

École polytechnique de Louvain

Quantification d'incertitude pour le traitement d'images par problèmes inverses

Auteur : Baptiste PIRLOT

Promoteurs : **Laurent JACQUES**, Nelly PUSTELNIK

Lecteurs : **François GLINEUR**

Année académique 2020–2021

Master [120] : ingénieur civil électricien

Remerciements

Je voudrais particulièrement remercier Professeur Laurent Jacques et Nelly Pustelnik pour les innombrables réunions, toutes plus virtuelles les unes que les autres, qui m'ont permis de rester motivé tout au long de cette année particulière. Leurs nombreux conseils ont été d'une aide précieuse.

Je voudrais également remercier mes parents qui ont toujours été présents pour m'aider durant tout la durée de mes études. Et surtout un grand merci à ma maman pour la relecture de ce texte.

Résumé

Ce travail traite de la quantification d'incertitude dans le cas de problèmes inverses en traitement d'images. On s'intéresse plus particulièrement à la formulation bayésienne de ces problèmes avec une approche convexe.

Les méthodes traditionnelles pour quantifier l'incertitude font usage des méthodes de Monte-Carlo par chaînes de Markov (MCMC). Ces méthodes sont réputées coûteuses en termes de temps de calcul. Récemment une méthode reposant sur l'estimation du maximum à postériori est apparue. Cette méthode, nommée BUQO suggère de considérer une forme log-concave du problème inverse et d'utiliser l'optimisation convexe pour le résoudre. Il est alors possible d'utiliser l'estimation MAP de l'image pour analyser l'incertitude autour d'une structure d'intérêt à l'aide d'une définition d'une région de crédibilité et d'un test d'hypothèse bayésien, ceci plus efficacement qu'avec les méthodes de Monte-Carlo.

Les deux méthodes présentées sont mises en parallèle pour mettre en lumière le lien entre celles-ci. Après une analyse dans une base de pixels, une autre analyse est effectuée dans la base des ondelettes. Nos résultats montrent une corrélation entre les méthodes MCMC et BUQO d'une part, et dans l'expression en base de pixels et d'ondelettes d'autre part.

Table des matières

1	Introduction et contexte	2
2	Introduction générale sur les problèmes inverses	3
2.1	Approche bayésienne	4
2.2	Distribution postérieure de notre modèle	5
2.2.1	Parcimonie et compressibilité dans la base Ψ	5
2.3	Formulation finale du problème	7
2.4	Opérateur proximal	8
2.5	Algorithme de Condat : primal-dual	10
2.5.1	Application sur l'estimateur MAP	11
2.6	Convolution	12
3	Sources d'incertitude et méthode classique de quantification	14
3.1	Quantifier l'incertitude et MCMC	14
3.2	Sources supplémentaires d'incertitudes	17
4	Méthode dite BUQO	20
4.1	Région HPD	20
4.2	Test d'hypothèse bayésien	23
4.3	Ensemble $\mathcal{S}$	23
4.4	Intersection entre les ensembles $\tilde{\mathcal{C}}_\alpha$ et $\mathcal{S}$	26
4.5	Interprétation de la distance entre les deux ensembles	28
4.6	BUQO en résumé	29
4.7	Ajout Artefact	30
4.8	Travail dans le domaine des ondelettes	31
5	Résultats	33
5.1	Déconvolution et débruitage	33
5.2	Quantification de l'incertitude via MCMC	37
5.3	Quantification de l'incertitude via BUQO	39
5.3.1	Influence de la hauteur de la structure	40
5.3.2	Influence de la taille de la structure	42
5.3.3	Influence de l'emplacement	43
5.4	Cartes d'incertitude	45
6	Conclusion et perspectives	48

1 Introduction et contexte

La quantification d'incertitude dans le cadre du traitement d'images est une question importante. De nombreux domaines nécessitent de prendre des décisions à la simple vue d'une image. Citons le domaine de la médecine, à la sortie d'un IRM ou d'autre autre examen d'imagerie médicale, le médecin devra analyser les résultats pour poser un diagnostic précis au patient. Il peut y avoir de l'incertitude autour de cette image et des artefacts de reconstruction peuvent apparaître. Quid si cet artefact prend l'apparence d'une forme suspecte ? Est-ce une tumeur ou une structure imaginaire ? Ce n'est pas une question anodine et il faut pouvoir quantifier la véracité de cette structure visible sur l'image. C'est ce qui va nous concerner tout le long de ce travail qui va se concentrer sur la mise en relation des différentes méthodes disponibles pour quantifier l'incertitude, en particulier la méthode BUQO dans son analyse approfondie en base de pixels et en base d'ondelettes.

La notion de problème inverse est d'abord abordée, dans sa formulation bayésienne et log-concave. Le chapitre suivant introduit la méthode classique de quantification (MCMC) et aborde les éléments qui amènent cette incertitude. Le chapitre 4 se concentre sur l'étude d'une méthode nommée BUQO qui utilise l'estimation à posteriori et des régions de crédibilité pour permettre de résoudre plus efficacement cette quantification. Enfin, dans le dernier chapitre des simulations sont menées pour illustrer le problème de la quantification d'incertitude.

2 Introduction générale sur les problèmes inverses

La notion de problème inverse est largement utilisée dans une multitude de domaines et dans le cas qui nous concerne, l'analyse ou le traitement d'image[19].

Selon [14] une définition opérationnelle d'un problème inverse est qu'il consiste à déterminer des *causes* connaissant des *effets*. La dualité du problème inverse est appelée problème direct et consiste à déterminer des *effets* connaissant des *causes*. La raison pour laquelle nous devons utiliser la notion de problème inverse est que l'électronique qui réalise la capture d'une image ne la reproduit pas sans bruit ou autre détérioration (convolution, décimation par exemple). En d'autres mots, les senseurs électroniques ne reproduisent pas exactement ce que nos yeux pourraient voir. Connaissant donc les effets produits sur la prise de vue par les senseurs, nous pouvons récupérer l'image sans bruit et sans détérioration. Ceci est vrai dans le cas le plus idéal, mais en réalité c'est presque chose impossible. Comment connaître avec exactitude le niveau de bruit, la convolution, ou tout autre opérateur induit par la capture de l'image ?

Une des difficultés principales dans la résolution des problèmes inverses d'images est le caractère mal posé de ceux-ci. Cette notion a été introduite par Jacques Hadamard dans [9] Il énonce qu'un problème est bien posé si :

- Une solution existe
- La solution est unique
- La solution dépend de façon continue des données.

Un problème qui ne respecte pas ces conditions est dit mal posé.

Plusieurs raisons mènent au fait que notre problème puisse être mal posé. La principale est qu'il est sous-déterminé, dans le sens où la solution n'est pas unique (dans le problème $y = \Theta x$ il y a une infinité de x possibles, par exemple si l'opérateur Θ réalise une décimation) . Une autre raison qui amène un problème à être mal posé est la présence du bruit, les données reçues ne sont pas les données réelles étant donné qu'elles diffèrent non continûment avec le bruit. Ces raisons sont en particulier vraies dans le cas du traitement d'images. Il y a plus d'inconnues que d'observations et l'image reçue est entachée de bruit, autrement dit une petite perturbation entraîne ici une possible grande différence entre la solution calculée et la solution réelle .

Le problème qui va nous intéresser pour la suite est l'estimation de l'image inconnue $x \in \mathbb{R}^N$ à partir des observations $y \in \mathbb{R}^M$ défini tel que :

$$y = \Phi x + b \tag{1}$$

avec $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^M$ un opérateur d'observation et $b \in \mathbb{R}^M$ une réalisation d'un bruit d'énergie limitée tel que $\|b\|^2 \leq \epsilon$.

Comme dit plus haut, ce problème est souvent mal posé et mal conditionné, ce qui

amène de l'incertitude autour de la solution x . Cependant les méthodes bayésiennes nous permettent de résoudre ce problème en utilisant une information *à priori* sur la distribution de x , ce qui permet de régulariser le problème et ainsi de réduire l'incertitude autour de x .

2.1 Approche bayésienne

L'approche bayésienne des problèmes inverses développée dans [13] et [10] stipule que toutes les quantités sont considérées comme aléatoires. Le caractère aléatoire reflète l'incertitude qui découle des valeurs d'un point de vue de l'observateur. De cette perspective, la solution d'un problème inverse est la distribution de probabilités des quantités auxquelles on s'intéresse quand toute l'information disponible a été incorporée au modèle. Cette distribution est appelée distribution postérieure et représente le niveau de confiance que l'on a par rapport aux quantités observées. On peut déjà remarquer la différence essentielle entre l'approche bayésienne et l'approche classique de la régularisation. Pendant que la régularisation produit une estimée *raisonnable* du problème, l'approche bayésienne nous donne une solution qui est une distribution de probabilités qui peut alors être utilisée pour produire des estimées. La méthode bayésienne résout le fait que le problème soit mal posé en reformulant le problème comme une extension bien posée dans un espace plus large de distributions de probabilité.

Supposons qu'on puisse obtenir des informations sur X avant la mesure de Y , on peut alors déterminer une densité de probabilité $p_{prior}(x)$ appelée *densité de probabilité à priori*. Celle-ci exprime ce que l'on sait à propos de l'inconnue, à priori des mesures. Supposons qu'après avoir analysé tous les paramètres de mesure ainsi que toutes les informations disponibles à propos des variables, nous puissions déterminer la densité de probabilité conjointe de X et Y , que l'on écrit $p(x, y)$. Alors la densité marginale de l'inconnue X est

$$\int_{\mathbb{R}^M} p(x, y) dy = p_{prior}(x). \quad (2)$$

Si d'un autre côté nous pouvions déterminer la valeur de l'inconnue, en d'autres mots $X = x$, la densité de probabilité conditionnelle de Y (appelée fonction de vraisemblance) étant donné cette information serait

$$p(y|x) = \frac{p(x, y)}{p_{prior}(x)}, \text{ si } p_{prior}(x) \neq 0. \quad (3)$$

Finalement, supposons que les mesures $Y = y_{observed}$ soient données. La densité de probabilité conditionnelle

$$p(x|y_{observed}) = \frac{p(x, y_{observed})}{p(y_{observed})}, \text{ si } p(y_{observed}) = \int_{\mathbb{R}^N} p(x, y_{observed}) dx \neq 0 \quad (4)$$

est appelée la distribution à posteriori de X . Elle exprime ce que l'on connaît de X après l'observation $Y = y_{observed}$.

Une formulation bayésienne du problème inverse qui nous concerne serait alors : étant donné les données $Y = y_{observed}$, trouver la distribution conditionnelle de probabilité $p(x|y_{observed})$ de la variable X . Et le problème s'écrit comme

$$p_{post}(x) = p(x|y) = \frac{p_{prior}(x)p(y|x)}{p(y)}. \quad (5)$$

Notons que y ici correspond aux observations donc $y = y_{observed}$. En effet, quand la densité de probabilité postérieure est utilisée ce sont évidemment les observations qui sont utilisées pour y . Notons également que dans 5 la densité marginale $p(y)$ est définie comme

$$p(y) = \int_{\mathbb{R}^N} p(x, y) dx = \int_{\mathbb{R}^N} p(y|x) p_{prior} dx. \quad (6)$$

2.2 Distribution postérieure de notre modèle

Plusieurs informations peuvent être déduites de notre modèle 1. Une première information dont on peut se servir est que le bruit est borné. En d'autres mots on a que

$$\|\Phi x - y\| \leq \epsilon. \quad (7)$$

De ceci on peut déterminer une fonction de vraisemblance uniforme qui va se traduire en une fonction indicatrice donnée par

$$p(y|x) \propto 1_{\mathcal{B}(y, \epsilon)}(\Phi x) \quad (8)$$

avec $\mathcal{B}_2(y, \epsilon)$ la boule l_2 centrée en y de rayon ϵ . Cette fonction peut être comprise dans ce sens, la probabilité est de 1 si l'observation y respecte (7) et 0 si elle ne le respecte pas. En d'autres mots, cette fonction s'assure que les résidus sont contenus dans la boule déterminée par la borne ϵ du bruit.

Une autre information dont on peut se servir est que dans un domaine particulier induit par l'opérateur Ψ , on peut promouvoir la parcimonie ou la régularité à l'aide d'une norme l_p ($p \geq 1$). Ψ peut se référer à un opérateur différentiel dans notre cas, une transformée en ondelettes [1] ou les gradients verticaux et horizontaux qui définissent la variation totale (TV) [8] à priori. Si l'on définit Ψ comme une transformée en ondelettes, on va se servir de la parcimonie.

2.2.1 Parcimonie et compressibilité dans la base Ψ

On dit qu'un signal $x \in \mathcal{R}^n$ est compressible [5] si il peut être défini par un vecteur parcimonieux (qui ne contient que quelques valeurs non nulles) $s \in \mathcal{R}^n$ dans une base de transformation $\Psi \in \mathcal{R}^{n \times n}$. Il peut alors être écrit comme

$$x = \Psi s. \quad (9)$$

On dira alors qu'il est K -parcimonieux dans Ψ si il y a exactement K coefficients différents de zéro dans s . Si Ψ est générique (dans le sens où toutes les images ou signaux sont parcimonieux dans cette base), dans notre cas si Ψ est une base d'ondelettes, alors seulement quelques éléments de s permettent de reconstruire le signal original x . Il est bon de savoir qu'un signal dans la base d'ondelettes est parcimonieux mais à quoi cela va-t-il nous servir ? La réponse est simple, le bruit n'est pas parcimonieux dans cette base ! Sur la Figure ?? on remarque que l'image (les illustrations sont obtenues via une image utilisée comme test dans les simulations du chapitre 5) est composée de presque 50000 coefficients d'ondelettes nuls, alors que les coefficients du bruit sont répartis de façon homogène sur toutes les valeurs.

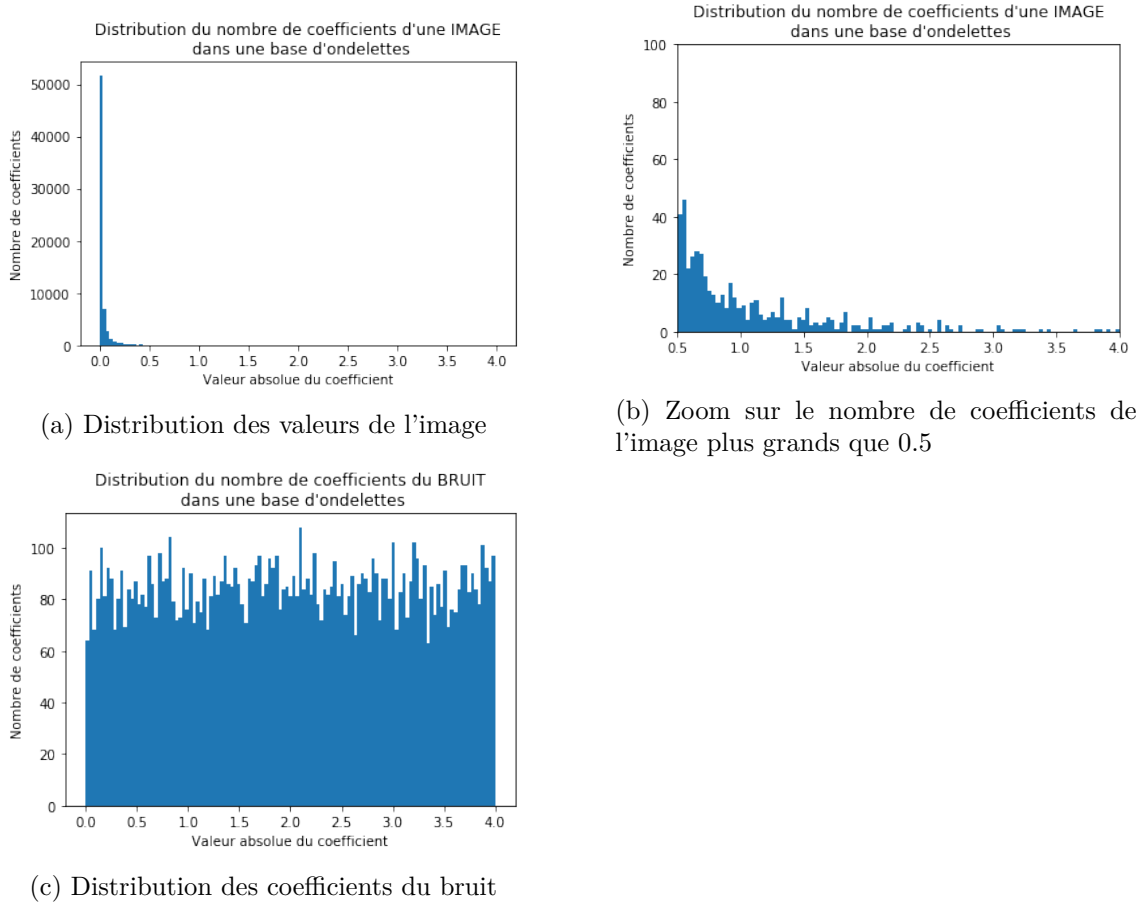


FIGURE 1

En promouvant la parcimonie du signal bruité, on peut alors retrouver une version non bruitée (ou moins bruitée) du signal. Pour réaliser ceci nous allons utiliser la norme l_1 . En effet comme on peut le voir sur la Figure 2 la norme l_1 (comme la norme l_0 mais la l_1 est plus facile à calculer) promeut la parcimonie étant donné que la minimisation du problème en cette norme ne nécessite qu'une seule coordonnée, pendant que la solution en norme l_2 nécessite les deux coordonnées (dans un cas à deux dimensions).

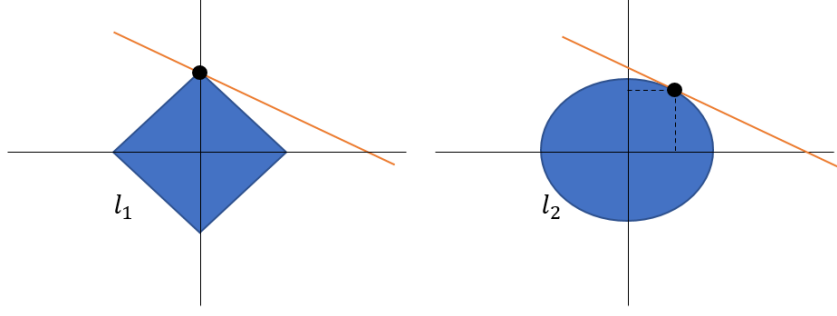


FIGURE 2 – Comparaison de la solution de la minimisation avec différentes normes

Cette condition sera alors exprimée comme

$$\|\Psi x\|_1 \text{ avec } \Psi : \mathbb{R}^N \rightarrow \mathbb{R}^L \quad (10)$$

Enfin, étant donné que notre image de base est composée de pixels dont la valeur est comprise entre 0 et 1, on va pouvoir rajouter cette information dans notre problème. La condition de parcimonie et la condition sur la valeur des pixels sont des conditions que l'on sait déduire de l'image originale x et constituent donc deux informations *à priori*.

2.3 Formulation finale du problème

Comme nous le verrons plus loin, une manière efficace de résoudre le problème inverse en image est d'utiliser des méthodes d'optimisation convexe. Une formulation pratique pour utiliser la convexité est la formulation log-concave [2].

Une fonction f non négative définie sur $[0, \infty[$ est dite log-concave [20] si elle respecte

$$f(\alpha + (1 - \alpha)y) \geq f^\alpha(x)f^{1-\alpha}(y) \text{ pour } 0 \leq x, y \leq \infty \text{ et } 0 < \alpha < 1. \quad (11)$$

La distribution postérieure prend alors la forme

$$p(x|y) \propto e^{-g_1(x,y) - g_2(x)} \quad (12)$$

avec

$$g_1(x, y) = \iota_{\mathcal{B}(y, \epsilon)}(\Phi x) \quad (13)$$

et

$$g_2(x) = -\log p(x) = \lambda \|\Psi x\|_1 + \iota_{[0,1]^N} \quad (14)$$

λ étant un paramètre de régularisation.

