Modélisations hybrides par méthode d'apprentissage supervisé pour le couplage multi-échelle d'écoulements de gaz faiblement raréfiés en conduites

Dahia Chibouti

▶ **To cite this version:**

Dahia Chibouti. Modélisations hybrides par méthode d'apprentissage supervisé pour le couplage multi-échelle d'écoulements de gaz faiblement raréfiés en conduites. Mécanique des fluides [physics.class-ph]. Université Gustave Eiffel, 2023. Français. NNT : 2023UEFL2032 . tel-04276699

HAL Id: tel-04276699

https://theses.hal.science/tel-04276699

Submitted on 9 Nov 2023

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

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

École doctorale n° 531 : Sciences, Ingénierie et Environnement

T H È S E

pour obtenir le grade de

Docteur en Science

de l'Université Gustave Eiffel

Spécialité : Mécanique des fluides

soutenue publiquement par

Dahia CHIBOUTI

Modélisations hybrides par méthode d'apprentissage supervisé pour le couplage multi-échelle d'écoulements de gaz faiblement raréfiés en conduites

le 20 juillet 2023

Devant le jury composé de :

Rapporteurs.es :	Irina MARTIN	-	Polytech'Marseille
	Guillaume GALLIÉRO	-	Université de Pau (LFCR)
Examinateurs.es :	Stéphane COLIN	-	INSA Toulouse (ICA)
	Céline LEONARD	-	Université Gustave Eiffel (MSME)
	Jean-Luc BATTAGLIA	-	Université de Bordeaux (I2M)
Directeur :	Eric CHÉNIER	-	Université Gustave Eiffel (MSME)
Co-Encadrant :	Benoît TROUETTE	-	Université Gustave Eiffel (MSME)

Laboratoire Modélisation et Simulation Multi Échelle (MSME)
Univ Gustave Eiffel, CNRS UMR 8208

UNIVERSITY GUSTAVE EIFFEL

DOCTORAL SCHOOL SIE
SCIENCES ENGINEERING ENVIRONMENT

PHD THESIS

to obtain the title of

Doctor of Science

of the University Gustave Eiffel

Field : Fluid Mechanics

Defended by

Dahia CHIBOUTI

Hybrid modeling by supervised machine learning for multiscale coupling of weakly rarefied gas flows in channels.

Thesis Advisors : Eric CHÉNIER & Benoît TROUETTE

prepared at MSME laboratory, TCM Team

defended on July 20^{th}, 2023

Jury :

Reviewers :	Irina MARTIN	-	Polytech'Marseille
	Guillaume GALLIÉRO	-	University of Pau (LFCR)
Examinators :	Stéphane COLIN	-	INSA Toulouse (ICA)
	Céline LEONARD	-	Gustave Eiffel University (MSME)
	Jean-Luc BATTAGLIA	-	University of Bordeaux (I2M)
Advisors :	Eric CHÉNIER	-	Gustave Eiffel University (MSME)
	Benoît TROUETTE	-	Gustave Eiffel University (MSME)

Laboratoire Modélisation et Simulation Multi Échelle (MSME)
Univ Gustave Eiffel, CNRS UMR 8208

The best preparation for tomorrow is doing your best today.

H. Jackson Brown, Jr.

Remerciements

Cher lecteur, bienvenue dans cette rubrique, je vous remercie de prendre le temps de lire les quelques remerciements adressés à tous les acteurs liés à ma thèse, des plus sympathiques et loufoques aux plus sérieux.

Je tiens particulièrement à remercier mes encadrants de thèse, Eric Chénier et Benoit Trouette pour leurs conseils avisés et leurs encouragements.
Je vous écris pour vous exprimer ma reconnaissance, vous avez été un guide précieux et un soutien considérable. Vous avez cru en moi et vous m'avez poussé à donner le meilleur de moi-même. Vous avez été patient, et vous m'avez prodigué vos précieux conseils. Votre savoir et votre expérience ont été pour moi des trésors inestimables et j'ai beaucoup appris grâce à vous. Je vous remercie du fond du cœur pour tout ce que vous avez fait pour moi. Merci d'avoir su rester à l'écoute, notamment dans les instants difficiles, les moments où la pression monte.

Eric, vous êtes un mentor exceptionnel et je suis heureuse de vous avoir comme directeur de thèse. Je vous donne également un franc « merci », d'être là dans l'urgence et à chaque fois, pour des conseils techniques, mais aussi pour des conversations plus personnelles.

Je tiens à remercier mes rapporteurs.es, Irina Martin et Guillaume Galliéro, qui ont accepté de lire ma thèse et donner leurs rapports. Je tiens à exprimer ma sincère reconnaissance envers vous, éminents membres du jury, pour votre expertise éclairée, votre soutien et vos contributions précieuses.
J'adresse également toute ma reconnaissance au directeur de notre équipe Stéphane Vincent pour m'avoir accepté dans le cercle réservé.
Maintenant, je passe aux choses moins sérieuses, mais qui ont forcément une importance considérable. Comme le lecteur pourra s'en douter, la liste est immense.
Je voudrais exprimer ma reconnaissance à ma famille et à mes proches pour leur amour et leur soutien indéfectibles pendant cette période de travail acharné. Je ne pourrais pas avoir terminé cette thèse sans leur amour et leur soutien.

Tout d'abord, un énorme merci à mes parents, qui n'ont JAMAIS cessé de croire en moi. Depuis que je suis toute petite, quand mon père m'apprenait et ma mère qui m'a ouvert la voie à tant de choses. Je voudrais remercier mes sœurs et à mon frère, qui sont toujours là quand je les appelle, même lorsqu'il fait nuit, ou ils sont occupés, ou autre. Rien ne les

empêche de me venir en aide.
Je voudrais dédier ce travail à mes neveux et nièces.
Mes remerciements vont à chacune des personnes qui ont certainement vécu ces soirées de travail, seul ou en groupe, acharnées sur une publication ou sur un bout de code qui « résiste », parfois jusqu'à l'aurore.

Je tiens également à remercier chaleureusement tous mes collègues et mes amis sans exception qui ont contribué à la réalisation de cette thèse : leur expertise m'a bcp aidé.
Je vous remercie tous sincèrement.
Mes remerciements s'envolent, vers tous ceux qui ont contribué, de près ou de loin, à réaliser ce rêve qui m'a tant tenu à cœur.

À tous ceux qui m'ont aidé
À longueur de jours et de nuits
À réaliser ce projet
Je vous dis Merci !

Résumé

Le travail présenté dans ce mémoire concerne les écoulements faiblement raréfiés de gaz dans des conduites dont les longueurs sont très grandes par rapport à leur diamètre. Ce problème porte aussi bien sur des micro-conduites à pression ambiante que sur des conduites de taille centimétrique à très faible pression, dès lors que le diamètre est compris entre 10 à 1000 fois le libre parcours moyen des molécules du gaz λ. Dans ces conditions, la mécanique des milieux continus peut s'appliquer dans le cœur de la conduite, ce qui n'est plus le cas au voisinage des parois sur une épaisseur de l'ordre $\mathcal{O}(\lambda)$. Plus particulièrement, un glissement dynamique et un saut de température sont observés entre les grandeurs du fluide et du mur solide.
Bien qu'il existe des modèles continus exprimant ces sauts en fonction du taux de cisaillement et du flux thermique, ces relations dépendent toujours de quantités qui résultent des interactions entre les molécules du gaz et les atomes du solide. De plus, ces modèles ne permettent pas de tenir compte d'une éventuelle physique particulière qui prendrait place à très petite échelle (adsorption, piégeage de molécules, rugosité atomique de la paroi, $\cdots$).

L'objectif de ce travail est de s'affranchir des modèles de proche paroi pour leur substituer une approche atomique/moléculaire locale et de la coupler avec la mécanique des milieux continus dans le cœur de l'écoulement. Ce couplage multi-échelle repose sur une approche numérique hybride, *volumes finis/dynamique moléculaire*, permettant de faire communiquer les grandes échelles de l'écoulement avec les petites échelles relatives aux interactions de proche paroi. Une simulation moléculaire étant trop coûteuse à réaliser dans le voisinage de la totalité de la paroi, des méthodes d'apprentissage supervisé sont utilisées pour modéliser la réponse moléculaire locale suite à une sollicitation macroscopique également locale. Les simulations de dynamique moléculaire sont alors effectuées dans des domaines très réduits, appelés plots virtuels, indépendamment des valeurs de communication du domaine continu vers celui discret.

À travers une méthode de couplage hybride explicite puis une méthode de coulage hybride implicite plus performante, nous montrons qu'il est possible de retrouver dans des cas académiques les écoulements obtenus avec les simples relations de glissement de Navier/Maxwell avec une très bonne précision. Notre méthode de couplage hybride est ensuite appliquée à géométries plus complexes et à des parois atomiques où le glissement peut notablement différer des modèles classiques.

Abstract

The work presented in this thesis focuses on weakly rarefied gas flows in channels whose lengths are very large compared to their diameter. This problem concerns both micro-pipes at ambient pressure and centimetric size pipes at very low pressure, since the diameter is between 10 and 1000 times the mean free path of the gas molecules λ. Under these conditions, continuum mechanics can be applied in the bulk, which is not the case in the vicinity of the walls over a thickness of the order of $\mathcal{O}(\lambda)$. In particular, a dynamic slip and a temperature jump are observed between the fluid and solid wall quantities.
Although continuous models exist expressing these jumps as a function of shear rate and heat flux, these relationships still depend on quantities that result from interactions between gas molecules and solid atoms. Moreover, these models do not allow to take a possible particular physics into account which would take place at very small scale (adsorption, trapping of molecules, atomic roughness of the wall, $\cdots$).

The objective of this work is to move away from near-wall models to a local atomic/molecular approach and to couple it with continuum mechanics in the flow environment. This multiscale coupling is based on a hybrid numerical approach, *finite volume/molecular dynamics*, allowing to communicate the large scales of the flow with the small scales related to the near wall interactions. Since a molecular simulation is too expensive to perform in the vicinity of the entire wall, supervised learning methods are used to model the local molecular response to a local macroscopic solicitation. Molecular dynamics simulations are then performed in very small domains, called virtual plots, independently of the communication values from the continuous to the discrete domain.

Through an explicit hybrid coupling method and then a more efficient implicit hybrid casting method, we show that it is possible to recover in academic cases the flows obtained with the simple Navier/Maxwell slip relations with a very good accuracy. Our hybrid coupling method is then applied to more complex geometries and to atomic walls, where the slip may differ significantly from classical models.

Table des matières

1 Introduction **16**
- 1.1 Effets d'échelles sur les systèmes physiques . . . 16
- 1.2 Modélisation d'écoulements micro/nano échelle . . . 19
  - 1.2.1 Modélisation approchée d'écoulements micro/nano échelle . . . 21
  - 1.2.2 Modélisation moléculaire d'écoulements micro/nano échelle . . . 22
  - 1.2.3 Modélisation multi-échelle d'écoulements micro/nano échelle . . . 23
- 1.3 Synthèse des méthodes hybrides et multi-échelles . . . 24
- 1.4 Plan de la thèse . . . 32

2 Modèles de la méthode hybride et multi-échelle **35**
- 2.1 Modèles physiques . . . 36
  - 2.1.1 Modèle moléculaire . . . 36
  - 2.1.2 Modèle continu . . . 41
- 2.2 Méthodes numériques . . . 43
  - 2.2.1 Solveur micro . . . 43
  - 2.2.2 Solveur macro . . . 54

3 Identification des conditions aux limites thermiques par dynamique moléculaire **56**
- 3.1 Introduction . . . 57
- 3.2 Continuum description . . . 57
- 3.3 Microscopic approach . . . 58
- 3.4 Results and discussions . . . 60
  - 3.4.1 Total energy flux analysis . . . 61
  - 3.4.2 Comparison between the molecular and continuum models . . . 63
- 3.5 Synthèse des cas étudiés . . . 68
- 3.6 Estimation des incertitudes sur le flux de conduction . . . 70
- 3.7 Conclusion . . . 71

4 Méthodes de régression pour l'apprentissage supervisé **72**
- 4.1 Généralités sur l'apprentissage . . . 73
  - 4.1.1 Régression linéaire ou logistique [42] . . . 75
  - 4.1.2 Régression par machines à vecteurs de support [6, 12, 82] . . . 75
  - 4.1.3 Réseaux de neurones [6, 76] . . . 76
  - 4.1.4 Méthodes de k plus proches voisins (k nearest neighbors, kNN) [4] . . . 78
- 4.2 Régression linéaire . . . 78
  - 4.2.1 Régression par la méthode des moindres carrées . . . 78
  - 4.2.2 Régression par la méthode bayésienne . . . 83
  - 4.2.3 Régression par les processus gaussiens . . . 91
- 4.3 Étude comparative des différentes techniques de mise à l'échelle . . . 96
  - 4.3.1 Outils de normalisation et de mise à l'échelle . . . 97

4.3.2 Choix de la fonction de base . . . 98
4.3.3 Résultats de comparaison . . . 99
4.4 Conclusion . . . 104

5 Méthodes hybrides et multi-échelles pour les écoulements en nano/micro-canaux de grande extension 109
5.1 Introduction . . . 110
5.2 Principe des méthodes de couplage . . . 111
5.3 Couplage fort par la méthode des plots . . . 112
5.3.1 Couplage par plots moléculaires . . . 112
5.3.2 Couplage hybride entre les solveurs continu et discret . . . 113
5.3.3 Inconvénients de la méthode hybride de couplage explicite par plots . 115
5.4 Couplage faible par la méthode des plots virtuel . . . 115
5.4.1 Du couplage fort au couplage faible : éléments principaux d'améliorations . . . 115
5.4.2 Couplage hybride faible explicite par la méthode des plots virtuels . . 117
5.4.3 Construction de la réponse du plot virtuel . . . 121
5.4.4 Validation de la procédure de couplage itératif, performance et coûts de calcul . . . 126
5.5 Couplage faible hybride multi-échelle implicite par la méthode des plots virtuels 134
5.5.1 Du couplage explicite au couplage implicite : éléments principaux d'amélioration . . . 134
5.5.2 Couplage hybride faible implicite par la méthode des plots virtuels . . 135
5.5.3 Construction de la réponse du plot virtuel . . . 139
5.5.4 Validation de la procédure de couplage itératif, performance et coûts de calcul . . . 142
5.5.5 Apprentissage de la fonction $f_{d_n u_\tau}$ dans une boite de dynamique moléculaire . . . 153
5.6 Conclusion . . . 155

6 Conclusion générale et perspectives 157

A Méthode de dynamique moléculaire 160
A.1 Propriétés physiques . . . 160
A.1.1 Conductivité thermique . . . 160
A.1.2 Viscosité dynamique . . . 162
A.2 Calcul du flux par la dynamique moléculaire . . . 164
A.3 Description du maillage co-localisé . . . 169

B Contexte sur l'apprentissage de hyperparamètres 171
B.1 Fonction de vraisemblance, priori, posteriori, production productive . . . 171
B.1.1 Fonction de vraisemblance . . . 171
B.1.2 Distribution *a priori* . . . 173
B.1.3 Distribution *a posteriori* . . . 173
B.1.4 Distribution prédictive . . . 175
B.2 Apprentissage des hyperparamètres : Bayésien/Gaussien . . . 176
B.2.1 Développement pour le cas de regression . . . 177
B.2.2 Développement pour le cas de classification . . . 179
B.2.3 Hyperparamètres de la méthode de régression par processus gaussiens 181
B.3 Calcul de dérivées des Noyaux des Processus Gaussiens . . . 181

Table des figures

1.1 Différentes applications des systèmes microélectromécaniques, micro-electro-mechanical-system (MEMS) Ejeian et al., (2019) [25]. 17

1.2 Représentation schématique du phénomène de fissuration radiale des pastilles dans le cas d'un espace fermé entre les pastilles (*pellet* en gris) et la gaine (*clad* en bleu clair), figure de [37, 67] modifiée. 18

1.3 Différents régimes de raréfaction Gad-el-Hak, (2005) [33] 20

1.4 Configuration d'écoulement entre deux plaques plates, montrant les conditions aux limites de glissement dynamique. 22

1.5 Configurations décrivant les différentes méthodes, hybrides et multi-échelle, existantes, ainsi que les diverses manières de couplage et de transmission des informations. Les zones en rouge représentent les simulations moléculaires, et les zones blanches maillées en bleu le domaine continu. 25

1.6 CGMD Sans recouvrement [85-87]. À droite : le maillage bleu correspondant aux éléments finis est raffiné (la zone jaune) de sorte à avoir une taille similaire au diamètre de la particule de dynamique moléculaire (zone rouge). À gauche : les grains de dynamique moléculaire en bleu sont de plus en gros (dans la zone jaune) pour que leurs tailles soient équivalentes au maillage éléments finis en vert. 25

1.7 DDM avec recouvrement O'Connell et Thompson, (1995) [78] et [19, 75, 84, 98, 99, 102, 103]. La ligne en pointillés délimite les intervalles utilisés pour calculer la moyenne des vitesses des particules dans la zone de recouvrement (overlap). 26

1.8 HMM par Ren et E, (2005) [83]. Les micro-simulations en rouge sont réparties dans tout le domaine. Les différentes configurations des micro-simulations sont présentées à droite de la figure. 28

1.9 IMM par Borg et al., (2013) [8]. Les micro-simulations en rouge sont réparties tout au long du canal. 28

1.10 Méthode multi-échelle de Ritos et al., (2015) [84], serial-network internal-flow multiscale method (SeN-IMM), pour le traitement des nanotubes. Les zones rouges correspondent à la dynamique moléculaire et les zones blanches aux zones résolues par le formalisme continu. 29

1.11 Méthode multi-échelle de Vu et al., (2016) [99], Méthode hybride atomique-continu, pour le traitement des nano/micro canaux. La région hachurée en bleu représente le domaine continu, les rectangles gris pour les plots de micro-simulations. 31

1.12 Couplage faible Zhang et al., (2017) [103] 31

2.1 Représentation du potentiel de Lennard-Jones en fonction de la distance intermoléculaire r_{ij}. Le paramètre ε est la profondeur du puits. σ est la distance de van der Walls et est la distance constante entre deux atomes pour laquelle le potentiel s'annule. Généralement σ est considérée comme le rayon de la molécule. Le cube est un domaine moléculaire où les molécules (en bleu) interagissent avec la molécule i (en rouge). 38
2.2 Représentation du potentiel de Lennard-Jones, et de la force en fonction de la distance intermoléculaire r_{ij}. 39
2.3 Représentation des conditions aux limites périodiques. La boîte de simulation est au centre en gris. Les boîtes imaginaires sont autour sur un fond blanc. La molécule (en rouge) qui quitte le domaine de simulation est réintroduite à l'opposé [97]. 45
2.4 Boite de simulation avec des murs atomiques en bleu qui interagissent avec les particules du fluide en rouge. 47
2.5 Moyenne en temps 49
2.6 Schéma présentant la procédure du calcul de la variance, M étant le nombre d'accumulations/transformations. Les blocs (en bleu) sont regroupés deux à deux, jusqu'à en avoir qu'un seul qui représentera l'incertitude finale sur une moyenne calculée. 50
2.7 Estimation de l'écart-type σ obtenue avec la méthode de moyenne par blocs. Après 7 transformations par bloc ($M = [4, 5, \ldots, 10]$, les estimations restent quasi-constantes à l'intérieur des barres d'erreur à 2×10^{-8}. 51
2.8 Profil de vitesse (en unité LJ) dans le cas d'utilisation des murs stochastiques. 52
2.9 Configuration de la boite de simulation. Les murs 1,2 sont thermostatés à une température fixe T_1, et dans les murs 3,4 les molécules sont des molécules fantômes et servent à maintenir les murs stables ainsi que des conditions limites périodiques pour l'ensemble de la boîte de stimulation. 52
2.10 Profils de vitesse et de température pour un écoulement de Couette avec les murs atomiques. 53
2.11 Maillage Delaunay, K, L, M sont des mailles voisines, dans cette figure la maille K est une maille de bord. les x_K, x_L, x_M sont les centres de mailles où les variables sont calculées. Une figure plus détaillée est en annexe (Fig. A.5). 55

3.1 Physical configuration 58
3.2 Normalized fluxes with respect to $|\bar{q}_k|/q_0 = 2.74 \times 10^{-2}$, the z-average of the conduction heat flux, for $Kn = 0.036$, $\rho/\rho_0 = 0.6$ and $g_x/g_0 = 0.01$ 62
3.3 Relative difference between the total energy fluxes $q_{tot}^{(m)}$ (Eq. 3.5) and $q_{tot}^{(c)}$ (Eq. 3.2) reconstructed with the molecular data. Filled (empty) symbols stand for points inside (outside) the Knudsen layer ($|z|/H - 1/2) < 2\lambda/H = 4Kn$. For clarity reason, the error bars have been omitted. Reference values are given in Tab. 3.2 63
3.4 Density $\rho(z)/\bar{\rho}$ for different volume forces g_x/g_0 and $Kn = 0.036$ (see Tab. 3.2). The number of bins for the molecular average calculation is $n_z = 100$ 64
3.5 Velocity $u_x(z)/\bar{u}_x$ for different volume forces g_x/g_0 and $Kn = 0.036$. The average values are provided in Tab. 3.2. The parameters for the analytical function is $\sigma_\mu = 2.46$ and $\xi = 36.0$. The number of bins for the molecular average calculation is $n_z = 25$, except the encapsulated figures which are based on $n_z = 100$ 64
3.6 Same caption as Fig. 3.5 – Reduced temperature $\theta(z)/\theta_b$ 65
3.7 Same caption as Fig. 3.5 – Total energy flux $q_{tot}^{(m)}(z)/|\bar{q}_{tot}^{(m)}|$. 66

3.8 Relative gap between the total energy fluxes stemmed from molecular dynamics simulations and the analytical expression (3.12). Filled symbols stand for points inside the Knudsen layer ($|z|/H - 1/2) < 2\lambda$, while empty ones denote measurement points in the bulk. For clarity reason, the error bars have been omitted. Reference values are given in Tab. 3.2 67
3.9 Corrélation des flux normalisés par le flux total calculé q_{tot}. (a) : variation du rayon de coupure r_c, (b) : variation de la grille (nombre de bins qui sert au calcul de moyenne), (e) : variation d'épaisseur du mur e, (d) : variation de forces volumiques appliquées g_x, (c) : variation de la profondeur du puits ε. Les symboles vides représentent des points dans la zone "intérieur" définie par $z \in [-H/4, H/4]$ tandis que les symboles remplis indiquent des points de mesure près des murs ($z \in \pm[-H/4, -H/2]$). 69

4.1 Représentation montrant l'hyperplan (y : en rouge) et les marges fonctionnelles $y \pm \epsilon$ [6, 12]. 75
4.2 Schéma d'un réseau de neurones à deux couches [6]. Les variables d'entrée, cachées et de sortie sont représentées par des nœuds, et les paramètres de poids (coefficients) sont représentés par des liens entre les nœuds. 77
4.3 Polynômes de différents ordres M en rouge qui corrèlent des données (points bleus) tirées à partir d'une fonction quadratique $g(x)$ en pointillé en vert. . . 80
4.4 Amplitudes des de différents ordres M en rouge qui corrèlent des données (points noirs) tirées à partir d'une fonction quadratique en pointillé en vert. . 81
4.5 Graphique d'un ensemble d'entraînement comprenant $N = 100$ points de données. la prédiction est réalisée avec un polynôme d'ordre 20. 81
4.6 Amplitudes des coefficients : Comparaison pour différents ordres $M = [\![1, 16]\!]$ entre les deux cas de régression LSE (symboles pleins) et RLS (symboles vides). 82
4.7 Amplitudes des coefficients pour différentes valeurs λ_r, cas $M = 16$. 82
4.8 Illustration synthétique, cas $1D$, d'une distribution conditionnelle gaussienne pour y compte tenu des données $\bar{x} = x$ (points *bleus*). Cette distribution est donnée par l'équation (4.20), dans laquelle la prédiction de régression est la moyenne de celle-ci présentée par la fonction polynomiale $f(x, \bar{w})$ (la courbe *rouge*), et la précision est donnée par le paramètre β. La gaussienne a une valeur plus élevée lorsqu'elle se rapproche de la courbe rouge ou des données et une valeur plus faible lorsqu'elle est plus éloignée. L'objectif est de maximiser la probabilité de y, ce qui revient à encourager la prédiction à s'adapter aux données. 86
4.9 Régression bayésienne en utilisant des fonctions de base polynomiales pour différents ordres M en rouge qui corrèlent des données (points noirs) tirées à partir d'une fonction quadratique en pointillé en vert. La zone colorée montre l'intervalle de confiance à 95% (la moyenne plus et moins 1.96 de la variance). 89
4.10 Représentation des incertitudes sur la prédiction σ_N pour différents ordres polynomiaux $M = [1, \ldots, 8]$, où le bruit aléatoire inclus dans les valeurs de sorties (cibles) y est fixé à $\sigma_n^2 = 1 \times 10^{-2}$. Pour $N \gg$ l'incertitude sur la prédiction $\sigma_N \to \sigma_n^2 \to \varepsilon$. 90
4.11 Graphique représentant les processus gaussiens avec 2 observations. La ligne noire pleine est la prédiction de la moyenne obtenue par $\mathcal{GP}$ et la zone colorée montre l'intervalle de confiance à 95% (la moyenne ± 1.96 de la variance). La gaussienne superposée correspond à la moyenne $m(.)$ et à l'écart type $\sigma(.)$ de la prédiction au point inconnu $x_\star$ 95
4.12 Représentation de l'apprentissage par enrichissement au maximum des incertitudes sur la prédiction σ_N. Pour $N \gg$ l'incertitude sur la prédiction $\sigma_N \to \varepsilon$. Les poinst noirs représente les données d'apprentisage. 96

4.13 Comparaison de la précision des calculs pour les différentes méthodes de mise à l'échelle (4.48) et (4.49), ainsi que les différentes méthodes de régression (Bayes et $\mathcal{GP}$). La différence entre les méthodes SE, MATERN32 et MATERN52 réside dans le noyau utilisé. 99

4.14 Variation des erreurs relatives entre l'approximation (la fonction prédite) et la solution analytique en fonction du nombre de points $N_B = 10^3$ dans la base de données d'apprentissage pour des ordres polynomiaux en pression ($M_p = \{1, 2, 3, 4, 5, 7, 10, 20\}$). 100

4.15 Représentation de l'influence de l'ordre en vitesse sur les erreurs relatives d'approximation entre la fonction prédite et la solution analytique en fonction du nombre de points $N_B = 10^3$ dans la base de données d'apprentissage. Les résultats sont faits sur une gamme des ordres de pression ($M_p = \{1, 2, 3, 4, 5, 7, 10, 20\}$), mais seulement les ordres ($M_p = \{3, 7\}$) sont présentés. 101

4.16 Comportement de la convergence et précision de calcul. Les symboles (carrés) : représentent les fonctions de base, $\phi^{(1)}(p)$, et les symboles (cercles) : représentent les fonctions de base sont, $\phi^{(2)}(p)$. N_B le nombre de points dans la base de données d'apprentissage et M_p l'ordre en pression. 102

4.17 Comparaison du comportement de la variation des erreurs relatives $\varepsilon_2^{(p)}$ et les incertitudes de prédiction σ_N en fonction du nombre de points N_B dans la base de données d'apprentissage. 103

4.18 Surfaces de réponses des cas d'apprentissage par la méthode de régression bayésienne polynomiale $M_u = 1$; $M_p = 5$ respectivement par ligne. Les points bleus représentent les données d'apprentissage et les croix rouges pour indiquer le nouveau point ajouté permettant d'enrichir la base de données. 103

4.19 Comparaison : enrichissement avec une grille uniforme comme celle des plans d'expériences, soit la grille parcours les deux variables p et u en bleu et pour p 105

4.20 Erreurs relatives entre l'approximation (la fonction prédite pour des données bruitées) et la solution analytique pour un nombre de points N_B dans la base de données d'apprentissage en fonction de l'ordre polynomial en (M_p,M_u). La première colonne : enrichissement avec Max(σ) pour la méthode bayésienne, deuxième colonne : enrichissement par ajout de points aléatoirement, troisième colonne : enrichissement avec Max(σ) pour les processus gaussiens. Par lignes, les bruits sont respectivement : $10^{-4}, 10^{-3}, 10^{-2}$. 106

4.21 Surfaces de réponses des cas d'apprentissage par la méthode de régression bayésienne polynomiale $M_u = 1$; $M_p = 5$ respectivement par ligne. 107

4.22 Surfaces de réponses des cas d'apprentissage par la méthode de régression bayésienne polynomiale $M_u = 1$; $M_p = 5$ respectivement par ligne. 108

5.1 Représentation multi-échelle d'un micro-canal de hauteur H. (a) : décomposition du domaine en une région bleue, de hauteur H_C, pilotée par les grandes échelles de l'écoulement (mécanique des milieux continus) et une région encadrée en rouge, de hauteur H_M, où les interactions fluide/paroi reposent sur une représentation atomistique de la dynamique. (b) : zoom sur une région de proche paroi. La région discrète est subdivisée en une région de paroi ($H_{M,w}$) et une région fluide ($H_{M,f}$). La zone bleue contenant les points discrets rouges indique la région de recouvrement où cohabitent les représentations continue et atomistique. 111

5.2 Représentation du couplage fort hybride multi-échelle explicite par la Méthode des plots [98, 99]. (a) : décomposition du domaine par en un milieu continu (zone bleue) et en 5 plots de dynamique moléculaire. (b) : région supérieure du plot moléculaire avec la zone de recouvrement H_0 subdivisée en 4 sous-zones permettant les communications entre les régions continu et moléculaire. . . . 113

5.3 Processus itératif de convergence d'un profil de vitesse pour un écoulement de Poiseuille. Les lignes $M \to C$ et $C \to M$ indiquent les hauteurs pour lesquelles les informations sont échangées entre les modèles continu (C) et moléculaire (M), et la flèche $\to$ le sens de la communication. Les solutions continues (**courbes bleues notées** $c^{(i)}$) et moléculaires (**courbes rouges notées** $m^{(i)}$) sont présentées pour les trois premières itérations du couplage. Les courbes notées (C) et (M) correspondent aux solutions convergées, c'est-à-dire à une demi-parabole. 114

5.4 Couplage explicite faible, par la méthode des plots virtuels. 118

5.5 Couplage explicite faible, par la méthode des plots virtuels. Les points bleus représentent les centres de mailles à la frontière. Les triangles rouges désignent l'information issue de dynamique moléculaire et imposée au continu comme étant la condition aux limites. La grandeur ϕ peut être la vitesse, pression, température, gradient de pression, etc. 118

5.6 Processus itératif de convergence dans le cas de couplage itératif faible avec des plots virtuels. Les lignes $M \to C$ et $C \to M$ indiquent les hauteurs pour lesquelles les informations sont échangées entre le modèle continu (C) et les réponses issues des plots moléculaires virtuels (M) ; la flèche $\to$ souligne le sens de la communication. Les solutions continues (courbes bleues notées $c^{(i)}$) et les valeurs issues des solutions moléculaires (cercles rouges notées $u^{(i+1)}_{\tau,H_{M\to C}} = f^{(i+1)}_{u_\tau}(u^{(i)}_\tau(\bar{x}_{C\to M}), p^{(i)}(\bar{x}_{C\to M}), (d_\tau p)^{(i)}(\bar{x}_{C\to M}), H_{C\to M}, H_{M\to C})$) sont présentées pour les trois premières itérations du couplage pour une coupe verticale donnée. 120

5.7 Dimensions du plot virtuel de dynamique moléculaire et représentation des boites ("bins") dans lesquelles sont calculées des valeurs moyennes des vitesses selon $\bar{\tau}_P$. 125

5.8 Canal plan $2D$ 128

5.9 Erreur relative de la solution hybride $\varepsilon^{(c)}_\infty$ en fonction de l'écart relatif entre deux itérés successifs $\varepsilon^{(i)}_\infty$ pour un cas écoulement de Couette incompressible, isotherme à $T_{in} = 300K$ avec $H_{M\to C}/H = 10\%$ et différentes valeurs de $(H_{C\to M} - H_{M\to C})/H$ allant de $1,25\%$ à 10%, avec 80 mailles dans la section de hauteur $H - H_{M\to C}$. 129

5.10 Erreur relative de la solution hybride $\varepsilon^{(c)}_\infty$ en fonction de l'écart relatif entre deux itérés successifs $\varepsilon^{(i)}_\infty$ pour, $H_{M\to C}/H = 1.25\%, 2.5\%, 5\%, 10\%$, cas d'un écoulement de Couette incompressible, isotherme à $T_{in} = 300K$, avec 80 mailles dans la section de hauteur $H - H_{M\to C}$. 129

5.11 Erreur relative à la solution exacte $\varepsilon^{(c)}_\infty$ en fonction des itérations de couplage i et différentes valeurs $(H_{C\to M} - H_{M\to C})/2H$. Écoulement de Couette avec $H_{M\to C}/2H = 5\%$ 130

5.12 Nombre d'itérations de couplage $N^{(i)}$ pour atteindre un écart relatif entre deux itérations $\varepsilon^{(i)}_\infty = 10^{-8}$. Les valeurs entre parenthèses indiquent l'erreur relative à la solution exacte $\varepsilon^{(c)}_\infty$. Écoulement de Couette avec $H_{M\to C}/2H = 5\%$. . . . 130

