Breaking cauchy model-based jpeg steganography with first order statistics », in *Computer Security - ESORICS 2004*, vol. 3193, p. 125–140, Springer Berlin / Heidelberg, 2004. (Cité dans la page 33.)
- [102] D. FU, Y. SHI, D. ZOU et G. XUAN, « Jpeg steganalysis using empirical transition matrix in block DCT domain », in *Multimedia Signal Processing, 2006 IEEE 8th Workshop on*, p. 310 –313, oct. 2006. (Cité dans la page 33.)
- [103] C. ULLERICH et A. WESTFELD, « Weaknesses of mb2 », in *Digital Watermarking*, vol. 5041, p. 127–142, Springer Berlin / Heidelberg, 2008. (Cité dans la page 33.)

- [104] K. SULLIVAN, K. SOLANKI, B. MANJUNATH, U. MADHOW et S. CHANDRASEKARAN, « Determining achievable rates for secure zero divergence steganography », in *IEEE International Conference on Image Processing*, 2006. (Cité dans les pages 33 et 194.)
- [105] K. SULLIVAN, U. MADHOW, S. CHANDRASEKARAN et B. MANJUNATH, « Steganalysis for markov cover data with applications to images », *Information Forensics and Security, IEEE Transactions on*, vol. 1, p. 275 – 287, june 2006. (Cité dans la page 33.)
- [106] Y. SUN, F. LIU, B. LIU et P. WANG, « Steganalysis based on difference image », in *Digital Watermarking*, vol. 5450in *Lecture Notes in Computer Science*, p. 184–198, Springer Berlin / Heidelberg, 2009. (Cité dans la page 33.)
- [107] E. LEHMAN et J. ROMANO, *Testing Statistical Hypotheses, Second Edition*. Springer, 3rd éd., 2005. (Cité dans les pages 36, 37, 47, 55, 56, 73, 130, 174, 175, et 177.)
- [108] O. ZEITOUNI, J. ZIV et N. MERHAV, « When is the generalized likelihood ratio test optimal? », *Information Theory, IEEE Transactions on*, vol. 38, p. 1597 –1602, sep 1992. (Cité dans les pages 37 et 175.)
- [109] G. V. MOUSTAKIDES, « Optimum joint detection and estimation », in *IEEE International Symposium on Information Theory*, p. 2984 – 2988, 2011. (Cité dans la page 37.)
- [110] Y. MICHE, *Developing Fast Machine Learning Techniques with Applications to Steganalysis Problems*. Thèse doctorat, Aalto University School of Science and Technology (Espoo, Finland) and INPG (Grenoble, France), November 2010. (Cité dans les pages 38 et 180.)
- [111] C. SCOTT et R. NOWAK, « A Neyman-Pearson approach to statistical learning », *IEEE Trans. Inform. Theory*, vol. 51, no. 8, p. 3806–3819, 2005. (Cité dans la page 38.)
- [112] C. SCOTT, « Performance measures for Neyman-Pearson classification », *IEEE Trans. Inform. Theory*, vol. 53, no. 8, p. 2852–2863, 2007. (Cité dans la page 38.)
- [113] R. RADKE, S. ANDRA, O. AL-KOFAHI et B. ROYSAM, « Image change detection algorithms : a systematic survey », *Image Processing, IEEE Transactions on*, vol. 14, p. 294 –307, march 2005. (Cité dans les pages 42 et 179.)
- [114] Y. Z. HSU, H. H. NAGEL et G. REKERS, « New likelihood test methods for change detection in image sequences », *Computer Vision, Graphics, and Image Processing*, vol. 26, no. 1, p. 73 – 106, 1984. (Cité dans les pages 42 et 179.)
- [115] L. FILLATRE, I. NIKIFOROV et F. RETRAINT, « -optimal non-bayesian anomaly detection for parametric tomography », *Image Processing, IEEE Transactions on*, vol. 17, p. 1985 –1999, nov. 2008. (Cité dans les pages 42, 56, 125, et 166.)
- [116] L. L. CAM, *Asymptotic Methods in Statistical Decision Theory*. New York : springer, 1986. (Cité dans la page 53.)
- [117] L. SCHARF et B. FRIEDLANDER, « Matched subspace detectors », *Signal Processing, IEEE Transactions on*, vol. 42, p. 2146 –2157, aug 1994. (Cité dans la page 56.)
- [118] A. SRIPAD et D. SNYDER, « A necessary and sufficient condition for quantization errors to be uniform and white », *Signal Processing, IEEE Transactions on*, vol. 25, p. 442 – 448, oct. 1977. (Cité dans les pages 62 et 92.)
- [119] B. WIDROW et I. KOLLÁR, *Quantization Noise : Roundoff Error in Digital Computation, Signal Processing, Control, and Communications*. Cambridge, UK : Cambridge University Press, 2008. (Cité dans la page 62.)
- [120] T. LÉPINE et J.-L. MEYZONNETTE, *Bases de radiométrie optique*. Paris, France : Cépaduès, 2nde éd., juin 2001. (Cité dans les pages 80, 81, 83, et 84.)
- [121] A. MARION, *Acquisition & visualisation des images*. Eyrolles, 1997. (Cité dans les pages 80, 82, 84, 86, et 99.)

