Approche par variance nulle afin d'orienter le choix des probabilités

Comme précisé en introduction de cette section, plusieurs grandeurs arbitraires doivent encore être déterminées. C'est en particulier le cas de trois probabilités : la densité de probabilité des nombres d'onde $p_H(\eta)$, la probabilité $\mathcal{P}_m(\mathbf{x})$ associée à chaque espèce moléculaire et la probabilité $\mathcal{P}_i(\mathbf{x})$ associée à chaque transition *i* pour une espèce moléculaire *m* donnée.

Pour orienter ces choix, nous avons recouru à une approche par variance nulle (voir Sec. 3.3.4.2), en partant de l'Eq. 5.18 dans laquelle nous avons libéré les choix de toutes les probabilités et densités de probabilité (à savoir $p_H(\eta)$, $\hat{p}_{\mathcal{L}_i}(l_j)$, $\mathcal{P}_m(\mathbf{x}_j)$,



FIGURE 5.5 – Benchmark proposé dans l'article [André et Vaillon, 2010]. Les dimensions, ainsi que les pressions, températures et fractions molaires des différentes espèces participantes y sont représentées.

 $\mathcal{P}_{i}(\mathbf{x}_{j})$ et $\mathcal{P}_{a,i}(\mathbf{x}_{j})$), dans le but de déterminer leur expression respective (indicée \emptyset) assurant une variance nulle du poids de Monte-Carlo. Les résultats obtenus par cette approche (décrite à l'Annexe C) sont présentés ci-après :

Nombres d'onde. La densité de probabilité des nombres d'onde assurant une variance nulle du poids de Monte-Carlo est donnée par le rapport entre la luminance monochromatique au point \mathbf{x}_0 dans la direction \mathbf{u}_0 et la luminance au même point, dans la même direction mais intégrée spectralement entre η_{\min} et η_{\max} :

$$p_{H,\emptyset}(\eta) = \frac{L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)}{L(\mathbf{x}_0, \mathbf{u}_0)} = \frac{L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)}{\int_{\eta_{\min}}^{\eta_{\max}} L_{\eta}(\mathbf{x}_0, \mathbf{u}_0) d\eta}$$
(5.26)

Nous rechercherons donc à approcher au mieux $L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)$ par un modèle simplifié de luminance, analytiquement intégrable entre η_{\min} et η_{\max} pour pouvoir assurer une inversion de la fonction de répartition et donc un échantillonnage des nombres d'onde.

Libres parcours. La densité de probabilité des libres parcours assurant une variance nulle est donnée par :

$$p_{\mathcal{L}_{j},\emptyset}(l_{j}) = \frac{\exp\left(-\int_{0}^{l_{j}} \hat{k}_{\eta}(\mathbf{x}_{j-1} - l_{j}'\mathbf{u}_{0})dl_{j}\right) \left[k_{a,\eta}(\mathbf{x}_{j})L_{\eta}^{eq}(\mathbf{x}_{j}) + k_{n,\eta}(\mathbf{x}_{j})L_{\eta}(\mathbf{x}_{j+1},\mathbf{u}_{0})\right]}{L_{\eta}(\mathbf{x}_{0},\mathbf{u}_{0})}$$
(5.27)

On remarque alors que la densité de probabilité

$$\hat{p}_{\mathcal{L}_j}(l_j) = \hat{k}_{\eta}(\mathbf{x}_j) \exp\left(-\int_0^{l_j} \hat{k}_{\eta}(\mathbf{x}_{j-1} - l'_j \mathbf{u}_0) dl_j\right)$$
(5.28)

que l'on a choisi de conserver, correspond à la densité de probabilité idéale $p_{\mathcal{L}_j,\emptyset}(l_j)$ pour le cas particulier d'un milieu à l'équilibre thermodynamique $(L_{\eta}^{eq}(\mathbf{x}_j) = L_{\eta}(\mathbf{x}_0, \mathbf{u}_0))$. En d'autres termes, la variance associée à la variable aléatoire des libres parcours, sera faible si le milieu est proche d'un équilibre thermodynamique et plus importante s'il en est éloigné.

Espèces moléculaires. La probabilité associée à chaque espèce moléculaire m garantissant une variance nulle du poids de Monte-Carlo est donnée par le rapport entre la participation $h_{a,m,i,\eta}(\mathbf{x})$ de l'ensemble des raies de l'espèce m et le coefficient d'absorption :

$$\mathcal{P}_{m,\emptyset}(\mathbf{x}_j) = \frac{\sum_{i=1}^{N_i(m)} h_{a,m,i,\eta}(\mathbf{x}_j)}{k_{a,\eta}(\mathbf{x}_j)} = \frac{\sum_{i=1}^{N_i(m)} h_{a,m,i,\eta}(\mathbf{x}_j)}{\sum_{m=1}^{N_m} \sum_{i=1}^{N_i(m)} h_{a,m,i,\eta}(\mathbf{x}_j)}$$
(5.29)

Il faudra donc chercher à modéliser, en fonction de la position \mathbf{x} et de la position η , la fraction du coefficient d'absorption $k_{a,\eta}(\mathbf{x})$ due à la molécule m.

Transitions moléculaires. La probabilité associée à chaque transition i d'une espèce m assurant une variance nulle est donnée par le rapport entre la participation de cette transition au coefficient d'absorption et la participation de l'ensemble des raies de l'espèce m à ce même coefficient d'absorption :

$$\mathcal{P}_{i,\emptyset}(\mathbf{x}_j) = \frac{h_{a,m,i,\eta}(\mathbf{x}_j)}{\sum_{i=1}^{N_i(m)} h_{a,m,i,\eta}(\mathbf{x}_j)}$$
(5.30)

Pour un gaz mono-moléculaire, nous chercherons donc à modéliser le rapport entre la participation d'une raie et le coefficient d'absorption global pour tout point \mathbf{x} du milieu et tout nombre d'onde η . Cette étape constitue certainement l'élément le plus délicat du choix des paramètres libres.

Type de collision. Enfin, la probabilité d'absorption assurant une variance nulle du poids de Monte-Carlo est donnée par :

$$\mathcal{P}_{a,i,\emptyset}(\mathbf{x}_j) = \frac{k_{a,\eta}(\mathbf{x}_j)L_{\eta}^{eq}(\mathbf{x}_j)}{k_{a,\eta}(\mathbf{x}_j)L_{\eta}^{eq}(\mathbf{x}_j) + k_{n,\eta}(\mathbf{x}_j)L_{\eta}(\mathbf{x}_{j+1},\mathbf{u}_0)}$$
(5.31)

S'il était possible d'utiliser ces probabilités idéales, une seule réalisation de l'algorithme de Monte-Carlo serait nécessaire pour estimer $L(\mathbf{x}_0, \mathbf{u}_0)$ avec une variance nulle, quelle que soit la valeur du champ de $\hat{k}_{\eta}(\mathbf{x})$. Celui ci n'aurait alors qu'une incidence sur le temps de calcul en conditionnant uniquement l'indice j de la collision à laquelle a lieu une absorption. Si l'on souhaitait minimiser ce temps de calcul, c'est-à-dire faire en sorte que l'absorption ait lieu toujours à la première collision, il faudrait alors que la probabilité $\mathcal{P}_{a,\imath,\emptyset}(\mathbf{x}_j)$ soit égale à 1, ce qui reviendrait à fixer $\hat{k}_{\eta}(\mathbf{x}) = k_{a,\eta}(\mathbf{x})$ pour tout \mathbf{x} .

