

LA TOMOGRAPHIE PAR ÉMISSION DE POSITONS

SOMMAIRE

2.1	LA TEP EN MÉDECINE NUCLÉAIRE	25
2.2	PHYSIQUE DE L'IMAGERIE TEP	26
2.2.1	Isotopes radioactifs	26
2.2.2	Annihilation positron-électron	27
2.2.3	Interactions photons-matière	29
2.2.4	Atténuation	30
2.3	DÉTECTION PAR UN SYSTÈME TEP	30
2.3.1	Détection des photons d'annihilation	31
2.3.2	Electronique d'acquisition	33
2.3.3	Le tomographe	34
2.3.4	Stockage des événements	36
2.3.5	Mesure de l'atténuation	38
2.4	RECONSTRUCTION TOMOGRAPHIQUE	39
2.4.1	Un problème inverse mal posé	39
2.4.2	Maximisation de l'espérance	42
2.5	IMAGES TEP QUANTITATIVES	43
2.5.1	Correction d'atténuation	44
2.5.2	Correction des coïncidences fortuites, diffusées, multiples	44
2.5.3	Normalisation	45
2.5.4	Temps mort et calibration	46
2.5.5	Implémentation des corrections	46
2.5.6	Caractérisation des images reconstruites	47
	SYNTHÈSE ET CONCLUSION	47

La tomographie par émission de positons exploite la notion de traceur d'une activité métabolique et permet d'étudier les mécanismes physiologiques d'une structure fonctionnelle. Elle représente sous forme d'image tridimensionnelle la distribution spatiale du composé administré au patient. Cette technique d'imagerie moléculaire associe un système de détection, mesurant les projections des produits de désintégration du composé radioactif, à un algorithme de reconstruction tomographique, permettant de les localiser dans

l'espace. La technique est intrinsèquement quantitative mais est affectée d'incertitudes liées au système d'acquisition et aux algorithmes de reconstruction.

Dans ce chapitre, nous exposons le principe de la Tomographie par Emission de Positons (TEP) et ses fondements physiques (paragraphe 2.1 et 2.2). Nous décrivons la détection des produits de désintégration du radiotracer émetteur de positons administré au patient (paragraphe 2.3) par un système TEP. Nous montrons ensuite comment les techniques numériques de reconstruction tomographique permettent de localiser le radiotracer au moment de sa désintégration (paragraphe 2.4). Nous montrons enfin que des mesures quantitatives peuvent être réalisées sur images TEP (paragraphe 2.5) mais sont affectées d'incertitudes liées à la résolution spatiale limitée et au bruit présent dans les images.

2.1 LA TEP EN MÉDECINE NUCLÉAIRE

La médecine nucléaire est une discipline médicale dont John Lawrence est le pionnier avec ses travaux réalisés en 1936 sur le traitement d'une forme de leucémie au phosphore 32. Comme nous l'avons évoqué au paragraphe 1.3, elle repose sur la notion de radiotraceur. Les premières images reposant sur la détection du radiotraceur ont été réalisées sur des patients atteints d'un cancer de la thyroïde et traités à l'iode ¹³¹. Il s'agissait alors de compteurs Geiger-Müller déplacés manuellement devant le cou du patient. L'acquisition des données fut automatisée lors de la mise au point du scanner à balayage par Benedict Cassen dans les années 1950. En 1951, Hal Anger développa un système composé de cristaux scintillants, la caméra à scintillation. En 1958, Anger décrivit le concept de l'imagerie par émission de positons. La Tomographie par Emission Mono-Photonique (TEMP) naît en 1963 des travaux de David Köhl et Roy Edwards et rend possible l'imagerie par émission en trois dimensions à l'aide d'acquisitions de scintigraphie par un système en rotation autour du patient. Il fallut attendre 1975 avec les travaux de Michel Ter-Pogossian pour voir construit le premier système Tomographie par Emission de Positons (TEP). Les systèmes hybrides couplant les acquisitions TEP avec des acquisitions de Tomodensitométrie (TDM) (figure 2.1) ont été développés en 1998 par David Townsend, permettant d'associer une région métaboliquement active à une région anatomique. Les systèmes couplant acquisitions TEP et d'Imagerie par Résonance Magnétique (IRM) ont fait leur apparition en 2011.

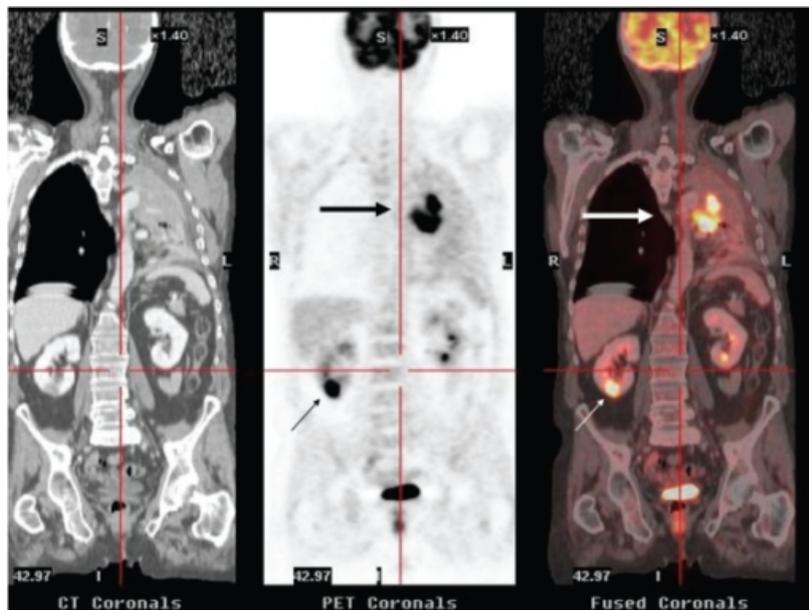


FIGURE 2.1 – Images TDM, TEP, TEP et TDM fusionnées d'un système hybride TEP-TDM (source : Pinilla et al. (2008)).

La Tomographie par Emission (TE) concerne l'ensemble des modalités d'imagerie utilisées en médecine nucléaire permettant de représenter par une image tridimensionnelle l'intérieur de l'organisme à partir de mesures externes à cet objet (tomographie, du grec *tomê* : coupe). Le terme d'émission provient du fait que le signal détecté est généré à l'intérieur du patient, par opposition

à l'imagerie de transmission pour laquelle le signal est généré à l'extérieur du patient puis modifié par celui-ci. Selon le radiotracteur utilisé, l'imagerie par émission est moléculaire ou cellulaire, elle est donc intrinsèquement fonctionnelle. De même que la TEMP, la TEP est une modalité de TE très utilisée en médecine nucléaire.

Le déroulement d'un examen de tomographie par émission de positons, dont le principe simplifié est schématisé sur la figure 2.2, se déroule en quatre étapes :

1. Le radiotracteur émetteur de positons est injecté au patient,
2. les positons issus de la désintégration des isotopes s'annihilent avec un électron de la matière, conduisant à l'émission d'une paire de photons à 511 keV dans des directions opposées,
3. les photons émis sont ensuite détectés par un tomographe composé de détecteurs élémentaires disposés autour du patient,
4. la mise en correspondance temporelle des photons détectés détermine un Tube De Réponse (TDR) le long desquels a eu lieu l'annihilation,
5. les TDR associés à de multiples annihilations sont stockés numériquement et un logiciel informatique implémentant un algorithme de reconstruction tomographique permet de retrouver les positions d'annihilation,

Le principe de la TEMP est similaire à celui de la TEP mais utilise des radiotraceurs émetteurs de photons simples (TEMP) collimatés mécaniquement.

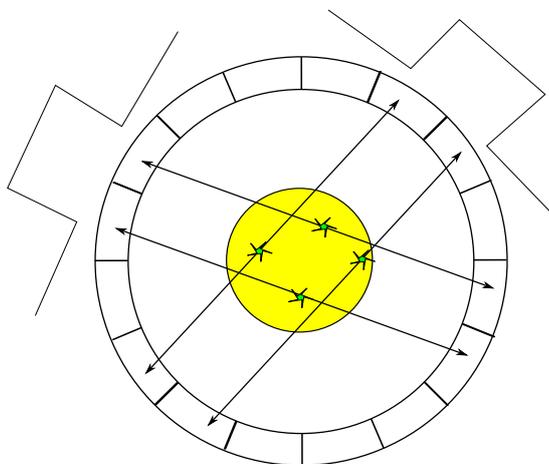


FIGURE 2.2 – Principe de la TEP.

Dans ce qui suit, nous décrivons les principes fondamentaux de la TEP. Pour plus de détails, nous renvoyons le lecteur à des ouvrages dédiés à cette modalité (Maisey *et al.* 2005, Cherry et Dahlbom 2006).

2.2 PHYSIQUE DE L'IMAGERIE TEP

2.2.1 Isotopes radioactifs

La TE repose sur un phénomène naturel par lequel des noyaux atomiques instables, présentant un excès d'énergie, retrouvent spontanément un état

stable en dégageant de l'énergie via l'émission d'un rayonnement à haute énergie. Ce phénomène, appelé désintégration radioactive, est un processus aléatoire suivant une loi de décroissance radioactive. Pendant un intervalle dt , la probabilité de désintégration suit approximativement une loi uniforme de paramètre λ homogène à une fréquence. En moyenne, le nombre de désintégrations se produisant dans un échantillon de $N(t)$ noyaux dans un intervalle de temps dt est donc égale à $-dN(t) \equiv N(t) - N(t + dt) = \lambda N(t)dt$. La résolution de cette équation différentielle conduit à l'expression de $N(t)$ en fonction de la taille initiale de l'échantillon $N(0)$ (équation 2.1a).