- [122] J. W. GOODMAN, *Introduction to Fourier Optics*. Roberts & Company Publishers, 3rd éd., January 2005. (Cité dans les pages 82, 85, et 86.)
- [123] A. JALOBÉANU, *Modèles, estimation bayésienne et algorithmes pour la déconvolution d'images satellitaires et aériennes*. Thèse doctorat, Université de Nice Sophia-Antipolis, 2001. (Cité dans les pages 85 et 86.)
- [124] M. BORN et E. WOLF, *Principles of Optics*. Oxford : Pergamon, 1970. (Cité dans la page 85.)
- [125] Y. G. LECLERC et S. ZUCKER, « The local structure of image discontinuities in one dimension », *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 9, no. 3, p. 341–355, 1987. (Cité dans les pages 87, 102, et 103.)
- [126] G. PAVLOVIC et A. TEKALP, « Maximum likelihood parametric blur identification based on a continuous spatial domain model », in *Proc. of Acoustics, Speech, and Signal Processing (ICASSP'91)*, p. 2489–2492 vol.4, apr 1991. (Cité dans la page 87.)
- [127] J. R. JANESICK, *Scientific Charge-Coupled Devices*. SPIE Press, 1st éd., January 2001. (Cité dans les pages 88, 89, et 90.)
- [128] G. C. HOLST et T. S. LOMHEIM, *CMOS/CCD Sensors and Camera Systems*. SPIE Press, 1st éd., June 2007. (Cité dans les pages 88, 89, et 90.)
- [129] G. HEALEY et R. KONDEPUDY, « Radiometric CCD camera calibration and noise estimation », *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 3, p. 267–276, 1994. (Cité dans les pages 89, 90, et 92.)
- [130] Y. REIBEL, M. JUNG, M. BOUHIFD, B. CUNIN et C. DRAMAN, « CCD or CMOS camera noise characterisation », *European Physical Journal Applied Physics*, vol. 21, p. 75–80, jan 2003. (Cité dans les pages 89, 90, 92, et 93.)
- [131] A. FOI, M. TRIMECHE, V. KATKOVNIK et K. EGIAZARIAN, « Practical poissonian-gaussian noise modeling and fitting for single-image raw-data », *Image Processing, IEEE Transactions on*, vol. 17, p. 1737–1754, Oct. 2008. (Cité dans les pages 89 et 93.)
- [132] R. RAMANATH, W. E. SNYDER, Y. YOO et M. S. DREW, « Color image processing pipeline », *IEEE Signal Processing Magazine*, p. 34–43, January 2005. (Cité dans la page 95.)
- [133] R. RAMANATH, W. E. SNYDER, G. L. BILBRO et W. A. S. III, « Demosaicking methods for bayer color arrays », *Journal of Electronic Imaging*, vol. 11, p. 306–315, 2002. (Cité dans la page 96.)
- [134] D. ALLEYSSON, « 30 ans de démosaïçage », *Traitement du Signal*, vol. 21, no. 6, 2004. (Cité dans les pages 96, 97, et 98.)
- [135] B. K. GUNTURK, J. GLOTZBACH, Y. ALTUNBASAK, R. W. SCHAFER et R. M. MERSEREAU, « Demosaicing : color filter array interpolation », in *IEEE Signal Processing Magazine*, p. 44–54, 2005. (Cité dans les pages 97 et 98.)
- [136] D. ALLEYSSON, S. SÚSTRUNK et J. HÉRAULT, « Linear demosaicing inspired by the human visual system », *IEEE Transactions on Image Processing*, vol. 14, no. 4, p. 439–449, 2005. (Cité dans la page 97.)
- [137] R. KIMMEL, « Demosaicing : image reconstruction from color CCD samples », *Image Processing, IEEE Transactions on*, vol. 8, p. 1221–1228, sept. 1999. (Cité dans la page 97.)
- [138] J. HADAMARD, « Sur les problèmes aux dérivées partielles et leur signification physique », *Princeton University Bulletin*, vol. 13, 1901. (Cité dans la page 97.)
- [139] J. IDIER, *Approche bayésienne pour les problèmes inverses*. Paris : Hermès science publ. : Lavoisier, 2001. (Cité dans la page 98.)
- [140] A. MOHAMMAD-DJAFARI, *Problèmes inverses*. Cours de Supélec et de DEA Automatique et traitement du signal., 2001. (Cité dans la page 98.)