Ces conclusions nous réconfortent également quant au choix fait pour le coefficient d'extinction $\hat{h}_{\eta}(\mathbf{x}_j) = \mathcal{P}_m(\mathbf{x}_j)\mathcal{P}_i(\mathbf{x}_j)\hat{k}_{\eta}(\mathbf{x}_j)$ associé à une transition particulière. En effet, si l'on considère les probabilités optimales assurant une variance nulle, cette égalité équivaut à $\hat{h}_{\eta}(\mathbf{x}_j) = \frac{h_{a,m,i,\eta}(\mathbf{x}_j)}{k_{a,\eta}(\mathbf{x}_j)}\hat{k}_{\eta}(\mathbf{x}_j)$. Et si l'on suppose que l'on parvienne à éviter toute collision nulle (si $\hat{k}_{\eta}(\mathbf{x}_j) = k_{a,\eta}(\mathbf{x}_j)$), nous arrivons alors à l'égalité $\hat{h}_{\eta}(\mathbf{x}_j) = h_{a,m,i,\eta}(\mathbf{x}_j)$.

Cette approche par variance nulle a été réalisée ici pour un milieu gazeux infini et non-diffusant. Toutefois la même approche peut être effectuée pour un milieu fini et/ou diffusant. Il est juste important de remarquer, pour les sections suivantes, que l'ajout d'une frontière n'entraîne aucune modification des probabilités idéales relatives aux indices d'espèce moléculaire $\mathcal{P}_{m,\emptyset}(\mathbf{x}_j)$, de transition $\mathcal{P}_{i,\emptyset}(\mathbf{x}_j)$ et au type de collision $\mathcal{P}_{a,i,\emptyset}(\mathbf{x}_j)$, dans la mesure où ces trois probabilités n'interviennent que dans le cas où la collision a lieu dans le milieu participant.

5.3.3 Choix des paramètres libres

Cette section a pour objectif d'introduire les choix faits pour les derniers paramètres libres : le champ de $\hat{k}_{\eta}(\mathbf{x})$, la densité de probabilité des nombres d'onde $p_H(\eta)$, la probabilité $\mathcal{P}_m(\mathbf{x})$ associée à chaque espèce moléculaire, la probabilité $\mathcal{P}_i(\mathbf{x})$ associée à chaque transition *i* pour une espèce moléculaire *m* donnée, ainsi que le critère ζ à partir duquel l'algorithme permute d'un traitement déterministe des collisions (*energy-partitioning*) à un traitement stochastique par test de Bernoulli. Ces différents termes arbitraires n'ont une incidence que sur les taux de convergence de la méthode proposée et n'influent en rien sur le caractère de solution de référence de l'algorithme présenté à la Fig. 5.4.

Les propositions faites ici résultent d'un travail d'optimisation qui a consisté à affiner petit à petit les modèles choisis par une approche de type essai-erreur, jusqu'à atteindre un taux de convergence satisfaisant pour chacun des 6 cas d'étude proposés dans [André et Vaillon, 2010]. Les choix proposés ne sont certainement pas optimaux et nous pensons qu'un important travail les concernant sera nécessaire dans la continuité de ces travaux. L'objectif de cette section n'est donc pas de faire des propositions conduisant à un algorithme plus rapide ou plus performant qu'un autre, mais de prouver la faisabilité pratique, dans des temps de calcul acceptables, de l'approche présentée dans ce chapitre.

La principale difficulté concerne très probablement la définition des champs de $\hat{k}_{\eta}(\mathbf{x})$ ainsi que des probabilités $\mathcal{P}_m(\mathbf{x})$ et $\mathcal{P}_i(\mathbf{x})$. En effet, comme mis en évidence à la Sec. 5.2.3, il est souhaitable de définir ces trois paramètres libres de sorte à garantir l'inégalité

$$k_{\eta}(\mathbf{x})\mathcal{P}_{m}(\mathbf{x})\mathcal{P}_{i}(\mathbf{x}) > h_{a,m,i,\eta}(\mathbf{x})$$
(5.32)

pour éviter d'éventuels problèmes d'augmentation brutale de la variance (tels que ceux présentés au Chap. 4) qui seraient causés par des coefficients $h_{n,m,i,\eta}(\mathbf{x})$ négatifs. Ces choix sont donc très critiques en termes de convergence. Trois leviers d'optimisation : $\hat{k}_{\eta}(\mathbf{x})$, $\mathcal{P}_m(\mathbf{x})$ et $\mathcal{P}_i(\mathbf{x})$ sont alors offerts pour que l'Eq. 5.32 soit vérifiée pour tout nombre d'onde, toute espèce moléculaire, toute transition et en tout point (*i.e.* pour toute température, pression et composition chimique).

Nous avons alors choisi de procéder de la façon suivante :

• nous définissons le champ $k_{\eta}(\mathbf{x})$ comme le produit d'un facteur α constant et d'un champ arbitraire $\tilde{k}_{\eta}(\mathbf{x})$ majorant idéalement pour tout η et tout \mathbf{x} le champ du coefficient d'absorption $k_{a,\eta}(\mathbf{x})$:

$$\hat{k}_{\eta}(\mathbf{x}) = \alpha \tilde{k}_{\eta}(\mathbf{x}) \tag{5.33}$$

- nous nous concentrons sur le choix des probabilités arbitraires $\mathcal{P}_m(\mathbf{x})$ et $\mathcal{P}_i(\mathbf{x})$;
- nous ajustons la constante α de sorte à vérifier dans la quasi-totalité des cas l'Eq. 5.32. La valeur du paramètre α sera donc conditionnée par la qualité du choix des probabilités $\mathcal{P}_m(\mathbf{x})$ et $\mathcal{P}_i(\mathbf{x})$. Si celles-ci respectaient les conclusions de l'approche par variance nulle alors un facteur $\alpha = 1$ serait suffisant, mais si

ces probabilités sont sous-estimées localement (pour un nombre d'onde et une transition donnés), la valeur de α devra être augmentée en conséquence.

Choix du champ de $\tilde{k}_{\eta}(\mathbf{x})$. En s'appuyant sur les conclusions de l'approche par variance nulle (voir Sec. 5.3.2), le champ de $\hat{k}_{\eta}(\mathbf{x}) = \alpha \tilde{k}_{\eta}(\mathbf{x})$ doit être le plus proche possible du champ de coefficient d'aborption $k_{a,\eta}(\mathbf{x})$. La constante α ne jouant un rôle que de facteur correctif permettant d'assurer la condition de l'Eq. 5.32 quelle que soit la qualité des probabilité $\mathcal{P}_m(\mathbf{x})$ et $\mathcal{P}_i(\mathbf{x})$, nous nous concentrons donc ici sur le champ de $\tilde{k}_{\eta}(\mathbf{x})$.

Dans un premier temps nous avons défini le champ de $k_{\eta}(\mathbf{x})$ comme uniforme et indépendant des nombres d'onde. Le paramètre α a été augmenté progressivement dans chacun des six cas d'étude, jusqu'à ce qu'aucun coefficient négatif de collision nulle $h_{n,m,i,\eta}(\mathbf{x})$ ne soit rencontré pendant le calcul de luminance $L(\mathbf{x}_0, \mathbf{u}_0)$ intégrée entre $\eta_{\min} = 10cm^{-1}$ et $\eta_{\max} = 15000cm^{-1}$. Les temps de calcul, assurant une erreur relative de 1%, obtenus suite à ce premier modèle de $\hat{k}_{\eta}(\mathbf{x})$, sont compris entre 1mn pour le cas C1 à 5h30 pour le cas C6³.