La radioactivité d'un échantillon est définie par le nombre de désintégrations se produisant par unité de temps, soit la quantité $-\frac{dN(t)}{dt}$. Elle s'exprime en Becquerel (Bq), équivalent à une désintégration par seconde, selon les unités du système international ou parfois en Curie ($1 \text{ Ci} = 3.7 \times 10^{10} \text{ Bq}$). Elle est entièrement déterminée par l'activité de l'échantillon au temps initial $A(0)$ et par λ (équation 2.1b) ou, de façon équivalente, par la demi-vie de l'isotope $t_{1/2}$. Cette grandeur correspond à la durée nécessaire pour qu'en moyenne la moitié des noyaux de l'échantillon se soit désintégrée.

$$N(t) = N(0)e^{-\lambda t} \text{ où } \lambda = \frac{\ln 2}{t_{1/2}} \quad (2.1a)$$

$$A(t) = -\frac{dN(t)}{dt} = \lambda N(0)e^{-\lambda t} = A(0)e^{-\lambda t} \quad (2.1b)$$

En TEP, on considère les isotopes émetteurs de rayonnements β^+ dont la désintégration provient de la conversion d'un neutron en un positon noté e^+ et un neutrino ν_e . En notant ${}^A_Z X$ l'isotope de numéro atomique Z et de nombre de masse A , cette réaction est décrite par l'équation 2.2. Le positon, de masse égale à celle de l'électron, est émis avec une énergie prenant un continuum de valeurs jusqu'à une énergie maximale E_{max} exprimée en joules (J) ou en électron-volt ($1 \text{ eV} \simeq 1.6 \times 10^{-19} \text{ J}$). Les noyaux émetteurs de positons peuvent également décroître par capture électronique selon l'équation 2.3.



Le mode de désintégration des isotopes considérés en TEP est principalement l'émission β^+ . Elle se produit avec une probabilité, ou rapport de branchement, proche de 1 pour ces isotopes.

2.2.2 Annihilation positron-électron

Après une désintégration β^+ , le positon émis traverse la matière et interagit par interactions coulombiennes avec le noyau et les électrons, induisant de larges déviations de son parcours. Il perd alors de l'énergie principalement par collisions avec les électrons. Lorsqu'il a perdu la plupart de son énergie cinétique, il s'apparie avec un électron pour former un positronium, peu stable (sa demi-vie est inférieure à 100 ns) et dont l'annihilation conduit à la formation de deux photons¹ (figure 2.3). Le positon se trouve alors à une certaine distance du lieu de désintégration dépendant de E_{max} et du milieu, suivant une loi de

1. Dans de très rares cas (moins de 0.003% des cas) l'annihilation aboutit à l'émission de trois photons (Cherry et Dahlbom 2006)

distribution approximativement exponentielle décroissante (Cherry et Dahlbom 2006) centrée en 0 et de moyenne d_{DA} . Le tableau 2.1 donne les valeurs de $t_{1/2}$, E_{max} , d_{DA} (dans l'eau) et du rapport de branchement des principaux isotopes utilisés en TEP.

Isotope	$t_{1/2}$ (min)	E_{max} (keV)	d_{DA} (mm)	RB
^{11}C	20.4	959	1.1	1
^{13}N	9.96	1197	1.5	1
^{15}O	2.03	1738	2.5	1
^{18}F	109.8	633	0.6	0.97
^{67}Ga	68.3	1898	2.9	0.89
^{82}Rb	1.25	340	5.9	0.96

TABLEAU 2.1 – Demi-vie $t_{1/2}$, énergie maximale du positon E_{max} , distance désintégration-annihilation d_{DA} (dans l'eau) et rapport de branchement (RB) des isotopes principalement utilisés en TEP (source : Bailey et al. (2005)).

Afin de satisfaire aux lois de conservation du moment cinétique et de l'énergie, provenant alors quasiment exclusivement de la masse du positon m_p et de l'électron m_e , les photons sont émis dans des directions opposées avec une même énergie égale à :

$$E = \frac{1}{2}(m_e + m_p)c^2$$

$$\simeq 8.2 \times 10^{-14} \text{ J} \simeq 511 \text{ keV}$$

où $m_e \simeq 9.1 \times 10^{-31} \text{ kg}$, $m_p = m_e$, $c \simeq 2.998 \times 10^8 \text{ m.s}^{-1}$ sont respectivement les masses de l'électron, du positon, et la vitesse de la lumière dans le vide.

L'électron et le positon n'étant par totalement au repos lors de l'annihilation, le moment cinétique total n'est pas nul et les photons sont émis avec un angle d'acolinéarité qui suit approximativement une distribution gaussienne de moyenne nulle et de Largeur à Mi-Hauteur (LMH) approximativement égale à 0.5 degrés, dépendant de l'énergie résiduelle du positon et indépendante de E_{max} (Cherry et Dahlbom 2006).

Les photons d'annihilation interagissent alors avec la matière mais les forces en jeu sont moins importantes que celles subies par les positons. Ils traversent ainsi facilement l'organisme et atteignent le détecteur avec une probabilité non négligeable.

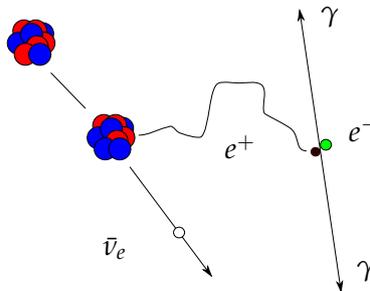


FIGURE 2.3 – Désintégration et annihilation d'un positon avec un électron.

2.2.3 Interactions photons-matière

Les photons d'annihilation interagissent essentiellement avec les électrons de la matière selon deux processus physiques dominants à 511 keV : l'effet photo-électrique et la diffusion Compton. La probabilité qu'a une interaction de se produire est donnée par la section efficace partielle σ , mesurée en cm^2 ou barn. Elle dépend de l'énergie du photon incident et du milieu de propagation. La probabilité d'interaction d'un photon dans un milieu de densité n est définie par le coefficient d'atténuation linéique μ (en cm^{-1}) :

$$\mu = n\sigma, \text{ avec } n = \frac{N_A}{m_a}\rho$$

où n , m_a , $N_A \simeq 6.02 \times 10^{23} \text{ mol}^{-1}$ et ρ sont respectivement la densité du milieu, la masse molaire atomique du milieu (en g.mol^{-1}), la constante d'Avogadro et la masse volumique du milieu (en g.cm^{-3}).

Effet photo-électrique

L'effet photo électrique se produit lorsque le photon transfère toute son énergie à un électron du cortège électronique d'un atome de la matière (figure 2.4a). L'électron est alors éjecté (et l'atome perd un électron : il est ionisé) avec une énergie égale à l'énergie du photon incident E_γ moins l'énergie de liaison de l'électron E_l (quelques dizaines de keV). La section efficace de cet effet dépend du numéro atomique du milieu Z et de l'énergie du photon incident E_γ par la relation approximativement valide autour de 511 keV (Levin 2004) :

$$\sigma_{pe} \propto \frac{Z^n}{E_\gamma^{3.5}}, n \in [3,4]$$

Ainsi, les détecteurs auront un numéro atomique élevé afin de maximiser l'effet photo-électrique.

Diffusion

Lorsque le photon interagit avec un électron libre ou faiblement lié de l'atome, seule une partie de l'énergie du photon incident est transférée à un électron de l'atome. Le photon est alors dévié avec un angle solide θ et une énergie E'_γ : c'est la diffusion Compton. La section efficace différentielle (pour un angle θ donné) de diffusion Compton d'un photon incident d'énergie E_γ est donnée par l'équation de Klein-Nishina :

$$\frac{d\sigma_{dc}}{d\Omega} = \frac{r_e^2}{2} \left(\frac{E_\gamma'}{E_\gamma} \right)^2 \left[\frac{E_\gamma}{E_\gamma'} + \frac{E_\gamma'}{E_\gamma} - \sin^2\theta \right] \quad (2.4)$$

où $r_e \simeq 2.82 \times 10^{-15}$ mètres est le rayon classique de l'électron et les rapports d'énergies entre photon émis et photon incident sont définis par l'équation de Coulomb :

$$\frac{E_\gamma'}{E_\gamma} = \frac{1}{1 + \frac{E_\gamma}{m_e c^2} (1 - \cos\theta)}$$

L'électron auquel le photon transfère une partie de son énergie est potentiellement éjecté du cortège électronique puis réabsorbé.

Un autre type de diffusion peut se produire avec un électron fortement lié à l'atome : la diffusion de Rayleigh. Ce mode de diffusion implique des déviations et pertes d'énergie faibles autour de 511 keV et est négligeable en TE (Levin 2004).

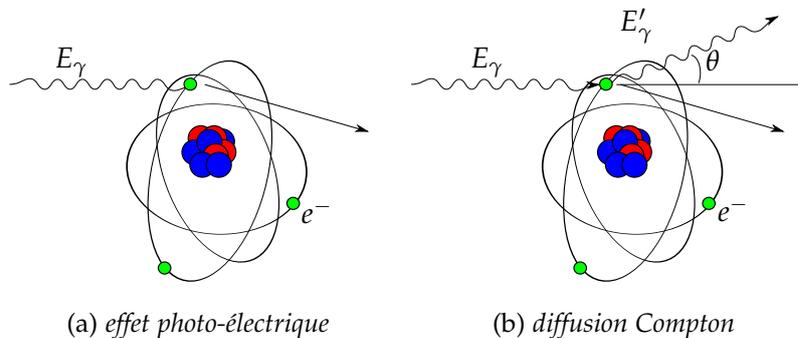


FIGURE 2.4 – Effets physiques dominants à 511 keV.