En utilisant l'estimateur du maximum à posteriori (MAP) on peut désormais résoudre le problème en profitant des propriétés convexes de notre formulation

$$x^\dagger \in \arg \max_{x \in \mathbb{R}^N} p(x|y). \quad (15)$$

$p(x|y)$ étant une exponentielle négative, cette expression devient

$$x^\dagger \in \arg \min_{x \in \mathbb{R}^N} g_1(x, y) + g_2(x). \quad (16)$$

Finalement, l'estimateur MAP est donné par

$$x^\dagger \in \arg \min_{x \in \mathbb{R}^N} \lambda f(\Psi x) + \iota_{\mathcal{C}} + \iota_{\mathcal{B}_2(y, \epsilon)}(\Phi x), \quad (17)$$

f étant la norme l_1 et $\mathcal{C} = [0, 1]^N$ comme expliqué plus haut.

Ou sous une autre forme

$$\text{minimiser } \|\Psi x\|_1 \text{ t.q. } x \in \mathbb{R}_+ \text{ et } \|\Phi x - y\| \leq \epsilon. \quad (18)$$

2.4 Opérateur proximal

Avant de rentrer plus en détail dans la résolution de l'estimateur MAP, il convient d'introduire quelques notions relatives aux opérateurs proximaux qui nous seront bien utiles pour la suite [4].

Définition

L'opérateur proximal $\mathbf{prox}_f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ de la fonction f est défini par

$$\mathbf{prox}_f(v) = \arg \min_x (f(x) + \frac{1}{2} \|x - v\|_2^2) \quad (19)$$

où $\|\cdot\|_2$ est la norme Euclidienne usuelle.

Il existe aussi l'opérateur proximal de la fonction mise à l'échelle λf , où $\lambda > 0$, exprimé comme

$$\mathbf{prox}_{\lambda f}(v) = \arg \min_x (f(x) + \frac{1}{2\lambda} \|x - v\|_2^2) \quad (20)$$

L'opérateur proximal peut être interprété comme un point qui fait le compromis entre minimiser f tout en étant proche de v . Dans l'opérateur proximal de la fonction mise à l'échelle, le paramètre λ peut être interprété comme un poids relatif ou un paramètre qui gère le compromis entre les deux termes.

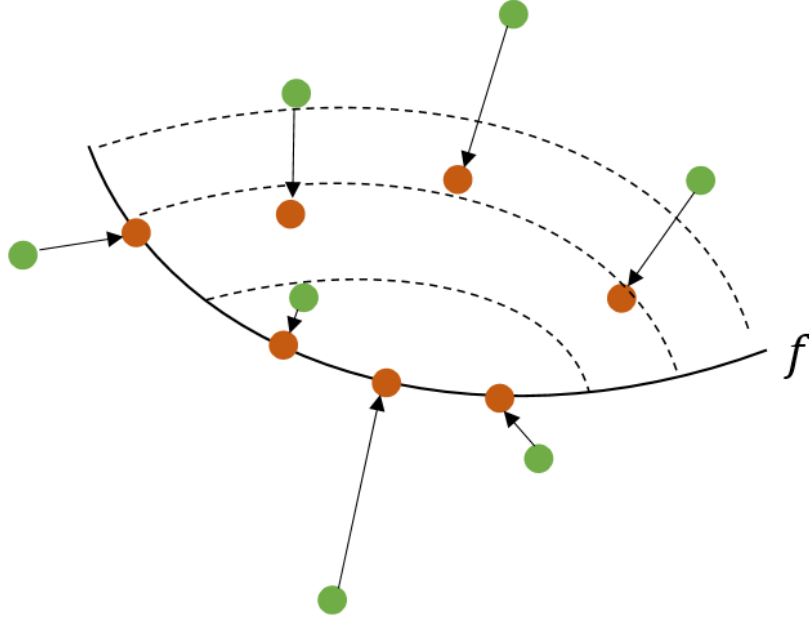


FIGURE 3 – Opérateur proximal[4]

La figure 3 illustre bien ce que fait l'opérateur proximal. Les lignes pointillées représentent les courbes de niveau de la fonction f et la ligne épaisse représente les limites du domaine de la fonction. Evaluer $\text{prox}_f(v)$ revient à faire bouger les points verts jusqu'aux positions représentées par les points oranges. Les trois points en dehors du domaine de f se déplacent donc vers la limite du domaine et également vers le minimum de la fonction. Les points situés à l'intérieur du domaine se déplacent aussi vers le minimum de la fonction. Augmenter le paramètre λ aura pour effet de favoriser le déplacement des points vers le minimum de la fonction, et à l'inverse diminuer ce paramètre aura tendance à favoriser les points à rester proches de leur position initiale.

Une autre interprétation suggère un lien avec la descente de gradient de la fonction f . En effet, on peut exprimer l'opérateur proximal comme

$$\text{prox}_f(v) \approx v - \Delta f(v) \quad (21)$$

quand λ est petit et f dérivable. Cette connexion entre l'opérateur proximal et la descente de gradient peut nous être utile dans notre cas étant donné que l'on veut réaliser la minimisation d'une fonction.

Enfin, les points qui minimisent la fonction f sont en réalité les points qui restent fixes étant donné qu'ils sont déjà positionnés dans le domaine de f et sur son minimum. En d'autres mots

$$\text{prox}_f(z) = z \quad (22)$$

si et seulement si z minimise f .

Lien entre la fonction indicatrice et la projection

Quand f est la fonction indicatrice

$$\iota_{\mathcal{C}}(x) = \begin{cases} 0 & x \in \mathcal{C} \\ +\infty & x \notin \mathcal{C} \end{cases} \quad (23)$$

l'opérateur proximal de f se réduit à la projection Euclidienne sur $\mathcal{C}$, que l'on note

$$\Pi_{\mathcal{C}}(v) = \arg \min_{x \in \mathcal{C}} \|x - v\|_2 \quad (24)$$

Décomposition de Moreau

Si la fonction f est la norme générale $\|\cdot\|$ sur $\mathbb{R}^N$, sa fonction duale f^* est la fonction indicatrice $\iota_{\mathcal{B}}$ où $\mathcal{B}$ est la boule unitaire pour la norme duale. Par la décomposition de Moreau on a

$$\begin{aligned} \text{prox}_{\lambda f}(v) &= v - \lambda \text{prox}_{f^*/\lambda}(v/\lambda) \\ &= v - \lambda \Pi_{\mathcal{B}}(v/\lambda) \end{aligned} \quad (25)$$

2.5 Algorithme de Condat : primal-dual

L'algorithme qui va nous permettre de résoudre 17 est l'approche primale-duale de Condat présentée dans [7].

Le problème général que l'on veut résoudre dans sa formulation primale est :

$$\text{Trouver } \hat{x} \in \arg \min_{x \in \mathcal{X}} \left[F(x) + G(x) + \sum_{i=1}^m H_i(L_i x) \right], \quad (26)$$

où

- $F : \mathcal{H} \rightarrow \mathbb{R}$ est convexe, dérivable sur $\mathcal{H}$ et son gradient ∇F est β -Lipschitz continu, pour un $\beta \in [0, +\infty[$; à savoir,

$$\|\nabla F(x) - \nabla F(x')\| \leq \|x - x'\|, \quad \forall (x, x') \in \mathcal{H}^2. \quad (27)$$

On peut noter que le cas $\beta = 0$ correspond à ∇F étant constant, ce qui est le cas par exemple quand le terme F doit être ignoré dans le problème (Donc $F = 0$).

- $G \in \Gamma_0(\mathcal{H})$ et $H \in \Gamma_0(\mathcal{Y})$ sont "simples", dans le sens où leur opérateur proximal a une représentation en forme close, où peuvent au moins être résolus efficacement avec haute précision.

— $L : \mathcal{H} \rightarrow \mathcal{Y}$ est un opérateur linéaire "fermé" avec un adjoint L^* et une norme induite

$$\|L\| = \sup\{\|Lx\| : x \in \mathcal{H}, \|x\| \leq 1\} < +\infty \quad (28)$$

Précisons que $\mathcal{H}$ et $\mathcal{Y}$ sont des espaces de Hilbert.

On obtient alors cet algorithme (algorithme 5.1 dans le papier de Condat) :

Algorithm 1 Algorithme primal-dual forward-backward pour résoudre 26

Choisir les paramètres proximaux $\tau > 0, \sigma > 0$, la séquence de paramètres positifs d'assouplissement $(\rho_n)_{n \in \mathbb{N}}$ et le schéma initial $(x_0, y_{1,0}, \dots, y_{m,0}) \in \mathcal{X} \times \mathcal{Y}_1 \times \dots \times \mathcal{Y}_m$. Ensuite itérer, pour tout $n \geq 0$,

- 1: $\tilde{x}_{n+1} := \mathbf{prox}_{\tau G}(x_n - \tau(\nabla F(x_n) + e_{F,n}) - \tau \sum_{i=1}^m L_i^* y_{i,n}) + e_{G,n}$
 - 2: $x_{n+1} := \rho_n \tilde{x}_{n+1} + (1 - \rho_n) x_n$
 - 3: $\forall i = 1, \dots, m, \tilde{y}_{i,n+1} := \mathbf{prox}_{\sigma H_i^*}(y_{i,n} + \sigma L_i(2\tilde{x}_{n+1} - x_n)) + e_{H_i,n}$
 - 4: $\forall i = 1, \dots, m, y_{i,n+1} := \rho_n \tilde{y}_{i,n+1} + (1 - \rho_n) y_{i,n}$
-

Les paramètres σ et τ doivent respecter $\frac{1}{\tau} - \sigma \|\sum_{i=1}^m L_i^* L_i\| \geq \frac{\beta}{2}$. Si $F = 0$, la relation devient $\sigma\tau \|\sum_{i=1}^m L_i^* L_i\| < 1$.

2.5.1 Application sur l'estimateur MAP

En reprenant 17 et l'algorithme général 1 on peut maintenant définir l'algorithme qui va résoudre notre problème inverse. Identifions les différents paramètres :

- F : Il n'y a pas de fonction dérivable ici, donc $F=0$.
- G correspond à la fonction indicatrice $\iota_{[0,1]}$. On se rappelle que l'opérateur proximal d'une fonction indicatrice est une projection sur l'ensemble donné. On a donc

$$\mathbf{prox}_G(x) = \Pi_{[0,1]}(x)$$

- Ligne 3 de l'algorithme nous devons calculer : $\mathbf{prox}_{\sigma H_i^*}(y_{i,n} + \sigma L_i(2\tilde{x}_{n+1} - x_n)) + e_{H_i,n}$. La première fonction H correspond à la fonction indicatrice $\iota_{\mathcal{B}_2(y,\epsilon)}(\Phi x)$. On a besoin ici de H^* , mais l'identité de Moreau nous permet d'éviter de calculer ce terme. En effet on a que $\mathbf{prox}_{\sigma H^*}(x) = x - \sigma \mathbf{prox}_{\frac{H}{\sigma}}(\frac{x}{\sigma})$. Enfin en utilisant la propriété de l'opérateur proximal d'une fonction indicatrice on a

$$\mathbf{prox}_{\gamma H^*}(x) = x - \gamma \Pi_{\mathcal{B}_2(y,\epsilon)}(\frac{x}{\gamma})$$

sans oublier l'opérateur L qui est dans notre cas Φ . Le résultat final se trouve aux lignes 6 et 7 de 2.

- Enfin, il reste le terme $\lambda f(\Psi x)$. Par les mêmes arguments on a donc les lignes 4 et 5 de 2.

L'algorithme complet est le suivant :

Algorithm 2 Algorithmme primal-dual forward-backward pour résoudre 17

```
1: Initialisation : Soit  $x^{(0)} \in \mathbb{R}^N$ ,  $y_1^{(0)} \in \mathbb{R}^N$ , et  $y_2^{(0)} \in \mathbb{R}^N$ . Soit  $\sigma > 0$  et  $\gamma > 0$  tels que  
    $\sigma\gamma(\|\Psi\Psi^*\| + \|\Phi\Phi^*\|) < 1$ .  
2: for  $i = 0, 1, \dots$  do  
3:    $x^{(i+1)} = \Pi_{[0,1]}(x^{(i)} - \sigma(\Psi^\dagger y_1^{(i)} + \Phi^\dagger y_2^{(i)}))$   
4:    $\tilde{y}_1^{(i)} = y_1^{(i)} + \gamma\Psi(2x^{(i+1)} - x^{(i)})$   
5:    $y_1^{(i+1)} = \tilde{y}_1^{(i)} - \gamma\mathbf{prox}_{\frac{\lambda}{\gamma}\|\cdot\|_1}(\gamma^{-1}\tilde{y}_1^{(i)})$   
6:    $\tilde{y}_2^{(i)} = y_2^{(i)} + \gamma\Phi(2x^{(i+1)} - x^{(i)})$   
7:    $y_2^{(i+1)} = \tilde{y}_2^{(i)} - \gamma\Pi_{\mathcal{B}_2(y, \epsilon)}(\gamma^{-1}\tilde{y}_2^{(i)})$   
8: end for
```

2.6 Convolution

L'opérateur Φ cité plus haut peut correspondre à une convolution. Il mérite qu'on s'y attarde. Deux convolutions vont être explorées, une Gaussienne et une LoG (Laplacien d'une Gaussienne). Pour les moins avertis, la fonction gaussienne est définie en toute généralité par

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (29)$$

Pour la suite nous allons utiliser une forme exprimée en 2D et simplifiée étant donné que le noyau sera normalisé et appliqué à une image elle-même en 2D. On a donc

$$G(x_1, x_2) = e^{-\frac{x_1^2 + x_2^2}{2\sigma^2}}. \quad (30)$$

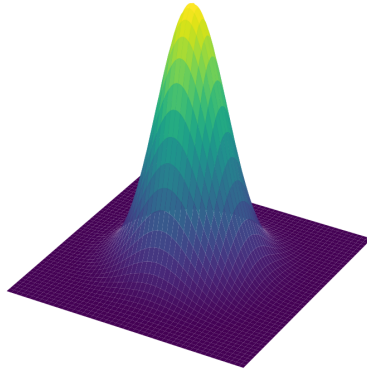


FIGURE 4 – Gaussienne

Le Laplacien se calcule comme suit

$$L(x, y) = \nabla^2 G(x_1, x_2) = \frac{\delta^2 G(x_1, x_2)}{\delta x_1^2} + \frac{\delta^2 G(x_1, x_2)}{\delta x_2^2}$$

et finalement en l'appliquant à la fonction Gaussienne nous obtenons le *LoG*

$$LoG(x_1, x_2) = -\frac{1}{\pi\sigma^4} \left(1 - \frac{x_1^2 + x_2^2}{2\sigma^2}\right) e^{-\frac{x_1^2 + x_2^2}{2\sigma^2}}.$$

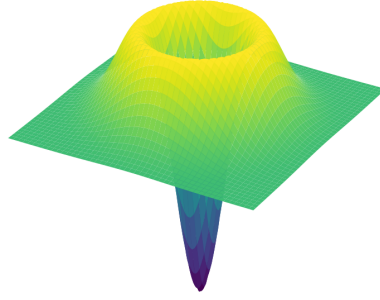


FIGURE 5 – LoG

Si on prend une image $f(n, m)$ et un filtre quelconque $h(n, m)$, la convolution de f par h est

$$(f * h)(n, m) = \sum_{n'=1}^N \sum_{m'=1}^M f(n', m') h(n' - n, m' - m)$$

Si on se rappelle du théorème de convolution [21],

$$f * h = \mathcal{F}^{-1}(F \cdot H)$$

avec $\mathcal{F}^{-1}$ l'opérateur de transformée de Fourier inverse et F et H respectivement l'image et le filtre exprimés dans le domaine de Fourier. On en conclut qu'une convolution dans le domaine spatial correspond à une multiplication dans le domaine de Fourier. Cette information nous sera utile pour réaliser la convolution d'une manière efficace. En effet la convolution réalisée en passant par le domaine de Fourier requiert bien moins de ressources calculatoires qu'une convolution dans le domaine spatial.

3 Sources d'incertitude et méthode classique de quantification

La solution obtenue via l'estimateur MAP $x^\dagger$ est une image unique résumant $x|y$ qui peut être facilement affichée et analysée. Cependant cette image unique ne permet pas de quantifier l'incertitude qui est inhérente aux problèmes mal posés et mal conditionnés. Quantifier l'incertitude est un problème important dans certains domaines du traitement d'images. Prenons comme exemple une image d'un IRM, il est important de savoir si une tache que l'on voit sur la solution affichable $x^\dagger$ est une tache réelle ou un artefact de reconstruction. En effet, dans le cas de cet IRM, le médecin doit être sûr de ce qu'il voit pour poser un diagnostic précis. La tache est elle une tumeur ? L'enjeu est important.

3.1 Quantifier l'incertitude et MCMC

La solution MAP est une façon parmi tant d'autres de caractériser la distribution de probabilité. On peut également vouloir connaître sa valeur moyenne (moyenne arithmétique), sa déviation standard ou des intervalles de confiance. La question est de savoir comment déterminer ce genre de valeur. C'est impossible par calcul conventionnel, cependant il existe une méthode qui permet de déduire ces valeurs sans connaître toutes les propriétés mathématiques de la distribution de probabilité de notre modèle, la Méthode de Monte-Carlo par chaînes de Markov ou MCMC très bien résumée dans [22]. Ceci se fait en échantillonnant au hasard des valeurs de la distribution. Cela repose sur deux principes distincts, la méthode de Monte-Carlo d'une part et les chaînes de Markov d'autre part. Le principe de Monte-carlo est de calculer les valeurs citées plus haut (moyenne, déviation standard entre autres) non pas via l'équation de la distribution mais en générant un grand nombre d'échantillons aléatoires à partir de cette distribution et d'ensuite à partir de ceux-ci obtenir les valeurs voulues. Il est en effet souvent plus facile de générer des échantillons aléatoires et de calculer les valeurs voulues que de les calculer via les équations de probabilité. On parle de chaîne de Markov parce que tous les échantillons forment une chaîne, chaque nouvel échantillon étant généré à partir de l'échantillon précédent et uniquement à partir de celui-ci.

La méthode MCMC est donc particulièrement utile dans notre cas étant donné l'expression de la densité de probabilité étudiée qui est difficilement traitable analytiquement. Rappelons ici que la distribution de probabilité postérieure définie comme

$$p_{post}(x) = p(x|y) = \frac{p_{prior}(x)p(y|x)}{p(y)}. \quad (31)$$

est de forme log-concave, soit

$$p(x|y) \propto e^{-g_1(x,y)-g_2(x)}. \quad (32)$$

Les méthodes traditionnelles utilisant MCMC pour l'analyse bayésienne sont réputées très coûteuses en terme de calcul. Il existe cependant une approche reposant sur l'optimisation proximale. Ces méthodes sont plus faciles à mettre en place et sont bien moins coûteuses en termes de calcul. Elles reposent sur l'utilisation de l'équation différentielle stochastique de Langevin suramortie définie par

$$dX_t = \nabla \log p_{post}^-(X_t) dt + \sqrt{2} dW_t$$

avec W un processus de Wiener. Cette équation converge vers la distribution p_{post} lorsque t tend vers l'infini et est donc utile pour générer des échantillons de Monte Carlo à partir de p_{post} . c'est encore plus vrai lorsque la distribution est de forme log-convexe comme dans notre cas. Cependant la forme continue n'est pas pratique et difficile à résoudre donc on introduit sa forme discrète, la méthode nommée ULA (pour algorithme de Langevin non ajusté)

$$X_{n+1} = X_n + \delta \nabla \log p_{post}^-(X_n) + \sqrt{2\delta} W_{n+1}$$

avec δ un pas donné et $W_{n+1} \sim \mathcal{N}(0, 1_N)$ une séquence de variables aléatoires gaussiennes. On reconnaît dans cette expression les caractéristiques d'une méthode MCMC, les échantillons forment une chaîne (séquence X) et chaque échantillon X_{n+1} est évalué à l'aide de celui qui le précède (X_n) et uniquement celui-ci. Une des limitations d'ULA (et de son homologue MULA) est que la fonction $\log p_{post}$ doit être continûment dérivable sinon la chaîne de Markov ne converge pas. Étant donné que dans de nombreux cas et en particulier le nôtre ce n'est pas le cas, il faut passer par d'autres méthodes. Citons entre autres la Moreau-Yosida regularised ULA (MYULA), Proximal MALA (Px-MALA) ou le schéma SK-ROCK développés dans [6] et [23]. Rappelons que ces méthodes nous permettent de générer des échantillons aléatoires provenant de la distribution de probabilité de notre problème.