Pour comprendre pourquoi ces temps sont si importants et différents, il est nécessaire d'analyser la dépendance spatiale et spectrale du coefficient d'absorption (voir Fig. 5.6, où trois spectres d'absorption produits pour 3 points du cas d'étude C3 sont représentés). En définissant $\tilde{k}_{\eta}(\mathbf{x})$ comme une constante majorant en tout



FIGURE 5.6 – Spectres d'absorption produits pour les conditions thermodynamiques et la composition chimique rencontrées en 3 points (x=0m, x=1.5m et x=8m) du cas d'étude C3. Le coefficient d'extinction $\hat{k}_{\eta}(\mathbf{x})$ défini comme une constante majorant le coefficient d'absorption maximal est également représenté.

^{3.} Les calculs ont été réalisés sur un seul cœur d'un processeur Intel Core i7 de 2.8GHz. Les temps de calcul excluent les étapes de preprocessing et de chargement des données.

point et pour tout nombre d'onde le coefficient d'absorption $k_{a,\eta}(\mathbf{x})$, on constate qu'excepté pour les nombres d'onde où le coefficient d'absorption est très important, $\tilde{k}_{\eta}(\mathbf{x})$ majore très largement (jusqu'à un facteur 10¹¹) le coefficient d'absorption. Cela se traduit algorithmiquement par une proportion de collisions nulles très conséquente. Puisque ces collisions nulles n'ont que peu d'effet sur la convergence de l'algorithme, et ne permettent pas de mettre fin à la récursivité d'une réalisation (voir Fig. 5.4), cela explique les temps de calcul importants, d'autant plus si les dimensions caractéristiques sont grandes.

On mesure alors les importantes marges d'amélioration qui peuvent être réalisées en détaillant de façon plus rigoureuse le champ de $\tilde{k}_{\eta}(\mathbf{x})$. Nous avons alors fait appel à des spectres d'absorption haute-résolution. Le point important ici, est que ces spectres n'ont pas besoin d'être en cohérence avec les données spectroscopiques ou avec les hypothèses spectrales faites pour la simulation d'intérêt. Ces spectres doivent simplement permettre de vérifier l'Eq. 5.32 tout en minimisant autant que possible la quantité de collisions nulles rencontrées. Idéalement, ils doivent être choisis de sorte à majorer légèrement les champs de coefficients d'absorption réels pour tout nombre d'onde. Si tel n'est pas le cas, une augmentation du facteur α permettra de corriger cette mauvaise estimation. Il est important de rappeler que, dans la mesure où le champ de \hat{k}_{η} est arbitraire, les choix faits ici n'influent en rien sur le caractère exact de la méthode de Monte-Carlo, il n'a pour seul but que d'accélérer les temps de calcul.

Dans l'objectif d'approcher au mieux la dépendance spatiale et spectrale du coefficient d'absorption, nous avons divisé les milieux d'étude en plusieurs sous-domaines (de 2 à 5 sous-domaines selon le cas d'étude) pour tenir compte de l'hétérogénéité des propriétés du milieu. À chaque sous-domaine a été affecté un spectre d'absorption approximativement représentatif des conditions thermodynamiques et chimiques rencontrées. Nous avons ici utilisé des spectres produits dans d'autres contextes à partir des bases de données spectroscopiques CDSD-1000 pour le CO_2 et Hitemp pour l'H₂O. Les raies les plus intenses (dont l'intensité était supérieure à $10^{-21}molec.cm$) de CDSD-4000 et de HITEMP 2010 ont également été rajoutées à ces spectres pour que l'Eq. 5.32 soit vérifiée dans les plages spectrales les plus critiques. Ces mêmes spectres ont été conservés pendant toute la durée de ces travaux, quelles que soient les bases de données spectroscopiques ou les hypothèses spectrales retenues (troncature d'ailes de raies, profils de raie, *etc.*) pour le calcul de $L(\mathbf{x}_0, \mathbf{u}_0)$. Ce nouveau choix de champ de k_{η} a alors conduit à une accélération des temps de calcul de 60 à 20000 fois selon le cas d'étude par rapport à un champ de \hat{k}_{η} uniforme. En effet, en perfectionnant ce choix, une quantité très importante de collisions nulles ont pu être ainsi évitées. Les résultats obtenus grâce à cette proposition sont décrits à la Sec. 5.3.4.

Toutefois, n'importe quel champ de $k_{\eta}(\mathbf{x})$ validant autant que possible l'Eq. 5.32 peut être accepté sans altérer le caractère de référence de la méthode, seul le taux de convergence de l'algorithme en sera modifié. Il est par exemple possible d'imaginer des solutions intermédiaires. Il pourrait notamment être possible de créer une banque de données de spectres approchés, de faible résolution spectrale, majorant les spectres réels d'absorption pour une large plage de conditions thermodynamiques. Les temps de calcul seraient alors certainement plus importants que ceux obtenus avec le choix présenté précédemment, mais le travail de définition des champs de \hat{k}_{η} en serait grandement simplifié.

Choix de la probabilité des transitions. Les résultats de l'approche par variance nulle énoncent que la probabilité idéale associée à chaque transition est définie comme le rapport entre la participation d'absorption $h_{a,m,i,\eta}(\mathbf{x})$ de la transition *i* et la somme des participations de toutes les transitions d'une molécule *m* donnée. Si l'on considère un gaz mono-moléculaire, cette probabilité idéale est définie comme le rapport entre $h_{a,m,i,\eta}(\mathbf{x})$ et le coefficient d'absorption. Bien entendu, cette proposition ne peut être retenue dans l'approche proposée ici, puisqu'il serait nécessaire d'évaluer le coefficient d'absorption à chaque collision (et donc de sommer, à chaque collision, les contributions de l'ensemble considérable de transitions, de façon déterministe).

Ce choix de probabilités est probablement le plus critique et le plus complexe à mettre en œuvre. En effet, il n'est pas imaginable de produire à chaque réalisation (et encore moins à chaque collision) un jeu complet de probabilités associées à chacune des transitions représentées dans la base de données spectroscopique d'intérêt. Les temps de calcul que cela impliquerait seraient trop importants. Or, toujours en se basant sur l'approche par variance nulle, il faudrait idéalement que ces jeux de probabilités soient définis en tout point du milieu et pour tout nombre d'onde. Cela n'est concrètement pas imaginable, même en négligeant les variations dues aux hétérogénéités des propriétés du milieu (température, pression et concentrations), cela signifierait créer autant de spectres de \mathcal{P}_i qu'il y a de transitions dans les bases de données spectroscopiques.

Notre première proposition de \mathcal{P}_i a alors consisté à imaginer un modèle très simplifié dans lequel les profils de toutes les raies $f(\eta)$ sont identiques, et où ces dernières sont centrées en un même nombre d'onde η_0 . En appliquant les résultats de l'approche par variance nulle pour ce modèle simplifié, on arrive à la proposition suivante :

$$\mathcal{P}_{i}(\mathbf{x}) \equiv \mathcal{P}_{m,i,\eta}(\mathbf{x}) = \frac{h_{a,m,i,\eta(\mathbf{x})}}{\sum_{i'=1}^{N_{i}(m)} h_{a,m,i',\eta}(\mathbf{x})} = \frac{S_{m,i}(\mathbf{x})f(\eta)}{\sum_{i'=1}^{N_{i}(m)} S_{m,i'}(\mathbf{x})f(\eta)} = \frac{S_{m,i}(\mathbf{x})}{\sum_{i'=1}^{N_{i}(m)} S_{m,i'}(\mathbf{x})}$$
(5.34)