5.13 Convergence itérative du couplage hybride vers la solution exacte de Poiseuille pour $p_{in} = 2.10^5$, $p_{out} = 5.10^4$, $H_{M\to C}/H = 10\%$ et $(H_{C\to M} - H_{M\to C})/H = 10\%$, pour un cas incompressible et un glissement constant, avec 80 mailles dans la section de hauteur $H - H_{M\to C}$. La valeur i indique le numéro de l'itération. 131
5.14 Nombre d'itérations de couplage $N^{(i)}$ pour atteindre un écart relatif entre deux itérations successives $\varepsilon_\infty^{(i)} = 10^{-8}$ en fonction de la différence entre les hauteurs de communication $\Delta H/H = (H_{C\to M} - H_{C\to M})/H$. Les valeurs entre parenthèses indiquent l'erreur relative à la solution exacte $\varepsilon_\infty^{(c)}$. Écoulement de Poiseuille incompressible et glissement constant avec 80 mailles sur une hauteur $H - H_{M\to C}$ (conditions de symétrie). 132
5.15 Erreur de convergence $\varepsilon_\infty^{(c)}$ en fonction de $\varepsilon_\infty^{(i)}$, pour $(H_{C\to M} - H_{M\to C})/H = 1.25\%, 2.5\%, 5\%, 10\%$ et $H_{M\to C}/H = 5\%$. Écoulement de Poiseuille incompressible et glissement constant avec 80 mailles sur une hauteur $H - H_{M\to C}$ (conditions de symétrie). 133
5.16 Erreur de convergence $\varepsilon_\infty^{(c)}$ en fonction de $\varepsilon_\infty^{(i)}$, pour $H_{M\to C}/H = 1.25\%, 2.5\%, 5\%, 10\%$ et $(H_{C\to M} - H_{M\to C})/H = 1.25\%$. Écoulement de Poiseuille incompressible et glissement constant avec 80 mailles sur une hauteur $H - H_{M\to C}$ (conditions de symétrie). 133
5.17 Couplage explicite faible, par la méthode des plots virtuels. 136
5.18 Représentation graphique du couplage faible hybride implicite par la méthode des plots virtuels. Les labels (1), (2) correspondent aux étapes **1** et **2** de l'algorithme, page 136. Les labels (3), (4), (5) (encadré en rouge) schématisent la construction de la réponse moléculaire dans le plot virtuel (**Étape 3**). Les labels (3) et 4 font référence à la dynamique moléculaire ((Sec. 5.5.3.2), page 141) et le label (5) se rapporte à l'algorithme d'apprentissage ((Sec. 5.5.3.1)). Le cadre vert (label (6)) se réfère à l'**Étape 4** de l'algorithme hybride implicite. Les flèches intermédiaires montrent les différentes possibilités qui existent : ainsi, lorsque les bornes [min, max] des valeurs extraites ne varient pas, les labels (2), (3) et (4) sont ignorés. Le processus d'apprentissage ("label (5)") est directement mis en œuvre (la flèche verte au milieu). Á chaque passage dans l'algorithme d'apprentissage (label (5)) un nouveau point est appris, il est alors opportun de définir (calculer) la cible d_{nu_K} par l'intermédiaire de la dynamique moléculaire (flèche rouges). 138
5.19 Écart relatif de la solution hybride à la solution volumes finis en fonction du nombre d'itérations de couplage pour l'apprentissage des fonctions f_{u_w} et $f_{d_n u_\tau}$ par des processus Gaussiens, avec les noyaux SE, Matern32 et Matern52 144
5.20 Données d'apprentissage et surfaces de réponse pour les deux premières itérations de couplage réalisées avec le processus Gaussien SE, $i = 0$ et $i = 1$. Figs. 5.20(a) et 5.20(c) : Les surfaces rouge et bleue correspondent à la solution exacte $d_n u_\tau$ et celle apprise $f_{d_n u_K}$. Les points bleus représentent les données de la base d'apprentissage. La surface colorée dans le plan (u_K, Kn_K) donne la distribution des incertitudes σ_N relatives à la fonction d'apprentissage et le cercle rouge localise la position où le maximum est atteint, $(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})$. Le cercle vert correspond à la donnée (u_K, Kn_K) où u_K résulte d'une simulation de Couette dans le plot virtuel avec une vitesse de la paroi supérieure $f_{u_w}(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})$ $(\mathrm{Kn}_K = \widetilde{\mathrm{Kn}_K})$. Figs. 5.20(b) et 5.20(d) : Le cercle de couleur verte est inséré dans la base d'apprentissage (point bleu) et une nouvelle surface de réponse est apprise (surface rouge), laquelle sera transmise à la méthode des volumes finis pour reconstruire la condition de glissement à la paroi. 145

5.21 Écart relatif de la solution hybride à la solution volumes finis, $\varepsilon_\infty^{(c)}$, en fonction de l'erreur d'apprentissage du cisaillement $\varepsilon_\infty^{(p)}$. 150
5.22 Solution hybride obtenue dans un canal de section variable. 150
5.23 Écart relatif de la solution hybride à la solution volumes finis pour l'écoulement dans un canal de section variable. Le noyau du processus Gaussien est SE. . . 151
5.24 Écarts relatifs de la solution hybride à la solution volumes finis, $\varepsilon_\infty^{(c)}$ (gauche), et incertitude σ_N sur la prédiction du cisaillement appris $f_{d_n u_\tau}$ (droite) en fonction du nombre d'itérations et du nombre M_{u_K} et M_{Kn_K} de polynômes dans la base d'apprentissage. 152
5.25 Écarts relatifs entre deux types de solutions hybrides. L'un est obtenu avec une boite moléculaire munie de parois stochastiques diffuses et le second s'appuie sur une solution analytique glissante. Gauche : modèle bayésien pour $M_{\mathrm{Kn}_K} = 3$ et $M_{u_K} = 1$. Droite : modèle gaussien pour le noyau SE. 153
5.26 Boite moléculaire avec des murs atomiques rugueux. 154
5.27 Écarts relatifs de la solution hybride à celle volumes finis avec des conditions de glissement. La comparaison se fait en utilisant les régressions bayésiennes polynomiales $M_{\mathrm{Kn}_K} = 3$ et $M_{u_K} = 1$ avec des murs atomiques lisses et rugueux. 155

A.1 Principe de calcul de conductivité thermique dans `LAMMPS` par la méthode de Müller-Plathe [69],(*mp*). 161
A.2 Principe de calcul de viscosité dynamique du fluide dans `LAMMPS` par la méthode de Müller-Plathe [69],(*mp*). 162
A.3 Variation de temps de calcul τ de la viscosité μ, en fonction du nombre de processus N_{proc} dans `LAMMPS` pour la méthode de Müller-Plathe [69](*mp*) et Green-Kubo (*gk*) [34]. 164
A.4 Erreurs max sur la viscosité μ, pour différentes fréquences d'échanges de molécules (N_{Ex}). 165
A.5 Illustration bidimensionnelle du maillage de Delaunay – Notations principales utilisées pour la discrétisation 169

B.1 Représentation des incertitudes sur la prédiction σ_N pour différents ordres polynomiaux $M = [1, \ldots, 8]$, où le bruit aléatoire inclus dans les valeurs de la variable cible $f_\star$ est fixé à $\varepsilon = 1 \times 10^{-2}$. Pour $N >>$ l'incertitude sur la prédiction $\sigma_N \to \varepsilon$. 176

Liste des tableaux

2.1 Échelles physiques, expressions et valeurs utilisées dans les simulations de dynamique moléculaire. 39

3.1 Physical scales, expressions and values used in the molecular dynamics simulations, with k_B is the Boltzmann constant 59
3.2 Normalization parameters for Figs. 3.3–3.8. The Knudsen number is $Kn = 0.036$, except the simulation with $g_x/g_0 = 0.002$ where $Kn = 0.0075$ 61
3.3 Dimensionless parameters for simulations presented in Figs. 3.3 and 3.8 ; σ_μ and ξ are used in the analytical functions (3.8), (3.11) and (3.12). 68

5.1 Erreur de couplage $\varepsilon_\infty^{(c)}$ à la solution exacte en fonction du maillage de la zone continue, pour une erreur $\varepsilon_\infty^{(i)} = 10^{-8}$, $H_{M\to C}/H = 10\%$ et $(H_{C\to M} - H_{M\to C})/H = 5\%$. Cas incompressible et glissement constant. 132
5.2 Tableau de données constituant la base d'apprentissage potentielle B_{pot} aux itérations de couplage $i = 0$ et $i = 1$. Pour une valeur de i donnée, l'entier j numérote les données dans la base d'apprentissage potentielle B_{pot} ; seules les lignes écrites en noir appartiennent à la base d'apprentissage B_{app}, celles en gris en sont exclues. La notation NS souligne que les données sont issues de Navier-Stokes (modèle continu) et App indique l'apprentissage du quadruplet $(u_K, \mathrm{Kn}_K, u_w = f_{u_w[j_0;\cdots;j_1]}(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K}), d_n u_K)$ pour lequel l'incertitude du modèle/données sur la base d'apprentissage contenant les lignes j_0, $\cdots$ et j_1 est maximale pour $(u_K, \mathrm{Kn}_K) \in [\mathrm{NS}^{(min,k)}, \mathrm{NS}^{(max;k)}]$. 146

A.1 Valeurs de conductivité κ, pour différentes densités ρ et températures T. Les simulations ont été effectuées sur 2×10^6 d'itérations. 161
A.2 Valeurs de la viscosité μ, pour différentes densités ρ et températures T. Les simulations ont été effectuées sur 5×10^7 d'itérations. 163
A.3 Valeurs de la viscosité μ, pour différents nombres de processus N_{proc}. Les simulations ont été effectuées sur 5×10^6 d'itérations. 164
A.4 Valeurs de la viscosité μ, pour différentes fréquences d'échanges de molécules (N_{Ex}). 165
A.5 Valeurs de la viscosité μ, pour différentes densités ρ et températures T. Les simulations ont été effectuées sur 5×10^7 d'itérations ; nombre de molécules échangées $N_{Ex} = 200$; nombre de couches cisaillées $N_l = 20$. 165

There is plenty of room at the bottom.

Richard Phillips Feynman

Chapitre 1

Introduction

Sommaire

1.1 Effets d'échelles sur les systèmes physiques 16
1.2 Modélisation d'écoulements micro/nano échelle 19
1.2.1 Modélisation approchée d'écoulements micro/nano échelle 21
1.2.2 Modélisation moléculaire d'écoulements micro/nano échelle 22
1.2.3 Modélisation multi-échelle d'écoulements micro/nano échelle 23
1.3 Synthèse des méthodes hybrides et multi-échelles 24
1.4 Plan de la thèse 32

Préambule

Ce chapitre présente toute la richesse, la diversité et la complexité des écoulements dans les systèmes à grandes extensions, et expose ainsi l'état de l'art sur les stratégies de modélisations classiques ou basées sur des méthodes hybrides et multi-échelles qui permettent de modéliser la physique à petite échelle. Enfin, le plan de ce manuscrit de thèse est annoncé à la fin de ce chapitre.

1.1 Effets d'échelles sur les systèmes physiques

Les dispositifs mécaniques, électroniques et biochimiques de taille micrométrique sont de plus en plus répandus, tant dans les applications commerciales que dans la recherche scientifique. La figure 1.1 dresse un panorama de différents cadres d'applications. Ils s'étendent des dispositifs de pointe, tels que les accéléromètres micrométriques utilisés pour déployer les systèmes d'airbag dans les automobiles, aux systèmes de dosages biologiques originaux constitués de réseaux microfluidiques conçus pour l'administration de médicaments.

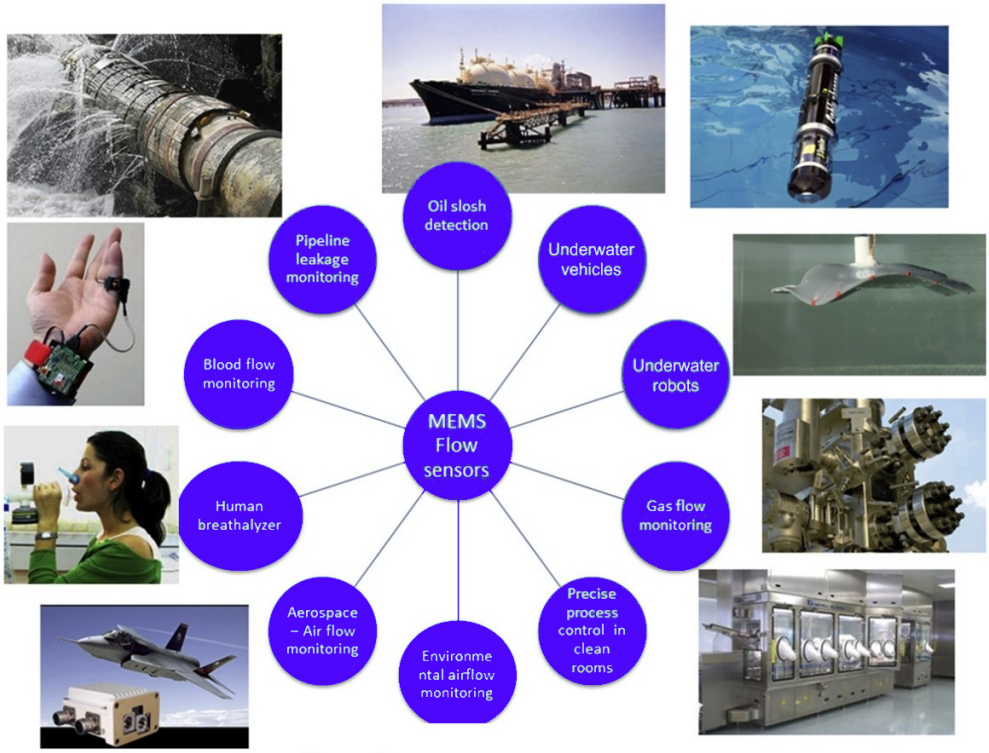


FIGURE 1.1 – Différentes applications des systèmes microélectromécaniques, micro-electro-mechanical-system (MEMS) EJEIAN et al., (2019) [25].

Bien que la conception et les applications de ces systèmes micro-électro-mécaniques ont connu des évolutions technologiques fulgurantes au cours des deux dernières décennies, ces progrès, comme l'indique GAD-EL-HAK [33], n'ont pas toujours été accompagnées d'avancées correspondantes dans la compréhension de la physique non conventionnelle mise en jeu pour leur fabrication et fonctionnement. En effet, les écoulements de fluides dans les micro-dispositifs diffèrent de ceux des objets macroscopiques auxquels nous sommes habituellement confrontés [33, 50]. Comme cela sera décrit et détaillé par la suite, la miniaturisation des systèmes inverse le rapport volume/surface et rend les effets proches parois dominants.

De ce fait, le fonctionnement des micro-conduites et d'autres micro-dispositifs, ne peut pas toujours être prédit à partir des modèles d'écoulements conventionnels (continus) tels que ceux décrits par les équations de Navier-Stokes. En raison des effets de raréfaction [33, 50], le comportement du fluide et, subséquemment, la dissipation induite sur de tels dispositifs peuvent ne pas être les mêmes qu'à l'échelle macroscopique. Les plus petites échelles de longueur des micro-dispositifs, au voisinage des parois fluide/solide, ne sont plus, dans une certaine mesure, directement modélisables avec la mécanique et des milieux continus. Cependant, il est possible d'étendre la mécanique des milieux continus à des écoulements de gaz faiblement raréfiés par l'introduction de modèles de glissement au niveau des parois.

Pour des conduites, il est facile de concevoir que la miniaturisation augmente le rapport surface/volume ; les effets de surface deviennent alors prépondérants vis-à-vis de ceux de volume. Les forces d'inertie, par exemple, y sont généralement assez faibles. De plus, les frottements et les effets visqueux deviennent de plus en plus importants à mesure que la

dimension des dispositifs se réduit, et doivent être quantifiés avec précision. Ces nouvelles technologies basées sur la miniaturisation nécessitent le développement de la science fondamentale dédiée aux petits dispositifs.

Un exemple de phénomène crucial à maîtriser est, par exemple, l'interaction pastille-gaine au sein des réacteurs nucléaires. En effet, l'espace entre la gaine (*clad* en bleu, Fig. (Fig. 1.2)) et la pastille (*pellet* en gris) se referme progressivement sous l'effet de plusieurs mécanismes. Lors de la mise sous tension initiale, l'expansion de la pastille et la pression exercée par le liquide de refroidissement sur la gaine réduisent l'espace initial entre l'oxyde et la gaine [37]. La tension causée par la pastille gonflée provoque une micro fissure dans la couche d'oxyde de la gaine et met à nu le matériau de la gaine. En conséquence, les micro fissures et les fragments de la plaquette acquièrent un aspect incurvé et entrent en contact avec la gaine et se propagent dans les zones interstitielles. Les fragments créés par les fissures peuvent se déplacer vers l'extérieur et provoquer des pics de contrainte locaux. En outre, les fissures facilitent la libération des produits de fission et peuvent influencer la nucléation des fissures sur la surface interne de la gaine.

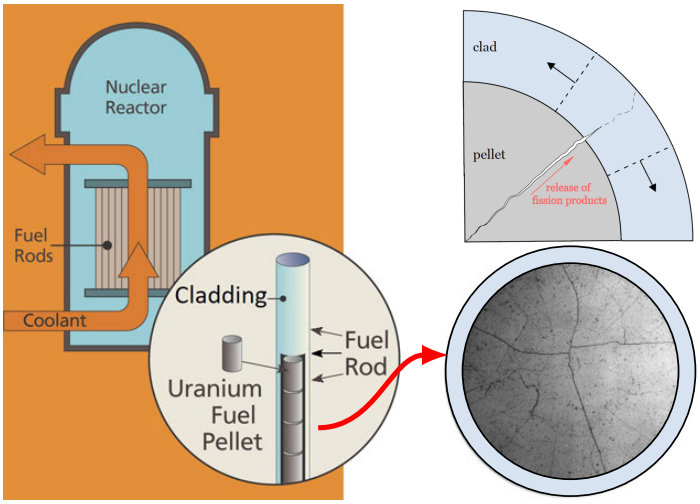


FIGURE 1.2 – Représentation schématique du phénomène de fissuration radiale des pastilles dans le cas d'un espace fermé entre les pastilles (*pellet* en gris) et la gaine (*clad* en bleu clair), figure de [37, 67] modifiée.

Outre cet exemple d'application, les micro-écoulements de gaz peuvent également être observés dans de nombreux dispositifs [18] :

- dans des micro-échangeurs de chaleur conçus pour des applications chimiques ou pour le refroidissement de composants électroniques,
- dans des micro-buses utilisées pour la micro-propulsion de nano- et pico-satellites
- dans des micro-chromatographes de gaz
- des analyseurs ou des séparateurs
- dans des micro-pompes Knudsen
- ...

Cette liste, non exhaustive, montre la diversité et la complexité des écoulements à l'échelle micrométrique. Au-delà des enjeux scientifiques tels que la détermination et la compréhension des phénomènes qui les régissent, leur étude revêt également une importante vocation industrielle pour l'obtention de paramètres macroscopiques pour la conception et l'optimisation de systèmes.

1.2 Modélisation d'écoulements micro/nano échelle

Comme présenté dans la section précédente, la modélisation des écoulements des fluides et des interactions gaz/paroi est cruciale pour de nombreuses technologies innovantes mettant en jeu des micros canaux [1, 18]. Toutefois, cette tâche se révèle être un défi, car, relativement au nombre de collisions entre particules de gaz, celui entre gaz et paroi augmente avec la réduction de l'échelle de longueur caractéristique de l'écoulement du fluide et/ou de la densité du gaz. De manière absolue, la diminution de la densité du gaz entraîne évidemment une diminution du nombre de collisions, ce qui affecte le mouvement des molécules, et qui est communément appelé "effets de non-équilibre".

Les équations de Navier-Stokes, qui sont basées sur la mécanique des milieux continus, supposent que le libre parcours moyen des molécules de gaz λ, (Eq. 1.1) est beaucoup plus petit que l'échelle de longueur caractéristique du système. Cependant, dans quelques applications techniques d'intérêt, cette hypothèse de continuum s'écarte, si l'échelle de longueur du système est de l'ordre du libre parcours moyen de la particule (écoulement de gaz à micro-échelle). L'une des expressions du libre parcours moyen est donnée par Cercignani, (1990) [15] :

$$\lambda = \frac{\mu}{\rho}\sqrt{\frac{\pi r T}{2}} \tag{1.1}$$

où μ est la viscosité, ρ la masse volumique, r la constante spécifique des gaz et T la température.
L'écoulement à travers les dispositifs à l'échelle nano/micro est dominé par des effets de non-équilibre tels que la raréfaction et les interactions molécule-surface du gaz. La formation d'une fine couche de non-équilibre, appelée couche de Knudsen est l'un de ces phénomènes. Elle se forme près de la surface solide dans les écoulements de gaz raréfiés/micro-échelle. En effet, les collisions fluide/paroi sont dominées par la présence d'une surface solide qui réduit le temps moyen entre les collisions, c'est-à-dire que le libre parcours moyen des molécules de gaz se réduit en présence d'une surface solide [18, 90]. Les molécules entrent en collision avec la paroi plus fréquemment qu'avec d'autres molécules, ce qui entraîne la formation de cette couche de Knudsen. Cette couche se réfère au nombre de Knudsen exprimé par :

$$Kn = \frac{\lambda}{\ell_c} \tag{1.2}$$

avec λ le libre parcours moyen moléculaire et ℓ_c une échelle de longueur caractéristique ℓ_c. Pour des conduites, ℓ_c est généralement le diamètre hydraulique.

Par ailleurs, le nombre de Knudsen est souvent utilisé comme un indicateur de non-équilibre.

Quatre différents régimes de raréfaction peuvent être définis en fonction de l'ordre de grandeur de Kn [18, 33, 101]. Ils sont représentés dans la figure 1.3 et décrits ci-dessous :

Le régime continu ($\mathrm{Kn} \leq 10^{-3}$) : l'écoulement est continu, il peut être par exemple modélisé avec précision par les équations compressibles de Navier-Stokes avec des conditions aux limites classiques d'adhérence à la paroi (sans glissement).

Le régime de glissement ($10^{-3} \leq \mathrm{Kn} \leq 10^{-1}$) : l'écoulement est glissant, et les équations de Navier-Stokes restent applicables, à condition de prendre en compte un glissement dynamique. Dans le cas où la thermique est modélisée, un saut de température ainsi qu'un saut de flux de conduction [17, 73] sont introduits aux parois. Ces conditions aux limites traduisent les effets de raréfaction qui deviennent significatifs à la paroi.

Le régime de transition ($10^{-1} \leq \mathrm{Kn} \leq 10$) : l'écoulement est un écoulement de transition, et l'approche continue des équations de Navier-Stokes n'est plus valable. Cependant, les collisions intermoléculaires ne sont pas encore négligeables et doivent être prises en compte.

Le régime moléculaire libre ($\mathrm{Kn} > 10$) : dans ce régime, l'occurrence de collisions intermoléculaires est négligeable par rapport aux collisions entre les molécules de gaz et celles des parois.

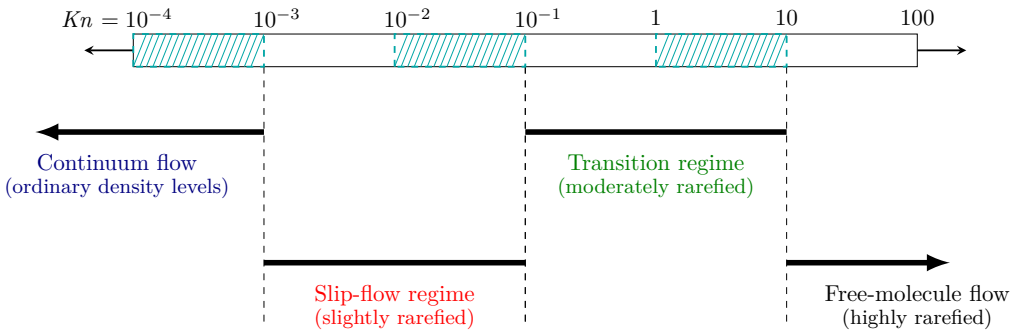


FIGURE 1.3 – Différents régimes de raréfaction GAD-EL-HAK, (2005) [33]

En raison de leurs tailles, dispositifs micro-électro-mécanique (MEMS) se positionnent généralement dans les régimes de glissement et de transition, pour des pression et température standard [14, 50]. Dans le cadre de ce travail, les régimes cibles seront restreints à $\mathcal{O}(Kn) < 10^{-1}$

Des expériences récemment menées, par exemple, sur des écoulements de revêtement et sur l'impact de gouttes de liquide (travaux de SPRITTLES, (2015) [89]) ont également démontré que les effets de non-équilibre dans le gaz peuvent expliquer les problèmes de mouillage causés par l'entraînement d'air. Ces problèmes peuvent être supprimés en réduisant la pression du gaz ambiant. Les réductions de la pression ambiante augmentant le libre parcours moyen des molécules de gaz ((Eq. 1.3)), et par conséquent le nombre de Knudsen. En supposant la relation des gaz parfaits, $p = \rho r T$, l'équation (1.1) devient

$$\lambda = \frac{\mu}{p} r T \sqrt{\frac{\pi r T}{2}} \tag{1.3}$$

Les questions auxquelles il faudra répondre portent sur les modèles et méthodes à utiliser pour résoudre et comprendre correctement le comportement des écoulements de fluides dans les micro-dispositifs. Il est également nécessaire de s'interroger sur les conditions aux limites à appliquer et sur la manière dont il faut procéder pour obtenir des solutions au problème.

1.2.1 Modélisation approchée d'écoulements micro/nano échelle

Dans les écoulements gazeux, les conditions d'adhérence aux frontières ne sont applicables que pour des valeurs de $Kn < 10^{-3}$ [33, 50]. A contrario, lorsque l'échelle de longueur caractéristique de l'écoulement est réduite à environ 10 fois le libre parcours moyen des molécules de gaz, c'est-à-dire pour $Kn \sim 10^{-1}$, les équations conventionnelles de la mécanique des fluides, présentés dans le chapitre 2, restent valides dans le cœur de l'écoulement, mais une attention particulière doit être accordée à la modélisation des conditions limites, fluide/paroi.

Dans toute cette gamme de Knudsen, les relations de continuité de la vitesse, de la température et du flux de conduction thermique entre le gaz et le solide ne s'appliquent plus aux micros/nano échelles du fait du non-équilibre thermodynamique entre les molécules de gaz dans la couche de Knudsen.

Pour les liquides, le libre parcours moyen λ est plus petit et la description avec un modèle continu et des conditions d'adhérence peut s'appliquer à l'échelle microscopique, pour les longueurs caractéristiques de l'ordre de 100 nm. Au-delà, comme pour les gaz, il faut prendre en compte les discontinuités de vitesse, température et du flux de conduction au niveau des interfaces liquide/solide.

Il est possible de représenter ces discontinuités de vitesse, température et de flux de conduction, en introduisant les concepts de *glissement dynamique*, et de *saut de température et de flux de conduction*, que nombreux auteurs comme Maslen, (1958) [63], Nicolas et al., (2019) [73] utilisent. Cela introduit un déphasage entre les vitesses tangentielles du fluide et de la surface solide à leur interface tels les modèles continus proposés indépendamment par Navier, (1823) [71] et plus tard par Maxwell, (1879) [65].

Du point de vue mathématique, ces modèles reviennent à remplacer la condition de non-glissement, par des relations de saut (Eq. 1.4) consistant par exemple à introduire une proportionnalité entre la vitesse de glissement et le cisaillement à l'interface solide, (Fig. 1.4). Des détails supplémentaires sont fournis dans le chapitre 2, section 2.1.2.2. On note ici, pour la partie dynamique,

$$u - u_w = L_s \left.\frac{\partial u}{\partial \bar{n}}\right|_w \tag{1.4}$$

où u est la vitesse tangentielle de glissement, u_w la vitesse de la paroi (généralement nulle) et $\left.\frac{\partial u}{\partial \bar{n}}\right|_w$ est le cisaillement à la paroi. Le paramètre L_s est appelé "longueur de glissement dynamique", et représente le rapport entre la vitesse et le cisaillement à la paroi (Fig. 1.4).

Pour les effets thermiques, le saut de température est donné par la résistance de Kapitza, (1971) [49], utilisé et modifié récemment par Kandlikar et al., (2013) [48], Nicolas et al.,

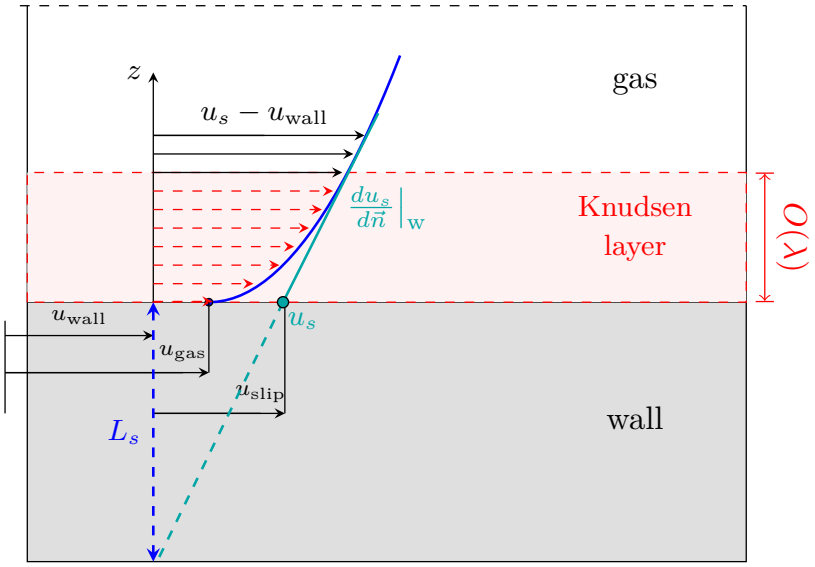


FIGURE 1.4 – Configuration d'écoulement entre deux plaques plates, montrant les conditions aux limites de glissement dynamique.

(2019) [73]. Ce saut s'exprime par,

$$T - T_w = L_T \left.\frac{\partial T}{\partial \bar{n}}\right|_w \tag{1.5}$$

où L_T est la longueur de glissement thermique. Ces deux longueurs (L_s et L_T) dépendent des paramètres à petite échelle, par exemple des interactions entre les atomes de gaz et les particules solides de la paroi.
La modélisation du saut de flux de conduction est proposé par MASLEN, (1958) [63]. Elle est donnée par,

$$k_f \nabla T - k_s \nabla T = -\mu u_\tau \nabla u_\tau \tag{1.6}$$

où k_f, k_s sont respectivement les conductivités du fluide et du solide. Le terme de droite de l'équation (1.6) représente l'énergie dissipée par frottements visqueux.

À ces échelles micrométriques, les effets de surface sur le volume dépassent les effets inertiels et les effets de glissement doivent être pris en compte [68, 88]. À ces échelles réduites, le fluide requiert une modélisation au niveau atomique par la dynamique moléculaire, [3, 81], et l'hypothèse relative au milieu continu est remplacée par une représentation discrète du fluide par le biais de particules en interaction.

1.2.2 Modélisation moléculaire d'écoulements micro/nano échelle

Fondamentalement, la condition réelle à imposer aux variables de l'écoulement à l'interface fluide/solide serait plus naturellement formulée en termes d'échange de quantité de mouvement et d'énergie entre les éléments adjacents du fluide et de la paroi, étant ainsi déterminé par l'état moléculaire de la région de contact. Les informations nécessaires pour spécifier ce type de condition aux limites est hors de portée du continu, mentionné par KARNIADAKIS et al., (2005) [50]. Elles ne sont accessibles que si la structure de l'interface est décrite à un niveau moléculaire/atomique [17, 98]. Il est alors nécessaire d'utiliser des méthodes adéquates.

Ces méthodes moléculaires permettent de prendre en considération tous les phénomènes présents à petite échelle, en regardant les interactions entre les molécules de gaz et de la paroi.

Les simulations de dynamique moléculaire [3, 81] sont en mesure de modéliser les nano/micro-écoulements de canaux dans les canaux des MEMS [18, 33, 50]. Les autres applications comprennent également la coalescence [54], l'étalement [103] et le mouillage [104], ainsi que la modélisation de l'écoulement dans les régions présentant des gradients élevés de vitesse et de densité [7]. Dans l'étude des fluides denses, les simulations de dynamique moléculaire se sont révélées en outre être un outil précieux. Ces techniques ont été employées dans l'étude des polymères liquides et colloïdaux, ainsi que dans la compréhension de phénomènes hydrodynamiques microscopiques complexes [7, 54, 98].

Dans le cas de microcanaux de grande longueur qui peuvent aller de 1 mm à 1 m, une modélisation complète de la paroi par une approche microscopique ne peut être envisageable. Il convient alors d'adopter une approche multi-échelle, combinant les approches moléculaire et macroscopique, afin de modéliser convenablement le cœur de l'écoulement par la mécanique des milieux continus et les régions dans lesquelles les effets de raréfaction sont présents.

1.2.3 Modélisation multi-échelle d'écoulements micro/nano échelle

Dans la section précédente, il était question de présenter l'intérêt d'utiliser des méthodes discrètes pour modéliser les écoulements dans des systèmes de taille micrométrique. Toutefois, l'extension de ces conduites rend donc l'étude plus délicate en raison du grand nombre de simulations à mettre en place.

Cette difficulté peut alors être surmontée en couplant une approche discrète avec une approche continue, permettant ainsi de bénéficier des avantages des deux approches. En effet, les discontinuités, au contact fluide/solide et liés aux effets de raréfaction, restent localisées dans la zone de proche paroi, nécessitant ainsi une description discrète du comportement des fluides, tandis que le reste de la conduite peut être traité par une approche continue et les informations extraites de cette approche peuvent être utilisées pour paramétrer l'approche discrète. La notion générale de ces méthodes hybrides est de coupler différentes descriptions macro et micro pour mieux modéliser des phénomènes sur des échelles de temps et de longueur variées, et ce, à des coûts moindres.

Dans la section suivante, un état de l'art des différentes méthodes de couplage utilisées dans le domaine de la mécanique des fluides sera présenté, avant de présenter la résolution du problème couplé ainsi que la mise en œuvre des approches couplées retenues.

L'approche couplant la modélisation "du continu", des équations de la dynamique des fluides, et la description discrète qui modélise l'écoulement à l'échelle moléculaire, a été développé dans le domaine de la physique pour la première fois pour l'étude du comportement thermique et mécanique des solides à l'échelle atomique par Kohlhoff et al., (1991) [52], dans lequel il a couplé entre la dynamique moléculaire et la méthode des éléments finis. Broughton

et al., (1999) [9], en s'appuyant sur ces travaux, ont présenté des validations de cette méthode pour l'étude de la propagation de fissures dans le silicium. Ensuite, ce couplage a été développé pour des physiques plus complexes, l'étude du comportement dynamique des fluides par O'Connell et Thompson, (1995) [78]. Depuis lors, elle suscite de plus en plus d'intérêt en tant que solution permettant de décrire les phénomènes d'écoulement à l'échelle du micro/nano mètre.

Les méthodes hybrides, couplant l'échelle atomistique et le continu, ont été utilisées dans des problèmes microfluidiques majeurs [68], notamment pour étudier le mouvement de l'interface bi-fluides (Hadjiconstantinou, (1999) [39]), les effets de glissement aux interfaces fluide-paroi (O'Connell et Thompson, (1995) [78], Ritos et al., (2015) [84]), les interfaces gaz/liquide (Bugel, (2009) [10]), les phénomènes de condensation (Sun et al., (2009) [91], Vu et al., (2018) [98]), et les microsystèmes complexes ([7, 60, 83]), et très récemment Lācis et al., (2022) [55] qui étudie des gouttelettes cisaillées à l'échelle nanométrique avec un faible glissement.

Par rapport à la méthodologie classique de la mécanique des milieux continus seule ou de la dynamique moléculaire seule, cette approche fournit des solutions précises près de la limite de la paroi dans un coût de calcul acceptable.

1.3 Synthèse des méthodes hybrides et multi-échelles

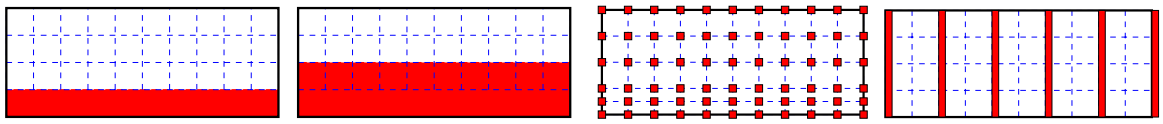
Une synthèse étendue des méthodes originales de couplage entre le discret et le continu, et des développements qui ont suivi, a été établie par Anciaux, (2007) [5], Bugel, (2009) [10], Mohamed et Mohamad, (2010) [68], Ritos et al., (2015) [84], Vu, (2016) [97] ... Ils abordent la décomposition des sous-domaines et la gestion de la transmission des informations entre un modèle qui résout de manière plus précise, au niveau discret et qui contient donc les informations les plus précises, et un modèle macroscopique qui représente une solution plus grossière. Le grand enjeu de ces méthodes hybrides est le raccordement approprié entre les descriptions très distinctes du fluide à l'interface qui sépare le discret et le continu. En effet, les avantages des deux méthodes sont conservés si le raccordement est également correctement réalisé.