- [141] M. NASHED, « Operator-theoretic and computational approaches to ill-posed problems with applications to antenna theory », *Antennas and Propagation, IEEE Transactions on*, vol. 29, p. 220–231, Mar 1981. (Cité dans la page 98.)
- [142] G. SHARMA et AL., *Digital Color Imaging Handbook*. Boca Raton, FL, USA : CRC Press, Inc., 2002. (Cité dans la page 98.)
- [143] MUMFORD, D. et SHAH, J., « Optimal approximations by piecewise smooth functions and associated variational problems », *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, p. 577–685, 1989. (Cité dans les pages 102, 103, et 131.)
- [144] C. BRUNI, A. DE SANTI, G. KOCH et C. SINISGALLI, « Identification of discontinuities in blurred noisy signals », *Circuit and systems-I, IEEE Transactions on*, vol. 44, p. 422 – 433, May 1997. (Cité dans les pages 102, 131, 132, 133, et 138.)
- [145] G. MATHERON, « Modèle séquentiel de partition aléatoire », *rapport technique*, 1968. (Cité dans la page 102.)
- [146] A. F. TIMAN, *Theory of Approximation of Functions of a Real Variable*. Pure and applied mathematics, Oxford : Pergamon Press, 1963. (Cité dans les pages 110, 190, et 191.)
- [147] G. G. LORENTZ, *Approximations of functions*. Holt, Rinehart & Winsto, 1966. (Cité dans les pages 110 et 190.)
- [148] L. VEIDINGER, « On the numerical determination of the best approximations in the chebyshev sense », *Numerische Mathematik*, vol. 2, p. 99–105, 1960. 10.1007/BF01386215. (Cité dans les pages 110 et 192.)
- [149] D. L. DONOHO et I. M. JOHNSTONE, « Ideal spatial adaptation by wavelet shrinkage », *Biometrika*, vol. 81, no. 3, p. 425–455, 1994. (Cité dans les pages 112, 113, et 114.)
- [150] D. DONOHO, « De-noising by soft-thresholding », *Information Theory, IEEE Transactions on*, vol. 41, p. 613 –627, may 1995. (Cité dans les pages 112, 113, et 114.)
- [151] D. L. DONOHO et I. M. JOHNSTONE, « Adapting to Unknown Smoothness via Wavelet Shrinkage », *Journal of the American Statistical Association*, vol. 90, p. 1200–1224, déc. 1995. (Cité dans les pages 112, 113, et 114.)
- [152] S. MALLAT, « A theory for multiresolution signal decomposition : the wavelet representation », *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, p. 674 –693, jul 1989. (Cité dans la page 112.)
- [153] S. MALLAT, *A wavelet tour of signal processing*. Academic Press, 3rd ed. éd., 2008. (Cité dans les pages 112 et 113.)
- [154] T. TAO et B. VIDAKOVIC, « Almost everywhere behavior of general wavelet shrinkage operators », *Applied and Computational Harmonic Analysis*, vol. 9, no. 1, p. 72 – 82, 2000. (Cité dans la page 114.)
- [155] G. DAVIS, S. MALLAT et M. AVELLANEDA, « Adaptive greedy approximations », *Constructive Approximation*, vol. 13, p. 57–98, 1997. 10.1007/BF02678430. (Cité dans la page 115.)
- [156] S. MALLAT et Z. ZHANG, « Matching pursuits with time-frequency dictionaries », *Signal Processing, IEEE Transactions on*, vol. 41, p. 3397 –3415, dec 1993. (Cité dans la page 115.)
- [157] Y. PATI, R. REZAIIFAR et P. KRISHNAPRASAD, « Orthogonal matching pursuit : recursive function approximation with applications to wavelet decomposition », in *Conf. Rec. of The 27th Asilomar Conference, Signals, Systems and Computers*, p. 40 –44 vol.1, nov 1993. (Cité dans la page 115.)
- [158] S. S. CHEN, D. L. DONOHO et M. A. SAUNDERS, « Atomic decomposition by basis pursuit », *SIAM J. Sci. Comput.*, vol. 20, p. 33–61, dec. 1998. (Cité dans la page 115.)
- [159] M. AHARON, M. ELAD et A. BRUCKSTEIN, « -svd : An algorithm for designing overcomplete dictionaries for sparse representation », *Signal Processing, IEEE Transactions on*, vol. 54, p. 4311 –4322, nov. 2006. (Cité dans la page 115.)