Les profils de raies étant identiques, cette probabilité est alors définie comme le ratio entre l'intensité $S_{m,i}$ de la raie *i* divisée par la somme des intensités de l'ensemble des raies de la molécule *m*. L'avantage de cette proposition réside dans le fait que les intensités de raies ne dépendent pas du nombre d'onde d'intérêt. En milieu homogène, ce cas d'étude est donc tout à fait envisageable d'un point de vue numérique. En effet, puisque $\mathcal{P}_i(\mathbf{x})$ ne dépend plus du nombre d'onde, seul un jeu de probabilités (pour toutes les transitions de l'espèce *m*) est à précalculer. Pour un milieu hétérogène, on peut imaginer le discrétiser suffisamment pour couvrir de façon approximative les hétérogénéités des propriétés et produire autant de jeux de $\mathcal{P}_i(\mathbf{x})$ qu'il y a de mailles. Cette proposition a donné des résultats et des temps de calcul acceptables dans le cas particulier où les domaines spectraux d'intégration étaient faibles (de la taille d'une bande étroite) et qu'une troncature des ailes de raie était mise en place. En effet, grâce à la troncature, seules les raies les plus proches de l'intervalle d'intégration sont prises en compte par l'algorithme qui les échantillonne alors uniquement selon leurs intensités respectives. Toutefois, dès que l'on supprime cette troncature où que l'on élargit le domaine d'intégration, les temps de calcul deviennent rapidement excessifs. Avec cette proposition, une raie très intense mais très éloignée du nombre d'onde d'étude η a une plus grande probabilité d'être échantillonnée qu'une raie plus faible mais centrée en η , même si son coefficient $h_{a,m,i,\eta}(\mathbf{x})$ en η est de loin supérieur à celui de la première raie.

Il est alors évident, qu'à cause de son indépendance spectrale, cette première proposition pose de sérieuses limitations. Toutefois, comme nous venons de le dire, il est difficilement imaginable de produire des jeux de $\mathcal{P}_i(\mathbf{x})$ couvrant les variations spectrales du coefficient d'absorption. La seule solution, pour prendre en compte l'effet dû à l'éloignement spectral des raies du nombre d'onde d'étude, consiste alors à proposer une probabilité permettant d'être calculée analytiquement au cours du calcul sans avoir à produire un jeu complet de $\mathcal{P}_i(\mathbf{x})$ pour toutes les transitions à chaque collision.

Nous avons alors retenu la proposition suivante pour prendre en compte cette dimension spectrale dans l'échantillonnage des transitions :

- Nous définissons $[\eta_{db,\min}, \eta_{db,\max}]$ comme l'intervalle englobant l'ensemble des nombres d'onde de centre de raie contenus dans la base de données spectroscopique d'intérêt. Puis, nous segmentons cet intervalle en bandes régulières (indicées β) de largeur $\delta_{\eta} = 0.3 cm^{-1}$ et de bornes $[\eta_{\min,\beta}, \eta_{\max,\beta}]$;
- Nous décomposons la probabilité *P_i* associée à chaque transition en un produit de deux probabilités : *P_i* = *P_β* × *P_{i,β}*
- Dans un souhait de représenter également les variations spatiales des propriétés radiatives, nous conservons les sous-domaines utilisés pour la définition du champ de $\hat{k}_{\eta}(\mathbf{x})$. Une valeur moyenne de température est alors affectée à chacun de ces sous-domaines.
- La première probabilité $\mathcal{P}_{\beta} \equiv \mathcal{P}_{\beta,\eta}(\mathbf{x})$ est associée à chaque bande spectrale β . Elle a pour objectif de rendre compte de l'éloignement d'une raie *i* par rapport au nombre d'onde d'intérêt η . Nous la définissons comme :

$$\mathcal{P}_{\beta} = \frac{\operatorname{atan}\left(\frac{\eta_{\max,\beta}(\beta)-\eta}{\gamma_{Q}}\right) - \operatorname{atan}\left(\frac{\eta_{\min,\beta}(\beta)-\eta}{\gamma_{Q}}\right)}{\operatorname{atan}\left(\frac{\eta_{db,\max}-\eta}{\gamma_{Q}}\right) - \operatorname{atan}\left(\frac{\eta_{db,\min}-\eta}{\gamma_{Q}}\right)}$$
(5.35)

où γ_Q est un paramètre libre (fixé ici à $\gamma_Q = 1cm^{-1}$ pour les sous-domaines de température T < 500K; à $\gamma_Q = 0.5cm^{-1}$ pour $T \in [500K, 1000K]$; à $\gamma_Q = 0.25cm^{-1}$ pour $T \in [1000K, 1500K]$ et à $\gamma_Q = 0.1cm^{-1}$ pour T > 1500K). Avec une telle définition, échantillonner une bande d'indice β revient à échantillonner un nombre d'onde η_β selon un profil lorentzien de demi-largeur à mi-hauteur γ_Q , centré sur le nombre d'onde d'intérêt η et de retenir la bande β dans laquelle est inclus η_{β} .⁴.

• La probabilité $\mathcal{P}_{i,\beta} \equiv \mathcal{P}_{i,\beta,m}(\mathbf{x})$ est associée aux transitions *i* centrées dans une bande β donnée. Elle permet de rendre compte de leur intensité par rapport à celles des autres transitions centrées dans la même bande β . Elle est définie comme :

$$\mathcal{P}_{i,\beta} = \frac{S_{m,i}(\mathbf{x})}{\sum_{i'=1}^{N_i(m,\beta)} S_{m,i'}(\mathbf{x})}$$
(5.36)

où $S_{m,i}(\mathbf{x})$ est l'intensité de la raie *i* centrée dans β à la température du sous-domaine auquel appartient \mathbf{x} et où $N_i(m,\beta)$ correspond au nombre de raies de l'espèce *m* centrées dans la bande β . La probabilité $\mathcal{P}_{i,\beta}$ peut ainsi être calculée en preprocessing pour chacune des températures moyennes des sous-domaines d'étude, pour chaque intervalle spectral β et pour chacune des transitions *i* centrées en β . Les temps de calcul, associés à la production d'un jeu complet de probabilités $\mathcal{P}_{i,\beta}$ pour une température donnée, varient entre 1 seconde (pour la base de données spectroscopique HITRAN) et 5mn (pour la base de données CDSD 4000). Bien que ces temps soient importants, les jeux de probabilités calculés (pour une base de données et une température) peuvent être stockés pour être réutilisés lors d'autres simulations.

Ces choix de probabilités conduisent alors à privilégier les raies les plus intenses, centrées près du nombre d'onde d'intérêt. La procédure d'échantillonnage (illustrée à la Fig. 5.7) est alors composée de deux étapes :

- l'échantillonnage analytique d'une bande β selon un profil de Lorentz centré en η et de demi-largeur à mi-hauteur γ_Q . La probabilité analytique associée à la bande échantillonnée est calculée (de façon quasi-instantanée) au cours de la simulation.
- l'échantillonnage d'une transition parmi les $N_i(m,\beta)$ transitions centrées dans β . Les probabilités associées à chacune de ces transitions ne dépendant que de la température et de la base de données spectroscopique, elles peuvent donc être précalculées et stockées.

Il n'est donc nullement nécessaire, avec la proposition faite ici, de recalculer à chaque réalisation de l'algorithme de Monte-Carlo ou à chaque collision un jeu complet de probabilités associées à l'ensemble des transitions moléculaires.

Choix de la probabilité des espèces moléculaires. Nous n'avons pas, au cours de ces travaux, cherché à optimiser la probabilité associée aux espèces moléculaires (seul le cas C3 comporte un mélange) et avons alors proposé une probabilité uniforme :

$$\mathcal{P}_m = 1/N_m \tag{5.37}$$

où N_m est le nombre d'espèces moléculaires considérées.