Les méthodes, hybrides et multi-échelle, évoquées peuvent être présentées sous quatre grandes philosophies pour traiter des écoulements dans des micro-conduites, voir la figure 1.5 :

1. les méthodes de couplage sans recouvrement où les sous-domaines sont disjoints (CGMD : Coarse-grained molecular dynamics and the atomic limit of finite elements) par Kohlhoff et al., (1991) [52] et [85-87],
2. les méthodes avec recouvrement où les deux approches se superposent dans une zone d'échange d'informations (DDM : Domain Decomposition Method) proposée par O'Connell et Thompson, (1995) [78] et [5, 19, 103],

3. les méthodes hétérogènes et multi-échelle (HMM : Heterogenous Multiscale Method) développée par Ren et E, (2005) [83], dans laquelle des micro-simulations sont réparties dans tout le domaine,

4. couplage avec recouvrement par micro-domaines discrétisés de manière régulière le long du domaine (IMM : Internal-flow Multiscale Method), dans laquelle seulement certaine partie sont superposées, où ces micro-domaines s'y trouvent par Borg et al., (2013) [8] et [7].



(a) CGMD [85-87] (b) DDM [5, 78] (c) HMM [83] (d) IMM [8]

Figure 1.5 – Configurations décrivant les différentes méthodes, hybrides et multi-échelle, existantes, ainsi que les diverses manières de couplage et de transmission des informations. Les zones en rouge représentent les simulations moléculaires, et les zones blanches maillées en bleu le domaine continu.

Méthode CGMD, coarse-grained molecular dynamics, sans recouvrement

En s'appuyant sur les travaux de Kohlhoff et al., (1991) [52], Rudd et Broughton, (1998) [86] ont proposé la première approche. Il s'agit de la technique de la dynamique moléculaire à gros grains, appelée CGMD (Fig. 1.6), qui est un modèle multi-échelle qui permet de coupler la dynamique moléculaire conventionnelle à une description à plus gros grains de la frontière. Le couplage dans cette technique se fait directement entre les deux régions : dynamique mo-

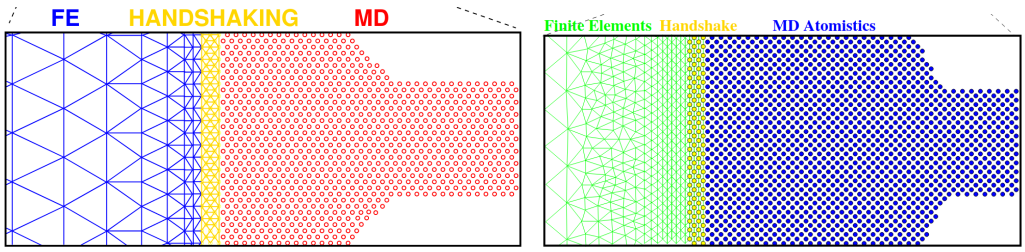


Figure 1.6 – CGMD Sans recouvrement [85-87]. À droite : le maillage bleu correspondant aux éléments finis est raffiné (la zone jaune) de sorte à avoir une taille similaire au diamètre de la particule de dynamique moléculaire (zone rouge). À gauche : les grains de dynamique moléculaire en bleu sont de plus en gros (dans la zone jaune) pour que leurs tailles soient équivalentes au maillage éléments finis en vert.

léculaire, et les régions à gros grains modélisées sur un maillage dans une formulation qui généralise la modélisation conventionnelle par éléments finis de l'élasticité en milieu continu. La particularité de cette méthode est qu'elle est dérivée uniquement du modèle dynamique moléculaire et ne comporte aucun paramètre du milieu continu. La stratégie est de remplacer les éléments finis conventionnels adaptés à un maillage de taille atomique dans certaines régions [87]. La procédure est basée sur une prescription statistique à grain grossier. Cela

permet aux régions modélisées en dynamique moléculaire d'être couplées de manière transparente aux régions de la FEM généralisée, en apportant toute la puissance de la dynamique moléculaire sur les parties importantes du système, sans la surcharge de calcul de la dynamique moléculaire dans d'autres régions importantes, mais physiquement moins complexes.

Cette simulation multi-échelle pose un certain nombre de difficultés, citées par [87] :

- la réponse physique du modèle à des champs de déformation stationnaires, variant lentement en position, qui s'étendent dans la région à gros grains, doit être dépourvue de forces fantômes ;
- le système doit avoir une dynamique réaliste, exempte de toute forme de dégradation due à des influences dans la région centrale de la dynamique moléculaire qui se propagent vers des interfaces non physiques ;
- sur le plan de la thermodynamique, il convient que ce modèle présente des réponses sensibles aux comportements hors équilibre.

Méthode DDM, Domain Decomposition Method, avec recouvrement

En outre, le couplage avec la superposition (Fig. 1.7) retient notre attention. La jonction spatiale entre les deux approches couplées se fait à travers une zone de recouvrement, permettant ainsi un transfert d'informations du discret vers le continu moins brutal que dans le cas d'un couplage sans recouvrement. Cette approche permet de ne pas lier de façon rigide les atomes discrets au maillage continu.
Les auteurs, Cosden et Lukes, (2013) [19], Nie et al., (2004) [75], O'Connell et Thompson,

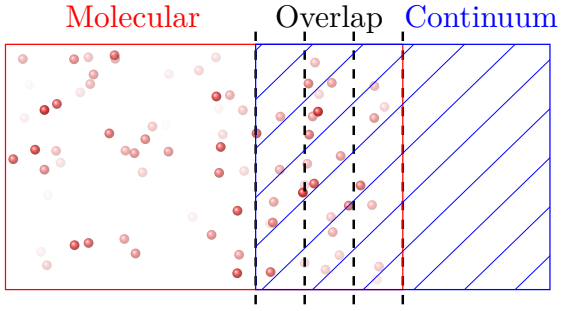


Figure 1.7 – DDM avec recouvrement O'Connell et Thompson, (1995) [78] et [19, 75, 84, 98, 99, 102, 103]. La ligne en pointillés délimite les intervalles utilisés pour calculer la moyenne des vitesses des particules dans la zone de recouvrement (overlap).

(1995) [78], Ritos et al., (2015) [84], Vu et al., (2018) [98], Vu et al., (2016) [99], Yen et al., (2007) [102], Zhang et al., (2017) [103], proposent une approche ou les domaines, discret et continu, se superposent dans une région de l'espace. Dans cette approche, le domaine fluide est subdivisé en sous-domaines et chaque sous-domaine est résolu soit par le solveur continu, soit par le solveur discret basé sur les molécules [10].

Le cœur de l'écoulement macroscopique où l'hypothèse de la mécanique des milieux continus est valide est résolue par la formulation du continu en utilisant une discrétisation appropriée. La zone contenant l'interface (fluide/solide ou fluide/fluide) ainsi que la zone proche paroi

sont analysées par une formulation moléculaire. En échangeant les quantités macroscopiques à travers la zone de recouvrement, les informations peuvent être synchronisées en temps et en espace. La continuité des flux est assurée si les coefficients de transport et de diffusion sont identiques entre les deux approches au niveau de la zone de recouvrement.

Concrètement, dans cette méthode, le domaine de simulation est décomposé en trois régions (Fig. 1.7) :

- le volume (cœur de l'écoulement) où les équations de mécanique des milieux continus (par exemple les équations de Navier-Stokes et d'énergie) sont résolues,
- le voisinage de la paroi et la paroi sont simulés par la dynamique moléculaire,
- la zone de recouvrement où les deux approches se superposent et dans laquelle s'effectue le transfert des informations, vitesse u, température T et masse volumique ρ etc, entre les deux premières régions.

Plusieurs études ont été publiées pour améliorer la technique hybride et/ou appliquer cette approche à divers secteurs d'écoulement à l'échelle micro/nanométrique. Ces recherches peuvent être classées en trois catégories :

- dynamique lagrangienne contrainte (Liu et al., (2008) [59], Nie et al., (2004) [75], O'Connell et Thompson, (1995) [77], Vu et al., (2018) [98], Yen et al., (2007) [102]),
- méthode de Schwarz alternée (Bugel, (2009) [10], Hadjiconstantinou, (1999) [39], Hadjiconstantinou et al., (2003) [40], T. Werder, (2005) [93])
- et échange direct de flux (Delgado-Buscalioni et Coveney, (2003) [20], Flekkøy et al., (2000) [29], Giupponi et al., (2007) [36]), selon la formulation des schémas hybrides et les caractéristiques des variables échangées dans la région de recouvrement.

Parmi ces méthodes, la dynamique lagrangienne consiste à contraindre les équations du mouvement afin de respecter la condition fournie par le domaine continu. Les solveurs de la méthode continue et moléculaire échangent des propriétés ou des variables, à savoir les composantes du vecteur vitesse $\bar{v}$, la température et la densité, ρ.

HMM, Heterogeneous Multiscale Method

Ren et E, (2005) [83] ont développé les méthodes hétérogènes multi-échelle (HMM), figure 1.9, dans laquelle l'objectif est de recourir à la micro-résolution au lieu de la décomposition micro-macro. En d'autres termes, le domaine est entièrement décrit par le macro-modèle avec de petits micro-modèles isolés répartis dans l'espace dans le domaine, qui fournissent, au niveau local, les informations manquantes à la solution macro.
La démarche générale à suivre pour la résolution des problèmes dynamiques au moyen des méthodes HMM est la suivante : Pour un champ de vitesse à l'instant t (macro-variable $u(t)$ connue), le calcul de la macro-variable à l'instant pas de temps $t + \Delta t$ se fait en estimant les données nécessaires dans le solveur macro, par exemple le flux de quantité de mouvement, à partir du micro-modèle tel que la dynamique moléculaire. Pour que le micro soit cohérent avec le macro local, certaines contraintes sont imposées dans le modèle micro, généralement

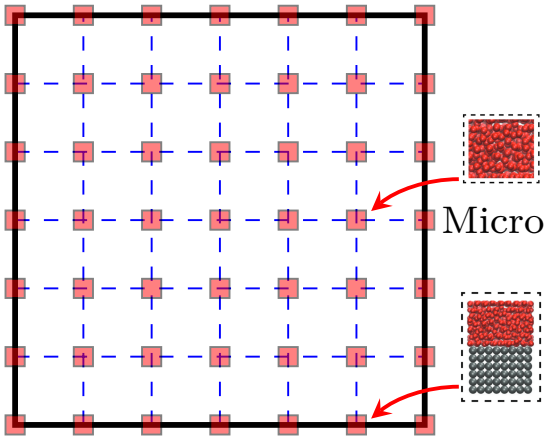


FIGURE 1.8 – HMM par REN et E, (2005) [83]. Les micro-simulations en rouge sont réparties dans tout le domaine. Les différentes configurations des micro-simulations sont présentées à droite de la figure.

par l'imposition de conditions aux limites. Les données microscopiques sont ensuite traitées à l'aide d'une moyenne d'ensemble, spatiale et temporelle, afin de fournir les données nécessaires à l'échelle macroscopique. Ces données sont finalement transmises au solveur macro et $u(t + \Delta t)$ est calculée.

IMM, Internal-flow Multiscale Method

BORG et al., (2013) [8] ont développé la méthode hybride et multi-échelle, appelée (IMM : Internal-flow Multiscale Method), figure 1.9, couplant entre les équations de Navier-Stokes-Fourier (NSF) et la dynamique moléculaire, pour résoudre des problèmes d'écoulements isothermes, incompressibles et stationnaires dans un canal nanométrique convergent/divergent axialement périodique.
En partant du principe des méthodes hétérogènes multi-échelle proposée par REN et E, (2005) [83], ils proposent alors une alternative, qui consiste à répartir les micro-sous-domaines dans de très petites tranches du canal.

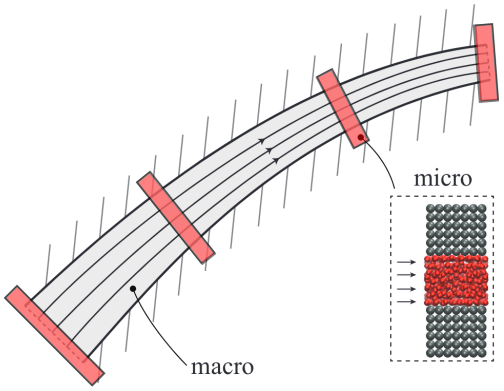


FIGURE 1.9 – IMM par BORG et al., (2013) [8]. Les micro-simulations en rouge sont réparties tout au long du canal.

Le couplage se fait par l'imposition itérative de la conservation de la masse et de la quantité

de mouvement à l'échelle macro, il n'y a aucune communication directe entre les micro-sous-domaines. La simplification majeure par rapport aux autres méthodes hybrides est que les micro-sous-domaines remplacent la condition limite de glissement ou les relations contrainte-déformation dans la description et que les simulations, dynamique moléculaire, utilisent simplement des conditions périodiques. Ce travail a été étendu et déployé pour résoudre cette fois-ci des écoulements compressibles et instationnaires de gaz confiné dans des géométries nanométriques à fort rapport d'aspect (unsteady-IMM) Borg et al., (2015) [7].

La première motivation de la méthode (IMM : Internal-flow Multiscale Method) [7, 8] était de résoudre des problèmes d'écoulements dans des canaux en série à fort rapport d'aspect, en particulier lorsque la séparation d'échelle de longueur (entre les effets moléculaires et le continu) est présente dans la direction du flux, sans pour autant qu'elle le soit dans la direction transversale.

SeN-IMM, Serial-Network Internal-flow Multiscale Method

Pour prendre en considération ces difficultés et pour mieux modéliser les pertes de débit et de pression à travers des membranes, Ritos et al., (2015) [84] ont poursuivi les travaux de [7, 8] et ont développé une méthode multi-échelle (serial-network internal-flow multiscale method (SeN-IMM)), où ils traitent le cas de conduite nanométrique de longueur infinie (Fig. 1.10). Dans cette méthode, les entrées et sorties et les régions proches de ces zones sont modélisées par dynamique moléculaire et le tube est modélisé en utilisant la méthode (IMM). Le modèle SeN-IMM a été utilisé pour calculer flux massique et le renforcement dans une membrane de 1 mm d'épaisseur, reproduisant des expériences. Les résultats par ce modèle ont été comparés à ceux obtenus par dynamique moléculaire seule, pour un cas de poiseuille et également avec les résultats expérimentaux. Les résultats étaient en bonne concordance, compte tenu des hypothèses et des erreurs commises au cours des processus expérimental et de calcul.

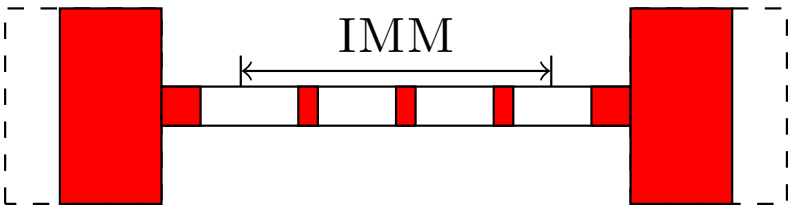


Figure 1.10 – Méthode multi-échelle de Ritos et al., (2015) [84], serial-network internal-flow multiscale method (SeN-IMM), pour le traitement des nanotubes. Les zones rouges correspondent à la dynamique moléculaire et les zones blanches aux zones résolues par le formalisme continu.

Méthode hybride des écoulements de gaz raréfiés

Docherty et al., (2014) [23] propose une méthode de calcul hybride pour simuler des écoulements de gaz raréfiés non-isothermes à l'échelle microscopique en s'appuyant sur les travaux de Borg et al., (2013) [8]. Leur stratégie consiste à coupler entre une description du fluide au

niveau continu, en utilisant les équations de la mécanique des milieux continus et un solveur de simulation directe Monte Carlo, au niveau discret. La méthodologie de départ considère le transfert de chaleur uniquement, avec une validation effectuée sur un écoulement micro Fourier $1D$.

Par la suite, dans un second travail [22], la stratégie est étendue pour prendre en compte le transport de masse, de quantité de mouvement et de chaleur, et des validations en $1D$ ont été effectuées sur le problème de l'écoulement de *micro-Couette* à grande vitesse. La résolution par de la méthode hybride permet un bon accord et a été équivalente à une simulation Monte Carlo pure. Cependant, cette méthode n'offre aucune amélioration en temps de calcul pour le problème $1D$ considéré.
Néanmoins, une amélioration significative est obtenue pour un problème $2D$. Ce dernier cas consiste à l'étude d'un écoulement de gaz à travers une micro-fissure, modélisée comme un micro-canal avec une section transversale à haut rapport d'aspect. Une différence de température a été imposée entre les parois de la section transversale, la méthode hybride a permis de prédire très précisément la variation de la vitesse et de la température, sur la section transversale, ainsi que du débit massique.

Méthode hybride atomique-continu

Vu et al., (2018) [98], Vu et al., (2016) [99] ont développé une méthode hybride atomique-continu dédiée à l'étude des problèmes multi-échelles dans des nano/micro canaux, en partant du principe de la méthode de décomposition du domaine proposée par O'Connell et Thompson, (1995) [78], voir (Fig. 1.7).
La simulation d'une paroi entière d'un millimètre par dynamique moléculaire, extrêmement coûteuse en termes de temps de calcul. Pour cela, dans cette méthode hybride atomique-continu [98], pour la simulation de nano/micro canaux de grandes extensions, la paroi et la zone qui se trouve aux voisinages de la paroi n'est pas entièrement décrite par des micro-simulations. Elle est alors substituée par quelques dizaines de micro-simulations de domaines moléculaires de longueur nanométrique.
En effet, des micro-sous-domaines sont adoptés et discrétisés le long de la direction de l'écoulement, voir (Fig. 1.11), ce qui permet la capture des variations de vitesse et de température de l'entrée à la sortie. Par conséquent, il est possible de simuler des micro-canaux de grande extension à un moindre coût. La communication entre les micro-sous-domaines se fait par interpolation.

La zone de recouvrement est subdivisée en quatre couches plus petites :

- la couche d'échange de variables comme la vitesse, température et densité du microscopique au continu,
- la couche de relaxation, où il n'y a pas d'échange d'informations.
- la couche transfert d'information entre le continu et dynamique moléculaire,

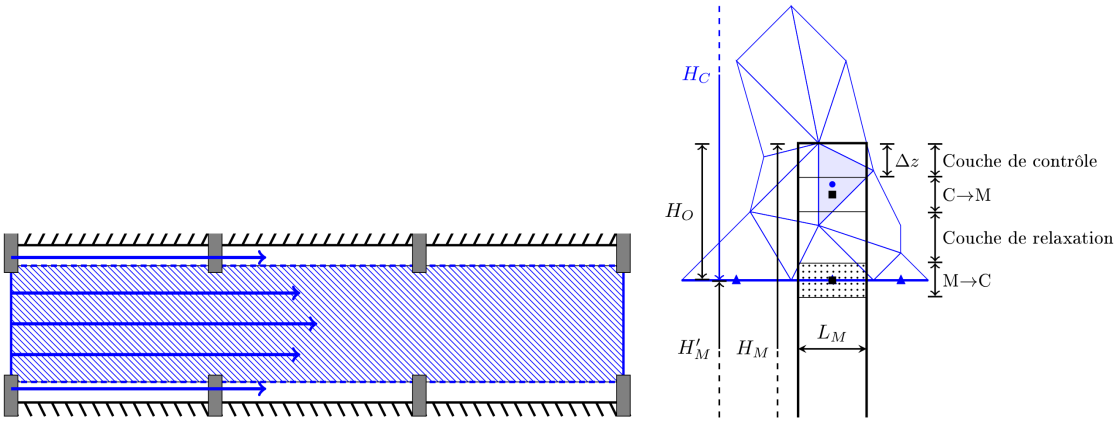


FIGURE 1.11 – Méthode multi-échelle de VU et al., (2016) [99], Méthode hybride atomique-continu, pour le traitement des nano/micro canaux. La région hachurée en bleu représente le domaine continu, les rectangles gris pour les plots de micro-simulations.

— et la couche de contrôle, son extrémité constitue la limite supérieure de la région moléculaire, permet d'atténuer les effets de bord et, éventuellement, de gérer la valeur de la densité locale.

La validité de la méthode hybride est démontrée par des comparaisons avec des solutions analytiques et des simulations par volumes finis. Cette méthode est ensuite appliquée à l'étude du développement hydrodynamique et thermique d'un écoulement liquide dans un long micro/nano canal. Ultérieurement [98], cette méthode est étendue pour des applications aux problèmes de compression et de condensation. La méthode, proposée par VU et al., (2016) [99], porte une attention particulière au contrôle de la densité dans les micro-domaines.

Méthodes multi-échelle : problèmes de mouillage

Les méthodes hybrides et multi-échelles ont été également utilisées dans des problèmes de mouillage [55, 58, 103, 104] et de coalescence [21] pour résoudre les phénomènes apparaissant dans la région proche du point de contact. Une des difficultés de la théorie du milieu continu est la représentation efficace du mouvement moléculaire à proximité de la ligne de contact. ZHANG et al., (2017) [103], ZHANG et XU, (2022) [104] ont proposé une stratégie multi-échelle qui combine la dynamique moléculaire avec des simulations de Navier-Stokes, *volume of fluid* (MD-VOF), qui est une variante de la méthode multi-échelle hétérogène (HMM) [83], pour étudier l'étalement de gouttelettes sur des surfaces solides, voir 1.12.

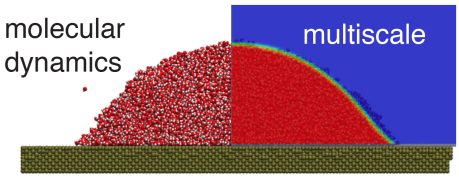


FIGURE 1.12 – Couplage faible ZHANG et al., (2017) [103]

Tout comme le principe de la méthode proposée par Vu et al., (2016) [99], les simulations de dynamique moléculaire sont distribuées le long de l'interface mouillée, fournissant les informations cruciales sur les frontières pour la dynamique de la ligne de contact triphasique et le glissement interfacial solide/liquide, qui sont transmises à la VOF qui elle est appliquée partout dans le domaine. Afin de calibrer le modèle de glissement de Navier, (1823) [71], des simulations d'écoulement de cisaillement moléculaire du liquide en contact avec le solide sont utilisées pour mesurer la longueur de glissement locale L_s. D'autres simulations de dynamique moléculaire de nano-gouttelettes sont exploitées afin d'avoir les informations sur l'angle de contacte et le glissement à cet endroit.

Ces simulations hybrides de Zhang et al., (2017) [103] ont été comparées avec la dynamique moléculaire pure. Les auteurs démontrent que pour les nano-gouttelettes qui se répandent sur des surfaces, la prise en compte d'un modèle d'angle de contact dynamique n'est pas suffisante pour capter les effets moléculaires ; il est impératif de considérer le glissement sur l'ensemble de l'interface solide/liquide, notamment le comportement de glissement important au niveau de la ligne de contact.

Très récemment, Liu et al., (2021) [58] ont utilisé la stratégie de simulation multi-échelle, (MD-VOF), qui s'est avérée capable de simuler le mouillage dynamique au-dessus de l'échelle nanométrique, où les simulations de dynamique moléculaire seules sont inaccessibles en raison du coût de calcul.

Dans cette méthode, deux sortes de conditions aux limites de glissement sont incluses :

- i) un modèle de glissement particulier basé sur la théorie de la cinétique moléculaire pour la ligne de contact triphasique afin de tenir compte de la physique moléculaire interfaciale,
- et ii) le modèle de glissement classique de Navier, (1823) [71], pour la partie restante de l'interface liquide-solide, et un nouveau terme source a été ajouté à l'équation de la quantité de mouvement dans VOF pour remplacer le modèle d'angle de contact dynamique par convection. Chaque modèle de glissement a été calibré par les simulations de dynamique moléculaire.

De même, Lācis et al., (2022) [55] ont développé une méthode multi-échelle qui permet la simulation des phénomènes de mouillage dans le cas des gouttelettes cisaillées, où le mouvement de la ligne de contact triphasique entre deux fluides non miscibles et une surface solide se manifeste. La dynamique moléculaire de la gouttelette cisaillée utilisée près de la ligne de contact permet de mieux comprendre la friction linéaire subie par les lignes de contact.

1.4 Plan de la thèse

Dans toutes les circonstances, et indépendamment de la méthode utilisée, la précision et le faible coût restent un élément clé dans la simulation des écoulements à petite échelle.

Une bonne description des écoulements dans des systèmes de taille micrométrique et des phénomènes qui se produisent à cette échelle est essentielle.
Par conséquent, pour décrire pleinement la physique dans ce type d'écoulements, des méthodes basées sur une approche hybride et multi-échelle sont mises en œuvre pour relier avec précision la solution proche de la paroi à l'écoulement global. Dans cette méthode multi-échelle, des simulations à l'échelle moléculaire sont utilisées près des parois solides et une méthode de volumes finis pour les équations de Navier-Stokes est appliquée ailleurs pour simuler l'écoulement global.

L'originalité de ce travail se tient dans le fait que le couplage est fait en utilisant une méthode d'apprentissage, dans laquelle la vitesse et le taux de cisaillement, sont fournis à une distance de l'ordre de quelques libres parcours moyens des molécules de gaz. C'est dans ce cadre que s'inscrit ce travail de recherche dont la finalité est le développent d'une méthode hybride et multi-échelle originale et l'établissement d'une base de données permettant la caractérisation et l'identification automatique des grandeurs (vitesse et cisaillement) aux interfaces et proches parois des micro-systèmes.

Pour y parvenir, nous avons structuré notre travail comme suit : Le développement d'une méthode hybride et multi-échelle repose alors sur le couplage entre deux approches :

- i) macroscopique, en utilisant les équations de Navier-Stokes, discrétisé par la méthode des volumes finis [79],
- et ii) microscopique, en résolvant molécule par molécule en utilisant la dynamique moléculaire [3, 81].

Ainsi, le chapitre 2 est dédié à la présentation des principaux éléments et modélisation sur lesquels ces méthodes hybrides sont basées. Les modèles physiques et numériques sont alors présentés dans ce chapitre.

Ensuite, des résultats obtenus en utilisant la dynamique moléculaire pour identifier les conditions aux limites thermiques dans les micros/nano conduites sont exposés dans le chapitre 3. Ces résultats ont fait l'objet d'un article dans *International Journal of Thermal Science*.
Notre choix s'est porté donc sur l'utilisation des méthodes d'apprentissage supervisé pour coupler les deux approches. Le chapitre 4 est alors consacré à l'introduction et à l'étude des différentes méthodes de régressions, à savoir les méthodes de régression Bayésienne et par processus Gaussiens. Des détails sur la manipulation de ces méthodes de régressions ainsi que des résultats de validations sur des cas physiques sont présentés également dans ce chapitre 4.

Ainsi vient le couplage entre ces chapitres pour construire la méthode hybride et multi-échelle. Dans le chapitre 5, les différentes méthodes de couplage développées dans cette thèse sont présentées, ainsi que des résultats de validation dans le cadre des simulations des écoulements glissants dans des nano/micro canaux de grandes extensions. Les validations sont démontrées d'abord par des comparaisons avec des solutions analytiques et des simulations par volumes finis, toujours au biais des méthodes d'apprentissage supervisé. Ensuite, la méthode implicite

est appliquée aux écoulements droits et courbée en utilisant des simulations de dynamique moléculaire.

Une théorie ne ressemble pas plus à un fait
qu'une photographie ne ressemble à son modèle.

Edgar Watson Howe

Chapitre 2

Modèles de la méthode hybride et multi-échelle

Sommaire

2.1 Modèles physiques 36
2.1.1 Modèle moléculaire 36
2.1.1.1 Énergies et potentiel d'interaction 37
2.1.1.2 Propriétés thermodynamiques 40
2.1.2 Modèle continu 41
2.1.2.1 Modélisation des écoulements de fluide 41
2.1.2.2 Modélisation des conditions aux limites de glissement . . . 42
2.2 Méthodes numériques 43
2.2.1 Solveur micro 43
2.2.1.1 Intégration des équations de mouvement 43
2.2.1.2 Condition aux limites périodiques 45
2.2.1.3 Rayon de coupure 46
2.2.1.4 Modèles de murs 47
2.2.1.5 Méthodes de thermalisation des murs 48
2.2.1.6 Calcul de moyenne et estimation des incertitudes 48
2.2.1.7 Écoulement de couette avec des murs stochastiques 51
2.2.1.8 Écoulement de couette avec des murs atomiques 52
2.2.2 Solveur macro 54
2.2.2.1 Discrétisation des équations de Navier-Stokes 54

Préambule

Ce chapitre a pour but d'exposer les descriptions physiques et numériques des différents outils et techniques qui constituent les éléments de la méthode de couplage développée dans cette thèse. Des résultats des cas test sont présentés à la fin de ce chapitre.

Pour des raisons de simplifications et de répétitions, et par abus de langage, le terme générique de "particule" sera parfois employé pour désigner les atomes ou molécules constituants à l'échelle moléculaire les fluides ou solides traités. Au même titre, pour la masse volumique, nous utiliserons souvent le terme densité. Les termes de paroi ou de mur sont utilisés pour désigner le réseau cristallin des atomes qui composent le solide.

2.1 Modèles physiques

2.1.1 Modèle moléculaire

Cette section présente les principaux éléments de la méthode de dynamique moléculaire classique. C'est une méthode de simulation standard en physique statistique permettant de calculer, à partir de la loi de Newton, les trajectoires de particules (atomes ou molécules). Leurs interactions sont déterminées à partir de forces issues de potentiels d'interactions. De nombreuses quantités macroscopiques peuvent être extraites à partir des différentes positions, vitesse et accélérations prises par les particules au cours du temps.

Dans le contexte de la mécanique des fluides, la dynamique moléculaire est par exemple mise en œuvre pour le calcul de propriétés effectives des fluides, par exemple leurs coefficients de diffusion, de viscosité ou entre de conductivité. Des propriétés plus globales, comme des flux de masse ou de chaleurs, peuvent également être extraits de simulations moléculaires pour des études de gaz en micro conduites. Des descriptions détaillées, à la fois des aspects théoriques et numériques, peuvent être trouvés dans les ouvrages de Allen et Tildesley, (1987) [3], Frenkel et Smit, (2002) [31] ou encore Rapaport, (2004) [81]. Dans la suite, les principaux éléments sont brièvement exposés.

Pour une particule i, les équations du mouvement de Newton peuvent s'écrire

$$m_i \, \ddot{\bar{r}}_i = \bar{f}_i \tag{2.1}$$

avec

$$\bar{f}_i = -\nabla_{\bar{r}_i} U^{pot}(\bar{r}_1, \ldots, \bar{r}_N) = -\frac{\partial U^{pot}(\bar{r}_1, \ldots, \bar{r}_N)}{\partial \bar{r}_i} \tag{2.2}$$

où m_i, $\bar{r}_i$ et $\ddot{\bar{r}}_i$, sont respectivement la masse et les vecteurs position et accélération de la particule i, le symbole $\nabla_{\bar{r}_i}$ l'opérateur de gradient par rapport à la position $\bar{r}_i$, et $\bar{f}_i$ représente la force qu'elle subit. Cette dernière peut provenir des interactions (attraction/répulsion) avec les particules avoisinantes, mais aussi du milieu extérieur (gravité, force magnétique, etc.). L'ensemble des potentiels dont dérive la force est contenu dans U^{pot}, dont une expression simplifiée sera détaillée dans la suite.

Les échelles de temps pouvant être considérées sont corrélées au temps d'interactions entre particules, les approches par dynamique moléculaire ne peuvent alors s'effectuer que sur des périodes de temps correspondant à quelques picosecondes jusqu'à quelques microsecondes. De nombreuses méthodes ont été développées pour l'intégration des équations de mouvement. Les domaines moléculaires traités étant généralement des systèmes fermés, un point

important concerne la conservation de l'énergie au cours du temps. Par exemple, un algorithme d'intégration communément utilisé est celui de *Verlet-vitesse* [92]. Plus de détails seront donnés dans la section dédiée aux méthodes numériques 2.2.1.

Divers logiciels de simulation reposent sur la méthode de dynamique moléculaire. Pour l'étude de simulation de ce travail de recherche, deux codes sont utilisés : l'un a été développé au laboratoire MSME et implémenté sur des ordinateurs parallèles à mémoire distribuée, et l'autre est `LAMMPS` (Large-scale Atomic/Molecular Massively Parallel Simulator) [56], qui est un logiciel de open source pouvant être utilisé pour simuler le comportement de systèmes plus complexes comme la croissance des cristaux, la fusion et la solidification, la plasticité, etc.

Dans la suite, les expressions de différentes grandeurs thermodynamiques sont détaillées.

2.1.1.1 Énergies et potentiel d'interaction

Énergies et potentielle, potentiel d'interaction Il existe une grande variété de modèles pour décrire les interactions entre les particules d'un système. Pour un système de N particules appartenant au volume V, le potentiel d'interaction, U^{pot}, peut se décomposer en somme de termes, chacun associé à un type d'interaction. Formellement,

$$U^{pot}(\bar{r}_1,\ldots,\bar{r}_N) = \sum_{i|\bar{r}_i\in V}\left[U_i^{(a)}(\bar{r}_i) + \sum_{\substack{j|j\neq i\\ \bar{r}_j\in V}} U_i^{(a/b)}(\bar{r}_{ij}) + \sum_{\substack{j|j\neq i\\ \bar{r}_j\in V}}\sum_{\substack{k|k\neq i,k\neq j\\ \bar{r}_k\in V}} U_i^{(a/b/c)}(\bar{r}_{ij},\bar{r}_{ik},\theta_{ijk}) + \ldots\right] \tag{2.3}$$

où a, b et c indiquent la nature des atomes, $U_i^{(a)}$ est l'énergie potentielle due à un champ extérieur, $U_i^{(a/b)}$ représente le potentiel d'interaction de paires entre i et j et dépend ici de leur distance $\bar{r}_{ij} = \|\bar{r}_{ij}\| = \|\bar{r}_j - \bar{r}_i\|$, et $U_i^{(a/b/c)}$ représente les interactions d'un triplet i et j et k, avec θ_{ijk} l'angle entre les vecteurs $\bar{r}_{ij}$ et $\bar{r}_{ik}$. Des interactions d'ordre plus élevé viennent ensuite s'ajouter.

Lors d'une simulation de dynamique moléculaire classique, la tâche la plus coûteuse est l'évaluation du potentiel. Afin de réduire le coût, il est généralement supposé que les interactions de paires sont en nombre prédominant par rapport à celles d'ordre supérieur et l'expression du potentiel est communément réduite à deux termes. Ainsi seules les contributions $U^{(a)}$ et $U^{(a/b)}$ sont prises en compte.

Dans la suite, nous nous intéressons à des écoulements de gaz monoatomique, généralement de l'argon, interagissant avec des parois composées d'atomes de platine. Le potentiel de paire de Lennard-Jones est adapté à la modélisation de gaz monoatomique, c'est-à-dire aux interactions entre atomes d'argon. Son utilisation est étendue ici à celles entre atomes de platine, mais aussi à celles entre argon et platine. L'expression du potentiel est donnée par

$$U_i^{(a/b)}(r_{ij}) = 4\varepsilon_{a/b}\left[\left(\frac{\sigma_{a/b}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{a/b}}{r_{ij}}\right)^{6}\right] \tag{2.4}$$

Les paramètres σ et ε du potentiel désignent respectivement la distance pour laquelle le potentiel s'annule et la profondeur du puits d'énergie à son minimum. L'évolution de ce potentiel en fonction de la distance r_{ij} est présentée en Fig. 2.1. L'intensité du potentiel dépend des natures (a et b) des atomes considérés.

