

Stage de fin d'études

Extension d'une méthode Monte-Carlo dans un cadre instationnaire

Stagiaire : Mathis Fitoussi (fitoussi@enst.fr)
Encadrement : Laetitia Laguzet (laetitia.laguzet@cea.fr)
Organisme d'accueil : CEA

Stage de fin d'études - Télécom Paris



INSTITUT
POLYTECHNIQUE
DE PARIS



DÉCLARATION D'INTÉGRITÉ RELATIVE AU PLAGIAT

Je soussigné Fitoussi Mathis, étudiant à Télécom Paris, déclare être pleinement conscient que le plagiat d'un document ou d'une partie de document publié sur toutes les formes existantes de support constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée. En conséquence, je m'engage à citer explicitement, à chaque fois que j'en fais usage, toutes les sources que j'ai utilisées pour écrire ce rapport.

En outre, j'atteste sur l'honneur que je suis l'auteur du présent travail et que tout ce qui a été emprunté est attribué à sa source et proprement référencé.

Fait à Arpajon, le 13 octobre 2022

Signature



Remerciements

Merci à mon encadrante, Laetitia Laguzet, pour m'avoir laissé la latitude d'explorer les problématiques de ce stage par moi-même. Cette liberté m'a permis de mieux m'appropriier les différents concepts qui ont été abordés. Son exigence a été pour moi un guide qui m'a forcé à rester rigoureux tout au long du stage. Elle a aussi facilité la coordination entre le stage, ma future thèse et mes colles.

Je remercie également Cédric Eaux, qui a su être à l'écoute et veiller au bon déroulement du stage sur le plan humain.

Je tiens également à saluer les collègues de mon open space et le personnel de Teratec, qui ont su être accueillants et m'intégrer pleinement à la vie sur le campus, et ont rendu ce stage encore plus entraînant.

Enfin, j'adresse une pensée à mes amis qui m'ont redonné la motivation dont j'avais besoin lorsque celle-ci manquait.

Table des matières

| | | |
|----------|--|-----------|
| 1 | Introduction | 5 |
| 1.1 | Positionnement du problème | 5 |
| 1.2 | Objectifs | 6 |
| 2 | Obtention d'une équation linéaire de transport | 7 |
| 2.1 | Approximation IMC | 7 |
| 2.2 | Procédure générale | 9 |
| 2.3 | Interprétation de l'équation de transport | 9 |
| 3 | Méthodes Monte Carlo pour la résolution d'équations de transport | 11 |
| 3.1 | Méthode Monte Carlo Adjointe | 11 |
| 3.2 | Méthode Monte Carlo Directe | 13 |
| 3.3 | État de l'art en méthode adjointe | 14 |
| 4 | Équation sans terme source ni dépendance spatiale : recherche de la meilleure approximation de $u(t_i, \cdot)$ | 15 |
| 4.1 | Résolution analytique | 15 |
| 4.2 | Estimateur méthode adjointe | 16 |
| 4.2.1 | Définition | 16 |
| 4.2.2 | Biais | 17 |
| 4.2.3 | Variance | 17 |
| 4.3 | Estimateur méthode directe | 18 |
| 4.3.1 | Définition | 18 |
| 4.3.2 | Biais | 18 |
| 4.3.3 | Variance | 19 |
| 4.4 | Comparaison des variances | 19 |
| 4.5 | Méthodes d'approximation de u_1 | 19 |
| 4.5.1 | Approximation par valeurs moyennes | 20 |
| 4.5.2 | Approximation par interpolation linéaire | 22 |
| 4.5.3 | Approximation par un polynôme de degré d | 23 |
| 4.5.4 | Approximation par des polynômes de Bernstein | 25 |
| 4.5.5 | Fonctions u_0 utilisées pour des tests | 26 |
| 4.5.6 | Configuration des tests | 26 |
| 4.5.7 | Résultats | 28 |
| 4.5.8 | Conclusion sur la meilleure méthode d'approximation | 30 |
| 4.6 | Comparaison MCA - MCD sans terme source ni dépendance spatiale | 30 |
| 4.7 | Estimateur adjoint sans approximation intermédiaire | 33 |
| 5 | Ajout d'un terme source puis d'un maillage | 34 |
| 5.1 | Ajout d'un terme source | 34 |
| 5.1.1 | Résolution analytique avec terme source | 34 |
| 5.1.2 | Estimateur adjoint associé | 36 |
| 5.1.3 | Estimateur direct associé | 37 |
| 5.1.4 | Discretisation déterministe contre FullMC | 38 |
| 5.1.5 | Comparaison MCA - MCD avec terme source, sans dépendance spatiale | 39 |
| 5.2 | Ajout d'une dépendance spatiale | 41 |
| 5.2.1 | Introduction du maillage | 41 |
| 5.2.2 | Particules qui sortent du maillage en méthode adjointe | 41 |
| 5.2.3 | Validation des méthodes - équation de transport avec dépendance spatiale | 41 |

| | | |
|----------|--|-----------|
| 5.2.4 | Comparaison MCA - MCD avec terme source et dépendance spatiale | 48 |
| 5.3 | Étude sur de nombreuses itérations d'IMC | 48 |
| 5.4 | Approximation par k -plus proches voisins | 51 |
| 5.4.1 | Présentation de la méthode | 51 |
| 5.4.2 | Paramètres de la méthode | 52 |
| 5.4.3 | Performance | 53 |
| 6 | Conclusion | 56 |
| 7 | Annexes | 57 |
| 7.1 | Présentation du CEA | 57 |
| 7.1.1 | Le CEA | 57 |
| 7.1.2 | La DAM | 57 |
| 7.2 | Preuve de la formule de Feynman-Kac | 58 |
| 7.3 | Conditions aux limites | 60 |
| 7.3.1 | Idée générale | 60 |
| 7.3.2 | Cas général | 60 |
| 7.3.3 | Exemple | 61 |

1 Introduction

1.1 Positionnement du problème

Dans le contexte des simulations de fusion par confinement inertiel, les équations du transfert radiatif modélisent les interactions entre la matière et le rayonnement :

$$\frac{1}{c} \frac{\partial u}{\partial t}(t, x, \omega) + \omega \cdot \nabla_x u(t, x, \omega) + \sigma_T u(t, x, \omega) = \frac{1}{4\pi} \sigma_T a c T^4(t, x) \quad (1)$$

$$\frac{\partial u_m}{\partial t}(T, t, x) = \sigma_T \left[\int_S u(t, x, \omega) d\omega - a c T^4(t, x) \right] + s(t, x, \omega) \quad (2)$$

Où :

- $t \in (0, t_{max})$
- $\omega \in S = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ représente une direction de propagation
- $x \in \mathcal{D} \subset \mathbb{R}^3$
- c est la vitesse de la lumière
- $\sigma_T = \sigma_T(x, t)$ est l'opacité totale
- a est la constante radiative
- $T = T(t, x)$ est la température du matériau
- $u_m(T(t, x))$ est la densité énergétique du matériau
- s est un terme source

Ces équations couplées et non linéaires en la température étant difficiles à résoudre, on va montrer comment dans le cadre de l'approximation Implicit Monte Carlo (IMC), décrit en premier par Fleck et Cummings en 1971 [1], on peut discrétiser en temps et en espace pour se ramener sur chaque pas de temps et chaque maille à une équation de transport de la forme :

$$\frac{1}{c} \frac{\partial u}{\partial t}(t, x, \omega) + \omega \cdot \nabla_x u(t, x, \omega) + \sigma_T(x) u(t, x, \omega) = \sigma_s(x) \int_S u(t, x, \omega') \frac{d\omega'}{4\pi} + f(t, x, \omega) \quad (3)$$

L'approche classique (directe) consiste à générer un grand nombre de particules représentant à chaque instant la densité énergétique u de l'équation de transport, puis de les laisser évoluer selon une loi reflétant l'équation (3). Cette méthode est réputée être la plus adaptée pour calculer u pour "tous les x, v ".

L'idée de base de la méthode adjointe est de générer les particules selon une répartition finale que l'on choisit, et de remonter le temps en simulant leur trajectoire en sens inverse depuis l'instant final vers l'instant initial.

De cette manière, il est possible de choisir le nombre de particules à destination de chaque zone de l'espace, alors qu'en méthode directe, celui-ci dépend de nombreux paramètres, pas tous contrôlables : la densité énergétique dans la maille, la discrétisation en espace, les paramètres physiques du milieu ou encore la méthode d'échantillonnage choisie.

On décrira ici les méthodes de Monte Carlo Direct (MCD) et Adjointe (MCA) pour les résolutions d'équations de transport générales, avant de les appliquer à (3). La description de ces méthodes est faite par Lapeyre, Pardoux et Sentis dans [2], mais l'analyse faite pour la méthode MCA y est limitée : il manque notamment des informations sur la manière d'implémenter un terme source ou les traitements des conditions aux limites. Nous tenterons donc de la compléter pour pouvoir implémenter une méthode adjointe dans le contexte IMC.

1.2 Objectifs

Ce stage est prospectif et est une étude de faisabilité pour la méthode adjointe. Classiquement, les équations de transport dans le cadre de l'IMC sont résolues par des méthodes directes, et le laboratoire souhaite savoir quels sont les avantages et inconvénients de la méthode adjointe comparée à une méthode directe.

Comme nous allons le voir, si on emploie une méthode MCA pour la résolution de l'équation de transport sur chaque intervalle de temps $[t_i, t_{i+1}]$, il est nécessaire de savoir approcher efficacement les $u(t_i, \cdot)$, i.e. la solution à (3) aux instants t_i . C'est le premier objectif de ce stage, qui sera étudié dans la section 4 en s'intéressant à une version simplifiée de l'équation de transport, sans dépendance spatiale ni terme source.

Dans [2], lorsque l'on ajoute un terme source en MCD, les différentes intégrales en temps des estimateurs sont calculées via des méthodes déterministes d'intégration. On se demandera s'il n'est pas plus efficace de remplacer cette approche par une approche de type Monte-Carlo en section 5.1.4.

Une autre question posée dans ce stage est celle de la gestion des conditions aux limites. Les deux méthodes permettent d'avoir facilement des conditions aux limites de type absorption ou réflexion. Comme on va le voir, la méthode directe permet naturellement de traiter une condition de type flux entrant. Ce n'est pas du tout le cas en MCA. [2] ne donne pas non plus de précisions pour leur gestion en méthode adjointe. Comme on le verra en partie 3.3, la littérature ne semble pas non plus répondre à cette question, et il va falloir proposer des méthodes pour simuler un flux entrant.

2 Obtention d'une équation linéaire de transport

2.1 Approximation IMC

Nous allons montrer comment, à partir des équations du transfert radiatif (1) et (2), on retrouve l'équation (3) évoquée dans l'introduction. Cette partie est en majorité une retranscription de [1].

Soient $u_r(x, t) = aT^4(t, x)$ et $\beta = \frac{\partial u_r}{\partial u_m}$. (2) peut alors s'écrire :

$$\frac{1}{\beta} \frac{\partial u_r}{\partial t}(t, x) = \sigma_T \left[\int_S u(t, x, \omega) d\omega - cu_r \right] + s(t, x, \omega)$$

Dans cette équation, toute la non-linéarité en température est contenue dans β , qui sera par la suite considéré constant sur chaque intervalle de temps où nous résoudrons l'équation, ce qui la rendra donc linéaire.

On va maintenant intégrer l'équation précédente sur un pas de temps $[t_n, t_{n+1}]$:

$$u_r(t_{n+1}, x) - u_r(t_n, x) = \int_{t_n}^{t_{n+1}} \beta \sigma_T \int_S u(t, x, \omega) d\omega dt - \int_{t_n}^{t_{n+1}} \beta \sigma_T cu_r(t, x) dt + \int_{t_n}^{t_{n+1}} s(t, x, \omega) dt$$

La stratégie de Monte-Carlo Implicite (IMC) consiste alors à fixer β , σ_T et T à leur valeur initiale¹ sur le pas de temps $[t_n, t_{n+1}]$. On prendra donc :

$$\begin{aligned} T_n(x) &= T(t_n, x) \\ \beta_n(x) &= \beta(x, t_n) = \frac{\partial u_r / \partial T}{\partial u_m / \partial T} = \frac{4aT_n^3(x)}{c_v(x, T_n)} \\ \sigma_{T,n}(x) &= \sigma_T(x, t_n) \end{aligned}$$

Où $c_v(x, T_n) := \partial u_m / \partial T$.

On va maintenant chercher une valeur de u_r , que l'on notera \tilde{u}_r , et que l'on va injecter dans (1).

On choisira cette valeur à partir de 2, afin que cette dernière soit respectée.

On introduit l'opérateur de moyenne temporelle sur un pas de temps :

$$\bar{(\cdot)} = \frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} (\cdot) dt$$

On fait l'hypothèse². que $\exists \alpha \in (0, 1)$:

$$\frac{1}{\Delta t} \int_{t_n}^{t_{n+1}} u_r(t, x) dt = \bar{u}_r(x) = \alpha u_r(t_{n+1}, x) + (1 - \alpha) u_r(t_n, x) \quad (4)$$

On obtient alors :

$$u_r(t_{n+1}, x) - u_r(t_n, x) = \Delta t \beta_n \sigma_{T,n} \left(\int_S \bar{u}(x, \omega) d\omega - c[\alpha u_r(t_{n+1}, x) + (1 - \alpha) u_r(t_n, x)] \right) + \Delta t \beta_n \bar{s}$$

1. Pour augmenter la précision, on peut choisir d'autres valeurs que la valeur initiale. Voir par exemple [3] et [4].
2. Un tel α existe toujours par égalité des accroissements finis, mais on le choisit ici de manière arbitraire. Pour des raisons de stabilité détaillées dans [5], il faut prendre $\alpha \geq \frac{1}{2}$, et on prendra le plus souvent $\alpha = 1$

$$\Leftrightarrow u_r(t_{n+1}, x) [1 + \alpha\beta_n c \Delta t \sigma_{T,n}] = u_r(t_n, x) [1 - (1 - \alpha)\beta_n c \Delta t \sigma_{T,n}] + \Delta t \beta_n \sigma_{T,n} \int_S \bar{u}(x, \omega) d\omega + \Delta t \beta_n \bar{s}$$

$$\Leftrightarrow u_r(t_{n+1}, x) = u_r(t_n, x) \frac{1 - (1 - \alpha)\beta_n c \Delta t \sigma_{T,n}}{1 + \alpha\beta_n c \Delta t \sigma_{T,n}} + \frac{\Delta t \beta_n \sigma_{T,n}}{1 + \alpha\beta_n c \Delta t \sigma_{T,n}} \int_S \bar{u}(x, \omega) d\omega + \frac{\Delta t \beta_n \bar{s}}{1 + \alpha\beta_n c \Delta t \sigma_{T,n}}$$

On note $F_n = \frac{1}{1 + \alpha\beta_n c \Delta t \sigma_{T,n}}$ le facteur de Fleck.

On injecte cette expression dans 4 :

$$\begin{aligned} \bar{u}_r &= \alpha u_r(t_{n+1}, x) + (1 - \alpha) u_r(t_n, x) \\ &= \alpha u_r(t_n, x) F_n (1 - (1 - \alpha)\beta_n c \Delta t \sigma_{T,n}) + F_n \Delta t \beta_n \sigma_{T,n} \int_S \bar{u}(x, \omega) d\omega + F_n \Delta t \beta_n \bar{s} + (1 - \alpha) u_r(t_n, x) \\ &= u_r(t_n, x) F_n (\alpha - (1 - \alpha)\alpha\beta_n c \Delta t \sigma_{T,n}) + F_n \alpha \Delta t \beta_n \sigma_{T,n} \int_S \bar{u}(x, \omega) d\omega + F_n \alpha \Delta t \beta_n \bar{s} + (1 - \alpha) u_r(t_n, x) \\ &= u_r(t_n, x) F_n [(1 - \alpha)(1 + \alpha\beta_n c \Delta t \sigma_{T,n}) + \alpha - (1 - \alpha)\alpha\beta_n c \Delta t \sigma_{T,n}] + F_n \alpha \Delta t \beta_n \left[\sigma_{T,n} \int_S \bar{u}(x, \omega) d\omega + \bar{s} \right] \\ &= u_r(t_n, x) F_n + F_n \alpha \Delta t \beta_n \left[\sigma_{T,n} \int_S \bar{u}(x, \omega) d\omega + \bar{s} \right] \end{aligned}$$

L'hypothèse faite en 4 est donc équivalente au fait que

$$\bar{u}_r(x) = u_r(t_n, x) F_n + F_n \alpha \Delta t \beta_n \left[\sigma_{T,n} \int_S \bar{u}(x, \omega) d\omega + \bar{s} \right]$$

Ainsi, une condition suffisante pour vérifier l'hypothèse est de prendre

$$\tilde{u}_r(t, x) = u_r(t_n, x) F_n + F_n \alpha \Delta t \beta_n \left[\sigma_{T,n} \int_S u(t, x, \omega) d\omega + s \right]$$

On injecte dans (1) :

$$\begin{aligned} \frac{1}{c} \frac{\partial u}{\partial t}(t, x, \omega) + \omega \cdot \nabla_x u(t, x, \omega) + \sigma_{T,n} u(t, x, \omega) &= \frac{c \sigma_{T,n}}{4\pi} \tilde{u}_r \\ &= \frac{F_n c \sigma_{T,n}}{4\pi} u_r(t_n, x) + \frac{F_n \sigma_{T,n} \alpha \Delta t \beta_n c}{4\pi} \left[\sigma_{T,n} \int_S u(t, x, \omega) d\omega + s \right] \\ &= \frac{F_n \sigma_{T,n} \alpha \Delta t \beta_n c}{4\pi} \sigma_{T,n} \int_S u(t, x, \omega) d\omega + \frac{F_n c \sigma_{T,n}}{4\pi} [u_r(t_n, x) + \alpha \Delta t \beta_n s] \end{aligned}$$

Posons $\sigma_{s,n} = F_n \sigma_{T,n} \alpha \Delta t \beta_n c \sigma_{T,n} = \sigma_{T,n} (1 - F_n) \in (0, \sigma_{T,n})$ et $f_n = \frac{F_n c \sigma_{T,n}}{4\pi} [u_r(t_n, x) + \alpha \Delta t \beta_n s]$.

$$\frac{1}{c} \frac{\partial u}{\partial t}(t, x, \omega) + \omega \cdot \nabla_x u(t, x, \omega) + \sigma_{T,n}(x) u(t, x, \omega) = \sigma_{s,n}(x) \int_S u(t, x, \omega') \frac{d\omega'}{4\pi} + f_n(t, x, \omega)$$

On va donc pouvoir résoudre de manière autonome cette équation de transport via des méthodes de Monte-Carlo que nous détaillerons par la suite, et à chaque itération en temps utiliser le résultat obtenu pour mettre à jour les paramètres pour l'itération suivante.

2.2 Procédure générale

Soit $k \in \mathbb{N}$. On considère connus $u_k = u(t_k, \cdot, \cdot)$, $T_k = T(t_k, x)$, et les fonctions $\sigma_T(x, t)$ et $c_v(x, T)$. On connaît donc $\beta_k = \frac{4aT_k^3(x)}{c_v(x, T_k)}$ et on peut, via la méthode de notre choix, approcher une trajectoire de u sur $[t_k, t_{k+1}]$ suivant l'équation

$$\frac{1}{c} \frac{\partial u}{\partial t}(t, x, \omega) + \omega \cdot \nabla_x u(t, x, \omega) + \sigma_{T,k}(x) u(t, x, \omega) = \sigma_{s,k}(x) \int_S u(t, x, \omega') \frac{d\omega'}{4\pi} + f_k(t, x, \omega)$$

Pour pouvoir simuler une trajectoire à l'étape suivante, il nous faut calculer les paramètres $\beta_{k+1} = \frac{4aT_{k+1}^3(x)}{c_v(x, T_{k+1})}$ et $u_r(t_{k+1}) = aT_{k+1}^4$, et donc il faut une estimation de T_{k+1} . On rappelle que l'équation (2) s'écrit :

$$\begin{aligned} \frac{\partial u_m}{\partial t}(t, x) &= \sigma_{T,k} \left[\int_S u(t, x, \omega) d\omega - c\tilde{u}_r \right] + \bar{s}(x, \omega) \\ \iff u_m(t_{k+1}) - u_m(t_k) &= \sigma_{T,k} \left[\int_{t_k}^{t_{k+1}} \int_S u(t, x, \omega) d\omega dt - c\Delta t \tilde{u}_r \right] + \Delta t \beta_k \bar{s}(x, \omega) \end{aligned}$$

D'autre part,

$$u_m(t_{k+1}) - u_m(t_k) = \int_{T_k}^{T_{k+1}} c_v(T') dt$$

Et on en déduit T_{k+1} puis β_{k+1} et $u_r(t_{k+1})$. Il sera également courant dans les applications que $s = 0$, ce que l'on supposera dorénavant. La quantité d'intérêt que l'on doit calculer avec nos méthodes de Monte Carlo est donc de la forme :

$$\int_{t_k}^{t_{k+1}} \int_S u(t, x, \omega) d\omega dt$$

Dans notre cas, on va de plus fixer toutes nos quantités par mailles : dans une maille M , on considère $\sigma_{T,k}$, $\sigma_{s,k}$ et β_k constants. La quantité d'intérêt devient donc :

$$\int_{t_k}^{t_{k+1}} \int_M \int_S u(t, x, \omega) d\omega dx dt$$

2.3 Interprétation de l'équation de transport

L'équation de transport

$$\frac{1}{c} \frac{\partial u}{\partial t}(t, x, \omega) + \omega \cdot \nabla_x u(t, x, \omega) + (\sigma_a + \sigma_s)(x) u(t, x, \omega) = \sigma_s(x) \int_S u(t, x, \omega') \frac{d\omega'}{4\pi} + f(t, x, \omega)$$

s'interprète de la manière suivante :

On va considérer que l'énergie est transportée par un grand nombre de photons qui peuvent chacun subir plusieurs événements. Ceux-ci se reflètent dans les termes de l'équation :

- Le terme $\frac{1}{c} \frac{\partial u}{\partial t}(t, x, \omega) + \omega \cdot \nabla_x u(t, x, \omega)$ indique que la direction de propagation est $\omega \in S$, avec une vitesse de c .³

3. Dans toutes les applications, on prendra $c = 1$ pour éviter d'obtenir des ordres de grandeur disproportionnés

- Le terme $\sigma_a(x)u(t, x, \omega)$ est le terme d'absorption. Lorsqu'un photon traverse la matière, il peut être absorbé et donner son énergie à la matière. On traduit cela dans l'équation à un niveau macroscopique par une déposition d'énergie dans la matière au cours du temps. Cela va induire un terme en $e^{-\sigma_a t}$ dans nos équations.
- Le terme $\sigma_s(x) \int_S u(t, x, \omega') \frac{d\omega'}{4\pi}$ est lié au "scattering" (dispersion) des photons. Un photon peut subir une collision et changer sa direction de propagation. On considère ici que le changement de direction est isotrope, et donc l'intégrale est contre une densité uniforme sur S .
- Enfin, on peut considérer un terme source f , qui agit comme une source créatrice de nouveaux photons et transcrit l'émission de photons par la matière chaude.

L'idée d'une simulation Monte-Carlo est alors de simuler un grand nombre de photons subissant ces événements afin d'estimer $u(t, x, \omega)$.

3 Méthodes Monte Carlo pour la résolution d'équations de transport

On suppose dans la suite que la condition initiale u_0 , le terme source f et la fonction ϕ contre laquelle on intègre u sont positives, et on pose :

$$\begin{aligned} \alpha &= \int_{\mathbb{R}^3} \int_S u_0(x, \omega) dx d\omega & \tilde{u}_0 &= \frac{u_0}{\alpha} \\ \beta &= \int_{\mathbb{R}^3} \int_S \phi(x, \omega) dx d\omega & \tilde{\phi} &= \frac{\phi}{\beta} \\ \gamma : t &\mapsto \int_{\mathbb{R}^3} \int_S f(t, x, \omega) dx d\omega & \tilde{f} &= \frac{f}{\gamma} \end{aligned}$$

On notera également $\sigma_a = \sigma_T - \sigma_s$.

3.1 Méthode Monte Carlo Adjointe

Définition 3.1. On appelle "processus adjoint" le processus markovien déterministe par morceaux $(X(t), \Omega(t))_{t \geq 0}$ sous-jacent à la méthode Monte Carlo adjointe, qui est défini de la manière suivante :

- La loi initiale est déterministe : $X(0) = x$ et $\Omega(0) = \omega$
- L'intensité des sauts est σ_s . Lors d'un saut, $\Omega(t)$ prend une valeur aléatoire selon une loi uniforme sur S et $X(t)$ ne change pas. Entre les sauts, le processus vérifie

$$\frac{d}{ds} \begin{pmatrix} X(s) \\ \Omega(s) \end{pmatrix} = - \begin{pmatrix} c\Omega(s) \\ 0 \end{pmatrix}$$

Dans la pratique, on le génère comme suit :

1. Une particule est initialisée en une position $X(0) = x$ et avec une direction de propagation $\Omega(0) = \omega$.
2. A partir de la position de la particule, on détermine dans quelle maille elle se trouve, puis on récupère les valeurs de σ_a et σ_s pour cette maille.
3. Jusqu'au prochain évènement (scattering, changement de maille, sortie de domaine ou fin d'itération), la particule se déplace selon l'équation

$$\frac{d}{ds} \begin{pmatrix} X(s) \\ \Omega(s) \end{pmatrix} = - \begin{pmatrix} c\Omega(s) \\ 0 \end{pmatrix}$$

4. Pour déterminer le prochain évènement, on génère une variable exponentielle de paramètre σ_s qui représente le temps avant le prochain évènement de scattering.