^{4.} Si les ailes de raie sont tronquées à une distance $\delta \eta_{tronc}$, alors $[\eta_{db,\min}, \eta_{db,\max}]$ est remplacé par $[\eta - \delta \eta_{tronc}, \eta + \delta \eta_{tronc}]$.



FIGURE 5.7 – Procédure d'échantillonnage des transitions moléculaires. Étape 1 : une bande β de largeur $\delta\eta$ est déterminée en échantillonnant un nombre d'onde selon un profil de Lorentz, de largeur γ_Q centrée en η . Étape 2 : Une raie i est échantillonnée parmi les $N_i(\beta)$ raies centrées dans β selon leurs intensités respectives.

Le choix de cette probabilité semble beaucoup moins critique que celui de \mathcal{P}_i . Toutefois, il est assez simplement possible d'imaginer une amélioration de cette probabilité d'espèces moléculaires. L'approche par variance nulle définit la probabilité \mathcal{P}_m idéale comme le rapport entre le coefficient d'absorption dû à l'espèce m et le coefficient d'absorption global (dû à toutes les molécules). Ici encore, il est souhaitable que la probabilité \mathcal{P}_m dépende de la position \mathbf{x} (pour tenir compte des hétérogénéités de composition moléculaire) et du nombre d'onde. On peut donc imaginer un précalcul basé sur des modèles approchés par bandes (ex : k-distributions ou modèle de Malkmus) et sur une discrétisation spatiale (telle que celle proposée pour la définition du champ $\hat{k}_{\eta}(\mathbf{x})$).

Impacts des choix de \mathcal{P}_i et \mathcal{P}_m sur le facteur correctif α . Comme introduit en début de section, les expressions des probabilités \mathcal{P}_m et \mathcal{P}_i conditionnent la valeur du facteur multiplicatif α . Avec les choix faits précédemment, il a été nécessaire d'ajuster la valeur de α à 50 pour éviter de rencontrer un nombre trop important de coefficients $h_{n,m,i,\eta}$ négatifs, susceptibles de provoquer une augmentation importante de la variance de l'estimation. Malgré la valeur très importante de ce facteur multiplicatif, nous rencontrons encore des coefficients de collision nulle négatifs, mais ceux-ci n'ont pas d'impact sensible sur le taux de convergence de l'algorithme.

Fixer le facteur α à 50 équivaut à définir un champ de coefficient d'extinction environ 50 fois plus important que celui du coefficient d'absorption. Aussi, cela se traduit d'un point de vue numérique par une quantité très importante de collisions nulles sans réel intérêt pour le calcul. Cela met donc bien en évidence l'importance du choix des probabilités \mathcal{P}_i et \mathcal{P}_m qui constituent indéniablement les paramètres libres pour lesquels on peut attendre les améliorations futures les plus significatives. Tout perfectionnement de ces deux probabilités entraînera de facto une réduction de la valeur du facteur correctif α et ainsi des temps de calcul.

L'amélioration du choix de ces paramètres libres constitue alors un enjeu im-

portant et une perspective motivante pour la suite de ces travaux. Nous sommes convaincus qu'elle ne pourra se produire de façon sensible que grâce à une meilleure compréhension et une meilleure modélisation de la physique relative à la spectroscopie moléculaire, qui permettraient d'approcher au mieux les probabilités optimales calculées lors de l'approche par variance nulle. Nous voyons également dans les modèles statistiques de bandes (de type modèle Malkmus) des pistes très intéressantes pour parvenir à ces améliorations.

Choix de la densité de probabilité des nombres d'onde. La densité de probabilité des nombres d'onde $p_H(\eta)$ ne joue un rôle que sur les temps de calcul. Plus elle sera proche de la probabilité assurant une variance nulle : $p_{H,\emptyset}(\eta) = L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)/L(\mathbf{x}_0, \mathbf{u}_0)$, plus les temps de calcul pour une erreur relative donnée seront faibles.

Dans un premier temps, nous avons défini $p_H(\eta)$ comme uniforme :

$$p_H(\eta) = \frac{1}{\eta_{\max} - \eta_{\min}} \tag{5.38}$$

Cette première proposition, la plus simple qui soit, a conduit à des temps de calcul plutôt satisfaisants quel que soit le cas d'étude considéré. Cependant, nous avons tenté d'améliorer cette procédure d'échantillonnage des nombres d'onde en proposant un modèle approché, inspiré des résultats de l'approche par variance nulle. Il semblait toutefois difficile de définir, pour tout nombre d'onde, un modèle approché de $L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)$ qui puisse être intégré analytiquement entre η_{\min} et η_{\max} pour permettre un échantillonnage des libres parcours. Nous nous sommes donc orientés vers une description par bandes étroites. Nous proposons alors un échantillonnage des nombres d'onde composé de deux étapes :

• Une bande étroite i_b de $25cm^{-1}$ de largeur est échantillonnée parmi les N_b bandes étroites comprises entre η_{\min} et η_{\max} , selon la probabilité

$$\mathcal{P}_{i_b} = \frac{L_b^m(i_b) + \frac{1}{N_b} \sum_{b=1}^{N_b} L_b^m(b)}{2 \sum_{b=1}^{N_b} L_b^m(b)}$$
(5.39)

où L_b^m est une luminance intégrée sur la bande i_b précalculée à partir d'un modèle approché de Malkmus, associé à une approximation de Curtis-Godson [Goody *et al.*, 1989]. Un offset : $\frac{1}{N_b} \sum_{b=1}^{N_b} L_b^m(b)$ a été rajouté dans cette probabilité (normalisée en conséquence) pour compenser les sous-estimations de certaines bandes par le modèle approché.

• Un nombre d'onde η est échantillonné de façon uniforme sur la bande étroite i_b d'intérêt.

La densité de probabilité des nombres d'onde $p_H(\eta)$ que nous proposons ici est alors donnée par :

$$p_H(\eta) = \frac{L_b^m(i_b) + \frac{1}{N_b} \sum_{b=1}^{N_b} L_b^m(b)}{2 \sum_{b=1}^{N_b} L_b^m(b)} \times \frac{1}{\eta_{b,\max} - \eta_{b,\min}}$$
(5.40)

où $\eta_{b,\min}$ et $\eta_{b,\max}$ représentent les bornes de la bande étroite i_b . Cette nouvelle proposition a conduit à accélérer de 2 à 5 fois les temps de calcul pour une même erreur relative par rapport à une densité de probabilité uniforme. Cependant, elle reste d'un intérêt limité, dans la mesure où seules les bandes étroites sont pondérées et non les nombres d'onde. Il pourrait alors être peut-être intéressant d'utiliser l'approche proposée par A. Feldick et M. Modest dans [Feldick et Modest, 2011] qui consiste à échantillonner les nombres d'ondes directement à partir des transitions moléculaires.

Choix du seuil ζ . Nous avons fixé la valeur de permutation algorithmique à $\zeta = 0.5$. C'est cette valeur qui a donné les meilleurs résultats en termes de temps de calcul pour une erreur relative donnée. Ce choix est en accord avec l'étude paramétrique réalisée dans le Chap. 4. Toutefois, dans ce même chapitre, on peut noter que la valeur optimale de ζ dépend fortement de l'épaisseur optique. Aussi, si l'on s'éloigne des six cas d'étude présents dans ce benchmark, il pourrait être opportun de repenser le choix de cette valeur. On pourrait également imaginer un algorithme qui adapterait automatiquement la valeur de ζ à partir d'une estimation grossière de l'épaisseur optique.