La quantification de l'incertitude requiert d'analyser des statistiques propres à la distribution $p(x|y)$. Les échantillons générés $x^{(j)}$ nous donnent justement accès à cette distribution. Comme on peut le voir sur la figure 6, la première étape de la procédure présentée dans [6] est de générer la densité postérieure complète de l'image et ensuite d'utiliser plusieurs formes de quantification d'incertitude.

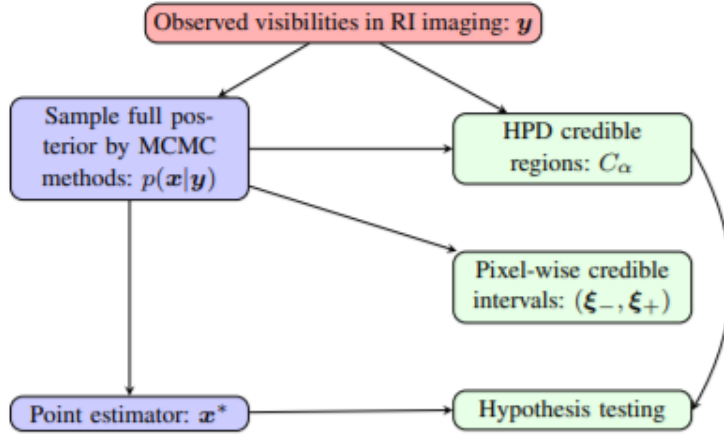


FIGURE 6 – Procédure présentée dans [6] pour la quantification d’incertitude en partant d’une méthode MCMC

La première chose que l’on peut faire est de déterminer des intervalles de confiance à l’échelle des pixels que l’on note I_i pour la borne inférieure de l’intervalle et S_i pour la borne supérieure. Ces intervalles indiquent la plage de valeurs que peuvent prendre ces pixels avec probabilité $(1 - \alpha)$, c’est-à-dire $p(x_i \in [I_i, S_i] | y) = 1 - \alpha$ pour $i \in [1, N]$. Pour obtenir ces intervalles, on calcule simplement

$$(I_i, S_i) = \text{quantile}(x_i^{(j)}, \left\{ \frac{\alpha}{2}, 1 - \frac{\alpha}{2} \right\}). \quad (33)$$

Ces intervalles sont utiles pour identifier des zones de haute incertitude, et peuvent facilement être visualisés en calculant la largeur de l’intervalle $S_i - I_i$.

Les intervalles de confiance évalués à l’échelle des pixels sont utiles pour analyser localement l’incertitude, mais pour une analyse plus approfondie il faut introduire une autre méthode de quantification. Celle-ci repose sur l’évaluation de régions de densités de probabilité postérieure avec niveau de confiance $(1 - \alpha)$ notées C_α . Il existe une région particulière qui est optimale et nommée HPD, elle sera développée plus en détail dans le prochain chapitre. Cette région est utilisée pour analyser le niveau de confiance concernant la présence de structures observées après la reconstruction de l’image.

L’étape finale de la procédure est de réaliser un test d’hypothèse à propos d’une structure observée dans l’estimateur x^* , en enlevant la structure de l’image et en vérifiant si la nouvelle image sans la structure fait toujours partie de la région de crédibilité C_α . Si la nouvelle image est acceptée dans C_α alors les données en présence nous permettent de conclure que la structure n’est pas présente dans la vraie image, ceci avec haute probabilité.

3.2 Sources supplémentaires d'incertitudes

La borne du bruit est définie comme $\epsilon = \sigma \sqrt{N + 2\sqrt{2N}}$. Ce résultat vient du fait que pour un bruit gaussien (celui qui sera utilisé) $\|b\|$ suit une distribution χ^2 avec N degrés de liberté. Quand N est grand, χ^2 se concentre autour de sa moyenne avec haute probabilité ce qui mène au résultat donné. Dans cette expression la dimension N est connue mais il n'en est pas de même pour σ . Précédemment nous avons utilisé le σ connu et utilisé dans le problème direct pour générer l'image bruitée, mais en pratique nous n'avons pas accès à cette information. Tout ce dont on a accès est l'image entachée de bruit. Pour déterminer le niveau de bruit on peut utiliser une information déjà évoquée plus tôt, le bruit est surreprésenté à hautes fréquences comparativement au signal. Cette méthode appelée *Écart absolu médian* définie dans [11] et [12] peut être utilisée en passant par la transformée en ondelettes.

1. Prendre la transformée en ondelettes de l'image.
2. A partir de la résolution la plus fine ($j_0 = -[\log_2(N) - 1]$), calculer l'écart médian absolu de la bande des coefficients ($D_k^{j_0}$)

$$\sigma_{estimé} = \frac{\text{Médiane}(|D_k^{j_0}|)}{0.6745}$$

Le choix de la base d'ondelettes à l'étape 1 est important étant donné qu'elle détermine le niveau de compression de l'énergie dans le domaine considéré. Il faut choisir une base qui compresse le signal en très peu de coefficients de grande valeur et en une grande majorité de coefficients de petite valeur. Dans notre cas, le bruit est représenté par les petits coefficients qui sont présents en grand nombre. L'étape 2 permet donc de calculer la déviation standard du bruit σ en assumant que la majorité des coefficients de la bande des détails sont des coefficients qui sont issus des composants du bruit.

Cette estimation, bien que relativement proche de la vraie valeur de σ , est en elle-même une source d'incertitude, ou tout du moins source de dégradation du résultat du débruitage. En effet les figures 7 et 8 montrent une claire différence entre l'estimation et la valeur exacte. Cette estimation a tendance à être surévaluée mais l'erreur tend à diminuer quand σ augmente. Prenons $\sigma_{exact} = 0.05$, l'erreur d'estimation est de 10 % ce qui est tout de même significatif. La solution du débruitage sera affectée négativement par cette mauvaise estimation.

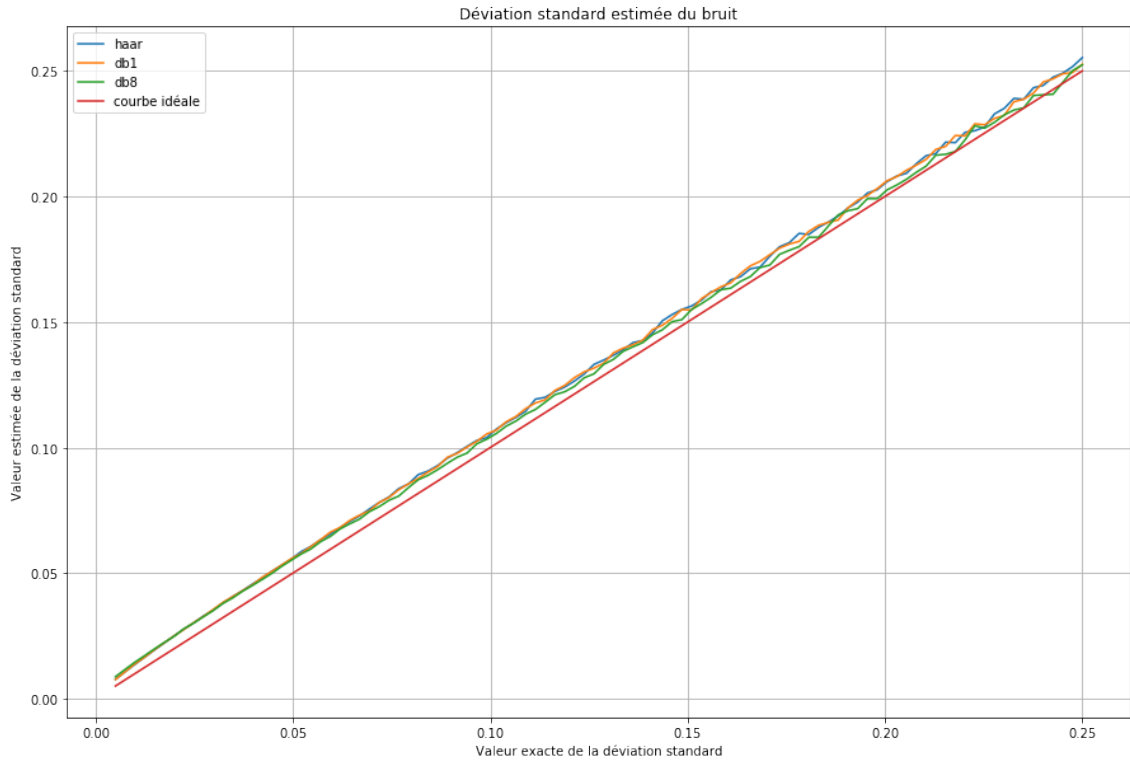


FIGURE 7 – Déviation standard estimée du bruit pour différentes bases d'ondelettes

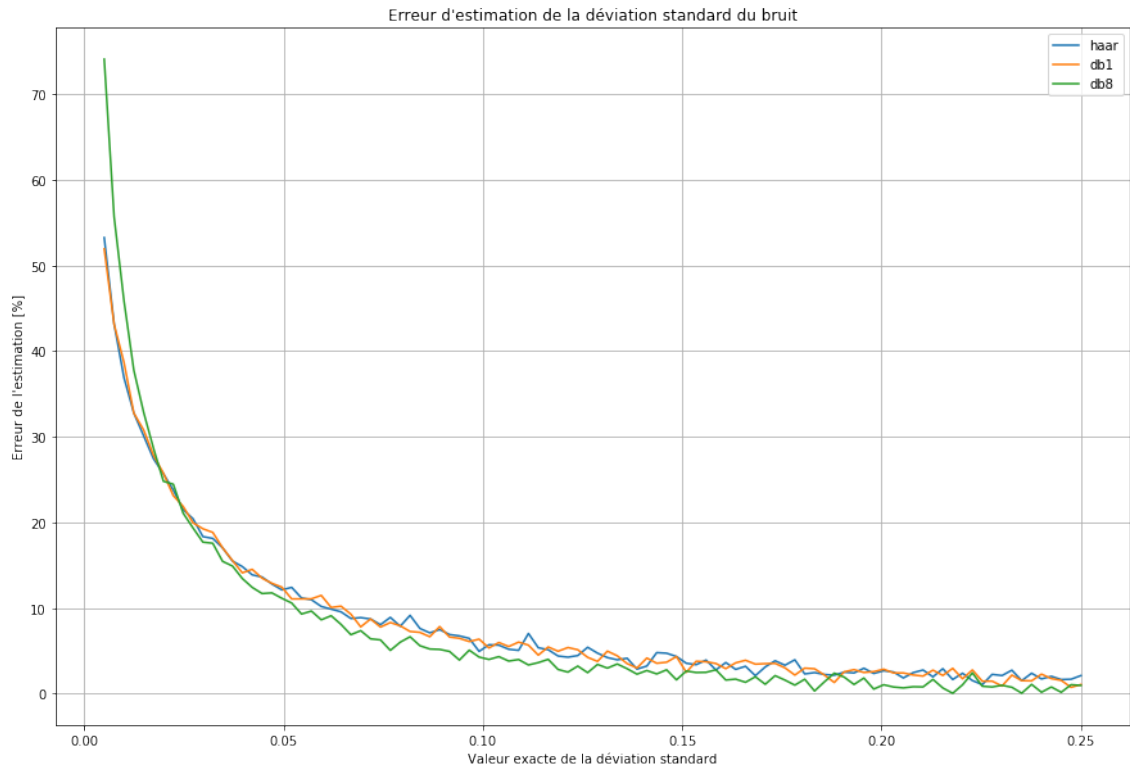


FIGURE 8 – Erreur d'estimation de la déviation standard du bruit pour différentes bases d'ondelettes

La figure 9 est encore plus explicite. La meilleure solution serait d'utiliser ce qu'on appelle la "ground truth", c'est à dire estimer le niveau de bruit en connaissant l'image x originale et l'image bruitée y . La seconde option est d'utiliser la valeur exacte de σ pour évaluer une estimation ϵ un peu plus large que la valeur obtenue via la "ground truth". La dernière option est celle présentée ici, c'est-à-dire d'utiliser l'estimation ϵ à partir d'une autre estimation σ . Étant donné notre *à priori* donné plus haut, on en déduit que plus ϵ est grand, plus l'espace des solutions de $\|\Phi x - y\| \leq \epsilon$ est grand. Autrement dit, plus grand espace de solution signifie plus d'incertitude.

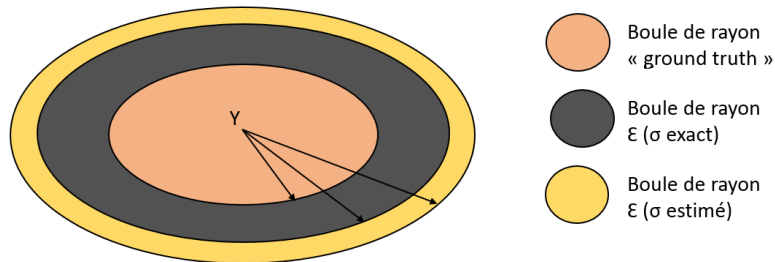


FIGURE 9 – Différentes boules de rayon ϵ centrées en y (image bruitée)

Une autre source d'incertitude peut provenir de l'opérateur ϕ . Plus haut nous avons utilisé l'opérateur ϕ qui nous a permis de générer le problème direct. Dans la réalité, on ne dispose pas de cet opérateur. Il faut l'estimer via une calibration par exemple. Une mauvaise estimation de cet opérateur peut elle aussi dégrader la solution du problème inverse et donc engendrer de l'incertitude.

4 Méthode dite BUQO

Comme on l’a vu au point précédent, la méthode de Monte-Carlo est utile dans le cadre de la quantification d’incertitude mais est bien plus coûteuse en termes de calculs que la méthode d’estimation ponctuelle MAP. De plus, est-il nécessaire de quantifier l’incertitude sur l’ensemble d’une image ? Ce qui pose question ce sont surtout les structures présentes dans $x^\dagger$ (l’estimation MAP). Et pas toutes les structures, plutôt certaines structures à propos desquelles des doutes sont présents. Un choix judicieux de méthode serait d’analyser seulement certaines structures qui posent question, et de profiter du caractère convexe du problème (déjà abordé dans la résolution de l’estimation MAP) qui permet de le résoudre en utilisant des algorithmes d’optimisation performants. C’est l’idée proposée dans [] avec la méthode dite BUQO (Bayesian uncertainty quantification by optimization) présentée dans [18] et [17]. Elle consiste en la quantification d’incertitude d’une structure en réalisant un test d’hypothèse bayésien, ceci en deux étapes. D’abord considérer que la structure n’est pas présente dans l’image réelle et ensuite utiliser nos informations à priori pour déterminer si l’hypothèse nulle est à rejeter avec haute probabilité. Avant d’entrer dans les détails de cette méthode il convient d’introduire un concept qui va nous permettre de mener à bien la méthode BUQO, la dite *Région postérieure de haute densité de crédibilité* (Highly Posterior Density credibility region or *HPD*).

4.1 Région HPD

La solution du problème MAP ne nous apportant aucune information sur l’incertitude il serait judicieux de déterminer un espace où la majeure partie de la masse de probabilité à posteriori de $x|y$ se trouve. On définit un ensemble C_α comme une région postérieure de crédibilité de degré de confiance $(1-\alpha)$ si

$$P(x \in C_\alpha | y) = \int p(x|y) 1_{C_\alpha}(x) dx = 1 - \alpha.$$

Il existe une région particulière qui est optimale (volume minimal) et pratique à utiliser, la *Région postérieure de haute densité de crédibilité* (HPD). Elle est définie comme

$$C_\alpha^* = \{x : g_y(x) \leq \gamma_\alpha\}$$

avec $\gamma_\alpha \in \mathbb{R}$ choisi tel que $\int_{C_\alpha^*} p(x|y) dx = 1 - \alpha$ soit valable.

Le problème rencontré est qu’il est difficile de calculer des ensembles de crédibilité. Cela requiert de résoudre des intégrales. Ces intégrales peuvent être approchées à l’aide de méthodes de Monte-Carlo mais le coût de calcul est important, coût qui est bien plus important que celui du calcul de l’estimation $x_{MAP}^\dagger$ (et qui augmente fortement en rapport avec la dimension du problème). Il serait justement intéressant de pouvoir profiter de l’estimation déjà calculée $x_{MAP}^\dagger$ pour définir l’ensemble de crédibilité voulu. Pour rappel notre distribution postérieure de x connaissant y , en considérant que $p(x|y)$ est log-concave,

est donnée en toute généralité par

$$p(x|y) = \exp\{-g_y(x)\}/Z$$

avec $g_y(x)$ une fonction convexe. L'estimation MAP est alors définie par

$$\begin{aligned}\hat{x}_{MAP} &= \arg \max p(x|y) \\ &= \arg \min g_y(x).\end{aligned}$$

La figure 10 résume la situation. L'estimation MAP vise à minimiser la fonction $g_y(x)$. Un seuil γ_α est défini, la région de crédibilité (zone verte) correspond à toute solution de $g_y(x)$ se trouvant en dessous de ce seuil.

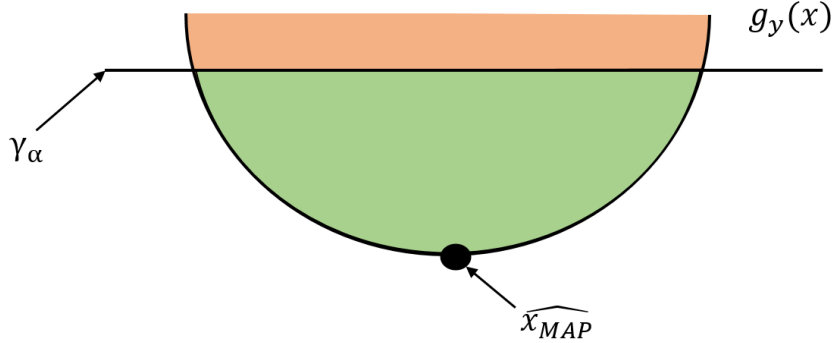


FIGURE 10 – Solution de l'estimation du MAP et illustration de la région de crédibilité HPD (zone verte)

Il est en réalité possible de profiter de la connaissance de $x_{MAP}^\dagger$ pour établir une nouvelle région $\tilde{C}_\alpha$ dans laquelle est contenue la région C_α^* définie plus haut. Cette approche est introduite dans [16]. Cette nouvelle région est définie pour tout $\alpha \in]4\exp -N/3, 1[$ comme

$$\tilde{C}_\alpha = \{x : g_y(x) \leq g_y(\hat{x}_{MAP}) + N(\tau_\alpha + 1)\}$$

avec $\tau_\alpha = \sqrt{16\log(\frac{3}{\alpha})/N}$ une constante positive indépendante de $p(x|y)$. On définit donc le nouveau seuil $\tilde{\gamma}_\alpha = g_y(\hat{x}_{MAP}) + N(\tau_\alpha + 1)$ et on a que

$$0 \leq \frac{\tilde{\gamma}_\alpha - \gamma_\alpha}{N} \leq 1 + \eta_\alpha \sqrt{\frac{1}{d}}$$

avec $\eta_\alpha = \sqrt{16\log(\frac{3}{\alpha})/N} + \sqrt{\frac{1}{\alpha}}$ une constante positive. Si la dimension n du problème augmente l'erreur tend vers $0 \leq \frac{\tilde{\gamma}_\alpha - \gamma_\alpha}{N} \leq 1$ (d'autant plus vrai lorsque $N > 10^3$).

L'ensemble $\tilde{C}_\alpha$ est une région bayésienne conservatrice, c'est à dire qu'elle vérifie

$$P(x \in \tilde{C}_\alpha | y) \geq 1 - \alpha.$$

En d'autres mots, le seuil $\tilde{\gamma}_\alpha$ est supérieur ou égal au seuil γ_α , et donc $\tilde{C}_\alpha$ englobe C_α^*

(Figure 11) tout en étant sa plus proche approximation possible obtenue en connaissant l'estimation MAP.