Si la particule sort du domaine avant cet évènement, on termine son historique.

Si la particule change de maille avant cet évènement, on retourne en 2.

Sinon, la particule subit un évènement de scattering (i.e. sa direction de propagation ω saute vers une nouvelle valeur tirée selon une loi uniforme sur la sphère unité S), et on retourne en 2.

La méthode MCA repose sur l'équation de Feynman-Kac, qui indique que pour un processus adjoint $(X(t), \Omega(t))$, la solution à (3) peut s'écrire de la manière suivante :

Théorème 3.1. ⁴ Pour une condition initiale u_0 donnée, on a :

$$u(t, x, \omega) = \mathbb{E}_{x, \omega} \left[u_0(X(t), \Omega(t)) e^{-\int_0^t \sigma_a(t-\zeta, X(\zeta), \Omega(\zeta)) d\zeta} + \int_0^t f(t-s, X(s), \Omega(s)) e^{-\int_0^s \sigma_a(t-\zeta, X(\zeta), \Omega(\zeta)) d\zeta} ds \right]$$

On peut alors évaluer des quantités de la forme :

$$I_{u_0}(t, \phi) := \int_S \int_{\mathbb{R}^3} u(t, x, \omega) \phi(x, \omega) dx d\omega$$

Il suffit pour cela de ne plus sélectionner la condition initiale de manière déterministe mais de l'échantillonner selon $\tilde{\phi}$, car on a alors :

$$I_{u_0}(t, \phi) = \mathbb{E}_{\tilde{\phi}} \left[u_0(X(t), \Omega(t)) e^{-\int_0^t \sigma_a(t-\zeta, X(\zeta), \Omega(\zeta)) d\zeta} + \int_0^t f(t-s, X(s), \Omega(s)) e^{-\int_0^s \sigma_a(t-\zeta, X(\zeta), \Omega(\zeta)) d\zeta} ds \right]$$

Si on se donne N réalisations indépendantes $(X_i(t), \Omega_i(t))$ du processus adjoint, on peut associer à cette espérance un "estimateur adjoint" :

$$\hat{I}_{u_0}^{adj}(t, \phi) = \sum_{i=1}^N \left[u_0(X_i(t), \Omega_i(t)) w_i(t) + \int_0^t f(t-s, X_i(s), \Omega_i(s)) e^{-\int_0^s \sigma_a(t-\zeta, X_i(\zeta), \Omega_i(\zeta)) d\zeta} ds \right]$$

Où $w_i(t)$ est le poids de la particule. Il suit l'équation différentielle :

$$w_i(0) = \frac{1}{N}$$

$$\forall s \in (0, t), w_i'(s) = \sigma_a(t-s, X_i(s), \Omega_i(s)) w_i(s)$$

Lorsque σ_a ne dépend pas du temps sur l'itération considérée et est constant sur chaque maille, ce poids peut être calculé à faible coût en l'initialisant à $\frac{1}{N}$ à l'instant initial pour chaque particule, puis en le mettant à jour à chaque événement (changement de maille, scattering ou sortie de domaine) en le multipliant par $e^{-\sigma_a \Delta t}$, où Δt est le temps écoulé depuis le dernier événement.

Dans notre cas, ϕ sera l'indicatrice d'une maille. On voit dès lors que cette méthode a pour avantage de nous laisser choisir combien de particules sont générées pour chaque maille. Ce ne sera pas le cas avec la méthode directe. C'est cette observation qui a initialement motivé le stage : dans le cas d'une méthode directe, les mailles trop petites ou celles recevant des quantités trop faibles de particules présentent des variances élevées, ce qui implique de mettre en place une gestion de la population de particules au cours du temps. L'espoir est de pouvoir corriger ce défaut à l'aide de la méthode adjointe. En effet, en méthode adjointe, il n'y a pas besoin de réserver des particules pour le terme source, car toute l'expression précédente est sous une seule espérance. En revanche, il est nécessaire de générer de nouvelles particules à chaque pas de temps $[t_i, t_{i+1}]$ selon la densité ϕ , puis de les évaluer en $u(t_i, \cdot, \cdot)$ ⁵. Il faut donc une estimation de $u(t_i, \cdot, \cdot)$ à chaque pas de temps.

Remarquons que si on veut retrouver $u(t, x, \omega)$, il suffit d'appliquer la méthode avec $\phi(x, \omega) = \delta_x \otimes \delta_\omega$ (Dirac pour choisir x et ω de manière déterministe). Dans la suite, on ne décrira donc cette

4. Cette formule diffère légèrement de celle donnée dans [2]. D'une part, il semblerait que [2] contienne une coquille, et d'autre part, la formule que je propose ici est une généralisation qui permet d'avoir σ_a dépendant du temps. La preuve donnée en annexe reprend donc les notations du chapitre 2 de [2].

5. La généralisation proposée de la formule de Feynman-Kac permet en théorie de se passer de ce besoin, mais on verra en partie 4.7 qu'elle implique un fort coût en mémoire

méthode que pour un ϕ quelconque. Par ailleurs, la méthode MCD ne permet pas le calcul de u en un point donné mais seulement de quantités intégrées, et il est donc plus simple pour les comparer de les exprimer de la même manière.

Remarque 3.1. *On ne fait pour l'instant pas mention des conditions aux limites, qui deviennent l'objet principal du stage à partir du 5ème mois.*

3.2 Méthode Monte Carlo Directe

Définition 3.2. *On appelle "processus direct" le processus markovien déterministe par morceaux $(X(t), \Omega(t))_{t \geq 0}$ sous-jacent à la méthode Monte Carlo directe, qui est défini de la manière suivante :*

- *La densité de la loi initiale pour $(X(0), \Omega(0))$ est \tilde{u}_0*
- *L'intensité des sauts est σ_s . Lors d'un saut, $\Omega(t)$ prend une valeur aléatoire selon une loi uniforme sur S et $X(t)$ ne change pas. Entre les sauts, le processus vérifie*

$$\frac{d}{ds} \begin{pmatrix} X(s) \\ \Omega(s) \end{pmatrix} = \begin{pmatrix} c\Omega(s) \\ 0 \end{pmatrix}$$

Dans la pratique, on le génère comme suit :

1. *Une particule est initialisée en une position $X(0) = x$ et avec une direction de propagation $\Omega(0) = \omega$ avec (x, ω) générés selon la densité \tilde{u}_0 .*
2. *A partir de la position de la particule, on détermine dans quelle maille elle se trouve, puis on récupère les valeurs de σ_a et σ_s pour cette maille.*
3. *Jusqu'au prochain évènement (scattering, changement de maille, sortie de domaine ou fin d'itération), la particule se déplace selon l'équation*

$$\frac{d}{ds} \begin{pmatrix} X(s) \\ \Omega(s) \end{pmatrix} = \begin{pmatrix} c\Omega(s) \\ 0 \end{pmatrix}$$

4. *Pour déterminer le prochain évènement, on génère une variable exponentielle de paramètre σ_s qui représente le temps avant le prochain évènement de scattering. Si la particule sort du domaine avant cet évènement, on termine son historique. Si la particule change de maille avant cet évènement, on retourne en 2. Sinon, la particule subit un évènement de scattering (i.e. sa direction de propagation ω saute vers une nouvelle valeur tirée selon une loi uniforme sur la sphère unité S), et on retourne en 2.*

La méthode MCD repose, elle, sur l'équation de Fokker-Plank :

Théorème 3.2. *Pour une condition initiale u_0 donnée, on a :*

$$\int_S \int_{\mathbb{R}^3} u(t, x, \omega) \phi(x, \omega) dx d\omega = \alpha \mathbb{E}_{0, \tilde{u}_0} \left[\phi(X(t), \Omega(t)) e^{-\int_0^t \sigma_a(\zeta, X(\zeta), \Omega(\zeta)) d\zeta} \right] \\ + \int_0^t \gamma(s) \mathbb{E}_{s, \tilde{f}(s, \cdot)} \left[\phi(X(t), \Omega(t)) e^{-\int_s^t \sigma_a(\zeta, X(\zeta), \Omega(\zeta)) d\zeta} \right] ds$$

Où $\mathbb{E}_{s, r(s, \cdot)}$ indique que la loi d'échantillonnage de la condition initiale est $r(s, \cdot)$ et que l'instant initial du processus est s (i.e. les particules sous-jacentes à l'espérance sont définies sur (s, t)).

On voit cette fois que :

- Le nombre de particules dans chaque maille n'est pas maîtrisé : on génère des particules selon la valeur de l'énergie initiale en chaque point, mais on ne peut pas contrôler la maille dans laquelle une particule finit son parcours.

- La présence d’une espérance dans l’intégrale implique de réserver un grand nombre de particules pour le terme source. Lorsque l’on discrétise cette intégrale en temps (que ce soit de manière déterministe ou Monte Carlo), il faudra calculer l’intégrande en un certain nombre de points, et donc calculer le terme en $\mathbb{E}_{s,f(s,\cdot)}$ autant de fois.
- Dans le cadre d’une discrétisation type IMC où l’on résout l’équation sur des intervalles $[t_i, t_{i+1}]$, on conserve les particules d’un pas de temps pour le pas de temps suivant, car à tout instant t , les particules sont un échantillonnage de $u(t, \cdot, \cdot)$.

3.3 État de l’art en méthode adjointe

Classiquement, la méthode adjointe est employée lorsque l’on a besoin d’une mesure précise sur une petite zone de l’espace, comme par exemple si on veut prédire u sur un détecteur (comme dans [6] pour le transport de muons ou [7], dans des cadres limités : stationnaires, pas de flux entrant).

Dans [8], la méthode adjointe est utilisée, cette fois non pas pour étudier les solutions à une équation de transport sur un détecteur, mais pour intégrer l’énergie sur différentes zones d’angles solide dans des applications à l’astrophysique. Elle a alors pour objectif de réduire la variance. Cet article ne couvre toujours qu’un cas stationnaire sans flux entrant ni terme source.

Dans [9], les auteurs s’attaquent à un problème de stabilité qui apparaît lorsque le scattering est anisotrope . Dans le cas anisotrope, les particules ont tendance à se diriger vers les énergies les plus faibles lors d’un évènement de scattering, donc en méthode adjointe, vers les plus fortes, ce qui entraîne des instabilités car les particules gagnent une grande quantité d’énergie au cours de la simulation.

Les techniques classiques de réduction de variance employées sont le "splitting" (scinder des particules dont le poids est grand en plusieurs particules) et la "roulette russe" (supprimer des particules dont le poids est trop faible) (voir [5], [10] et [11]). Les versions basiques de ces méthodes entraînent une non-conservation de l’énergie (lors de la suppression d’une particule par exemple). Certains auteurs proposent des adaptations conservant l’énergie ([12]).

En plus de leur emploi en réduction de variance, ces techniques sont également employées à des fins de régulation de la population de particules. Cette problématique de la gestion du nombre de particules est centrale en méthode directe, car, en présence d’un terme source, il est nécessaire d’émettre en continu de nouvelles particules, et il faut donc un moyen de gérer leur population.

Pour les approximations IMC, il est également possible d’utiliser une estimation "a priori" de T_{n+1} afin de gagner en précision (voir [3] et [4]). On peut également envisager d’utiliser ces quantités "a priori" pour déterminer la taille du prochain pas de temps, comme suggéré dans [3].

Il semblerait cependant que la méthode adjointe n’ait pour l’instant pas été utilisée dans un contexte IMC où des approximations successives des $u(t_i, \cdot, \cdot)$ sont nécessaires, ni dans une situation impliquant un flux entrant comme conditions aux limites.

4 Équation sans terme source ni dépendance spatiale : recherche de la meilleure approximation de $u(t_i, \cdot)$

Comme expliqué plus haut, on va appliquer les méthodes MCA et MCD sur des intervalles de temps $0 = t_0 < \dots < t_p = T$. Cela ne pose pas de problèmes en MCD car on peut, entre les pas de temps, conserver les particules que l'on génère pour nos estimations et simplement les évaluer aux différents instants t_i . À l'inverse, en MCA, il faut générer en début de chaque intervalle des particules dans la maille étudiée pour les évaluer plus tard selon $u(t_i, \cdot)$, ce qui nécessite d'approcher efficacement cette fonction. Dans cette section, on va étudier une version simplifiée de l'équation de transport (3) en ignorant la dépendance en espace et en prenant $f = 0$ dans le but de proposer des approximations de cette densité. On remplace donc, dans le cadre de cette étude, la dépendance spatiale par une dépendance en angle, afin de pouvoir exploiter une solution analytique.

On commence par l'équation à résoudre en $u \in C^1$ en temps et C^0 en ω , avec σ_s et σ_a constants (étude 0D-plan) :

$$\begin{aligned} \frac{\partial u(t, \omega)}{\partial t} + \sigma_T u(t, \omega) &= \sigma_s \int_{-1}^1 u(t, \omega') \frac{d\omega'}{2} \\ u_0 &= u(0, \cdot) \end{aligned} \quad (5)$$

L'étude de cette équation a deux avantages :

- En supprimant la dépendance en espace, on est capable de trouver une solution analytique
- Dans le cadre de l'approximation IMC, on cherche à calculer des expressions du type :

$$\int_M \int_S \int_{t_i}^{t_{i+1}} u(t, x, \omega) dt d\omega dx$$

Ici, on peut transposer le problème de l'intégrale "sur une maille" en un problème d'intégrale sur "une partie de sphère" D en ω et donc s'intéresser à trouver de bonnes estimations de

$$\int_D \int_{t_i}^{t_{i+1}} u(t, \omega) dt d\omega$$

Avec cette problématique en tête, on va comparer différentes méthodes d'approximation de la fonction $\omega \mapsto u(t_1, \omega)$ pour t_1 fixé.

Pour rappel, on fait cette discrétisation car l'approximation IMC demande de changer les σ ainsi que f à chaque intervalle de temps. On considèrera donc que sur chaque intervalle les σ sont constants. On verra en fin de section comment on peut se passer de cette contrainte et proposer un estimateur sans instant intermédiaire, au prix d'un fort coût en mémoire.

4.1 Résolution analytique

Proposition 4.1. *La solution à l'équation (5) est donnée par :*

$$u(t, \omega) = \frac{W(0)}{2} e^{-\sigma_a t} + e^{-\sigma_T t} \left(u_0(\omega) - \frac{W(0)}{2} \right)$$

Avec $\forall t > 0, W : t \mapsto \int_{-1}^1 u(t, \omega) d\omega$

Démonstration. On intègre (5) sur $[-1, 1]$:

$$\frac{\partial}{\partial t} \int_{-1}^1 u(t, \omega) d\omega + \sigma_T \int_{-1}^1 u(t, \omega) d\omega = \sigma_s \int_{-1}^1 \left(\int_{-1}^1 u(t, \omega') \frac{d\omega'}{2} \right) d\omega = \sigma_s \int_{-1}^1 u(t, \omega) d\omega$$

On a posé $W : t \mapsto \int_{-1}^1 u(t, \omega) d\omega$, en sorte que :

$$W'(t) + \underbrace{(\sigma_T - \sigma_s)}_{\sigma_a} W(t) = 0$$

$$\implies \forall t \geq 0, W(t) = W(0) \exp(-\sigma_a t) = \left(\int_{-1}^1 u_0(\omega) d\omega \right) \exp(-\sigma_a t)$$

On injecte dans (5) :

$$\frac{\partial u(t, \omega)}{\partial t} + \sigma_T u(t, \omega) = \frac{\sigma_s W(0)}{2} \exp(-\sigma_a t)$$

On multiplie par $e^{\sigma_T t}$:

$$e^{\sigma_T t} \frac{\partial u(t, \omega)}{\partial t} + \sigma_T e^{\sigma_T t} u(t, \omega) = \frac{\sigma_s W(0)}{2} e^{-\sigma_a t} e^{\sigma_T t}$$

$$\iff e^{\sigma_T t} \frac{\partial u(t, \omega)}{\partial t} + \frac{\partial e^{\sigma_T t}}{\partial t} u(t, \omega) = \frac{\sigma_s W(0)}{2} e^{\sigma_s t}$$

$$\iff \frac{\partial (u(t, \omega) e^{\sigma_T t})}{\partial t} = \frac{W(0)}{2} \frac{\partial e^{\sigma_s t}}{\partial t}$$

$$\iff u(t, \omega) e^{\sigma_T t} = \frac{W(0) e^{\sigma_s t}}{2} + K$$

Avec $K = u(0, \omega) e^0 - \frac{W(0)}{2} e^0 = u_0(\omega) - \frac{W(0)}{2}$

$$\implies u(t, \omega) = \frac{W(0)}{2} e^{-\sigma_a t} + e^{-\sigma_T t} \left(u_0(\omega) - \frac{W(0)}{2} \right)$$

□

Remarque 4.1. Soit ϕ une fonction positive sur $[-1, 1]$. Notons $\beta = \int_{-1}^1 \phi(\omega) d\omega$. On peut alors calculer :

$$I_{u_0}(t, \phi) = \int_{-1}^1 u(t, \omega) \phi(\omega) d\omega = \frac{W(0)}{2} e^{-\sigma_a t} \beta + e^{-\sigma_T t} \left(\int_{-1}^1 u_0(\omega') \phi(\omega) d\omega - \frac{W(0)}{2} \beta \right)$$

4.2 Estimateur méthode adjointe

4.2.1 Définition

Soient $(\Omega_i(t))_{i \in \{1, \dots, N\}}$ des réalisations i.i.d. du processus adjoint. L'estimateur adjoint s'écrit :

$$\hat{I}_{N, u_0}^{adj}(t, \phi) = \frac{1}{N} \sum_{i=1}^N \beta u_0(\Omega_i) \exp(-t\sigma_a)$$

4.2.2 Biais

D'après la formule de Feynman-Kac (3.1), cet estimateur est sans biais. On va retrouver cela ici en utilisant la solution analytique de l'équation et la loi de $\Omega(t)$.

Pour X définie sur un espace de probabilité, on note sa densité $f_X(x)$.

On note N_t le processus de poisson décrivant l'évolution de $\Omega(t)$, d'intensité σ_s . On a alors :

$$\begin{aligned}
f_{\Omega(t)}^{adj}(\omega) &= \sum_{n=0}^{\infty} f_{N_t, \Omega(t)}(n, \omega) \\
&= \sum_{n=0}^{\infty} f_{\Omega(t) | N_t=n}(\omega) f_{N_t}(n) \\
&= \sum_{n=0}^{\infty} f_{\Omega(t) | N_t=n}(\omega) e^{-\sigma_s t} \frac{(\sigma_s t)^n}{n!} \\
&= \tilde{\phi}(\omega) e^{-\sigma_s t} + \sum_{n=1}^{\infty} \frac{\mathbb{1}_{[-1,1]}(\omega)}{2} e^{-\sigma_s t} \frac{(\sigma_s t)^n}{n!} \\
&= \tilde{\phi}(\omega) e^{-\sigma_s t} + \frac{\mathbb{1}_{[-1,1]}(\omega)}{2} (1 - e^{-\sigma_s t})
\end{aligned}$$

On en déduit :

$$\begin{aligned}
\mathbb{E}(u_0(\Omega(t))) &= e^{-\sigma_s t} \mathbb{E}_{\Omega \sim \phi}(u_0(\Omega)) + (1 - e^{-\sigma_s t}) \mathbb{E}_{\Omega \sim \mathcal{U}([-1,1])}(u_0(\Omega)) \\
\implies \mathbb{E}(\hat{I}_{N, u_0}^{adj}(t, \phi)) &= \beta e^{-\sigma_a t} \mathbb{E}(u_0(\omega)) \\
&= \beta e^{-\sigma_a t} \left(e^{-\sigma_s t} \mathbb{E}_{\Omega \sim \tilde{\phi}}(u_0(\Omega)) + (1 - e^{-\sigma_s t}) \mathbb{E}_{\Omega \sim \mathcal{U}([-1,1])}(u_0(\Omega)) \right) \\
&= \beta e^{-\sigma_a t} \int_{-1}^1 \frac{\phi(\omega)}{\beta} u_0(\omega) d\omega + \beta e^{-\sigma_a t} (1 - e^{-\sigma_s t}) \frac{W(0)}{2} \\
&= \beta e^{-\sigma_a t} \left(\int_{-1}^1 \frac{\phi(\omega)}{\beta} u_0(\omega) d\omega - \frac{W(0)}{2} \right) + \beta e^{-\sigma_a t} \frac{W(0)}{2} \\
&= I_{u_0}(t, \phi)
\end{aligned}$$

On conclut avec la loi des grands nombres, qui nous donne la convergence presque sûre de cet estimateur vers $I_{u_0}(t, \phi)$.

4.2.3 Variance

On calcule également la variance pour se donner une idée de la performance de l'estimateur :

$$\begin{aligned}
\mathbb{V}(\hat{I}_{N, u_0}^{adj}(t, \phi)) &= \frac{1}{N} \left[\mathbb{E}((\beta u_0(\Omega) e^{-t\sigma_a})^2) - I_{u_0}(t, \phi)^2 \right] \\
&= \frac{1}{N} \left[\beta^2 e^{-\sigma_s t} \int_{-1}^1 e^{-2\sigma_a t} u_0^2(\omega) \phi(\omega) d\omega + \beta^2 (1 - e^{-\sigma_s t}) \int_{-1}^1 e^{-2\sigma_a t} u_0^2(\omega) \frac{d\omega}{2} - I_{u_0}(t, \phi)^2 \right] \\
&= \frac{1}{N} \left[\beta^2 e^{-\sigma_s t} \int_{-1}^1 e^{-2\sigma_a t} u_0^2(\omega) \left[\phi(\omega) - \frac{1}{2} \right] d\omega + \beta^2 \int_{-1}^1 e^{-2\sigma_a t} u_0^2(\omega) \frac{d\omega}{2} - I_{u_0}(t, \phi)^2 \right]
\end{aligned}$$

Le théorème central limite indique alors que :

$$\sqrt{N} \frac{\hat{I}_{N,u_0}^{adj}(t, \phi) - I_{u_0}(t, \phi)}{\sqrt{\mathbb{V}(\hat{I}_{N,u_0}^{adj}(t, \phi))}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

On pourra confirmer ce théorème en vérifiant que l'erreur que l'on obtient décroît en $\frac{1}{\sqrt{N}}$ (valable pour toutes les méthodes de type Monte Carlo).

4.3 Estimateur méthode directe

4.3.1 Définition

Soient $(\Omega_i(t))_{i \in \{1, \dots, N\}}$ des réalisations i.i.d. du processus direct. L'estimateur direct s'écrit :

$$\hat{I}_{N,u_0}^{dir}(t, \phi) = \frac{1}{N} \sum_{i=1}^N \alpha \phi(\Omega_i(t)) e^{-\sigma_a t}$$

Soit, pour l'intégrale sur $[a, b]$:

$$\begin{aligned} \hat{I}_{N,u_0}^{dir}(t, \mathbb{1}_{[a,b]}) &= \frac{1}{N} \sum_{i=1}^N \alpha \mathbb{1}_{[a,b]}(\Omega_i(t)) e^{-\sigma_a t} \\ \implies \mathbb{E} \left(\hat{I}_{N,u_0}^{dir}(t, \mathbb{1}_{[a,b]}) \right) &= \mathbb{E} \left(\alpha e^{-\sigma_a t} \mathbb{1}_{[a,b]}(\Omega_i(t)) \right) \end{aligned}$$

On voit alors déjà un des inconvénients de cette méthode : si on veut calculer $I_{\mathbb{1}_{[a,b]}}(t)$, un certain nombre de particules va être simulé sans apporter de contribution, car on perd les particules qui arrivent en dehors du support de ϕ ou n'y passent pas. Cela est d'autant plus un problème lorsque l'on intègre sur de petits intervalles (a proche de b , i.e. des mailles petites), car très peu de particules vont finir ou passer dans cet intervalle. Si on compare cela à la méthode adjointe, cette dernière permet de choisir le nombre de particules qui contribueront dans chaque maille. On s'attend donc à ce que pour de petits intervalles, la variance de la méthode adjointe soit plus faible.