5.3.4 Résultats obtenus pour les cas d'étude considérés

L'algorithme présenté à la Fig. 5.4 a été appliqué aux six cas d'étude rassemblés dans le benchmark proposé dans [André et Vaillon, 2010] avec les choix de paramètres libres présentés précédemment. Sauf indication contraire, les calculs ont été réalisés en considérant des profils de raie lorentziens et en utilisant une troncature des ailes de raie à une distance de $25cm^{-1}$ de leur nombre d'onde central.

Les Fig. 5.8, 5.9 et 5.10 ont pour but d'illustrer le type de résultats que l'on peut attendre de l'approche présentée dans ce chapitre. Il devient ainsi possible de modifier les hypothèses spectrales (troncature de raies, profil des raies, *etc.*) ou les données spectroscopiques à partir desquelles est décrit le coefficient d'absorption (bases de données, intensité en dessous de laquelle on néglige les raies, *etc.*) et de constater les effets de ces modifications sur une observable radiative sans avoir à produire de nouveaux spectres de haute-résolution en cohérence avec ces modifications.

Ces figures illustrent, pour le cas C2, les luminances $L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)$ moyennes pour plusieurs bandes étroites de $25cm^{-1}$ (définies de 1175 à $1925cm^{-1}$) estimées par l'algorithme introduit à la Fig. 5.4 avec 10^4 réalisations indépendantes. Les intervalles de confiance sont également fournis pour chaque estimation. À titre de validation, un calcul raie-par-raie basé sur des spectres haute-résolution (produits à partir de HITEMP 2010) est également représenté en trait plein sur chacun des trois graphiques. Quelles que soient les hypothèses ou données spectroscopiques considérées, chacune des simulations, dont les résultats sont illustrés par ces trois graphiques, a utilisé strictement le même algorithme et le même jeu de paramètres libres (en particulier de k_{η} , de \mathcal{P}_i et de $p_H(\eta)$).

La figure Fig. 5.8 rassemble les résultats de simulations lancées à partir de différentes bases de données : Hitemp 2010, Hitemp, Hitran 2008. La figure Fig. 5.9 illustre le fait qu'il est possible de mesurer l'effet des seuils minimums d'intensité (en dessous desquels les raies sont négligées) sans avoir à produire de nouveaux spectres. Elle représente les luminances moyennées par bandes étroites calculées avec différentes valeurs de seuils minimums d'intensité à 1500K : 0molec.cm (toutes les raies sont sélectionnées), $3.10^{-27}molec.cm$, $10^{-22}molec.cm$ et $10^{-21}molec.cm$. Enfin, la Fig. 5.10 souligne le fait qu'il est également possible d'évaluer simplement l'impact qu'ont les paramètres de raies, en l'occurrence la distance à laquelle sont tronquées les ailes de raie (par rapport à leur nombre d'onde central), sur les luminances moyennées par bandes. Les résultats relatifs à plusieurs distances de troncature : ∞cm^{-1} (pas de troncature), $25cm^{-1}$, $5cm^{-1}$ et $0.5cm^{-1}$ y sont représentés.



FIGURE 5.8 – Luminances $L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)$ moyennées par bandes étroites pour le cas C2. Ce calcul a été réalisé grâce à l'algorithme de la Fig. 5.4 pour différentes bases de données spectroscopiques : HITEMP 2010, HITEMP et HITRAN 2008. Les résultats sont fournis avec leurs intervalles de confiance. Chaque point a été obtenu avec 10⁴ réalisations indépendantes. Un calcul déterministe, réalisé grâce à des spectres d'absorption produits à partir de la base de données HITEMP 2010, est également représenté en trait plein.

Les luminances $L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)$ moyennes représentées dans ces trois graphiques ont été estimées dans des temps de calcul assurant une erreur relative de 1% compris entre 0.1s et 2.4s (avec un processeur Intel Core i7 - 2.8GHz sans parallélisation). La taille de la base de données, le nombre de transitions sélectionnées ou encore le niveau de troncature des raies n'ont que très peu d'influence sur ces temps de calcul (les différences de temps de calcul sont principalement dues à la bande étroite considérée). Ces temps de calcul affichés ne prennent pas en compte les temps de preprocessing ou de chargement des données en mémoire qui peuvent durer de quelques secondes à plusieurs minutes pour les bases de données les plus importantes. Cependant, une



FIGURE 5.9 – Luminances $L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)$ moyennées par bandes étroites pour le cas C2. Ce calcul a été réalisé grâce à l'algorithme de la Fig. 5.4 pour différents seuils minimums d'intensité (en dessous desquels, les raies sont négligées) : 0molec.cm (toutes les raies sont sélectionnées), $3.10^{-27}molec.cm$, $10^{-22}molec.cm$ et $10^{-21}molec.cm$. Les résultats sont fournis avec leurs intervalles de confiance. Chaque point a été obtenu avec 10^4 réalisations indépendantes. Un calcul déterministe, réalisé grâce à des spectres d'absorption produits à partir de la base de données HITEMP 2010 et un seuil minimum d'intensité de $3.10^{-27}molec.cm$ à 1500K, est également représenté en trait plein.



FIGURE 5.10 – Luminances $L_{\eta}(\mathbf{x}_0, \mathbf{u}_0)$ moyennées par bandes étroites pour le cas C2. Ce calcul a été réalisé grâce à l'algorithme de la Fig. 5.4 pour différentes distances de troncature de raie : ∞cm^{-1} (pas de troncature), $25cm^{-1}$, $5cm^{-1}$ et $0.5cm^{-1}$. Les résultats sont fournis avec leurs intervalles de confiance. Chaque point a été obtenu avec 10^4 réalisations indépendantes. Un calcul déterministe, réalisé grâce à des spectres d'absorption produits à partir de la base de données HITEMP 2010 et une troncature de raie à $25cm^{-1}$, est également représenté en trait plein.

fois que le preprocessing et le chargement des données en mémoire sont effectués, il est possible de lancer autant de simulations différentes que désirées (pour tester différentes hypothèses spectrales ou encore pour évaluer différentes observables radiatives).

La Tab. 5.1, rassemble les luminances $L(\mathbf{x}_0, \mathbf{u}_0)$ intégrées spectralement entre $10cm^{-1}$ et $15000cm^{-1}$ calculées à partir de l'algorithme introduit à la Fig. 5.4 et avec les choix des paramètres libres énoncés dans la Sec. 5.3.3. Ces calculs ont été réalisés avec 10^6 réalisations indépendantes, pour chacun des 6 cas d'étude et pour différentes bases de données spectroscopiques. Ces résultats sont fournis avec leur écart-type σ assimilable à un intervalle de confiance, et avec le temps de calcul $t_{1\%}$ requis pour obtenir une erreur relative de 1% (ces temps de calcul excluent ici aussi les étapes de preprocessing et de chargement des données spectroscopiques à partir d'un calcul déterministe (notées $L_{hr}(\mathbf{x}_0, \mathbf{u}_0)$) basé sur une production de spectres de haute-résolution (voir Sec. 5.1.1) utilisant les mêmes modèles spectraux et données d'entrée que pour l'algorithme à collisions nulles et aux résultats obtenus avec une approche raie-par-raie par F. André et R. Vaillon dans l'article [André et Vaillon, 2010] (notés $L_{a,v}(\mathbf{x}_0, \mathbf{u}_0)$).