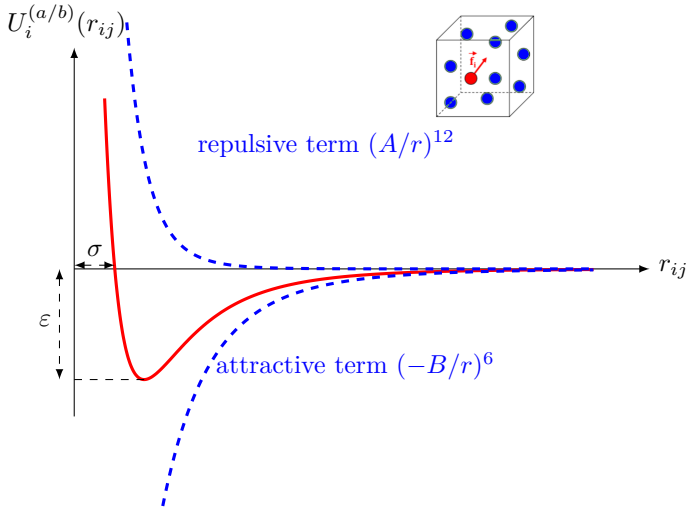


FIGURE 2.1 – Représentation du potentiel de Lennard-Jones en fonction de la distance intermoléculaire r_{ij}. Le paramètre ε est la profondeur du puits. σ est la distance de van der Walls et est la distance constante entre deux atomes pour laquelle le potentiel s'annule. Généralement σ est considérée comme le rayon de la molécule. Le cube est un domaine moléculaire où les molécules (en bleu) interagissent avec la molécule i (en rouge).

Chaque interaction entre les particules de nature a et/ou b se voit ainsi attribuer un jeu de paramètres "direct" $(\sigma_{a/a}, \varepsilon_{a/a})$, $(\sigma_{b/b}, \varepsilon_{b/b})$ ou "croisé" $(\sigma_{a/b}, \varepsilon_{a/b})$. Les interactions entre particules de même nature sont connues et largement référencées dans la littérature [81]. Par exemple pour de l'argon $\sigma_{Ar/Ar} = 3.405 \times 10^{-10}$ m et $\varepsilon_{Ar/Ar} = 1.656 \times 10^{-21}$ J. En revanche, les interactions croisées, comme $(\sigma_{a/b}, \varepsilon_{a/b})$, sont évaluées par des lois de mélange, ici en utilisant la loi de combinaison de Lorentz-Berthelot [81].

$$\begin{cases} \sigma_{a/b} = \dfrac{1}{2}\left(\sigma_{a/a} + \sigma_{b/b}\right) \\ \varepsilon_{a/b} = \sqrt{\left(\varepsilon_{a/a}\ \varepsilon_{b/b}\right)} \end{cases} \tag{2.5}$$

Compte tenu des propriétés du potentiel d'interaction, des échelles de référence et leurs quantités dérivées peuvent être construites et permettent de travailler avec des valeurs sans dimension. Avec pour référence les propriétés de l'argon, les différentes échelles de références sont données dans le tableau 2.1.

Une fois le potentiel choisi, la force d'interaction dérive du potentiel. Pour le couple de particules i et j, distantes de r_{ij}, respectivement de natures a et b, la force $\bar{f}_{ij}$, représentée en figure 2.2, et peut s'écrire :

$$\bar{f}_{ij}\left(r_{ij}\right) = \left(\frac{48\ \varepsilon_{a/b}}{\sigma_{a/b}^2}\right)\ \left[\left(\frac{\sigma_{a/b}}{r_{ij}}\right)^{14} - \frac{1}{2}\left(\frac{\sigma_{a/b}}{r_{ij}}\right)^{8}\right] \bar{r}_{ij} \tag{2.6}$$

Échelle	Expression	Valeur (Argon)	
masse	m	6.633×10^{-26}	kg
longueur	σ	3.405×10^{-10}	m
énergie	ε	1.656×10^{-21}	J
temps, t_0	$\sqrt{m\sigma^2/\varepsilon}$	2.155×10^{-12}	s
température, T_0	ε/k_B	120	K
vitesse, v_0	$\sqrt{\varepsilon/m}$	158.03	m/s
densité, ρ_0	m/σ^3	1680.31	kg/m^3
force, g_0	ε/σ	4.86×10^{-12}	N
viscosité, μ_0	$\sqrt{m\varepsilon}/\sigma^2$	9.0345×10^{-5}	kg/m/s
conductivité, k_0	$k_B\sqrt{\varepsilon/m}/\sigma^2$	0.0188	W/m/K
flux de chaleur, q_0	$\sqrt{\varepsilon^3/m}/\sigma^3$	6.62×10^9	W/m^2

TABLE 2.1 – Échelles physiques, expressions et valeurs utilisées dans les simulations de dynamique moléculaire.

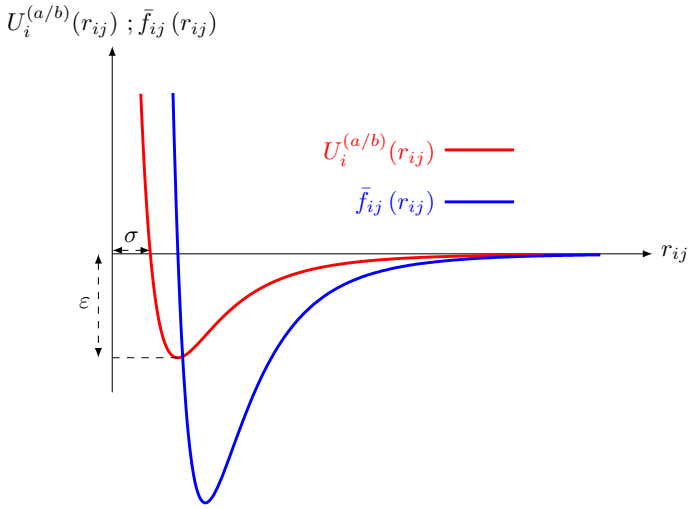


FIGURE 2.2 – Représentation du potentiel de Lennard-Jones, et de la force en fonction de la distance intermoléculaire r_{ij}.

La force d'interaction totale s'appliquant sur la particule i s'obtient par sommation,

$$\bar{f}_i = \sum_{\substack{j|j\neq i \\ \bar{r}_j \in V}}^{N} \bar{f}_{ij} \tag{2.7}$$

où la somme porte sur les N atomes du volume V, à l'exclusion de l'atome i lui-même. La troisième loi de Newton implique que $\bar{f}_{ji} = -\bar{f}_{ij}$, donc chaque paire d'atomes n'a besoin en pratique d'être examinée qu'une seule fois.

Énergie cinétique Toujours pour un système de N molécules dans un volume V, l'énergie cinétique U^c est exprimée classiquement [31] par

$$U^c = \sum_{i|\bar{r}_i \in V} \frac{1}{2} m_i \bar{\dot{r}}_i^2 \tag{2.8}$$

où U^c est l'énergie cinétique totale, $\bar{\dot{r}}_i$ est la vitesse de la molécule i.

Énergie totale À partir de ces deux dernières définitions, on peut déduire l'énergie totale qui n'est que la somme des dernières potentielle et cinétique :

$$U^{tot} = U^c + U^{pot} \tag{2.9}$$

2.1.1.2 Propriétés thermodynamiques

Température La température est un paramètre important dans les simulations de dynamique moléculaire, car elle reflète l'agitation thermique des particules dans un système.

Il existe un lien direct entre la température du système et la composante cinétique de son énergie. En effet, elle est mesurée en calculant la moyenne de l'énergie cinétique par degré de liberté [32], Son expression est donnée par

$$k_B T = \frac{2}{3N - N_c} \left\langle \frac{1}{2} m_i \left(\bar{\dot{r}}_i - \langle v \rangle \right)^2 \right\rangle \tag{2.10}$$

où, k_B, est la constante de Boltzmann, $N_{dl} = 3N - N_c$ le nombre de degrés de liberté pour un système N particules avec une quantité de mouvement totale fixe, N_c est le nombre de contraintes externes exercées sur le système. Les symboles $\langle\rangle$ représentent les moyennes de la grandeur entre ces symboles, utilisés pour montrer le passage à des grandeurs macroscopiques.

Pression La pression, comme définie d'un point de vue macroscopique, est due aux chocs des particules du fluide sur la surface considérée, par exemple la pression artérielle est due à la force exercée par le sang sur les parois des artères lorsqu'il circule dans le corps. En dynamique moléculaire, elle est calculée à partir du théorème du Viriel [62]

$$p = n k_B T - \frac{1}{3V} \left\langle \sum_{i=1} \sum_{j \neq i} \bar{f}_{ij} \cdot \bar{r}_{ij} \right\rangle \tag{2.11}$$

où $n = N/V$ est la densité volumique de particules, $\bar{f}_{ij}$ et $\bar{r}_{ij}$ représentent respectivement la force d'interaction et le vecteur de position entre les deux molécules i et j. La pression s'exprime alors comme la somme de deux termes : le premier terme représente principalement le mécanisme de convection de particules correspondant alors à une contribution cinétique, et le second terme désigné par le terme Viriel, qui est la moyenne de la résultante des interactions entre les molécules.

Flux d'énergie total Différents flux (masse, chaleur, énergie, ...) peuvent être extraits de simulations selon les besoins. Ici, seul le flux total d'énergie $\bar{q}_{tot}$ est présenté, car il est utilisé dans la suite de ce travail afin de discuter du modèle de glissement thermique sur une paroi solide. À l'équilibre thermodynamique, issu du transport interne entre les molécules [27, 96],

il est défini par,

$$\begin{aligned}\bar{q}_{tot} &\equiv \frac{1}{V_c}\left(\frac{d}{dt}\sum_i \bar{r}_i\, U_{int}\right) \\ &= \frac{1}{V_c}\left(\frac{d}{dt}\sum_i \bar{r}_i \left\{\underbrace{\frac{1}{2}m_i(\bar{\dot{r}}_i)^2}_{U^c_{int}} + \underbrace{\left[U_i^{(a)}(\bar{r}_i) + \sum_{j|j\neq i} U_i^{(a/b)}(r_{ij})\right]}_{U^{pot}_{int}}\right\}\right)\end{aligned} \tag{2.12}$$

où $U_{int} = U^c_{int} + U^{pot}_{int}$ est l'énergie interne de la particule i correspondant à l'énergie totale interne, $\bar{r}_i$ est la position de la particule i et $\bar{\dot{r}}_i$ sa vitesse et V_c le volume de contrôle dans lequel les moyennes sont calculées. Après plusieurs combinaisons et simplifications (plus de détails dans l'annexe A), l'expression du flux, comme dérivé dans Irving et Kirkwood [47], finalement devient

$$\bar{q}_{tot} = \frac{1}{2V_c}\left(\sum_i m_i \bar{\dot{r}}_i^2 \bar{\dot{r}}_i + \sum_i \sum_{j|j\neq i} U_{ij}\bar{\dot{r}}_i - \sum_i \sum_{j|j\neq i} \bar{r}_{ij}\left[\bar{f}_{ij}\cdot\bar{\dot{r}}_i\right] - 2\sum_i\left(\bar{f}\cdot\bar{r}_i\right)\bar{\dot{r}}_i\right) \tag{2.13}$$

2.1.2 Modèle continu

2.1.2.1 Modélisation des écoulements de fluide

Dans le domaine continu et pour un écoulement compressible d'un gaz, les équations de Navier-Stokes (équations de conservation de la masse, de la quantité de mouvement) et l'équation de l'énergie, ainsi que l'équation d'état reliant la pression, la densité et la température, sont exprimées par les équations (2.14) à (2.17), où le vecteur vitesse, la température, la densité et la pression sont respectivement notées $\bar{u}$, T, ρ et p.

$$\frac{\partial \rho}{\partial t} + \bar{\nabla}\cdot(\rho\bar{u}) = 0 \tag{2.14}$$

$$\frac{\partial \rho\bar{u}}{\partial t} + \bar{\nabla}\cdot(\rho\bar{u}\otimes\bar{u}) = -\bar{\nabla}p + \bar{\nabla}\cdot\bar{\bar{\tau}} \tag{2.15}$$

$$c_p\left(\frac{\partial(\rho T)}{\partial t} + \bar{\nabla}\cdot(\rho\bar{u}T)\right) = \bar{\nabla}\cdot(k\nabla T) - \left(\frac{\partial \ln\rho}{\partial \ln T}\right)\frac{Dp}{Dt} + \bar{\bar{\tau}} : \bar{\bar{d}} \tag{2.16}$$

$$p - z(\rho,T)\rho r T = 0 \tag{2.17}$$

Le symbole $\otimes$ désigne le produit tensoriel, $\bar{\bar{d}} = \frac{1}{2}\left(\nabla\bar{u} + (\nabla\bar{u})^t\right)$ est le tenseur des taux de déformation, $\bar{\bar{\tau}} = 2\mu\bar{\bar{d}} - \frac{2}{3}\mu(\nabla\cdot\bar{u})\bar{\bar{I}}$ le tenseur des contraintes visqueuses, les propriétés physiques ρ, μ, k et c_p sont respectivement la masse volumique, la viscosité dynamique, la conductivité thermique et la chaleur spécifique du fluide et $r = \frac{R}{M}$ représente la constante spécifique des gaz.

2.1.2.2 Modélisation des conditions aux limites de glissement

Dans les écoulements de dimensions très réduites et à faible pression, c'est-à-dire lorsque le nombre de Knudsen (définit en Sec. 1.2 est grand), des conditions aux limites de glissement dynamique et thermique doivent être prises en compte au niveau des parois solides. Ces conditions sont dans la suite exprimée dans un repère bidimensionnel dans lequel $\bar{n}$ est le vecteur normal extérieur à la paroi et où $\bar{t}$ est le vecteur tangent à la paroi.

- **Conditions dynamiques**. Originalement proposée par Maxwell [65] et Navier [71], les conditions dynamiques sont

$$\bar{u} \cdot \bar{t} = \sigma_u \lambda \, \bar{n} \cdot \bar{\bar{d}} \cdot \bar{t} \tag{2.18}$$

où λ représente le libre parcours moyen des molécules et σ_u est un coefficient d'accommodation. Plusieurs expressions du libre parcours moyen peuvent être utilisées. Dans la suite, nous retiendrons l'expression du libre parcours moyen (1.1).

$$\lambda = \frac{\mu}{\rho}\sqrt{\frac{\pi r T}{2}} \tag{2.19}$$

où μ la viscosité, ρ la masse volumique, r la constante spécifique des gaz et T la température.

En introduisant l'expression du tenseur de déformation, et pour une paroi plane, l'expression se réduit à sa forme simplifiée

$$u - u_w = L_s \left.\frac{\partial u}{\partial \bar{n}}\right|_w \tag{2.20}$$

dans laquelle la longueur de glissement L_s regroupe l'ensemble des paramètres (le libre parcours moyen, le coefficient d'accommodation, etc.)

- **Conditions thermiques.** Elles regroupent deux sauts, l'un portant sur la température, l'autre portant sur le saut de flux de conduction.

Le saut de température à l'interface solide-fluide, causé par la résistance de Kapitza [49][48, 73], indique le processus thermique selon lequel la chaleur latente du gaz est transférée dans la paroi solide par conduction thermique. Un saut de température à la paroi est alors exprimé par

$$T - T_w = \xi_\tau \lambda \nabla T \cdot \bar{n} \quad ; \quad T - T_w = L_T \left.\frac{\partial T}{\partial \bar{n}}\right|_w \tag{2.21}$$

Toujours au niveau d'une paroi séparant solide (s) et fluide (f), le saut de flux de conduction peut s'écrire [17, 63, 73]

$$k_f \nabla T \cdot \bar{n} - k_s \nabla T \cdot \bar{n} = \mu \bar{u} \cdot \bar{t} \cdot \bar{\bar{d}} \cdot \bar{n}$$

Dans le cas d'une surface plane et d'un glissement, il apparaît que le saut du flux thermique de conduction est égal à l'énergie dissipée par le frottement visqueux, $-\mu u \nabla u \cdot \bar{n}$.

Il convient alors d'ajouter à ce terme une contribution proportionnelle à l'énergie cinétique du fluide pour tenir compte du saut de flux de conduction.

2.2 Méthodes numériques

2.2.1 Solveur micro

2.2.1.1 Intégration des équations de mouvement

Une variété de méthodes numériques différentes est disponible, pour intégrer les équations du mouvement. En particulier, deux schémas numériques sont largement utilisés en dynamique moléculaire, connus sous le nom de méthodes de saut de mouton "*leapfrog*" [28] et l'algorithme de "*Verlet*" [57] impliquant une approche prédictive-corrective. Ces deux schémas sont algébriquement équivalents et permettent d'obtenir des coordonnées précises au troisième ordre en t. En termes de conservation de l'énergie, lorsque des potentiels de type Lennard-Jones sont impliqués, tendent à être considérablement meilleures que les méthodes d'ordre supérieur, comme Predictor–Corrector (PC), discutées dans [81]. En matière de stockage, ils demandent peu de mémoire.

Les facteurs qui ont un impact sur le temps total calcul nécessaire à une simulation est la taille du système, le pas de temps d'intégration Δt qui représente la durée entre les évaluations du potentiel et surtout le calcul des forces. Ce coût peut être réduit en définissant des paramètres numériques spécifiques, comme le rayon de coupure r_c détaillé dans la section 2.2.1.3.

L'intégration de ces équations de mouvement permet alors d'obtenir les positions, les vitesses et les accélérations de chaque particule i à chaque temps $t^{(n)} = n\Delta t$ ou à pas intermédiaires pour certains algorithmes. L'algorithme de *Verlet* [57] dérive immédiatement du développement de Taylor comme suit :

$$\bar{r}_i\left(t+\Delta t\right) = \bar{r}_i\left(t\right) + \Delta t\ \bar{\dot{r}}_i\left(t\right) + \frac{1}{2}\Delta t^2\ \bar{\ddot{r}}_i\left(t\right) + \mathcal{O}\left(\Delta t^3\right) \tag{2.22}$$

où $\bar{r}_i(t)$, $\bar{\dot{r}}_i(t)$ et $\bar{\ddot{r}}_i(t)$ sont respectivement les vecteurs position, vitesse et accélération de la particule i à l'instant (t), et Δt est le pas de temps. Notez bien que $\bar{\dot{r}}_i(t)$ soit exprimé en fonction de t, il s'agit en fait d'une fonction connue via la loi de Newton en connaissant les positions (Eq. 2.1). L'accélération peut alors être déduite à partir du calcul des forces, ce calcul comme mentionné précédemment est considéré le calcul le plus coûteux dans ce schéma

$$\bar{\ddot{r}}_i\left(t\right) = \frac{1}{m_i}\bar{f}_i\left(t\right) = -\frac{1}{m_i}\nabla_{\bar{r}_i(t)}\ U_i^{pot}\left(t\right) \tag{2.23}$$

En introduisant le développement des positions à l'instant $t - \Delta t$ dans (Eq. 2.22) et en réarrangeant l'équation, les positions sont obtenues à l'ordre plus élevé

$$\bar{r}_i\left(t+\Delta t\right) = 2\bar{r}_i\left(t\right) - \bar{r}_i\left(t-\Delta t\right) + \Delta t^2\ \bar{\ddot{r}}_i\left(t\right) + \mathcal{O}\left(\Delta t^4\right) \tag{2.24}$$

L'erreur de troncature est de l'ordre $\mathcal{O}\left(\Delta t^4\right)$. Le seul inconvénient de cette formule (2.24) est que pour des machines de faible précision le Δt^2 peut engendrer des erreurs. Les vitesses n'interviennent dans cette formule, mais pour le calcul de l'énergie cinétique du système, elle peut être obtenue par

$$\dot{\bar{r}}_i\left(t\right) = \frac{\bar{r}_i\left(t+\Delta t\right) - \bar{r}_i\left(t-\Delta t\right)}{2\Delta t} + \mathcal{O}\left(\Delta t^2\right) \tag{2.25}$$

Bien que l'algorithme *Verlet* est assez efficace, ils présentent un défaut. En effet, l'utilisation des valeurs à des instants $t+\Delta t$ et $t-\Delta t$, ce qui correspond à deux pas de temps, double les incertitudes. L'algorithme *Verlet-vitesse* [92] vient corriger cette anomalie. Il procède en deux principales étapes. D'abord, les nouvelles positions sont calculées

$$\bar{r}_i\left(t+\Delta t\right) = \bar{r}_i\left(t\right) + \Delta t\ \dot{\bar{r}}_i\left(t\right) + \frac{1}{2}\Delta t^2\ \ddot{\bar{r}}_i\left(t\right) + \mathcal{O}\left(\Delta t^3\right) \tag{2.26}$$

Ensuite, après le calcul des forces à $t+\Delta t$ entre deux pas de temps, les vitesses, à $t+\Delta t$, sont obtenues par

$$\dot{\bar{r}}_i\left(t+\Delta t\right) = \dot{\bar{r}}_i\left(t\right) + \frac{1}{2}\Delta t\left[\ddot{\bar{r}}_i\left(t\right) + \ddot{\bar{r}}_i\left(t+\Delta t\right)\right] \tag{2.27}$$

L'algorithme du *leapfrog* (ou saute-mouton) [28] est tout aussi simple à obtenir en réécrivant le développement de Taylor où il utilise des vitesses intermédiaires au demi pas de temps

$$\begin{aligned}
\bar{r}_i\left(t+\Delta t\right) &= \bar{r}_i\left(t\right) + \Delta t\left[\ \dot{\bar{r}}_i\left(t\right) + \frac{\Delta t}{2}\ \ddot{\bar{r}}_i\left(t\right)\right] + \mathcal{O}\left(\Delta t^3\right) \\
&= \bar{r}_i\left(t\right) + \Delta t\ \dot{\bar{r}}_i\left(t+\frac{\Delta t}{2}\right) + \mathcal{O}\left(\Delta t^3\right) \\
&= \bar{r}_i\left(t\right) + \Delta t\ \dot{\bar{r}}_i\left(t-\frac{\Delta t}{2}\right) + \Delta t\ \ddot{\bar{r}}_i\left(t\right) + \mathcal{O}\left(\Delta t^3\right)
\end{aligned} \tag{2.28}$$

avec,

$$\begin{aligned}
\dot{\bar{r}}_i\left(t+\frac{\Delta t}{2}\right) &= \dot{\bar{r}}_i\left(t\right) + \frac{\Delta t}{2}\ \ddot{\bar{r}}_i\left(t\right) \\
\dot{\bar{r}}_i\left(t+\frac{\Delta t}{2}\right) &= \dot{\bar{r}}_i\left(t-\frac{\Delta t}{2}\right) + \Delta t\ \ddot{\bar{r}}_i\left(t\right)
\end{aligned} \tag{2.29}$$

Finalement, la vitesse aux différents instants est obtenue [81]

$$\begin{aligned}
\dot{\bar{r}}_i\left(t\right) &= \dot{\bar{r}}_i\left(t \pm \frac{\Delta t}{2}\right) \mp \frac{\Delta t}{2}\ \ddot{\bar{r}}_i\left(t\right) \\
\dot{\bar{r}}_i\left(t+\Delta t\right) &= \dot{\bar{r}}_i\left(t+\frac{\Delta t}{2}\right) + \frac{\Delta t}{2}\ \ddot{\bar{r}}_i\left(t+\Delta t\right)
\end{aligned} \tag{2.30}$$

Cet algorithme permet à la fois de conserver l'énergie totale et la quantité de mouvement totale à des temps suffisamment longs.

Dans toutes les simulations de ce travail, l'algorithme *Verlet-vitesse* a été utilisé pour intégrer les équations de mouvement un pas de temps (en unité réduite) $\Delta t = 0.005$.

2.2.1.2 Condition aux limites périodiques

Les conditions aux limites périodiques sont souvent utilisées dans les simulations de dynamique moléculaire pour modéliser des systèmes infinis ou à bords périodiques. Elles permettent de simuler le mouvement de particules dans un système qui se répète dans l'espace de manière périodique, comme une structure cristalline ou un liquide dans un système de grande taille.

En pratique, les conditions aux limites périodiques sont mises en œuvre en utilisant des techniques de "réplication périodique" de la simulation, comme représenté dans la figure 2.3. Cela consiste à reproduire virtuellement le système de particules à l'intérieur des "boîtes

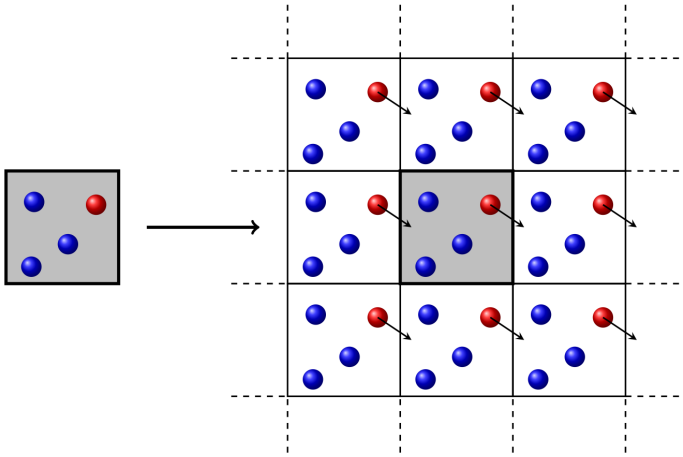


FIGURE 2.3 – Représentation des conditions aux limites périodiques. La boîte de simulation est au centre en gris. Les boîtes imaginaires sont autour sur un fond blanc. La molécule (en rouge) qui quitte le domaine de simulation est réintroduite à l'opposé [97].

imaginaires", sur un fond blanc dans la figure 2.3, de manière que les particules, la particule rouge, qui sortent de l'une des faces de la boîte, réapparaissent à l'opposé. Cela permet de simuler un système infini en utilisant une simulation finie.

L'avantage de ces conditions périodiques est lorsqu'elles sont utilisées pour simuler des systèmes en équilibre thermodynamique, comme des gaz ou des liquides, elles permettent de reproduire les propriétés thermodynamiques de ces systèmes en utilisant des simulations finies, ce qui est plus efficace que de simuler l'ensemble du système.

Il est important de noter que les conditions aux limites périodiques ont un impact sur les propriétés physiques et chimiques du système simulé. Elles ne sont, par exemple, pas recommandées pour les systèmes, où l'on s'intéresse à des phénomènes de diffusion ou de transport à longue distance, car elles peuvent entraîner des effets de bord qui peuvent fausser les résultats de la simulation.

Dans le cas de simulation d'un écoulement de *Couette* ou de *Couette-Poiseuille*, les conditions dans deux directions sont périodiques, cependant, dans une troisième direction, elles sont fixes pour permettre d'imposer une condition du mouvement sur une des parois.

2.2.1.3 Rayon de coupure

Le rayon de coupure, r_c, est une notion utilisée en dynamique moléculaire pour décrire la distance à partir de laquelle l'influence de deux particules sur l'une de l'autre est considérée comme négligeable. Cette distance est généralement déterminée en fonction de la force de Van der Waals, qui est la force attractive ou répulsive, représentée 2.1, qui agit entre les particules.

Le rayon de coupure, défini comme $r_c = C\sigma_{ij}$, est généralement utilisé pour déterminer la taille de la zone de simulation dans les modèles de dynamique moléculaire, ainsi que pour limiter le coût de calcul. Dans la littérature publiée, la distance de coupure la plus souvent utilisée dans les simulations de dynamique moléculaire est $rc = 2.5\sigma_{ij}$ [43]. D'autres distances de coupure, par exemple $r_c = 2\sigma_{ij}, 3\sigma_{ij}$ et $5\sigma_{ij}$, ont été utilisées par d'autres auteurs. Par exemple, si le rayon de coupure est défini comme étant de $2.5\sigma_{ij}$, alors toutes les particules qui sont à une distance supérieure à $2.5\sigma_{ij}$ ne seront pas prises en compte dans la simulation, car leur influence sur les autres particules sera considérée comme négligeable. Par conséquent, le fait de négliger les forces des particules qui agissent au-delà de la zone délimitée par le rayon de coupure permettra de réduire le coût de calcul.

La force d'interaction se réécrit finalement,

$$\bar{f}_{ij}\left(r_{ij}\right) = \bar{f}_{a/b}(r_{ij}) = \begin{cases} \left(\dfrac{48\ \varepsilon_{a/b}}{\sigma_{a/b}^2}\right) \left[\left(\dfrac{\sigma_{a/b}}{r_{ij}}\right)^{14} - \dfrac{1}{2}\left(\dfrac{\sigma_{a/b}}{r_{ij}}\right)^{8}\right] \bar{r}_{ij} & \text{si} \quad r_{ij} \leq r_c \\ 0 & \text{si} \quad r_{ij} > r_c \end{cases} \tag{2.31}$$

Lorsque r augmente vers r_c, la force tend vers zéro, de sorte qu'il n'y a pas de discontinuité à r_c (à la fois dans la force et le potentiel) ; ∇f et les dérivés supérieurs sont discontinus, bien que cela n'ait pas d'impact réel sur la solution numérique.

Bien que la force d'interaction $\bar{f}_{ij}$ soit négligeable, le produit $(\bar{f}_{ij} \cdot \bar{r}_{ij})$ ne l'est pas forcément. Il faut alors prendre en compte les interactions au-delà d'une certaine distance requise, en utilisant donc la notion du rayon de coupure r_c. La contribution longue distance de la pression et de l'énergie interne peuvent s'écrire, pour un potentiel de Lennard-Jones, respectivement, où C est une constante, $C \in [2, 5]$,

$$p_{ld} = -16\pi \left(\frac{N}{V}\right)^2 \left(\frac{\varepsilon\sigma^3}{3C^3}\right)\left(1 - \frac{2}{3C^6}\right) \tag{2.32a}$$

$$U_{ld} = -8\pi \left(\frac{N^2}{V}\right) \left(\frac{\varepsilon\sigma^3}{3C^3}\right)\left(1 - \frac{2}{3C^6}\right) \tag{2.32b}$$

Le choix du rayon de coupure est crucial pour garantir la qualité et la précision des résultats de la simulation. Il détermine à la fois la taille de la zone de simulation et la manière dont les forces entre les particules sont prises en compte. Idéalement, il est préférable de sélectionner une valeur de rayon qui permet de ne pas tenir compte des particules ayant une influence négligeable sur les résultats, tout en évitant de passer à côté des particules pouvant avoir des effets significatifs sur ces derniers. Si le rayon de coupure est trop grand, le coût de la simulation peut devenir très important. En revanche, si le rayon de coupure est trop petit,

alors certains effets peuvent être manqués et fausser l'analyse. Par exemple, dans le cas d'extraction de propriétés macroscopiques à partir de simulation moléculaires, des études sur l'impact de la troncature peuvent être trouvées dans [11, 34].

2.2.1.4 Modèles de murs

Dans ce travail de recherche, il s'agit d'étudier un fluide confiné entre deux parois solides et parallèles. Ainsi, prendre en compte les conditions aux limites est primordial pour une bonne modélisation. Cependant, une des difficultés liées aux systèmes moléculaires est de prendre en compte l'effet des frontières solides sur un système de taille finie et de mettre en œuvre la modélisation des conditions aux limites.

Deux différents types de parois sont considérés :
Mur atomique : Il s'agit d'une paroi explicitement constituée d'atomes, de platine dans notre cas, répartis le long d'un réseau cristallin de type cubique à faces centrées (FCC111). La figure 2.4, présente une boite de simulation en utilisant ces murs atomiques. La modélisation de tous les atomes solides prend une part importante de la simulation, ce qui rend alors cette dernière plus coûteuse en termes de temps de calcul.

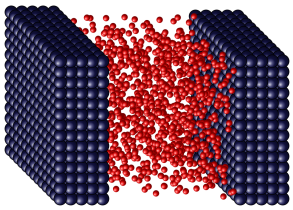


FIGURE 2.4 – Boite de simulation avec des murs atomiques en bleu qui interagissent avec les particules du fluide en rouge.

Mur stochastique : C'est une paroi fictive qui interagit avec les atomes de fluides comme un mur atomique. Contrairement à la première description, pour ce type de paroi, les atomes composant le solide ne sont pas explicitement décrits. Les interactions entre les particules composant le fluide et le "mur" sont modélisées par une combinaison entre lois de probabilités (partie diffuse) et réflexion spéculaire.
Dans `LAMMPS`, les murs stochastiques sont utilisés pour simuler les effets des parois sur les particules en mouvement à l'intérieur d'un système. Ils sont utilisés pour imiter les effets des collisions des particules avec les parois, comme si elles étaient réellement présentes. Les murs stochastiques sont souvent utilisés dans les simulations de dynamique moléculaire à l'aide du logiciel de simulation pour étudier les propriétés thermodynamiques des systèmes fermés, comme la pression, la température et l'entropie.

Un mur stochastique est considéré comme une frontière solide en mouvement avec une température finie, capable d'échanger de l'énergie avec les particules qui entrent en collision avec lui. C'est un mur réflectif et diffusif qui renvoie les particules lorsqu'elles se déplacent à l'extérieur de la paroi pendant un pas de temps et tentent de les traverser. Ces atomes

sont replacés à l'intérieur avec une vitesse modifiée. La vitesse de chaque particule après la collision est calculée de manière stochastique. Ce caractère aléatoire peut provenir du mouvement thermique des atomes du mur, de la rugosité de la surface, et plusieurs d'autres sources.

Il existe trois modèles de réflexion stochastique implémentés dans `LAMMPS`. Chaque type de mur stochastique a ses propres caractéristiques et peut être utilisé de manière appropriée en fonction du système et de l'objectif de la simulation.
Le premier de ces modèles, celui utilisé dans ce travail de recherche, est le modèle DIFFUSIF. Ce dernier a la particularité de ne pas spécifier de coefficient d'accommodation contrairement aux deux autres modèles de parois stochastiques. Dans le cadre de ce modèle, les particules sont réfléchies de manière diffuse, et la répartition de leur vitesse correspond à une distribution d'équilibre des particules à la température de la paroi.
Dans le modèle MAXWELL [65], les particules sont réfléchies de façon maxwellienne, c'est-à-dire une certaine proportion des particules est thermalisée, dans ce cas la réflexion est diffuse et une autre proportion est réfléchie de manière spéculaire.
Dans le cas du modèle CERCIGNANI et LAMPIS [13], les collisions de particules sont calculées par le modèle Cercignani/Lampis [13].

2.2.1.5 Méthodes de thermalisation des murs

Les particules solides sont généralement thématisées. Il existe plusieurs méthodes pour maintenir la température constante dans les simulations de dynamique moléculaire :

- (i) Méthode du thermostat : elle consiste à ajouter un terme de friction à l'équation de mouvement des particules, de manière à maintenir la température à un niveau constant. Il existe plusieurs types de thermostats disponibles, tels que le thermostat de Nosé-Hoover et le thermostat de Langevin [44, 91, 98], qui ont chacun leurs propres avantages et limites.
- (ii) Méthode de régulation de la température : la méthode de régulation de la température consiste à ajuster la vitesse des particules de manière à maintenir la température à un niveau constant. Cette méthode est généralement utilisée lorsque l'on souhaite maintenir la température à un niveau précis, mais elle peut être coûteuse en termes de temps de calcul.
- (iii) Méthode de contrôle de l'énergie cinétique : la méthode de contrôle de l'énergie cinétique consiste à ajuster l'énergie cinétique des particules de manière à maintenir la température à un niveau constant. Cette méthode est souvent utilisée lorsque l'on souhaite maintenir la température à un niveau précis, car elle permet de contrôler directement l'énergie cinétique des particules.

2.2.1.6 Calcul de moyenne et estimation des incertitudes

Le calcul des positions et des vitesses des particules à chaque instant permet d'obtenir la plupart des propriétés macroscopiques associées au système.