- [160] C. ZITZMANN, R. COGRANNE, F. RETRAINT, I. NIKIFOROV, L. FILLATRE et P. CORNU, « Statistical decision methods in hidden information detection », in *Information Hiding*, p. 163 – 177, LNCS vol.6958, Springer, 2011. (Cité dans la page 125.)
- [161] L. FILLATRE, *Détection et localisation d'anomalies à partir d'un nombre limité de projections tomographiques*. Thèse doctorat, Université de Technologie de Troyes, décembre 2004. (Cité dans les pages 125, 166, 169, et 178.)
- [162] G. SEBER et C. WILD, *Nonlinear Regression*. Wiley, 1989. (Cité dans les pages 127, 128, 129, et 136.)
- [163] S. GEMAN et D. GEMAN, « Stochastic relaxation, gibbs distributions and the bayesian restoration of images », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, p. 721–741, November 1984. (Cité dans les pages 127 et 179.)
- [164] H. LACRESSE, *Détection de pannes en présence de paramètres de nuisance non-linéaires bornés*. Thèse doctorat, Université de Technologie de Troyes, mars 2004. (Cité dans la page 130.)
- [165] I. NIKIFOROV, « Eléments de théorie de la décision statistique II : compléments, Chap. 2 », in *Décision et reconnaissance des formes en signal*, p. 43–84, Hermes Science Publications (Traité IC2), Edited by R.Lengellé, 2002. (Cité dans les pages 130 et 169.)
- [166] D. V. HINKLEY, « Inference about the change-point in a sequence of random variables », *Biometrika*, vol. 57, no. 1, p. 1–17, 1970. (Cité dans les pages 130 et 143.)
- [167] M. BASSEVILLE et I. V. NIKIFOROV, *Detection of abrupt changes : theory and application*. Prentice Hall, 1993. (Cité dans les pages 130 et 143.)
- [168] I. NIKIFOROV, « A generalized change detection problem », *Information Theory, IEEE Transactions on*, vol. 41, no. 1, 1995. (Cité dans les pages 130 et 143.)
- [169] H. STRYHN, *Spatial change point models applied to image segmentation*. Thèse doctorat, Department of Mathematics and Physics, Royal Veterinary and Agricultural University, Copenhagen, August 1994. (Cité dans les pages 130, 131, et 134.)
- [170] M. BASSEVILLE et I. NIKIFOROV, « Fault isolation for diagnosis : nuisance rejection and multiple hypotheses testing », rapport de recherche, INRIA, 2002. (Cité dans les pages 130 et 143.)
- [171] M. RUDEMO et H. STRYHN, « Approximating the distribution of maximum likelihood contour estimators in two-region images », *Scandinavian Journal of Statistics*, vol. 21, no. 1, p. 41–55, 1994. (Cité dans la page 131.)
- [172] R. JAIN, R. KASTURI et B. G. SCHUNCK, *Machine vision*. New York, NY, USA : McGraw-Hill, Inc., 1995. (Cité dans la page 131.)
- [173] I. PITAS, *Digital Image Processing Algorithms and Applications*. New York, NY, USA : John Wiley & Sons, Inc., 1st éd., 2000. (Cité dans la page 131.)
- [174] R. C. GONZALEZ et R. E. WOODS, *Digital Image Processing*. Prentice Hall, 3rd edition éd., 2008. (Cité dans la page 131.)
- [175] F. CATTÉ, P. L. LIONS, J. M. MOREL et T. COLL, « Image selective smoothing and edge detection by nonlinear diffusion », *SIAM Journal on Numerical Analysis*, vol. 29, p. 182–193 & 845–866, Feb. 1992. (Cité dans la page 131.)
- [176] P. PERONA et J. MALIK, « Scale-space and edge detection using anisotropic diffusion », *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, p. 629–639, jul 1990. (Cité dans les pages 131 et 132.)
- [177] S. MALLAT et W. HWANG, « Singularity detection and processing with wavelets », *Information Theory, IEEE Transactions on*, vol. 38, p. 617–643, mar. 1992. (Cité dans la page 131.)
- [178] L. ZHANG et P. BAO, « Edge detection by scale multiplication in wavelet domain », *Pattern Recognition Letters*, vol. 23, no. 14, p. 1771 – 1784, 2002. (Cité dans les pages 131 et 133.)