2.2.4 Atténuation

L'atténuation désigne la quantité de photons d'un faisceau collimaté qui sont absorbés ou déviés par un milieu. A 511 keV, cette quantité dépend de la probabilité d'interaction par effet photo-électrique ou diffusion Compton $\sigma_{tot} = \sigma_{pe} + \sigma_{dc}$. Ainsi, si $N(x + dx)$ photons d'un faisceau incident collimaté de $N(x)$ photons traversent un milieu absorbant d'épaisseur dx composé de densité n sans aucune interaction, alors $-dN(x) \equiv N(x) - N(x + dx)$ est égal à $N(x)n\sigma_{tot}(x)$. La résolution de cette équation différentielle conduit à l'expression 2.5.

$$N(X) = N(0)e^{-\int_0^X \mu_{tot}(x)dx}, \text{ où } \mu_{tot} = n\sigma_{tot} \quad (2.5)$$

La figure 2.5 été produite à partir de la base de données XCOM (Berger *et al.* 2010) et donne les coefficients d'atténuation massiques partiels μ/ρ dans l'eau en fonction de l'énergie incidente ainsi que l'atténuation massique totale. A 511 keV, on remarque que l'atténuation est principalement due à la diffusion Compton. Le tableau 2.2 présente l'atténuation résultante pour des milieux de dimension et composition variables. Ainsi, plus de 75% des photons de 511 keV traversent une distance de 3 cm d'eau sans être atténués alors que ce rapport vaut 38% pour 10 cm d'eau (dimension approximative d'un cerveau) et 5% pour 30 cm (épaisseur d'un abdomen). A l'inverse, 3 cm de GSO sont suffisants pour arrêter près de 90% des photons incidents.

2.3 DÉTECTION PAR UN SYSTÈME TEP

Tous les systèmes TEP reposent sur le même principe : détecter la droite selon laquelle l'annihilation a lieu. La différence de temps de parcours entre les deux photons détectés (appelée temps de vol) peut, en théorie, permettre de retrouver la position d'annihilation mais ne peut pas être mesurée avec

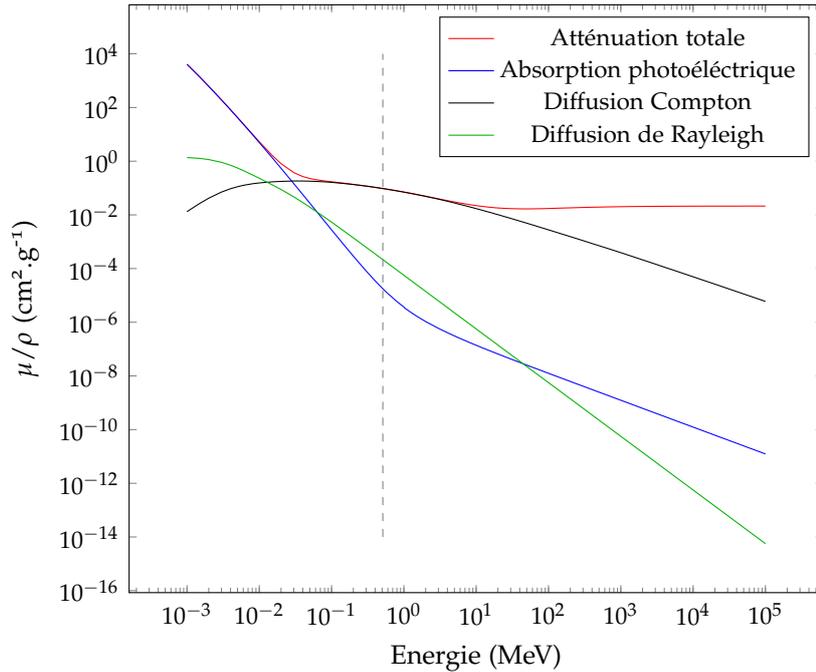


FIGURE 2.5 – Coefficients d'atténuation massiques partiels dans l'eau en fonction de l'énergie incidente (source : (Berger et al. 2010)).

Milieu	Composition	ρ (g/cm ³)	μ_{dc}/ρ (cm ² /g)	μ_{pe}/ρ (cm ² /g)	A(3 cm)	A(10 cm)	A(30 cm)
Eau	H ₂ O	1.0	0.09663	0.00002	24.2	62.0	94
BGO	Bi ₄ Ge ₃ O ₁₂	7.13	0.08121	0.01874	88.2	99.9	>99.99

TABLEAU 2.2 – Atténuation $A(X) = 100 \left(1 - e^{-(\mu_{pe} + \mu_{dc})X}\right)$ à 511 keV pour plusieurs épaisseurs X de milieu traversé en fonction de la composition, de la masse volumique ρ et des sections efficaces massiques du milieu à 511 keV associées à l'effet photoélectrique μ_{pe} et à la diffusion Compton μ_{dc} .

suffisamment de précision : une résolution temporelle aussi fine que 100 picosecondes conduirait à une résolution spatiale de 1.5 centimètres (Cherry et Dahlbom 2006) (en revanche, cette information peut être utilisée afin d'améliorer la résolution des systèmes la mesurant). Il s'agira donc de détecter les droites sur lesquelles ont lieu des annihilations, et d'utiliser un algorithme de reconstruction tomographique pour estimer la distribution spatiale 3D des lieux d'annihilation.

2.3.1 Détection des photons d'annihilation

Détection des photons à 511 keV

Les rayonnements γ sont détectés par des instruments mesurant les ionisations de la matière par interaction Compton ou effet photo-électrique. Ces systèmes sont majoritairement composés de cristaux à scintillation, qui ont la propriété d'émettre de la lumière dans le domaine visible lorsqu'une ionisation se produit. Les cristaux sont composés de matériaux inorganiques. Les principaux matériaux utilisés en TEP sont le germanate de bismuth (BGO) ou les oxyorthosilicates de lutécium (LSO) et de gadolinium (GSO). Ils sont choisis pour maximiser le rendement lumineux, leur atténuation à 511 keV et leur rap-

port des sections efficaces de l'effet photoélectrique et de la diffusion Compton et pour minimiser le temps de réponse à un dépôt d'énergie, leur résolution en énergie à 511 keV et le temps de décroissance d'une scintillation.

Contrairement aux détecteurs gazeux (Missimer *et al.* 2004), ils présentent de bonnes performances en terme de pouvoir d'arrêt et de résolution en énergie (Bailey *et al.* 2005). Les détecteurs à base de semi-conducteurs (Shiga *et al.* 2009) suscitent un intérêt croissant puisqu'ils présentent une excellente résolution en énergie. Leur pouvoir d'arrêt est cependant inférieur à celui des scintillateurs (Bailey *et al.* 2005).

Conversion en courant électrique

La lumière est ensuite convertie en signal électrique généralement par un réseau de Tubes Photo-Multiplicateurs (TPM). Un TPM convertit les photons en électrons sur une première dynode, puis, une succession de dynodes amplifient le signal électrique. En sortie du TPM, le courant électrique est ainsi amplifié et fournit un courant électrique mesurable. L'énergie déposée est calculée comme l'intégrale du signal électrique sur une période donnée. Le signal est alors traité par une carte électronique. Certains systèmes récents sont constitués de détecteurs à base de photodiodes en silicium permettant de mesures du temps de vol et de la profondeur d'interaction (Del Guerra *et al.* 2011, Seifert *et al.* 2013).

Positionnement des événements

La localisation de l'interaction à partir des dépôts d'énergie peut être réalisée en utilisant des matrices de cristaux de petite taille couplés à un réseau de TPM. Afin de réduire la taille des cristaux à quelques millimètres de côté, donc la résolution spatiale sans augmenter de façon majeure le coût du système (les TPM sont les éléments les plus coûteux), on utilise aujourd'hui un modèle en bloc où M petits cristaux sont taillés dans un cristal plus large et sont couplés à $N < M$ TPM. La figure 2.6) montre un exemple de bloc de 8 cristaux de côté. Des coupures à profondeurs variables sont faites entre les cristaux avec un matériau réfléchissant pour diffuser la lumière de façon optimisée aux TPMs. Les coordonnées du cristal sont déterminées par l'utilisation d'une table de correspondance pré-calculée associant une position de cristal à un barycentre des dépôts d'énergie (Cherry et Dahlbom 2006). Un guide de lumière ou des graisses optiques sont parfois interfacées entre le cristal et le TPM pour permettre une diffusion de la lumière efficace. L'usage d'un même TPM pour plusieurs cristaux, ou partage de lumière, induit cependant une perte de sensibilité puisque plusieurs interactions simultanées seront considérées comme une unique interaction localisée sur le barycentre des multiples dépôts.

Ce système permet de déterminer le cristal dans lequel un photon incident a déposé de l'énergie mais aucune information n'est donnée sur la profondeur d'interaction. Or, les cristaux ont généralement plusieurs centimètres de profondeur pour assurer un pouvoir d'arrêt satisfaisant. L'incertitude sur la profondeur d'interaction introduit des erreurs de localisation appelées effet de parallaxe. Une mesure de la profondeur d'interaction est possible sur certains tomographes et peut être réalisée, par exemple, grâce à un empilement de plusieurs matrices 2D de cristaux composés de différents matériaux

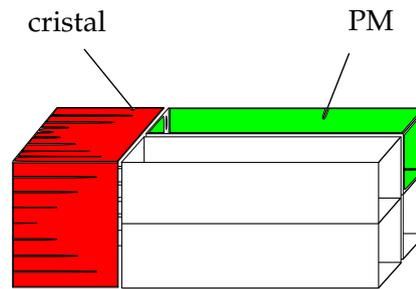


FIGURE 2.6 – Bloc de 8x8 cristaux couplés 4 tubes photo-multiplieurs.

de constantes de temps différentes ou des photodiodes à avalanche (Cherry et Dahlbom 2006).

2.3.2 Electronique d'acquisition

Lorsqu'un dépôt d'énergie suffisant est détecté, un signal binaire, appelé événement simple, est généré. Une fenêtre temporelle, appelée fenêtre de coïncidence, est alors déclenchée (environ 3 à 4 nanosecondes). Une coïncidence a lieu lorsqu'un deuxième dépôt d'énergie est détecté pendant la fenêtre de coïncidence, à condition que le TDR formé par les cristaux soit valide et que les énergies soient comprises dans la fenêtre en énergie du système (réduisant de ce fait les photons ayant diffusé et les phénomènes d'empilement) (Bailey 2005).