La procédure par laquelle les grandeurs macroscopiques $\bar{v}$, ρ, T et $\bar{q}_{tot}$ sont obtenues est décrite dans [30, 31]. Une technique de moyenne par blocs est alors utilisée. Cette technique selon Flyvbjerg et Petersen, (1989) [30], Frenkel et Smit, (2002) [31] permet d'estimer les erreurs statistiques des quantités statiques en étudiant simplement le comportement de la moyenne temporelle sur un temps fini τ_a. Les moyennes sont calculées dans, n_z, cellules uniformes définies par une discrétisation spatiale verticale de ℓ_z et sur des intervalles de temps τ_a (Fig. 2.5). Pour une définition plus étendue, afin d'obtenir une évaluation plus

Figure 2.5 – Moyenne en temps

précise des grandeurs physiques, nous pouvons calculer la moyenne de la variable A à un intervalle de temps τ_a :

$$\bar{A}_a = \frac{1}{\tau_a}\int_0^{\tau_a} A(t)dt \tag{2.33}$$

Dans une simulation, on peut facilement accumuler les moyennes par blocs pour un intervalle de temps τ_a. Une fois la simulation finie, on calcule la moyenne par blocs pour $n_b \times \tau_a$
Tout résultat de simulation est sujet à des erreurs systématiques et statistiques. Les erreurs systématiques sont par exemple causées par la troncature potentielle au rayon de coupure, les effets de taille finie dus à l'utilisation de petits nombres de particules, les imprécisions numériques dans l'intégration des équations de mouvement ou l'équilibrage insuffisant du système avant la phase de production d'une simulation. Les erreurs systématiques peuvent être éliminées dans une certaine mesure par un choix judicieux des paramètres de simulation. Une estimation de la variance de la moyenne de l'ensemble est donnée au biais de cette technique :

$$\sigma^2(A) = \frac{1}{n_b}\sum_{n=1}^{n_b}\left(\bar{A}_n - \langle A\rangle\right)^2 \tag{2.34}$$

où n_b représente le nombre de blocs sur lesquels la moyenne est effectuée, $\langle A\rangle$ désigne la moyenne d'ensemble de la variable A obtenue lorsque $\tau_a \to \infty$.

Une méthode alternative pour estimer l'erreur statistique dans une simulation a été développée par [30]. Admettons $\{A_1, A_2, \ldots, A_{N_E}\}$ soient N_E échantillons consécutifs d'une certaine quantité fluctuante A dont une moyenne d'ensemble et une erreur statistique seront calculées. Les N_E échantillons sont prélevés après que le système aura été équilibré. La moyenne d'ensemble est obtenue par,

$$\langle A\rangle = \frac{1}{N_E}\sum_{i=1}^{N_E} A_i \tag{2.35}$$

La variance de la moyenne de l'ensemble est estimée alors,

$$\sigma^2(A) = \left\langle A^2\right\rangle - \langle A\rangle^2 = \frac{1}{N_E}\sum_{i=1}^{N_E}\left[\bar{A}_i - \langle A\rangle\right]^2 \tag{2.36}$$

Lorsque tous les échantillons ne sont pas corrélés, il suffit d'utiliser les formules standard de

la statistique pour calculer cette variance. Toutefois, en pratique, les échantillons peuvent corréler dans une simulation et il est nécessaire d'en tenir compte.

Un schéma descriptif de la procédure de calcul la variance est représenté dans la figure 2.6.

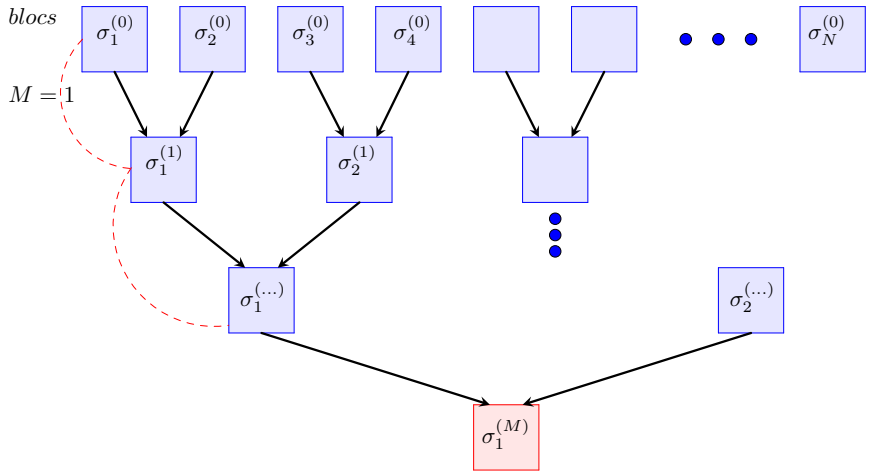


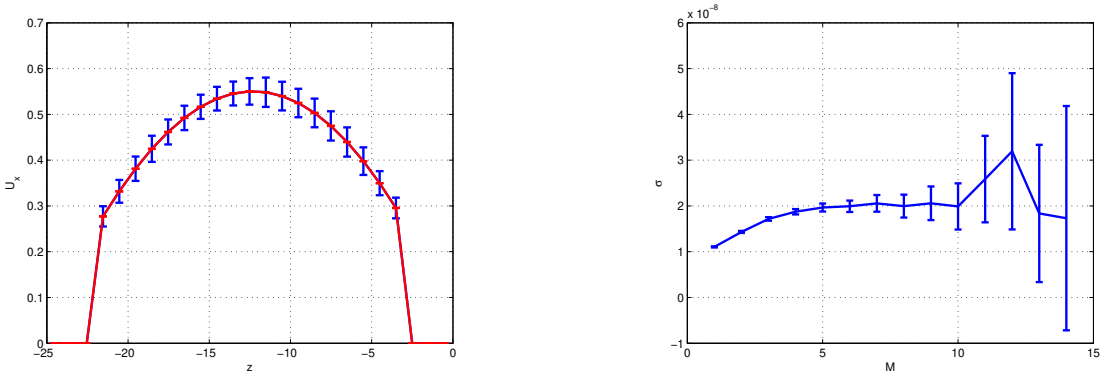
FIGURE 2.6 – Schéma présentant la procédure du calcul de la variance, M étant le nombre d'accumulations/transformations. Les blocs (en bleu) sont regroupés deux à deux, jusqu'à en avoir qu'un seul qui représentera l'incertitude finale sur une moyenne calculée.

— (i) En premier, il s'agit de regrouper les données de simulation en blocs consécutifs (deux à deux) et de calculer une moyenne et un écart-type pour chacun de ces blocs.

— (ii) En second, il faut déterminer à nouveau l'écart type ce nouvel ensemble de blocs créé. Ce dernier contient la moitié du nombre de blocs de l'ensemble original. Ces moyennes par bloc montreront une corrélation de plus en plus faible entre deux blocs consécutifs si la taille du bloc augmente.

— (iii) La procédure est répétée jusqu'à ce qu'il n'y ait plus assez de blocs pour calculer un écart-type. Le nombre de fois dont nous effectuons cette opération est appelé M.

La figure 2.7 représente les résultats de calcul des écarts types au biais de la méthode de moyenne par blocs. Dans le cadre d'une simulation, le délai entre deux blocs est généralement trop court pour obtenir un échantillon statistiquement indépendant. Par conséquent, les blocs consécutifs sont fortement corrélés. Lorsque le calcul de l'écart-type s'appuie sur ces données, l'écart-type est trop optimiste. L'opération par blocs aura pour effet qu'après avoir regroupé deux blocs consécutifs, la corrélation entre les deux nouveaux blocs sera moindre. Toutefois, l'écart-type s'en trouvera augmenté ; les données auront plus de sens puisque les blocs consécutifs ne se ressembleront plus autant. Cette diminution de la précision en fonction du nombre d'opérations par bloc se poursuivra jusqu'à ce que les blocs regroupés soient si nombreux que deux points consécutifs ne sont plus vraiment corrélés. Il est important de noter que les écarts types examinés doivent être significatifs, et que par conséquent, il convient de déterminer en même temps l'écart type de l'erreur.



(a) Profil de vitesse

(b) Incertitudes σ

FIGURE 2.7 – Estimation de l'écart-type σ obtenue avec la méthode de moyenne par blocs. Après 7 transformations par bloc ($M = [4, 5, \ldots, 10]$, les estimations restent quasi-constantes à l'intérieur des barres d'erreur à 2×10^{-8}.

2.2.1.7 Écoulement de couette avec des murs stochastiques

Description simulation

Dans ce qui suit, toutes les unités sont exprimées en unité réduite (unité de Lennard-Jones). Cette section indique comment simuler l'écoulement de *Couette* dans un canal en utilisant `LAMMPS`. Pour réaliser cette simulation, le script a été construit en plusieurs étapes : une seule boite de simulation a été créée pour le fluide. Ce fluide est compris entre deux murs stochastiques. Les deux murs sont maintenus à la même température par le biais de thermostat Langevin [44, 91] qui engendre des fluctuations de température canoniques en agissant sur les forces intermoléculaires du système. L'écoulement de *Couette* est engendré en imposant au mur de dessus une vitesse non-nulle $v_0 = 0.28$. Ensuite, le type d'atomes à densité $\rho = 1.26 \times 10^{-3}$ à structure atomique cubique à faces centrées (FCC111) est affecté pour la région fluide.

Extraction des propriétés macroscopiques

Le profil de vitesse peut être obtenu en enregistrant la vitesse moyenne dans chacune des couches, appelée bins. La moyenne est établie sur un nombre fini de quantités calculées dans la simulation à chaque pas de temps.
Comme l'épaisseur d'une couche dans la $z-$direction est de 0.025, le nombre de blocs de moyennes, équivalent au nombre de points dans la figure 2.8, est de $n_z = 40$. À partir de ce profil de vitesse, il est possible de remonter au gradient de vitesse, $\frac{\partial u}{\partial z}$.
La figure 2.8 présente le profil de vitesse, montrant un écoulement de Couette, obtenu à partir d'une simulation de dynamique moléculaire avec des murs stochastiques. Sur cette figure, il est également possible de constater qu'il y a un glissement près des parois (la vitesse $u \sim 3.2 \times 10^{-3}$).

Les barres d'erreurs représentent les incertitudes calculées en utilisant le principe de moyennes par blocs expliqué dans le livre de FRENKEL et SMIT [31], et dont la procédure est détaillée

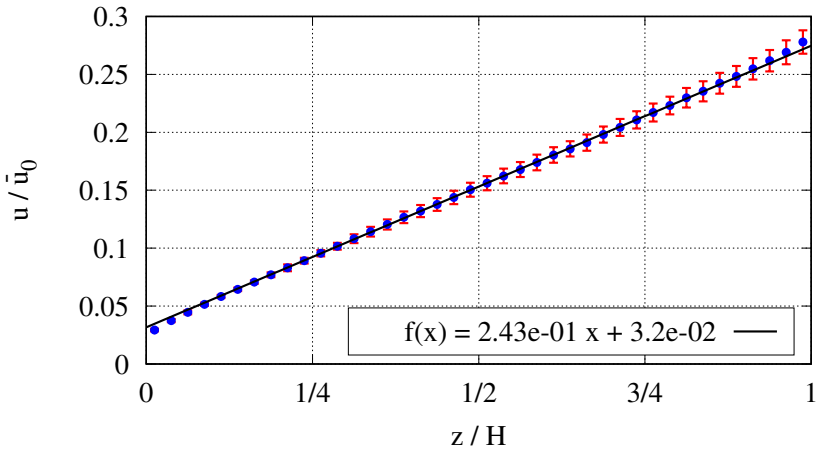


FIGURE 2.8 – Profil de vitesse (en unité LJ) dans le cas d'utilisation des murs stochastiques.

dans la section 2.2.1.6, page 48. Bien que les murs soient stochastiques diffus, Cette figure montre que les incertitudes augmente en se rapprochant du mur mobile (à vitesse $u \sim 0.3$). Cela est expliqué par la grande agitation des particules fluides près de cette paroi.
Dans les calculs ultérieurs, il est question de trouver le cisaillement $\frac{\partial u}{\partial z}$, aussi bien pour calculer le saut de flux de conduction dans le chapitre 3, la viscosité et encore afin de pouvoir mieux paramétrer la procédure de couplage (chapitre 5). Le cisaillement est présenté par la pente obtenue par le fit du profil de couette dans la figure 2.8.

2.2.1.8 Écoulement de couette avec des murs atomiques

Description simulation

Le but est de simuler l'écoulement de *Couette* dans un canal en utilisant le logiciel de simulation de dynamique moléculaire `LAMMPS`. Pour réaliser cette simulation, le code a été construit en plusieurs étapes : deux boites de simulation sont créées, une pour le fluide et l'autre pour les deux murs (haut et bas). Ensuite, deux types d'atomes de différentes densités à structure atomique cubique à faces centrées sont affectés pour les deux régions, respectivement ($\rho_{fluide} = 0.6$, $\rho_{solide} = 0.8$) représentés dans la figure 2.9. Le modèle de paroi en

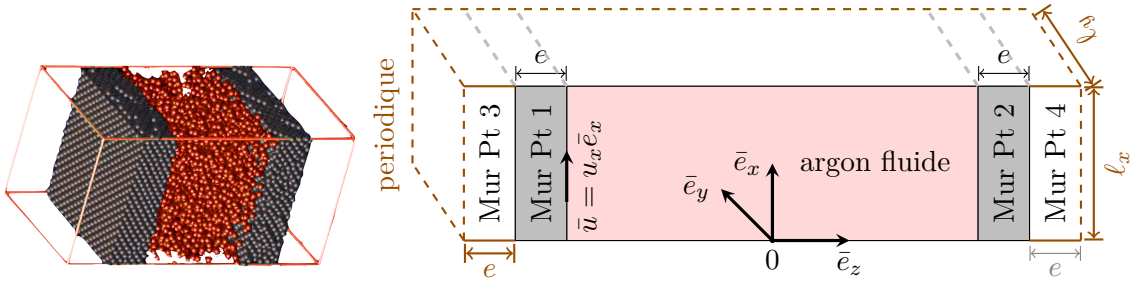


FIGURE 2.9 – Configuration de la boite de simulation. Les murs 1,2 sont thermostatés à une température fixe T_1, et dans les murs 3,4 les molécules sont des molécules fantômes et servent à maintenir les murs stables ainsi que des conditions limites périodiques pour l'ensemble de la boîte de stimulation.

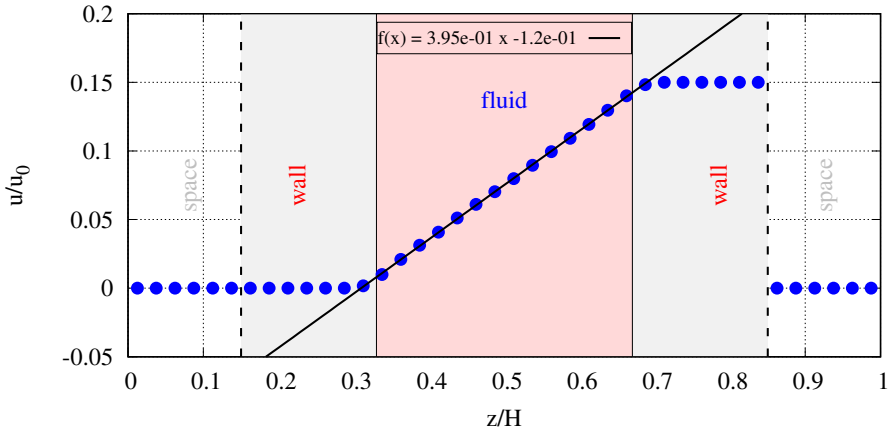
réseau transmet de l'énergie au système. Par conséquent, il est impératif d'utiliser un thermostat pour dissiper la chaleur visqueuse induite par le mouvement de cisaillement [51]. Le

thermostat choisi est alors celui de Langevin [44, 91, 98] qui engendre des fluctuations de température en agissant sur les forces intermoléculaires du système.

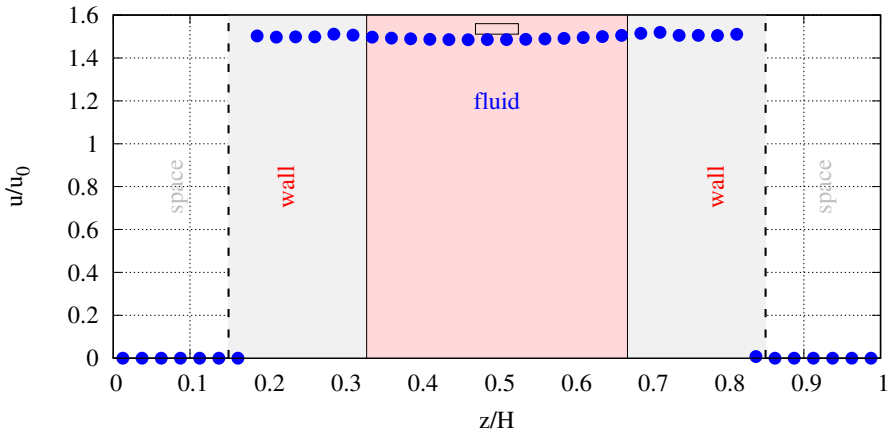
Afin de modéliser les interactions entre les atomes, le potentiel d'interaction le plus commode a été utilisé dans celle simulation, à savoir le potentiel de Lennard-Jones, voir (Eq. 2.4). La température des murs a été fixée à $180K$, (équivalent à $T = 1.5$ en unité LJ). Une vitesse $u = 0.5$ est ensuite imposée au mur d'au-dessus, ce qui entraînera le fluide pour un écoulement de Couette.

Profils de vitesse et de température

Les profils de vitesse et de température sont respectivement illustrés dans les figures 2.10(a) et 2.10(b). Elles illustrent le cas des simulations avec des murs atomiques en rajoutant des couches fantômes aux extrémités, proposé par [62], dans lesquelles un amortissement, issu d'un thermostat, permet de contrôler la température de l'ensemble en thermalisant uniquement ces couches fantômes à l'aide d'un thermostat de Langevin.



(a) Profil de Vitesse u



(b) Profil de température T

FIGURE 2.10 – Profils de vitesse et de température pour un écoulement de Couette avec les murs atomiques.

Tout comme les simulations utilisant les murs stochastiques, ces résultats (Fig. 2.10(a)) montrent le glissement observé près des parois (haut et bas). La pente représente ainsi le cisaillement. Dans la figure 2.10(b), la température du fluide est devenue constante à l'équilibre.

2.2.2 Solveur macro

2.2.2.1 Discrétisation des équations de Navier-Stokes

La discrétisation des équations repose sur une méthode de volumes finis non-structurés, avec des variables co-localisées au centre des volumes de contrôle, comme décrit dans l'article de Chénier et al. [16]. Ce maillage co-localisé est utilisé dans de nombreuses applications, notamment en mécanique des structures, en thermodynamique et en transport de matière. Il est particulièrement utile lorsque les données sont hétérogènes et que des propriétés physiques doivent être modélisées sur une zone très détaillée de l'espace. Les maillages co-localisés permettent de modéliser de manière précise les variations de ces propriétés dans l'espace et de simuler de manière plus précise le comportement des systèmes physiques.

L'expression discrète des opérateurs spatiaux requiert l'utilisation d'un maillage admissible qui doit vérifier des propriétés d'orthogonalité. De ce fait, ce maillage co-localisé est construit sur la base d'un maillage Delaunay qui est généralement structuré à partir d'un ensemble de points dans l'espace, en reliant ces points de manière à former des triangles qui respectent une certaine condition de Delaunay. Cette condition garantit que pour chaque triangle du maillage, le cercle circonscrit (c'est-à-dire le cercle qui passe par les trois sommets du triangle) ne contient aucun autre sommet du maillage à l'intérieur de lui. Pour les problèmes bidimensionnels traités dans cette thèse, le centre x_K de la maille K est le centre du cercle circonscrit à la maille (Fig. 2.11).Cette condition est importante, car elle permet de garantir que les triangles du maillage sont aussi réguliers que possible, ce qui peut être utile dans de nombreuses applications.

Les maillages Delaunay ont de nombreux avantages, notamment le fait qu'ils sont généralement plus réguliers et plus équilibrés que d'autres types de maillages, ce qui peut rendre les simulations plus précises. Ils ont également l'avantage de permettre un contrôle précis de la qualité des triangles formant le maillage, ce qui peut être utile pour améliorer la précision des simulations.

Ainsi, parmi les maillages respectant de manière satisfaisante toutes ces conditions (maillages acceptables), figurent bien entendu les maillages d'hexaèdres à faces parallèles deux à deux, et aussi ceux, prismatiques, basés sur des triangles de Delaunay. Dans notre cas, les maillages de Delaunay sont réalisés avec le logiciel `gmsh` [35].

La stabilisation des modes parasites de pression est obtenue par l'utilisation de regroupements de mailles dans lesquelles les oscillations de pression sont contrôlées. L'ensemble des équations non-linéaires discrètes couplées, regroupant les équations de bilan de la masse et

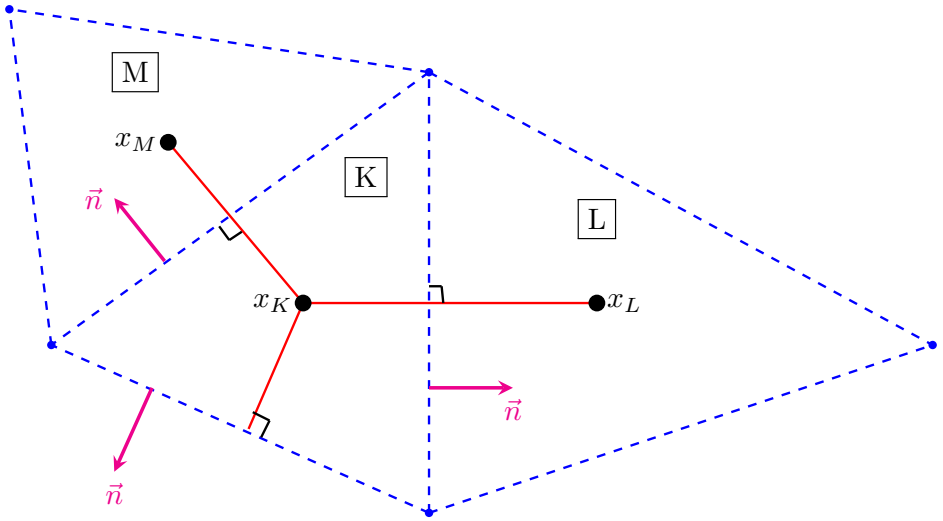


FIGURE 2.11 – Maillage Delaunay, K, L, M sont des mailles voisines, dans cette figure la maille K est une maille de bord. les x_K, x_L, x_M sont les centres de mailles où les variables sont calculées. Une figure plus détaillée est en annexe (Fig. A.5).

de la quantité de mouvement, munies des conditions initiales et aux limites appropriées, et de la loi d'état, est résolue par une méthode Newton-Raphson.

In theory there is no difference between theory and practice, but in practice there is.

Richard Phillips Feynman ou Yogi Berra

Chapitre 3

Identification des conditions aux limites thermiques par dynamique moléculaire

Sommaire

3.1 Introduction 57
3.2 Continuum description 57
3.3 Microscopic approach 58
3.4 Results and discussions 60
3.4.1 Total energy flux analysis 61
3.4.2 Comparison between the molecular and continuum models 63
3.5 Synthèse des cas étudiés 68
3.6 Estimation des incertitudes sur le flux de conduction 70
3.7 Conclusion 71

Préambule

Ce chapitre est dédié à la présentation des résultats de l'article publié dans le journal international (International Journal of Thermal Science, IJTS). Dans cet article, des simulations de dynamique moléculaire sont réalisées pour identifier les conditions aux limites thermiques. En l'occurrence, cet article est utilisé pour vérifier et valider le concept de saut de flux de conduction à la paroi démontrée par Maslen, (1958) [63], Nicolas et al., (2018) [74]. Le flux de chaleur dans la boite de simulation remplie du fluide de Lennard-Jones. Par la suite, des résultats supplémentaires, portant sur une étude plus approfondie des paramètres numériques/physiques de simulation de dynamique moléculaire, sont présentés et discutés. Ces résultats ne sont pas publiés dans le journal IJTS.

Jump in the conduction heat flux at the gas/solid interface in micro-channels

Dahia Chibouti* ; Benoît Trouette ; Eric Chénier
{dahia.chibouti ; benoit.trouette; eric.chenier} @univ-eiffel.fr

3.1 Introduction

The understanding of fluid flows and heat transfer at scales of few microns is of paramount importance since the development of Micro-electro-mechanical Systems (MEMS). When the characteristic length scale of the flow decreases to reach about 10 times the free mean path λ of the gas molecules, the usual Navier-Stokes and energy equations remain valid in the core flow, but careful attention must be paid on the gas/solid boundary conditions. In a fluid layer of order λ close to the wall, the so-called Knudsen layer, the gas molecules are no more in a local thermodynamic equilibrium : the macroscopic model based on the continuity of the velocity, temperature and heat flux at the interface between the gas and the wall falls and slip or jump conditions have to be used, like those independently proposed by Navier [71] and later on by Maxwell [65]. Whereas the boundary conditions applied on the velocity and temperature are quite well established in the 'gaseous microfluidic' community, that associated to the jump in the conduction heat flux at the wall due to the viscous friction, and first proposed by Maslen [63] in 1958, has often been neglected, forgotten or misunderstood, including in late papers. In the recent works by X. Nicolas et al. [73], a review of few papers using or ignoring the Maslen's flux condition is performed, followed by its proof which relies on the energy flux conservation principle at the macro-scale. The same authors also showed that the use of the Maslen's flux condition explains the very small Nusselt number values in experiments performed on microfluidic gaseous flows [74]. However, their mathematical approach is questionable since it relies on the assumption of the continuum mechanics which is valid far from the boundaries but seems no more applicable in the Knudsen layer.
This short contribution aims to check if the Maslen's flux condition at the gas/solid interface is correct by getting rid of the continuum model and by simulating the fluid flow and energy transfer at the molecular scale.

3.2 Continuum description

An established argon (Ar) fluid flow between two infinite (x, y)-parallel solid walls of platinum (Pt) is considered (Fig. 3.1). Walls are distant from H and their thickness is e. The flow is created by a volume force $\vec{g} = g_x \vec{e}_x$ acting as a uniform pressure gradient in x-direction. The platinum wall starting at $z = H/2$ is thermally isolated, i.e. the conduction heat flux $q_z = 0$, whereas the one ending at $z = -H/2$ is isotherm, at the temperature T_1.
For the compressible fluid flow in the micro-channel of length L, driven by the pressure difference $p_{in} - p_{out}$, the density decreases all along the channel. From the mass flow rate conservation, the velocity magnitude increases, and consequently, the viscous heating also. If the length scale associated with the pressure gradient is assumed much larger that the channel height, namely $(p_{in} + p_{out})/(p_{in} - p_{out}) \gg 2H/L$, then the density variations over a length scale H of the channel can be disregarded. Therefore the axial dependence of all the

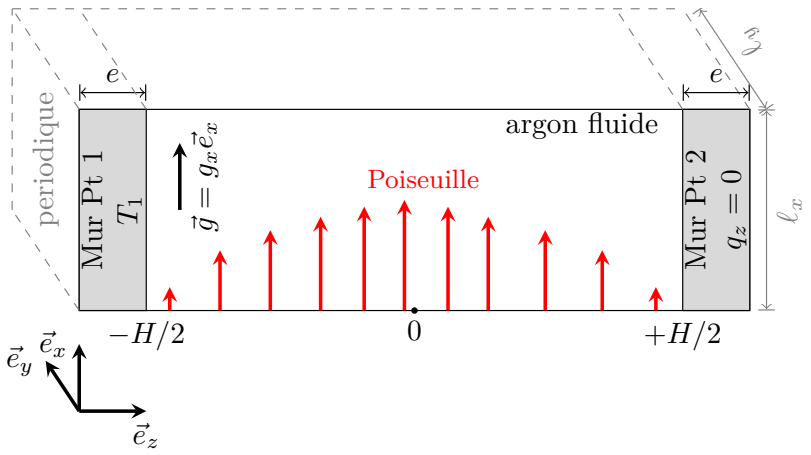


FIGURE 3.1 – Physical configuration

macroscopic variables are locally negligible, at least far from the inlet and outlet regions [74]. Thus, in what follows, the thermodynamics quantities are assumed to depend only on the transverse coordinate z.

From a macroscopic point-of-view, and as discussed in [73], the total energy flux $\vec{q}_{tot}^{(c)}$ in the fluid region can be split into three contributions :

$$\begin{aligned} \vec{q}_{tot}^{(c)} &= \vec{q}_k + \vec{q}_\mu + \vec{q}_h \\ &= -k_f\vec{\nabla}T - (\bar{\bar{\tau}} \cdot \vec{u}) + \rho h\vec{u} \end{aligned} \tag{3.1}$$

where $\vec{q}_k = -k_f\vec{\nabla}T$ is the conduction heat flux with k_f the fluid conductivity, $\vec{q}_\mu = -(\bar{\bar{\tau}} \cdot \vec{u})$ is the energy flux produced by the viscous friction and $\vec{q}_h = \rho h\vec{u}$ is the convection enthalpy flux. Since the velocity simply writes $\vec{u} = u_x(z)\vec{e}_x$, the total energy flux in z-direction, $q_{tot}^{(c)}$, reduces into two terms in z-direction :

$$q_{tot}^{(c)} = -k_f\frac{\partial T}{\partial z} - u_x\mu\frac{\partial u_x}{\partial z} \tag{3.2}$$

The Maslen's flux boundary condition [63] is to equalize the total energy fluxes on both sides of the fluid/solid interface. In the wall region, the total energy flux reduces in its conduction contribution :

$$-k_s\frac{\partial T}{\partial z}\bigg|_{H/2} = -k_f\frac{\partial T}{\partial z}\bigg|_{H/2} - u_x\mu\frac{\partial u_x}{\partial z}\bigg|_{H/2} \tag{3.3}$$

where k_s is the solid conductivity and μ the fluid dynamic viscosity. In order to check the validity of Eq. (3.3) in the Knudsen layer, simulations are carried out at the molecular scale.

3.3 Microscopic approach

Solid walls are composed of platinum atoms disposed on a FCC111 lattice. The dynamics of all atoms, either argon confined between the solid walls or platinum, relies on the classical Newton's law. The trajectory of each atom i, or particle, is governed by $\vec{F}_i = m_i\ddot{\vec{r}}_i$ where $\ddot{\vec{r}}_i$ is the acceleration vector, m_i is the mass and $\vec{F}_i$ is the resultant of external forces acting on the particle i. With $\vec{r}_i$ is the particle position, $\vec{g}$ is the volume force which drives the

fluid flow, and assuming only binary interactions given by the pairwise potential $U_{\alpha/\beta}$, the resultant $\vec{F}_i$ is developed as $\vec{F}_i = -\vec{\nabla}\left[\vec{g}\cdot(\vec{r}_i - \langle\vec{r}_i\rangle) + \sum_j U_{\alpha/\beta}(r_{ij})\right]$ for an argon particle and is reduced to $\vec{F}_i = -\vec{\nabla}\sum_j U_{\alpha/\beta}(r_{ij})$ for a platinum one. The classical Lennard-Jones model [81] is chosen to describe the argon/argon, platinum/platinum and argon/platinum interactions :

$$U_{\alpha/\beta}(r_{ij}) = 4\varepsilon_{\alpha/\beta}\left[\left(\frac{\sigma_{\alpha/\beta}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{\alpha/\beta}}{r_{ij}}\right)^{6}\right] \tag{3.4}$$

where α and β denote the Ar- or Pt-atoms, $r_{ij} = \|\vec{r}_{ij}\|$ is the intermolecular distance between the i and j particles, $\vec{r}_{ij} = \vec{r}_j - \vec{r}_i$, $\varepsilon_{\alpha/\beta}$ is the potential well-depth and $\sigma_{\alpha/\beta}$ is the distance where the potential cancels out. Taking argon as reference for the physical properties, the interaction coefficients write $\varepsilon_{Ar/Ar} = 1\,\varepsilon$, $\sigma_{Ar/Ar} = 1\,\sigma$, $m_{Ar/Ar} = m$, $\varepsilon_{Pt/Pt} = 31.3\,\varepsilon$, $\sigma_{Pt/Pt} = 0.73\,\sigma$, $\varepsilon_{Ar/Pt} = 0.66\,\varepsilon$, $\sigma_{Ar/Pt} = 0.865\,\sigma$, $m_{Pt} = 4.8833\,m$ with $\varepsilon = 1.656 \times 10^{-21}$ J, $\sigma = 3.405$ Å and $m = 6.633 \times 10^{-26}$ kg. The reference scale expressions and their derived quantities are given in Tab. 3.1.

Scale	Expression	Value	
mass (Ar)	m	6.633×10^{-26}	kg
length	σ	3.405×10^{-10}	m
energy	ε	1.656×10^{-21}	J
time, t_0	$\sqrt{m\sigma^2/\varepsilon}$	2.155×10^{-12}	s
temperature, T_0	ε/k_B	120	K
velocity, v_0	$\sqrt{\varepsilon/m}$	158.03	m/s
density, ρ_0	m/σ^3	1680.31	kg/m^3
force, g_0	ε/σ	4.86×10^{-12}	N
viscosity, μ_0	$\sqrt{m\varepsilon}/\sigma^2$	9.0345×10^{-5}	kg/m/s
conductivity, k_0	$k_B\sqrt{\varepsilon/m}/\sigma^2$	0.0188	W/m/K
heat flux, q_0	$\sqrt{\varepsilon^3/m}/\sigma^3$	6.62×10^9	W/m^2

TABLE 3.1 – Physical scales, expressions and values used in the molecular dynamics simulations, with k_B is the Boltzmann constant

The choice of argon as a fluid is justified by the binary interactions and spherical approximation induced by the pairwise Lennard-Jones potential choice. This latter allows a good restitution of the dynamical and thermal properties of mono-atomic gases, and to some extend, of non-spherical molecules for dense cases [11, 34].
Like in continuum mechanics (Sec. 3.2), the total energy flux $\vec{q}^{(m)}_{tot}$ can be defined from a microscopic description. Using the Irving-Kirkwood formulation [27, 81], it reads :

$$\begin{aligned}\vec{q}^{(m)}_{tot} = &\frac{1}{2V}\sum_i^N m_i\dot{\vec{r}}_i^2\dot{\vec{r}}_i \\ &+ \frac{1}{2V}\sum_i^N\sum_{j\neq i}^N U_{\alpha/\beta}(r_{ij})\dot{\vec{r}}_i \\ &- \frac{1}{2V}\sum_i^N\sum_{j\neq i}^N \vec{r}_{ij}\left[\vec{f}_{ij}\cdot\dot{\vec{r}}_i\right]\end{aligned} \tag{3.5}$$

where $\dot{\vec{r}}_i$ is the velocity of particle i, $\vec{f}_{ij} = -\vec{\nabla}U_{\alpha/\beta}(r_{ij})$ is the interaction force between

particles i and j and N is the total number of particles in the volume V where $\vec{q}_{tot}^{(m)}$ is evaluated. The first and second terms of the r.h.s. are related to summations of kinetic and potential energy carried by a particle i, while the third term represents the energy transfer by the pressure work [62]. It can be noticed that no contribution involving the volume force $\vec{g}$ exists, simply because the velocity and position are independent variables.