		Monte-Carlo $(10^6 \text{ realisations})$		Haute résol.	[André et Vaillon, 2010]	
Cas	Bases de données	$L(\mathbf{x}_0, \mathbf{u}_0)$	σ	$t_{1\%}$	$L_{hr}(\mathbf{x}_0,\mathbf{u}_0)$	$L_{a,v}(\mathbf{x}_0,\mathbf{u}_0)$
d'étude	spectroscopiques	$(W/m^2/sr)$	$(W/m^2/sr)$	(s)	$(W/m^2/sr)$	$(W/m^2/sr)$
C1	CDSD-1000	3125.61	4.42	0.97	3126.06	3105
	CDSD-4000	3146.25	4.53	1.10	3150.32	
C2	HITEMP	3315.11	8.15	1.38	3311.88	4161
	HITEMP 2010	4545.05	9.83	1.11	4558.68	
C3	CDSD-1000 & HITEMP	39223.87	51.56	1.75	39202.5	39331
C4	CDSD-1000	12325.99	16.16	1.26	12320.1	11956
C5	HITEMP	38240.31	49.58	1.27	38215.0	39144
C6	HITEMP	885.93	3.93	9.86	886.55	-
	HITEMP 2010	1066.92	4.30	7.39	1069.81	-

TABLE 5.1 – Luminances intégrées de $10cm^{-1}$ à $15000cm^{-1}$ pour les six cas d'étude présentés à la Sec. 5.3.1 et pour différentes bases de données spectroscopiques (CDSD-1000 et CDSD-4000 pour le CO₂; HITEMP et HITEMP 2010 pour l'H₂O). Ces luminances $L(\mathbf{x}_0, \mathbf{u}_0)$ ont été estimées à partir de l'algorithme introduit à la Fig. 5.4 et sont fournies avec leur écart-type σ et le temps de calcul nécessaire à l'obtention d'une erreur relative de 1%. Les luminances $L(\mathbf{x}_0, \mathbf{u}_0)$ peuvent être comparées à celles obtenues à partir d'un calcul déterministe (notées $L_{hr}(\mathbf{x}_0, \mathbf{u}_0)$) basé sur une production de spectres de haute-résolution utilisant les mêmes hypothèses et données d'entrée que pour l'algorithme à collisions nulles et aux résultats obtenus avec une approche raie-par-raie par F. André et R. Vaillon dans l'article [André et Vaillon, 2010] (notés $L_{a,v}(\mathbf{x}_0, \mathbf{u}_0)$). Les temps de calcul excluent les étapes de preprocessing et de chargement des données spectroscopiques en mémoire.

Les résultats de l'algorithme proposé dans ce chapitre concordent parfaitement avec les luminances $L_{hr}(\mathbf{x}_0, \mathbf{u}_0)$ calculées de façon déterministe (pour les mêmes données spectroscopiques et un même modèle de raie). Les différences constatées avec les résultats de [André et Vaillon, 2010] sont inférieures aux effets mesurés lors du changement de bases de données, soulignant probablement une disparité dans les données spectroscopiques utilisées. Les temps de calcul requis par l'algorithme de Monte-Carlo pour assurer une erreur relative de 1% sont raisonnables dans le positionnement qui est le nôtre. Ils sont compris entre 1s et 8s selon le cas d'étude (sur un seul cœur d'un processeur Intel Core i7 - 2.8GHz). On constate alors que les temps de calcul relatifs à une intégration de la luminance sur une bande étroite de $25cm^{-1}$ ou sur tout le domaine infrarouge sont du même ordre de grandeur, ce qui est directement dû au caractère statistique de l'approche proposée. Enfin on notera, que grâce à la nature de la méthode utilisée (algorithme à collisions nulles), les champs de température continus des cas C3, C4 et C5 ont pu être pris en compte sans recourir à la moindre discrétisation volumique des propriétés du milieu (ce qui n'est pas le cas pour le calcul déterministe ou encore dans [André et Vaillon, 2010]).

5.3.5 Coefficients $h_{n,m,i,\eta}$ négatifs et événements rares

Au cours des travaux menés pour proposer des choix satisfaisants de $\hat{k}_{\eta}(\mathbf{x})$, $\mathcal{P}_{m}(\mathbf{x})$ et $\mathcal{P}_{i}(\mathbf{x})$ et pour prouver la faisabilité de l'approche faisant l'objet de ce chapitre, la principale difficulté a été d'éviter de rencontrer des plages spectrales et spatiales dans lesquelles les coefficients de collision nulle $h_{n,m,i,\eta}(\mathbf{x})$ étaient négatif. Cependant, même avec les choix proposés à la Sec. 5.3.3 et un facteur correctif α fixé à 50, nous n'avons pas pu assurer en toute généralité $h_{n,m,i,\eta}(\mathbf{x}) > 0$. Nous avons simplement réussi à réduire la quantité de ces événements à un nombre suffisamment faible pour qu'ils ne posent plus de problème d'augmentation brutale de variance telle que rencontrée dans le Chap. 4. En effet, assurer la condition $h_{n,m,i,\eta}(\mathbf{x}) > 0$ équivaut à garantir que le champ arbitraire de $\hat{k}_{\eta}(\mathbf{x})$ soit plus grand que $h_{a,m,i,\eta}(\mathbf{x})/[\mathcal{P}_{m}(\mathbf{x})\mathcal{P}_{i}(\mathbf{x})]$ pour tout nombre d'onde η , en tout point \mathbf{x} , pour toute espèce moléculaire m et pour toute transition i. Il est donc très difficile d'assurer de façon exhaustive le caractère majorant du champ de \hat{k}_{η} .

Pour un champ de $\hat{k}_{\eta}(\mathbf{x})$ fixé, majorant largement le champ du coefficient d'absorption, le non-respect de la condition $h_{n,m,i,\eta}(\mathbf{x}) > 0$ est alors dû à une mauvaise définition des probabilités $\mathcal{P}_m(\mathbf{x})$ et $\mathcal{P}_i(\mathbf{x})$ qui sont localement (pour un nombre d'onde, une position et un nombre d'onde donnés) sous-estimées de façon importante. Or, puisque cinq des six cas d'étude constituant le benchmark sont composés de gaz mono-moléculaires, ces coefficients négatifs de collision nulle sont causés par un choix imparfait de \mathcal{P}_i . Avec les choix faits pour cette probabilité, décomposée comme le produit de \mathcal{P}_{β} et de $\mathcal{P}_{i,\beta}$ (voir Sec. 5.3.3), nous rencontrons deux principaux types de cas pathologiques (illustrés à la Fig. 5.11) :

- Ceux dus à une sous-estimation de $\mathcal{P}_{i,\beta}$. On les rencontre lorsque la bande spectrale β échantillonnée est celle à laquelle appartient le nombre d'onde η d'intéret et dans laquelle est centrée une raie très intense (en vert sur la Fig. 5.11a). Si, une raie de plus faible intensité (en rouge sur la Fig. 5.11a) est centrée à une proximité directe du nombre d'onde η , sa contribution $h_{a,m,i,\eta}(\mathbf{x})$ en η peut être très importante, mais sa probabilité $\mathcal{P}_{i,\beta}$ est très sous-estimée à cause de la raie très intense présente également dans β .
- Ceux dus à une sous-estimation de \mathcal{P}_i à cause de \mathcal{P}_{β} . Ces événements sont rencontrés lorsque la bande β échantillonnée est très distante du nombre d'onde η d'intérêt et donc a une probabilité \mathcal{P}_{β} très faible, mais qu'au sein de cette bande est présente une raie très intense (en rouge sur la Fig. 5.11b) avec une largeur de raie importante. Il est possible que dans cette configuration, ce soit

cette raie qui constitue l'essentiel du coefficient d'absorption $k_{a,\eta}(\mathbf{x})$ global en η , rendant alors la probabilité \mathcal{P}_i largement sous-estimée à cause de la très faible valeur de \mathcal{P}_{β} associée à β .