- [179] C. BRUNI, A. DE SANTIS, D. IACOVIELLO et G. KOCH, « Modeling for edge detection problems in blurred noisy images », *Image Processing, IEEE Transactions on*, vol. 10, p. 1447–1453, oct 2001. (Cité dans les pages 131 et 138.)
- [180] J. CANNY, « A computational approach to edge detection », *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-8, p. 679–698, nov. 1986. (Cité dans la page 132.)
- [181] M. ABRAMOWITZ et I. A. STEGUN, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York : Dover, ninth dover printing, tenth gpo printing éd., 1964. (Cité dans la page 132.)
- [182] D. MARR et E. HILDRETH, « Theory of Edge Detection », *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 207, no. 1167, p. 187–217, 1980. (Cité dans la page 133.)
- [183] W. FELLER, *An Introduction to Probability Theory and Its Applications*, vol. 2. Wiley, 3rd éd., 1968. (Cité dans la page 146.)
- [184] BOSS CONTEST, « Break Our Steganographic System », 2010. (Cité dans les pages 150, 151, 152, 155, et 156.)
- [185] G. SCHAEFER et M. STICH, « UCID - an uncompressed colour image database », in *SPIE Proceedings*, vol. 5307, (San Jose, USA), p. 472–480, <http://vision.cs.aston.ac.uk/datasets/UCID/>, 2004. (Cité dans les pages 151 et 152.)
- [186] R. COGRANNE, C. ZITZMANN, L. FILLATRE, I. NIKIFOROV, F. RETRAINT et P. CORNU, « A cover image model for reliable steganalysis », in *Information Hiding*, p. 178 – 192, LNCS vol.6958, Springer, 2011. (Cité dans les pages 159 et 193.)
- [187] S. BAYRAM, İsmail AVCIBAŞ, B. SANKUR et N. MEMON, « Image manipulation detection », *Journal of Electronic Imaging*, vol. 15, no. 4, p. 041102, 2006. (Cité dans la page 165.)
- [188] S. BAYRAM, H. SENCAR, N. MEMON et I. AVCIBAS, « Source camera identification based on cfa interpolation », in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 3, p. III – 69–72, sept. 2005. (Cité dans la page 166.)
- [189] S. CHANG, B. YU et M. VETTERLI, « Adaptive wavelet thresholding for image denoising and compression », *Image Processing, IEEE Transactions on*, vol. 9, p. 1532–1546, sep 2000. (Cité dans la page 166.)
- [190] F. GUSTAFSSON et R. KARLSSON, « Statistical results for system identification based on quantized observations », *Automatica*, vol. 45, no. 12, p. 2794 – 2801, 2009. (Cité dans la page 166.)
- [191] I. NIKIFOROV et J. RAGOT, « Surveillance des systèmes dynamiques : approches statistique et géométrique », in *formation de l'école doctorale de l'UTT*, 2007-2008. (Cité dans la page 169.)
- [192] M. FOULADIRAD et I. NIKIFOROV, « Optimal statistical fault detection with nuisance parameters », *Automatica*, vol. 41, no. 7, p. 1157–1171, 2005. (Cité dans la page 178.)
- [193] X. YU, I. REED et A. STOCKER, « Comparative performance analysis of adaptive multispectral detectors », *Signal Processing, IEEE Transactions on*, vol. 41, p. 2639–2656, aug 1993. (Cité dans la page 179.)
- [194] T. HORPRASERT, D. HARWOOD et L. S. DAVIS, « A statistical approach for real-time robust background subtraction and shadow detection », p. 1–19, 1999. (Cité dans la page 179.)
- [195] R. M. HARALICK et L. G. SHAPIRO, « Image segmentation techniques », *Computer Vision, Graphics, and Image Processing*, vol. 29, no. 1, p. 100 – 132, 1985. (Cité dans la page 179.)
- [196] M. STYNER, C. BRECHBUHLER, G. SZCKELY et G. GERIG, « Parametric estimate of intensity inhomogeneities applied to mri », *Medical Imaging, IEEE Transactions on*, vol. 19, p. 153–165, march 2000. (Cité dans la page 179.)