La procédure ci dessus permet donc de retrouver le TDR caractéristique de la coïncidence. Elle enregistre également un certain nombre de coïncidences parasites : les coïncidences diffusées, fortuites, multiples. Une coïncidence est dite diffusée lorsqu'au moins un photon a dévié de sa trajectoire initiale par une ou plusieurs diffusions Compton (figure 2.7b). Une coïncidence est dite fortuite lorsque deux photons issus de deux annihilations différentes sont détectées pendant la même fenêtre de coïncidence (figure 2.7c). Enfin, une coïncidence multiple est définie par plus de deux coïncidences détectées pendant la même fenêtre temporelle (figure 2.7d). Par opposition à ces coïncidences parasites, lorsque deux photons issus de la même annihilation sont détectés sans avoir été diffusés, on parle de coïncidence vraie (figure 2.7a). En négligeant le temps mort et les coïncidences multiples, les coïncidences fortuites augmentent proportionnellement au carré de l'activité dans le champ de vue, tandis que le nombre de coïncidences vraies et diffusées augmentent linéairement avec l'activité. Les coïncidences multiples sont de même nature que les fortuites et augmentent avec les taux d'événements simples. Ces coïncidences sont généralement éliminées. Comme il n'existe pas de moyen direct de discriminer les coïncidences vraies des coïncidences fortuites, dominantes aux hautes activités et des coïncidences diffusées survenant avec une probabilité élevée (voir paragraphe 2.2), il est nécessaire d'élaborer des stratégies permettant d'éliminer ces coïncidences parasites. Nous les présenterons dans le paragraphe 2.5.

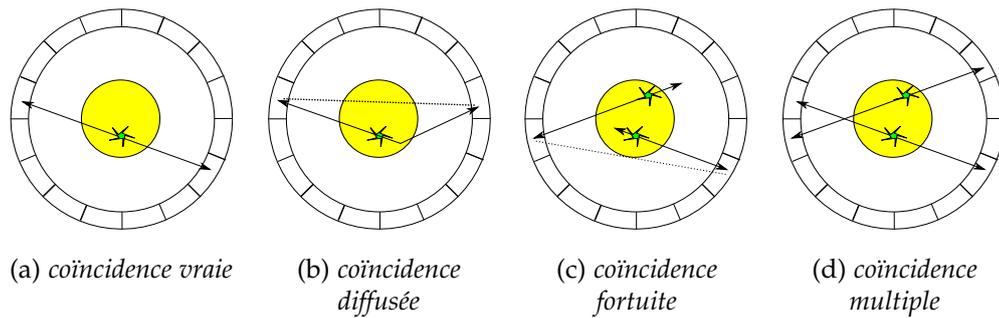


FIGURE 2.7 – Types de coïncidences détectées par un tomographe.

2.3.3 Le tomographe

De nombreux cristaux sont utilisés et arrangés en une ou plusieurs couronnes juxtaposées selon une géométrie cylindrique (figure 2.8) ou polygonale si une architecture en bloc est employée. Le cylindre inscrit mesure typiquement 80 à 90 cm de diamètre et 20 cm de long (Cherry et Dahlbom 2006).

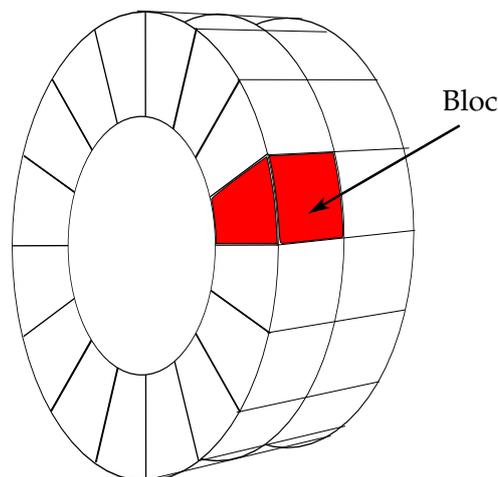


FIGURE 2.8 – Architecture d'un système TEP.

Cette géométrie présente cependant le désavantage d'introduire un échantillonnage radial irrégulier (figure 2.9a), appelé effet d'arc, induisant des hétérogénéités de performance selon la position de la source dans le détecteur.

Plus le nombre de TDRs intersectant l'élément de volume est élevé, plus la résolution spatiale sera élevée. En pratique, une distance minimale entre les cristaux est imposée pour qu'un TDR soit valide, limitant la dégradation de la résolution spatiale. Le maillage couvert par les TDR couvre alors un volume réduit appelé champ de vue.

Caractérisation des projections

Les projections sont issues de désintégrations radioactives aléatoires et deux acquisitions d'une même distribution d'activité donneront généralement des projections différentes. Les projections mesurées sont décrites par un vecteur aléatoire $\mathbf{g} = (g_i)_{i \in [1, n]}$ dont les composantes sont indépendantes et

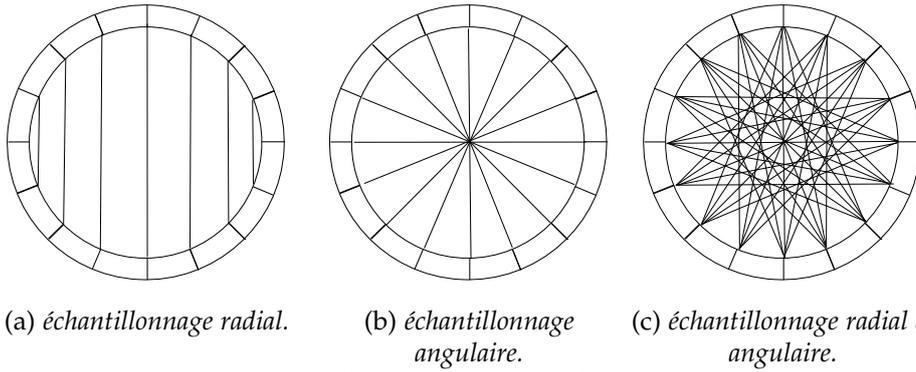


FIGURE 2.9 – Echantillonnage de l'espace par les tubes de réponse. Un écart minimum de 6 cristaux a été imposé sur la figure c.

suivent des lois de Poisson, décrivant la statistique de comptage. La probabilité d'observer g_i sur l'élément de détection i sera liée à son espérance par la relation 2.6 :

$$P(g_i) = \frac{E[g_i]^{g_i} e^{-E[g_i]}}{g_i!} \quad (2.6)$$

L'espérance $E[g_i]$ sera aussi affectée d'erreurs de mesures, ou biais, liés à l'atténuation par les tissus, à la contribution des coïncidences parasites et aux sensibilité et résolution spatiale limitées du système. La majorité de ces effets peut être corrigée (voir le paragraphe 2.5). En revanche, la résolution spatiale reste intrinsèquement limitée par la physique du signal mesuré (parcours du positon, acolinéarité des photons d'annihilation), les caractéristiques des détecteurs (effets de parallaxe) et la géométrie du tomographe.

En supposant que les effets sont modélisables par une opération de convolution de la source par des fonctions gaussiennes indépendantes, la Fonction de Dispersion du Point (FDP) peut être approchée par une fonction gaussienne et les LMH sommées en quadrature. Ainsi, la résolution spatiale intrinsèque Γ au centre du champ de vue d'un tomographe de géométrie cylindrique de diamètre D composé de cristaux dont la face intérieure au cylindre est un carré de côté d et une source de fluor 18 sera :

$$\Gamma = \sqrt{\Delta_{nc}^2 + \Delta_{pos}^2 + \Delta_{det}^2}$$

où Δ_{nc} , Δ_{pos} , Δ_{det} sont les LMH associées aux effets d'acolinéarité, de parcours du positon et valent approximativement (la source placée à proximité du centre du champ de vue) :

$$\begin{aligned} \Delta_{nc} &\simeq \tan(0.25 \text{ deg}) \times 0.5 \times D \\ \Delta_{pos} &\simeq 0.102 \text{ mm pour le fluor 18} \\ \Delta_{det} &\simeq \frac{d}{2} \end{aligned}$$

ce qui conduit à une résolution spatiale minimum de 2.7 mm environ pour $D=80$ cm et $d=4$ mm.

L'échantillonnage non régulier de l'espace par les TDRs implique que la résolution spatiale dépend de la position de la source dans le champ de vue, et de la direction radiale, tangentielle ou axiale dans laquelle elle est mesurée. Elle est dite non stationnaire et non isotrope.

2.3.4 Stockage des événements

A ce stade, nous pouvons présenter un premier modèle de la mesure. Dans le paragraphe ci-dessous, nous omettons l'atténuation, que nous introduisons au paragraphe 2.3.5.

Modèle de la mesure

Notons f une fonction continue de \mathbb{R}^3 représentant la concentration d'activité dans l'objet à imager et g les données mesurées par le tomographe, supposé linéaire. Nous considérons ici que f et g sont des fonctions déterministes. Si on néglige le parcours du positon, l'acolinéarité des photons, et la sensibilité limitée du tomographe, le signal détecté par une paire de cristaux (C_1, C_2) est l'intégrale volumique de f suivant le parallélépipède formé par les deux cristaux. Si, de plus, nous supposons que les photons d'annihilation ont été émis le long de la droite joignant les centres des faces des cristaux, notés X_1 et X_2 , la projection se réduit à une intégrale linéique et le TDR est appelé Ligne De Réponse (LDR). Ce modèle est appelé le modèle ligne-intégrale.