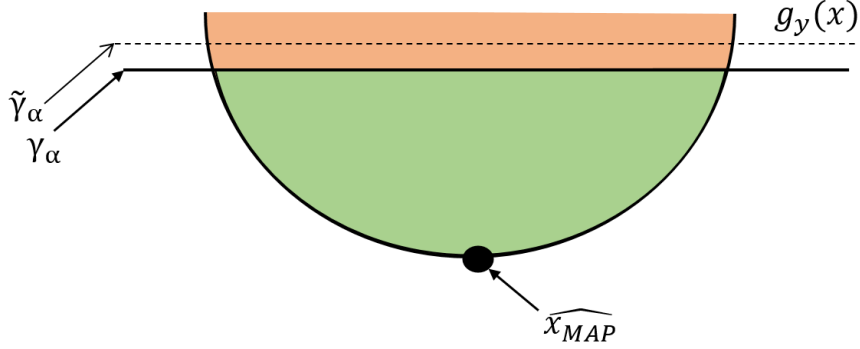


FIGURE 11 – Nouveau seuil $\tilde{\gamma}_\alpha$

Dans le cas de notre modèle, la région de crédibilité est donnée par

$$\tilde{\mathcal{C}}_\alpha = \{x \in \mathcal{C} \mid \Phi x \in \mathcal{B}_2(y, \epsilon) \text{ et } \lambda f(\Psi x) \leq \tilde{\eta}_\alpha\} \quad (34)$$

où $\tilde{\eta}_\alpha = \lambda f(\Psi x_{MAP}^\dagger) + N(\tau_\alpha + 1)$. La projection sur cet ensemble est donnée par

$$\Pi_{\tilde{\mathcal{C}}_\alpha}(x^{(k)}) = \arg \min_{x \in \mathbb{R}^N} F(x, x^{(k)}) \quad (35)$$

avec pour tout $(x, x^{(k)}) \in \mathbb{R}^N \times \mathbb{R}^N$

$$F(x, x^{(k)}) = \iota_{\mathcal{B}_2(y, \epsilon)}(\Phi x) + \iota_{lev_{\leq \tilde{\eta}_\alpha/\lambda}(f)}(\Psi x) + \iota_{[0,1]^N}(x) + \frac{1}{2} \|x - x^{(k)}\|^2. \quad (36)$$

La fonction 36 est similaire à la fonction 17 qui permet de trouver l'estimation MAP à deux termes près. Premièrement une attache aux données $\frac{1}{2} \|x - x^{(k)}\|^2$ apparaît et le terme $\iota_{lev_{\leq \tilde{\eta}_\alpha/\lambda}(f)}(\Psi x)$ diffère de 17. f représente la norme l_1 et $lev_{\leq \beta}(f)$ est défini comme $lev_{\leq \beta}(f) = \{u \in \mathcal{R}^N \mid f(u) \leq \beta\}$.

Par le même raisonnement établi pour générer l'algorithme de l'estimation MAP, on peut déterminer l'algorithme pour effectuer la projection sur $\tilde{\mathcal{C}}_\alpha$. Il est exprimé comme

Algorithm 3 Algorithme primal-dual forward-backward pour résoudre 35-36

- 1: **Initialisation** : Soit $u^{(0)} \in \mathbb{R}^N$, $v_1^{(0)} \in \mathbb{R}^N$, et $v_2^{(0)} \in \mathbb{C}^M$. Soit $\sigma > 0$ et $\gamma > 0$ tels que $\sigma(\frac{1}{2} + \gamma(\|\Psi\|^2 + \|\Phi\|^2)) < 1$.
 - 2: **for** $i = 0, 1, \dots$ **do**
 - 3: $u^{(i+1)} = \Pi_{[0,1]^N}(u^{(i)} - \sigma(u^{(i)} - x^{(k)} + \Psi^\dagger v_1^{(i)} + \Phi^\dagger v_2^{(i)}))$
 - 4: $\tilde{v}_1^{(i)} = v_1^{(i)} + \gamma \Psi(2u^{(i+1)} - u^{(i)})$
 - 5: $v_1^{(i+1)} = \tilde{v}_1^{(i)} - \gamma \Pi_{lev_{\leq \tilde{\eta}_\alpha/\lambda}(f)}(\gamma^{-1} \tilde{v}_1^{(i)})$
 - 6: $\tilde{v}_2^{(i)} = v_2^{(i)} + \gamma \Phi(2u^{(i+1)} - u^{(i)})$
 - 7: $v_2^{(i+1)} = \tilde{v}_2^{(i)} - \gamma \Pi_{\mathcal{B}_2(y, \epsilon)}(\gamma^{-1} \tilde{v}_2^{(i)})$
 - 8: **end for**
-

4.2 Test d'hypothèse bayésien

Pour déterminer si oui ou non une structure est bien présente dans l'image réelle, un test d'hypothèse bayésien introduit dans [17] est effectué à partir de deux hypothèses

- H_0 : La structure est **absente** dans la vraie image
- H_1 : La structure est **présente** dans la vraie image.

Le but du test est de déterminer si l'image observée y supporte H_0 ou H_1 c'est à dire si l'image supporte le fait que la structure est bien réelle ou si c'est un artefact de reconstruction. Le test divise donc l'espace de l'image en deux régions, un ensemble $\mathcal{S} \subset \mathbb{R}^N$ associé à l'hypothèse H_0 contenant toutes les images sans la structure d'intérêt, et son complément $\mathbb{R}^N \setminus \mathcal{S}$ associé à l'hypothèse H_1 . La théorie bayésienne nous indique que l'on rejette l'hypothèse H_0 en faveur de H_1 avec un niveau de confiance $\alpha \in]0, 1[$ si

$$P[H_0|y] = P[x \in \mathcal{S}|y] = \int_{\mathcal{S}} p(x|y) dx \leq \alpha, \quad (37)$$

ou par équivalence si le ratio

$$\frac{P[H_1|y]}{P[H_0|y]} = \frac{P[x \in \mathbb{R}^N \setminus \mathcal{S}|y]}{P[x \in \mathcal{S}|y]} \geq \frac{1 - \alpha}{\alpha}. \quad (38)$$

Grâce à la région de crédibilité développée au point précédent, on peut désormais développer un schéma de travail. Ce que l'on veut déterminer c'est si l'ensemble $\mathcal{S}$ contient une intersection avec l'ensemble $\tilde{C}_\alpha$. En d'autres mots on veut déterminer si il existe une image qui se trouve dans la région de crédibilité *ne contenant pas* la structure d'intérêt. Si ceci n'est pas vérifié alors il n'existe pas d'intersection entre les deux ensembles

$$\tilde{C}_\alpha \cap \mathcal{S} = \emptyset$$

et l'hypothèse H_0 est rejetée en faveur de H_1 avec un niveau de confiance α , les conditions 37 et 38 étant respectées. Si α est suffisamment petit on peut en déduire avec forte conviction que la structure est bien présente dans l'image réelle.

A l'inverse, si l'on trouve que $\tilde{C}_\alpha \cap \mathcal{S} \neq \emptyset$ alors cela implique que la structure d'intérêt n'est pas présente dans toutes les images que la densité postérieure $p(x|y)$ suggère comme solution à notre problème inverse. La présence de la structure d'intérêt est donc de faible évidence.

4.3 Ensemble $\mathcal{S}$

Afin de réaliser le test d'hypothèse, il faut introduire une nouvelle image représentant $\mathcal{S}$, $x_{\mathcal{S}}^\dagger \in \mathcal{S}$. Cette image est donc $x^\dagger$ à laquelle on a enlevé la structure d'intérêt, telle que $x^\dagger - x_{\mathcal{S}}^\dagger$ correspond à la structure d'intérêt seule. Pour enlever la structure d'intérêt de l'image $x^\dagger$, on définit un opérateur $\mathcal{M} : \mathbb{R}^N \rightarrow \mathbb{R}^{N_{\mathcal{M}}}$ qui sélectionne la structure d'intérêt et son complément $\mathcal{M}^c : \mathbb{R}^N \rightarrow \mathbb{R}^{N - N_{\mathcal{M}}}$ qui sélectionne toute l'image *sauf* la structure d'intérêt. $\mathcal{M}(x)$ représente donc la région de l'image où la structure d'intérêt est présente.

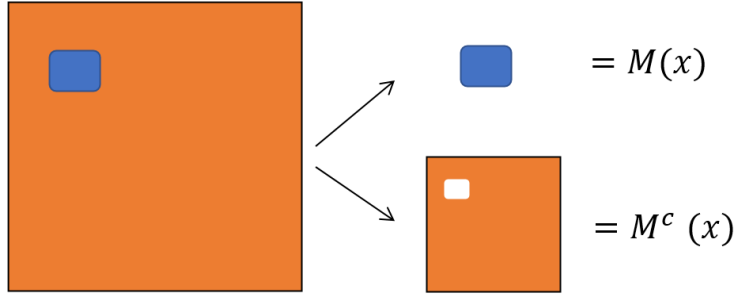


FIGURE 12 – Opérateurs $\mathcal{M}$ et $\mathcal{M}^c$

Il convient également d'introduire l'opérateur $\mathcal{L} : [0, +\infty[^{N-N_{\mathcal{M}}} \rightarrow [0, +\infty[^{N_{\mathcal{M}}}$ qui remplit les pixels $\mathcal{M}(x)$ par les informations contenues dans l'image complémentaire $\mathcal{M}^c(x)$. Pourquoi cet opérateur ? On ne peut pas retirer la structure d'intérêt en produisant $x_{\mathcal{S}}^{\dagger}$ par une simple mise à zéro des pixels à l'endroit de la structure. Cela produirait un trou dans l'image et ne représenterait pas correctement l'allure générale de l'image, il faut au contraire retirer la structure en laissant le *background* intact, similaire à la hauteur des pixels qui se trouvent dans les environs de la structure d'intérêt. Dans cette optique, on définit $x_{\mathcal{S}}^{\dagger}$ telle que $\mathcal{M}(x_{\mathcal{S}}^{\dagger}) = \mathcal{L}(\mathcal{M}^c(x^{\dagger}))$ et $\mathcal{M}^c(x_{\mathcal{S}}^{\dagger}) = \mathcal{M}^c(x^{\dagger})$.

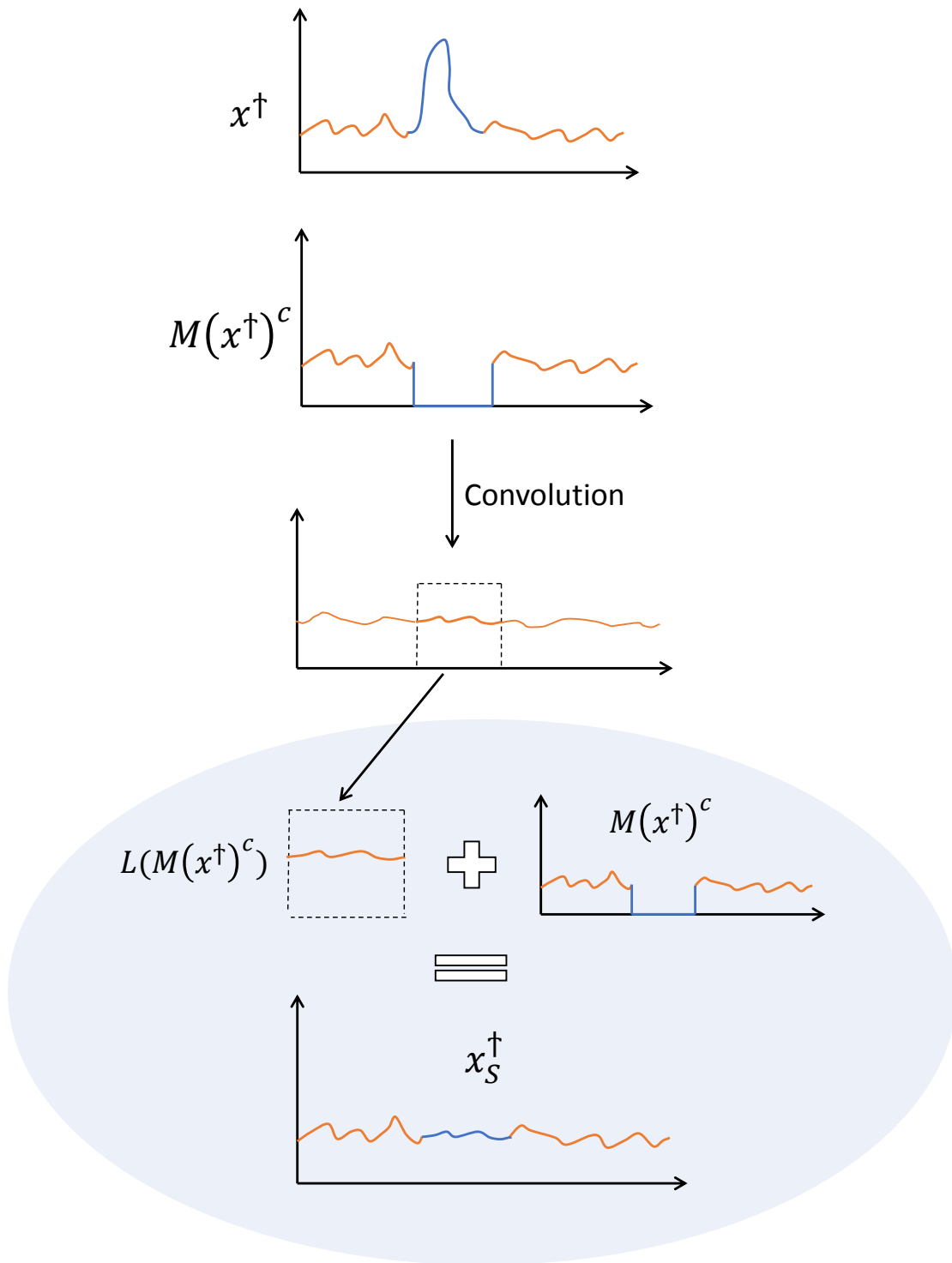


FIGURE 13 – Combinaison des opérateurs $\mathcal{M}, \mathcal{M}^c$ et $\mathcal{L}$ pour produire l'image $x_S^\dagger$

4.4 Intersection entre les ensembles $\tilde{\mathcal{C}}_\alpha$ et $\mathcal{S}$

Pour rappel, on veut vérifier si $P[H_0|y] \leq \alpha$, c'est-à-dire si l'on peut rejeter ou non l'hypothèse H_0 . Il faut donc déterminer si

$$\tilde{\mathcal{C}}_\alpha \cap \mathcal{S} = \emptyset.$$

On veut donc trouver une image que l'on notera $x^\dagger$ où la structure n'est pas présente

$$x^\dagger \in \tilde{\mathcal{C}}_\alpha \cap \mathcal{S}.$$

Pour déterminer si l'intersection entre les deux ensembles est vide ou non, on va alterner les projections sur chacun d'eux. On se référera à l'algorithme de von Neumann pour résoudre ce problème nommé POCS (projection sur des ensembles convexes) qui est utilisé dans [17].

Algorithm 4 Algorithme POCS pour résoudre le problème suivant : trouver $x^\dagger \in \tilde{\mathcal{C}}_\alpha \cap \mathcal{S}$

- 1: **Initialisation** : Soit $x^{(0)} \in \mathcal{S}$
 - 2: **for** $i = 0, 1, \dots$ **do**
 - 3: $x^{(k+\frac{1}{2})} = \Pi_{\tilde{\mathcal{C}}_\alpha}(x^{(k)})$
 - 4: $x^{(k+1)} = \Pi_{\mathcal{S}}(x^{(k+\frac{1}{2})})$
 - 5: **end for**
-

Deux résultats sont possibles, d'une part une intersection existe, et d'autre part l'intersection n'existe pas. $(x^{k+\frac{1}{2}})_{k \in \mathbb{N}}$ et $(x^k)_{k \in \mathbb{N}}$ sont les projections successives générées par l'algorithme.

1. La première possibilité est la convergence des deux projections $(x^{k+\frac{1}{2}})_{k \in \mathbb{N}}$ et $(x^k)_{k \in \mathbb{N}}$ vers un même point $x^\dagger \in \tilde{\mathcal{C}}_\alpha \cap \mathcal{S}$. On a alors que $\tilde{\mathcal{C}}_\alpha \cap \mathcal{S} \neq \emptyset$. Le résultat est visible à la figure ??.
2. La deuxième possibilité est la convergence des projections $(x^{k+\frac{1}{2}})_{k \in \mathbb{N}}$ vers un point $x_{\tilde{\mathcal{C}}_\alpha}^\dagger$ et des projections $(x^k)_{k \in \mathbb{N}}$ vers un point $x_{\mathcal{S}}^\dagger$ et donc $\tilde{\mathcal{C}}_\alpha \cap \mathcal{S} = \emptyset$. Le résultat est visible à la figure 14.

Pour la deuxième possibilité, on va définir une distance entre les deux ensembles. En toute généralité la distance entre les ensembles $\tilde{\mathcal{C}}$ et $\mathcal{S}$ est définie comme

$$\text{dist}(\tilde{\mathcal{C}}_\alpha, \mathcal{S}) = \inf \|\tilde{\mathcal{C}}_\alpha - \mathcal{S}\|,$$

C'est-à-dire que la distance entre les deux ensembles est déterminée par les deux points les plus proches des deux. C'est justement ce que POCS détermine, on a alors

$$\text{dist}(\tilde{\mathcal{C}}_\alpha, \mathcal{S}) = \|x_{\tilde{\mathcal{C}}_\alpha}^\dagger - x_{\mathcal{S}}^\dagger\|.$$

Cette distance va nous permettre d'interpréter le résultat des projections successives sur les deux ensembles, et par conséquent de quantifier l'intensité de la structure d'intérêt.

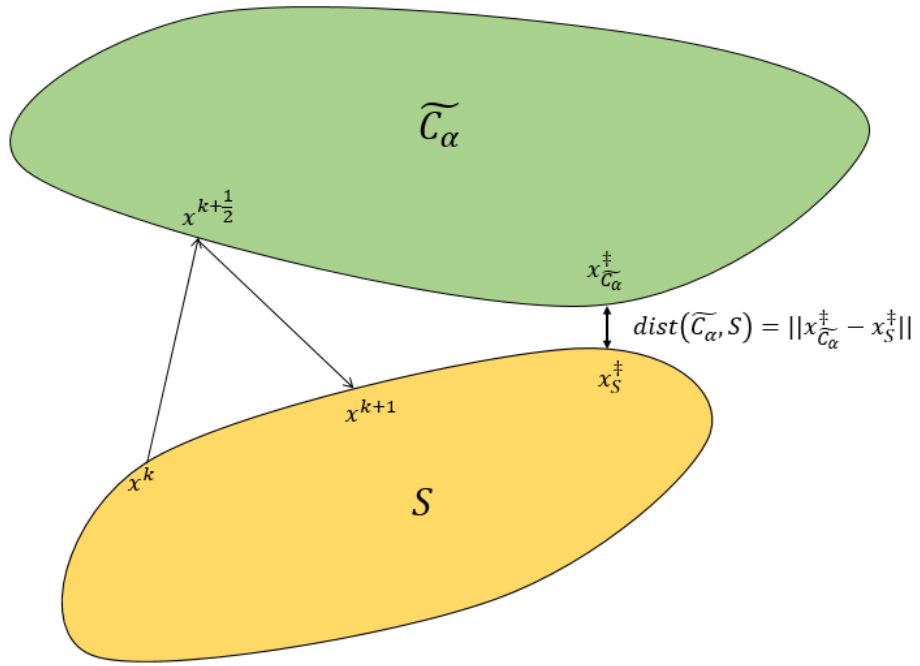


FIGURE 14 – Cas où les projections ne convergent pas vers un même point.

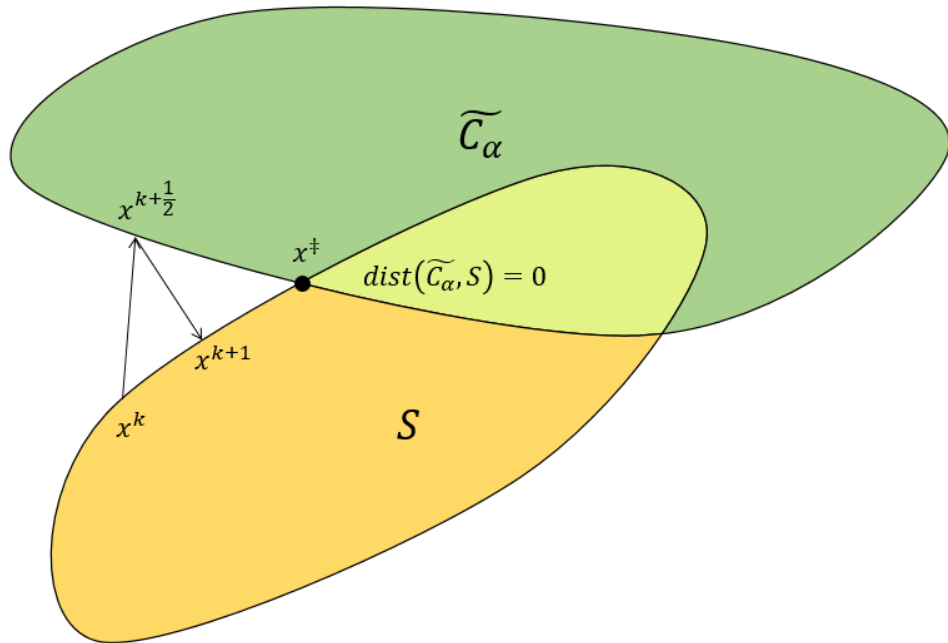


FIGURE 15 – Cas où les projections convergent vers un même point.