- [197] J. KYBIC et M. UNSER, « Fast parametric elastic image registration », *Image Processing, IEEE Transactions on*, vol. 12, p. 1427 – 1442, nov. 2003. (Cité dans la page 179.)
- [198] J. MODESTINO et J. ZHANG, « A markov random field model-based approach to image interpretation », *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 14, p. 606 –615, jun 1992. (Cité dans la page 179.)
- [199] A. BLAKE, C. ROTHER, M. BROWN, P. PEREZ et P. TORR, « Interactive image segmentation using an adaptive gmmrf model », in *Computer Vision - ECCV 2004*, vol. 3021, p. 428–441, Springer Berlin / Heidelberg, 2004. (Cité dans la page 179.)
- [200] E. SIMONCELLI et E. ADELSON, « Noise removal via bayesian wavelet coring », in *Image Processing, 1996. Proceedings., International Conference on*, vol. 1, p. 379 –382 vol.1, sep 1996. (Cité dans la page 179.)
- [201] J. LIU et P. MOULIN, « Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients », *Image Processing, IEEE Transactions on*, vol. 10, p. 1647 –1658, nov 2001. (Cité dans la page 179.)
- [202] J. BIOUSCAS-DIAS, « Bayesian wavelet-based image deconvolution : a GEM algorithm exploiting a class of heavy-tailed priors », *Image Processing, IEEE Transactions on*, vol. 15, p. 937 –951, avril 2006. (Cité dans la page 179.)
- [203] C. RICHARD, *Une méthodologie pour la détection à structure imposée. Applications au plan temps-fréquence*. Thèse doctorat, Université de Technologie de Compiègne, décembre 1998. (Cité dans les pages 180, 181, 183, et 184.)
- [204] R. LENGELLÉ, « Apprentissage de règles de décisions à structure imposée et contrôle de complexité, Chap. 4 », in *Décision et reconnaissance des formes en signal*, p. 109–140, Hermes Science Publications (Traité IC2), Edited by R.Lengellé, 2002. (Cité dans la page 180.)
- [205] B. SCHOLKOPF et A. J. SMOLA, *Learning with Kernels : Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA : MIT Press, 2001. (Cité dans la page 183.)
- [206] I. HERO, A.O., J. FESSLER et M. USMAN, « Exploring estimator bias-variance tradeoffs using the uniform cr bound », *Signal Processing, IEEE Transactions on*, vol. 44, p. 2026 –2041, aug 1996. (Cité dans la page 185.)
- [207] L. L. CAM, « Maximum likelihood : An introduction », *Intl. Stat. Rev*, vol. 58, p. 153–171, 1990. (Cité dans la page 186.)
- [208] J. HARMSSEN et W. PEARLMAN, « Steganalysis of additive noise modelable information hiding », in *Security, Steganography, and Watermarking of Multimedia Contents V, volume 5020 of Proc. SPIE*, 2003. (Cité dans la page 193.)
- [209] R. COGRANNE, C. ZITZMANN, L. FILLATRE, I. NIKIFOROV, F. RETRAINT et P. CORNU, « Reliable detection of hidden information based on a non-linear local model », in *Statistical Signal Processing, Proc. of IEEE Workshop on*, p. 493 – 496, 2011. (Cité dans la page 193.)
- [210] Y. WANG et P. MOULIN, « Perfectly secure steganography : Capacity, error exponents, and code constructions », *IEEE Trans. Information Theory special issue on Security*, 2008. (Cité dans la page 194.)
- [211] A. D. KER, « A capacity result for batch steganography », *Signal Processing Letters*, vol. 14, no. 8, p. 525–528, 2007. (Cité dans la page 195.)
- [212] A. D. KER, T. PEVNÝ, J. KODOVSKÝ et J. FRIDRICH, « The square root law of steganographic capacity », in *Proceedings of the 10th ACM workshop on Multimedia and security, MM&Sec '08*, (New York, NY, USA), p. 107–116, ACM, 2008. (Cité dans la page 195.)
- [213] T. FILLER, A. D. KER et J. FRIDRICH, « The square root law of steganographic capacity for markov covers », in *Media Forensics and Security XI, Proc SPIE 7254*, SPIE, p. 801–0811, 2009. (Cité dans la page 195.)
- [214] A. KER, « The square root law in stegosystems with imperfect information », in *Information Hiding*, vol. 6387 in *Lecture Notes in Computer Science*, p. 145–160, Springer Berlin / Heidelberg, 2010. (Cité dans la page 195.)