Soit $\mathcal{B} = (O, \mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z)$ une base orthonormale de l'espace telle que l'axe (O, \mathbf{u}_z) est confondu avec l'axe du tomographe et $\mathbf{u} = \frac{\overrightarrow{X_1 X_2}}{\|X_1 X_2\|}$ un vecteur directeur de $(X_1 X_2)$. On définit une nouvelle base $\mathcal{B}'' = (O, \mathbf{u}_{x''}, \mathbf{u}_{y''}, \mathbf{u}_{z''})$ par rotations successives de \mathcal{B} autour de \mathbf{u}_z et $\mathbf{u}_{x'}$ d'angle azimutal $\phi \in [-\frac{\pi}{2}, +\frac{\pi}{2}]$ et copolaire $\theta \in [-\Theta_{max}, +\Theta_{max}]$ de sorte que $\mathbf{u}_{y''} = \mathbf{u}$ (figure 2.10). La matrice de passage de \mathcal{B} à \mathcal{B}'' , notée $R_{\phi, \theta}$, est donnée par l'expression 2.7 et \mathbf{u} est alors paramétré par les angles ϕ et θ . Nous noterons ce vecteur $\mathbf{u}(\phi, \theta)$.

$$R_{\phi, \theta} \equiv \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\cos \theta \sin \phi & \cos \theta \cos \phi & \sin \theta \\ \sin \theta \sin \phi & -\sin \theta \cos \phi & \cos \theta \end{pmatrix} \quad (2.7)$$

Soit $\mathbf{s} = (s_{x''}, s_{z''})$ la projection de (X_1, X_2) dans le plan $(O, \mathbf{u}_{x''}, \mathbf{u}_{z''})$. La projection de f suivant (X_1, X_2) peut s'écrire :

$$p(\mathbf{s}, \mathbf{u}(\phi, \theta)) = \int_{-\infty}^{+\infty} f(\mathbf{s} + \lambda \mathbf{u}(\phi, \theta)) d\lambda, \quad \mathbf{s} \in \mathbf{u}(\phi, \theta)^\perp \quad (2.8)$$

correspondant à la transformée en rayons X de f selon $(X_1 X_2)$.

Lorsque $\theta = 0$, cette expression se réduit à :

$$p(\mathbf{s}, \mathbf{u}(\phi, 0)) = \int_{-\infty}^{+\infty} f(s_{x''}(\cos \phi \mathbf{u}_x + \sin \phi \mathbf{u}_y) + s_{z''} \mathbf{u}_z) + \lambda(-\sin \phi \mathbf{u}_x + \cos \phi \mathbf{u}_y) d\lambda \quad (2.9)$$

Il s'agit de la projection dans un plan transaxial (ou plan droit) situé à une distance signée $s_{z''}$ du centre du tomographe et correspondant à la transformée de Radon 2D de f selon la droite de vecteur directeur $\mathbf{u}(\phi, 0)$ localisée à une distance $s_{x''}$ de l'axe du tomographe (figure 2.11).

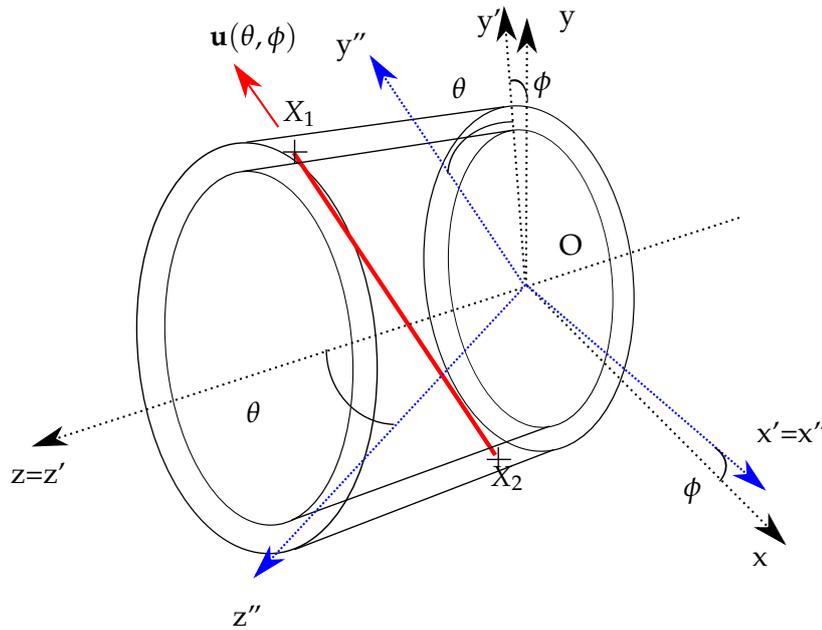


FIGURE 2.10 – Paramétrage 3D d'une ligne de réponse.

Le sinogramme

Suivant le modèle ligne-intégrale ci-dessus, les événements sont généralement stockés sous forme de projections parallèles à \mathbf{u} et indexées par $\mathbf{s} = (x'', z'')$, ϕ et θ . Les projections d'un point localisé en (x_0, y_0, z_0) ainsi stockées vaudront 1 pour :

$$\begin{pmatrix} x'' \\ \lambda \\ z'' \end{pmatrix} = R_{\phi, \theta} \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

Ainsi, à \mathbf{s} et θ constants, x'' et z'' sont des combinaisons linéaires de $\cos \phi$ et $\sin \phi$. Le sinogramme d'un point est donc une sinusoïde, d'où le terme de sinogramme.

Une variante de ce mode de stockage consiste à indexer la projection par ses coordonnées x' , z' , ϕ et θ dans la base non orthogonale $(0, \mathbf{u}_{x'}, \mathbf{u}_z, \mathbf{u}(\phi, \theta))$. On appelle ce format un sinogramme oblique. x' et z' sont appelées coordonnées radiales et axiales de la LDR et les plans de projections de mêmes coordonnées θ et z' sont appelés plans obliques dont une représentation est donnée sur la figure 2.12.

Trois coordonnées, par exemple x', z' et ϕ , suffisent à parcourir l'espace. Les données sont dites suffisantes. L'introduction du paramètre θ ajoute des données non nécessaires à la reconstruction mais réduit le bruit. Les données sont dites redondantes. En revanche, la longueur finie du tomographe induit une troncature des données d'autant plus sévère que y'' et θ sont grands.

Dans un format sinogramme classique, les projections sont interpolées radialement afin d'échantillonner uniformément l'espace. Cela induit une perte de résolution et modifie la nature du bruit dans les données. Par opposition, nous appellerons sinogramme brut un format sinogramme pour lequel aucune interpolation radiale n'est effectuée.

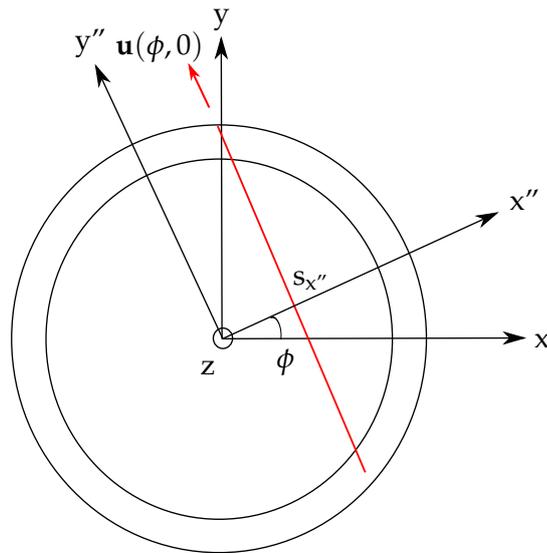
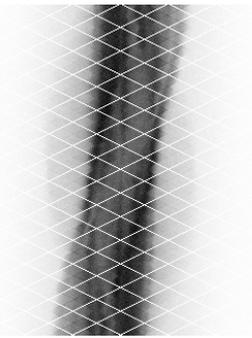


FIGURE 2.11 – Paramétrage 2D d'une ligne de réponse.

FIGURE 2.12 – Plan oblique d'un sinogramme (les positions radiale x'' et azimutale ϕ varient respectivement horizontalement et verticalement).

Format mode liste

Le format mode liste est un mode de stockage des coïncidences sous forme de liste contenant les coordonnées des détecteurs et les informations d'énergie et de temps associées à chaque coïncidence. C'est un format qui permet de gagner en efficacité de stockage lorsque le nombre de coïncidences est relativement faible et de conserver l'information temporelle, ce qui peut être utile quand on veut mesurer le temps de vol ou synchroniser les données aux mouvements respiratoires et cardiaques. De plus, ce format conserve la nature du bruit dans les données puisqu'aucune interpolation n'est faite.

2.3.5 Mesure de l'atténuation

Afin de prendre en compte l'atténuation par les tissus, il est nécessaire de disposer d'une carte des coefficients d'atténuation à 511 keV. Ces coefficients sont mesurés au moyen d'une acquisition en transmission. Nous explicitons ici les moyens de réaliser ces acquisitions et verrons au paragraphe 2.5 comment utiliser ces données. Avant que les systèmes hybrides TEP-TDM et TEP-IRM ne soient commercialisés, la mesure d'atténuation se faisait le plus souvent

au moyen d'une source externe de germanium 68 émettrice de positons ou de césium 137 émettrice de photons à 662 keV. Deux acquisitions étaient alors réalisées : la première sans et la deuxième avec le patient. Pour chaque LDR un facteur correctif égal au rapport du nombre de coïncidences entre les deux acquisitions est calculé. Le logarithme de ce rapport donne le facteur multiplicatif décrivant l'impact de l'atténuation sur la LDR.

Les acquisitions TDM réalisées en même temps que les acquisitions TEP dans les systèmes hybrides TEP-TDM fournissent un moyen direct de mesurer de l'atténuation par les tissus. L'acquisition TDM fournit les coefficients d'atténuation d'un faisceau de rayons X à environ 70 keV exprimés en unités Hounsfield $HU = 1000 \times \frac{\mu - \mu_{eau}}{\mu_{eau}}$. Les images sont converties en coefficients à 511 keV à l'aide d'une relation bilinéaire entre les coefficients aux différentes énergies (Kinahan *et al.* 1998).