4.3.2 Biais

Un calcul similaire à celui de la méthode MCA donne :

$$f_{\Omega(t)}^{dir}(\omega) = \tilde{u}_0(\omega) e^{-\sigma_s t} + \frac{\mathbb{1}_{[-1,1]}(\omega)}{2} (1 - e^{-\sigma_s t})$$

Avec le même calcul que pour l'estimateur adjoint et en inversant les rôles de u et ϕ , on obtient :

$$\begin{aligned} \mathbb{E}(\phi(\Omega(t))) &= e^{-\sigma_s t} \mathbb{E}_{\Omega \sim \tilde{u}_0}(\phi(\Omega)) + (1 - e^{-\sigma_s t}) \mathbb{E}_{\Omega \sim \mathcal{U}(S)}(\phi(\Omega)) \\ \implies \mathbb{E}(\hat{I}_{N,u_0}^{dir}(t, \phi)) &= e^{-\sigma_a t} \alpha \left[e^{-\sigma_s t} \int_{-1}^1 \phi(\omega) \frac{u_0(\omega)}{\alpha} d\omega + (1 - e^{-\sigma_s t}) \int_{-1}^1 \frac{\phi(\omega)}{2} d\omega \right] \\ &= e^{-\sigma_a t} \int_{-1}^1 u_0(\omega) \phi(\omega) d\omega + \alpha e^{-\sigma_a t} (1 - e^{-\sigma_s t}) \frac{1}{2} \\ &= e^{-\sigma_a t} \int_{-1}^1 u_0(\omega) \phi(\omega) d\omega + e^{-\sigma_a t} \int_{-1}^1 u_0(\omega) \frac{d\omega}{2} - e^{-\sigma_a t} \frac{W(0)}{2} \\ &= I_{u_0}(t, \phi) \end{aligned}$$

Cette fois-ci, en revanche, il est impossible d'obtenir directement une estimation de $u(t, \omega)$ en prenant pour ϕ un dirac en ω car l'estimateur n'aurait plus de sens.

4.3.3 Variance

Via un calcul similaire à la partie précédente :

$$\mathbb{V}(\hat{I}_{N,u_0}^{dir}(t)) = \frac{1}{N} \left[e^{-\sigma_s t} e^{-2\sigma_a t} \alpha^2 \int_{-1}^1 \phi^2(\omega) \left[\tilde{u}_0(\omega) - \frac{1}{2} \right] d\omega + \alpha^2 \int_{-1}^1 e^{-2\sigma_a t} \phi^2(\omega) \frac{d\omega}{2} - I_{u_0}(t, \phi)^2 \right]$$

De même,

$$\sqrt{N} \frac{\hat{I}_{N,u_0}^{dir}(t, \phi) - I_{u_0}(t, \phi)}{\sqrt{\mathbb{V}(\hat{I}_{1,u_0}^{dir}(t, \phi))}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

4.4 Comparaison des variances

On compare ici directement avec les versions normalisées de u_0 et ϕ .

$$\begin{aligned} N e^{2\sigma_a t} \Delta \mathbb{V}(t) &= N e^{-2\sigma_a t} \left(\mathbb{V}(\hat{I}_{N,u_0}^{adj}(t)) - \mathbb{V}(\hat{I}_{N,u_0}^{dir}(t)) \right) \\ &= e^{-\sigma_s t} \int_{-1}^1 \tilde{u}_0^2(\omega) \left[\tilde{\phi}(\omega) - \frac{1}{2} \right] - \tilde{\phi}^2(\omega) \left[\tilde{u}_0(\omega) - \frac{1}{2} \right] d\omega + \int_{-1}^1 \tilde{u}_0^2(\omega) - \tilde{\phi}^2(\omega) d\omega \end{aligned}$$

Supposons qu'on ait une situation du type u_0 constante et $\phi = \mathbb{1}_{[a,b]}$ (ce qui est une situation classique correspondant à l'estimation du premier moment angulaire de u dans certaines directions seulement). Dans ce cas, $\tilde{u}_0 = \frac{1}{2}$.

$$\begin{aligned} N e^{2\sigma_a t} \Delta \mathbb{V}(t) &= e^{-\sigma_s t} \int_{-1}^1 \tilde{u}_0^2(\omega) \left[\tilde{\phi}(\omega) - \frac{1}{2} \right] - \tilde{\phi}^2(\omega) \left[\tilde{u}_0(\omega) - \frac{1}{2} \right] d\omega + \int_{-1}^1 \tilde{u}_0^2(\omega) - \tilde{\phi}^2(\omega) d\omega \\ &= e^{-\sigma_s t} \left[\frac{1}{4} - \frac{1}{4} - \frac{\frac{1}{2} - \frac{1}{2}}{b-a} \right] + \frac{1}{2} - \frac{1}{b-a} \\ &= \frac{(b-a) - 2}{2(b-a)} \\ &\leq 0 \end{aligned}$$

On retrouve alors que l'estimateur adjoint est toujours meilleur (du moins dès que $[a, b] \neq [-1, 1]$), et l'est d'autant plus lorsque a se rapproche de b , ce qui correspond à l'intuition guidée par le nombre de particules exploitées par chaque méthode. Cette conclusion repose néanmoins sur le fait d'utiliser autant de particules pour les deux méthodes. On verra plus loin que pour la méthode adjointe, il faut attribuer un nombre de particules à chaque maille, et donc partager le nombre total de particules entre les différentes parties de l'espace.

4.5 Méthodes d'approximation de u_1

On va maintenant proposer plusieurs méthodes pour l'approximation des densités successives. On se place dans un premier temps dans le cadre d'un u_0 connu de manière exacte, et on va chercher à quantifier l'erreur que commet chaque type d'approximation lors de l'estimation de $u(t_1, \cdot)$.

Les méthodes testées sont :

- Approximation par valeurs moyennes

- Approximation par interpolation linéaire
- Approximation par un polynôme de degré arbitraire
- Approximation par des polynômes de Bernstein

4.5.1 Approximation par valeurs moyennes

On découpe $[-1, 1]$ en n intervalles de longueurs égales sur lesquels on sait calculer la valeur moyenne de $u(t_1, \cdot)$, et y on approche la valeur de $u(t_1, \omega)$ par sa valeur moyenne :

$$\begin{aligned}
u_1^{moy}(\omega) &= \sum_{k=0}^{n-1} \frac{n}{2} \int_{-1+\frac{2k}{n}}^{-1+\frac{2(k+1)}{n}} u(t_1, \omega') d\omega' \mathbb{1}_{[-1+\frac{2k}{n}, -1+\frac{2(k+1)}{n}]}(\omega) \\
&= \sum_{k=0}^{n-1} \frac{n}{2} \int_{-1}^1 u(t_1, \omega') \mathbb{1}_{[-1+\frac{2k}{n}, -1+\frac{2(k+1)}{n}]}(\omega') d\omega' \mathbb{1}_{[-1+\frac{2k}{n}, -1+\frac{2(k+1)}{n}]}(\omega) \\
&= \sum_{k=0}^{n-1} I_{u_0}(t_1, \phi_k) \mathbb{1}_{[-1+\frac{2k}{n}, -1+\frac{2(k+1)}{n}]}(\omega)
\end{aligned}$$

Avec $\forall k, \phi_k = \frac{n}{2} \mathbb{1}_{[-1+\frac{2k}{n}, -1+\frac{2(k+1)}{n}]}(\omega)$. Cela donne l'estimateur :

$$\hat{u}_{1,N}^{moy}(\omega) = \sum_{k=0}^{n-1} \hat{I}_{N,u_0}^{adj}(t_1, \phi_k) \mathbb{1}_{[-1+\frac{2k}{n}, -1+\frac{2(k+1)}{n}]}(\omega)$$

Remarque 4.2. Cette méthode est par ailleurs économe car les $I_{u_0}(t_1, \phi_k)$ sont de toutes façons utilisés pour le calcul de

$$\int_D \int_{t_i}^{t_{i+1}} u(t, \omega) dt d\omega$$

qui est nécessaire pour calculer T_{n+1} dans le cadre de l'approximation IMC.

Proposition 4.2. Erreur de l'estimateur par valeurs moyennes

Soit $M = \sup_{\omega \in [-1,1], t \in [0, t_1]} \left| \frac{\partial u}{\partial \omega}(t, \omega) \right|$. $\forall \omega \in [-1, 1]$,

$$|u_1^{moy}(\omega) - u(t_1, \omega)| \leq \frac{M}{n}$$

Démonstration. Pour $k \in \{0, \dots, n-1\}$, si $\omega \in \left[-1 + \frac{2k}{n}, -1 + \frac{2(k+1)}{n}\right]$,

$$\begin{aligned}
|u_1^{moy}(\omega) - u(t_1, \omega)| &= \left| \frac{n}{2} \int_{-1+\frac{2k}{n}}^{-1+\frac{2(k+1)}{n}} u(t_1, \omega') d\omega' - u(t_1, \omega) \right| \\
&= \left| \frac{n}{2} \int_{-1+\frac{2k}{n}}^{-1+\frac{2(k+1)}{n}} [u(t_1, \omega') - u(t_1, \omega)] d\omega' \right| \\
&\leq \frac{n}{2} \int_{-1+\frac{2k}{n}}^{-1+\frac{2(k+1)}{n}} |u(t_1, \omega') - u(t_1, \omega)| d\omega' \\
&\leq \frac{n}{2} \int_{-1+\frac{2k}{n}}^{-1+\frac{2(k+1)}{n}} M |\omega' - \omega| d\omega' \\
&\leq \frac{Mn}{2} \left(\frac{2}{n}\right)^2 \frac{1}{2} \\
&\leq \frac{M}{n}
\end{aligned}$$

□

Ce résultat renseigne sur la partie "déterministe" de cette méthode d'approximation. En l'absence d'erreur stochastique, on est capable de majorer l'erreur d'approximation par M/n .

Proposition 4.3. Soit $\epsilon = \max_{i \in \{1, \dots, n\}} |\hat{I}_{N, u_0}^{adj}(t_1, \phi_i) - I_{u_0}(t_1, \phi_i)|$. $\forall \omega \in [-1, 1]$,

$$\left| \hat{u}_{1, N}^{moy}(\omega) - u_1^{moy}(\omega) \right| \leq \epsilon$$

Ce second résultat concerne l'impact d'une erreur stochastique sur le résultat : si on commet une erreur d'au plus ϵ sur les différents estimateurs, alors on commet une erreur d'au plus ϵ par rapport à une approximation par valeurs moyennes ne souffrant d'aucune erreur stochastique.

Remarque 4.3. *Conservation de l'énergie*

On peut constater que cette méthode a pour avantage de conserver l'énergie totale :

$$\begin{aligned} \int_{-1}^1 u_1^{moy}(\omega) d\omega &= \sum_{k=0}^{n-1} \int_{\omega_k}^{\omega_{k+1}} u_1^{moy}(\omega) d\omega \\ &= \sum_{k=0}^{n-1} \int_{\omega_k}^{\omega_{k+1}} \sum_{j=0}^{n-1} \frac{n}{2} \int_{\omega_j}^{\omega_{j+1}} u(t_1, \omega') d\omega' \mathbb{1}_{[\omega_j, \omega_{j+1}]}(\omega) d\omega \\ &= \sum_{j=0}^{n-1} \frac{n}{2} \int_{\omega_j}^{\omega_{j+1}} u(t_1, \omega') \left[\sum_{k=0}^{n-1} \underbrace{\int_{\omega_k}^{\omega_{k+1}} \mathbb{1}_{[\omega_j, \omega_{j+1}]}(\omega) d\omega}_{= \frac{2}{n} \delta_{j,k}} \right] d\omega' \\ &= \sum_{j=0}^{n-1} \int_{\omega_j}^{\omega_{j+1}} u(t_1, \omega') d\omega' \\ &= \int_{-1}^1 u_1(\omega') d\omega' \end{aligned}$$

Cette méthode est la seule à présenter cette propriété (bien qu'au moyen d'une renormalisation on puisse forcer les autres méthodes à conserver l'énergie également).

6. $x \mapsto \int_a^b |x - y| dy$ est maximale pour $x = a$ ou $x = b$, et vaut alors $(b - a)^2/2$

4.5.2 Approximation par interpolation linéaire

On se donne n points équidistants $-1 = \omega_1 < \dots < \omega_n = 1$ et on interpole entre deux points de manière linéaire :

$$\begin{aligned} u_1^{ipo}(\omega) &= \sum_{k=0}^{n-1} \left[\frac{u(t_1, \omega_{i+1}) - u(t_1, \omega_i)}{\omega_{i+1} - \omega_i} \omega + \frac{u(t_1, \omega_i)\omega_{i+1} - u(t_1, \omega_{i+1})\omega_i}{\omega_{i+1} - \omega_i} \right] \mathbb{1}_{[\omega_i, \omega_{i+1}[}(\omega) \\ &= \sum_{k=0}^{n-1} \left[\frac{u(t_1, \omega_{i+1})(\omega - \omega_i) - u(t_1, \omega_i)(\omega - \omega_{i+1})}{\omega_{i+1} - \omega_i} \right] \mathbb{1}_{[\omega_i, \omega_{i+1}[}(\omega) \\ &= \sum_{k=0}^{n-1} \left[I_{u_0}(t_1, \delta_{\omega_{i+1}}) \frac{\omega - \omega_i}{\omega_{i+1} - \omega_i} - I_{u_0}(t_1, \delta_{\omega_i}) \frac{\omega - \omega_{i+1}}{\omega_{i+1} - \omega_i} \right] \mathbb{1}_{[\omega_i, \omega_{i+1}[}(\omega) \end{aligned}$$

Et on peut alors prendre comme estimateur :

$$\hat{u}_1^{ipo}(\omega) = \sum_{k=0}^{n-1} \left[\hat{I}_{N, u_0}^{adj}(t_1, \delta_{\omega_{i+1}}) \frac{\omega - \omega_i}{\omega_{i+1} - \omega_i} - \hat{I}_{N, u_0}^{adj}(t_1, \delta_{\omega_i}) \frac{\omega - \omega_{i+1}}{\omega_{i+1} - \omega_i} \right] \mathbb{1}_{[\omega_i, \omega_{i+1}[}(\omega)$$

Proposition 4.4. *Erreur de l'estimateur par interpolation linéaire*

Soit $M = \sup_{\omega \in [-1, 1], t \in [0, t_1]} \left| \frac{\partial^2 u}{\partial \omega^2}(t, \omega) \right|$. $\forall \omega \in [-1, 1]$,

$$\left| u_1^{ipo}(\omega) - u(t_1, \omega) \right| \leq \frac{M}{2} \left(\frac{2}{n} \right)^2$$

Démonstration. Dans cette preuve seulement, on note $u(t_1, \omega) = u(\omega)$.

Soit $\omega \in [\omega_i, \omega_{i+1}[$. On écrit $u(\omega) = u(\omega) \frac{\omega - \omega_i}{\omega_{i+1} - \omega_i} - u(\omega) \frac{\omega - \omega_{i+1}}{\omega_{i+1} - \omega_i}$.

On utilisera également les développements de Taylor : $\exists \xi_1, \xi_2 \in [\omega_i, \omega_{i+1}[^2$ tels que :

$$\begin{aligned} u(\omega_{i+1}) &= u(\omega) + (\omega_{i+1} - \omega)u'(\omega) + \frac{(\omega_{i+1} - \omega)^2}{2}u''(\xi_1) \\ u(\omega_i) &= u(\omega) + (\omega_i - \omega)u'(\omega) + \frac{(\omega_i - \omega)^2}{2}u''(\xi_2) \end{aligned}$$

$$\begin{aligned} \left| u_1^{ipo}(\omega) - u(\omega) \right| &= \frac{1}{\omega_{i+1} - \omega_i} \left| (u(\omega_{i+1}) - u(\omega))(\omega - \omega_i) - (u(\omega_i) - u(\omega))(\omega - \omega_{i+1}) \right| \\ &= \frac{n}{2} \left| \left((\omega_{i+1} - \omega)u'(\omega) + \frac{(\omega_{i+1} - \omega)^2}{2}u''(\xi_1) \right) (\omega - \omega_i) - \left((\omega_i - \omega)u'(\omega) + \frac{(\omega_i - \omega)^2}{2}u''(\xi_2) \right) (\omega - \omega_{i+1}) \right| \\ &\leq \frac{n}{2} \left| \frac{\partial u}{\partial \omega}(t_1, \omega) \right| \underbrace{\left| (\omega_{i+1} - \omega)(\omega - \omega_i) - (\omega_i - \omega)(\omega - \omega_{i+1}) \right|}_{=0} + \frac{n}{2} \frac{M_i}{2} \left(\frac{2}{n} \right)^3 \\ &\leq \frac{M_i}{2} \left(\frac{2}{n} \right)^2 \end{aligned}$$

□

Proposition 4.5. Soit $\epsilon = \max_{i \in \{1, \dots, n\}} \left| \hat{I}_{N, u_0}^{adj}(t, \delta_{\omega_i}) - u(t_1, \omega_i) \right|$. $\forall \omega \in [-1, 1]$,

$$\left| \hat{u}_{1, N}^{ipo}(\omega) - u_1^{ipo}(\omega) \right| \leq \epsilon$$

4.5.3 Approximation par un polynôme de degré d

On se donne n points équidistants $-1 = \omega_1 < \dots < \omega_n = 1$ et on prend pour $u_1^{pol}(\omega)$ le polynôme en ω qui minimise l'erreur quadratique en ces points :

$$u_1^{pol}(\omega) = \arg \min_{P \in \mathbb{R}_d[X]} \sum_{i=0}^n |P(\omega_i) - u(t_1, \omega_i)|^2$$

Ce qui correspond à l'estimateur :

$$\hat{u}_1^{pol}(\omega) = \arg \min_{P \in \mathbb{R}_d[X]} \sum_{i=0}^n \left| P(\omega_i) - \hat{I}_{N, u_0}^{adj}(t_1, \delta_{\omega_i}) \right|^2$$

Proposition 4.6. Soit $\epsilon = \max_{i \in \{1, \dots, n\}} |\hat{I}_{N, u_0}^{adj}(t, \delta_{\omega_i}) - u(t_1, \omega_i)|$. Soit C le conditionnement pour la norme 2 de la matrice de Vandermonde de rang d associée aux points $\omega_1, \dots, \omega_n$. $\forall \omega \in [-1, 1]$,

$$\left| \hat{u}_{1, N}^{pol}(\omega) - u_1^{pol}(\omega) \right| \leq \epsilon C \frac{\|\beta\|_2}{\|Y\|_2} (n+1)(d+1)$$

(cf. preuve pour définitions de $\|\beta\|_2$ et $\|Y\|_2$)

Démonstration. On note $\forall i \in \{1, \dots, n\}, y_i = u(t_1, \omega_i)$ en sorte que le modèle de régression s'écrit :

$$\forall i \in \{1, \dots, n\}, y_i = \sum_{k=0}^d \beta_k \omega_i^k + \epsilon_i$$

Avec $(\epsilon_1, \dots, \epsilon_n)^T$ le vecteur dont on cherche à minimiser la norme 2. Matriciellement, cela donne :

$$\underbrace{\begin{pmatrix} y_0 \\ \vdots \\ y_n \end{pmatrix}}_Y = \underbrace{\begin{pmatrix} 1 & \omega_1 & \omega_1^2 & \dots & \omega_1^d \\ 1 & \omega_2 & \omega_2^2 & \dots & \omega_2^d \\ 1 & \omega_3 & \omega_3^2 & \dots & \omega_3^d \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega_n & \omega_n^2 & \dots & \omega_n^d \end{pmatrix}}_{V_d(\omega_0, \dots, \omega_n)} \underbrace{\begin{pmatrix} \beta_0 \\ \vdots \\ \beta_d \end{pmatrix}}_{\beta} + \begin{pmatrix} \epsilon_0 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

où $V_d(\omega_0, \dots, \omega_n)$ est la matrice de Vandermonde de rang d aux points ω_i .

On se ramène à un problème de moindres carrés pour minimiser $\|Y - V_d \beta\|_2$, dont la solution est $\beta = (V_d^T V_d)^{-1} V_d^T Y$ (voir [13] page 40).

Soit $C = \text{cond}(V_d)$. On note de plus $\forall i \in \{0, \dots, n\}, \hat{y}_i = \hat{I}_{N, u_0}^{adj}(t, \delta_{\omega_i})$. On sait alors que pour $\Delta = \beta - \hat{\beta}$, $\|\Delta\|_2 \leq \epsilon C \frac{\|\beta\|_2}{\|Y\|_2}$. On a donc :

$$\begin{aligned}
\left| \hat{u}_{1,N}^{pol}(\omega) - u_1^{pol}(\omega) \right| &= \left| \sum_{k=0}^d (\hat{\beta}_k - \beta_k) X^k \right| \\
&\leq \sum_{k=0}^d \left| \hat{\beta}_k - \beta_k \right| |X|^k \\
&\leq \sum_{k=0}^d \|\Delta\|_\infty |X|^k \\
&\leq \sum_{k=0}^d (n+1) \|\Delta\|_2 |X|^k \\
&\leq \frac{\|\beta\|_2}{\|Y\|_2} C\epsilon(n+1) \sum_{k=0}^d |X|^k \\
&\leq \frac{\|\beta\|_2}{\|Y\|_2} C\epsilon(n+1)(d+1)
\end{aligned}$$

□

Remarque 4.4. *Le problème de cette méthode est que même si on a un bon contrôle de l'écart entre $u_1^{pol}(\omega)$ et $u(t_1, \omega)$, il est difficile de maîtriser l'erreur entre $\hat{u}_1^{pol}(\omega)$ et $u_1^{pol}(\omega)$. Cette erreur dépend essentiellement du conditionnement de la matrice de Vandermonde, qui est rapidement grand. Dans tous les cas, on perd un facteur $C(n+1)(d+1)$ du point de vue théorique comparé aux autres méthodes. Même si on ignore le facteur $n+1$ qui vient de la mauvaise majoration de la norme 1 par la norme 2, il reste l'erreur due au conditionnement et au degré. Si on veut changer le critère de minimisation pour prendre la norme 1 et pas la norme 2, on n'a plus affaire à un problèmes de moindres carrés dont la solution est analytique et donc on perd du temps de calcul à trouver le polynôme adéquat.*

L'évaluation de la fonction coûte ici $O(n)$ par méthode de Horner.

4.5.4 Approximation par des polynômes de Bernstein

On se donne n points équidistants $-1 = \omega_1 < \dots < \omega_n = 1$ et on prend pour $u_1^{bern}(\omega)$ le polynôme en ω donné par :

$$u_1^{bern}(\omega) = \sum_{k=0}^n u\left(t_1, \frac{2k}{n} - 1\right) B_k^n\left(\frac{\omega+1}{2}\right)$$

où les B_k^n sont les polynômes de Bernstein (voir [14]) :

$$B_k^n(X) = \binom{n}{k} x^k (1-x)^{n-k}$$

On résout le problème de la partie précédente car pour ces polynômes, d'une part, si on note h le module de continuité de $u(t_1, \cdot)$, on a :

$$\|u_1^{bern}(\omega) - u(t_1, \omega)\|_\infty \leq \frac{9}{4} h \left(\frac{1}{\sqrt{n}}\right)$$

Et d'autre part, si on note $\epsilon = \max_{i \in \{1, \dots, n\}} |\hat{I}_{N, u_0}^{adj}(t, \delta_{\omega_i}) - u(t_1, \omega_i)|$,

$$\begin{aligned} |\hat{u}_1^{bern}(\omega) - u_1^{bern}(\omega)| &\leq \sum_{k=0}^n \epsilon \left| B_k^n\left(\frac{\omega+1}{2}\right) \right| \\ &\leq \epsilon \sum_{k=0}^n \binom{n}{k} \left| \left(\frac{\omega+1}{2}\right)^k \left(1 - \left(\frac{\omega+1}{2}\right)\right)^{n-k} \right| \\ &= \epsilon \end{aligned}$$

Néanmoins, même si la convergence théorique est assurée, elle est très lente. Par exemple, si $u(t_1, \cdot)$ est Lipchitzienne de constante de Lipchitz k , son module de continuité est $y \mapsto k|y|$ et donc la convergence est en $\frac{9k}{4\sqrt{n}}$, alors que le coût d'évaluation de ces polynômes est en $O(n)$ (par méthode de Horner, plus rapide mais instable pour n grand - supérieur à environ 40) voire $O(n^2)$ (évaluation classique, lente mais stable), ce qui est donc insuffisant.

Dans la pratique, l'approximation par un polynôme de degré d qui minimise l'erreur quadratique en un certain nombre de points semble plus efficace. Cette méthode est sur le banc d'essai essentiellement car elle offre une convergence théorique.