In order to impose a zero conduction heat flux in the solid platinum wall ranging from $z = H/2$ to $z = H/2 + e$, the fluid region is replicated at the wall ending, at $z = H/2 + e$. That trick also allows keeping the periodicity of molecular simulations in all space directions. Since the potential interaction between two particles decreases quickly with their distance, to become negligible in the total amount of energy, the computation cost is highly reduced by introducing a truncation length r_c. In practice, the interaction force derived from Eq. (3.4) is set to zero, if $r_{ij} > r_c$. The truncation radius r_c is set to $r_c = 2.5\ \sigma$.
The simulations are carried out on boxes of dimensions $\ell_x \times \ell_y \times \ell_z$. With the fluid zone replica, the length is finally $\ell_z = 2(H + e)$. The lengths ℓ_i, $i = x, y$, are multiples of the gaps of the FCC111 solid lattice. They also fulfil the condition $\ell_i \geq 3 \times r_c$ in order to manage efficiently the neighbourhood of each particle (Cell and Verlet lists [81]). The height H is such that $H \geq C \times \min(\ell_x, \ell_z)$ and is adjusted to fit the desired fluid density ρ in the volume defined by $V = \ell_x \times \ell_y \times H$. The constant C, initially fixed at $C = 10$, can be adapted to change the Kn number. In order the particles in both fluid regions do not interact between each other, the wall thickness e is such that $e \geq r_c$ and e is a multiple of the vertical gap of the solid FCC111 lattice.
The temperature of the left wall is maintained at T_1 using a Langevin thermostat [45], while that of the opposite wall is left free. The Velocity-Verlet algorithm [81] is used to perform the time integration with a time step fixed at $\Delta t/t_0 = 5 \times 10^{-3}$.

To calculate the macroscopic quantities ρ, $\vec{u}$, T and $\vec{q}_{tot}^{(m)}$, time and space averages are performed on n_z uniform sub-domains covering the channel height H and over time intervals τ_a. The converged estimate of the variance of the ensemble average is computed by a block average technique [30, 32].

3.4 Results and discussions

The molecular dynamics simulations are based on a Lennard-Jones potential interactions. Therefore, the pressure, the temperature and the density are solutions of a specific equation of state [53]. Likewise, the dynamic viscosity μ and the thermal conductivity k_f are estimated through the correlations by Galliéro et al. [34], and Bugel and Galliéro [11]. The results are then presented as a function of the physical quantities, namely the mean density ρ and the volume force g_x. The magnitude order of the slip velocity and jump temperature is controlled by the Knudsen number

$$Kn = \frac{\lambda}{2H} \tag{3.6}$$

with λ is the mean free path for molecules interacting through long-range potentials, which is expressed with the formulation by Cercignani [14] :

$$\lambda = \frac{\mu}{\rho}\sqrt{\frac{\pi m}{2k_B T}} \tag{3.7}$$

where m is the molecular mass, T is the temperature of the gas, and k_B is the Boltzmann constant.

For the sake of simplicity, the physical parameters and the dynamic and thermal variables, averaged over the channel height, are gathered in Tab. 3.2 as a function of the Knudsen

ρ/ρ_0	g_x/g_0	H/σ	$\bar{u}_x/v_0$	$\bar{\theta}_b/T_0$	$\lvert\bar{q}_{tot}^{(m)}\rvert/q_0$
0.6	0.02	19.87	0.882	1.61	0.105
0.6	0.01	19.87	0.452	0.534	0.027
0.6	0.005	19.87	0.225	0.152	0.0067
0.6	0.0025	19.87	0.112	0.042	0.0011
0.6	0.002	96.65	1.171	3.135	0.066
0.1	0.01	28.05	0.719	2.59	0.0097
0.05	0.005	51.28	0.713	2.42	0.0045

TABLE 3.2 – Normalization parameters for Figs. 3.3–3.8. The Knudsen number is $Kn = 0.036$, except the simulation with $g_x/g_0 = 0.002$ where $Kn = 0.0075$

number Kn, the mean fluid density ρ and the volume force g_x. The notation $\bar{\bullet}$ denotes a spatial average over the channel height. All molecular dynamics simulations were carried out for a thermostated wall fixed at $T_1/T_0 = 1.5$. The mean free path (Eq. 3.7), and therefore the Knudsen number (Eq. 3.6), is evaluated for a fluid flow without volume force $g_x = 0$ ($u_x = 0$), at the temperature equilibrium ($T = T_1$) and with a density measured in the fluid core, far from the boundaries.

3.4.1 Total energy flux analysis

The first simulation is carried out for a fluid density $\rho/\rho_0 = 0.6$ and a volume force $g_x/g_0 = 0.01$. Based on the expression of the mean free path (Eq. 3.7), the Knudsen number (3.6) is $Kn = 0.036$. According to the fluid flow regimes, a velocity slip and a temperature jump are expected at the fluid/solid interface.

In order to compare the total energy flux at the molecular scale $q_{\text{tot}}^{(\text{m})}$ (Eq. 3.5) with its continuous counterpart $q_{\text{tot}}^{(\text{c})}$ (Eq. 3.2), the conduction heat flux q_k and the viscous power density q_μ have to be evaluated. To this end, the macroscopic velocity and temperature are first calculated by averaging the particle velocities and the local kinetic energy fluctuations (see Fig. 3.5 and Fig. 3.6 for the velocity and temperature profiles). Then, the macroscopic fluxes are approximated using the second order centred finite difference scheme to get the z-derivatives, combined with the dynamic viscosity and thermal conductivity provided by correlations [11, 34]. In Fig. 3.2 are presented the different fluxes $q_{\text{tot}}^{(\text{m})}$, $q_{\text{tot}}^{(\text{c})}$, q_k and q_μ the viscous contribution to the energy equation. Each of them is normalized with respect to the absolute value of the mean conduction heat flux over the channel height. The error bars

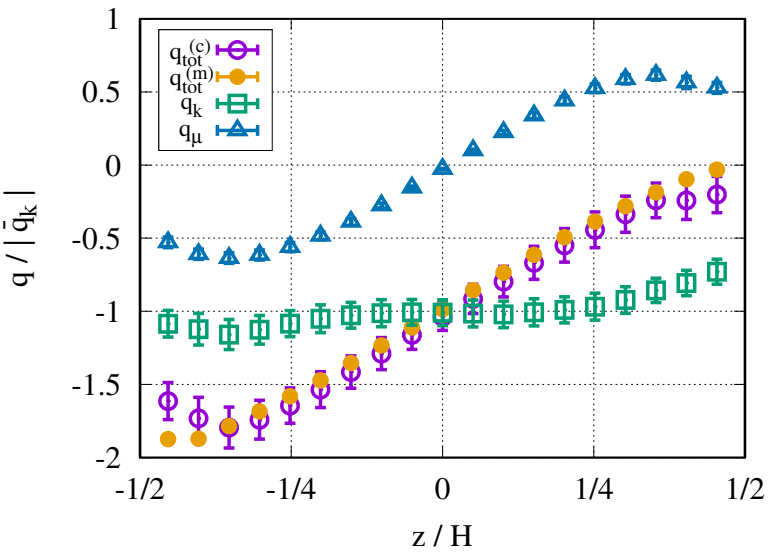


FIGURE 3.2 – Normalized fluxes with respect to $|\bar{q}_k|/q_0 = 2.74 \times 10^{-2}$, the z-average of the conduction heat flux, for $Kn = 0.036$, $\rho/\rho_0 = 0.6$ and $g_x/g_0 = 0.01$

take into account the uncertainties on both, the calculated macroscopic data, the conductivity and viscosity correlations and the approximations of the velocity and temperature derivatives as well. The analysis of the curves confirms that the total energy flux computed with the molecular dynamics, $q_{tot}^{(m)}$, gets to zero at the fluid/solid interface $z/H = 1/2$, which is consistent with the conduction heat flux transmitted to the adiabatic solid wall. The reconstructed macroscopic flux $q_{tot}^{(c)}$ exhibits a very good agreement with the molecular one $q_{tot}^{(m)}$ in the core channel, and to a lesser extent, near the fluid/solid interfaces. In this problem, the viscous contribution q_μ is far from being negligible, since it reaches almost half of the conduction heat flux. Its spatial distribution is anti-symmetric with respect to $z = 0$, which results from the symmetry of the nearly parabolic velocity profile (Fig. 3.5). Because the temperature profile is almost linear (Fig. 3.6), the conduction heat flux q_k is nearly constant ; the variations are mainly due to the conductivity dependence on the density and temperature. Finally, the macroscopic energy flux decomposition (3.2) is, as expected, accurate in the fluid core domain. On the other hand, differences are observed in the vicinity of the walls. At the adiabatic fluid/solid interface, the reconstructed flux is not as small as the molecular one, what may suggest that the macroscopic expression (3.2) fails to represent the right total energy flux, at least locally.

In order to examine thoroughly the departures between the total energy fluxes $q_{tot}^{(m)}$ and $q_{tot}^{(c)}$, calculated by either the Irving-Kirkwood formulation (Eq. 3.5) or deduced from relations of the continuum mechanics (Eq. 3.2), their relative difference to $\bar{q}_{tot}^{(m)}$ is drawn for different densities, external forces and Knudsen numbers (Fig. 3.3). The reconstructed fluxes $q_{tot}^{(c)}$ deviate by less than 10% from the molecular ones $q_{tot}^{(m)}$ in the bulk fluid domain. As discussed before, the discrepancies are mainly due to the reconstructed flux $q_{tot}^{(c)}$ which relies on, both correlation laws for the thermal conductivity and dynamic viscosity evaluations, and approximations of the derivatives by the second order centred difference scheme. However, in

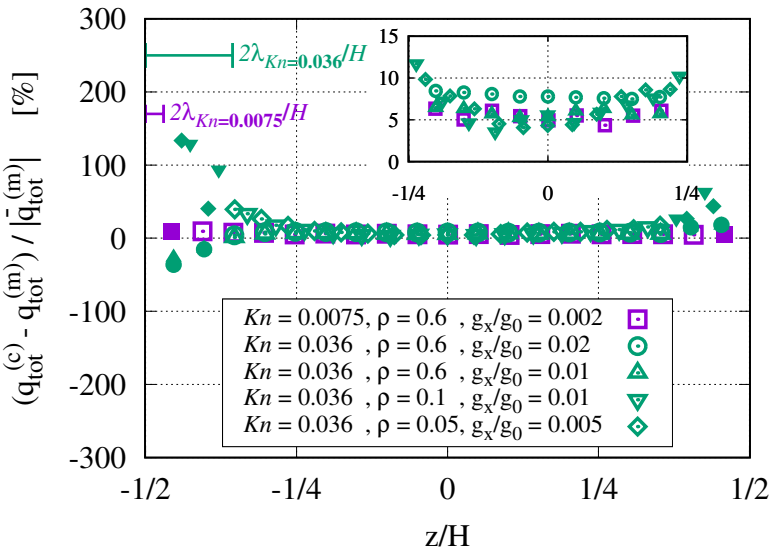


FIGURE 3.3 – Relative difference between the total energy fluxes $q_{tot}^{(m)}$ (Eq. 3.5) and $q_{tot}^{(c)}$ (Eq. 3.2) reconstructed with the molecular data. Filled (empty) symbols stand for points inside (outside) the Knudsen layer ($|z|/H - 1/2) < 2\lambda/H = 4Kn$. For clarity reason, the error bars have been omitted. Reference values are given in Tab. 3.2

a thickness region of about two mean free paths, the relative gap increases substantially, all the more that one gets closer to the wall. Looking at the density distributions (Fig. 3.4), growing oscillations develop towards the solid wall, approximately over the same length scale. These oscillations result from the clustering of the argon atoms close to the wall. Since the physical properties are density dependent, the dynamic viscosity μ and the thermal conductivity k_f fluctuate in this zone as well. This atomic patterning breaks the local thermodynamics equilibrium, which is required to get a macroscopic model for the fluid flow. The aim of the slip/jump relations is precisely to connect the macroscopic temperature and velocity variables on both side of the Knudsen layer, namely between the wall and the outer region of the Knudsen layer. The use of these boundary conditions, and more specifically the Maslen's flux condition, is the motivation of what follows.

3.4.2 Comparison between the molecular and continuum models

This section is devoted to the comparisons of the molecular dynamics results with those derived from the continuum model with constant physical properties, an incompressible fluid flow assumption and suitable boundary conditions. The simulations are carried out for volume force ranging from $g_x/g_0 = 0.0025$ to 0.02 and a mean density $\rho/\rho_0 = 0.6$. The Knudsen number, evaluated at T_1, is equal to $Kn = 0.036$.

The assumption of a constant density is acceptable in the core flow, where a weak decrease is observed due to the heating by the viscous friction (Fig. 3.4). On the other hand, at the vicinity of the boundaries, the density oscillates and departs up to about 75% of its bulk value. As already mentioned, this usual behaviour is related to the clustering of the argon atoms near solid wall due to the interaction parameters that make the Lennard-Jones

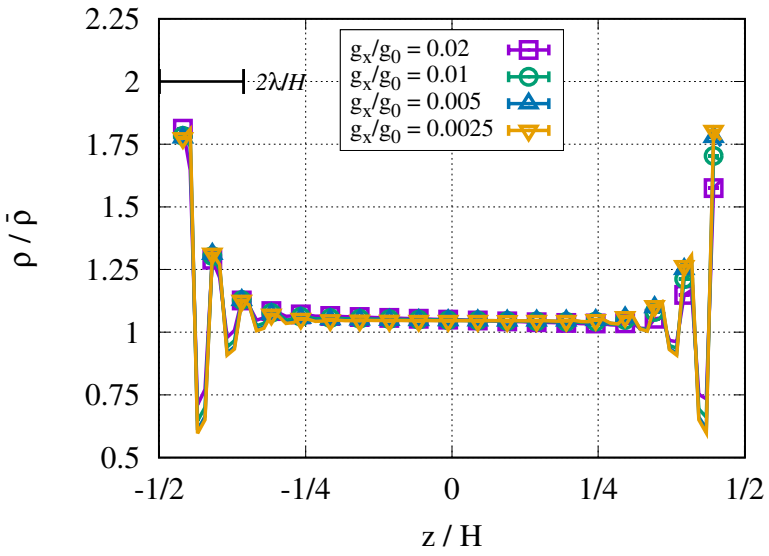


FIGURE 3.4 – Density $\rho(z)/\bar{\rho}$ for different volume forces g_x/g_0 and $Kn = 0.036$ (see Tab. 3.2). The number of bins for the molecular average calculation is $n_z = 100$

potential attractive [72]. The oscillations are damped over a length of approximately 2λ that roughly corresponds to the Knudsen layer.

The normalized velocity profiles, presented in Fig. 3.5, exhibit a Poiseuille like flow with slip at walls. Whereas the magnitude of the velocity and the dynamical slip depend on the

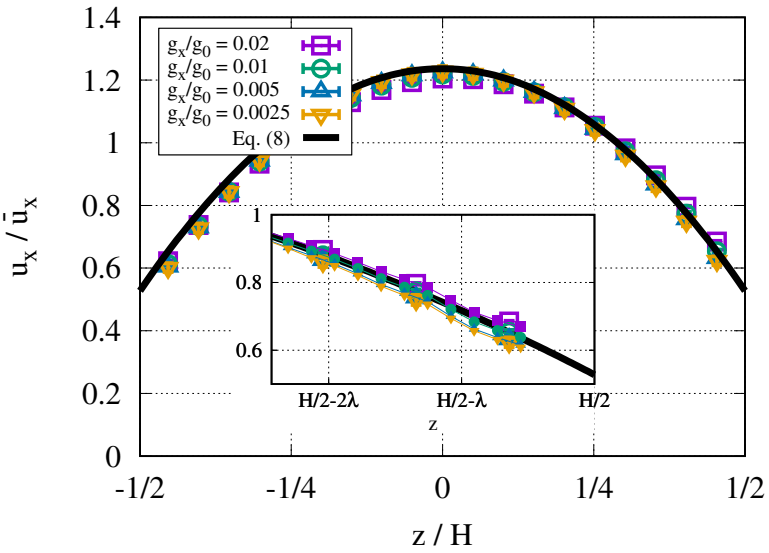


FIGURE 3.5 – Velocity $u_x(z)/\bar{u}_x$ for different volume forces g_x/g_0 and $Kn = 0.036$. The average values are provided in Tab. 3.2. The parameters for the analytical function is $\sigma_\mu = 2.46$ and $\xi = 36.0$. The number of bins for the molecular average calculation is $n_z = 25$, except the encapsulated figures which are based on $n_z = 100$

applied volume force g_x, they become quite identical once they are normalized by $\bar{u}_x$, the averaged value of u_x over the channel height (Tab. 3.2). The superposition of the molecular

data are fully consistent with the analytical function derived from the continuum model :

$$\frac{u_x(z)}{\bar{u}_x} = \frac{3(1+8\sigma_\mu Kn) - 12(z/H)^2}{2+24\sigma_\mu Kn} \tag{3.8}$$

with σ_μ is the viscous slip coefficient classically used in the expression of the dynamical slip boundary condition :

$$u_f - u_w = \sigma_\mu \lambda \left(\nabla u \cdot \vec{n}_{wf}\right)\big|_f \tag{3.9}$$

The indexes f and w indicate fluid and wall quantities evaluated at the fluid/solid interface and $\vec{n}_{wf}$ its unit normal vector, directed from the wall to the fluid. In the present problem, $u_w = 0$.

Since the velocity profile is not constant in the channel section, the viscous friction leads to mechanical energy dissipation and a temperature increase by getting away from the lower wall at T_1. This heating and the temperature jump are even more pronounced as the external force g_x increases. However, like for the velocity, the dimensionless reduced temperature θ/θ_b, with $\theta = T - T_1$ and $\theta_b = \int \rho u_x \theta \, dz / \int \rho u_x \, dz$ the dimensionless bulk temperature (Tab. 3.2), turns out to be independent on the external force g_x (Fig. 3.6). Such a result is recovered by

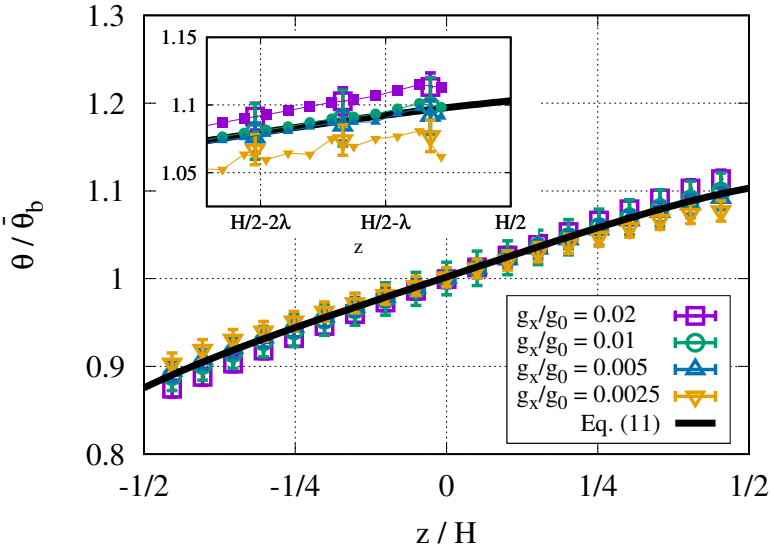


FIGURE 3.6 – Same caption as Fig. 3.5 – Reduced temperature $\theta(z)/\theta_b$

integrating the energy equation with the velocity field (3.8), the temperature jump at the lower boundary $z/H = -1/2$

$$T_f - T_w = \xi \lambda \left(\nabla T \cdot \vec{n}_{wf}\right)\big|_f \tag{3.10}$$

with ξ the thermal jump coefficient, and the Maslen's flux condition (3.3) at the adiabatic wall $z/H = 1/2$. The analytical reduced temperature $\theta(z)/\theta_b$ then reads :

$$\frac{\theta(z)}{\theta_b} = A + B\frac{z}{H} + C\left(\frac{z}{H}\right)^4 \tag{3.11}$$

with

$$A = \frac{a_0 + a_1 Kn + a_2 Kn^2}{D},\ B = \frac{96\sigma_\mu Kn + 8}{D},\ C = \frac{-16}{D},$$
$$D = \frac{2304(d_0 + d_1 Kn + d_2 Kn^2 + d_3 Kn^3)}{(12\sigma_\mu Kn + 1)}$$

and

$$a_0 = 5,\ a_1 = 48\sigma_\mu + 32\xi,\ a_2 = 192\sigma_\mu\xi,$$
$$d_0 = \frac{43}{20160},\ d_1 = \frac{\xi}{72} + \frac{11\sigma_\mu}{240},$$
$$d_2 = \frac{\sigma_\mu}{4}(\xi + \sigma_\mu),\ d_3 = \sigma_\mu^2\xi$$

This analytical solution, which is based on the Maslen's flux condition, fits perfectly the molecular data. This observation suggests that the flux boundary condition (3.3), which can be interpreted as a jump in the conduction heat flux, permits to connect correctly the thermal quantities on both part of the Knudsen layer.

As well as the velocity and temperature profiles, the total energy flux becomes independent of the volume force g_x, once the suitable normalization $q_{tot}^{(m)}/|\bar{q}_{tot}^{(m)}|$ is considered (Fig. 3.7). An analytical approximation $q_{tot}^{(a)}(z)$ of the molecular data is developed by substituting the

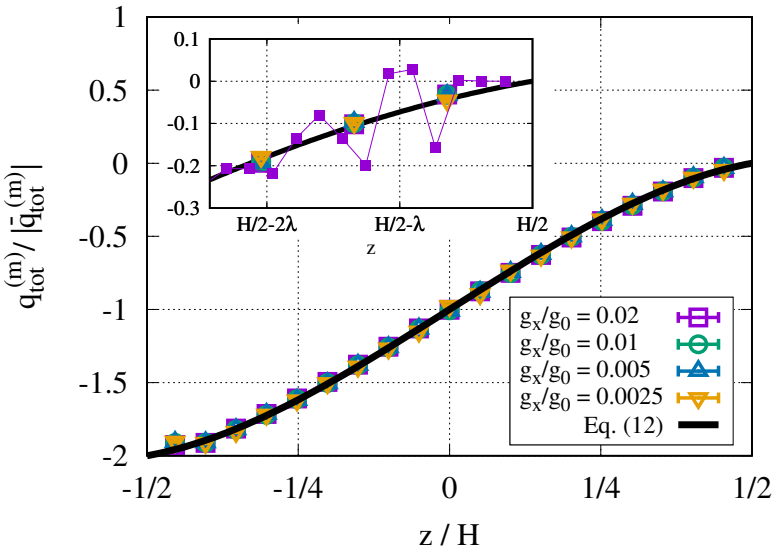


FIGURE 3.7 – Same caption as Fig. 3.5 – Total energy flux $q_{tot}^{(m)}(z)/|\bar{q}_{tot}^{(m)}|$.

velocity (3.8) and temperature (3.11) functions into the macroscopic expression of the total energy flux (3.2) :

$$\frac{q_{tot}^{(a)}(z)}{|\bar{q}_{tot}^{(a)}|} = \frac{(24\sigma_\mu Kn + 3)(z/H) - 4(z/H)^3}{12\sigma_\mu Kn + 1} - 1 \tag{3.12}$$

Again, the analytical expression provides a very good approximation of the results coming from the molecular dynamics simulations. It is worth to noticed that, as for the density profiles (Fig. 3.4), flux oscillations are observed at the molecular scale, over a length of about 2λ.

To point out the robustness of the Maslen's boundary conditions, the total heat fluxes $q_{tot}^{(m)}$, achieved by the molecular simulations and presented in Fig. 3.3, are reconsidered and com-

pared to the analytical solutions $q_{tot}^{(a)}(z)$ (Fig. 3.8). Taking into account the approximation

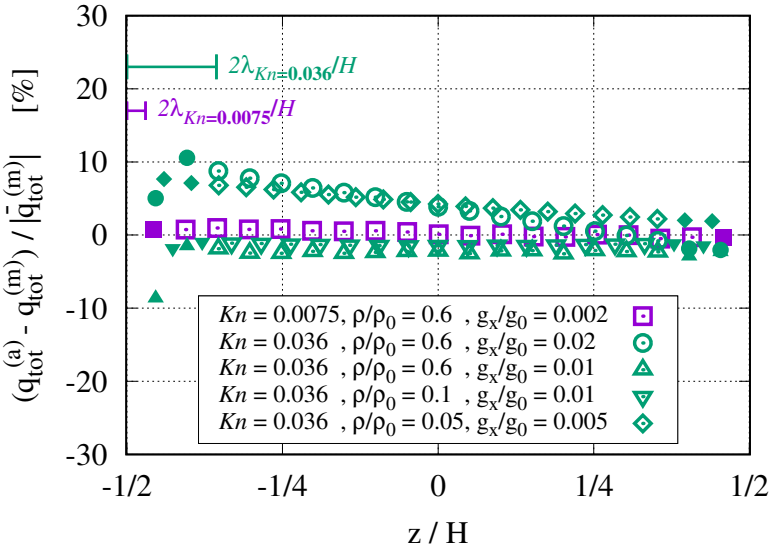


FIGURE 3.8 – Relative gap between the total energy fluxes stemmed from molecular dynamics simulations and the analytical expression (3.12). Filled symbols stand for points inside the Knudsen layer ($|z|/H - 1/2) < 2\lambda$, while empty ones denote measurement points in the bulk. For clarity reason, the error bars have been omitted. Reference values are given in Tab. 3.2

level used to develop the continuum model, the analytical function $q_{tot}^{(a)}(z)$ gives satisfying results, with relative gaps less than 10%. The accuracy is much better close to the adiabatic wall $z/H = 1/2$ ($< 2\%$), where the total energy flux is controlled by the Maslen's condition. The slight deviation observed by moving away from the adiabatic wall is probably the result of error accumulations induced by the assumptions used to develop the analytical solution. Again, the comparisons between the analytical and molecular total energy fluxes emphasize the robustness of the Maslen's flux boundary condition (3.3) to achieve a correct approximation.

From the dimensionless writing of the Navier-Stokes and energy equations with their boundary conditions, 6 dimensionless parameters appear : the heat conductivity ratio k_s/k_f, the ratio between the volume and viscous forces measured by the Jeffreys number $Je = g_x H^2/(\mu U)$, the energy flux ratio between the viscous friction and the heat conduction, namely the Brinkman number $Br = \mu U^2/(k_f \Delta T)$, and obviously the above introduced Knudsen number Kn and the slip and jump coefficients σ_μ and ξ. Setting $Je = 1$ leads to define the scale for the velocity : $U = g_x H^2/\mu$. Likewise, the equilibrium between the viscous dissipation and the heat conduction allows to define the scale for the temperature variation : $\Delta T = \mu U^2/k_f = g_x^2 H^4/\mu k_f$. Tab. 3.3 summarizes the Jeffreys, Brinkman and Knudsen parameters, and the slip and jump coefficients as well. For each fixed Knudsen, the numerical values σ_μ and ξ result from a minimization process between the equations (3.8), (3.11) and (3.12) and the averaged molecular data, using the non-linear least-squares Marquardt-Levenberg algorithm. Since the fluid properties results from the Lennard-Jones interaction potential, the ranges of variation of the Jeffreys and Brinkman parameters are moderate : the ratio between the maximal and minimal values are close to 2.

ρ/ρ_0	Kn	Je	Br	σ_μ	ξ
0.6	0.0075	9.6	0.13	2.63	36
0.6	0.036	5.1	0.24	2.46	36
0.1	0.036	5.5	0.21	4.13	37
0.05	0.036	5.2	0.15	5.72	25

TABLE 3.3 – Dimensionless parameters for simulations presented in Figs. 3.3 and 3.8 ; σ_μ and ξ are used in the analytical functions (3.8), (3.11) and (3.12).

The numerical parameters involved in the molecular dynamics simulations, such as the cut-off radius r_c, the time step Δt, the lengths ℓ_x and ℓ_y, the interaction potential of the platinum atoms of the solid wall or the (σ, ε)-values of the Ar/Pt interaction coefficients do not alter the analyses presented below. Obviously, the changes in the Lennard-Jones interaction parameters may modify slightly the fluid flow and heat transfer, especially the dynamical slip and temperature jump magnitudes, but the velocity, temperature, and flux profiles remain similar ; the total energy fluxes computed by the molecular simulation and the analytical model, using the Maslen's flux boundary condition (3.3), always cancel at the adiabatic solid wall $z = H/2$.

3.5 Synthèse des cas étudiés

Comme l'influence *a priori* des paramètres de simulation, et comment ils agissent sur les propriétés macroscopiques, il est alors nécessaire d'étudier leur influence sur les résultats pour les flux.

Les résultats d'une simulation peuvent dépendre du nombre de particules N dans le système simulé, du rayon de coupure r_c, auquel la fonction du potentiel intermoléculaire est tronquée, du pas de temps Δt utilisé dans l'algorithme d'intégration des équations de mouvement, du nombre de moyennes dans la simulation et de la taille des cellules, etc.

Dans toutes les simulations de ce travail, l'algorithme vitesse-Verlet a été utilisé pour intégrer les équations de mouvement avec la taille de pas $\Delta t = 0.005$. L'énergie totale du système a toujours été maintenue constante à l'intérieur du système. Toutes les simulations ont été équilibrées sur au moins 10000 pas de temps avant de commencer la phase de production pour atteindre l'équilibre thermodynamique dans le système.

L'influence des paramètres numériques a été largement étudiée sur configuration de masse volumique $\rho = 0.6$. Le nombre de Knudsen pour l'ensemble des cas est $Kn \approx 3.6 \times 10^{-2}$.

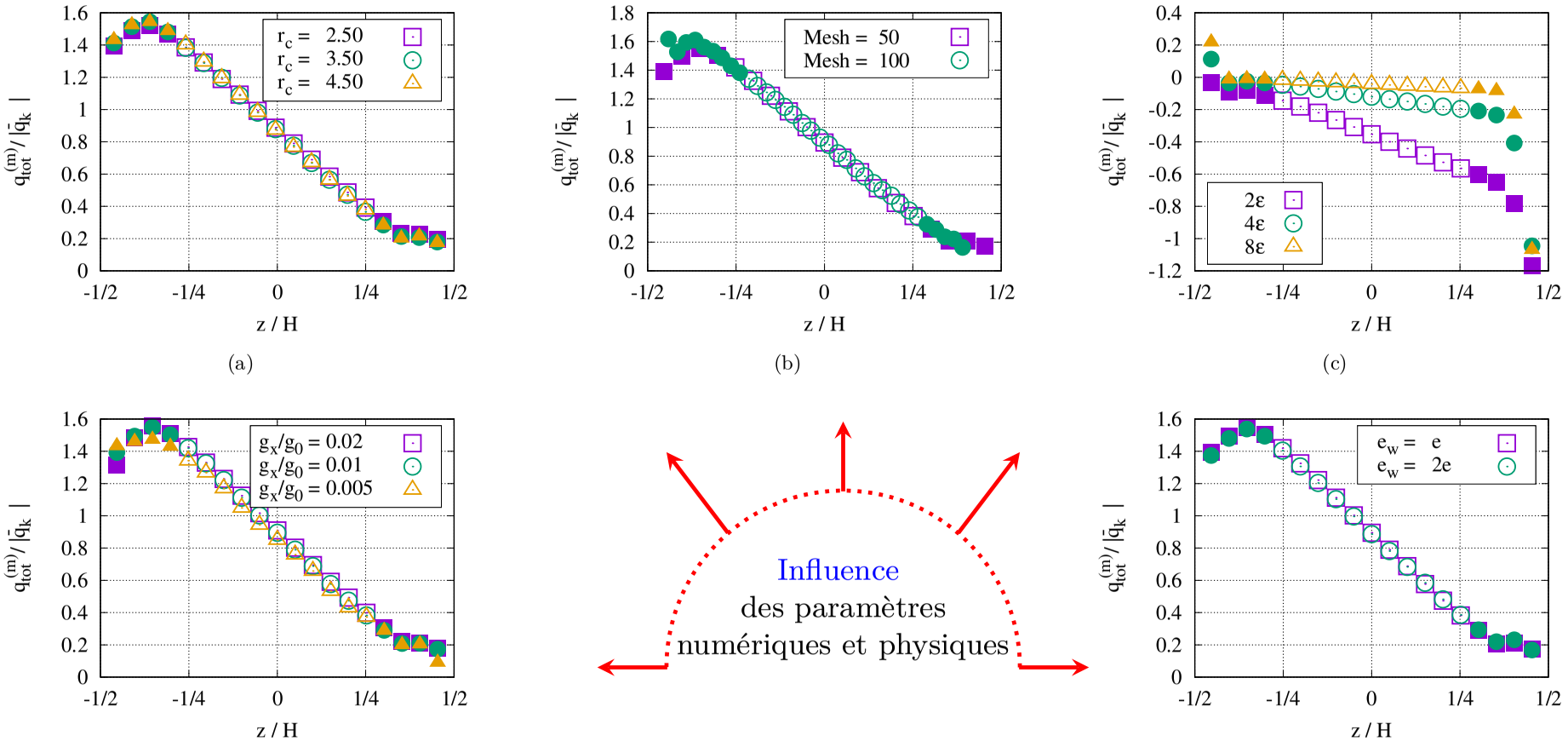


FIGURE 3.9 – Corrélation des flux normalisés par le flux total calculé q_{tot}. (a) : variation du rayon de coupure r_c, (b) : variation de la grille (nombre de bins qui sert au calcul de moyenne), (e) : variation d'épaisseur du mur e, (d) : variation de forces volumiques appliquées g_x, (c) : variation de la profondeur du puits ε. Les symboles vides représentent des points dans la zone "intérieur" définie par $z \in [-H/4, H/4]$ tandis que les symboles remplis indiquent des points de mesure près des murs ($z \in \pm[-H/4, -H/2]$).

Les conclusions présentées sur la valeur du flux de chaleur total q_{tot}, qui est annulée à l'interface du mur 2 ou le saut de flux au mur 1, ainsi que celles tirées pour la décomposition de ce flux de chaleur total en deux flux $q_k + q_\mu$ ne sont sensibles à aucun des paramètres numériques, que ce soit la force volumique (d), le rayon de coupure r_c (a), le pas de temps, la taille des cellules (b) ou même les intervalles de temps utilisés dans les moyennes.
La seule influence remarquée est lorsque la paramètres du potentiel d'interaction des atomes de platine constituant la paroi solide ou la valeur du couple (σ, ε) des coefficients d'interaction Ar/Pt. La figure (c) montre que plus la profondeur du puits augmente le flux d'énergie totale diminue. Pour tout synthétiser, plusieurs cas sont superposés dans les figures (Fig. 3.9) pour différentes valeurs de la force de volume g_x et différentes valeurs du nombre de Knudsen Kn.

3.6 Estimation des incertitudes sur le flux de conduction

L'influence des paramètres de simulation sur les résultats est discutée dans la section précédente. Ici, l'estimation des erreurs statistiques est traitée. Les erreurs statistiques sont dues au fait que les moyennes de simulation sont prises sur des trajectoires dans l'espace de phase de longueur finie. Pour obtenir des données de simulation précises avec de faibles incertitudes, le système doit expérimenter un grand nombre de processus caractéristiques à l'échelle moléculaire, qui sont responsables du comportement macroscopique, tout au long de la simulation.