Les nouveaux systèmes hybrides TEP-IRM actuellement commercialisés nécessitent d'adapter la mesure de l'atténuation puisqu'il n'existe pas de conversion directe entre le signal mesuré par l'IRM et les coefficients d'atténuation à 511 keV.

2.4 RECONSTRUCTION TOMOGRAPHIQUE

Le calcul des positions des annihilations à partir des projections mesurées se fait à l'aide d'un logiciel de reconstruction tomographique. Les positions d'annihilation sont représentées sous forme d'images tridimensionnelles, composées généralement de voxels (de l'anglais VOlumatic ELements par analogie avec le terme PICTURE ELement ou pixel en deux dimensions). Ainsi, après correction de la décroissance radioactive et du parcours du positon (voir le paragraphe 2.5), ces programmes permettent d'estimer la distribution du radiotracteur au moment de l'acquisition. Le problème de reconstruction tomographique est une branche des mathématiques appliquées très vaste dont nous synthétisons ci-dessous les applications en TEP. Pour plus de détails, le lecteur pourra se référer aux ouvrages de Defrise *et al.* (2005) et Grangeat (2009).

2.4.1 Un problème inverse mal posé

Reconstruction analytique

On distingue deux classes de méthodes de reconstruction tomographique : les méthodes analytiques, formulées dans un cadre continu et les méthodes discrètes. Les hypothèses sur lesquelles reposent les méthodes analytiques sont celles définies au paragraphe 2.3 : f et g sont supposées continues et déterministes, et on considère un modèle ligne-intégrale pour les projections.

Historiquement, les méthodes analytiques étaient utilisées avec des données acquises en mode 2D ou, afin d'augmenter le nombre de coïncidences, en autorisant une faible variation angulaire et en ré-échantillonnant les données par des méthodes de ré-échantillonnage en plans droits par exemple (méthode Single Slice ReBinning, FOurier REbinning). Ainsi, les premières méthodes analytiques consistaient à inverser la transformé de Radon 2D (équation 2.9). Dans la suite de ce paragraphe, z est fixé et on notera simplement f_z et p_z les fonctions $(x, y) \rightarrow f(x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z)$ et $(x'', \phi) \rightarrow p(x''\mathbf{u}_{x''} + z\mathbf{u}_z, \mathbf{u}(0, \phi))$.

Les reconstructions par rétroprojection filtrée reposent sur le théorème de la coupe centrale. Si on note F et P la transformée de Fourier 2D de f_z et la transformée de Fourier 1D de p_z selon la première variable :

$$\begin{aligned}
 P(v_{x'}, \phi) &= \int_{-\infty}^{+\infty} p_z(x', \phi) e^{-i2\pi x' v_{x'}} dx' \\
 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_z(x', y') e^{-i2\pi(x' v_{x'})} dx' dy' \\
 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_z(x', y') e^{-i2\pi(x \cos \phi + y \sin \phi) v_{x'}} dx' dy' \\
 &= F(v_{x'} \cos \phi, v_{x'} \sin \phi)
 \end{aligned}$$

Ainsi, le théorème de la coupe centrale énonce que la transformée 1D des projections selon un angle ϕ est égale à la section de la transformée 2D de f_z passant par l'origine et d'angle ϕ .

Une inversion de la transformée de Fourier par une méthode de type transformée de Fourier rapide nécessite cependant un échantillonnage régulier des paramètres. Ceci peut être fait en exprimant v_x et v_y selon des coordonnées polaires v et ϕ : $v_x = v \cos \phi$ et $v_y = v \sin \phi$. Une expression de f_z peut alors être dérivée :

$$\begin{aligned}
 f_z(x, y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(v_x, v_y) e^{i2\pi(xv_x + yv_y)} dv_x dv_y \\
 &= \int_0^{2\pi} \int_{-\infty}^{+\infty} F(v \cos \phi, v \sin \phi) e^{i2\pi(xv \cos \phi + yv \sin \phi)} v dv d\phi \\
 &= \int_0^{\pi} \int_{-\infty}^{+\infty} P(v_{x'}, \phi) e^{i2\pi(xv_{x'})} |v| dv d\phi \\
 &= \int_0^{\pi} p_z(x', \phi) * h(x') d\phi \text{ avec } h(x') = \int_{-\infty}^{+\infty} |v| e^{2\pi i x' v} dv
 \end{aligned}$$

où $*$ est le produit de convolution selon la variable x' . Cette opération s'appelle une rétroprojection filtrée, Filtered BackProjection en anglais (FBP), puisqu'elle somme les contributions pour tous les angles des projections préalablement filtrées. Le bruit haute fréquence est amplifié par le filtre rampe h , mettant en évidence le caractère mal posé du problème. Un filtrage supplémentaire est donc généralement réalisé à l'aide d'une fenêtre d'apodisation (par exemple filtre de Hamming en 2D).

En pratique, les projections sont mesurées selon des variables x' et ϕ échantillonnées selon un maillage irrégulier en x' , comme le montre la figure 2.9. Les projections sont d'abord filtrées avec une convolution discrète par le filtre h préalablement échantillonné sur le même maillage. Puis elles sont interpolées sur un maillage régulier en x' . Si f est représentée par une image échantillonnée régulièrement en x et en y , l'élément pour chacune des valeurs désormais prises par $x' = x \cos \phi + y \sin \phi$ est égal à la somme des projections sur tous les échantillons de ϕ .

Le théorème de la coupe centrale et la rétroprojection filtrée sont généralisables en 3D par l'opération $f(\mathbf{r}) = \int_{\Omega} \int_{\mathbf{u}^{\perp}} p(\mathbf{s}', \mathbf{u}) h(\mathbf{s} - \mathbf{s}', \mathbf{u}) d\mathbf{s}' d\mathbf{u}$ où h est, par exemple, le filtre de Colsher. L'intégration se fait alors sur toutes les directions $\mathbf{u} \in \Omega$ telles que le plan perpendiculaire à \mathbf{u} contient des données complètes (ce sont les conditions de Orlov). Les plans pour lesquels les données sont

incomplètes peuvent être reconstruits en estimant les données manquantes à partir d'une reconstruction des plans droits reprojctée dans l'espace des sinogrammes. Cet algorithme est appelé reprojction (Kinahan et Rogers 1989).

Méthodes discrètes

Les méthodes analytiques ont l'inconvénient de poser des hypothèses fortes sur la nature des données et la géométrie du système. A l'inverse, les méthodes discrètes permettent une représentation plus flexible de l'acquisition grâce à une formulation générique sous forme matricielle :

$$\mathbf{g} = \mathbf{H}\mathbf{f} \quad (2.10)$$

où \mathbf{f} et \mathbf{g} sont les formulations discrètes de la distribution du traceur f et des projections p . Ce sont des vecteurs aléatoires de dimensions respectivement égales au nombre de voxels dans l'image n et au nombre d'éléments de projections m . \mathbf{H} est la réponse impulsionnelle du tomographe de dimension $m \times n$. Un élément $h_{i,j}$ de \mathbf{H} peut être vu comme la probabilité qu'une annihilation ayant eu lieu dans un voxel j donne lieu à un signal détecté par le détecteur élémentaire i . Le problème consiste alors à retrouver \mathbf{f} connaissant \mathbf{g} .

Ce problème inverse peut être résolu avec les méthodes algébriques, négligeant alors la nature stochastique du signal. Cependant, la matrice \mathbf{H} n'est généralement pas carrée donc n'est pas inversible. Une solution consisterait à employer la solution inverse généralisée mais la nature bruitée des données rend la solution très instable mettant de nouveau en évidence le caractère mal posé du problème. On cherche donc généralement à résoudre un problème d'optimisation dont le principe est de minimiser itérativement une fonction de coût. Parmi ces techniques d'optimisation, figurent les techniques algébriques pures, ne faisant aucune hypothèse sur la nature du bruit dans les données, comme les méthodes ART et SIRT. Les méthodes statistiques à l'inverse comprennent un modèle de bruit et consistent généralement à maximiser la probabilité a posteriori.

D'après le théorème de Bayes, la probabilité a posteriori de \mathbf{f} sachant qu'on observe \mathbf{g} , notée $P(\mathbf{f}|\mathbf{g})$, est liée à la fonction de vraisemblance $P(\mathbf{g}|\mathbf{f})$ et aux probabilités marginales $P(\mathbf{f})$ et $P(\mathbf{g})$ par la relation :

$$P(\mathbf{f}|\mathbf{g}) = \frac{P(\mathbf{g}|\mathbf{f})P(\mathbf{f})}{P(\mathbf{g})}$$

Si on suppose que \mathbf{f} est déterministe et que $P(\mathbf{g})$ suit une distribution uniforme, alors maximiser $P(\mathbf{f}|\mathbf{g})$ revient à maximiser la fonction de vraisemblance de \mathbf{f} sachant qu'on observe \mathbf{g} , ou de façon équivalente la log-vraisemblance :

$$\mathcal{L}(\mathbf{f}) \equiv \log(P(\mathbf{g}|\mathbf{f}))$$

Le critère des moindres carrés et des moindres carrés pondérés sont des cas particuliers de maximisation de la vraisemblance lorsque le bruit est supposé gaussien et les algorithmes à base de gradient produisent des solutions minimisant ces critères. L'approximation gaussienne n'est pas toujours une bonne approximation pour la TE et on préfère un modèle de distribution selon une loi de Poisson.