4.5.5 Fonctions u_0 utilisées pour des tests

Pour nos tests, on va utiliser différentes répartitions d'énergie initiale, qui sont plus ou moins continues.

| Expression | $\sup_{\omega \in [-1,1]} \left \frac{\partial u}{\partial \omega}(t, \omega) \right $ | $\sup_{\omega \in [-1,1]} \left \frac{\partial^2 u}{\partial \omega^2}(t, \omega) \right $ |
|---|---|---|
| $\omega \mapsto \frac{\mathbb{1}_{[0,1]}(\omega)}{2}$ | $+\infty$ | $+\infty$ |
| $\omega \mapsto -3 \frac{(\omega+1)(\omega-1)}{4}$ | $\frac{3}{2}$ | $\frac{3}{2}$ |
| $\omega \mapsto \frac{17}{2} \omega^{16}$ | 136 | 2040 |
| $\omega \mapsto \frac{e^{-40\omega^2}}{\int e^{-40x^2} dx}$ | ≈ 285 | ≈ 22837 |

4.5.6 Configuration des tests

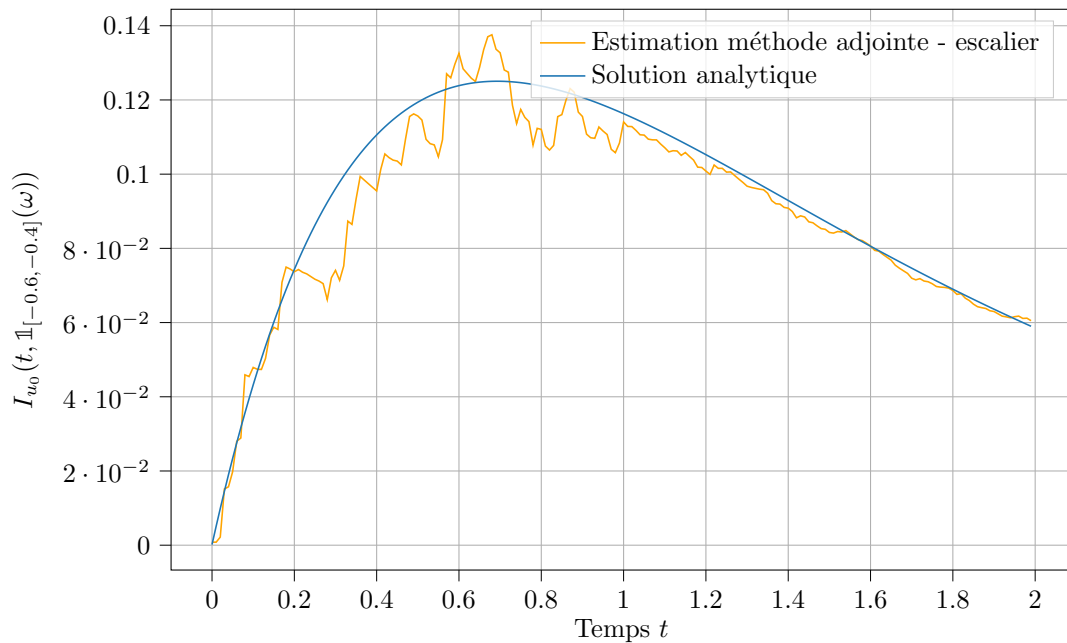
Toutes nos méthodes fonctionnent de manière optimale pour $n \approx 10$ et on retient donc cette valeur pour les tests. Si on sait que l'on traite avec une fonction u_0 à fortes variations, on peut choisir d'augmenter n .

On va chercher à calculer $I_{u_0}(t, \mathbb{1}_{[-0.6, -0.4]}(\omega))$ pour $t \in (0, 2)$ (que l'on évaluera en 60 points), avec un instant intermédiaire $t_1 = 1$ et en faisant varier le nombre de particules utilisées. L'erreur calculée est l'erreur en norme 1 sur ces points :

$$erreur = \sum_{k=1}^{60} \left| \hat{I}_{N, u_0}^{adj} \left(\frac{2k}{60}, \phi \right) - I_{u_0} \left(\frac{2k}{60}, \phi \right) \right|$$

Les courbes d'erreur sont tracées à partir d'un échantillon de 1000 trajectoires. Ci-dessous un exemple de trajectoire tracée avec la méthode d'approximation par valeurs moyennes :

Exemple de trajectoire en 200 points avec instant intermédiaire en $t = 1$



Pour l'approximation par valeurs moyenne, si on se donne N particules au total, on les répartit équitablement sur les n intervalles. Il y a donc N/n particules pour le calcul de $I_{u_0}(t, \mathbb{1}_{[-0.6, -0.4]}(\omega))$.

Pour les autres méthodes, on ne peut pas utiliser les calculs intermédiaires pour le calcul de $I_{u_0}(t, \mathbb{1}_{[-0.6, -0.4]}(\omega))$. Dans le but de conserver un nombre identique de particules pour l'estimation de $u(t_1, \cdot)$, chaque estimateur se verra également alloué N/n particules.

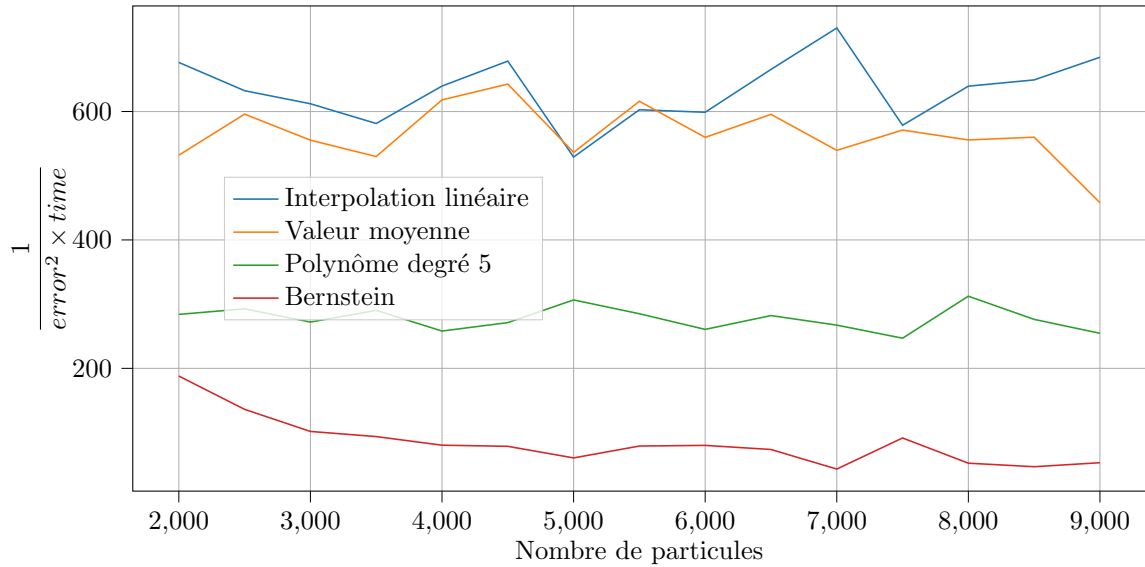
Lors de chaque test, on applique une méthode adjointe une première fois pour calculer une trajectoire entre 0 et t_1 , puis une seconde fois pour la trajectoire entre t_1 et $t_2 = 2$ à partir de la densité approchée u_1 , et on calcule également u_2 (en effet, en vue d'une discrétisation avec plusieurs intervalles de temps, il faut prendre en compte les différences de temps de calcul pour cette phase aussi).

Pour chaque méthode d'approximation, on testera la convergence avec le nombre de particules. On tracera donc des figures de mérite $\frac{1}{\text{erreur}^2 \times \text{tempsDeCalcul}}$ (plus la valeur est grande, meilleure est la performance de la méthode). Si on considère que le temps d'exécution de l'algorithme est proportionnel au nombre de particules utilisées, le théorème central limite indique que ces courbes doivent approximativement être constantes.

4.5.7 Résultats

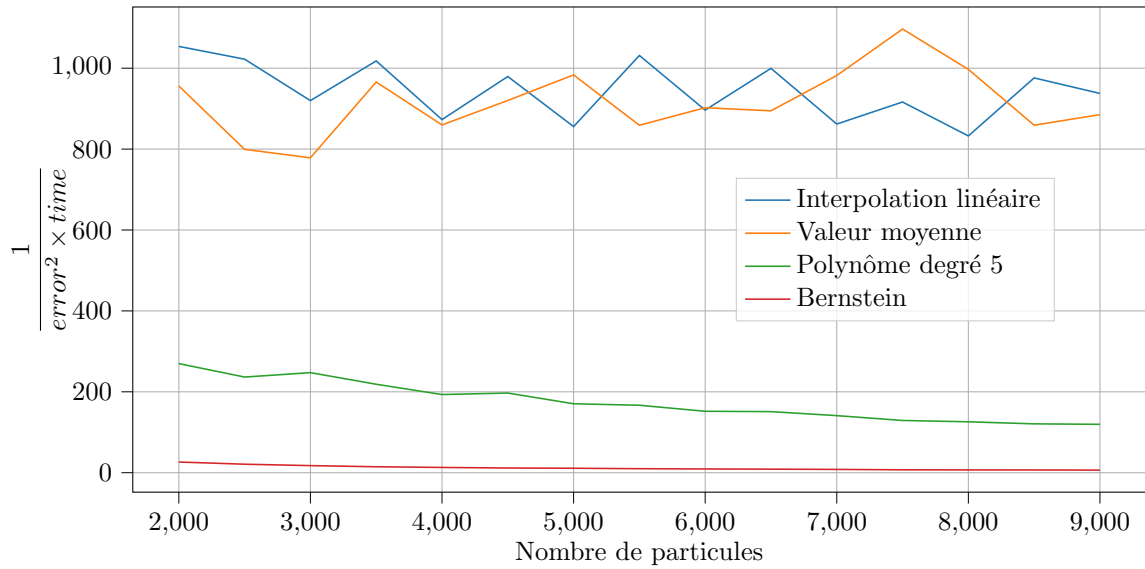
Approximation pour $u_0(\omega) = \frac{\mathbb{1}_{[0,1]}(\omega)}{2}$

Figure de mérite $u_0 : \omega \mapsto \frac{\mathbb{1}_{[0,1]}(\omega)}{2}$



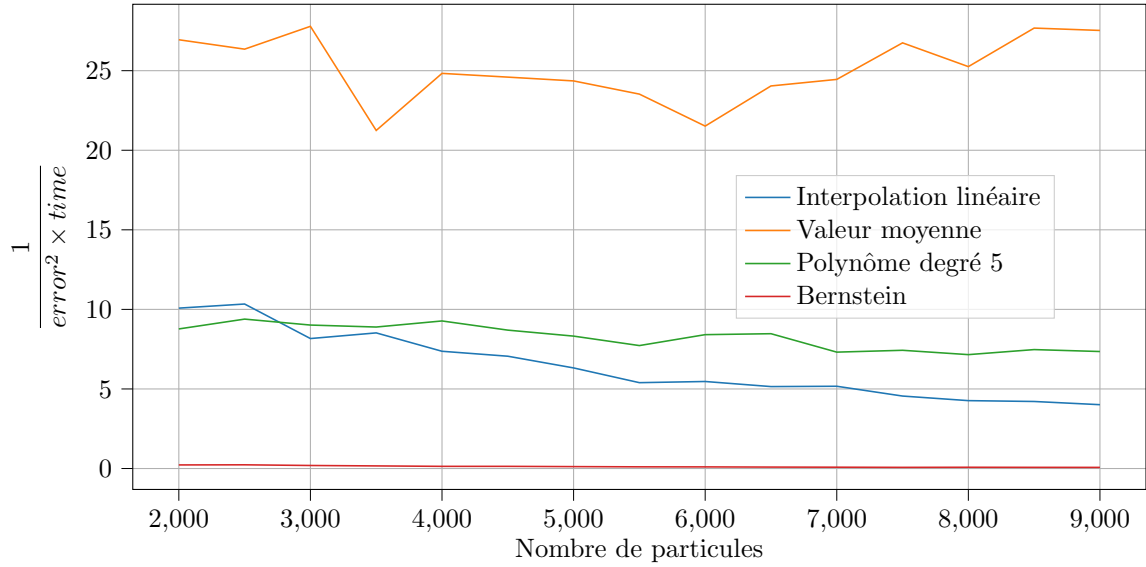
Approximation pour $u_0(\omega) = \frac{-3}{4}(\omega + 1)(\omega - 1)$

Figure de mérite $u_0 : \omega \mapsto \frac{-3}{4}(\omega + 1)(\omega - 1)$



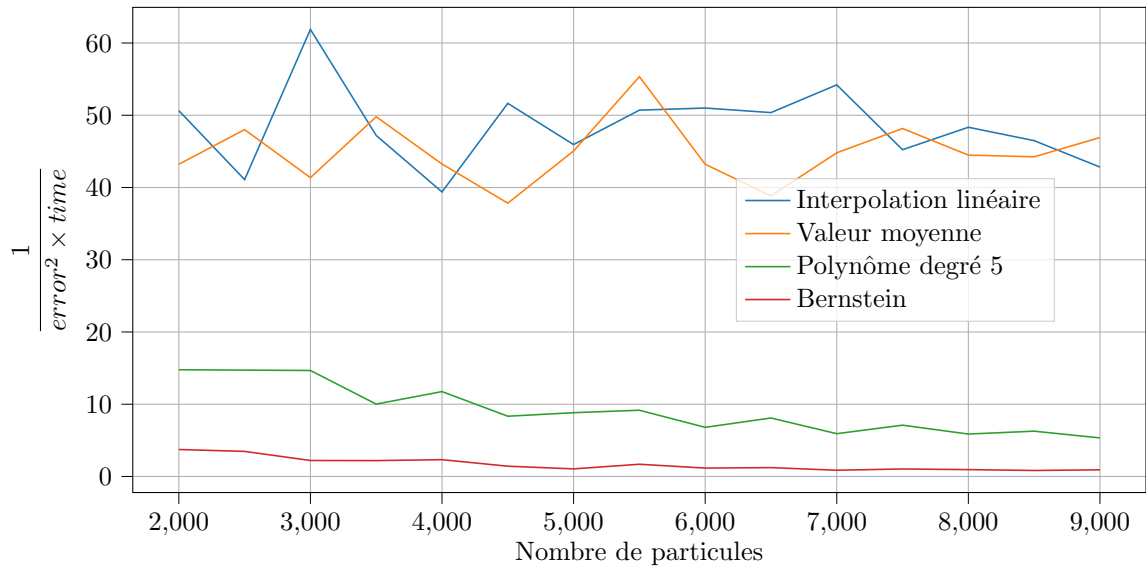
Approximation pour $u_0(\omega) = \frac{17}{2}\omega^{16}$

Figure de mérite $u_0 : \omega \mapsto \frac{17}{2}\omega^{16}$



Approximation pour $u_0(\omega) = \frac{e^{-40\omega^2}}{\int e^{-40x^2} dx}$

Figure de mérite $u_0 : \omega \mapsto \frac{e^{-40\omega^2}}{\int e^{-40x^2} dx}$



4.5.8 Conclusion sur la meilleure méthode d'approximation

Dans toutes les configurations testés, les méthodes les plus efficaces sont l'interpolation linéaire et l'approximation par valeurs moyennes, qui sont à la fois rapides d'exécution et dont la convergence est assurée et rapide. De plus, on appliquera souvent ces méthodes avec des fonctions u_0 constantes par morceaux, et donc, en adaptant la discrétisation aux intervalles sur lesquels u_0 est constante, la méthode par valeurs moyennes est extrêmement efficace (c'est ce que l'on observe dans le test avec f_0). Elle permet en plus de réutiliser les intégrales calculées dans l'approximation pour le calcul de l'intégrale sur une maille dans le cadre de l'IMC. Enfin, bien que toutes les méthodes citées soient généralisables en dimension supérieure pour $x \in \mathbb{R}^3$ (voir [15] pour l'interpolation en dimension quelconque, voir [16] pour les généralisations de polynômes de Bernstein et une simple régression polynomiale multivariée pour la méthode d'un polynôme de degré fixé), la méthode dont la généralisation est la plus simple reste la méthode par valeurs moyennes. On utilisera donc par la suite cette méthode principalement, et on conservera la méthode par interpolation linéaire pour certains usages, notamment car elle est simple à mettre en place.

La propagation de l'erreur due à cette approximation est difficile à estimer car une erreur dans un estimateur modifie les coefficients de l'EDP à résoudre à l'étape suivante, or il existe peu de résultats sur les répercussions d'une incertitude sur les coefficients d'une EDP. Au mieux, on pourra étudier la question dans un cas à coefficients constants. Cette étude sera réalisée sur des cas-tests en partie 5.3.

4.6 Comparaison MCA - MCD sans terme source ni dépendance spatiale

Pour comparer les deux méthodes, on teste avec les paramètres suivants :

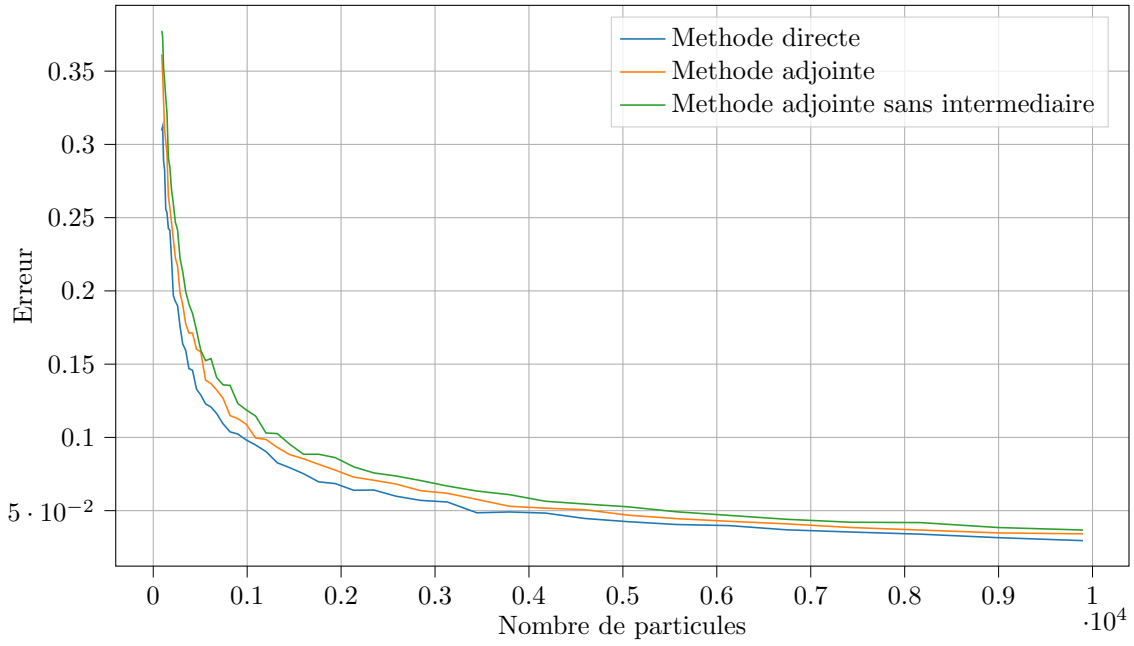
- Énergie initiale $u_0 : \omega \mapsto \mathbb{1}_{[0,1]}(\omega)/2$
- On mesure l'erreur sur la valeur moyenne de u sur l'intervalle $(-0.6, -0.4)$ en différents instants :

$$I_{u_0, N}(t, 5\mathbb{1}_{[-0.6, -0.4]}) = 5 \int_{-0.6}^{-0.4} u(t, \omega) d\omega$$

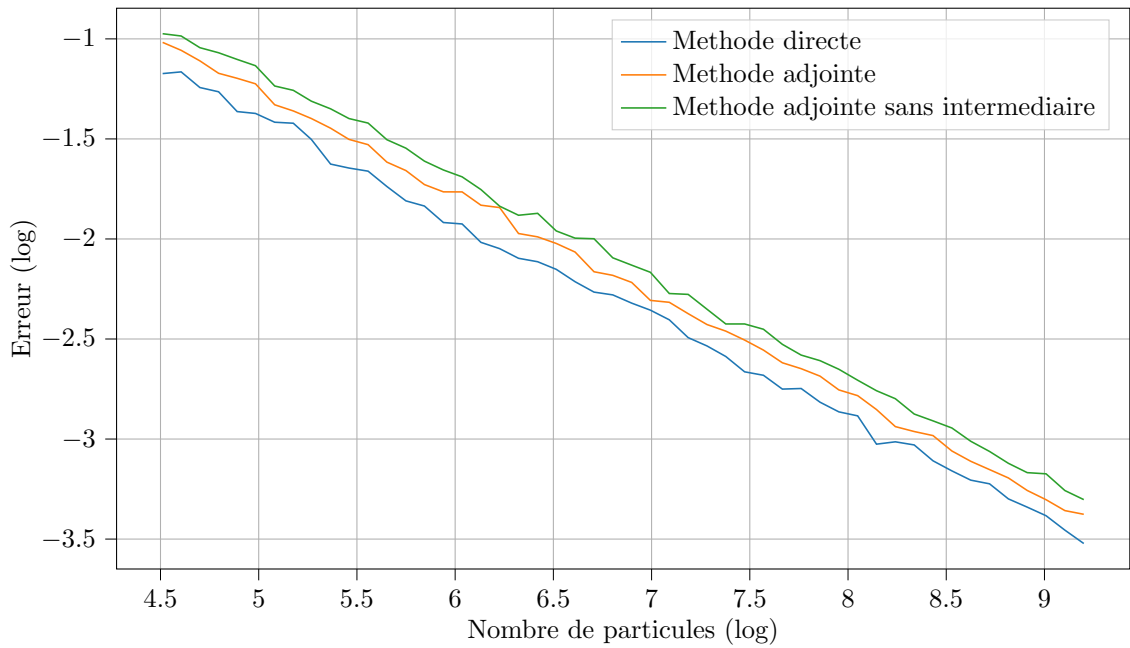
On calcule l'erreur en 60 instants $t \in (0, 4)$, avec un instant intermédiaire en $t = 2$.

- Nombre total de particules testées : $N \in \{100, 300, \dots, 9900\}$
- Pour la méthode adjointe, on utilise comme précédemment la méthode d'approximation par valeurs moyennes avec 10 sous-intervalles de $[-1, 1]$ pour u_1 . On réutilise les particules utilisées pour le calcul de u_1 pour calculer $I_{u_0, N}(t, 5\mathbb{1}_{[-0.6, -0.4]})$. En particulier, cela signifie que si on a N particules au total, on a $N/10$ particules pour le calcul de chaque valeur moyenne et également $N/10$ particules pour le calcul de $I_{u_0, N}(t, 5\mathbb{1}_{[-0.6, -0.4]})$.
- Chaque point de cette courbe représente l'erreur moyenne sur 100 trajectoires-test.

Comparaison MCA/MCD - 0D Sans source



Comparaison MCA/MCD - 0D Sans source (log-log)



Conclusions

La performance en fonction du nombre total de particules est équivalente, quoique très légèrement meilleure pour la méthode directe.

On observe bien une décroissance de l'erreur en $\frac{1}{\sqrt{N}}$ (pente -1/2 sur le graphique log-log).

Il faut cependant garder en tête qu'à nombre de particules égal, la méthode adjointe est sensiblement plus longue à exécuter (d'environ 5%) car elle nécessite plus de calculs intermédiaires (par exemple pour la définition et les appels à $u(t_1, \cdot)$). Ce désavantage sera compensé par la suite par une exécution plus rapide lorsque l'on ajoute un terme source à l'équation.

Enfin, on a ajouté une courbe pour une méthode adjointe sans instant intermédiaire, qui semble légèrement moins performante. En effet, dans ce cas, ajouter un instant intermédiaire a globalement pour effet de diminuer l'erreur. On retient donc qu'ici, il est plus efficace de régénérer des particules qu'on évaluera avec une densité énergétique approchée plutôt que de garder d'anciennes particules que l'on évalue avec la fonction u_0 exacte (du moins lorsque u_1 est une bonne approximation de $u(t_1, \cdot)$, comme c'est le cas ici).

4.7 Estimateur adjoint sans approximation intermédiaire

Les estimateurs utilisés durant tout le stage découlent d'une méthodologie générale décrite dans [2], qui elle-même repose sur une formule de Feynman-Kac. Néanmoins, celle-ci est généralisable par rapport à la version énoncée dans [2]. On peut notamment ajouter une dépendance temporelle aux σ , comme énoncé et démontré dans (3.1). Cela permet de s'affranchir théoriquement du besoin de discrétisation.

Pour cela, il faut générer à $t = 0$ N particules. On laisse ensuite évoluer indépendamment les particules et on les évalue au temps adéquat. Néanmoins, si les σ_a changent au cours du temps, le calcul des termes en $e^{-\int_0^t \sigma_a(t-s, X(s), \Omega(s)) ds}$ ne peut pas se faire tout au long du parcours comme pour les poids. Il faut nécessairement garder l'intégralité de l'historique de la particule en mémoire pour évaluer ses pertes par absorption, ce qui n'est pas faisable sur de nombreuses itérations IMC. Cela représente un grand coût en mémoire (il faut garder beaucoup de particules pendant très longtemps) ainsi qu'un grand coût computationnel.