4.5 Interprétation de la distance entre les deux ensembles

Le but premier de la méthode mise en place est de déterminer si oui ou non une structure analysée est présente dans l'image réelle. Pour bien comprendre cette notion de distance, il faut d'abord se rappeler que l'algorithme POCS détermine la distance minimale entre les ensembles $\tilde{\mathcal{C}}_\alpha$ et $\mathcal{S}$. Il y a deux possibilités. Soit cette distance est nulle et l'ensemble $\mathcal{S}$ touche ou est contenu partiellement dans $\tilde{\mathcal{C}}_\alpha$. Soit elle est non nulle et les deux ensembles n'ont aucun point en commun. On a cependant besoin de quantifier l'incertitude autour de la présence de la tache. On va donc introduire une nouvelle valeur définie par la distance entre l'estimation $x_{MAP}^\dagger$ et l'image initiale sans la structure $x_S^\dagger$ et notée $\|x_{MAP}^\dagger - x_S^\dagger\|$. Elle représente l'intensité de la structure présente dans l'estimation MAP. Évaluons le rapport entre les deux distances précitées (notion introduite dans [17])

$$\rho_\alpha = \frac{\|x_{\tilde{\mathcal{C}}_\alpha}^\dagger - x_S^\dagger\|}{\|x_{MAP}^\dagger - x_S^\dagger\|} = \frac{\text{dist}(\tilde{\mathcal{C}}_\alpha, \mathcal{S})}{\|x_{MAP}^\dagger - x_S^\dagger\|}. \quad (39)$$

Deux cas sont possibles

1. $\rho_\alpha = 0$, qui indique que $\text{dist}(\tilde{\mathcal{C}}_\alpha, \mathcal{S}) = 0$ et $\tilde{\mathcal{C}}_\alpha \cap \mathcal{S} \neq \emptyset$ donc l'hypothèse H_0 ne peut être rejetée. La présence de la structure est hautement incertaine.
2. $\rho_\alpha > 0$, la distance $\text{dist}(\tilde{\mathcal{C}}_\alpha, \mathcal{S})$ étant non nulle et $\tilde{\mathcal{C}}_\alpha \cap \mathcal{S} = \emptyset$, l'hypothèse H_0 est rejetée et la structure confirmée.

Dans le cas $\rho_\alpha > 0$ la valeur de ρ_α indique le pourcentage de certitude vis à vis de l'énergie de la structure contenue dans $x_{MAP}^\dagger$. Plus ρ_α est proche de 1 (ou 100%), plus on est certain de voir la vraie structure dans l'image $x_{MAP}^\dagger$.

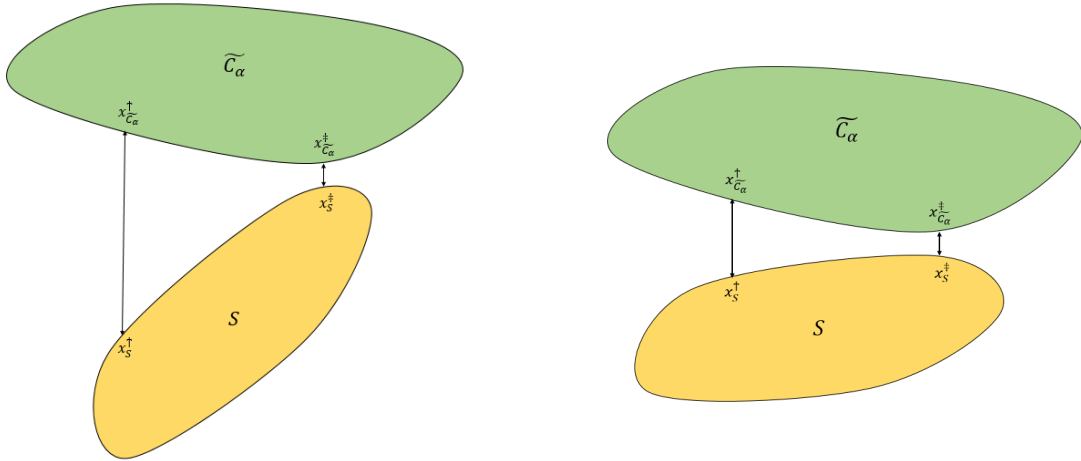


FIGURE 16 – L'image de gauche correspond au cas où ρ_α est relativement petit et l'image de droite correspond à un cas où ρ_α est sensiblement plus proche de 1 (sa valeur maximale)

4.6 BUQO en résumé

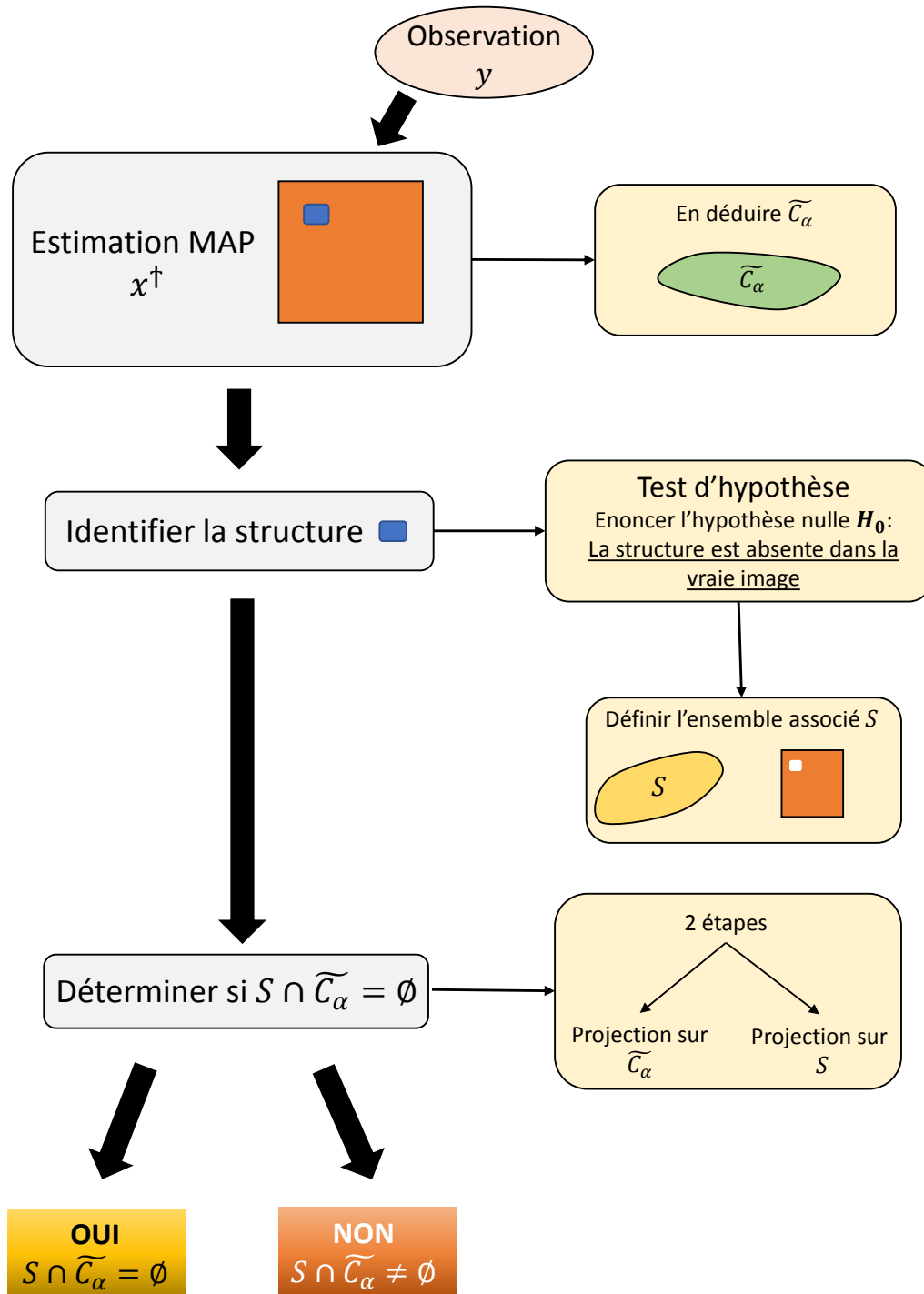


FIGURE 17

4.7 Ajout Artefact

Bien que cela ne fasse pas partie de l'objet du raisonnement de la méthode présentée, on peut pour des raisons d'approfondissement du lien entre les ensembles $\tilde{\mathcal{C}}_\alpha$ et $\mathcal{S}$ illustrer le cas où l'on rajoute une structure à un endroit où il n'y en a pas initialement.

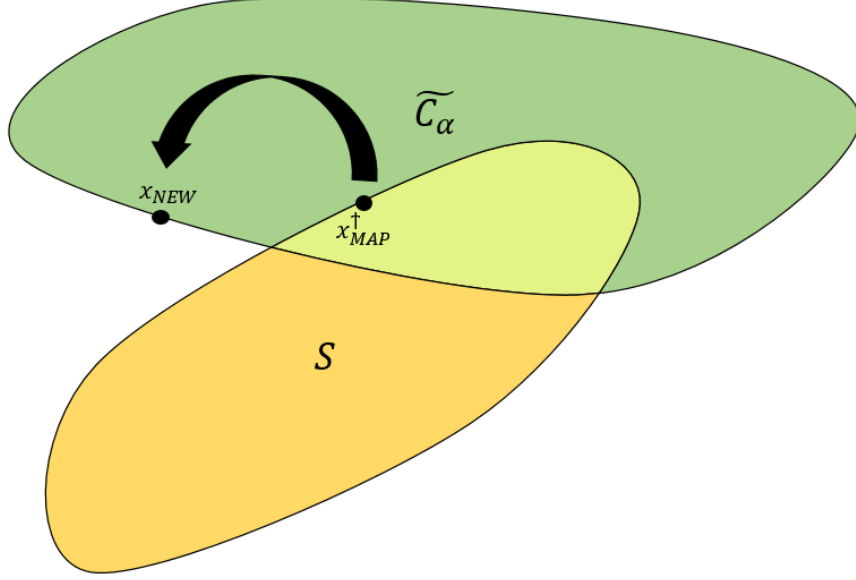


FIGURE 18

L'équation à résoudre est la suivante

$$F(x, x^{(k)}) = \iota_{\mathcal{B}_2(y, \epsilon)}(\Phi x) + \iota_{\mathcal{B}_1(0, \theta)}(\Psi x) + \iota_{[0, 1]^N}(x) + \frac{1}{2} \|x - x^{(k)}\|^2. \quad (40)$$

Elle s'explique comme suit. Partons de la fonction 17 permettant de calculer l'estimation MAP. Cette estimation fait partie de l'ensemble $\tilde{\mathcal{C}}_\alpha$ et est par définition l'élément premier pour le définir. On veut que notre nouvelle image avec une structure ajoutée fasse aussi partie de $\tilde{\mathcal{C}}_\alpha$. On part donc de l'image $x_{MAP}^\dagger$, on lui ajoute une tache à un emplacement voulu en gardant en tête qu'au mieux l'image finale doit garder les mêmes propriétés que $x_{MAP}^\dagger$. On ajoute donc une attache aux données $\frac{1}{2} \|x - x^{(k)}\|^2$ pour s'assurer que la projection conservera la tache ajoutée et le terme $\iota_{\mathcal{B}_1(0, \theta)}(\Psi x)$ s'assure que l'on reste au plus proche des propriétés de $x_{MAP}^\dagger$. La borne θ étant définie comme

$$\lambda f(\Psi x_{MAP}^\dagger) < \theta \ll \tilde{\eta}_\alpha = \lambda f(\Psi x_{MAP}^\dagger) + N(\tau_\alpha + 1)$$

avec η_α la borne de l'ensemble $\tilde{\mathcal{C}}_\alpha$.

L'algorithme suivant permet de résoudre 40

Algorithm 5 Algorithmme primal-dual forward-backward pour résoudre 40

```

1: Initialisation : Soit  $u^{(0)} \in \mathbb{R}^N$ ,  $v_1^{(0)} \in \mathbb{R}^N$ , et  $v_2^{(0)} \in \mathbb{C}^M$ . et  $\sigma > 0$  et  $\gamma > 0$  tels que
    $\sigma(\frac{1}{2} + \gamma(\|\Psi\|^2 + \|\Phi\|^2)) < 1$ .
2: for  $i = 0, 1, \dots$  do
3:    $u^{(i+1)} = \Pi_{[\ell, \infty]^N}(u^{(i)} - \sigma(u^{(i)} - x^{(k)} + \Psi^\dagger v_1^{(i)} + \Phi^\dagger v_2^{(i)}))$ 
4:    $\tilde{v}_1^{(i)} = v_1^{(i)} + \gamma \Psi(2u^{(i+1)} - u^{(i)})$ 
5:    $v_1^{(i+1)} = \tilde{v}_1^{(i)} - \gamma \Pi_{\mathcal{B}_1(0, \theta)}(f)(\gamma^{-1} \tilde{v}_1^{(i)})$ 
6:    $\tilde{v}_2^{(i)} = v_2^{(i)} + \gamma \Phi(2u^{(i+1)} - u^{(i)})$ 
7:    $v_2^{(i+1)} = \tilde{v}_2^{(i)} - \gamma \Pi_{\mathcal{B}_2(y, \epsilon)}(\gamma^{-1} \tilde{v}_2^{(i)})$ 
8: end for

```

Avec cet algorithme, il est donc possible de générer une image x_{NEW} contenant une structure non présente dans l'image réelle, tout en restant dans la région de crédibilité $\tilde{\mathcal{C}}_\alpha$ et en quittant l'ensemble $\mathcal{S}$ des images ne contenant pas la structure. Cela démontre qu'il est possible qu'une image appartenant à $\tilde{\mathcal{C}}_\alpha$ contienne une structure non présente dans l'image réelle.

4.8 Travail dans le domaine des ondelettes

Tout les travaux de quantification d'incertitude étudiés dans les sections précédentes ont été réalisé dans le domaine des pixels. Il serait intéressant d'étudier l'incertitude dans le domaine des ondelettes. Il se pourrait qu'un artefact de reconstruction soit dû à un ou plusieurs coefficients dans la base d'ondelettes qui seraient incertains, l'expression de la reconstruction par l'estimateur MAP contient en effet une optimisation sur les ondelettes. On ne peut pas exclure de l'incertitude dans cette base.

Connaissant la région de crédibilité $\tilde{\mathcal{C}}_\alpha$, il est possible de déterminer des intervalles de confiance, comme ceux énoncés dans le chapitre 2 à l'aide des échantillons générés par MCMC dans le domaine des pixels, mais cette fois-ci dans le domaine des ondelettes et sans avoir besoin d'utiliser MCMC. Le principe est le suivant, partir de $\Psi x_{MAP}^\dagger$, c'est-à-dire la représentation de l'estimation MAP dans le domaine des ondelettes et ensuite diminuer progressivement les valeurs des coefficients jusqu'à ce que l'image associée $x_{\mathcal{W}}^\dagger$ (en définissant le nouveau domaine associé $\mathcal{W}$ des images dont la valeurs des coefficients en ondelettes a changé) ne fasse plus partie de $\tilde{\mathcal{C}}_\alpha$. On aura alors déterminé la valeur minimale de chaque coefficient avec niveau de confiance $(1 - \alpha)$.

La méthode est illustrée sur la figure 19. En pratique on ne jouera que sur les coefficients du premier niveau de détail d_{J-1} (tout du moins dans ce travail). En d'autres mots il faut déterminer

$$\text{Minimum} \left\{ d_{J-1}^{H,V,D}(i, j) \right\} \text{ tel que } x_{\mathcal{W}}^\dagger \in \tilde{\mathcal{C}}_\alpha$$

avec $i, j \in [1, \frac{n}{2}]$.

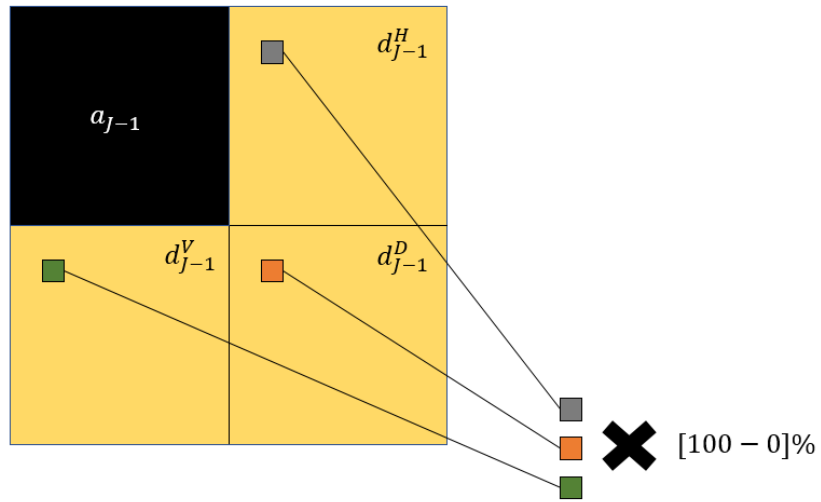


FIGURE 19 – Méthode de quantification de l'incertitude dans le domaine des ondelettes dans une représentation des coefficients de premier niveau de détail

5 Résultats

5.1 Déconvolution et débruitage

Deux types de convolutions ont été abordées dans le chapitre 2, une avec un kernel gaussien et l'autre avec un kernel LoG (le laplacien de la gaussienne). Elles seront donc toutes les deux explorées dans le cadre de l'estimation du MAP. Les simulations sont évaluées jusqu'à la convergence de l'algorithme 2. On considère que cet algorithme a convergé si la condition suivante est remplie

$$\|x^{k+1} - x^k\| < 10^{-5} \|x^{k+1}\|.$$

L'image utilisée est la représentation d'une fleur d'Hibiscus. Les ondelettes utilisées sont de type Haar. Lorsque l'expression *SNR du bruit* est rencontrée, elle traduit le SNR entre l'image convoluée et l'image convoluée et bruitée de sorte que l'on puisse analyser les résultats indépendamment du bruit apporté par la convolution. On rappelle que le SNR est inversement proportionnel au bruit, donc un haut niveau de SNR se traduit par un faible bruit. L'expression du SNR est

$$SNR = 20 \log_{10} \left(\frac{\|a\|}{\|a - b\|} \right)$$

avec a un signal de référence (image sans bruit) et b le signal à analyser.

Tant pour la déconvolution gaussienne que pour la déconvolution LoG, une tendance claire apparaît, l'écart entre le SNR d'entrée (entre y l'image bruitée et convoluée et x la vraie image, courbe bleue) et le SNR de sortie (entre $x^\dagger$ l'image débruitée et déconvoluée et x la vraie image, courbe orange) diminue en fonction du bruit ajouté (quand le SNR bruit diminue). Cela s'explique par le fait qu'à un petit niveau de bruit, l'algorithme ne doit lutter que contre la déconvolution alors qu'au fur et à mesure que celui-ci augmente, l'algorithme doit de plus en plus lutter pour à la fois résoudre le problème de la déconvolution et le problème du débruitage avec un effet de bruit de plus en plus fort.

Une autre tendance est la meilleure performance de la reconstruction de la convolution LoG. Alors que la déconvolution LoG atteint un écart maximum de 20 dB entre le SNR de y et le SNR de $x^\dagger$, pour la déconvolution gaussienne cet écart tombe à maximum 10 dB.