## Publications

---

Ce mémoire reprend des concepts, des idées et des supports utilisés précédemment dans les publications suivantes :

Articles de revues internationales (avec comité de lectures) :

Un article en fin de rédaction pour IEEE trans. on Information Forensics and Security

Présentations dans des conférences internationales (avec actes et comité de lectures) :

• **R.Cogranne**, C.Zitzmann, L.Fillatre, F. Retraint, I. Nikiforov et P. Cornu “*Statistical decision by using quantized observations*”. IEEE Intl Symposium on Information Theory, pp. 1135-1139, 31 juillet - 5 août 2011. [\[doi\]](#) [\[article au format pdf\]](#)

• **R.Cogranne**, C.Zitzmann, L.Fillatre, I. Nikiforov, F. Retraint et P. Cornu “*Reliable Detection of Hidden Information Based on a Non-Linear Local Model*”. IEEE Workshop on Statistical Signal Processing, pp. 493-496, 28-30 juin 2011. [\[doi\]](#) [\[article au format pdf\]](#)

• C.Zitzmann, **R.Cogranne**, F. Retraint, I. Nikiforov, L.Fillatre et P. Cornu “*Hypothesis Testing by Using Quantized Observations*”. IEEE Workshop on Statistical Signal Processing, pp. 501-504, 28-30 juin 2011. [\[doi\]](#) [\[article au format pdf\]](#)

• **R.Cogranne**, C.Zitzmann, L.Fillatre, I. Nikiforov, F. Retraint et P. Cornu “*A Cover Image Model For Reliable Steganalysis*”. Information Hiding conference\*, Springer LNCS vol.6958, pp. 178-192, 18-20 mai 2011 (**Obtention d’une bourse de 700\$**). [\[doi\]](#) [\[article au format pdf\]](#) .