C'est pourquoi, dans la plupart des cas, il reste plus efficace de discrétiser (d'autant qu'avec un maillage assez fin, on peut se retrouver avec une approximation de $u(t_i, \cdot, \cdot)$ de bonne qualité). De plus, l'estimateur non discrétisé avec terme source s'écrit :

$$I_{N, u_0}^{adj}(t, \phi) = \frac{1}{N} \sum_{i=1}^N \left[u_0(X_i(t), \Omega_i(t)) e^{-\int_0^t \sigma_a(t-\zeta) d\zeta} + \int_0^t f(t-s, X_i(s), \Omega_i(s)) e^{-\int_0^s \sigma_a(t-\zeta) d\zeta} ds \right]$$

Or le calcul de l'intégrale $\int_0^t f(t-s, X_i(s), \Omega_i(s)) e^{-\int_0^s \sigma_a(t-\zeta) d\zeta}$ demande une discrétisation fine, et est donc très cher sur des temps longs.

Cet estimateur "en un coup" sort du cadre initial de ce stage, et ne fera pas l'objet d'une étude aussi approfondie que les estimateurs précédents.

5 Ajout d'un terme source puis d'un maillage

On va commencer par ajouter un terme source à notre équation précédente, et montrer qu'on peut encore trouver une solution analytique à l'équation. Cela permettra de confirmer expérimentalement que les résultats des parties précédentes sont toujours valables avec un terme source.

Une fois notre algorithme validé avec une dépendance spatiale, on pourra établir une comparaison finale entre la méthode adjointe et la méthode directe. L'objectif n'est pas forcément de montrer qu'une méthode est meilleure que l'autre (car cela pourrait être problème-dépendant), mais de vérifier que la méthode adjointe donne des résultats au moins équivalents, tout en étudiant l'impact de l'ajout d'un terme source.

5.1 Ajout d'un terme source

5.1.1 Résolution analytique avec terme source

On considère à présent, sur le même espace, l'équation :

$$\begin{aligned} \frac{\partial u(t, \omega)}{\partial t} + \sigma_T u(t, \omega) &= f(t, \omega) + \sigma_s \int_{-1}^1 u(t, \omega') \frac{d\omega'}{2} \\ u_0 &= u(0, \cdot) \end{aligned} \quad (6)$$

Où f est un terme source quelconque.

Proposition 5.1. *La solution à (6) est donnée par :*

$$\begin{aligned} u(t, \omega) &= \int_0^t e^{-\sigma_T(t-s)} \int_{-1}^1 (f(s, \omega) - f(s, \omega')) \frac{d\omega'}{2} ds + \frac{W(0)e^{-\sigma_a t}}{2} - \frac{W(0)e^{-\sigma_T t}}{2} \\ &\quad + u_0(\omega)e^{-\sigma_T t} + \int_0^t e^{-\sigma_a(t-s)} \int_{-1}^1 f(s, \omega') \frac{d\omega'}{2} ds \end{aligned}$$

Démonstration. On intègre (6) sur $[-1, 1]$:

$$\begin{aligned} \frac{\partial}{\partial t} \int_{-1}^1 u(t, \omega) d\omega + \sigma_T \int_{-1}^1 u(t, \omega) d\omega &= \int_{-1}^1 f(t, \omega) d\omega + \sigma_s \int_{-1}^1 \left(\int_{-1}^1 u(t, \omega') \frac{d\omega'}{2} \right) d\omega \\ &= \int_{-1}^1 f(t, \omega) d\omega + \sigma_s \int_{-1}^1 u(t, \omega) d\omega \end{aligned}$$

On pose $W : t \mapsto \int_{-1}^1 u(t, \omega) d\omega$, en sorte que :

$$\begin{aligned} W'(t) + \underbrace{(\sigma_T - \sigma_s)}_{\sigma_a} W(t) &= \int_{-1}^1 f(t, \omega) d\omega \\ \implies \forall t \geq 0, W(t) &= W(0)e^{-\sigma_a t} + e^{-\sigma_a t} \int_0^t \int_{-1}^1 f(s, \omega) d\omega ds \end{aligned}$$

Avec

$$W(0) = \int_{-1}^1 u_0(\omega) d\omega$$

On injecte dans 6 :

$$\frac{\partial u(t, \omega)}{\partial t} + \sigma_T u(t, \omega) = f(t, \omega) + \frac{\sigma_s e^{-\sigma_a t}}{2} \left(W(0) + \int_0^t e^{\sigma_a s} \int_{-1}^1 f(t, \omega') d\omega' \right)$$

On multiplie par $e^{\sigma_T t}$:

$$e^{\sigma_T t} \frac{\partial u(t, \omega)}{\partial t} + \sigma_T e^{\sigma_T t} u(t, \omega) = e^{\sigma_T t} f(t, \omega) + e^{\sigma_T t} \frac{\sigma_s e^{-\sigma_a t}}{2} \left(W(0) + \int_0^t e^{\sigma_a s} \int_{-1}^1 f(s, \omega') d\omega' ds \right)$$

$$\iff e^{\sigma_T t} \frac{\partial u(t, \omega)}{\partial t} + \frac{\partial e^{\sigma_T t}}{\partial t} u(t, \omega) = e^{\sigma_T t} f(t, \omega) + \frac{\sigma_s e^{\sigma_s t}}{2} \left(W(0) + \int_0^t e^{\sigma_a s} \int_{-1}^1 f(s, \omega') d\omega' ds \right)$$

$$\iff \frac{\partial u(t, \omega) e^{\sigma_T t}}{\partial t} = e^{\sigma_T t} f(t, \omega) + \frac{1}{2} \frac{\partial e^{\sigma_s t}}{\partial t} \left(W(0) + \int_0^t e^{\sigma_a s} \int_{-1}^1 f(s, \omega') d\omega' ds \right)$$

$$\iff u(t, \omega) e^{\sigma_T t} = \int_0^t e^{\sigma_T s} f(s, \omega) ds + \frac{W(0)}{2} [e^{\sigma_s t} - 1] + \underbrace{\frac{1}{2} \int_0^t \frac{\partial e^{\sigma_s \zeta}}{\partial \zeta} \int_0^\zeta e^{\sigma_a s} \int_{-1}^1 f(s, \omega') d\omega' ds d\zeta}_{B} + u_0(\omega)$$

Par intégration par parties,

$$\begin{aligned} B &= \frac{1}{2} \left[e^{\sigma_s \zeta} \int_0^\zeta e^{\sigma_a s} \int_{-1}^1 f(s, \omega') d\omega' ds \right]_{\zeta=0}^t - \frac{1}{2} \int_0^t e^{(\sigma_s + \sigma_a) s} \int_{-1}^1 f(s, \omega') d\omega' ds \\ &= \frac{1}{2} e^{\sigma_s t} \int_0^t e^{\sigma_a s} \int_{-1}^1 f(s, \omega') d\omega' ds - \int_0^t e^{\sigma_T s} \int_{-1}^1 f(s, \omega') \frac{d\omega'}{2} ds \end{aligned}$$

$$\implies u(t, \omega) e^{\sigma_T t} = \int_0^t e^{\sigma_T s} \int_{-1}^1 (f(s, \omega) - f(s, \omega')) \frac{d\omega'}{2} ds + \frac{W(0)}{2} [e^{\sigma_s t} - 1]$$

$$+ u_0(\omega) + e^{\sigma_s t} \int_0^t e^{\sigma_a s} \int_{-1}^1 f(s, \omega') \frac{d\omega'}{2} ds$$

$$\implies u(t, \omega) = \int_0^t e^{-\sigma_T(t-s)} \int_{-1}^1 (f(s, \omega) - f(s, \omega')) \frac{d\omega'}{2} ds + \frac{W(0) e^{-\sigma_a t}}{2} - \frac{W(0) e^{-\sigma_T t}}{2}$$

$$+ u_0(\omega) e^{-\sigma_T t} + \int_0^t e^{\sigma_a(s-t)} \int_{-1}^1 f(s, \omega') \frac{d\omega'}{2} ds$$

□

Remarque 5.1. Soit ϕ une fonction positive sur $[-1, 1]$. Notons $\beta = \int_{-1}^1 \phi(\omega) d\omega$ On peut alors calculer :

$$\begin{aligned}
I_\phi(t) &= \beta \frac{e^{-\sigma_a t}}{2} W(0) + \int_{-1}^1 \int_0^t e^{-\sigma_T(t-s)} f(s, \omega) ds \phi(\omega) d\omega \\
&\quad + \beta \int_0^t e^{-\sigma_a(t-s)} \left(1 - e^{-\sigma_s(t-s)}\right) \int_{-1}^1 f(s, \omega) \frac{d\omega}{2} ds \\
&\quad + \left(\int_{-1}^1 u_0(\omega) \phi(\omega) d\omega - \frac{\beta}{2} \int_{-1}^1 u_0(\omega) d\omega \right) e^{-\sigma_T t}
\end{aligned}$$

Remarquons que si f ne dépend pas de ω ,

$$\begin{aligned}
I_\phi(t) &= \beta \frac{e^{-\sigma_a t}}{2} W(0) + \beta \int_0^t e^{-\sigma_T(t-s)} f(s) ds \\
&\quad + \beta \int_0^t e^{-\sigma_a(t-s)} \left(1 - e^{-\sigma_s(t-s)}\right) f(s) ds \\
&\quad + \left(\int_{-1}^1 u_0(\omega) \phi(\omega) d\omega - \frac{\beta}{2} W(0) \right) e^{-\sigma_T t} \\
&= \beta \frac{e^{-\sigma_a t}}{2} W(0) + \beta \int_0^t e^{-\sigma_a(t-s)} f(s) ds + \left(\int_{-1}^1 u_0(\omega) \phi(\omega) d\omega - \frac{\beta}{2} W(0) \right) e^{-\sigma_T t}
\end{aligned}$$

5.1.2 Estimateur adjoint associé

L'estimateur tiré de (3.1) est :

$$\hat{I}_N^{adj}(t, \phi) = \frac{1}{N} \sum_{i=1}^N \left[\beta u_0(\Omega_i(t)) e^{-\sigma_a t} + \beta \int_0^t f(t-s, \Omega_i(s)) e^{-\sigma_a s} ds \right]$$

Avec $\forall i, \forall \zeta > 0, \Omega_i(\zeta) \sim f_{\Omega(\zeta)}^{adj}$

La génération des particules ne change pas, et la méthode n'est pas plus complexe que pour $f = 0$.

En reprenant les calculs précédents, pour montrer que l'estimateur est sans biais, il faut montrer que :

$$\begin{aligned}
\mathbb{E} \left(\beta \int_0^t f(t-s, \Omega_i(s)) e^{-\sigma_a s} ds \right) &= \int_{-1}^1 \int_0^t e^{-\sigma_T(t-s)} f(s, \omega) ds \phi(\omega) d\omega \\
&\quad + \beta \int_0^t e^{-\sigma_a(t-s)} \left(1 - e^{-\sigma_s(t-s)}\right) \int_{-1}^1 f(s, \omega) \frac{d\omega}{2} ds
\end{aligned}$$

Calculons :

$$\begin{aligned}
\mathbb{E} \left(\beta \int_0^t f(t-s, \Omega_i(s)) e^{-\sigma_a s} ds \right) &= \beta \int_0^t \mathbb{E}(f(t-s, \Omega(s))) e^{-\sigma_a s} ds \\
&= \beta \int_0^t \mathbb{E}(f(y, \Omega(t-y))) e^{-\sigma_a(t-y)} dy
\end{aligned}$$

Or

$$\mathbb{E}(f(y, \Omega(t-y))) = \int_{-1}^1 e^{-\sigma_s(t-y)} f(y, \omega) \tilde{\phi}(\omega) d\omega + \int_{-1}^1 (1 - e^{-\sigma_s(t-y)}) f(y, \omega) \frac{d\omega}{2}$$

$$\begin{aligned} \implies \mathbb{E} \left(\beta \int_0^t f(s, \Omega_i(s)) e^{-\sigma_a s} ds \right) &= \int_{-1}^1 \int_0^t e^{-\sigma_T(t-y)} f(y, \omega) dy \phi(\omega) d\omega \\ &+ \beta \int_0^t e^{-\sigma_a(t-y)} \left(1 - e^{-\sigma_s(t-y)} \right) \int_{-1}^1 f(y, \omega) \frac{d\omega}{2} dy \end{aligned}$$

L'estimateur est donc bien sans biais⁷.

5.1.3 Estimateur direct associé

L'estimateur direct tiré de (3.2) est, avec un terme source :

$$\hat{I}_{N, u_0}^{dir}(t, \phi) = \frac{1}{N} \sum_{i=1}^N \alpha \phi(\Omega_i(t)) e^{-\sigma_a t} + \int_0^t \gamma(s) \mathbb{E}_{s, \bar{f}(s, \cdot)} \left[\phi(\Omega(t)) e^{-\sigma_a(t-s)} \right] ds$$

La loi sous-jacente au processus $\Omega(t)$ s'écrit :

$$f_{\Omega(t)}(\omega) = \bar{f}(s, \omega) e^{-\sigma_s(t-s)} + \frac{\mathbb{1}_{[-1,1]}(\omega)}{2} (1 - e^{-\sigma_a(t-s)})$$

Et on a ainsi :

$$\begin{aligned} \int_0^t \gamma(s) \mathbb{E}_{s, \bar{f}(s, \cdot)} \left[\phi(\Omega(t)) e^{-\sigma_a(t-s)} \right] ds &= \int_0^t \int_{-1}^1 \phi(\omega) \gamma(s) \bar{f}(s, \omega) e^{-\sigma_s(t-s)} e^{-\sigma_a(t-s)} d\omega ds \\ &+ \int_0^t \gamma(s) \int_{-1}^1 (1 - e^{-\sigma_s(t-s)}) e^{-\sigma_a(t-s)} \phi(\omega) \frac{d\omega}{2} \\ &= \int_0^t \int_{-1}^1 \phi(\omega) f(s, \omega) e^{-\sigma_s(t-s)} e^{-\sigma_a(t-s)} d\omega ds \\ &+ \beta \int_0^t (1 - e^{-\sigma_s(t-s)}) e^{-\sigma_a(t-s)} \int_{-1}^1 f(s, \omega) \frac{d\omega}{2} \end{aligned}$$

Ce qui assure que l'estimateur est non-biaisé (dans la pratique, si on discrétise l'intégrale de manière déterministe, on introduit un biais : l'erreur de troncature).

Cette méthode demande néanmoins de calculer $\int_0^t \gamma(s) \mathbb{E}_{s, \bar{f}(s, \cdot)} \left[\phi(\Omega(t)) e^{-\sigma_a(t-s)} \right] ds$, ce qui implique d'une part une discrétisation de l'intégrale, et, d'autre part, de calculer l'espérance de l'intégrale par un estimateur de Monte Carlo direct en chaque point de discrétisation. De plus, on compare nos méthode à nombre de particules émisés égaux. Dans le cas de la MCD, ces particules doivent être réparties entre particules pour calculer l'intégrale du terme source et particules pour calculer le reste de l'estimateur.

En revanche, l'avantage est qu'avec la méthode directe, il est aisé de faire changer les σ au cours du temps.

7. Ce qui vient confirmer la coquille dans la formule donnée par [2]

5.1.4 Discrétisation déterministe contre FullMC

Que ce soit en méthode adjointe ou directe, on a une intégrale en temps à estimer. On a alors deux possibilités : une discrétisation déterministe (intégrale approchée par des trapèzes par exemple) ou une discrétisation aléatoire de type Monte Carlo (que l'on notera "FullMC").

En méthode directe, on cherche à calculer

$$\int_0^t \gamma(s) \mathbb{E}_{s, \tilde{f}(s, \cdot)} \left[\phi(\Omega(t)) e^{-\sigma_a(t-s)} \right] ds$$

Si on prend une méthode déterministe, on doit choisir un nombre de points de discrétisation, puis, en chaque point, faire une estimation de $\mathbb{E}_{s, \tilde{f}(s, \cdot)} \left[\phi(\Omega(t)) e^{-\sigma_a(t-s)} \right]$ ⁸.

Si on prend une méthode FullMC, on peut remarquer que pour $s \sim \mathcal{U}(0, t)$,

$$\int_0^t \gamma(s) \mathbb{E}_{s, \tilde{f}(s, \cdot)} \left[\phi(\Omega(t)) e^{-\sigma_a(t-s)} \right] ds = t \mathbb{E} \left[\gamma(s) \mathbb{E}_{s, \tilde{f}(s, \cdot)} \left[\phi(\Omega(t)) e^{-\sigma_a(t-s)} \right] \right]$$

Ce qui revient à prendre une fonction ϕ dépendant du temps. Chaque particule allouée à l'intégrale "naît" donc en un instant aléatoire, avec une direction de propagation ω aléatoire. Cette méthode FullMC ne doit donc pas souffrir d'une erreur de troncature car lorsque le nombre de particule total augmente, on augmente également le nombre de particules Monte Carlo de cette intégrale, et donc on doit bien observer une décroissance de l'erreur en $\frac{1}{\sqrt{N}}$.

Pour la méthode adjointe, on cherche à calculer

$$\frac{1}{N} \sum_{i=1}^N \beta \int_0^t f(t-s, \Omega_i(s)) e^{-\sigma_a s} ds$$

La problématique est alors de nature différente : le nombre de particules est choisi à l'avance et le choix de la méthode de discrétisation de l'intégrale n'impacte que les instants auxquels on évalue chaque particule.

Si on opte pour une méthode déterministe, il suffit de choisir un nombre de points de discrétisation et d'évaluer l'intégrande en chaque point.

La première méthode "FullMC" testée durant ce stage consistait à choisir un nombre de points de discrétisation (par exemple, une vingtaine) que l'on prend selon une loi uniforme sur $(0, t)$, et d'évaluer toutes les particules en ces points. Cette méthode va alors faire subsister une erreur de troncature quel que soit le nombre de particules. En effet, même en augmentant le nombre de particules, on n'augmente pas le nombre de points d'évaluation de l'intégrale, et donc la loi forte des grands nombre ne s'applique pas. Les résultats présentés dans la partie suivante utilisent cette méthode, qui, bien qu'elle soit précise, est lente d'exécution.

On verra en section 5.2.3 une seconde version de la méthode FullMC, la méthode FullMC-1, qui consiste à tirer un point d'évaluation aléatoire uniformément sur $(0, t)$ pour chaque particule, et qui règle le problème précédent.

8. C'est la seule méthode donnée dans [2]. En effet, pour des applications en neutronique (dont [2] est le sujet), on ne peut calculer le terme source qu'en un nombre limité de points déterministes. Dans nos applications en photonique, il convient donc d'étudier si cette méthode est efficace en comparaison à la méthode FullMC proposée ici.

5.1.5 Comparaison MCA - MCD avec terme source, sans dépendance spatiale

On va comparer en même temps les méthodes adjointes et directe, et les avantages et inconvénients d'adopter une discrétisation déterministe ou Monte Carlo (FullMC) de l'intégrale liée au terme source.

Le test est dans la même configuration qu'en partie précédente, avec en plus un terme source constant $f = 0.05$.

Pour comparer les deux méthodes à nombre de particules égal, il faut d'abord optimiser l'utilisation des particules de la méthode directe-déterministe, i.e. savoir en combien d'intervalles diviser l'intégrale en temps dans l'équation (3.2) et combien de particules utiliser en chaque point de discrétisation.

Pour ce faire, on procède, toutes choses égales par ailleurs, à une recherche du meilleur jeu de ces paramètres de manière extensive (non détaillé ici). Par ailleurs, puisque la contribution en énergie de u_0 est comparable à celle de f dans le calcul de u , il convient d'avoir autant de particules pour le calcul du terme associé à u_0 que pour le calcul des intégrandes en différents points. Ainsi, si on discrétise l'intégrale en p points et que l'on a N particules au total, on réserve $N/(p + 1)$ particules pour le calcul de \mathbb{E}_{0,u_0} et $N/(p + 1)$ pour chaque calcul de $\mathbb{E}_{s,f(s,\cdot)}$.

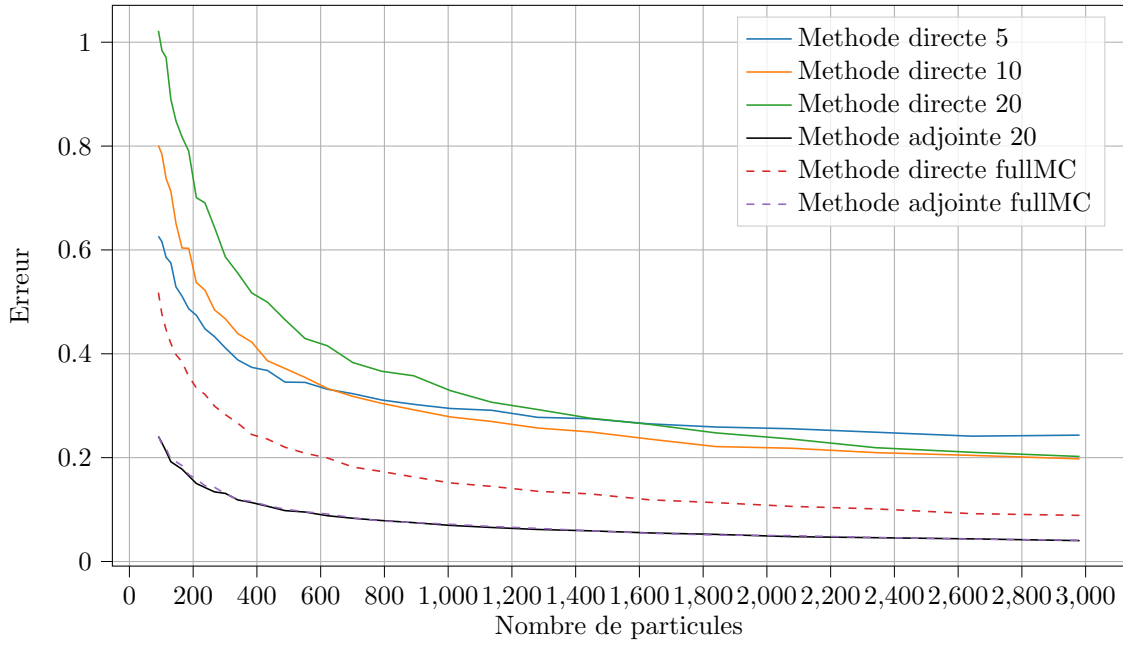
Côté méthode adjointe, on trace deux courbes : la première avec une estimation Monte Carlo de l'intégrale liée au terme source (20 instants aléatoires) et la seconde avec une discrétisation déterministe de cette dernière (20 instants déterministes).

On rappelle que même si en méthode directe, on peut faire tendre le nombre de particules (et donc le nombre de points de discrétisation) dédiées à l'intégrale vers $+\infty$, ce n'est pas possible en méthode adjointe. C'est pourquoi on spécifie en méthode adjointe le nombre de points de discrétisation de l'intégrale.

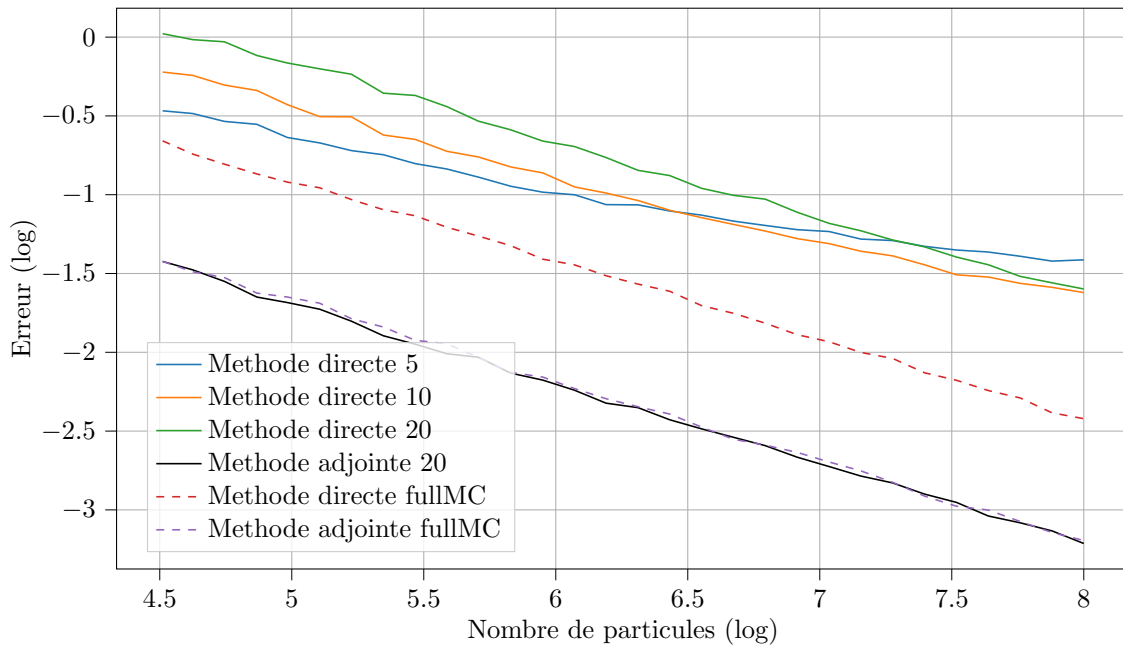
On réalise la comparaison pour des nombres de particules similaires à précédemment.