FIGURE 5.11 – Principaux cas dans lesquels la condition $h_{n,m,i,\eta}(\mathbf{x}) > 0$ n'est pas respectée à cause d'une sous-estimation de la probabilité $\mathcal{P}_{i,\beta}$ (voir Fig. (A)) ou d'une sous-estimation de la probabilité \mathcal{P}_{β} (voir Fig. (B)).

Si ces coefficients négatifs de collision nulle $h_{n,m,i,\eta}$ ont un impact important sur le taux de convergence de l'algorithme, c'est qu'ils entraînent, à cause du caractère récursif de l'algorithme, une augmentation importante de la valeur absolue du poids de Monte-Carlo. Cette augmentation est d'autant plus grande qu'il y a, au cours d'une même réalisation indépendante, plusieurs collisions caractérisées par $h_{n,m,i,\eta} < 0$ (voir Sec. 5.2.3). Les réalisations, pour lesquelles un nombre très important de coefficients négatifs de collision nulle est rencontré, se comportent alors statistiquement comme des événements rares (voir Sec. 3.2.3.2) : ces réalisations ont de très faibles probabilités d'occurrence, mais des poids de Monte-Carlo très importants, conduisant à une augmentation importante de l'écart-type de l'estimation de la grandeur d'intérêt. Bien que problématiques, ces événements particuliers doivent être échantillonnés pour que la statistique du problème soit bien prise en compte par l'algorithme de Monte-Carlo. Dans le cas contraire, les résultats produits par ce type d'algorithme peuvent être biaisés.

De tels événements rares sont couramment rencontrés lorsque l'on explore de nouvelles approches statistiques, s'éloignant des pratiques habituelles. Il y a toujours un risque que la statistique soit telle que des événements rares ne soient pas échantillonnés (ici une succession de collisions caractérisées par $h_{n,m,i,\eta} < 0$) et induisent une mauvaise évaluation de l'écart-type associé à l'estimation de la grandeur d'intérêt. Typiquement, un intervalle de confiance peut indiquer que les résultats sont précis à 1%, alors qu'à cause d'événements rares qui n'ont pas été échantillonnés, l'estimation d'une observable radiative est différente à plus de 10% de la valeur de la solution exacte.

Au cours des travaux présentés dans ce chapitre, nous avons rencontré de telles évaluations erronées. Pour nous assurer de la validité statistique des résultats, la première étape a été de les comparer à ceux obtenus par un calcul déterministe. Puis, nous avons augmenté le nombre de réalisations pour chacune des six simulations afin d'augmenter le nombre d'éventuels événements rares. Nous pouvons alors vérifier, que malgré cette augmentation du nombre d'échantillons, les calculs convergent bien et l'écart-type associé aux estimations évolue bien comme l'inverse de la racine carré du nombre de réalisations. Pour accroître notre confiance dans les écarts-types estimés, nous avons systématiquement calculé la variance de la variance de ces estimations (voir Sec. 3.2.3.2) et avons vérifié que cette dernière était toujours inférieure à 0.1 (comme conseillé dans [X-5 Monte Carlo Team, 2008]). Dans le cas présent, la variance de la variance est toujours inférieure à 0.005, quel que soit le cas d'étude du Benchmark. Enfin, pour éviter que les coefficients négatifs de collision nulle n'entraînent une erreur relative trop importante, nous avons augmenté la valeur du champ de $\hat{k}_{\eta}(\mathbf{x})$ pour la maille à laquelle appartient \mathbf{x} et pour un petit intervalle spectral de $1cm^{-1}$, lorsque la condition $h_{n,m,i,\eta}(\mathbf{x}) > 0$ n'est pas respectée. Cela permet ainsi d'éviter que, dans la suite de la réalisation ou même de la simulation, cet événement particulier, source de variance, ne se reproduise. Dans la mesure où l'on ne modifie pas le poids associé à la collision pour laquelle $h_{n,m,i,\eta}(\mathbf{x}) < 0$, cette procédure n'entraîne aucun biais (puisque le champ de $k_{\eta}(\mathbf{x})$ est entièrement arbitraire) et permet de contrôler d'éventuelles augmentations importantes de variance. En poussant ce raisonnement un peu plus loin, on peut imaginer des spectres de \hat{k}_{η} qui s'enrichiraient au cours des simulations pour réduire petit à petit le nombre de coefficients $h_{n,m,i,\eta}(\mathbf{x})$ négatifs rencontrés lors d'un calcul.

Aussi, grâce à ce processus de validation, nous sommes confiants quant à la précision et au caractère non-biaisé de l'approche introduite dans ce chapitre pour des configurations typiques de celles proposées dans le benchmark de André et Vaillon, 2010 (en termes de compositions chimiques, de propriétés thermodynamiques et de tailles caractéristiques). Mais pour d'autres champs applicatifs, nous suggérons que la variance de la variance soit systématiquement calculée et que des simulations déterministes soient associées aux premiers exercices d'exploration. Nous pensons en particulier aux applications atmosphériques et astrophysiques dans lesquelles les distances d'intérêt sont beaucoup plus importantes et pour lesquelles les faibles pressions induisent des profils de raies très étroits (qui pourraient être susceptibles de complexifier la procédure d'échantillonnage des transitions ou des nombres d'onde). Mais pour des applications usuelles de combustion, la statistique semble être bien maîtrisée avec les propositions faites ici, et il n'y a pas de surprise que 10^5 échantillons soient suffisants pour évaluer avec une précision d'1% une grandeur radiative qui implique des millions de transitions. En effet, un grand nombre de transitions moléculaires sont prises en compte pour le calcul d'un seul poids de Monte-Carlo, et les probabilités associées à chacune de ces transitions ont été choisies de sorte à ce que les transitions non-échantillonnées aient une contribution similaire à celles échantillonnées.

Résumé du chapitre

Une conséquence directe de l'introduction de collisions nulles dans l'équation du transfert radiatif est que les coefficients d'absorption n'apparaissent plus dans le terme d'extinction exponentielle, mais seulement de façon linéaire dans les albédos. À travers ce chapitre, nous avons montré comment, grâce à cette propriété, il est possible de décrire de façon statistique les coefficients d'absorption à partir des transitions moléculaires, directement au sein de l'équation du transfert radiatif. Le calcul d'une grandeur radiative est alors ramené à une simple estimation d'espérance. Cela conduit alors au développement d'algorithmes de Monte-Carlo de référence, permettant d'estimer une observable radiative directement à partir de bases de données spectroscopiques, sans passer par une coûteuse production de spectres d'absorption. La mise en pratique de cette approche requiert cependant de définir un grand nombre de paramètres arbitraires ayant des conséquences importantes sur la qualité de convergence de ces algorithmes. Nous nous sommes alors appuyé sur un benchmark de six cas d'étude caractéristiques de configurations de combustion pour proposer un choix concernant ces paramètres libres. Ces propositions sont probablement loin d'être optimales, mais elles permettent d'assurer de bons taux de convergence quel que soit le cas d'étude : quelques secondes sont nécessaires pour estimer, sans approximation de modèle ou numérique, avec une précision de 1%, une luminance intégrée sur tout le domaine infrarouge. Les perspectives qu'offre une telle méthode sont nombreuses du fait qu'elle ne nécessite plus de production rigoureuse de spectres d'absorption haute-résolution. D'un point de vue numérique, il n'est plus nécessaire à chaque changement d'hypothèse spectrale ou de base de données spectroscopique de produire à nouveau un jeu complet de spectres. Cette méthode a également des conséquences en termes d'analyse, il devient notamment possible de calculer de façon exacte la sensibilité d'une grandeur radiative intégrée spectralement à un paramètre du modèle de raie, aux variables d'état thermodynamiques ou encore aux fractions molaires d'espèces présentes dans le milieu gazeux d'étude.