\* taux d’acceptation donné à titre indicatif  $\approx 30\%$

• C.Zitzmann, **R.Cogranne**, F. Retraint, I. Nikiforov, L.Fillatre et P. Cornu “*Statistical Decision Methods in Hidden Information Detection*”. Information Hiding conference\*, Springer LNCS vol.6958, pp. 163-177, 18-20 mai 2011. [\[doi\]](#) [\[article au format pdf\]](#)

Présentations dans des conférences nationales (avec actes et comité de lectures) :

• **R.Cogranne**, C.Zitzmann, L.Fillatre, F. Retraint, I. Nikiforov et P. Cornu “*Détection quasi-optimale d’informations cachées basée sur un modèle local non-linéaire*”. XXIII<sup>ème</sup> colloque du GRETSI, pp. xxx-xxx, 5-8 septembre 2011(**Présentation orale**). [\[article au format pdf\]](#)

• C.Zitzmann, **R.Cogranne**, I. Nikiforov, L.Fillatre, F. Retraint et P. Cornu “*Détection optimale à base de distribution laplacienne quantifiée*”. XXIII<sup>ème</sup> colloque du GRETSI, pp. xxx-xxx, 5-8 septembre 2011. [\[article au format pdf\]](#)

• **R.Cogranne**, L.Fillatre, F. Retraint et C.Zitzmann “*Détection Optimale d’information cachées indépendante du contenu des images*”, JCS’11, janvier 2011, Saint-Cyr Coëtquidan. [\[article au format pdf\]](#)

• **R.Cogranne**, L.Fillatre, F. Retraint et C.Zitzmann “*Modélisation des images naturelles pour la criminalistique numérique*”, workshop 3SGS, Reims, septembre 2010 (**Prix de la meilleure présentation**).

## Activités d'enseignements

---

«Enseigner, c'est apprendre deux fois.»  
— Joseph Joubert, essayiste, 1754 – 1824

- Fonctions d'une variable réelle (MATH01) : participation à la mise en place des TPs & 144h de TP.
- Fonctions de plusieurs variables réelles et applications (MATH02), niveau L2 : 85h de TD.
- Algèbre linéaire (MATH23), niveau L2 : 34h de TD et 24h de TD.
- Éléments de probabilités (SY01), niveau L2 : 34h de TD (51 UTP).
- Techniques mathématiques de l'ingénieur (MT12), niveau L3 : 34h de TD.
- Encadrements de 2 étudiants dans une TITS, "*Travail d'Investigation Technologique et Scientifique*" (8UTP).
- Co-Encadrements de unités de valeurs AC, "*Acquisition de Connaissance*" (16 UTP)
- Architecture et administration des systèmes (LO03) à l'UTSEUS, Shanghai, niveau L3 : 65h de cours/TD.
- Participation à l'encadrement d'un stage de Master II (T. H. Thai) et des 2 stages d'ingénieur (L. Hoppenot et S. Savarraradjaj).
- Evaluation des rapports et des soutenances de 8 stages d'ingénieur (en remplacement de Cédric Richard et Nada Matta).

## Animation de la recherche

---

- Reviewer pour la revue *IEEE Trans. on Signal Processing* et la conférence *IEEE SSP 2011*.
- Membre du *Technical Program Committee* du workshop *Information hiding techniques for Internet anonymity and privacy* (IHTIAP 2012), Venise, 24-29 juin 2012.
- Membre du *Comité de programme* des Journées Codes et Stéganographie (JCS 2012), Saint-Cyr, Paris, 19-20 mars 2012.
- Co-organisateur de la journée "*Stéganographie et Stéganalyse*" du GDR-ISIS, 25 mars 2010, Paris.

## Implication dans la vie de l'UTT

---

- Membre du conseil d'administration de l'UTT, 2008-2010, représentant des doctorants.
- Membre du conseil de l'école doctorale de l'UTT, 2009-2011, représentant des doctorants.
- Membre du conseil de la documentation de l'UTT, 2008-2010, représentant des doctorants.
- Secrétaire général (2009-2010), puis trésorier (2010-2012) d'Ellidoc, l'association des doctorants de l'UTT.
- Participation aux fêtes de la science, en 2008, 2009, 2010 et 2011.
- Présentations de nos travaux de recherches aux journées portes ouvertes de l'UTT en 2009 et 2010.
- Présentations aux forum des doctorants de 2009 et 2010.
- Participation à l'organisation du forum des doctorants de 2010.

- Présentation des à des étudiants de master I (spécialité OSS) et à des étudiants de terminales.