Les différentes courbes tracées pour la méthode directe indiquent le nombre p de points de discrétisation utilisés pour l'intégrale.

Comparaison MCA/MCD - 0D Avec source



Comparaison MCA/MCD - 0D Avec source (log-log)



Conclusions

Premièrement, on observe bien une décroissance de l'erreur en $\frac{1}{\sqrt{N}}$ pour la méthode directe "FullMC". C'est aussi le cas pour les deux méthodes adjointes, ce qui signifie qu'avec 20 points de

discrétisation (qu'ils soient aléatoires ou déterministes), l'erreur de troncature de l'intégrale n'est pas dominante.

On observe qu'en méthode adjointe et avec autant de points de discrétisation, opter pour une discrétisation déterministe ou FullMC ne change pas la performance au niveau de l'erreur.

Il est à noter que la méthode adjointe est néanmoins notablement plus longue à exécuter, notamment à cause des nombreuses évaluations des particules nécessaires pour le calcul de l'intégrale.

Pour la méthode directe, on observe ce à quoi on s'attend : si on prend une discrétisation déterministe, on reste relativement efficace avec peu de particules, mais on ne parvient pas à éliminer l'erreur de troncature liée à la discrétisation de l'intégrale, alors que si on prend une discrétisation FullMC, l'erreur tend bien vers 0.

Enfin, le fait de devoir allouer un très grand nombre de particules au calcul de la contribution du terme source en méthode directe (courbes du haut) rend celle-ci moins efficace que la méthode adjointe (courbes du bas).

5.2 Ajout d'une dépendance spatiale

5.2.1 Introduction du maillage

Dorénavant, on va introduire la dépendance spatiale en considérant que l'espace est sectionné en différentes mailles. On va travailler en 1D ($x \in [0, 5]$), et le maillage-test utilisé est le suivant :

| x | (0,1) | (1,2) | (2,3) | (3,4) | (4,5) |
|------------|-------|-------|-------|-------|-------|
| σ_a | 1 | 2 | 1 | 3 | 2 |
| σ_s | 1 | 2 | 1 | 3 | 2 |

On validera les méthodes avec un maillage constant (toutes les valeurs à 1), puis on les comparera avec ce maillage-test.

5.2.2 Particules qui sortent du maillage en méthode adjointe

Lorsque l'on introduit un maillage, on ajoute une difficulté supplémentaire en méthode adjointe : le principe général est qu'on génère des particules qui arrivent "en fin de parcours" dans la maille qui nous intéresse, puis on remonte leur trajectoire. Néanmoins, si on a des conditions aux limites de type absorption (ce qui est ici notre cas), cela signifie que des particules risquent de disparaître.

C'est le contrecoup de la méthode adjointe : en méthode directe, on génère les particules où il y a de l'énergie, mais on ne maîtrise pas leur destination et en version adjointe, on génère les particules là où on veut en voir arriver, mais il est possible que les particules arrivant à cet endroit ne puissent venir que de l'extérieur du maillage, où u n'est pas définie. Physiquement, ces particules qui viennent de l'extérieur du maillage doivent donc être d'énergie en lien avec ce qui s'y trouve (par exemple, si le maillage est une cuve isolée de l'extérieur, on peut considérer $u = 0$ en dehors du maillage).

5.2.3 Validation des méthodes - équation de transport avec dépendance spatiale

On va dorénavant s'attaquer à l'équation de transport complète, avec une dépendance en espace de dimension 1 et avec $\omega \in [-1, 1]$:

$$\frac{\partial u}{\partial t}(t, x, \omega) + \omega \cdot \frac{\partial u}{\partial x}(t, x, \omega) + (\sigma_a + \sigma_s)u(t, x, \omega) = \sigma_s \int_{-1}^1 u(t, x, \omega') \frac{d\omega'}{2} + f(t, x, \omega) \quad (7)$$

Dans un premier temps, on valide nos méthodes en prenant $\sigma_s = 0$, cas dans lequel on a une solution analytique (cf. infra). Une fois nos méthodes validées, on pourra les comparer.

Proposition 5.2. *Pour des conditions au limites de type absorption, la solution à l'équation*

$$\frac{\partial u}{\partial t}(t, x, \omega) + \omega \cdot \frac{\partial u}{\partial x}(t, x, \omega) + \sigma_a u(t, x, \omega) = f(t)$$

est donnée par

$$u(t, x, \omega) = u_0(x - \omega t, \omega) e^{-\sigma_a t} + e^{-\sigma_a t} \int_0^t e^{\sigma_a s} f(s) ds$$

Démonstration. Pour simplifier les notations, on note $u = u(t, x)$ en omettant le ω , qui n'intervient pas au cours de la preuve.

On cherche une ligne caractéristique $(s, x(s))$ le long de laquelle cette EDP se réduit en une EDO :

$$\begin{aligned} \frac{du}{ds}(s, x(s)) &= \frac{\partial u}{\partial t}(s, x(s)) + x'(s) \frac{\partial u}{\partial x}(s, x(s)) \\ &= \frac{\partial u}{\partial t}(s, x(s)) + \omega \frac{\partial u}{\partial x}(s, x(s)) \\ &= -\sigma_a u(s, x(s)) + f(s) \end{aligned}$$

Par variation de la constante (que l'on n'aurait pas pu utiliser si $\sigma_s \neq 0$, car la variation de la constante ne permet pas de traiter le terme $\int_{-1}^1 u(t, x, \omega) d\omega$), on obtient :

$$u(s, x(s)) = u(s, x(0) + \omega t) = u(0, x(0)) e^{-\sigma_a s} + e^{-\sigma_a s} \int_0^s e^{\sigma_a \xi} f(\xi) d\xi$$

Si on impose de plus $x(s) = x = x(0) + \omega t$, on obtient :

$$\implies u(s, x) = u_0(x - \omega t) e^{-\sigma_a s} + e^{-\sigma_a s} \int_0^s e^{\sigma_a \xi} f(\xi) d\xi$$

On vérifie bien que ce u est solution de l'équation. □

Ce résultat analytique nous permet de vérifier nos résultats en prenant σ_a indépendant de la maille dans laquelle on se trouve, et d'utiliser par la suite les algorithmes validés avec un maillage.

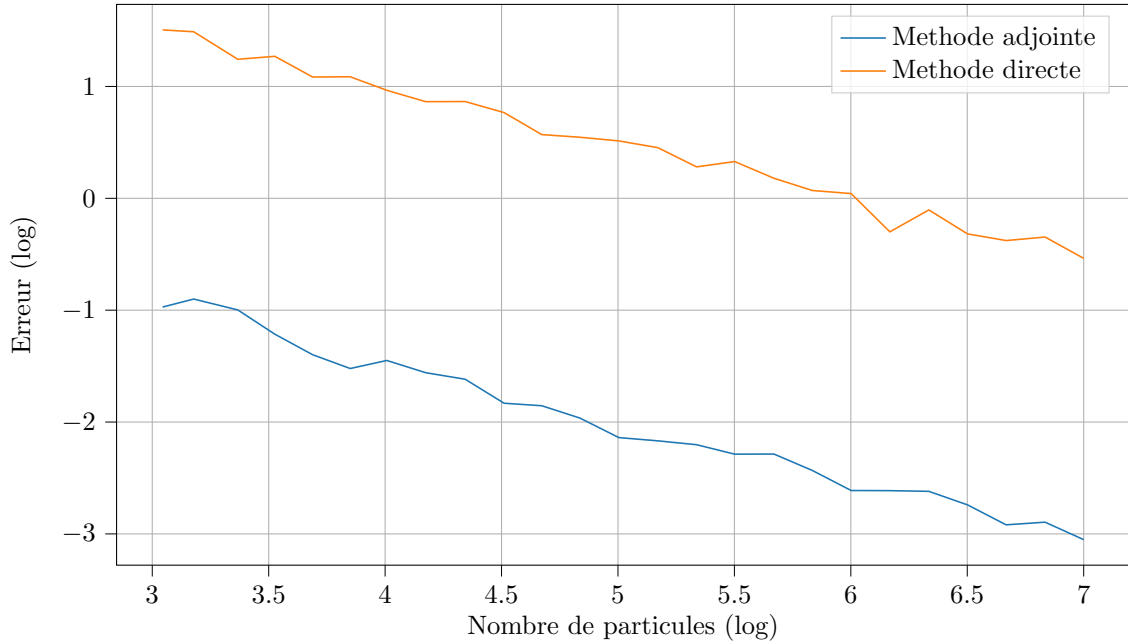
On valide nos méthodes en vérifiant que l'erreur diminue en $\frac{1}{\sqrt{N}}$.

Sans terme source

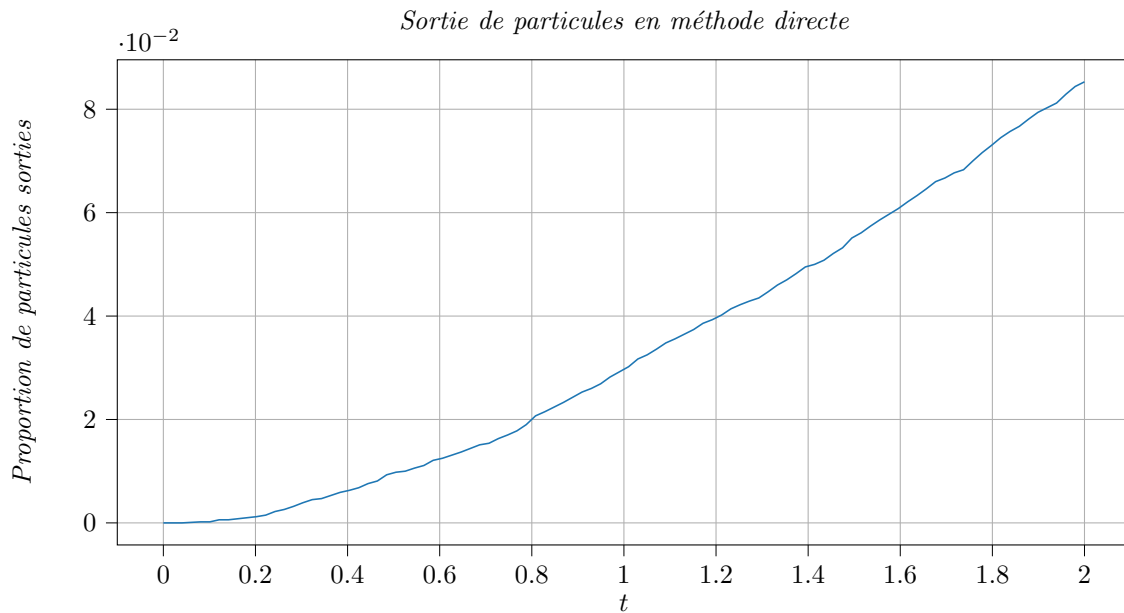
On commence par valider sans terme source, avec les paramètres suivants :

- $u_0 : (x, \omega) \mapsto x(1 - \frac{x}{5})$
- $\phi : (x, \omega) \mapsto \mathbb{1}_{[2,3]}(x)$
- On trace une trajectoire en 15 instants $t \in (0, 2)$ sans instant intermédiaire.
- On calcule l'erreur en norme 1 sur les 15 points.
- Comme précédemment, puisque l'espace $(0, 5)$ est divisé en 5 mailles de tailles égales, on utilise 5 fois moins de particules pour la méthode adjointe.

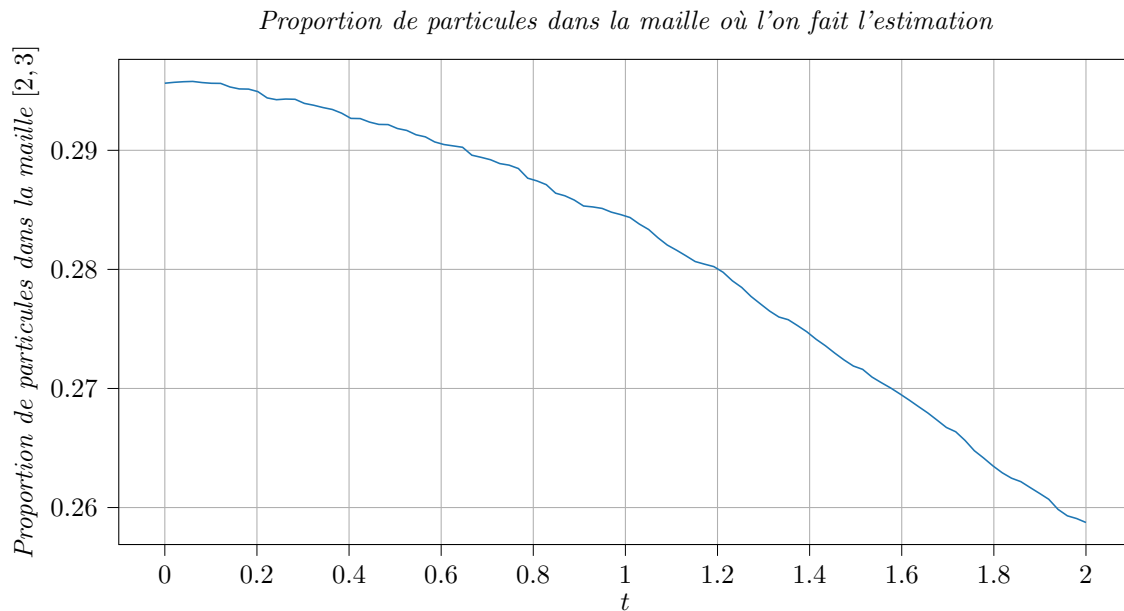
Comparaison MCA/MCD - 1D Sans source (log-log)



Remarque 5.2. On peut également expliquer la faible précision de la méthode directe par le fait que lorsque un maillage entre en jeu, les particules sont susceptibles d'en sortir. Ici, en méthode adjointe, les paramètres ont été choisis tels qu'aucune particule n'ait le temps de sortir du maillage (en effet, elles naissent en $x \in (2, 3)$, ont une vitesse maximale de 1, et le temps de fin de simulation est 2). En méthode directe, des particules sont générées selon la densité \tilde{u}_0 , et les particules qui naissent près du bord ont donc de grandes chances de sortir rapidement. Pour cette simulation, on peut tracer la proportion moyenne de particules en dehors du maillage au cours du temps :



On peut également tracer la proportion de particules qui se trouvent dans la maille [2,3] et qui contribuent à l'estimateur :



Avec terme source

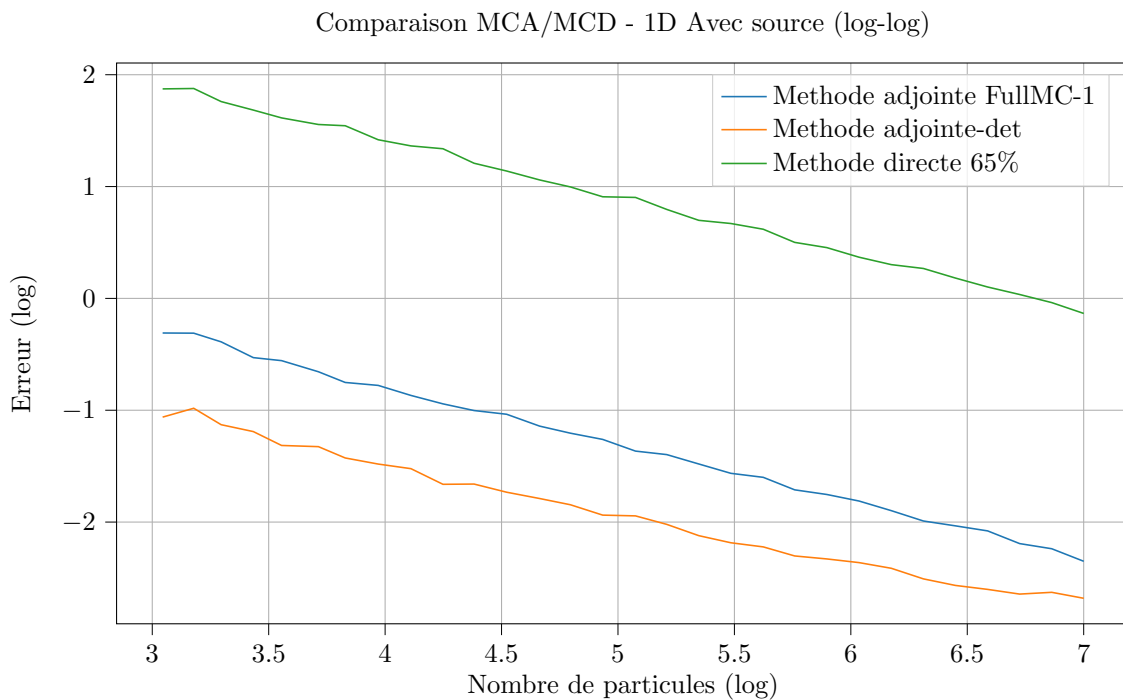
On fait un test identique au précédent avec en plus un terme source constant $f = 0.25$. Pour la méthode directe, on emploie une méthode FullMC avec 65% des particules allouées à la condition initiale et 35% au terme source (répartition optimale pour ce cas-test).

Pour la méthode adjointe, on va tester une nouvelle méthode de discrétisation de l'intégrale. Jusque là, on évaluait chaque particule en un certain nombre de points (le nombre de points

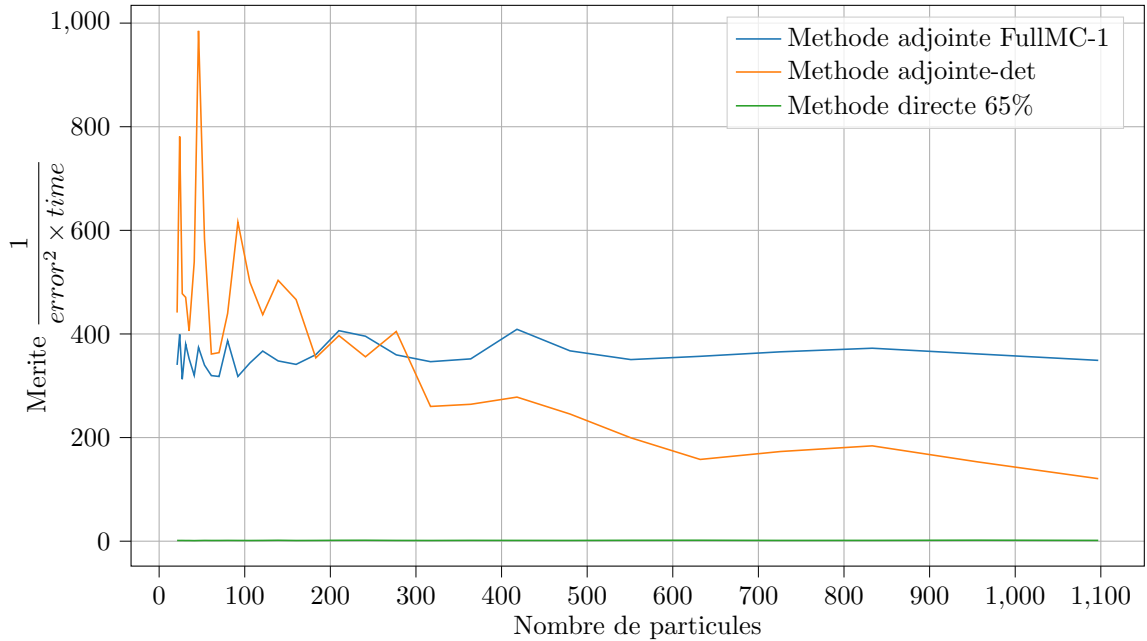
de discrétisation choisi pour l'intégrale, que ces points soient déterministes ou aléatoires). Pour économiser du temps de calcul, on peut à la place évaluer chaque particule une seule fois en un instant aléatoire (ce qui revient à une méthode FullMC avec 1 point de discrétisation différent pour chaque particule, d'où l'appellation FullMC-1 par la suite). Ainsi, avec l'augmentation du nombre de particules, on garde une convergence de l'intégrale par la loi des grands nombres. Cela est motivé par le fait que sans cela, même si on compare les méthodes à nombre total de particules égal, le coût des nombreuses évaluations des particules en méthode adjointe rend celle-ci significativement plus longue d'exécution.

On va donc tracer ici une courbe adjointe avec 15 points de discrétisation déterministes (adjointe-det), et une courbe FullMC avec 1 point de discrétisation aléatoire par particule.

Pour prendre en compte l'aspect "temps d'exécution", on donne aussi la figure de mérite associée à ces calculs.



Comparaison mérite MCA/MCD - 1D Avec Source



Conclusions

On remarque tout d'abord que la méthode hybride adjointe-déterministe ne décroît pas tout à fait en $\frac{1}{\sqrt{N}}$, ce qui signifie que dans ce test, une erreur de troncature significative subsiste.

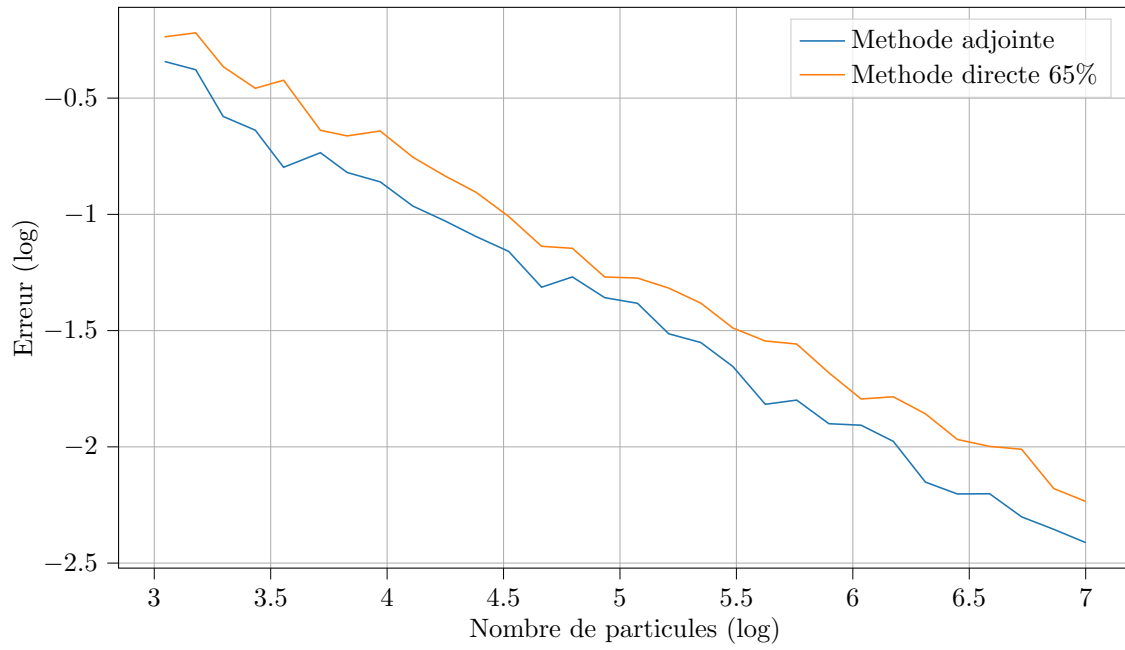
Pour les méthodes dont la décroissance de l'erreur est en $\frac{1}{\sqrt{N}}$, on observe bien un plateau pour les figures de mérite.

Dans notre cas, la méthode adjointe FullMC-1 est plus rapide d'exécution (car on n'évalue chaque particule qu'une seule fois), mais présente une erreur légèrement supérieure. La figure de mérite tracée nous indique que, assez rapidement, la méthode FullMC-1 est plus efficace. Enfin, on peut noter qu'avec la méthode FullMC-1, l'erreur tend bien vers 0 avec l'augmentation du nombre de particules. C'est pour cela que l'on va dorénavant éliminer la méthode adjointe-déterministe au profit d'une méthode FullMC-1 avec un seul point de discrétisation de l'intégrale.

Validation du scattering

Dans un second temps, on trace les courbes d'erreur pour un maillage constant et une condition initiale ne dépendant pas de x , mais avec un scattering non nul, afin de vérifier que l'implémentation du scattering n'est pas source d'erreurs.