En vue de comparaison, les incertitudes associées à chacun des termes calculés ou modélisés sont dérivés. Les incertitudes, sur le flux de conduction et le flux associé aux termes visqueux, sont données par les expressions,

$$\Delta q_k = \left|\frac{\partial T}{\partial z}\right| \Delta k + |k| \, \Delta \frac{\partial T}{\partial z} \tag{3.13}$$

et

$$\Delta q_\mu = \left|u \frac{\partial u}{\partial z}\right| \Delta \mu + \left|\mu \frac{\partial u}{\partial z}\right| \Delta u + |\mu u| \, \Delta \frac{\partial u}{\partial z} \tag{3.14}$$

où la dérivée partielle $\partial\phi/\partial z$ sont exprimées par différences finies (DF). h étant le pas spatial et inclut l'erreur de troncature

$$\left.\frac{\partial \phi}{\partial z}\right|_z = \frac{\phi(z+h) - \phi(z-h)}{2h} + \Delta \frac{\partial \phi}{\partial z} \tag{3.15}$$

avec

$$\Delta \frac{\partial \phi}{\partial z} \approx \underbrace{\frac{\Delta \phi}{h}}_{T_1} + \underbrace{\frac{h^2}{3} \left.\frac{\partial^3 \phi}{\partial z^3}\right|_{DF}}_{T_2} \tag{3.16}$$

où T_1 représente l'incertitude sur les valeurs numériques et T_1 l'erreur d'approximation par différences finies.
Les expressions de $\Delta\mu$ et Δk dérive des corrélations de viscosité et de conductivité du fluide de Lennard-Jones respectivement, $\mu = \mu(\rho, T)$ [34] et $k = k(\rho, T)$ [11]. Cette dernière nécessite

l'expression du facteur de compressibilité donné par l'équation d'état [53]. Ces dérivées sont données respectivement :

$$\Delta\mu_{corr} = \left|\frac{\partial\mu}{\partial T}\right| \Delta T + \left|\frac{\partial\mu}{\partial\rho}\right| \Delta\rho \tag{3.17}$$

et

$$\Delta k_{corr} = \left|\frac{\partial k}{\partial T}\right| \Delta T + \left|\frac{\partial k}{\partial\rho}\right| \Delta\rho \tag{3.18}$$

Les corrélations varient autour de $\pm 5\%$ de la valeur de viscosité ou de conductivité calculées à partir des valeurs de densité et de température, comme expliqué dans [11, 34]. Au final, la formule qui donne les incertitudes sur la viscosité et la conductivité est donné par la relation (3.19) où ϕ représente l'une des variables μ ou k, avec $\phi = \phi_{corr} + \Delta\phi$.

$$\Delta\phi = \Delta\phi_{corr} \pm 5\% \times \phi_{corr} \tag{3.19}$$

3.7 Conclusion

Results of molecular dynamics simulation, obtained for a Lennard-Jones fluid, have shown that the Maslen's boundary condition [63], demonstrated in the framework of the continuum mechanics (see [73]), although questionable in the Knudsen layer, is proved to be efficient to connect the conduction heat fluxes between the wall and the fluid domain. Therefore, the total energy flux conservation at the fluid/solid interface introduces effectively a jump in the conduction heat flux which is equal to the energy flux produced by the viscous friction.

The key to artificial intelligence has always been the representation.

Jeff Hawkins

Chapitre 4

Méthodes de régression pour l'apprentissage supervisé

Sommaire

4.1 Généralités sur l'apprentissage 73
4.1.1 Régression linéaire ou logistique [42] 75
4.1.2 Régression par machines à vecteurs de support [6, 12, 82] 75
4.1.3 Réseaux de neurones [6, 76] 76
4.1.4 Méthodes de k plus proches voisins (k nearest neighbors, kNN) [4] . 78
4.2 Régression linéaire 78
4.2.1 Régression par la méthode des moindres carrées 78
4.2.2 Régression par la méthode bayésienne 83
4.2.2.1 Modèle de régression linéaire bayésienne 85
4.2.2.2 Apprentissage sur les hyperparamètres 90
4.2.3 Régression par les processus gaussiens 91
4.3 Étude comparative des différentes techniques de mise à l'échelle 96
4.3.1 Outils de normalisation et de mise à l'échelle 97
4.3.1.1 Méthode de mise à échelle par les valeurs minimales et maximales 97
4.3.1.2 Méthode de normalisation par la moyenne 97
4.3.2 Choix de la fonction de base 98
4.3.3 Résultats de comparaison 99
4.3.3.1 Comparaison entre les différentes méthodes de mise à échelle 99
4.3.3.2 Erreurs de prédiction et incertitudes sur la prédiction . . . 103
4.3.3.3 Modalités d'enrichissement 104
4.3.3.4 Résultats de calcul sur des données bruitées 104
4.4 Conclusion 104

Préambule

Ce chapitre expose les différentes méthodes d'apprentissage utilisées au cours de cette thèse. Les bases de ces techniques sont brièvement introduites, discutées et analysées afin de permettre un choix pour les cas envisagés. Ce choix, étant la régression par les deux méthodes bayésienne et processus gaussiens, est détaillé dans ce chapitre. Des résultats préliminaires sont également présentés à la fin de ce chapitre.

4.1 Généralités sur l'apprentissage

L'apprentissage automatique regroupe un ensemble de méthodes qui consistent à construire des modèles statistiques ou déterministes d'une réponse à partir des données. Ces modèles peuvent ensuite être utilisés à des fins de prédiction, pour l'identification des comportements ou encore pour la prise de décisions. Pour y parvenir, il faut avant tout mettre à disposition de l'algorithme d'apprentissage un ensemble de données, appelé observations, afin qu'il puisse "s'entraîner" et "s'améliorer", ce qui explique le mot apprentissage.

Au cours du processus, il est opportun de vérifier par l'analyse des erreurs que le modèle devient de plus en plus efficace. Si les résultats s'améliorent lorsque les données d'apprentissage sont fournies, alors le modèle converge.

Il existe deux principales classes : apprentissage supervisé et non supervisé.
Un problème d'apprentissage supervisé est une application dans laquelle les données d'apprentissage comprennent des entrées ainsi que leurs cibles correspondantes. Il peut être subdivisé en problèmes de régression et de classification. Si la sortie souhaitée est constituée d'une ou plusieurs variables continues, ce problème est appelé régression (la sortie est alors quantitative).
Les cas tels que l'exemple de la reconnaissance des chiffres, où le but est d'assigner chaque tableau d'entrée à une parmi un nombre fini de catégories discrètes, sont appelés des problèmes de classification (la sortie est par conséquent qualitative). Pour notre étude où on cherchera des lois pour la mécanique des fluides en micro-conduites, les variables sont continues, la régression est alors la meilleure approche à suivre.

Au contraire dans certains problèmes, les données d'apprentissage ne comprennent que les entrées sans aucune valeur cible correspondante. Ces problèmes d'apprentissage sont appelés apprentissage non supervisé, ils ont pour objectif de mettre en évidence des groupes d'exemples similaires dans les données, ou de déterminer la distribution des données dans l'espace d'entrée, appelée estimation de la densité.

Dans le contexte de l'apprentissage, des observations $(\bar{x}, y)$, contenant des entrées $\bar{x}$ et des cibles y, sont utilisées comme ensemble de données d'apprentissage pour créer la meilleure fonction d'approximation correspondant à ces points, c'est-à-dire, la fonction qui représente au mieux y en fonction des $\bar{x}$.

Pour toutes les sections qui suivent, et sauf précision contraire, les notations suivantes sont utilisées :

* $\bar{x}$: vecteur regroupant les différentes variables du problème, $\bar{x} = (x_j | j \in [\![1, D]\!])$

> **Exemple** x_j peut être par exemple la pression, la vitesse, la température, etc. Pour deux variables $\bar{x}$ s'écrit,
>
> $$\bar{x} = (p, u) \tag{Ex. 4.1}$$

* y : la sortie ou la cible scalaire, et $\bar{y}$: le vecteur regroupant toutes les cibles de N_B points d'observations.
* $\bar{w}$: coefficients de la régression, $\bar{w} = (w_0, \ldots, w_C)^T$
* C : Nombre de coefficients
* M : ordre du polynôme, $M_{j=1,\ldots,D}$ pour D variables
* $\bar{\bar{X}}$: les entrées du vecteur colonne pour les N_B observations, $\bar{\bar{X}} = (\bar{x}_i | i \in [\![1, N_B]\!])$

> **Exemple** pour $\bar{x} = (p, u)$ la matrice de 6 points de données s'écrit,
>
> $$\bar{\bar{X}} = \begin{pmatrix} u_1 & p_1 \\ u_2 & p_2 \\ \vdots & \vdots \\ u_6 & p_6 \end{pmatrix} \tag{Ex. 4.2}$$

* $\bar{\phi}$: regroupe les fonctions de base, $\bar{\phi}(\bar{x}) = (\phi_0(\bar{x}), \ldots, \phi_C(\bar{x}))$, et représente la forme

> **Exemple** le vecteur $\bar{\phi}$ peut s'écrire par exemple $\bar{\phi}(\bar{x}) = \left(1, u, p, up, p^2\right)$ avec $\phi_0(\bar{x}) = 1$, $\phi_1(\bar{x}) = u$, $\phi_2(\bar{x}) = p$, etc.

* $\bar{\bar{\Phi}}$: la matrice qui regroupe les fonctions de base pour N_B points d'observations,

> **Exemple** $\bar{\bar{\Phi}}$ pour un ensemble de 6 points d'observation pour le même $\bar{x}$ que l'exemple (Ex. 4.1) en ordre 2 avec les fonctions de base $\bar{\phi}$,
>
> $$\bar{\bar{\Phi}} = \begin{pmatrix} 1 & u_1 & p_1 & u_1 p_1 & p_1^2 \\ 1 & u_2 & p_2 & u_2 p_2 & p_2^2 \\ 1 & u_3 & p_3 & u_3 p_3 & p_3^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & u_6 & p_6 & u_N p_6 & p_6^2 \end{pmatrix} \tag{Ex. 4.3}$$

De nombreux algorithmes peuvent être choisis selon l'objectif final, et selon le type de données mises à disposition. Autrement dit, selon la nature de l'entrée $\bar{x}$ avec laquelle l'algorithme commence à apprendre et la sortie y pour laquelle l'algorithme est programmé. Voici quelques exemples d'algorithmes d'apprentissage dans les cas de régression ou de classification :

4.1.1 Régression linéaire ou logistique [42]

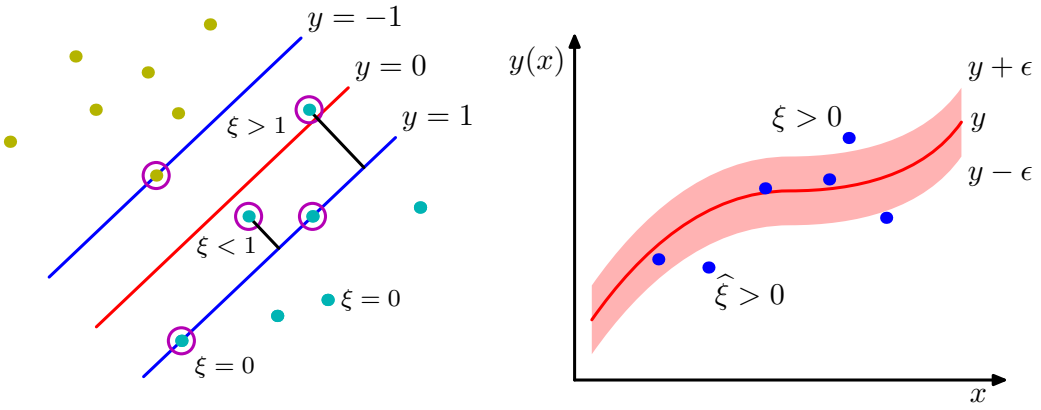
Cette technique est l'algorithme de régression le plus simple, basé sur un modèle. Le principe consiste à trouver les deux paramètres w_0 et w_1 reliant les N_B entrées $x_i, i = 1, \ldots, N_B$ à leurs N_B sorties $y_i, i = 1, \ldots, N_B$, en minimisant la somme des carrées des écarts ($\varepsilon^{(\text{lse})}$),

$$\varepsilon^{(lse)}(w_0, w_1) = \sum_{i=1}^{N_B} ((w_0 + w_1 x_i) - y_i)^2 \tag{4.4}$$

Cet algorithme est certes simple, mais il permet d'obtenir de bons résultats, en particulier lorsque les données suivent une relation linéaire.

4.1.2 Régression par machines à vecteurs de support [6, 12, 82]

La régression par machines à vecteurs de support (Support Vector Machine SVM) a été initialement établi pour les problèmes de classification, puis étendu pour traiter le cas de la régression. Cette méthode, contrairement à la précédente, permet de minimiser les erreurs sans tenir compte de certaines données aberrantes. Pour ce faire, il faut de pénaliser la norme de w_1 de manière à maximiser la distance entre l'hyperplan, $(y = 0)$ dans la figure 4.1, et chaque donnée d'apprentissage. Le raisonnement ci-après est valable pour les deux cas



(a) Cas de classification (b) Cas de régression

FIGURE 4.1 – Représentation montrant l'hyperplan (y : en rouge) et les marges fonctionnelles $y \pm \epsilon$ [6, 12].

(classification et régression). Prenons l'exemple de la classification par exemple. Dans la figure 4.1(a), l'hyperplan est la frontière entre deux classes différentes, représenté par $y = 0$. elle indique que la fonction de décision (Eq. 4.5) est nulle. Pour un cas $1D$, la fonction de décision est donnée par,

$$f(x) = w_0 + w_1 x_i \tag{4.5}$$

Lorsque la valeur de la fonction de décision est significativement élevée et positive, cela signifie que le vecteur des données est orienté dans la direction du vecteur normal de l'hyperplan. Les données sont alors situées du côté positif de la classe $y = 1$. À l'inverse, quand la valeur est notablement élevée et est négative, le vecteur des données est opposé à celle de la normale à l'hyperplan et les données se placent alors du côté de la classe $y = -1$. La frontière est établie par un paramètre w_0, qui définit le déplacement de l'hyperplan par rapport à l'origine. L'objectif est alors de déterminer les valeurs optimales des paramètres $\bar{w}$, afin de définir efficacement l'hyperplan qui sépare les deux classes.

Le problème d'optimisation associé à l'approche SVM à des fins d'optimisation par la méthode de descente de gradient peut être exprimée initialement par,

$$\min_{w_0,w_1} \frac{1}{2} \|w_1\|^2 \tag{4.6}$$

Il est possible de rencontrer un problème d'optimisation en utilisant ce modèle de descente gradient. Pour y remédier, il existe un modèle qui permet certaines erreurs de classification. Le concept est illustré dans la figure 4.1. Le principe est d'introduire une variable ξ associée à chaque paire de données (x_i, y_i). Cette procédure permet aux données de ne pas forcément appartenir à leur classification. Le problème d'optimisation après l'ajout de la contrainte d'erreur ξ prend la forme,

$$\begin{aligned}
&\text{En classification} \quad \min_{w_0,w_1,\xi_i,\hat{\xi}_i} \frac{1}{2} \|w_1\|^2 + C \sum_{i=1}^{N_B} \xi_i \\
&\qquad \text{tels que} \begin{cases} y_i - (w_1 x_i + w_0) \leq 1 - \xi_i \\ \xi_i, \hat{\xi}_i \geq 0 \end{cases} \\
&\text{En régression} \quad \min_{w_0,w_1,\xi_i,\hat{\xi}_i} \frac{1}{2} \|w_1\|^2 + C \sum_{i=1}^{N_B} \left(\xi_i + \hat{\xi}_i\right) \\
&\qquad \text{tels que} \begin{cases} y_i - (w_1 x_i + w_0) \leq \xi_i \\ y_i - (w_1 x_i + w_0) \geq \hat{\xi}_i \\ \xi_i, \hat{\xi}_i \geq 0 \end{cases}
\end{aligned} \tag{4.7}$$

où $C = 1/\lambda_r$ est une constante de régularisation, un hyperparamètre qui permet d'ajuster la distance entre l'hyperplan et les données d'apprentissage, qui alors contrôle le compromis entre la marge de tolérance et la classification erronée des données. Plus la valeur de C est élevée, plus la classification des données est rigoureuse, ce qui réduit les erreurs ed classification mais peut entraîner une marge plus étroite. N_B le nombre total d'entrées x_i, et ξ_i la distance du point de la catégorie à l'hyperplan correspondant.

4.1.3 Réseaux de neurones [6, 76]

Les algorithmes d'apprentissage basés sur des modèles incluent également les réseaux de neurones artificiels, qui sont inspirés du mode de fonctionnement des neurones biologiques. Le modèle du neurone peut être décrit par une série de transformations fonctionnelles. Il est caractérisé par un état interne, des données d'entrée $\bar{\bar{X}} = (\bar{x}_i | i \in [\![1, N_B]\!])$ (pour un cas $1D$

par exemple, $\bar{x} = x$), et une fonction d'activation notée h. La fonction d'activation réalise une transformation d'une combinaison affine a_j des données d'entrée [6].

$$f(x, \bar{w}) = h(a_j) = h\left(w_0 + \sum_{i=1}^{N_B} w_i x_i\right) \tag{4.8}$$

Concrètement, un réseau de neurones artificiels est composé de plusieurs couches de neurones [76]. Trois types de couches peuvent être distingués (Fig. 4.2) :

— **Couches d'entrée :** Cette couche sert uniquement à introduire les données d'entrée, x_i, dans le réseau. À ce niveau, le nombre de neurones est équivalent à la dimension du vecteur d'entrée, $\bar{\bar{X}}$.

— **Couches cachées :** Il peut y avoir plusieurs couches selon la complexité du problème. Ces différentes couches constituent le centre du réseau de neurones. La fonction d'activation h est associée à chaque neurone de ces couches. Le choix de la fonction d'activation est déterminé par la nature des données et la distribution supposée des variables cibles. Si deux couches sont présentes comme sur l'exemple de la figure 4.2 , la sortie, y_i^j de ce neurone s'écrira :

$$y_k(\bar{x}, \bar{w}) = h'\left(\sum_{j=1}^{M} \bar{w}_{j,k}^{(2)} h\left(\sum_{i=1}^{N_B} \bar{w}_{i,j}^{(1)} + \bar{w}_{i0}^{(1)}\right) + \bar{w}_{j0}^{(2)}\right) \tag{4.9}$$

Les exposants $^{(1)}$ et $^{(2)}$, précisent la couche dans laquelle les paramètres se trouvent.

— **Couches de sortie :** Cette couche correspond au vecteur de sortie de prédiction $\bar{y}$ obtenu après une série de transformations des données d'entrée.

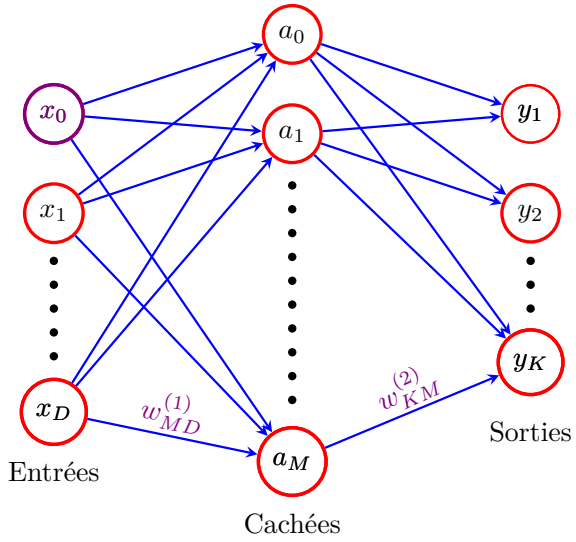


FIGURE 4.2 – Schéma d'un réseau de neurones à deux couches [6]. Les variables d'entrée, cachées et de sortie sont représentées par des nœuds, et les paramètres de poids (coefficients) sont représentés par des liens entre les nœuds.

4.1.4 Méthodes de k plus proches voisins (k nearest neighbors, kNN) [4]

Brièvement, la méthode des k plus proches voisins est l'une des méthodes d'apprentissage basées sur l'observation. L'objectif est toujours celui de prédire une sortie $y_\star$ associée à une nouvelle entrée $x_\star$. Cette dernière se fait en deux étapes :

— (i) les k entrées les plus proches de x sont sélectionnées dans l'ensemble d'apprentissage.

— (ii) la moyenne des k sorties associées aux k entrées sélectionnées est calculée.

Il existe également une approche probabiliste d'un plus proche voisin (P1NN)[82]. Elle est une extension simple de la méthode du plus proche voisin qui fournit des prédictions probabilistes. Brièvement, elle calcule la prédiction d'un plus proche voisin sur l'ensemble d'apprentissage et enregistre la fraction des cas où les prédictions étaient correctes.

Le choix de l'algorithme d'apprentissage est basé sur les critères les plus pratiques pour notre cas. Avec l'utilisation de la dynamique moléculaire, les variables macroscopiques (telles que : vitesse, densité et température, etc.) sont calculées par le biais de la prise de moyenne. Cette démarche génère des incertitudes sur les variables macroscopiques extraites. De plus, comme nous le verrons par la suite lors de la procédure de couplage, la condition aux limites transmise aux équations de Navier-Stokes devra être sous la forme d'une fonction implicite. L'utilisation de méthodes bayésiennes permet d'identifier une fonction reliant les entrées à la réponse.

Compte tenu de ces différentes considérations, la solution qui semble la plus adaptée est l'utilisation de la régression par apprentissage. Cette approche permet à la fois de prendre en compte l'incertitude et de construire une fonction implicite à partir de la régression. Cela justifie donc le choix qui a été retenu. Dans la suite, seules les méthodes de régression (bayésiennes et par processus gaussiens) sont détaillées.

4.2 Régression linéaire

L'ensemble $\mathcal{D}$ considéré par la suite de ce travail comprend N_B observations et peut s'écrire sous la forme $\mathcal{D} = \{(\bar{x}_i, y_i)\,|i \in [\![1, N_B]\!]\}$, où $\bar{x}$ désigne un couple d'entrée (de plusieurs variables) de dimension D, y une variable dépendante représentant une sortie ou une cible scalaire. Les entrées du vecteur colonne pour les N_B cas sont regroupés dans la matrice $\bar{\bar{X}}$ de dimension $D \times N_B$, et les cibles sont rassemblées dans le vecteur $\bar{y}$, cela permet de réécrire l'ensemble d'apprentissage comme $\mathcal{D} = \left(\bar{\bar{X}}, \bar{y}\right)$.

Il est important de noter qu'afin de pouvoir mieux trouver les prédictions, pour les deux approches bayésienne et gaussienne, respectivement les bases de données / les données sont normalisées en utilisant des méthodes spécifiques, plus de détails dans la section (Sec. 4.3).

4.2.1 Régression par la méthode des moindres carrées

Dans un premier temps, une approche basée sur la méthode des moindres carrées (LSE) est examinée. Considérons des sorties y_i générées à partir d'une fonction connue, $y_i = g(x_i)$ avec,

pour l'exemple, $g(x) = -\frac{3}{2}x - \frac{1}{7}x^2$ dans la suite. De plus, un bruit aléatoire ϵ est inclus dans les valeurs y_i. Il peut, par exemple, provenir de processus d'intégration de la méthode de dynamique moléculaire qui permet de remonter aux propriétés macroscopiques, en calculant les moyennes ou des fluctuations déjà existantes.

Un ajustement des données pour D variables notés $\bar{x}$ a la forme,

$$y = f(\bar{x}, \bar{w}) + \varepsilon(\bar{x}) \tag{4.10}$$

$$f(\bar{x}, \bar{w}) = \sum_{k=0}^{C} w_k \phi_k(\bar{x}) \tag{4.11}$$

où $\bar{w} = (w_0, \ldots, w_C)^T$ représente l'ensemble des coefficients du polynôme, M est l'ordre de ce polynôme. Ce polynôme dépend ici que d'une seule variable $\bar{x}$ qui représente l'ensemble des données d'entrée d'apprentissage. La fonction polynomiale $f(\bar{x}, \bar{w})$ est certes une fonction non linéaire de $\bar{x}$, mais elle est une fonction linéaire des coefficients $\bar{w}$, le modèle est par conséquent considéré comme linéaire, de part sa dépendance linéaire aux coefficients $\bar{w}$.

Les $\bar{\phi}_k(\bar{x})$ sont des fonctions de base dépendantes des variables $\bar{x}$, $\bar{x}$ désigne les différentes variables qui sont par exemple la pression p et la vitesse u, dans ce cas $\bar{x} = \{p, u\}$, $\bar{\phi}(p, u) = \bar{\phi}(\bar{x}) = (\phi_0(\bar{x}), \ldots, \phi_C(\bar{x}))^T$, avec C le nombre de coefficients. Ces fonctions de base transforment un vecteur d'entrée $\bar{x}$ de dimension D dans un espace d'entrée de dimensions N. Typiquement, $\phi_0(\bar{x}) = 1$, de sorte que w_0 agit comme un biais. D'autres fonctions de base existent notamment en sciences expérimentales, les plus classiques sont : splines, bases de Fourier, gaussienne, etc. Pour les N_B points d'apprentissage, la matrice regroupant les $\bar{\phi}(\bar{x})$ est donnée par $\bar{\bar{\Phi}}$ (Eq. (4.12)) et est le regroupement des colonnes $\bar{\phi}(\bar{x})$ pour toutes les variables et toutes les données d'apprentissage. Elle contient alors les valeurs des fonctions de bases $\bar{\phi}_i^{(j)}$ pour chaque point de données et elle est de dimension $N_B \times C$, avec généralement $N_B \gg C$.

$$\bar{\bar{\Phi}} = \begin{pmatrix} \phi_0(\bar{x}_1) & \phi_1(\bar{x}_1) & \cdots & \phi_M(\bar{x}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(\bar{x}_i) & \phi_1(\bar{x}_i) & \cdots & \phi_M(\bar{x}_i) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(\bar{x}_{N_B}) & \phi_1(\bar{x}_{N_B}) & \cdots & \phi_M(\bar{x}_{N_B}) \end{pmatrix} \tag{4.12}$$

Bien corréler le polynôme avec les données d'apprentissage se traduit par le fait de minimiser l'erreur (fonction coût) qui mesure le mauvais ajustement entre la fonction $f(\bar{x}, \bar{w})$, pour $\bar{w}$ donnée et les points de données de l'ensemble d'apprentissage.

$$\varepsilon^{(lse)}(\bar{w}) = \frac{1}{2} \sum_{i=1}^{N_B} \left(f(\bar{x}_i, \bar{w}) - y_i \right)^2 \tag{4.13}$$

Cette fonction coût donnée par (Eq. 4.13) est une quantité non négative qui ne serait nulle que si, et seulement si, la fonction $f(\bar{x}, \bar{w})$ passait exactement par chaque point de données d'apprentissage. D'une manière générale, plus la valeur de cette fonction coût $\varepsilon^{(lse)}(\bar{w})$ est petite, meilleure est l'approximation.

La fonction coût (Eq. 4.13) étant une fonction quadratique des coefficients $\bar{w}$, ses dérivées par rapport aux coefficients seront alors linéaires par rapport à $\bar{w}$. Ce qui signifie que la minimisation de cette fonction coût, $(\nabla_{\bar{w}}\varepsilon^{(lse)}(\bar{w}) = 0)$ a une solution unique donnée par,

$$\bar{w} = (\bar{\bar{\Phi}}^T\bar{\bar{\Phi}})^{-1}\bar{\bar{\Phi}}^T\bar{y} \tag{4.14}$$

La figure 4.3 montre un graphique d'un ensemble d'entraînement comprenant $N = 20$ points de données générées par une fonction connue. Elle illustre trois exemples de résultats d'approximation polynomiale à des ordres $M = 1, 2$, et 17 de cet ensemble.

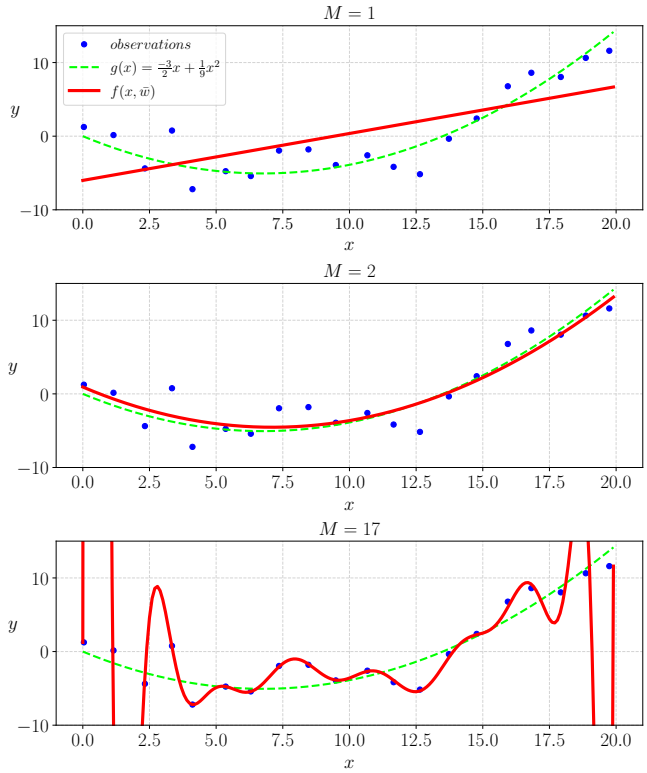


FIGURE 4.3 – Polynômes de différents ordres M en rouge qui corrèlent des données (points bleus) tirées à partir d'une fonction quadratique $g(x)$ en pointillé en vert.

D'après la figure 4.3, le polynôme d'ordre $(M = 2)$ semble donner la meilleure corrélation de la courbe verte. En utilisant un ordre plus élevé, nous pouvons avoir plus de capacité à ajuster des points de données plus complexes.

Il faut noter que dans le cas d'un polynôme d'ordre plus élevé (par exemple $M = 17$), la corrélation aux données d'apprentissage est excellente. En effet, le polynôme passe pratiquement par chaque point de données d'apprentissage, ce qui signifie que la fonction d'erreur tend vers zéro $\varepsilon^{(lse)}(\bar{w}) \rightarrow 0$. Cependant, la courbe obtenue par la régression présente de grandes oscillations, notamment près des extrémités, et donne ainsi une très mauvaise représentation de la fonction, en particulier des nouvelles données, par exemple les $x = \{0.5, 1.5, 19, \dots\}$. Ce

dernier comportement est connu sous le nom du sur-apprentissage et il n'est pas souhaitable. Il est possible d'avoir un aperçu du problème en examinant les valeurs des coefficients $\bar{w}$ obtenus à partir de polynômes de divers ordres. En effet, la figure 4.4 montre que lorsque M augmente, l'amplitude des coefficients devient plus grande.

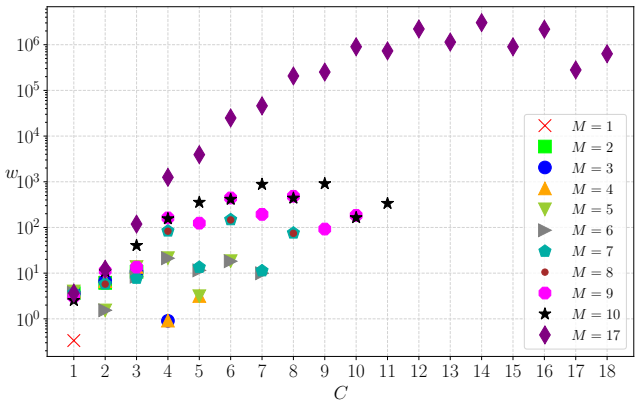


FIGURE 4.4 – Amplitudes des de différents ordres M en rouge qui corrèlent des données (points noirs) tirées à partir d'une fonction quadratique en pointillé en vert.

Il est à souligner que le problème de sur-apprentissage devient moins grave à mesure que la taille de l'ensemble de données augmente (Fig. 4.5). Autrement dit, plus l'ensemble de données est grand, plus le modèle que l'on peut utiliser pour approximer les données peut être complexe [6].

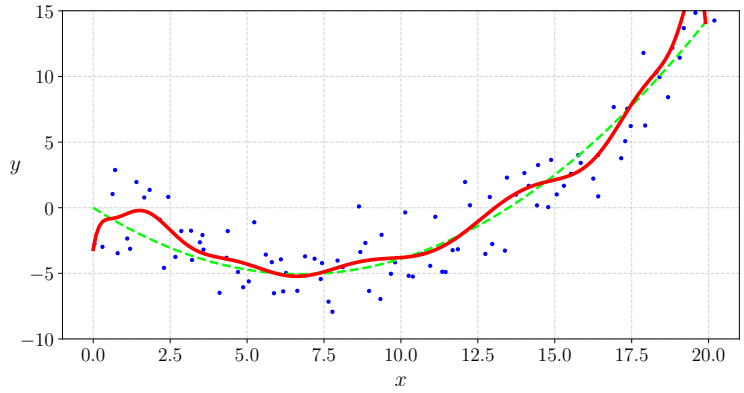


FIGURE 4.5 – Graphique d'un ensemble d'entraînement comprenant $N = 100$ points de données. la prédiction est réalisée avec un polynôme d'ordre 20.

Pour éviter le problème du sur-apprentissage, il existe une technique de régularisation appelée, Ridge (RLS) (Ridge Lest Square) [41]. En général, une fonction est susceptible d'être sur-ajustée lorsque ses coefficients $\bar{w}$ (valeurs de pondération) ont une grande valeur et ne sont pas bien distribués. Par conséquent, il faut forcer le processus de formation à réduire ces coefficients en ajoutant un terme de pénalisation à la fonction d'erreur, Eq. (4.15). Ce processus rend également ces coefficients mieux distribués.

La fonction d'erreur correspondante à cette technique de régularisation s'écrit :

$$\varepsilon^{(rls)}(\bar{w}) = \frac{1}{2}\left[\underbrace{\sum_{i=1}^{N}\left(f\left(x_i,\bar{w}\right)-y_i\right)^2}_{[T_1]} + \underbrace{\lambda_r \sum_{j=1}^{M-1} w_j^2}_{[T_2]}\right] \tag{4.15}$$

où le terme $[T_1]$ permet de rendre la valeur de l'erreur carrée aussi petite possible. Le terme $[T_2]$, quant à lui, sert à contraindre les coefficients à être plus petite. La figure 4.6 confirme ce comportement. Elle illustre justement une comparaison entre les valeurs des coefficients avec la méthode des moindres carrés LSE et celle de régularisation RLS. Cette figure montre alors que la valeur des coefficients est largement plus petite dans le cas de RLS par rapport au cas de LSE. Ce second terme $[T_2]$ est appelé la régularisation dans laquelle la constante λ_r

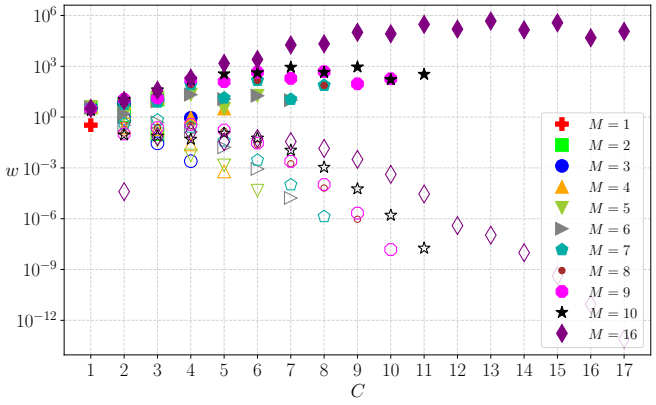


FIGURE 4.6 – Amplitudes des coefficients : Comparaison pour différents ordres $M = [\![1, 16]\!]$ entre les deux cas de régression LSE (symboles pleins) et RLS (symboles vides).

permet de déterminer l'ampleur de la régularisation. Le choix de la bonne valeur λ_r s'avère être délicat, il faut alors essayer plusieurs valeurs de λ_r et choisir celle qui donne la meilleure approximation et donc la meilleure prédiction. La figure 4.7 montre l'influence du choix de

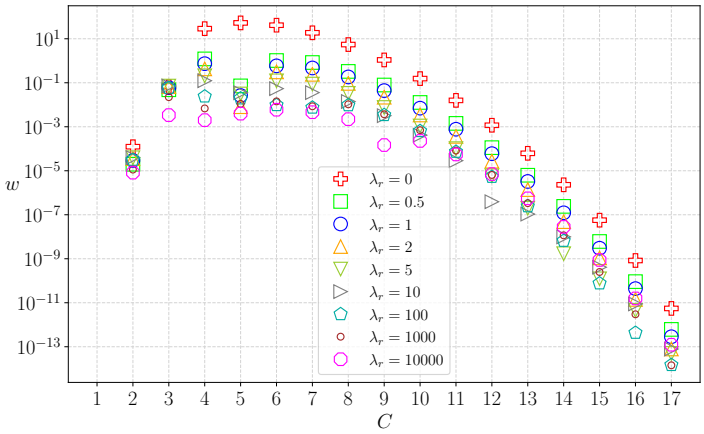


FIGURE 4.7 – Amplitudes des coefficients pour différentes valeurs λ_r, cas $M = 16$.

la valeur de la constante de régularisation λ_r sur la valeur des coefficients. en effet, pour des ordres moins élevés $M = [1, 2, \ldots, 10]$, en augmentant la valeur de λ_r, la valeur de w diminue. En revanche, pour des ordres plus élevés encore, à $\lambda_r = 10$ les coefficients ont de plus petites valeurs qu'à des λ_r plus grandes. Cette procédure est alors à la fois laborieuse et difficile, du fait qu'il faille essayer plusieurs valeurs pour avoir une approximation jugée meilleure.