Les mêmes observations peuvent être déduites par l'analyse visuelle des images. Lorsque le bruit est très faible (SNR bruit = 70 [dB]), la déconvolution est presque parfaite visuellement dans les deux cas. En revanche lorsque le bruit est plus important (40 [dB]), dans le cas de la déconvolution gaussienne il est clair que le résultat est moins bon. Dans le cas de la déconvolution LoG le résultat est meilleur, avec cependant l'apparition d'artefacts de reconstruction (sorte de carrés) propres aux ondelettes. Lorsque le bruit est très im-

portant (18 [dB]) le résultat est loin d'être satisfaisant, et ceci tant pour la déconvolution Gaussienne que la LoG avec des artefacts de reconstruction très marqués (surtout dans le cas LoG).

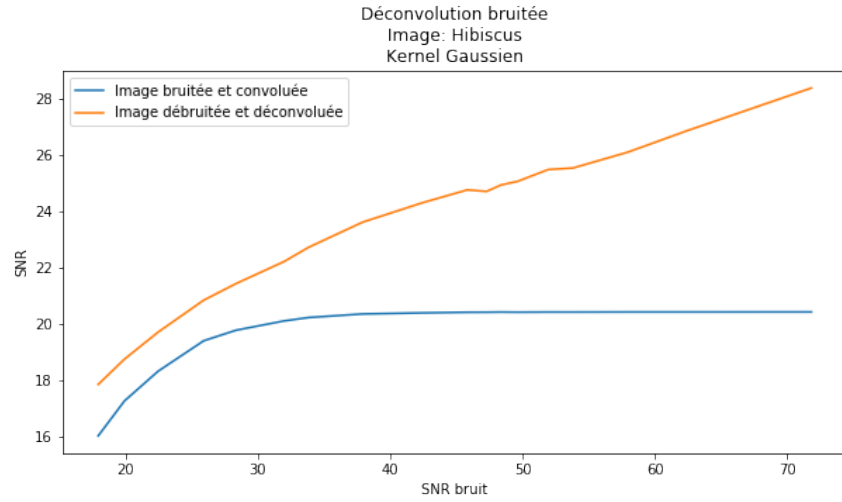


FIGURE 20 – Résultat de la déconvolution gaussienne en fonction du bruit d'entrée

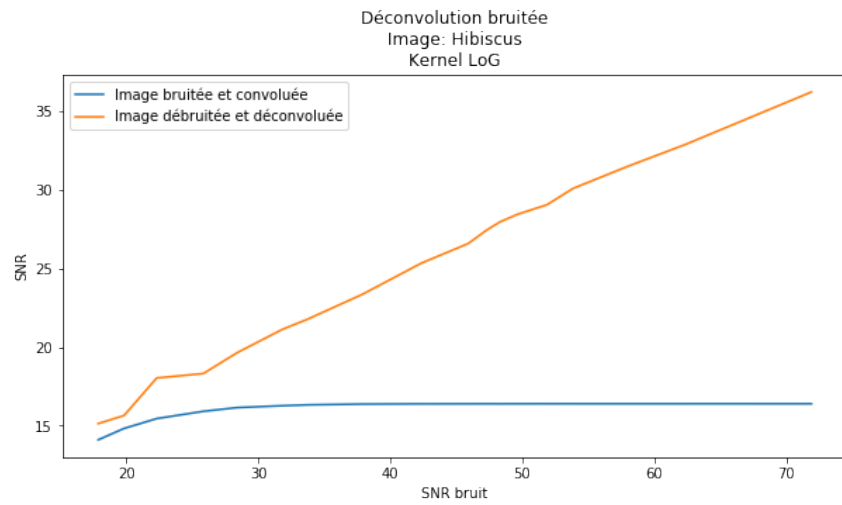
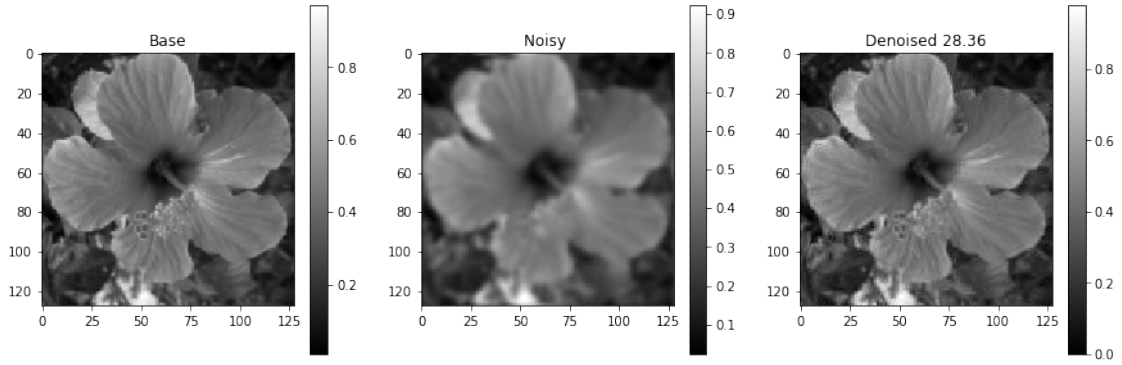
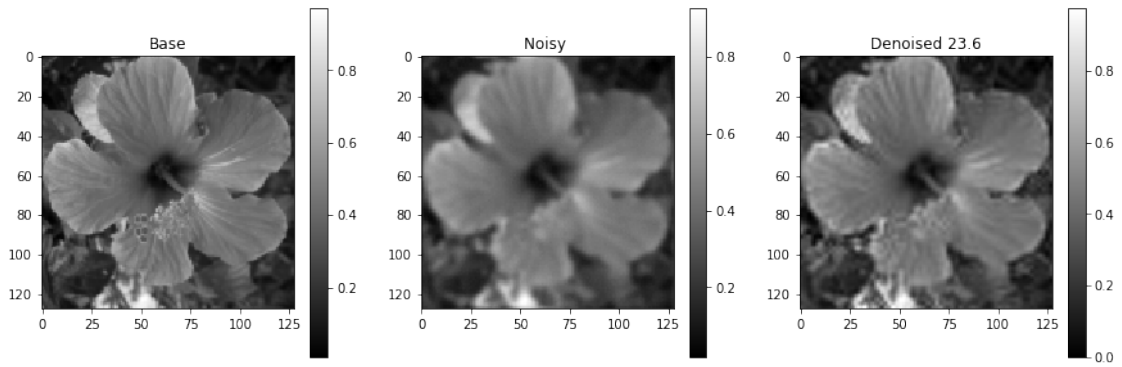


FIGURE 21 – Résultat de la déconvolution LoG en fonction du bruit d'entrée

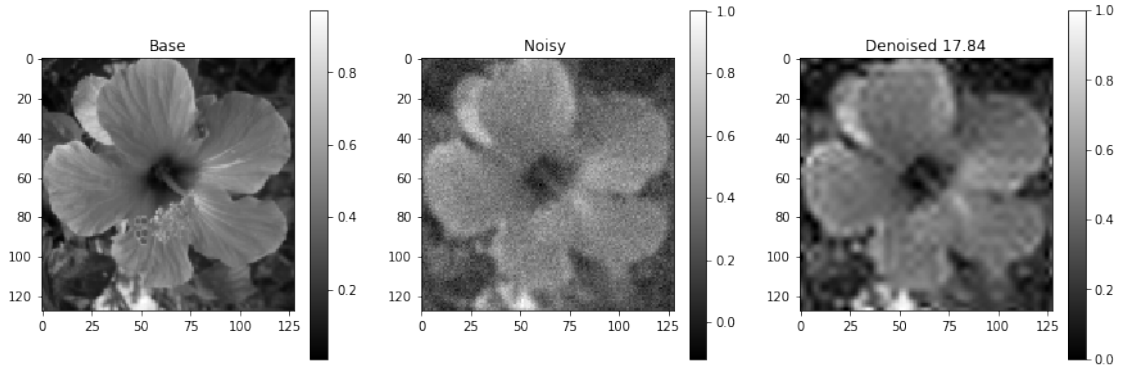
Convolution kernel Gaussien



(a) SNR bruit = 70 [dB]



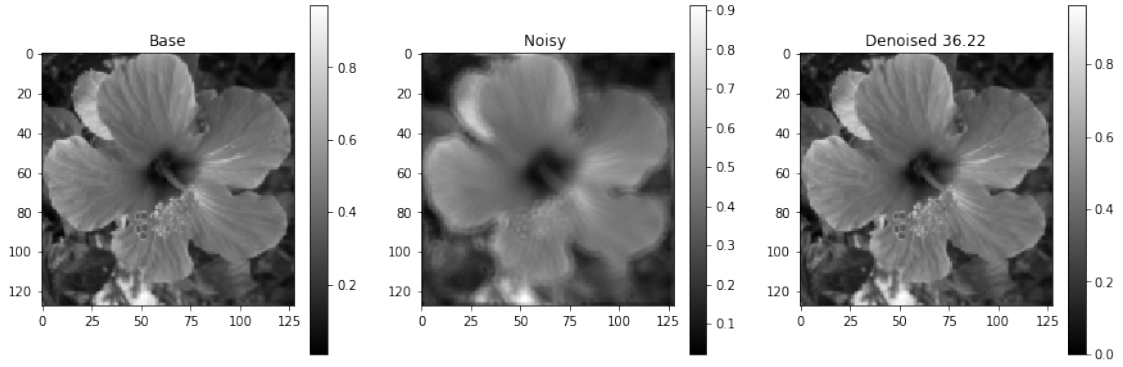
(b) SNR bruit = 40 [dB]



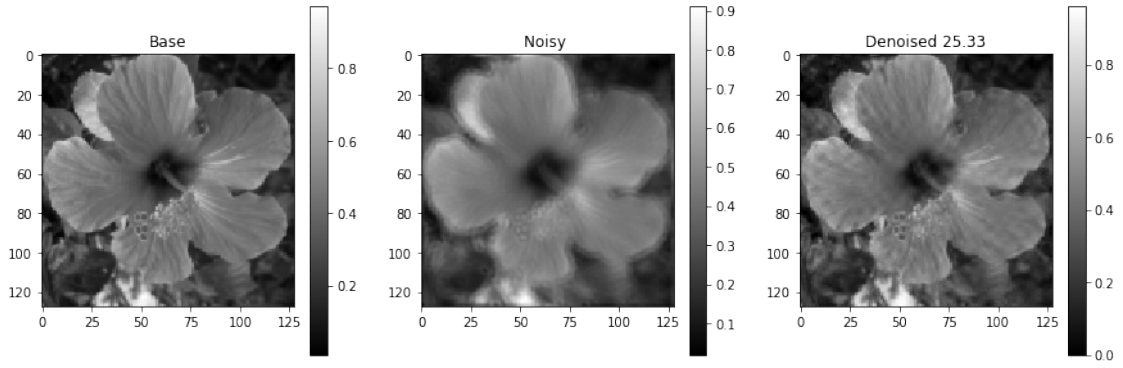
(c) SNR bruit = 18 [dB]

FIGURE 22 – Déconvolution Gaussienne pour différents niveaux de bruit. De gauche à droite respectivement, l'image réelle, l'image bruitée et déconvoluée, et l'estimation MAP $x^\dagger$.

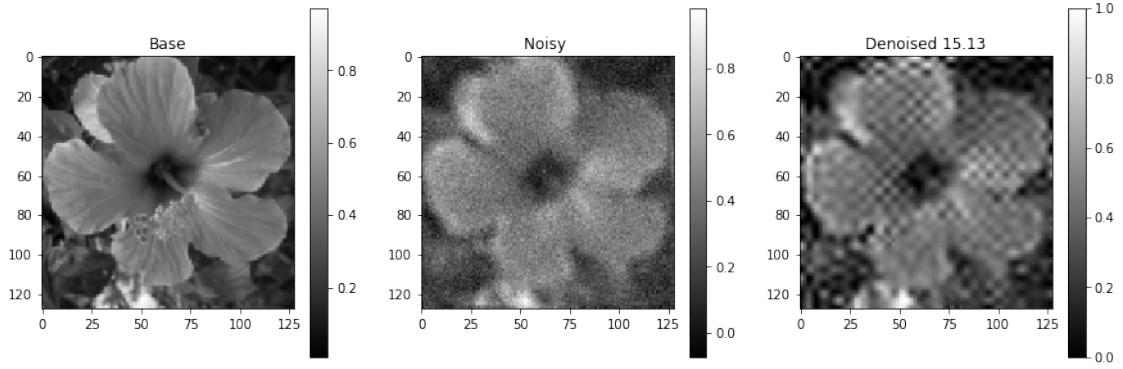
Convolution kernel LoG



(a) SNR bruit = 70 [dB]



(b) SNR bruit = 40 [dB]



(c) SNR bruit = 18 [dB]

FIGURE 23 – Déconvolution LoG pour différents niveaux de bruit. De gauche à droite respectivement, l'image réelle, l'image bruitée et déconvoluée, et l'estimation MAP $x^\dagger$.

5.2 Quantification de l'incertitude via MCMC

Pour la suite de l'analyse, on considère une image (Figure 24) générée artificiellement par un algorithme développé pour la reconstruction d'image en microscopie présenté dans [3]. Cette image est composée de fibres plus ou moins longues et d'intensités variables, et de taches ponctuelles de forme ronde. Il y a également un léger "brouillard" dans certaines zones, le fond de l'image n'y est donc pas nul. Différentes analyses sont menées, à différents niveaux de bruit et toujours avec une convolution de type LoG.

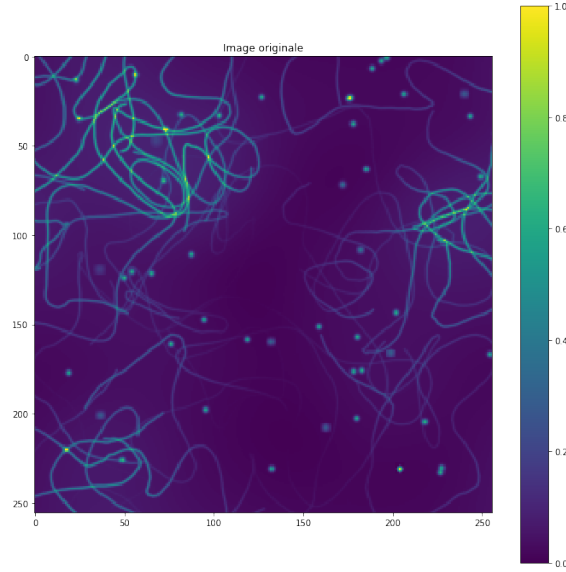
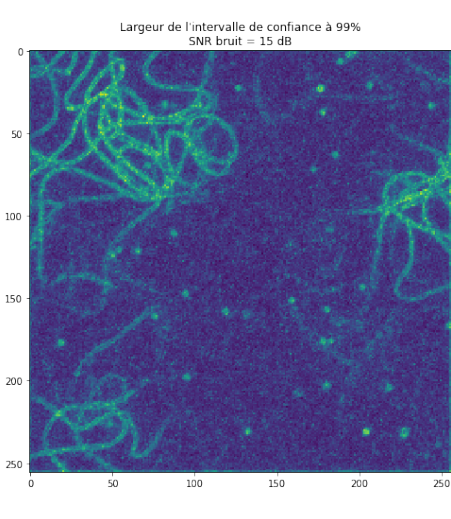
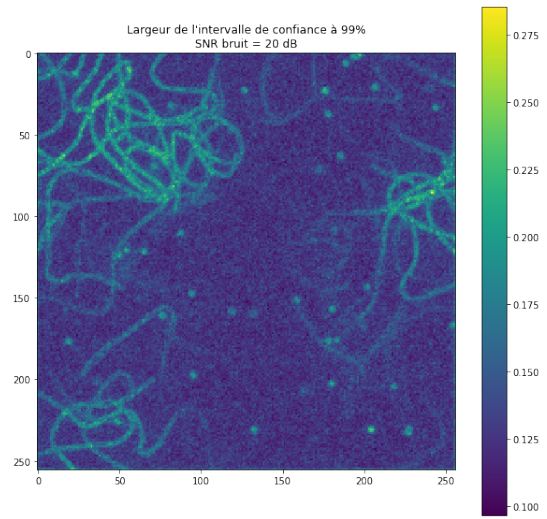


FIGURE 24 – Image utilisée pour l'analyse de l'incertitude

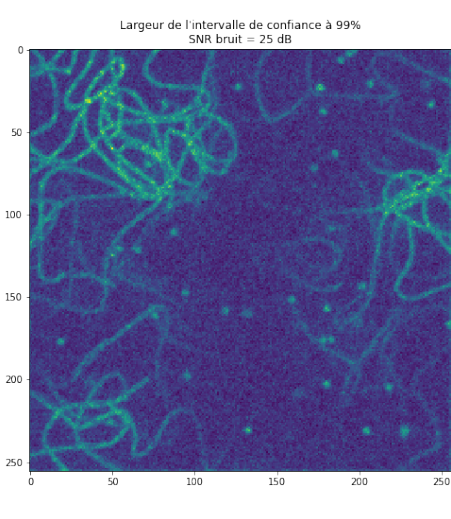
Comme abordé dans les précédents chapitres, il y a plusieurs méthodes pour quantifier l'incertitude. En utilisant la méthode de MCMC, on peut calculer des intervalles de confiance pour chaque pixel. C'est réalisé ici en utilisant une version légèrement modifiée de l'algorithme SK-ROCK disponible ici [15]. Le résultat est visible sur la figure 25 pour différents niveaux de bruits indiqués. Il est clair que plus le bruit augmente, plus la largeur des intervalles de confiance sont larges. D'une largeur comprise entre 0.009 et 0.027 à 50 [dB], on atteint une largeur comprise entre 0.1 et 0.35 à 15 [dB]. Étant donné que la valeur des pixels est comprise entre 0 et 1, on passe donc d'intervalles de confiance qui valent entre 0.9 % et 2.7 % de la valeur que peut prendre un pixel à 50 [dB] à entre 10 et 35 % à 15 [dB]. La différence est significative. On remarque également que plus le bruit augmente, plus l'incertitude est localisée là où il y a des structures. Ces intervalles de confiance étant calculés pixel par pixel, ils ne nous donnent pas d'indication claire quant à la présence d'une structure ou non. Cependant ils nous donnent déjà une première tendance quant à l'évolution de l'incertitude en fonction du niveau de bruit et nous permettent de déterminer des zones où l'incertitude est plus forte.



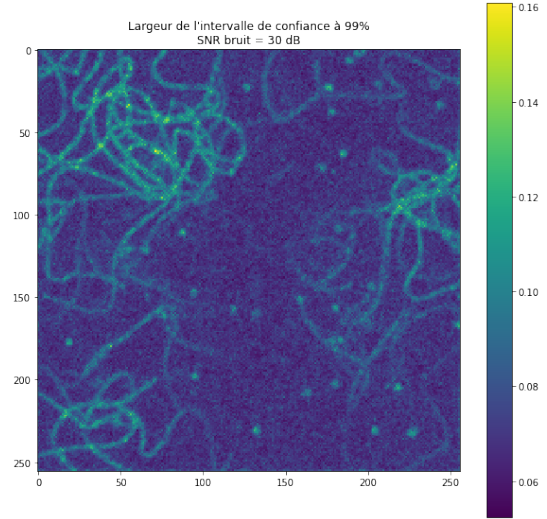
(a) SNR bruit = 15 [dB]



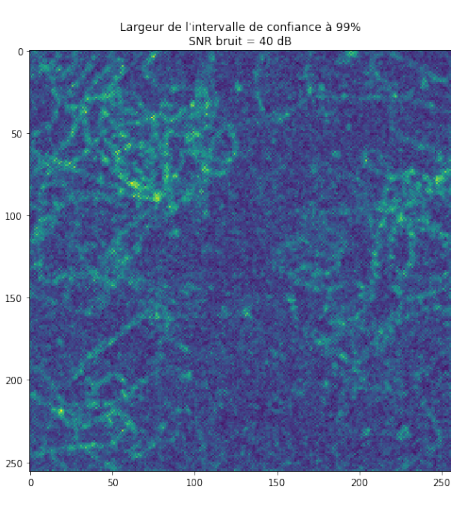
(b) SNR bruit = 20 [dB]



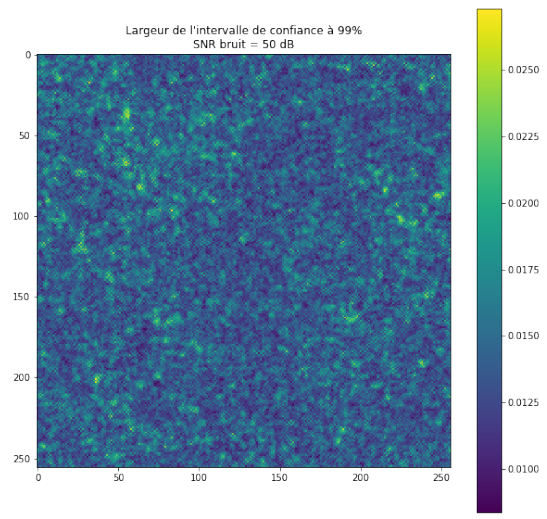
(c) SNR bruit = 25 [dB]



(d) SNR bruit = 30 [dB]



(e) SNR bruit = 40 [dB]



(f) SNR bruit = 50 [dB]

FIGURE 25 – Largeurs des intervalles de confiance générés par une méthode de MCMC évaluées pixel par pixel.