2.4.2 Maximisation de l'espérance

Algorithme de maximisation de l'espérance

L'algorithme maximisation de l'espérance, Expectation Maximization en anglais (EM) est un algorithme itératif proposé par Dempster *et al.* (1977) permettant de trouver le maximum de vraisemblance. Il a été appliqué à la TE par Shepp et Vardi (1982) et repose sur l'usage d'une variable aléatoire intermédiaire $\mathbf{s} = (s_{ij})$ suivant une loi de Poisson et représentant le nombre total d'émissions dans le voxel i détectées par l'élément de détection j dont l'espérance vaut $E[s_{i,j}] = h_{ij}f_j$:

$$P(\mathbf{s}|\mathbf{f}) = \prod_{i,j} \frac{E[s_{i,j}]^{s_{i,j}} e^{-E[s_{i,j}]}}{s_{i,j}!} \text{ avec } E[s_{i,j}] = h_{ij}f_j$$

L'algorithme EM consiste à maximiser l'espérance de $\mathcal{L}(\mathbf{f}) \equiv \log(P(\mathbf{s}|\mathbf{f}))$ sachant qu'on observe \mathbf{g} , par l'expression itérative :

$$\begin{aligned} \mathbf{f}^{(n+1)} &= \text{Argmax} Q(\mathbf{f}, \mathbf{f}^{(n)}) \\ Q(\mathbf{f}, \mathbf{f}^{(n)}) &= E \left[\mathcal{L}(\mathbf{f}) | \mathbf{g}, \mathbf{f}^{(n)} \right] \end{aligned}$$

$$\begin{aligned} Q(\mathbf{f}, \mathbf{f}^{(n)}) &= \sum_{i,j} p_{ij} \log(h_{ij}f_j) - h_{ij}f_j - E[\log(s_{ij}!) | \mathbf{g}, \mathbf{f}^{(n)}] \\ \text{avec } p_{ij} \equiv E[s_{i,j} | \mathbf{g}, \mathbf{f}^{(n)}] &= \frac{h_{ij}f_j^{(n)}}{\sum_k h_{ik}f_k^{(n)}} g_i \end{aligned}$$

$\mathbf{f}^{(n+1)}$ est ensuite obtenue simplement en annulant la dérivée de Q :

$$\begin{aligned} \frac{\partial Q(\mathbf{f}, \mathbf{f}^{(n)})}{\partial f_j} &= \sum_i \frac{p_{ij}}{f_j} - h_{ij} \\ f_j^{(n+1)} &= \frac{f_j^{(n)}}{\sum_i h_{ij}} \sum_i h_{ij} \frac{g_i}{\sum_k h_{ik}f_k^{(n)}} \end{aligned}$$

ou de façon synthétique :

$$\mathbf{f}^{(n+1)} = \frac{\mathbf{f}^{(n)}}{\mathbf{H}^t \mathbf{1}} \mathbf{H}^t \frac{\mathbf{g}}{\mathbf{H} \mathbf{f}^{(n)}} \quad (2.11)$$

où les opérations de multiplications et divisions vectorielles se font point à point et les étapes de multiplications par \mathbf{H} (projection) et \mathbf{H}^t (rétroprojection) sont des multiplications matricielles classiques. $\mathbf{f}^{(n)}$ est initialisée arbitrairement. Outre la tomographie, cet algorithme est utilisé dans de nombreux domaines : classification de données, apprentissage automatique, vision artificielle.

Accélération par sous-ensembles

L'algorithme maximisation de l'espérance par sous-ensembles ordonnés, Ordered Subset Expectation Maximization en anglais (OSEM) est une méthode introduite par (Hudson et Larkin 1994), utilisée principalement en imagerie médicale. Elle consiste à réaliser les opérations de projection et rétroprojection sur S sous-ensembles $E_s, s \in [1, \dots, S]$ selon l'algorithme :

$$\begin{aligned} f_j^{(n,s+1)} &= \frac{f_j^{(n,s)}}{\sum_{i \in E_s} h_{ij}} \sum_{i \in E_s} h_{ij} \frac{g_i}{\sum_k h_{ik} f_k^{(n,s)}} \\ f_j^{(n+1,1)} &= \frac{f_j^{(n,S)}}{\sum_{i \in E_s} h_{ij}} \sum_{i \in E_s} h_{ij} \frac{g_i}{\sum_k h_{ik} f_k^{(n,S)}} \end{aligned} \quad (2.12)$$

où $f_j^{(n,s)}$ est l'élément j de l'estimée de \mathbf{f} au sous-ensemble s de l'itération n .

Alors que l'algorithme EM converge vers le maximum de vraisemblance, l'algorithme OSEM n'est en général pas convergent. Une possibilité est alors d'inclure un paramètre de relaxation et de restreindre le nombre de sous-ensembles à être égal au nombre de projections. La convergence de cette méthode appelée Row Action Maximum Likelihood Algorithm en anglais (RAMLA) a été démontrée (Browne et De Pierro 1996).

Matrice système

La mise en œuvre de l'algorithme ci dessus nécessite de disposer de la matrice \mathbf{H} , modélisant la réponse du tomographe. Cette matrice est de taille $m \times n$, m étant le nombre de projections et n le nombre de voxels dans l'image. Elle peut être calculée une fois pour toute et stockée ou calculée à la volée pendant la reconstruction. La deuxième approche est généralement employée. Le choix de cette matrice est crucial puisque son réalisme détermine la qualité de la reconstruction de même que le temps de calcul. Les techniques les plus simples consistent à utiliser une opération de projection géométrique du voxel j vers l'élément de détection i . Deux stratégies sont possibles pour définir le projecteur géométrique : l'approche orientée voxel, où la projection du voxel sur le détecteur est calculée et les approches par lancer de rayon, où le détecteur est projeté sur l'objet. La probabilité h_{ij} est dans ce cas calculée comme l'intégrale de l'intersection entre la droite joignant les centres des faces des cristaux et le voxel (Siddon 1985). Des modèles plus élaborés de \mathbf{H} sont disponibles sur les systèmes récents et reposent sur les techniques de modélisation de la résolution. Nous les présenterons au paragraphe 5.3.

2.5 IMAGES TEP QUANTITATIVES

Le terme quantification est employé pour désigner l'extraction de valeurs numériques interprétables par l'observateur. Il s'agit donc dans un premier temps de produire des images dont les pixels représentent une quantité d'activité absolue, exprimée par exemple en Bq par unité de volume. L'étape ultime consistera en l'extraction de paramètres physiologiques comme la perfusion sanguine. Cette étape dépend essentiellement de la tâche applicative et nous nous limiterons dans ce paragraphe aux stratégies permettant de produire

des images reflétant une activité absolue. Pour avoir un aperçu des méthodes employées selon la tâche applicative, le lecteur pourra consulter l'ouvrage de Zaidi *et al.* (2006). Nous présenterons les méthodes d'analyse quantitative des images dans le contexte de l'évaluation de l'atteinte athéromateuse dans les chapitres 3 à 5 de ce manuscrit.

2.5.1 Correction d'atténuation

Comme nous l'avons vu au paragraphe 2.3.5, les photons subissent un phénomène d'atténuation variable suivant la quantité de matière traversée de sorte que les photons émis en surface d'un milieu sont détectés en quantité plus importante que les photons émis au centre de l'objet, faisant apparaître artificiellement un contraste dans les régions homogènes. Pour une LDR donnée, l'atténuation subie par les photons ne dépend pas des lieux d'annihilation. En effet, une coïncidence est détectée si aucun des deux photons d'annihilation n'a interagi avec la matière. En négligeant les variations de sensibilité des détecteurs, le nombre $N(C_i, C_j)$ de coïncidences détectées par des cristaux C_i et C_j de coordonnées x_i et x_j issues de N_0 annihilations en x_0 vaudra :

$$\begin{aligned} N(C_i, C_j) &= N_0 e^{+\int_{x_0}^{x_i} \mu(x) dx} e^{-\int_{x_0}^{x_j} \mu(x) dx} \\ &= N_0 e^{-\int_{x_i}^{x_j} \mu(x) dx} \end{aligned}$$

Ainsi, comme nous l'avons vu au paragraphe 2.2, le nombre de coïncidences vraies détectées provenant d'une distribution homogène est réduit de 24% lorsque l'épaisseur du médium traversé (composé d'eau) vaut 3 cm tandis que moins de 6% des photons traversent 30 cm de ce médium. Pour chaque paire de détecteurs (C_i, C_j) , une fois déterminés les coefficients d'atténuation, ceux ci pourront être utilisés pour compenser les valeurs mesurées de l'atténuation, en multipliant la valeur mesurée par l'inverse du facteur d'atténuation affectant la LDR ou en les intégrant à la matrice système (paragraphe 2.4.2). Le problème revient donc à utiliser une carte de coefficients d'atténuation à 511 keV, généralement fournie par une acquisition en transmission (voir le paragraphe 2.3.5) bien qu'il existe des méthodes approximatives de calcul simultané ou séquentiel des coefficients d'atténuation et de la distribution d'activité sans données complémentaires (Zaidi et Hasegawa 2006).

2.5.2 Correction des coïncidences fortuites, diffusées, multiples

Si on néglige la contribution des coïncidences vraies et diffusées, les coïncidences fortuites R_{ij} détectées par une paire de cristaux données (C_i, C_j) pendant une fenêtre temporelle de longueur τ peuvent être estimées en fonction des taux d'événements simples $(r_i(t), r_j(t))$ des deux cristaux et de la largeur de la fenêtre par l'expression $R_{ij}(t) = 2\tau r_i(t)r_j(t)$. Cette expression est applicable lorsque le nombre de coïncidences vraies et diffusées est dominé par le nombre de coïncidences fortuites. Une première méthode de correction consiste donc à intégrer les taux d'événements simples sur la durée totale de l'acquisition où le nombre de coïncidences fortuites est généralement dominant et en déduire une expression moyenne égale à $2\tau S_i S_j$, S_i et S_j sont les taux d'événements simples des deux détecteurs. Une alternative permettant

de conserver la dynamique de ce taux consiste à déclencher une fenêtre retardée (après la fenêtre en coïncidence), lors de l'occurrence d'un événement simple, pendant laquelle les coïncidences détectées sont supposées être décorréélées temporellement, donc fortuites. Parmi les méthodes ci-dessus, la technique par fenêtre retardée est la plus exacte et la plus couramment employée (Meikle et Badawi 2005). En revanche, cette méthode fournit une estimée plus bruitée et impose de filtrer les données par lissage simple ou en appliquant des techniques de réduction de variance.