Comparaison MCA/MCD - 0D avec algo 1D (log-log)



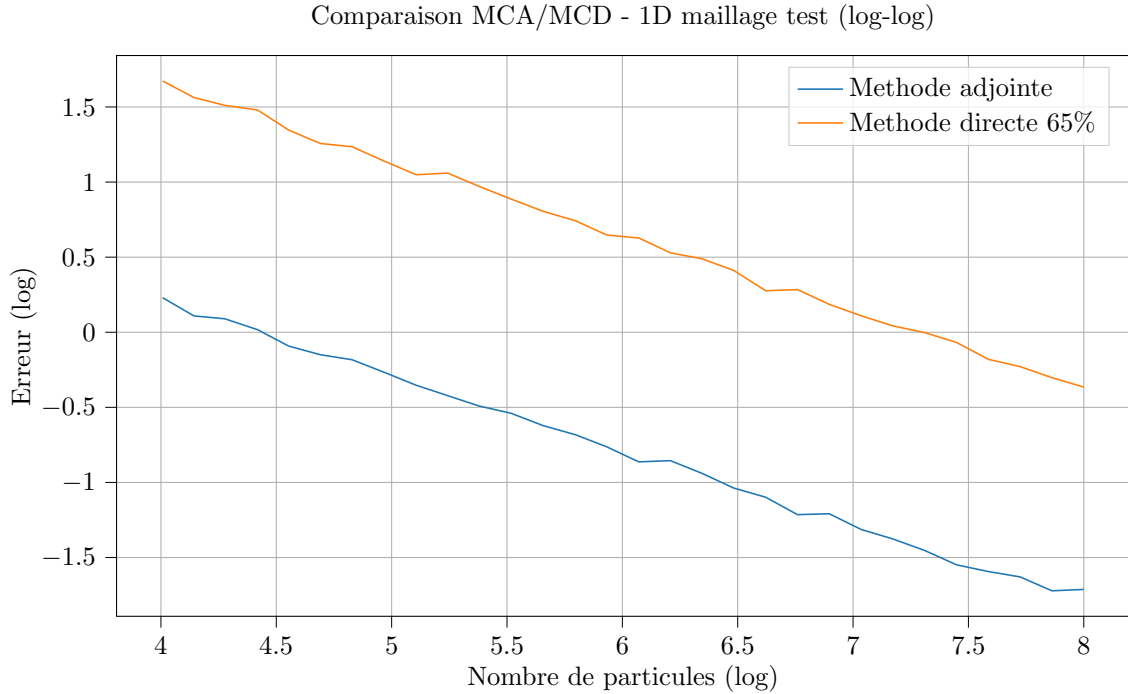
On observe à nouveau la décroissance voulue, et on peut donc considérer que nos algorithmes sont valides.

5.2.4 Comparaison MCA - MCD avec terme source et dépendance spatiale

On commence avec un terme source nul. On ne dispose plus de la solution analytique, on va donc prendre à la place une trajectoire calculée par méthode adjointe⁹ avec 10^6 particules.

On reprend le maillage décrit précédemment, et on réalise des tests avec les paramètres suivants :

- $u_0 : (x, \omega) \mapsto x(1 - \frac{x}{5})\mathbb{1}_{[0,5]}(x)$
- $\phi : (x, \omega) \mapsto \mathbb{1}_{[2,3]}(x)$
- $f = 0.25$
- Pour chaque N , on calcule un millier de trajectoires en 20 instants $t \in (0, 2)$ pour chaque méthode et on affiche l'erreur en norme 1 moyenne de ces trajectoires.



On rappelle qu'ici, puisque 5 mailles sont en jeu, on a 5 fois plus de particules pour la méthode directe. Cette courbe sert de base de comparaison pour le paragraphe suivant : ici, on n'impose pas d'instant intermédiaire à la méthode adjointe, d'où sa grande précision. On verra ensuite ce qu'il se passe si on ajoute un ou plusieurs instants intermédiaires.

5.3 Étude sur de nombreuses itérations d'IMC

Une question à laquelle nous n'avons pas répondu en partie précédente est celle de la propagation de l'erreur. Si on considère de nombreuses itérations d'IMC (une centaine, par exemple) et donc autant d'approximations de $u(t_i, \cdot, \cdot)$, il est possible que les erreurs d'approximations s'accumulent.

Nous nous contenterons de considérer plusieurs itérations en reprenant à chaque fois les mêmes valeurs pour σ_s et σ_a , même si, en pratique, ils peuvent être fonction de la température et dépendent donc du résultat de l'itération IMC précédente. On teste ainsi uniquement l'ac-

9. Les mêmes tests ont été réalisés en prenant la méthode directe pour ce calcul et les résultats restent les mêmes.

cumulation des erreurs d'approximation sur les $u(t_i, \cdot, \cdot)$, mais pas l'impact d'une erreur sur les coefficients de l'équation de transport de l'itération suivante.

On va réaliser des tests avec l'algorithme 1D sur deux fonctions test¹⁰ pour cette étude : la fonction $u_0 : (x, \omega) \mapsto \mathbb{1}_{[3,5]}(x)$ et la fonction $u_0 : (x, \omega) \mapsto x(1 - \frac{x}{5})$.

Pour la première fonction, l'approximation par valeurs moyennes est exacte, et la seule erreur contenue dans u_1^{moy} est d'ordre aléatoire. On sait qu'au cours du temps, les particules ont tendance à sortir du domaine, ce qui diminue la précision de la méthode adjointe. En ajoutant des instants intermédiaires, on produit à chaque instant intermédiaire une nouvelle génération de particules qui sont toutes dans le domaine, et donc la population de particules utile à chaque instant est plus grande. Dans ce cas où l'approximation par valeurs moyennes est efficace, on s'attend donc à ce qu'un grand nombre d'instants intermédiaires diminue l'erreur. Cette diminution devrait être assez visible pour un faible nombre de particules mais plus faible pour un grand nombre de particules.

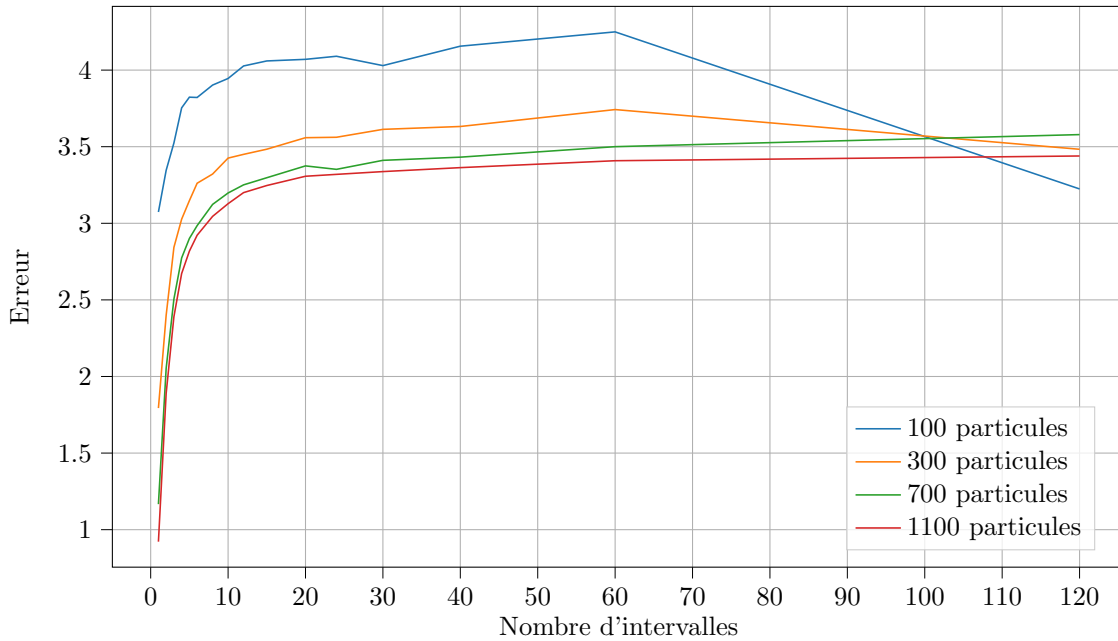
Pour la seconde fonction, on s'attend à une erreur augmentant avec le nombre d'intervalles car même avec un grand nombre de particules, la méthode d'approximation par valeurs moyennes sera peu efficace. L'erreur dominante n'est alors pas d'ordre stochastique mais déterministe, régie par la proposition 4.2.

On va calculer des trajectoires (1920 trajectoires pour chaque méthode et pour chaque nombre d'intervalles) pour $0 \leq t \leq 2$ en 120 points, en sorte de pouvoir faire varier le nombre d'intervalles de la manière suivante :

- 1 intervalle où la trajectoire est calculée en 120 points (i.e. aucun instant intermédiaire)
- ...
- 60 intervalles où la trajectoire est calculée en 2 points (i.e. 59 instants intermédiaires)
- 120 intervalles où la trajectoire est calculée en 1 point (i.e. 119 instants intermédiaires)

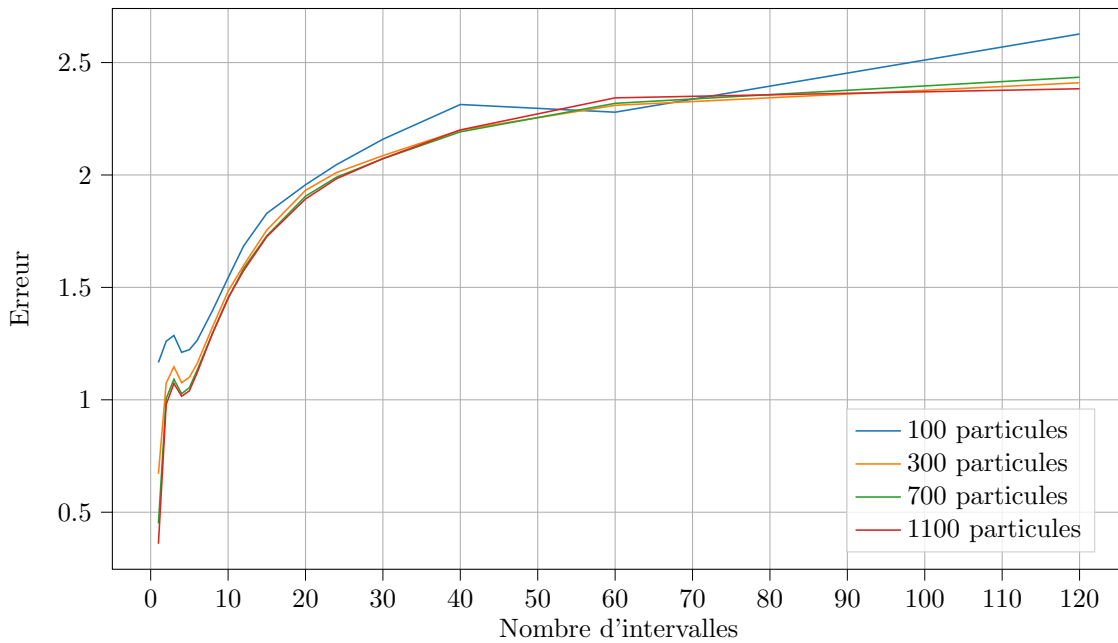
10. Attention, ces fonctions ne sont pas normalisées, et, selon la maille concernée par ϕ , on peut avoir des ordres de grandeur différents pour ces deux fonctions. On ne peut donc comparer directement la valeur de l'erreur, il faut se contenter d'étudier la forme des courbes obtenues.

Erreur MCA en fonction du nombre d'intervalles - u_0 constante par morceaux



En comparaison, une méthode directe avec 10 000 particules donne une erreur de 3.5 en moyenne (20 échantillons).

Erreur MCA en fonction du nombre d'intervalles - $u_0 : (x, \omega) \mapsto x(1 - x/5)$



En comparaison, une méthode directe avec 10 000 particules donne une erreur de 2.8 en moyenne (20 échantillons).

Conclusions

On observe bien ce à quoi on s'attendait au niveau de la forme des courbes, quoi que même avec une bonne approximation de u_0 , l'erreur ne rechute que peu. Cela est sûrement dû au fait que, avec peu de particules, les $\hat{u}_i^{moy}(\cdot, \cdot)$ restent de mauvaise qualité.

Même au bout d'un grand nombre d'itérations, l'erreur observée est comparable à l'erreur que l'on aurait obtenue avec une méthode directe, et on en conclut donc que la méthode adjointe aurait été viable dans ces deux cas.

On peut cependant remarquer que la convergence vers la bonne trajectoire avec le nombre de particules n'a plus lieu, comme prédit par la proposition 4.2. L'erreur qui subsiste a une double nature : d'une part, en espace, puisque l'on moyenne u sur la maille, une partie de l'énergie est "téléportée". D'autre part, le maillage en espace dont on dispose pour l'instant ne prend pas compte le fait que même si u_0 ne dépend pas de ω , il est possible que u en dépende ultérieurement. Lorsque l'on ajoute un instant intermédiaire, on génère des particules uniformément sur la sphère unité, et on ajoute donc un évènement de scattering artificiel pour toutes nos particules. Ainsi, on observe bien une convergence de la trajectoire estimée, mais vers une trajectoire distincte (bien que proche) de la trajectoire cible. C'est pourquoi l'erreur se stabilise aussi bien. Pour pallier ce problème, il faudrait aussi mailler l'espace des ω .

5.4 Approximation par k -plus proches voisins

5.4.1 Présentation de la méthode

Le problème précédemment cité peut être résolu en remplaçant l'approximation par valeurs moyennes par une de ses variantes : l'approximation par les k -plus proches voisins (noté KNN). Le principe est le suivant : lors d'un pas de temps en méthode adjointe, on a besoin d'évaluer, pour chaque particule, $u(t_i, \cdot, \cdot)$. On dispose, d'autre part, des particules du pas de temps précédent : on connaît leur position à t_i , leur direction de propagation, ainsi que leur énergie (nos données). On peut donc, via une méthode KNN, faire une prédiction pour chaque particule.

Remarquons que, puisqu'en méthode adjointe, les particules "remontent" le temps, la donnée à retenir pour l'algorithme KNN est leur position et angle à l'instant 0 de la simulation (soit au moment de leur génération), qui correspond à l'instant t_1 de l'interprétation. L'énergie à retenir, en revanche, est bien celle à t_1 , c'est à dire celle estimée par la méthode adjointe pour cette particule. Cela a un avantage : puisque les particules sont générées de manière uniforme sur chaque maille, on sait que les données sur lesquelles l'algorithme KNN va se baser sont également réparties de manière uniforme, et donc on ne risque pas d'avoir de zones "creuses" de l'espace, où aucune donnée ne serait présente.

Cette méthode est une "variation" de la méthode d'approximation par valeurs moyennes, car elle consiste en réalité à faire une approximation par valeurs moyennes avec une sous-maille spécifique à chaque particule (la maille contenant les k voisins les plus proches). On conserve donc les propriétés de l'approximation par valeurs moyennes. Mieux encore, si on note $k(N)$ le nombre de voisins utilisés par la méthode (avec N le nombre de particules par maille), on peut faire la remarque suivante :

- Si $k(N)/N \xrightarrow{N \rightarrow \infty} 1$, la méthode va se rapprocher de l'approximation par valeurs moyennes car les particules utilisées pour la moyenne seront les mêmes.
- Si $k(N) = \alpha N, \alpha < 1$, la méthode va faire une moyenne avec moins de particules que la méthode par valeurs moyennes, mais ces particules seront plus proche de la particule cible.

- On observera donc une variance plus élevée, mais aussi une erreur déterministe plus faible.
- Si $k(N)/N \xrightarrow{N \rightarrow \infty} 0$ et $k(N) \xrightarrow{N \rightarrow \infty} \infty$, alors le théorème centrale limite s'applique, et, lorsque $N \rightarrow \infty$, on va bien observer la convergence de la méthode, à la vitesse $\frac{1}{\sqrt{k(N)}}$.

Ces remarques sont liées au fait que cette méthode est une méthode de type Monte-Carlo. En effet, on fait une estimation par valeurs moyennes (i.e. on estime une intégrale) de u sur un intervalle variable. La taille de cet intervalle est proportionnelle au rapport $0 \leq k(N)/N \leq 1$. Ainsi, si ce rapport tend vers 1, cela revient à estimer la valeur moyenne sur toute la maille, si ce rapport tend vers une valeur $0 < \alpha < 1$, cela revient à faire une valeur moyenne sur une maille dont le volume serait réduit d'un facteur α par rapport à la maille originale, et si ce rapport tend vers 0, cela revient à une estimation ponctuelle de u .

Cette méthode est beaucoup plus coûteuse à deux niveaux :

- En termes de temps de calcul, il est nécessaire de faire une recherche de plus proches voisins pour chaque particule. En utilisant des méthodes heuristiques, on peut gagner du temps de calcul et conserver un temps de calcul proportionnel au nombre de particules, mais il restera grand. La différence est moins visible avec l'augmentation de la dimension, car la méthode KNN se scale sans avoir à faire d'adaptations.
- En termes de mémoire, cette méthode implique de garder en mémoire toutes les particules du pas de temps précédent.¹¹

Elle permet néanmoins de s'affranchir du besoin d'un maillage en ω , est robuste, et assure la convergence avec le nombre de particules pour peu que le nombre de voisins soit choisi de manière adéquate.

5.4.2 Paramètres de la méthode

Pour mettre en place cette méthode, on a plusieurs degrés de liberté :

- (1) Le nombre k de voisins à choisir. Notons que dans nos méthodes, augmenter k ne fait pas augmenter le temps de calcul.
- (2) On peut choisir de garder ou non un maillage en espace.
Si on garde un maillage en espace, la recherche des k plus proches voisins se fait parmi un jeu de données nettement plus petit. Par exemple, si une particule arrive dans une certaine maille, on peut rechercher ses voisins parmi les particules de cette maille ainsi que des voisines, sans avoir à chercher au delà, et sans sacrifier la précision.
Si on ne garde pas de maillage en espace, on doit faire une recherche parmi beaucoup plus de particules, mais, dans certains cas, cela permet d'éviter un biais.
- (3) Une fois les k voisins choisis pour une particule, on peut choisir de prédire l'énergie de la particule en faisant une moyenne classique de l'énergie des k voisins ou bien de faire une moyenne pondérée (par exemple par l'inverse de la distance à la particule cible dans l'espace x, ω).

Pour (2), il semblerait que l'économie computationnelle de conserver un maillage l'emporte sur le reste.

Pour (3), dans tous les cas que nous avons pu tester, la méthode la plus simple et la plus performante à la fois était de pondérer par l'inverse de la distance en norme 2 (i.e. on calcule pour chaque voisin la distance à la particule cible, et on pondère par l'inverse de cette distance). Néanmoins, une moyenne classique conserve un avantage : on peut imaginer une méthode qui

¹¹. Dans le cadre des méthodes Monte-Carlo pour le transfert radiatif, la mémoire est le principal facteur limitant le nombre de particules utilisées.

consiste à diviser cette maille en sous mailles telles que dans une sous-maille, les particules ont toutes les mêmes k plus proches voisins, et donc la même valeur pour l'énergie (ce qui n'est pas possible si l'énergie prédite est pondérée par la distance). Cette variante permet d'économiser un coût computationnel et en mémoire.

Pour (1), nous avons cherché à travers de nombreux tests une heuristique permettant de dire combien de voisins prendre, en vain. Pour cela, nous avons fait des tests similaires aux tests réalisés sur de nombreuses itérations d'IMC, en cherchant à minimiser l'erreur obtenue. Sur ces tests, on a pu remarquer que :

- Le nombre optimal de voisins dépend des variations de la fonction : si la fonction varie fortement, il est préférable de prendre moins de voisins pour éviter de prendre des valeurs trop éloignées.¹²
- Le nombre optimal de voisins ne semble pas tendre vers $+\infty$ avec le nombre de particules. Dans nos exemples, le nombre idéal de voisins était souvent autour de 14 à partir de 700 particules. Cela signifie qu'à partir de ce nombre de voisins, l'erreur en $\frac{1}{\sqrt{k(N)}}$ n'est plus dominante, i.e. l'estimation est assez précise. L'erreur dominante devient alors l'erreur liée à la méthode adjointe.

5.4.3 Performance

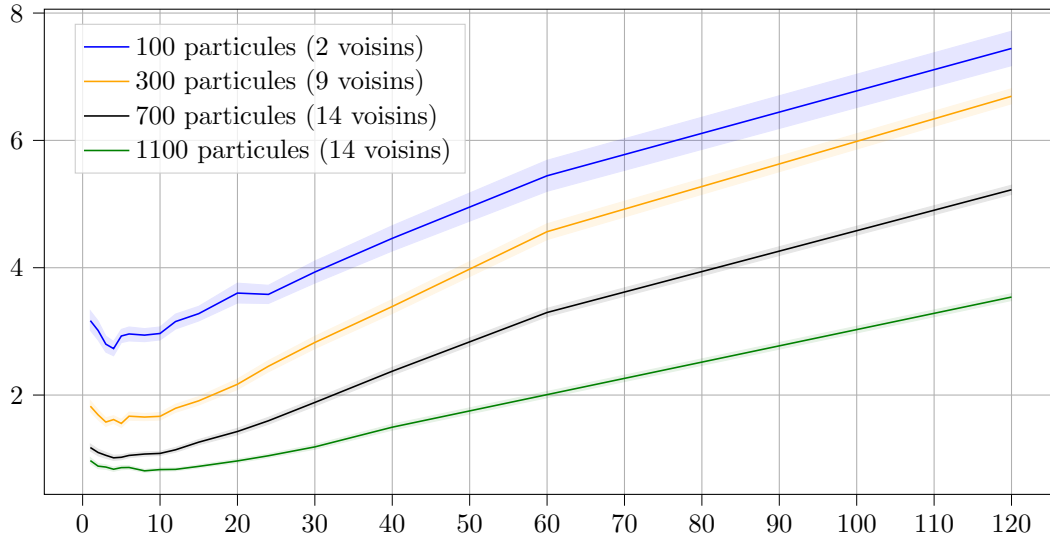
On procède aux mêmes tests que précédemment. Les paramètres utilisés pour la méthode KNN sont les suivants :

- Le nombre de voisins (cf. graphes) est choisi de manière optimale grâce à une recherche extensive réalisée préalablement
- Un maillage est conservé
- La moyenne des k plus proches voisins est pondérée (norme 2 de la distance à la particule cible)

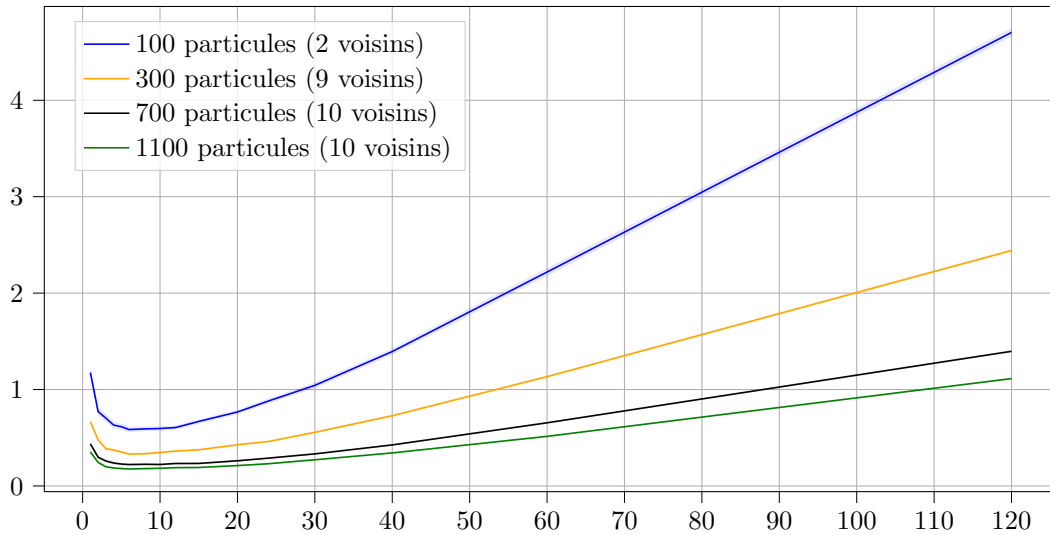
Pour ce test, on trace également l'intervalle de confiance à 95%.

12. Cette remarque suggère que, dans (3), on pourrait aussi prendre en compte la variabilité de la fonction en ajoutant dans la pondération le sup de la dérivée de u , par exemple

Erreur MCA-KNN en fonction du nombre d'intervalles - u_0 escalier



Erreur MCA-KNN en fonction du nombre d'intervalles - $u_0 : (x, \omega) \mapsto x(1 - x/5)$



Commentaires

On constate cette fois que la trajectoire calculée ne converge pas vers une trajectoire voisine (car l'erreur ne se stabilise pas autour d'une certaine valeur), mais l'erreur continue à diminuer lorsque l'on augmente le nombre de particules. La forme des courbes suggère cette fois que plus on a d'instantanés intermédiaires, plus on commet une erreur élevée.

Le niveau d'erreur que l'on parvient à obtenir, pour peu que l'on puisse utiliser assez de particules, est plus faible qu'avec une simple méthode d'approximation par valeurs moyennes. En revanche, avec très peu de particules, puisque cette méthode augmente la variance, l'erreur devient très grande.

Le temps de calcul est environ 7 fois plus grand avec une méthode KNN qu'avec une méthode par valeurs moyennes (ce qui est à nuancer avec le fait que cette méthode n'est pas plus longue pour $(x, \omega) \in \mathbb{R}^3 \times \mathcal{S}^2$, contrairement à l'approximation par valeurs moyennes).

On retiendra que l'avantage principal de cette méthode est d'offrir une convergence avec le nombre de particules, propriété essentielle des calculs par méthodes de Monte-Carlo. Néanmoins, cette convergence est lente, et implique un fort coût en mémoire.