5.3 Quantification de l'incertitude via BUQO

Pour rappel la base d'ondelettes Ψ est de type Haar. Comme dans l'article [17], on choisit $\alpha = 1\%$ et $\lambda = \frac{N}{\|\Psi x^\dagger\|_1}$ pour $\tilde{\mathcal{C}}_\alpha$. Les simulations sont menées jusqu'à la convergence le l'algorithme 4. On considère qu'il a atteint la convergence si les conditions

$$\|x^{k+1} - x^k\| < 10^{-5} \|x^{k+1}\|$$

et

$$\|x^{k+\frac{1}{2}} - x^{k-\frac{1}{2}}\| < 10^{-5} \|x^{k+\frac{1}{2}}\|$$

sont simultanément respectées. Ces critères vérifient la relative convergence des séquences des projections sur $\tilde{\mathcal{C}}_\alpha$ et $\mathcal{S}$. Un autre critère évalue si l'algorithme converge vers une distance $\text{dist}(\tilde{\mathcal{C}}_\alpha, \mathcal{S})$

$$|\delta^{k+1} - \delta^k| < 10^{-5} \delta^{k+1}$$

avec $\delta^{k+1} = \|x^{k+\frac{1}{2}} - x^{k+1}\|$.

Il est important de remarquer que vu les critères d'arrêt définis ici, l'algorithme ne peut pas converger exactement vers $\text{dist}(\tilde{\mathcal{C}}_\alpha, \mathcal{S}) = 0$. Il faut alors revoir le critère de rejet de l'hypothèse H_0 étant donné que ρ_α ne peut (presque) jamais atteindre 0. On choisit donc une valeur arbitraire de 3% (et plus 0) au dessus de laquelle l'hypothèse H_0 sera rejetée. En d'autres mots si $\rho_\alpha > 3\%$ on confirme la structure.

On choisit également l'opérateur $\mathcal{L}$ associé à $\mathcal{S}$ défini dans le chapitre 4, comme $\mathcal{L} = \frac{1}{3}(\mathcal{L}_{3x3} + \mathcal{L}_{7x7} + \mathcal{L}_{11x1})$ avec $\mathcal{L}_{3x3}$ représentant une convolution entre l'image sans la structure (0 à cet endroit) et un kernel gaussien de taille 3×3 .

La figure 26 montre le résultat de la méthode BUQO pour une structure en particulier (la boule de droite en l'occurrence sur l'image). Ce test a été réalisé pour un niveau de bruit de 20 [dB]. On obtient $\rho_\alpha = 9.26\%$, cette valeur étant supérieure à 3%, on en conclut que $\tilde{\mathcal{C}}_\alpha \cap \mathcal{S} = \emptyset$. Donc l'hypothèse H_0 est rejetée et la structure est confirmée avec niveau de confiance $(1 - \alpha)$.

La figure 27 montre un autre résultat pour une structure située à un autre emplacement pour un bruit de 34 [dB]. On obtient $\rho_\alpha = 0.81\%$, cette valeur étant inférieure à 3%, on en conclut que $\tilde{\mathcal{C}}_\alpha \cap \mathcal{S} \neq \emptyset$ et l'hypothèse H_0 ne peut être rejetée. La structure est donc considérée comme hautement incertaine.

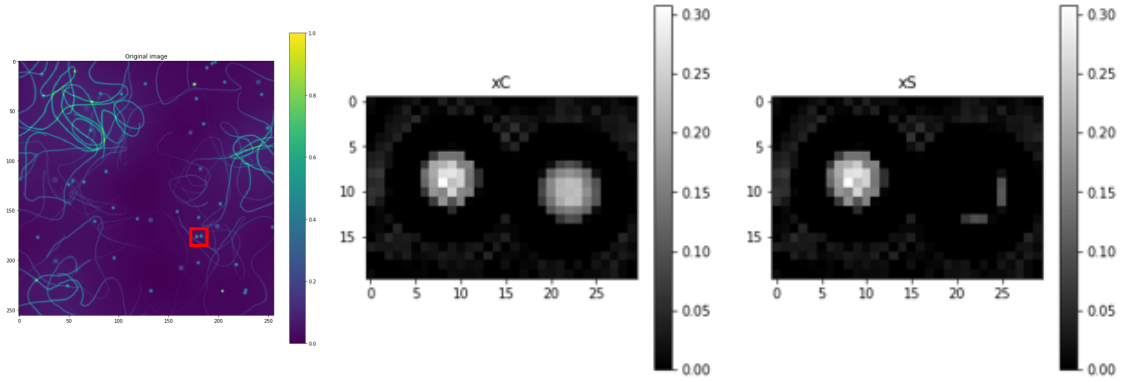


FIGURE 26 – À gauche, l’emplacement de la structure d’intérêt, seule la boule de droite est analysée. Au milieu $x_{\tilde{\mathcal{C}}_\alpha}^\dagger$, le résultat de la projection sur $\tilde{\mathcal{C}}_\alpha$ et à droite $x_{\mathcal{S}}^\dagger$, le résultat de la projection sur $\mathcal{S}$. Ici $\rho_\alpha = 9.23\%$ et la structure est confirmée.

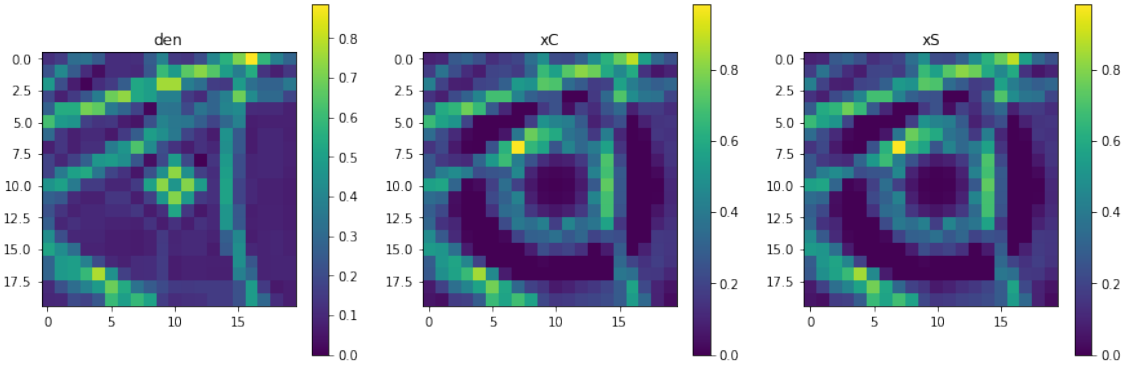


FIGURE 27 – À gauche, l’emplacement de la structure dans l’image de l’estimation MAP $x^\dagger$. Au milieu $x_{\tilde{\mathcal{C}}_\alpha}^\dagger$, le résultat de la projection sur $\tilde{\mathcal{C}}_\alpha$ et à droite $x_{\mathcal{S}}^\dagger$, le résultat de la projection sur $\mathcal{S}$. Ici $\rho_\alpha = 0.81\%$ et la structure n’est pas confirmée.

Plusieurs paramètres pourraient influencer la validation d’une structure, sa hauteur, sa taille et enfin son emplacement.

5.3.1 Influence de la hauteur de la structure

On propose d’étudier la hauteur de la structure à un emplacement donné sur l’image. La structure utilisée comme banc de tests est une structure similaire aux structures déjà présentes. On fait varier la hauteur pour différents niveaux de bruits. Ces hauteurs sont comprises dans l’intervalle $h \in [0.1, 1]$. Le tableau des résultats est disponible à la figure 28. Pour les bruits de 50, 40 et 33.5 [dB], la valeur de ρ_α est toujours supérieure à 3%, l’hypothèse H_0 est toujours rejetée et la structure confirmée. Cependant à partir de 30 [dB], les structures de hauteurs 0.2 et 0.1 ne sont plus validées.

		SNR BRUIT							
		50	40	33,5	30	25,5	21	16	13,5
HAUTEUR DE LA STRUCTURE	0,1	74	36,5	3,84	1,3	0,5	0,027	0,0093	0
	0,2	83,34	57	23,65	2,94	0,75	0,26	0,015	0,01
	0,3	87,45	67,27	40,53	17,93	1,03	0,347	0,014	0,117
	0,4	89,31	73,12	51,47	31,52	3,35	0,48	0,18	0,125
	0,5	91,05	71,72	59,3	42,8	14,76	0,61	0,135	0,086
	0,6	91,26	80,36	63,94	48,95	24,32	0,91	0,2	0,135
	0,7	91,79	82	67,81	54,7	30,97	4,48	0,178	0,099
	0,8	92,24	83	70,8	59,1	37,5	10,7	0,23	0,093
	0,9	92,7	84,3	73,1	62,6	42,98	17,2	0,238	0,158
	1	93,12	85,25	75,15	65,61	47,46	23,82	0,3	0,16

FIGURE 28 – Tableau des valeurs de ρ_α en fonction de la hauteur de la structure et du bruit.

La tendance concernant l'évolution du ρ_α en fonction de la hauteur et du bruit est disponible à la figure 29. Il est clair que la hauteur de la structure étudiée a un impact significatif sur la conclusion qui sera donnée concernant sa présence ou non dans l'image réelle (rejet de H_0 ou non). Plus la hauteur est importante, plus le ρ_α associé est élevé.

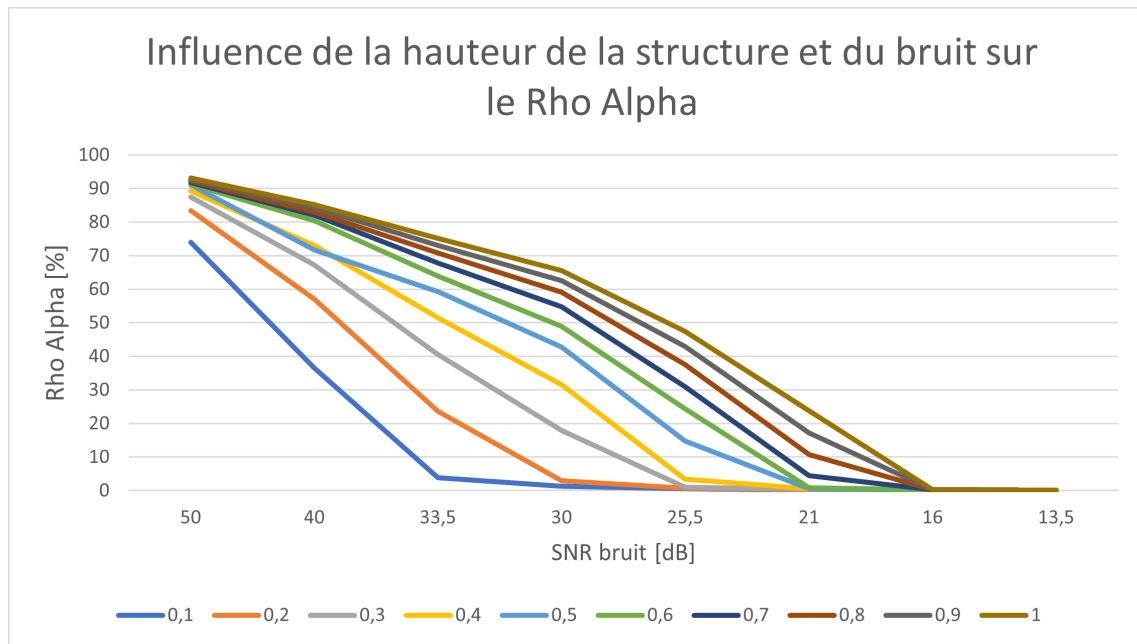


FIGURE 29 – Graphe de l'évolution du ρ_α en fonction de la hauteur de la structure et du bruit

La figure 30 montre combien de fois la structure a été validée sur le banc de test (les différents niveaux de bruit) en fonction de sa hauteur. Plus la hauteur de la structure est haute, plus elle est validée.

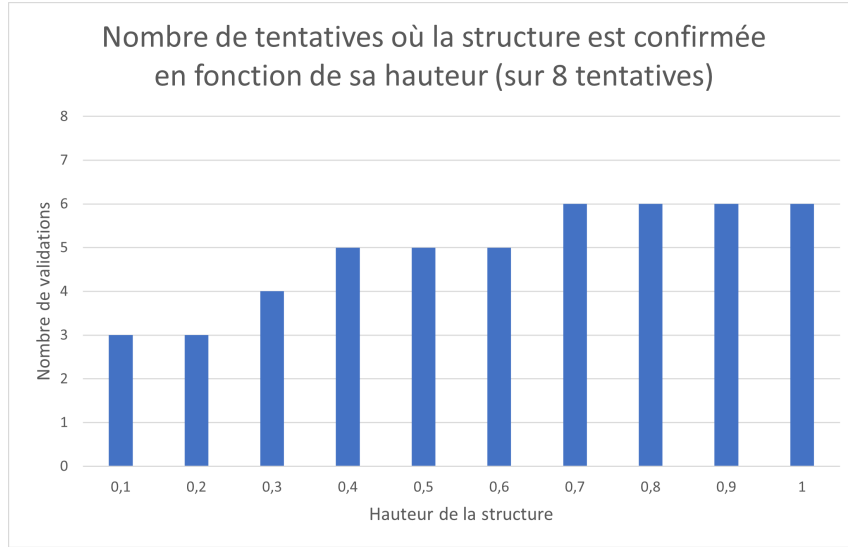


FIGURE 30 – Statistiques de validation de la structure en fonction de sa hauteur

5.3.2 Influence de la taille de la structure

La même expérience qu’au point précédent est réalisée cette fois-ci concernant la taille de la structure. On analyse une structure (un disque) de rayons compris entre 1 et 10 pixels et de hauteur fixe $h = 0.6$. Le tableau des résultats est disponible à la figure 31. Rappelons que la structure est confirmée si $\rho_\alpha > 3\%$.

		SNR BRUIT							
		50	40	33,5	30	25,5	21	16	13,5
RAYON DE LA STRUCTURE	1	71,96	30	0,52	0,13	0,012	0,007	0,004	0,003
	2	84,5	61,05	32,66	9,13	0,34	0,165	0,077	0,052
	3	92,87	81,66	68,25	55,58	33,48	5,83	0,16	0,075
	4	95,12	87,9	79,01	70,93	57,48	37,796	1,735	0,184
	5	96,13	90,81	84,5	78,97	68,83	54,38	29,13	15,6
	6	97,06	92,75	87,57	82,86	74,28	62,03	42,06	30,95
	7	97,63	94,07	89,67	85,6	78,18	68,38	50,84	41,11
	8	97,8	94,8	91,13	87,73	81,5	73,07	58,1	50,16
	9	98,15	95,58	92,36	89,44	83,97	76,64	63,45	56,42
	10	98,49	96,23	93,38	90,75	85,9	79,4	67,84	61,68

FIGURE 31 – Tableau des valeurs de ρ_α en fonction du rayon de la structure et du bruit.

La même conclusion qu’au point précédent peut être tirée via la figure 32, plus la taille de la structure est grande, plus le ρ_α sera élevé.

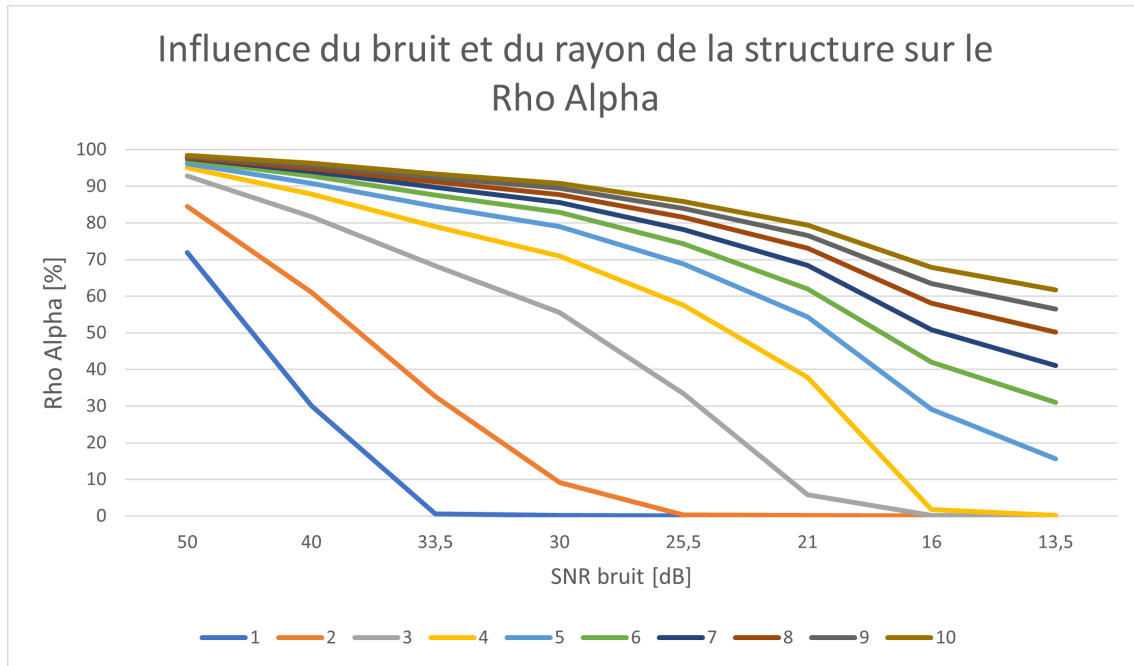


FIGURE 32 – Graphe de l'évolution du ρ_α en fonction du rayon de la structure et du bruit

Sur la figure 33 on remarque que tous les tests sont validés à partir d'un rayon de 5 pixels.

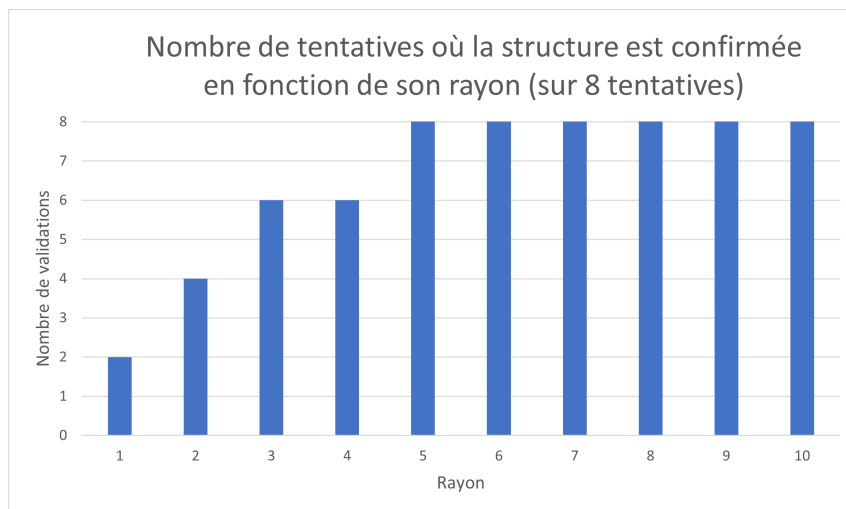


FIGURE 33 – Statistiques de validation de la structure en fonction de son rayon

5.3.3 Influence de l'emplacement

On l'a vu au début de ce chapitre, la méthode MCMC nous permet de générer des intervalles de confiance. Il est clair que sur la figure 25 on peut identifier des zones plus incertaines que d'autres (des zones où l'intervalle de confiance est plus large). On identifie sur la figure 35 trois emplacements, un dans lequel il y a peu d'incertitude et deux où il y a une haute incertitude. On y place la même structure, rayon de 2 pixels et hauteur

$h = 0.6$ et on teste pour différents niveau de bruits.