L'estimation des diffusés peut être réalisée en ajustant, simultanément ou après correction des fortuits, une fonction gaussienne ou polynômiale en dehors du patient dans l'espace des projections ou en convoluant les coïncidences mesurées avec un noyau dont les coefficients sont préalablement déterminés empiriquement. Les diffusés peuvent être estimés de façon potentiellement plus exacte (Meikle et Badawi 2005) par simulation à partir d'une première image de l'activité reconstruite corrigée de tous les effets sauf de la diffusion. La simulation peut être analytique, en utilisant la formule de Klein-Nishina, ou Monte-Carlo (MC).

Les coïncidences multiples sont généralement supprimées. Dans certains cas, il peut être avantageux de sélectionner une paire de photons, par exemple celle maximisant l'énergie totale déposée.

La contribution des coïncidences fortuites et diffusées aux coïncidences vraies est additive et une soustraction des estimées peut-être réalisée pour corriger de ses effets ou en les intégrant à la matrice système (paragraphe 2.4.2).

2.5.3 Normalisation

Les détecteurs élémentaires possèdent des sensibilités différentes liées aux variations d'efficacité des détecteurs simples, d'origine matérielle ou électronique et à la géométrie du tomographe. Les variations de sensibilité sont liées à l'angle solide du cristal mesuré au centre du tomographe. Plus l'angle solide sera réduit (position radiale en valeur absolue et angle copolaire élevés), plus la probabilité que l'interaction se produise dans un TDR adjacent au TDR dans lequel s'est réellement produite l'annihilation augmente. D'autre part, la réduction de l'angle solide réduit le volume couvert par les TDR et la sensibilité de la paire de détecteurs associée. Cet effet peut être prédit analytiquement et inclus dans la reconstruction.

La normalisation consiste à mesurer la réponse du système à un signal uniforme puis à corriger les coïncidences mesurées sur chaque élément de détection par application d'un facteur correctif multiplicatif η_{ij} égal au rapport du nombre de photons détectés en moyenne sur l'ensemble des éléments de détection sur le nombre de photons détectés par l'élément de détection (C_i, C_j). Ceci peut être fait à l'aide d'une acquisition de lignes sources rotatives ou de sources planaires à haute activité en incluant une correction analytique de la non uniformité de l'exposition des LDR. Cette méthode de normalisation directe a pour désavantage d'immobiliser le tomographe pendant une longue durée et de nécessiter l'utilisation d'une source très homogène à haute activité qui peut induire des temps morts, effets d'empilements et fractions de diffusés non réalistes. Une alternative consiste à décomposer les facteurs de normalisation η_{ij} en produit de facteurs $\eta_{ij} = \epsilon_i \epsilon_j g_{ij}$ où ϵ_i représente le facteur correctif

des hétérogénéités d'efficacité du détecteur i simples et g_{ij} représente les variations de sensibilité liées à l'interférence entre les deux détecteurs, d'origine géométrique (Cherry et Dahlbom 2006). Les coefficients ϵ_i sont calculés à partir de moyennes du nombre de coïncidences détectées sur un ensemble de LDR joignant le détecteur i à un ensemble de détecteurs opposés. Les facteurs géométriques sont calculés une seule fois à haute activité et sont supposés constants au cours du temps. De même que pour les autres corrections, les facteurs multiplicatifs peuvent être appliqués sur les sinogrammes ou pendant la reconstruction.

2.5.4 Temps mort et calibration

Le temps mort induit un biais de mesure dépendant de l'activité dans le champ de vue. Il peut être corrigé en mesurant le « temps de vie » du détecteur ou en ajustant sur les données un modèle associant deux types de réponses du détecteur lors d'un phénomène d'empilement. Dans un cas, le temps mort est prolongé, on parle de temps mort paralysable, dans l'autre il reste le même et l'événement est simplement ignoré. La correction du temps mort peut aussi être incluse dans l'étape de calibration.

Une fois que les différents effets ont été explicitement corrigés, les données reconstruites sont calibrées. Cette étape consiste à établir une table de correspondance donnant le taux de coïncidences détectées en fonction de l'activité. Ceci peut être fait simplement en réalisant plusieurs acquisitions d'une même source homogène d'activité connue qu'on laisse décroître. Enfin, l'activité au temps de l'injection pourra être retrouvée à partir de l'intégrale mesurée entre le temps de l'injection t_{inj} et le temps de fin d'acquisition $t_{inj} + \Delta t$ par la relation :

$$\int_{t_{inj}}^{t_{inj}+\Delta t} A(t)dt = A(t_{inj}) \int_{t_{inj}}^{t_{inj}+\Delta t} e^{-\lambda t} dt = A(t_{inj}) \frac{e^{-\lambda t_{inj}} - e^{-\lambda(t_{inj}+\Delta t)}}{\lambda}$$

2.5.5 Implémentation des corrections

D'après ce que nous venons de voir, l'opération de projection peut être considérée comme l'opération

$$\mathbf{f} \rightarrow \mathbf{A}\mathbf{N}\mathbf{H}\mathbf{f} + \mathbf{s} + \mathbf{r} \quad (2.13)$$

où \mathbf{A} et \mathbf{N} sont les matrices d'atténuation et de normalisation, diagonales de taille $m \times m$, \mathbf{H} est la matrice système et \mathbf{s} et \mathbf{r} sont les estimées des coïncidences fortuites et diffusées. Les mesures brutes $\mathbf{g} = \mathbf{A}\mathbf{N}\mathbf{H}\mathbf{f} + \mathbf{s} + \mathbf{r}$ contiennent donc les coïncidences vraies, fortuites et diffusées. On pourra alors appliquer ces facteurs correctifs directement sur les données mesurées puis procéder à l'inversion par \mathbf{H} de $\mathbf{A}^{-1}\mathbf{N}^{-1}(\mathbf{g} - \mathbf{s} - \mathbf{r})$. Cette façon de procéder a le désavantage de modifier les propriétés statistiques du signal. A l'inverse, on pourra conserver ces propriétés en intégrant le modèle de projection complet décrit par l'équation 2.13 dans l'algorithme. Il s'agira alors de remplacer l'opérateur de projection $\mathbf{f} \rightarrow \mathbf{H}\mathbf{f}$ dans l'équation 2.11 par l'opération 2.13 (Michel *et al.* 1998) soit :

$$\mathbf{f}^{(n+1)} = \frac{\mathbf{f}^{(n)}}{\mathbf{A}^t \mathbf{N}^t \mathbf{H}^t \mathbf{1}} \mathbf{A}^t \mathbf{N}^t \mathbf{H}^t \frac{\mathbf{g}}{\mathbf{A}\mathbf{N}\mathbf{H}\mathbf{f}^{(n)} + \mathbf{s} + \mathbf{r}} \quad (2.14)$$

On dit que les opérations réalisées à gauche de \mathbf{H} sont faites dans l'espace des projections, tandis que les opérations réalisées à droite de \mathbf{H} sont faites dans l'espace image. L'opérateur \mathbf{H} fait généralement l'hypothèse de projections suivant un modèle ligne-intégrale et la FDP est supposée parfaite. Les systèmes récents proposent d'inclure dans l'opérateur \mathbf{H} la FDP, exprimée dans l'espace image ou dans l'espace des projections (Rahmim *et al.* 2013), afin d'améliorer son réalisme. Ces méthodes permettent d'améliorer significativement la résolution spatiale dans les images reconstruites (Panin *et al.* 2006). De la même manière, le parcours du positon peut être modélisé dans l'espace image (Rahmim *et al.* 2008).

2.5.6 Caractérisation des images reconstruites

La qualité des images, par nature stochastiques, est généralement définie en termes de biais et de variance. Le biais dans les images dépend de la résolution spatiale intrinsèque du système, de l'exactitude des corrections, de l'échantillonnage des images reconstruites et des propriétés de l'algorithme utilisé. Sans considérer les corrections appliquées (voir le paragraphe 2.5), les algorithmes analytiques présentés ici sont linéaires. Le biais est ainsi entièrement défini par leur fonction de réponse et sont indépendants de l'objet. À l'inverse, l'algorithme EM est non linéaire et le biais dépend de l'objet reconstruit. D'autre part, l'algorithme est souvent arrêté avant convergence, conduisant à des variations supplémentaires de résolution spatiale. La variance dans les images reconstruites est liée au faible nombre de coups par LDR (moins de 10 en conditions cliniques) et aux propriétés de l'algorithme. Avec EM, la variance est approximativement proportionnelle à l'image elle-même. Compte tenu de la nature mal posée du problème de reconstruction, la variance diverge pour les nombres d'itérations élevés. Celle-ci doit être contrôlée de façon explicite (filtrage) ou implicite (arrêt des algorithmes itératifs après un nombre d'itérations fixe).

SYNTHÈSE ET CONCLUSION

Nous avons exposé dans ce chapitre les principes physiques et informatiques sur lesquels reposent la TEP. Nous avons vu que la distribution spatiale des positions d'annihilations des positons émis lors de la désintégration du radiotracer était mesurable grâce à l'association d'un système de détection des projections et d'un algorithme de reconstruction. Nous avons montré qu'une mesure quantitative de cette distribution était possible mais celle-ci reste affectée d'un biais de mesure et d'une variabilité importante, causés par la résolution spatiale limitée du système d'acquisition et de reconstruction et par le bruit dans l'image. Ces erreurs peuvent significativement affecter les résultats et leur interprétation, en particulier sur des petites structures fonctionnelles comme les lésions athéromateuses.

Dans le chapitre suivant, nous proposons une analyse systématique de la littérature dédiée à l'interprétation quantitative des images TEP au Fluorodésoxyglucose marqué au fluor 18 ([^{18}F]-FDG) de l'athérosclérose afin d'identifier les paramètres d'imagerie, liés à l'acquisition, à la reconstruction et à l'analyse d'image, susceptibles d'induire des variations dans le signal mesuré indépendamment de changements d'origine physiopathologique.