6 Conclusion

On a résolu en partie 4 la question d'obtenir une approximation des $u(t_i, \cdot, \cdot)$: la méthode la plus efficace est l'approximation par valeurs moyennes, qui consiste à approcher u par sa valeur moyenne sur une maille ou une sous-maille.

La méthode MCA répond aux attentes : on peut effectivement choisir le nombre de particules par maille et donc il est plus aisé de contrôler la variance de nos résultats. En revanche, à nombre de particules égal, elle n'est pas significativement meilleure (voir partie 4.6). On voit la différence lorsqu'un terme source entre en jeu, car cela impose pour la méthode directe d'y allouer des particules, ce qui n'est pas le cas en méthode adjointe (voir parties 5.1.5 et 5.2.4).

Pour le traitement de l'intégrale du terme source en méthode adjointe, la méthode d'intégration la plus efficace est la méthode "FullMC-1", qui consiste à évaluer chaque particule en un instant aléatoire différent (voir parties 5.1.4 et 5.2.3). Même s'il est plus simple d'obtenir une grande précision avec une discrétisation déterministe (même avec peu de particules), celle-ci est trop coûteuse.

Pour prendre en considération des conditions aux limites de types flux entrant, une méthode est donnée en annexe 7.3.

Dans tous les cas que nous avons pu considérer, nous sommes arrivés à la conclusion que la méthode adjointe était donc viable.

Enfin, nous avons également introduit une variante de la méthode d'approximation par valeurs moyenne, l'approximation par les k plus proches voisins. Cette méthode présente une variance plus grande, demande plus de calcul et plus de mémoire, mais permet de récupérer la propriété de convergence des trajectoires avec le nombre de particules. Elle permet également de s'affranchir du besoin de mailler en angle le domaine, et ne souffrira pas lors du passage en 3D.

7 Annexes

7.1 Présentation du CEA

Cette partie est extraite du "Livret d'accueil des stagiaires du DSSI sur le site de Bruyères-le-Châtel" du CEA.

7.1.1 Le CEA

Acteur majeur de la recherche, du développement et de l'innovation, le Commissariat à l'Énergie Atomique et aux Énergies Alternatives intervient dans quatre domaines :

- La défense et la sécurité
- Les énergies bas carbone (nucléaires et renouvelables)
- La recherche technologique pour l'industrie
- La recherche fondamentale (sciences de la matière et sciences de la vie)

S'appuyant sur une capacité d'expertise reconnue, le CEA participe à la mise en place de projets de collaboration avec de nombreux partenaires académiques et industriels.

Le CEA est implanté sur 9 centres répartis dans toute la France. Il développe de nombreux partenariats avec les autres organismes de recherche, les collectivités locales et les universités. À ce titre, le CEA est partie prenante des alliances nationales coordonnant la recherche française dans les domaines de l'énergie (ANCRE), des sciences de la vie et de la santé (AVIESAN), des sciences et technologies du numérique (ALLISTENE), des sciences de l'environnement (AllEnvi) et des sciences humaines et sociales (ATHENA).

Reconnu comme un expert dans ses domaines de compétence, le CEA est pleinement inséré dans l'espace européen de la recherche et exerce une présence croissante au niveau international.

Le CEA compte 19 925 techniciens, ingénieurs, chercheurs et collaborateurs pour un budget de 5 milliards d'euros (chiffres publiés fin 2018).

7.1.2 La DAM

Le CEA/DAM - Île de France (DIF) est l'une des directions opérationnelles de la DAM (Direction des Applications Militaires).

Le site de la DIF compte environ 2000 salariés CEA et accueille quotidiennement environ 600 salariés d'entreprises extérieures. Il est situé à Bruyères-le-Châtel à environ 40 km au sud de Paris, dans l'Essonne.

Les missions de la DIF comprennent :

- La conception et garantie des armes nucléaires, grâce au programme Simulation. L'enjeu consiste à reproduire par le calcul les différentes phases du fonctionnement d'une arme nucléaire et à confronter ces résultats aux mesures des tirs nucléaires passés et aux résultats expérimentaux obtenus sur les installations actuelles (machine radiographique, lasers de puissance, accélérateurs de particules).
- La lutte contre la prolifération et le terrorisme, en contribuant notamment au programme de garantie du Traité de Non-Prolifération et en assurant l'expertise technique française

- pour la mise en œuvre du Traité d'Interdiction Complète des Essais Nucléaires (TICE) .
- L'expertise scientifique et technique, dans le cadre de la construction et du démantèlement d'ouvrages complexes ainsi que pour la surveillance de l'environnement et les sciences de la terre.
- L'alerte des autorités, mission opérationnelle assurée 24h sur 24, 365 jours par an, en cas d'essai nucléaire, de séisme en France ou à l'étranger, et de tsunami dans la zone Euro-méditerranéenne. La DIF fournit aux autorités les analyses et synthèses techniques associées.

Depuis 2003, le centre DAM-Île-de-France héberge le complexe de calcul scientifique du CEA, qui regroupe l'ensemble des supercalculateurs du CEA, et qui comprend à ce jour :

- Le supercalculateur EXA1 pour les besoins du programme Simulation du CEA/DAM, successeur de TERA 1000, dispose d'une puissance de calcul de 23,2 petaflops, c'est à dire capable d'effectuer 23,2 millions de milliards d'opérations par seconde.
- Les ordinateurs du Centre de Calcul pour la Recherche et la Technologie (CCRT), ouverts à la communauté de la recherche et de l'industrie, pour une puissance globale de 8,8 petaflops.
- Le supercalculateur Joliot-Curie, d'une puissance de 22 petaflops, deuxième élément d'un réseau de supercalculateurs de classe petaflopique destiné aux chercheurs de la communauté scientifique européenne. Ce supercalculateur est hébergé au TGCC (Très Grand Centre de Calcul) et exploité par les équipes du CEA, qui apporte ainsi sa contribution à la participation de la France au projet PRACE (Partnership for Advanced Computing in Europe).

7.2 Preuve de la formule de Feynman-Kac

On rappelle le théorème (3.1) :

Théorème. *Pour une condition initiale u_0 donnée, on a :*

$$u(t, x, \omega) = \mathbb{E}_{x, \omega} \left[u_0(X(t), \Omega(t)) e^{-\int_0^t \sigma_a(t-\zeta, X(\zeta), \Omega(\zeta)) d\zeta} + \int_0^t f(t-s, X(s), \Omega(s)) e^{-\int_0^s \sigma_a(t-\zeta, X(\zeta), \Omega(\zeta)) d\zeta} ds \right]$$

Pour en faire la preuve, on va reprendre les notations de [2]. Le processus $(X(t), \Omega(t))$ est remplacé par un processus $(X(t), V(t))$ à valeurs dans $\mathbb{R}^3 \times E$, la fonction σ_a est remplacée par $-c$, et la fonction σ_s par une fonction A , la condition initiale est notée g . On rappelle la définition du générateur infinitésimal d'un processus :

Définition 7.1. *Soit $\{(X(t), V(t)), t \geq 0\}$ un processus de Markov. Son générateur infinitésimal \mathcal{A} est l'opérateur qui, à toute fonction mesurable bornée f , associe*

$$\mathcal{A}f : (x, v) \mapsto \lim_{h \rightarrow 0^+} \frac{\mathbb{E}_{x, v} [f(X(h), V(h)) - f(x, v)]}{h}$$

Proposition 7.1. *Le générateur du processus adjoint est :*

$$\mathcal{A}f : (x, v) \mapsto -\nabla f(x, v) + \int_E A(t, x, v, v') f(x, v') dv'$$

(Dans notre cas, A ne dépend pas de v' et sort donc de l'intégrale)

On rappelle également le lemme 2.3.5 de la page 40 de [2] :

Lemme 7.1. *Soit $h \in C_b^1(\mathbb{R}_+ \times \mathbb{R}^3 \times E)$. La fonction*

$$t \mapsto \mathbb{E}_{x, v} \left[\exp \left(\int_0^t c(s, X(s), V(s)) ds \right) h(t, X(t), V(t)) \right]$$

est continûment dérivable, et, $\forall t > 0$, sa dérivée est :

$$\mathbb{E}_{x,v} \left[\exp \left(\int_0^t c(s, X(s), V(s)) ds \right) \left(\frac{\partial h}{\partial t}(t, X(t), V(t)) + \mathcal{A}h(t, X(t), V(t)) + (ch)(t, X(t), V(t)) \right) \right]$$

L'équation de transport que l'on cherche à résoudre peut s'écrire :

$$\begin{aligned} \frac{\partial u}{\partial t}(t, x, v) &= (\mathcal{A}u)(t, x, v) + c(t, x, v)u(t, x, v) + f(t, x, v) \\ u(0, x, v) &= g(x, v), \forall (x, v) \in \mathbb{R}^d \times E \end{aligned} \quad (8)$$

et il faut montrer que $\forall f \in C_b^1(\mathbb{R}_+ \times \mathbb{R}^d \times E)$ et $\forall g \in C_b^1(\mathbb{R}^d \times E)$,

$$\begin{aligned} u(t, x, v) &= \mathbb{E}_{x,v} \left[\exp \left(\int_0^t c(t - \zeta, X(\zeta), V(\zeta)) d\zeta \right) g(X(t), V(t)) \right. \\ &\quad \left. + \int_0^t \exp \left(\int_0^s c(t - \zeta, X(\zeta), V(\zeta)) d\zeta \right) f(t - s, X(s), V(s)) ds \right] \end{aligned}$$

est solution de (8).

Démonstration. Soit $t \in \mathbb{R}_+$ et soit $\bar{u} : (s, x, v) \mapsto u(t - s, x, v)$ avec u solution de l'équation. On a alors :

$$\frac{\partial \bar{u}}{\partial s}(s, x, v) + (\mathcal{A}\bar{u})(s, x, v) + c(t - s, x, v)\bar{u}(s, x, v) + f(t - s, x, v) = 0$$

D'après le lemme 7.1, on a (en considérant donc t comme un paramètre et en remplaçant la fonction $\zeta \mapsto c(\zeta, x, v)$ par $\zeta \mapsto c(t - \zeta, x, v)$) :

$$\begin{aligned} & \frac{d}{ds} \mathbb{E}_{x,v} \left[\exp \left(\int_0^s c(t - \zeta, X(\zeta), V(\zeta)) d\zeta \right) \bar{u}(s, X(s), V(s)) \right] \\ &= \mathbb{E}_{x,v} \left[\exp \left(\int_0^s c(t - \zeta, X(\zeta), V(\zeta)) d\zeta \right) \left(\frac{\partial \bar{u}}{\partial s}(s, X(s), V(s)) + (\mathcal{A}\bar{u})(s, X(s), V(s)) \right. \right. \\ &\quad \left. \left. + c(t - s, X(s), V(s))\bar{u}(s, X(s), V(s)) \right) \right] \\ &= -\mathbb{E}_{x,v} \left[\exp \left(\int_0^s c(t - \zeta, X(\zeta), V(\zeta)) d\zeta \right) f(t - s, X(s), V(s)) \right] \end{aligned}$$

Or

$$\begin{aligned} & \frac{d}{ds} \mathbb{E}_{x,v} \left[\int_0^s \exp \left(\int_0^z c(t - \zeta, X(\zeta), V(\zeta)) d\zeta \right) f(t - z, X(z), V(z)) dz \right] \\ &= \frac{d}{ds} \int_0^s \mathbb{E}_{x,v} \left[\exp \left(\int_0^z c(t - \zeta, X(\zeta), V(\zeta)) d\zeta \right) f(t - z, X(z), V(z)) \right] dz \\ &= \mathbb{E}_{x,v} \left[\exp \left(\int_0^s c(t - \zeta, X(\zeta), V(\zeta)) d\zeta \right) f(t - s, X(s), V(s)) \right] \end{aligned}$$

Et donc, en sommant les deux calculs précédents,

$$\begin{aligned} & \mathbb{E}_{x,v} \left[\exp \left(\int_0^s c(t-\zeta, X(\zeta), V(\zeta)) d\zeta \right) \bar{u}(s, X(s), V(s)) \right. \\ & \left. + \int_0^s \exp \left(\int_0^z c(t-\zeta, X(\zeta), V(\zeta)) d\zeta \right) f(t-z, X(z), V(z)) dz \right] \end{aligned}$$

est constante. En $s = 0$, cette expression vaut :

$$\bar{u}(0, X(0), V(0)) = u(t, x, v)$$

En $s = t$, elle vaut :

$$\begin{aligned} & \mathbb{E}_{x,v} \left[\exp \left(\int_0^t c(t-\zeta, X(\zeta), V(\zeta)) d\zeta \right) \bar{u}(t, X(t), V(t)) \right. \\ & \left. + \int_0^t \exp \left(\int_0^z c(t-\zeta, X(\zeta), V(\zeta)) d\zeta \right) f(t-z, X(z), V(z)) dz \right] \\ & = \mathbb{E}_{x,v} \left[\exp \left(\int_0^t c(t-\zeta, X(\zeta), V(\zeta)) d\zeta \right) g(X(t), V(t)) \right. \\ & \left. + \int_0^t \exp \left(\int_0^z c(t-\zeta, X(\zeta), V(\zeta)) d\zeta \right) f(t-z, X(z), V(z)) dz \right] \end{aligned}$$

□

7.3 Conditions aux limites

On explique brièvement ici comment traiter un flux entrant en méthode adjointe.

7.3.1 Idée générale

On se place en 1D plan.

On sait qu'en méthode directe, un flux entrant correspond à une émission de particules au cours de la simulation. Ces particules sont émises à des instants t_i aléatoires suivant une certaine loi (par exemple une loi uniforme pour un flux ne dépendant pas du temps), et avec des directions de propagation ω_i aléatoires suivant également une loi (par exemple une loi uniforme dans le cas d'un flux qui émet dans toutes les directions).

Cela correspond à une certaine densité d'émission $(t, \omega) \mapsto f(t, x, \omega)$. L'idée est de faire correspondre à cette densité un terme supplémentaire dans u_0 (non-nul exclusivement à l'extérieur du domaine), i.e. trouver le u_0 artificiel qui crée le flux que l'on veut à la frontière du domaine.

7.3.2 Cas général

On suppose qu'à la frontière $x = 0$, on impose que les instants d'émission suivent une loi à densité f_1 et que les directions de propagation suivent une loi à densité f_2 . On note $f_3(x, \omega)$ la densité jointe de $X(0)$ et $\Omega(0)$, que l'on recherche.

Soit τ le temps d'atteinte de la frontière pour une particule. On a $\tau = -X(0)/\Omega(0)$. Soit h une fonction continue bornée. On veut donc :

$$\mathbb{E}(\tau) = \mathbb{E}(h(-X(0)/\Omega(0))) = \int h(t) f_1(t) dt$$

Avec la contrainte : $\forall g$ continue bornée,

$$\mathbb{E}(g(\Omega)) = \int g(t)f_2(t)dt$$

On va alors chercher une condition suffisante sur f_3 .

Premièrement, il faut prendre f_3 de la forme $f_2 \times \tilde{f}_3$, avec \tilde{f}_3 la densité conditionnelle de $X(0)$ sachant $\Omega(0)$ pour assurer qu'au moment de traverser la frontière du domaine, la répartition en angle suit bien f_2 .

Ensuite,

$$\begin{aligned} \mathbb{E}(h(-X(0)/\Omega(0))) &= \int_{\mathbb{R}_+^2} h\left(-\frac{x}{\omega}\right)f_3(x, \omega)dx d\omega \\ &= \int_{\mathbb{R}_+^2} h(t)f_3(-t\omega, \omega)\omega dt d\omega \\ &= \int_{\mathbb{R}_+} h(t) \left[\int_{\mathbb{R}_+} f_3(t\omega, \omega)\omega d\omega \right] dt \end{aligned}$$

Il faut donc aussi choisir f_3 tel que $\forall t > 0$

$$\int_{\mathbb{R}_+} f_3(-t\omega, \omega)\omega d\omega = f_1(t)$$

7.3.3 Exemple

Soit $s > 0$. On considère un flux entrant constant par morceaux, strictement positif pour $0 < t < s$ et nul pour $t > s$. On a donc :

$$f_1 : t \mapsto \frac{\mathbb{1}_{[0,s]}(t)}{s}$$

$$f_2 : \omega \mapsto \mathbb{1}_{[0,1]}(\omega)$$

On peut prendre

$$f_3 : (x, \omega) \mapsto \frac{\mathbb{1}_{[0,1]}(\omega)\mathbb{1}_{[-\omega s, 0]}(x)}{\omega s}$$

Car on a alors :

$$\begin{aligned} \int_{\mathbb{R}_+} f_3(-t\omega, \omega)\omega d\omega &= \int_{\mathbb{R}_+} \frac{\mathbb{1}_{[0,1]}(\omega)\mathbb{1}_{[-\omega s, 0]}(-t\omega)}{\omega s} \omega d\omega \\ &= \int_{\mathbb{R}_+} \frac{\mathbb{1}_{[0,1]}(\omega)\mathbb{1}_{[0,s]}(t)}{s} d\omega \\ &= \frac{\mathbb{1}_{[0,s]}(t)}{s} \\ &= f_1(t) \end{aligned}$$

On va donc, en méthode adjointe, ajouter à u_0 un terme $\eta f_3(x, \omega)$, où η est une constante de normalisation telle que

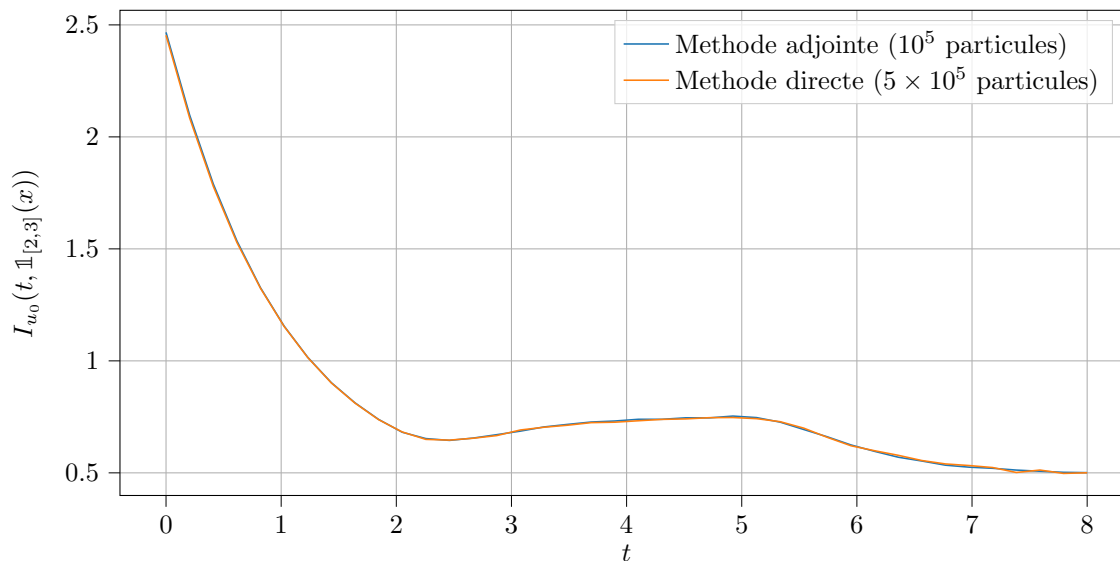
$$\int_{\mathbb{R}_+^2} \eta f_3(x, \omega) dx d\omega = \eta = \int \int f(t, \omega) d\omega dt$$

(où f désigne le flux entrant en $x = 0$)

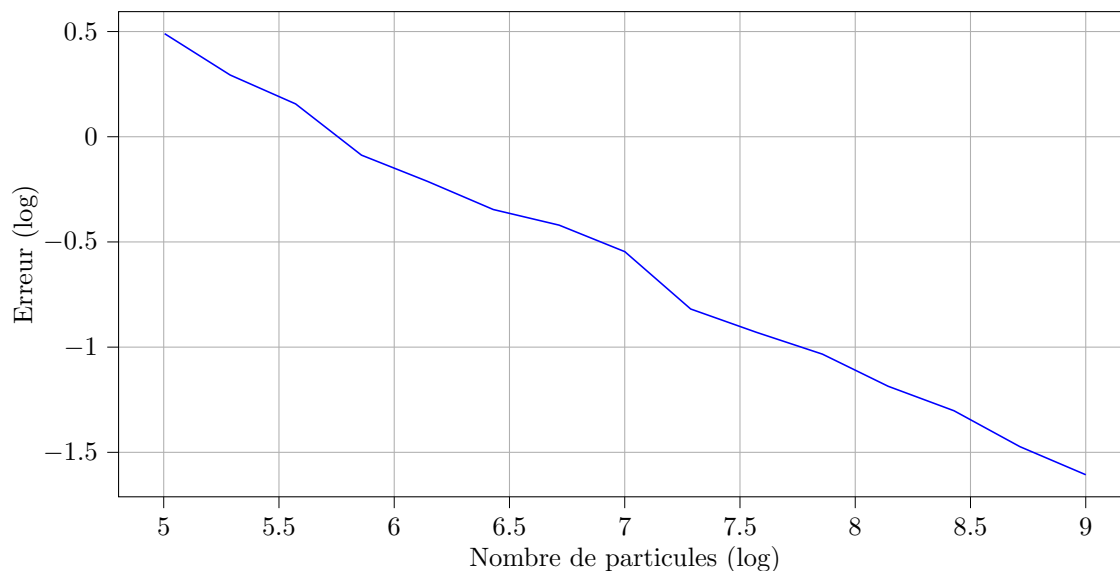
Dans notre cas, en prenant $f(t, \omega) = 10 \times \mathbb{1}_{[0,1]}(\omega) \mathbb{1}_{[0,s]}(t)$, on a $\eta = 10s$.

Test ci-dessous réalisé avec $u_0 : x \mapsto x(1 - x/5)$, un terme source constant et $s = 3$. La trajectoire est calculée en 40 points.

Flux entrant en methode adjointe et directe



On peut également tracer l'erreur de cette méthode en fonction du nombre de particules (on prend comme trajectoire de référence une trajectoire calculée par méthode adjointe avec 5 millions de particules) :



L'erreur est bien en $\frac{1}{\sqrt{n}}$

Références

- [1] J.A. Fleck and J.D. Cummings. An implicit monte carlo scheme for calculating time and frequency dependent nonlinear radiation transport. *Journal of Computational Physics*, 8(3) :313–342, 1971.
- [2] R. Sentis B. Lapeyre, E. Pardoux. *Méthodes de Monte-Carlo pour les équations de transport et de diffusion*. Springer, 1998.
- [3] Allan Wollaber and Edward Larsen. A hybrid monte carlo-deterministic approach to improve the accuracy and efficiency of monte carlo calculations for thermal radiative transfer. 01 2009.
- [4] Allan Wollaber and James Warsa. A new, coupled transport-diffusion method for radiative transfer calculations. 01 2013.
- [5] Allan Wollaber. Four decades of implicit monte carlo. *Journal of Computational and Theoretical Transport*, 2016.
- [6] Valentin Niess, Anne Barnoud, Cristina Cârloganu, and Eve Le Ménédeu. Backward monte-carlo applied to muon transport. *Computer Physics Communications*, 229 :54–67, 2018.
- [7] Huang Yong, Shi Guo-Dong, and Zhu Ke-Yong. Backward and forward monte carlo method in polarized radiative transfer. *The Astrophysical Journal*, 820(1) :9, mar 2016.
- [8] García Muñoz, A. and Mills, F. P. Pre-conditioned backward monte carlo solutions to radiative transport in planetary atmospheres - fundamentals : Sampling of propagation directions in polarising media. *A&A*, 573 :A72, 2015.
- [9] Markus Kampl and Hans Kosina. The backward monte carlo method for semiconductor device simulation. *Journal of Computational Electronics*, 17(4) :1492–1504, 2018.
- [10] Laetitia Laguzet and Gabriel Turinici. A cell-based population control of monte carlo particles for the global variance reduction for transport equations. *Journal of Computational Physics*, 467 :111373, 2022.
- [11] David Legrady, Mate Halasz, Jozsef Kophazi, Balazs Molnar, and Gabor Tolnai. Population-based variance reduction for dynamic monte carlo. *Annals of Nuclear Energy*, 149 :107752, 2020.
- [12] Allan Wollaber. *Advanced Monte Carlo Methods for Thermal Radiation Transport*. PhD thesis, 08 2008.
- [13] Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, 1 edition, 2007.
- [14] G. G. Lorentz. *Bernstein polynomials*. AMS Chelsea Publishing. American Mathematical Society, 2 sub edition, 1997.
- [15] William T. Vetterling Brian P. Flannery William H. Press, Saul A. Teukolsky. *Numerical Recipes in C : The Art of Scientific Computing*. Cambridge University Press, 1992.
- [16] Mama Foupouagnigni and Merlin Mouafo Wouodjié. On multivariate bernstein polynomials. *Mathematics*, 8, 2020.