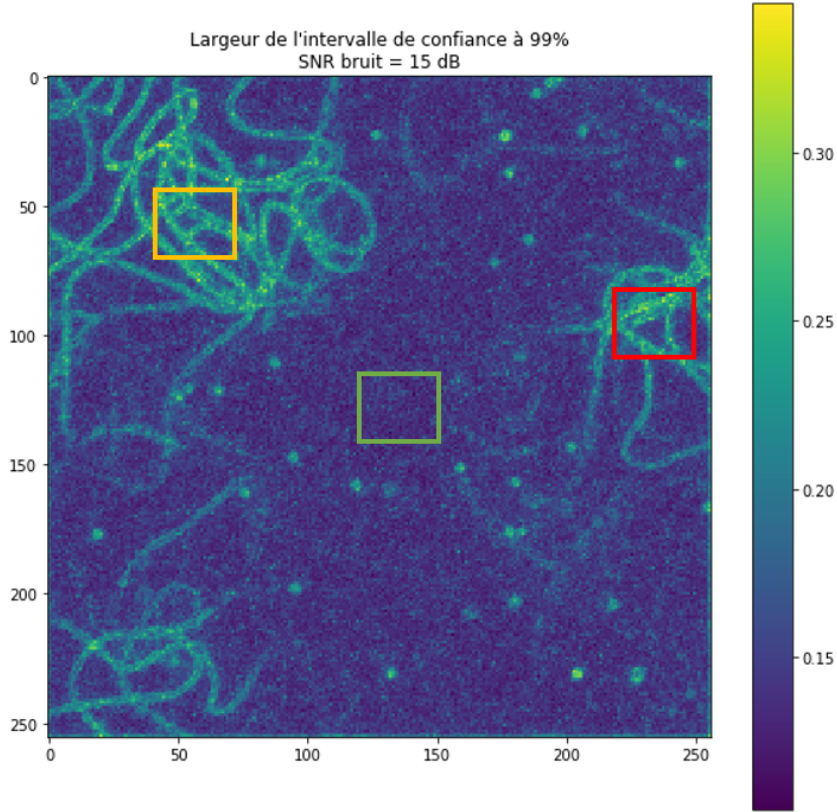


FIGURE 34 – Identification de zones d’incertitudes et emplacement des zones de test. En vert la zone de faible incertitude, et en rouge et orange les zones de forte incertitude

La figure 35 indique une claire différence entre les résultats obtenus dans une zone de faible incertitude et les zones de haute incertitude. Étant donné que dans une zone de plus haute incertitude l’intervalle de confiance est plus large, il est logique de constater que le ρ_α réagit dans le même sens vu qu’il représente le pourcentage de certitude vis à vis de l’énergie de la structure contenue dans $x_{MAP}^\dagger$. Plus haute incertitude mène à une plus petite valeur de ρ_α . On en conclut également que la structure n’est plus confirmée que dans deux cas sur huit dans les zones de haute incertitude alors qu’elle était validée 4 fois dans la zone de basse incertitude.

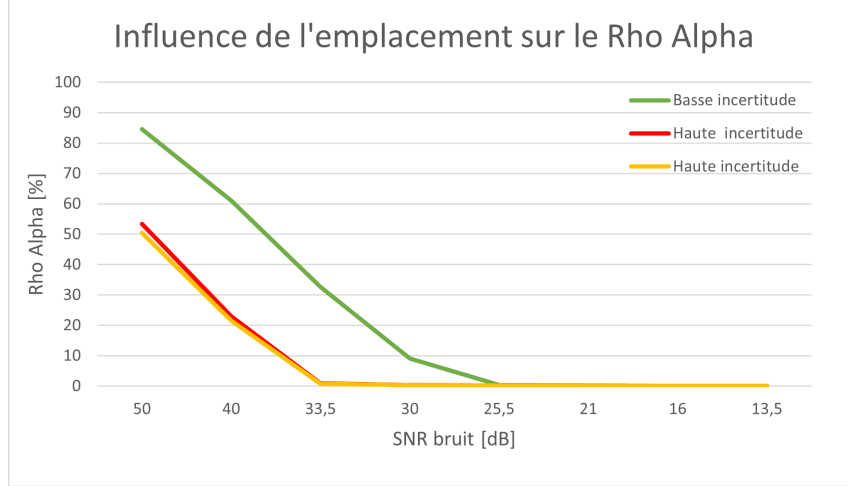


FIGURE 35 – Graphique de l’influence de l’emplacement des structures sur la valeur de ρ_α . Le code couleur est le même que celui de la figure 35

5.4 Cartes d’incertitude

A la lumière de la méthode proposée dans le chapitre 4 concernant la quantification d’incertitude dans le domaine des ondelettes, on réalise sur plusieurs niveaux de bruit ce qu’on appellera des *cartes d’incertitude*. Elles représentent pour chaque coefficient dans la base d’ondelettes le pourcentage **minimal** de sa valeur initiale tel que l’image associée $x_{\mathcal{W}}^\dagger$ se trouve toujours dans la région de crédibilité $\tilde{\mathcal{C}}_\alpha$.

Les différentes cartes pour différents niveaux de bruit se trouvent aux figures 36 et 37. Visuellement on en déduit que plus le bruit augmente, plus on peut diminuer la valeur des coefficients. Cela rejoint l’observation faite avec les intervalles de confiance générés avec MCMC vus précédemment. Augmenter le bruit, augmente l’incertitude.

On calcule également la moyenne des pourcentages sur la totalité des coefficients pour chaque carte afin de mettre des chiffres sur l’incertitude. Ces valeurs vont de 3.23 % (qui exprime donc un pourcentage de **certitude**) de moyenne à 13 [dB] de bruit jusque 54.3% de moyenne à 50 [dB].

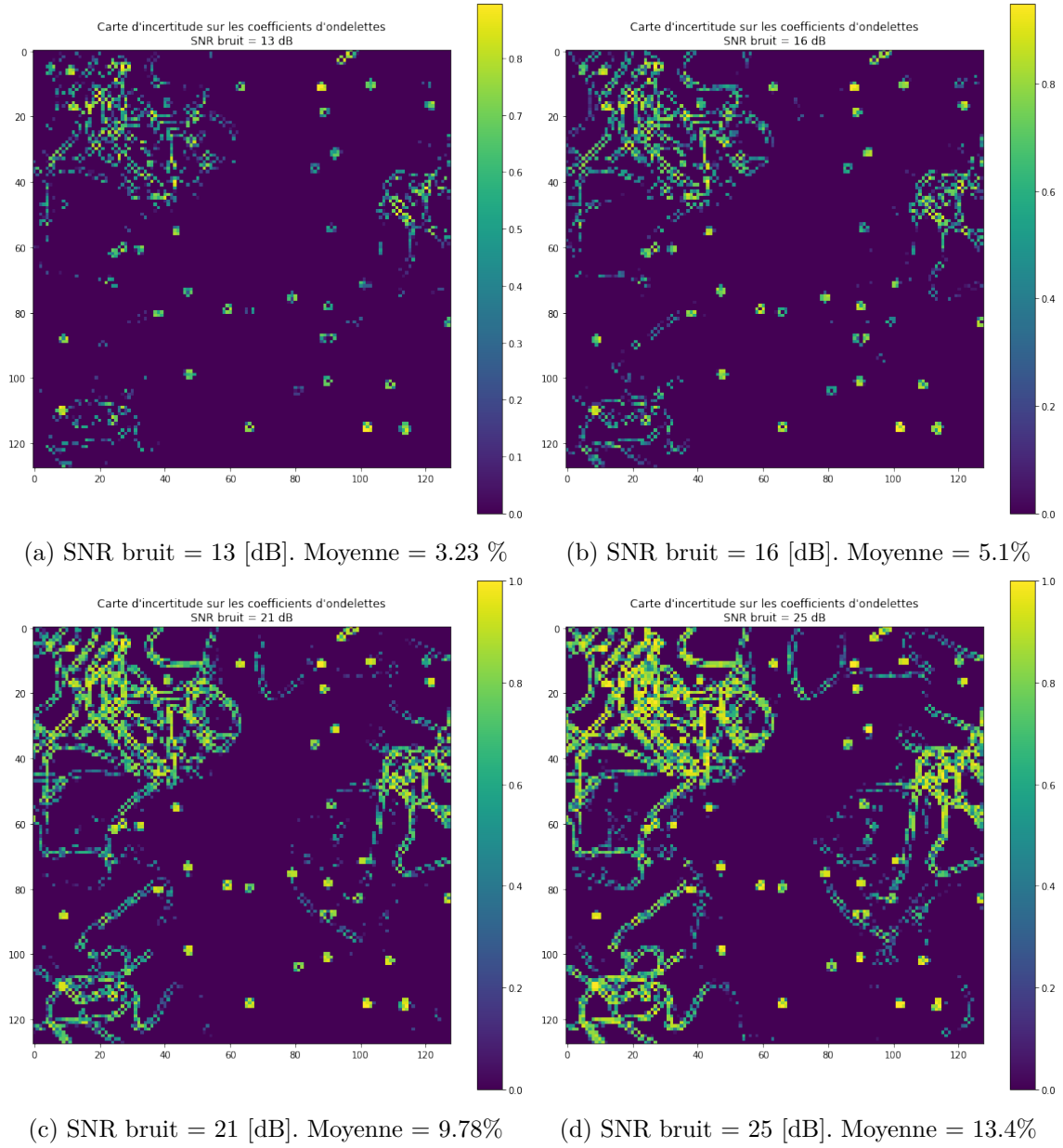


FIGURE 36 – Cartes d'incertitude dans le domaine des ondelettes. La couleur jaune indique une très forte certitude à propos d'un coefficient et la couleur bleue indique une très forte incertitude. La valeur 1 correspond à un pourcentage de 100 %

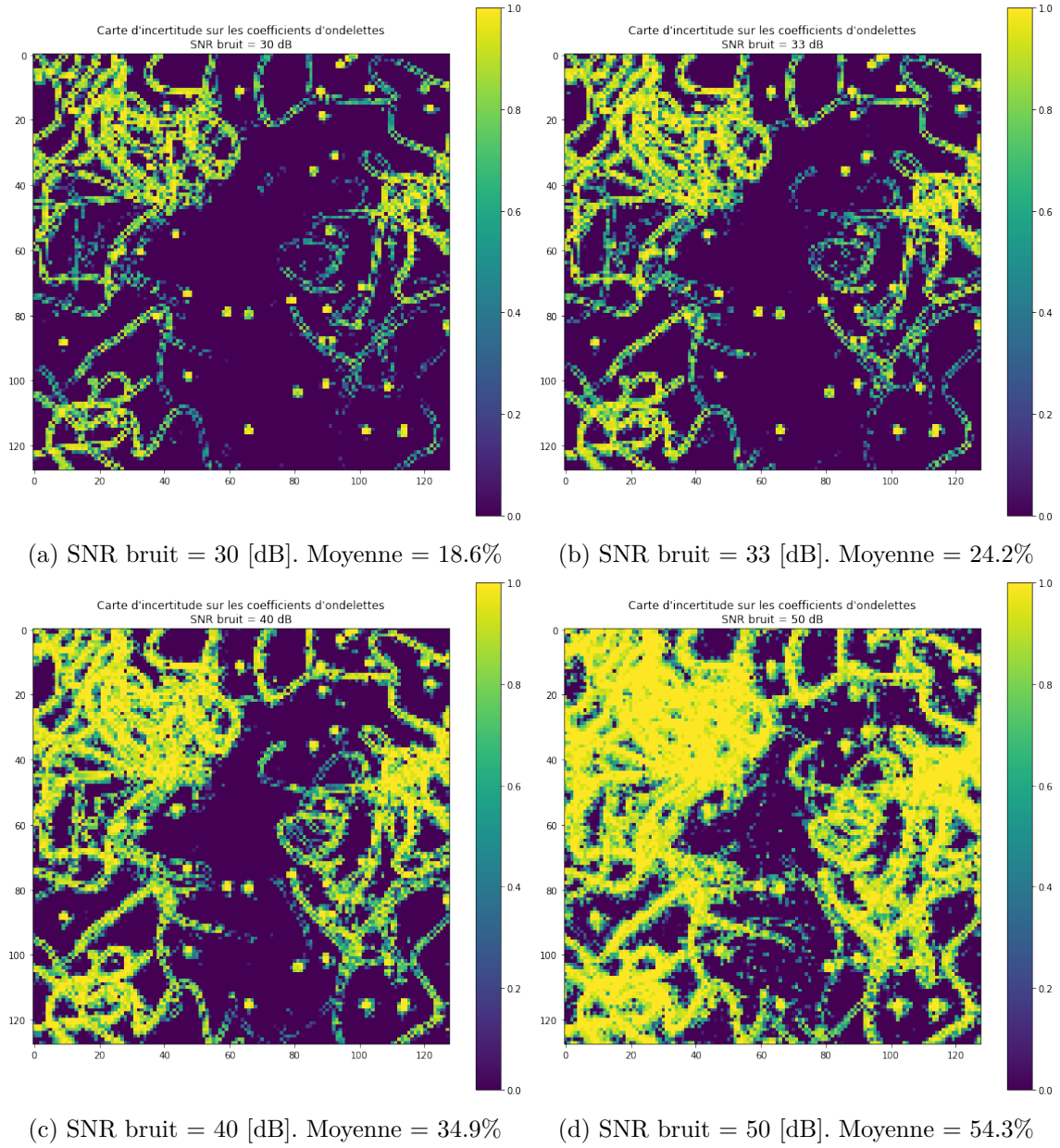


FIGURE 37 – Cartes d'incertitude dans le domaine des ondelettes. La couleur jaune indique une très forte certitude à propos d'un coefficient et la couleur bleue indique une très forte incertitude. La valeur 1 correspond à un pourcentage de 100 %

6 Conclusion et perspectives

Le but de ce travail était de mettre en lumière les méthodes de quantification d'incertitude en partant de la définition du problème inverse sous formulation bayésienne et log-concave jusqu'à sa résolution et implémentation grâce à l'optimisation convexe. La méthode BUQO amène une relative facilité d'appréhension comparativement aux méthodes classiques de Monte-Carlo (générer des échantillons avec SK-ROCK demande un temps de calcul d'environ une heure alors que l'estimation de $x_{MAP}^\dagger$ ne demande pas plus de quelques minutes) puisque elle repose sur l'estimation du maximum à posteriori pour générer la région de haute densité de crédibilité qui permet de réaliser un test d'hypothèse. Ce test, on l'a vu, permet de déterminer si oui ou non une structure présente dans l'estimation $x_{MAP}^\dagger$ est présente dans la vraie image.

Des intervalles de confiance ont été implémentés grâce à une méthode MCMC, ce qui a permis de confirmer le lien entre la quantification de l'incertitude avec la méthode classique dite MCMC et la nouvelle méthode BUQO. L'influence de la taille, de la hauteur et surtout de l'emplacement de la structure ont été mis en évidence. Les régions de haute incertitude indiquées par le méthode de MCMC via les intervalles de confiance ont été validées par l'évaluation du pourcentage de certitude de la structure avec BUQO (la valeur ρ_α) et inversement.

Finalement la nouvelle définition du problème dans la base d'ondelettes a également permis de montrer l'évolution de l'incertitude dans une image et ceci dans une nouvelle base, autre que celle des pixels. Cela ouvre la voie vers une amélioration ou une méthode alternative à BUQO qui déterminerait l'incertitude non plus dans la base de pixels mais dans la base d'ondelettes. Une amélioration de la méthode de détermination des cartes d'incertitudes pourrait également être amenée. Jusqu'ici tous les coefficients sont testés mais dans un but d'optimisation il serait judicieux de présenter une méthode similaire qui ne teste que les coefficients d'intérêt particulier.

Références

- [1] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2) :205–220, April 1992.
- [2] Mark Bagnoli and Ted Bergstrom. Log-concave probability and its applications. volume 26, pages 217–241. January 2006.
- [3] Jérôme Boulanger, Nelly Pustelnik, Laurent Condat, Lucie Sengmanivong, and Tristan Piolot. Nonsmooth convex optimization for structured illumination microscopy image reconstruction. *Inverse Problems*, 34(9), july 2018.
- [4] Stephen Boyd and Parikh, Neal. Proximal Algorithms. *Foundations and Trends in Optimization*, 1(3) :123–231, 2013.
- [5] Steven L. Brunton and J. Nathan Kutz. *Data-Driven Science and Engineering : Machine Learning, Dynamical Systems, and Control*. Cambridge University Press, February 2019. Google-Books-ID : gNcRuQEACAAJ.
- [6] Xiaohao Cai, Marcelo Pereyra, and Jason D. McEwen. Uncertainty quantification for radio interferometric imaging : I. proximal MCMC methods. *Monthly Notices of the Royal Astronomical Society*, 480(3) :4154–4169, November 2018. arXiv : 1711.04818.
- [7] Laurent Condat. A Primal–Dual Splitting Method for Convex Optimization Involving Lipschitzian, Proximable and Linear Composite Terms. *Journal of Optimization Theory and Applications*, 158(2) :460–479, August 2013.
- [8] Joan Duran, Bartomeu Coll, and Catalina Sbert. Chambolle’s Projection Algorithm for Total Variation Denoising. *Image Processing On Line*, 3 :311–331, December 2013.
- [9] Hadamard, J. Sur les Problèmes aux Dérivées Partielles et Leur Signification Physique. Number 13 in Princeton University Bulletin, pages 49–52. 1902.
- [10] Kenneth M. Hanson. Introduction to Bayesian image analysis. pages 716–731, Newport Beach, CA, September 1993.
- [11] Ibraheem Ibraheem and Waleed Al-Jawher. Noise Detection and Estimation Using Discrete Wavelet Transform. *AMSE*, 48, June 2005.
- [12] Laurent Jacques. *Ondelettes, repères et couronne solaire*. PhD thesis, Université catholique de Louvain, June 2004.
- [13] Jari Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*. Springer Science & Business Media, March 2006.
- [14] Michel Kern. Problèmes inverses : aspects numériques. *Engineering school. 1999 à 2002*, 2002.
- [15] Luis Vargas Miele. luisvargasmieles/2020-SIIMS-AcceleratingMCMCmethods. <https://github.com/luisvargasmieles/2020-SIIMS-AcceleratingMCMCmethods>, February 2021. commit : 7c5e009.
- [16] Marcelo Pereyra. Maximum-a-posteriori estimation with bayesian confidence regions, 2016.

- [17] Audrey Repetti, Marcelo Pereyra, and Yves Wiaux. Scalable Bayesian uncertainty quantification in imaging inverse problems via convex optimization. *arXiv :1803.00889 [astro-ph, stat]*, November 2018. arXiv : 1803.00889.
- [18] Audrey Repetti, Marcelo Pereyra, and Yves Wiaux. Uncertainty Quantification in Imaging : When Convex Optimization Meets Bayesian Analysis. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 2668–2672, Rome, September 2018. IEEE.
- [19] Alejandro Ribes and Francis Schmitt. Linear inverse problems in imaging. *IEEE Signal Processing Magazine*, 25(4) :84–99, July 2008.
- [20] Debasis Sengupta and Asok K. Nanda. Log-concave and concave distributions in reliability. *Naval Research Logistics (NRL)*, 46(4) :419–433, 1999.
- [21] R. Tolimieri, Myoung An, and Chao Lu. *Algorithms for Discrete Fourier Transform and Convolution*. Signal Processing and Digital Filtering. Springer New York, New York, NY, 1989.
- [22] Don van Ravenzwaaij, Pete Cassey, and Scott D. Brown. A simple introduction to Markov Chain Monte–Carlo sampling. *Psychonomic Bulletin & Review*, 25(1) :143–154, February 2018.
- [23] Luis Vargas, Marcelo Pereyra, and Konstantinos C. Zygalakis. Accelerating proximal Markov chain Monte Carlo by using an explicit stabilised method. *arXiv :1908.08845 [cs, math, stat]*, March 2020. arXiv : 1908.08845.

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
École polytechnique de Louvain

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/epl