En adoptant une approche bayésienne, non seulement le problème du sur-apprentissage peut être évité, mais encore, elle permet de s'affranchir du choix de la constante de régularisation λ_r de l'équation (4.15). En outre, dans un modèle de régression bayésienne, le nombre effectif de paramètres s'adapte systématiquement au nombre de points de données. En d'autres termes, il n'est pas exigé d'avoir au moins autant de points que de coefficients comme c'est le cas pour les moindres carrés. C'est alors plus accessible à employer des modèles pour lesquels le nombre de paramètres dépasse largement le nombre de points de données. Cette approche est plus détaillée dans la section suivante (Sec. 4.2.2).

4.2.2 Régression par la méthode bayésienne

Rappel notations

* $\bar{x}$ désigne un couple d'entrée (de plusieurs variables) de dimension D, $\bar{x} = (x_j | j \in [\![1, D]\!])$
* y une variable dépendante représentant une sortie ou une cible scalaire
* $\bar{\bar{X}}$ matrice de dimension $D \times N_B$, regroupant toutes les entrées.
* $\bar{y}$: vecteur rassemblant toutes les cibles de N_B points d'observations.

En probabilité, il existe généralement deux types d'approches de raisonnement : fréquentiste et bayésienne. Les statistiques fréquentistes supposent que les données sont un échantillon aléatoire et visent à identifier les paramètres fixes qui ont généré les données. Les statistiques bayésiennes, quant à elles, considèrent les données telles qu'elles et les paramètres comme des variables aléatoires avec une distribution qui peut être déduite des données. Par conséquent, les approches fréquentistes nécessitent au moins autant de points de données que de paramètres à estimer. Les approches bayésiennes, en revanche, sont compatibles avec des ensembles de données plus petits et bien adaptés à l'apprentissage à partir d'un échantillon à la fois.

Le théorème, élaboré par *Bayes*, utilise la théorie fondamentale des probabilités pour prescrire comment les probabilités ou les croyances doivent changer à mesure que de nouvelles informations pertinentes arrivent. L'inférence dans le modèle linéaire bayésien est fondée sur

la distribution *a posteriori* sur les poids, calculée par le théorème de Bayes (4.16).

$$\underbrace{p(\bar{w}|\bar{y},\bar{\bar{X}})}_{a\ posteriori} = \frac{\overbrace{p(\bar{y}|\bar{\bar{X}},\bar{w})}^{\text{vraisemblance}}\ \overbrace{p(\bar{w})}^{a\ priori}}{\underbrace{p(\bar{y}|\bar{\bar{X}})}_{\text{évidence = régularisation}}} \tag{4.16}$$

La distribution *a priori*, $p(\bar{w})$, indique la probabilité que nous considérons chaque hypothèse possible. Elle traduit également les hypothèses que nous formulons sur les coefficients $\bar{w}$, avant d'observer les données. Si la distribution *a priori* n'est pas connue avec certitude, il est impératif de faire un choix, souvent parmi plusieurs options raisonnables. En général, les bonnes pratiques consistent à justifier le *a priori* et à vérifier la robustesse en testant si les alternatives conduisent à la même conclusion. La fonction vraisemblance $p(\bar{y}|\bar{\bar{X}},\bar{w})$ produit la probabilité d'observer un ensemble de données lorsque certaines valeurs des paramètres $\bar{w}$ sont données, c'est-à-dire pour une hypothèse spécifique. L'évidence, également connue sous le nom de vraisemblance marginale, mesure la probabilité des données observées, compte tenu de toutes les hypothèses possibles. Par conséquent, elle est la même pour toutes les valeurs des paramètres et sert à normaliser et régulariser le numérateur.

La vraisemblance n'est pas une distribution de probabilité sur $\bar{w}$, et son intégrale par rapport à $\bar{w}$ n'est pas forcément égale à l'unité [70]. Compte tenu de cette définition de la vraisemblance, le théorème de Bayes peut être exprimé en quelques mots : la distribution *a posteriori* d'un paramètre sachant qu'on observe les données est proportionnelle à la distribution *a priori* sur ce paramètre avant d'avoir observé les données, multipliée par la vraisemblance.

$$\text{distribution } a\ posteriori \ \propto\ \text{vraisemblance} \ \times\ a\ priori \tag{4.17}$$

L'estimation de la probabilité *a posteriori* maximale (MAP) s'appuie sur le fait que l'évidence est un facteur constant qui met à l'échelle la distribution *a posteriori* pour répondre aux exigences d'une distribution de probabilité. Puisque l'évidence ne dépend pas de $\bar{w}$, la distribution *a posteriori* est proportionnelle au produit de la vraisemblance et de l'*a priori*. Par conséquent, l'estimation MAP choisit la valeur de w qui maximise la distribution *a posteriori* étant donné les données observées et la croyance *a priori*. Contrairement à la MAP, l'estimation du maximum de vraisemblance (MLE) sélectionne la valeur de paramètre w qui maximise la fonction de vraisemblance pour les données d'apprentissage observées.

$$\bar{w}_{\text{MLE}} = \arg\max_{\bar{w}}\ p(y|\bar{\bar{X}},\bar{w}) \tag{4.18}$$

Les distributions *a priori*, $p(\bar{w})$, sont un ingrédient essentiel des modèles bayésiens. Elles doivent refléter les connaissances sur la distribution des paramètres.

Une distribution *a priori* est conjuguée par rapport à la vraisemblance lorsque la distribution *a posteriori* résultante est de la même classe ou famille que la distribution *a priori*. Par exemple, lorsque l'*a priori* et la vraisemblance suivent une distribution normale, la distribution *a posteriori* suit également une distribution normale.

La distribution *a priori* et de vraisemblance conjuguée implique une solution de forme fermée pour la distribution *a posteriori*, ce qui facilite le processus de mise à jour et évite le besoin d'utiliser des méthodes numériques pour approcher l'*a posteriori*. De plus, la distribution *a posteriori* résultante peut être utilisée comme *a priori* pour la prochaine étape de mise à jour.
Pour un modèle complexe, la distribution *a priori* est utilisée pour pénaliser l'éloignement d'un paramètre par rapport aux valeurs plausibles pour le modèle étudié. Elle sert ainsi à réduire les coefficients et à les rendre mieux distribués. Il est important de souligner que l'influence de la distribution diminue lorsque le nombre d'observations augmente.

4.2.2.1 Modèle de régression linéaire bayésienne

Dans la section précédente (Sec. 4.2.1), il était question de dériver une régression polynomiale à l'aide de la méthode des moindres carrés. Dans cette section, la discussion sera portée sur la régression linéaire d'un point de vue *bayésien*.

La méthode des moindres carrés (LSE) est une approche visant à ajuster une courbe ou une ligne aux données en minimisant la somme des carrés des écarts entre les points de données et la courbe. Les méthodes de régression bayésiennes, en revanche, utilisent les lois de probabilité bayésiennes pour estimer les paramètres de la courbe ou de la ligne de régression, en tenant compte de l'incertitude et du biais potentiel des données.

Il faut savoir que le résultat final de la dérivation de la régression en utilisant (LSE) est égal au résultat de la dérivation de la régression linéaire en utilisant l'estimation maximale de la vraisemblance MLE dans la méthode bayésienne ($\bar{w}_{MLE} = \bar{w}_{LSE}$). En outre, le résultat de la dérivation de la régression, en utilisant la méthode de régularisation (RLS) (Eq. 4.15), est égal au résultat de la dérivation en utilisant le maximum *a posteriori* (MAP) dans la méthode bayésienne ($\bar{w}_{MAP} = \bar{w}_{RLS}$). La différence est que le paramètre de pénalisation $\lambda_r = \alpha/\beta$ est mis à jour itérativement au cours de l'apprentissage, ce qui est appelé par apprentissage des hyperparamètres (Sec. 4.2.2.2). La démonstration de ces égalités est faite en annexe B.

Pour mieux comprendre le concept d'apprentissage, cette section passe en revue l'analyse bayésienne du modèle standard de régression linéaire avec un bruit gaussien.
Rappelons que l'ensemble $\mathcal{D} = \{(\bar{x}_i, y_i) \,|i = 1, \ldots, N_B\}$, comprend N_B observations, $\bar{x}$ désigne un couple d'entrée de dimension D et y représente la sortie/cible scalaire. $\bar{\bar{X}}$ est une matrice de dimension $D \times N_B$, et le vecteur $\bar{y}$ rassemble toutes les cibles, l'ensemble d'apprentissage peut se réécrire finalement $\mathcal{D} = \left(\bar{\bar{X}}, \bar{y}\right)$.

L'objectif est de faire des inférences sur la relation entre les entrées et les cibles, c'est-à-dire la distribution conditionnelle des cibles compte tenu des entrées. Pour cela, il suffit de déterminer les coefficients ou vecteur paramètre $\bar{w} = (w_0, \ldots, w_C)^T$ afin de pouvoir prédire $f(\bar{x}_\star, \bar{w})$ pour une nouvelle entrée $\bar{x}_\star$, noté simplement $f_\star \triangleq f(\bar{x}_\star, \bar{w})$ et illustré dans la

figure 4.8. Le modèle de régression linéaire standard avec bruit gaussien est donné par,

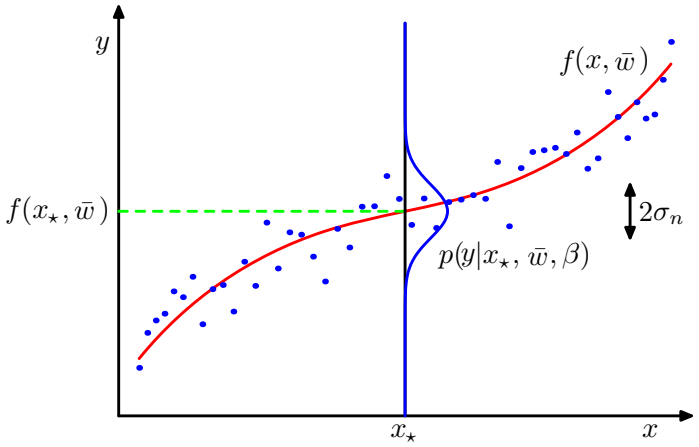


FIGURE 4.8 – Illustration synthétique, cas $1D$, d'une distribution conditionnelle gaussienne pour y compte tenu des données $\bar{x} = x$ (points *bleus*). Cette distribution est donnée par l'équation (4.20), dans laquelle la prédiction de régression est la moyenne de celle-ci présentée par la fonction polynomiale $f(x, \bar{w})$ (la courbe *rouge*), et la précision est donnée par le paramètre β. La gaussienne a une valeur plus élevée lorsqu'elle se rapproche de la courbe rouge ou des données et une valeur plus faible lorsqu'elle est plus éloignée. L'objectif est de maximiser la probabilité de y, ce qui revient à encourager la prédiction à s'adapter aux données.

$$f(\bar{x}, \bar{w}) = \bar{\bar{X}}^T \bar{w} \quad ; \qquad y = f(\bar{x}, \bar{w}) + \varepsilon \tag{4.19}$$

où $\bar{w}$ le vecteur paramètre du modèle. La variable cible des observations est représentative du modèle $f(\bar{x}, \bar{w})$ avec l'ajout d'un bruit aléatoire, ε, qui est supposé être régi d'un processus gaussien de moyenne nulle et de variance σ_n^2, ce qui revient à dire que $\varepsilon \sim \mathcal{N}\left(0, \sigma_n^2\right)$ [6, 82].

L'hypothèse de ce bruit, qui accompagne le modèle, engendre une conséquence directe qui est la vraisemblance, c'est-à-dire la densité de probabilité des observations compte tenu des paramètres. Elle est donnée par l'équation (4.20) et est présentée dans la figure 4.8,

$$p(y|\bar{x}, \bar{w}, \beta) = \mathcal{N}(y|f(\bar{x}, \bar{w}), \beta^{-1}) \tag{4.20}$$

où $\beta = \sigma_n^{-2}$, est l'inverse de la variance. Cette vraisemblance est le produit de tous les cas dans l'ensemble d'apprentissage (en raison de l'hypothèse d'indépendance). La probabilité de vraisemblance pour N_B points s'écrit alors,

$$\begin{aligned} p(\bar{y}|\bar{\bar{X}}, \bar{w}) &= \prod_{i=1}^{N_B} p(y_i|\bar{x}_i, \bar{w}) \\ &= \prod_{i=1}^{N_B} \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{1}{2\sigma_n^2}\left(y_i - \bar{x}_i^T \bar{w}\right)^2\right) \\ &= \frac{1}{(2\pi\sigma_n^2)^{N_B/2}} \exp\left(-\frac{1}{2\sigma_n^2}\left|\bar{y} - \bar{\bar{X}}^T \bar{w}\right|^2\right) \\ &= \mathcal{N}(\bar{\bar{X}}^T \bar{w}, \sigma_n^2 \bar{\bar{I}}) \end{aligned} \tag{4.21}$$

Nous rappelons que $\left\{\bar{\bar{X}} = \bar{x}_i | i \in [\![1, N_B]\!]\right\}$, qui est de dimension $D \times N_B$, regroupe tous les vecteurs des variables dont dépend la réponse. Dans le cadre du concept bayésien, il est nécessaire de spécifier une distribution *a priori* sur les paramètres, de manière à exprimer les croyances sur les paramètres avant d'examiner les observations et à limiter la complexité et donc la croissance des coefficients (poids $\bar{w}$). Pour y parvenir, cette probabilité *a priori* doit être d'une forme particulière afin de simplifier le traitement. Des fonctions plus lisses et donc moins complexes sont privilégiées. Nous utilisons un *a priori* de type d'une gaussienne de moyenne nulle avec une matrice de covariance ou de précision $\Sigma_p = \alpha^{-1}\bar{\bar{I}}$ sur les poids : $\bar{w} \sim \mathcal{N}(\bar{0}, \Sigma_p)$; où α décrit l'inverse de la variance et donc la précision sur la distribution des coefficients $\bar{w}$. Il contrôle alors la distribution des paramètres du modèle. Il est également utilisé afin de faciliter la dérivation. Cet hyperparamètre α, est optimisé ou appris au cours des calculs. Les détails sur l'apprentissage de cet hyperparamètre sont donnés dans l'annexe B.2.

La distribution *a posteriori* sur tous les paramètres inconnus est proportionnelle au produit de la fonction de vraisemblance et de la distribution *a priori*

$$p(\bar{w}|\bar{\bar{X}}, \bar{y}) \propto \; p(\bar{y}|\bar{\bar{X}}, \bar{w}) \;\; p(\bar{w}) \tag{4.22}$$

En conséquence du choix d'une distribution *a priori* gaussienne conjuguée, la distribution *a posteriori* sera également gaussienne

$$\begin{aligned} p(\bar{w}|\bar{\bar{X}}, \bar{y}) &\propto \exp\left(-\frac{1}{2\sigma_n^2}(\bar{y} - \bar{\bar{X}}^T\bar{w})^T(\bar{y} - \bar{\bar{X}}^T\bar{w})\right)\exp\left(-\frac{1}{2}\bar{w}^T\Sigma_p^{-1}\bar{w}\right) \\ &\propto \exp\left(-\frac{1}{2}(\bar{w} - \bar{m}_N)^T(\frac{1}{\sigma_n^2}\bar{\bar{X}}\bar{\bar{X}}^T + \Sigma_p^{-1})(\bar{w} - \bar{m}_N)\right) \end{aligned} \tag{4.23}$$

où, $\bar{m}_N = \beta\bar{\bar{S}}_N\bar{\bar{X}}\bar{y}$, avec, $\bar{\bar{S}}_N = (\beta\bar{\bar{X}}\bar{\bar{X}}^T + \Sigma_p^{-1})^{-1}$, la matrice de covariance pour la distribution *a posteriori*.

$$p(\bar{w}|\bar{\bar{X}}, \bar{y}) = \mathcal{N}\left(\bar{m}_N = \beta\bar{\bar{S}}_N\bar{\bar{X}}\bar{y}, \bar{\bar{S}}_N\right) \tag{4.24}$$

Il est à souligner par ailleurs que, étant donné que la distribution *a posteriori* $p(\bar{w}|\bar{\bar{X}}, \bar{y})$ est gaussienne, sa moyenne correspond également à son mode [6]. Par conséquent, le vecteur de paramètre *a posteriori* est obtenu par l'estimation maximale a posteriori (MAP) $\bar{w}_{\text{MAP}} = \bar{m}_N$. En considérant une distribution *a priori* infiniment grande $\bar{\bar{S}}_0 = \alpha^{-1}\bar{\bar{I}}$, avec α tendant vers 0, la moyenne de la distribution *a posteriori* $\bar{m}_N = \bar{w}_{ML}$. Cette moyenne est ensuite réduite à la valeur de vraisemblance maximale obtenue en fixant le gradient de la vraisemblance logarithmique à zéro et en résolvant l'équation (B.4) par rapport à $\bar{w}$.

Contrairement au raisonnement dans LSE, la prédiction, dans le cas bayésien, est donnée par la moyenne sur toutes les valeurs possibles des paramètres, pondérées par leur probabilité *a posteriori*. Ainsi, la distribution prédictive est donnée en faisant la moyenne de la sortie de

tous les modèles linéaires possibles par rapport à la distribution *a posteriori*.

$$\begin{aligned} p(f_\star|\bar{x}_\star, \bar{\bar{X}}, \bar{y}) &= \int p(f_\star|\bar{x}_\star, \bar{w})\ p(\bar{w}|\bar{\bar{X}}, \bar{y}) d\bar{w} \\ &= \mathcal{N}\left(\beta \bar{x}_\star^T \bar{\bar{S}}_N \bar{\bar{X}} \bar{y}, \bar{x}_\star^T \bar{\bar{S}}_N \bar{x}_\star\right) \end{aligned} \tag{4.25}$$

Par conséquent, cette distribution prédictive est à nouveau gaussienne, de moyenne obtenue en multipliant la moyenne postérieure des poids (B.8) et la nouvelle entrée $\bar{x}_\star$, et la variance prédictive résulte de la combinaison de la matrice de covariance postérieure et de la forme quadratique de la nouvelle entrée $\bar{x}_\star^T \bar{x}_\star$. En analysant cette formule, on peut s'attendre à ce que les incertitudes prédictives augmentent avec l'amplitude de la nouvelle entrée $\bar{x}_\star$.

Il est possible de surmonter le problème, lorsqu'il s'agit de relations non linéaires, par exemple celle dans l'équation (4.25). Cela se fait en projetant d'abord les entrées dans un espace à dimension plus élevée en utilisant un ensemble de fonctions de base $\phi(x)$, puis en appliquant le modèle linéaire dans cet espace plutôt que directement sur les entrées elles-mêmes [6, 70, 82].

$$f(\bar{x}, \bar{w}) = \sum_{j=0}^{M} w_j \phi_j(x) = \bar{\phi}(x)^T \bar{w}, \qquad y = f(\bar{x}, \bar{w}) + \varepsilon \tag{4.26}$$

où, $\bar{\phi}(\bar{x}) = (\phi_0(x), \ldots, \phi_M(x))^{\mathrm{T}}$ est appelée la fonction de base et transforme un vecteur d'entrée $\bar{x}$ de dimension D dans un espace d'entrée de dimensions N. Par exemple, l'entrée scalaire x pourrait être projetée dans l'espace des puissances de x et sera donnée dans le cas de la régression polynomiale par, $\bar{\phi}(x) = \left(1, x, x^2, \ldots, x^M\right)^{\mathrm{T}}$, avec M le degré polynomial.

La figure 4.9 représente le même exemple de la section 4.2.1 en utilisant cette fois-ci les méthodes de régression polynomiale bayésiennes. Cette figure montre que le problème de sur-apprentissage est résolu. Elle illustre également qu'il existe de nombreux autres choix possibles pour les fonctions de base, un autre exemple est celui des fonctions de base gaussiennes : $\phi_j(x) = \exp\left\{-\frac{(x-\mu_j)^2}{2s^2}\right\}$. En outre, la matrice $\bar{\bar{\Phi}}(\bar{x}_i)$ est le regroupement des colonnes $\bar{\phi}(\bar{x})$ pour toutes les variables et toutes les données d'apprentissage.

Par analogie au modèle précédent, la distribution prédictive donnée par l'Eq. (4.25) peut prendre une forme cette fois-ci en substituant la matrice de conception $\bar{\bar{X}}$ par la matrice $\bar{\bar{\Phi}}(X)$.

$$\begin{aligned} p(f_\star|\bar{x}_\star, \bar{\bar{X}}, \bar{y}) &= \mathcal{N}\left(\beta \bar{\phi}(\bar{x}_\star)^T \bar{\bar{S}}_N \bar{\bar{\Phi}} \bar{y}, \bar{\phi}(\bar{x}_\star)^T \bar{\bar{S}}_N \bar{\phi}(\bar{x}_\star)\right) \\ &= \mathcal{N}\left(\bar{m}_N^T \bar{\phi}(\bar{x}_\star), \sigma_N^2(\bar{x}_\star)\right) \end{aligned} \tag{4.27}$$

avec,

$$\bar{\bar{S}}_N^{-1} = \alpha \bar{\bar{I}} + \beta \bar{\bar{\Phi}}^{\mathrm{T}} \bar{\bar{\Phi}} \tag{4.28a}$$

$$\bar{m}_N = \beta \bar{\bar{S}}_N \bar{\bar{\Phi}}^{\mathrm{T}} \bar{y} \tag{4.28b}$$

$$\sigma_N^2(\bar{x}_\star) = \frac{1}{\beta} + \bar{\phi}(\bar{x}_\star)^T \bar{\bar{S}}_N \bar{\phi}(\bar{x}_\star) \tag{4.28c}$$

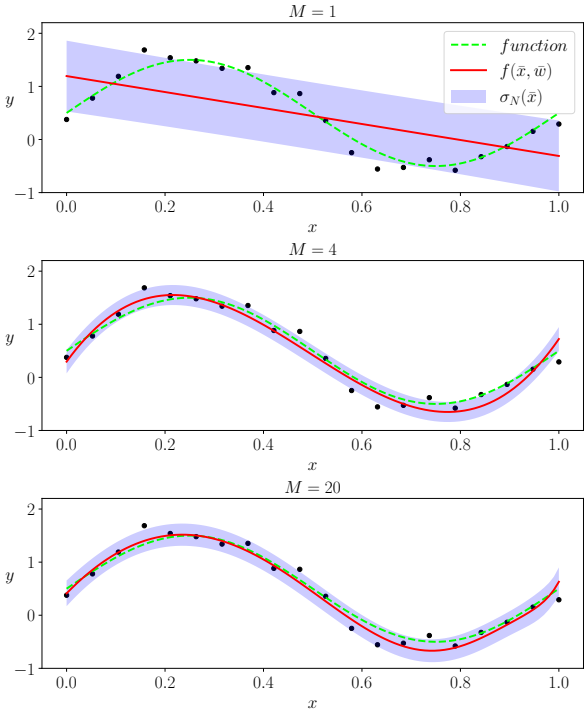


FIGURE 4.9 – Régression bayésienne en utilisant des fonctions de base polynomiales pour différents ordres M en rouge qui corrèlent des données (points noirs) tirées à partir d'une fonction quadratique en pointillé en vert. La zone colorée montre l'intervalle de confiance à 95% (la moyenne plus et moins 1.96 de la variance).

avec $\sigma_N^2(\bar{x}_\star)$ est l'incertitude sur la prédiction. Son expression est composée de deux termes :

— (i) le premier représente le bruit sur les données,

— (ii) et le second correspond à l'incertitude associée aux paramètres $\bar{w}$.

Du fait que le bruit et la distribution des paramètres soient des distributions gaussiennes indépendantes, leurs variances sont alors additives.

Il est également important de souligner qu'avec l'enrichissement de la base de données, c'est-à-dire après plusieurs apprentissages, la distribution *a posteriori* devient plus mince. Par conséquent, l'incertitude diminue avec l'augmentation de la base de données, c'est-à-dire $\sigma_{N+1}(\bar{x}_\star) \leq \sigma_N(\bar{x}_\star)$ [6, 80]. Il en résulte que quand $N \to \infty$, le second terme de l'équation (4.28c), $\bar{\phi}(\bar{x}_\star)^T \bar{\bar{S}}_N \bar{\phi}(\bar{x}_\star) \to 0$. L'incertitude sur la prédiction se réduit alors qu'aux incertitudes dues au bruit et sa valeur sera égale à l'inverse de β.
La figure 4.10, présente l'incertitude sur la prédiction pour différents ordres polynomiaux. Cette figure montre que les incertitudes se réduisent avec l'augmentation de la base de données. De nouveau selon la Fig. 4.10, les incertitudes tendent vers le bruit imposé, ici $\varepsilon = 10^{-2}$ en l'occurrence.
Dans l'équation (4.27), il est indispensable d'inverser la matrice $\bar{\bar{S}}_N$ de taille $N_B \times N_B$ pour faire des prédictions à l'aide de cette équation, ce qui peut ne pas être pratique si

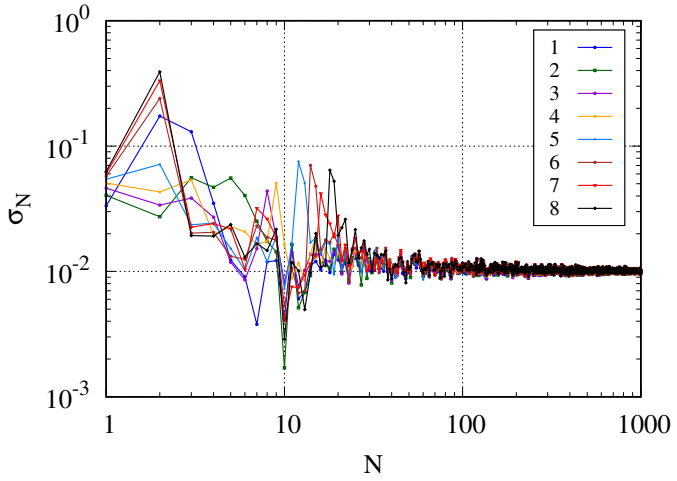


FIGURE 4.10 – Représentation des incertitudes sur la prédiction σ_N pour différents ordres polynomiaux $M = [1, \ldots, 8]$, où le bruit aléatoire inclus dans les valeurs de sorties (cibles) y est fixé à $\sigma_n^2 = 1 \times 10^{-2}$. Pour $N \gg$ l'incertitude sur la prédiction $\sigma_N \to \sigma_n^2 \to \varepsilon$.

N_B est grand. Toutefois, nous pouvons réécrire l'Eq. (4.27) par formule (4.29) pour pouvoir s'affranchir de l'inversion de la matrice. En posant, $\bar{\phi}(\bar{x}_\star) = \bar{\phi}_\star$ et $\bar{\bar{\Phi}}(\bar{x}) = \bar{\bar{\Phi}}$, pour alléger l'écriture, elle peut s'écrire [82],

$$p(f_\star|\bar{x}_\star, \bar{\bar{X}}, \bar{y}) = \mathcal{N}\left(\bar{\phi}_\star^T \bar{\bar{M}} \bar{y}\ , \quad \bar{\phi}_\star^T \Sigma_p \bar{\phi}_\star - \bar{\phi}_\star^T \bar{\bar{M}} \bar{\bar{\Phi}}^T \Sigma_p \bar{\phi}_\star\right) \tag{4.29}$$

avec $\bar{\bar{M}} = \Sigma_p \bar{\bar{\Phi}} (\bar{\bar{K}} + \sigma_n^2 \bar{\bar{I}})^{-1}$ et $\bar{\bar{K}} = \bar{\bar{\Phi}}^T \Sigma_p \bar{\bar{\Phi}}$.

donc,	$\sigma_n^{-2} \bar{\bar{\Phi}} (\bar{\bar{K}} + \sigma_n^2 \bar{\bar{I}})$	$= \sigma_n^{-2} \bar{\bar{\Phi}} (\bar{\bar{\Phi}}^T \Sigma_p \bar{\bar{\Phi}} + \sigma_n^2 \bar{\bar{I}})$
en utilisant Eq. (4.28a)	$\sigma_n^{-2} \bar{\bar{\Phi}} (\bar{\bar{\Phi}}^T \Sigma_p \bar{\bar{\Phi}} + \sigma_n^2 \bar{\bar{I}})$	$= \bar{\bar{S}}_N^{-1} \Sigma_p \bar{\bar{\Phi}}$
alors,	$\beta \bar{\bar{S}}_N \bar{\bar{\Phi}} = \Sigma_p \bar{\bar{\Phi}} (\bar{\bar{K}} + \sigma_n^2 \bar{\bar{I}})^{-1}$	; $\bar{\bar{M}} = \Sigma_p \bar{\bar{\Phi}} (\bar{\bar{K}} + \sigma_n^2 \bar{\bar{I}})^{-1} = \beta \bar{\bar{S}}_N \bar{\bar{\Phi}}$

4.2.2.2 Apprentissage sur les hyperparamètres

Étant donné que l'approche bayésienne doit faire des hypothèses *a priori* sur la forme du modèle, sur les hyperparamètres α et β, et si celles-ci ne sont pas correctes, les résultats peuvent être erronés. Notamment, l'évidence du modèle peut être sensible au choix *a priori*. Il est alors judicieux de faire des prédictions en marginalisant par rapport à ces hyperparamètres ainsi que par rapport aux paramètres $\bar{w}$. Cette procédure est appelée approximation de l'évidence [38, 61]. Il existe deux manières explicatives sur l'apprentissage des hyperparamètres qui sont décrites dans les références BISHOP, (2006) [6] et TIPPING, (2001) [94] (voir annexe B.2).

Brièvement, l'apprentissage sur les paramètres se fait en deux étapes :

1. Premièrement, il faut maximiser de $p(\bar{f}|\alpha, \beta)$ en fonction de α. Il faudra d'abord définir l'équation du vecteur des valeurs propres

$$\left(\beta \bar{\bar{\Phi}}^{\mathrm{T}} \bar{\bar{\Phi}}\right) \bar{u}_i = \lambda_i \bar{u}_i \tag{4.30}$$

2. Deuxièmement, de calculer la dérivée de la log-vraisemblance marginale, plus détaillé dans l'annexe B.2.

$$\nabla_\alpha \ln p(\bar{y}|\alpha,\beta) = \nabla_\alpha \left[\frac{M}{2}\ln\alpha + \frac{N}{2}\ln\beta - E\left(\bar{m}_N\right) - \frac{1}{2}\ln|\bar{\bar{S}}_N^{-1}| - \frac{N}{2}\ln(2\pi) \right] = 0 \tag{4.31}$$

Cette opération conduira à la valeur optimale de α et de β.

La valeur de l'hyperparamètre α qui maximise la vraisemblance obtenue est,

$$\alpha = \frac{\gamma}{\bar{m}_N^{\mathrm{T}}\bar{m}_N} \tag{4.32}$$

où γ est le nombre effectif de paramètres, $\bar{m}_N$ est donné par l'équation (4.28). Une fois la valeur de α est connue, la maximisation de $p(\bar{y}|\alpha,\beta)$ en fonction de β permettra d'obtenir le deuxième hyperparamètre.

$$\frac{1}{\beta} = \frac{1}{N-\gamma}\sum_{i=1}^{N}\left\{y_i - \bar{m}_N^{\mathrm{T}}\bar{\phi}\left(x_i\right)\right\}^2 \tag{4.33}$$

Il s'agit d'une solution implicite pour β et elle peut être résolue itérativement en choisissant une valeur initiale de β, puis en l'utilisant pour calculer la moyenne m_N et γ et ensuite estimer à nouveau β en utilisant (4.33), en répétant jusqu'à convergence. Si α et β doivent être déterminés à partir des données, alors leurs valeurs peuvent être ré-estimées ensemble après chaque mise à jour de γ. En résumé, l'objectif de la régression bayésienne est de calculer la distribution *a posteriori* des paramètres du modèle, en utilisant des données observées et les distributions *a priori* probabilistes comme informations. Cette distribution *a posteriori* peut être utilisée pour prédire la valeur de la variable cible pour de nouvelles observations en utilisant la loi conditionnelle.

Pour calculer la distribution *a posteriori*, on utilise la règle de Bayes qui permet de calculer la probabilité d'un événement (ici les paramètres du modèle) en fonction des données observées et des priors probabilistes. Cette distribution *a posteriori* est généralement approximée par une distribution gaussienne, ce qui permet d'utiliser des algorithmes de maximisation de vraisemblance pour trouver les paramètres optimaux du modèle.

Une fois que les paramètres du modèle sont estimés, on peut utiliser la loi conditionnelle pour prédire la valeur de la variable cible pour de nouvelles observations en utilisant la distribution *a posteriori* des paramètres. Cette approche permet de tenir compte de l'incertitude dans les paramètres du modèle et d'obtenir des prédictions plus robustes et précises.

4.2.3 Régression par les processus gaussiens

Il est possible d'obtenir des résultats similaires à ceux de la section précédente, en considérant cette fois l'inférence directement dans l'espace des fonctions. Il s'agit des *processus gaussiens* $\mathcal{GP}$: Modèle dans lequel on considère une distribution de probabilité sur toutes les fonctions possibles correspondantes à un ensemble de points.

La régression gaussienne fut introduite pour la première fois par Wiener et Kolmogorov dans les années 1940, dans la théorie du filtrage des séries temporelles [24]. Par la suite, elle est devenue de plus en plus populaire en géostatistique dans les années 1970, connue sous le nom de krigeage. Puis récemment, la régression par les processus gaussiens est revenue dans le domaine de l'apprentissage automatique [6, 82].

La régression bayésienne et la régression par les processus gaussiens sont deux approches de la régression statistique qui utilisent des hypothèses sur la distribution des paramètres pour prédire une variable cible. La principale différence entre ces deux approches est que la régression bayésienne utilise une distribution *a priori* pour les paramètres du modèle de régression, tandis que la régression par les processus gaussiens utilise une distribution *a priori* pour les fonctions de régression.
En suivant l'approche bayésienne, il faut avant tout exprimer la connaissance *a priori* du problème sous la forme d'une distribution de probabilité. Pour les processus gaussiens, cette démarche se résume à déterminer une fonction de covariance $k(\bar{x}, \bar{x}')$, permettant de contrôler la forme de la fonction recherchée, et une fonction moyenne $m(\bar{x})$, qui spécifie le sous-espace de sortie où elle est attendue.

$$f(\bar{x}) \sim \mathcal{GP}\left(m(\bar{x}), k(\bar{x}, \bar{x}')\right) \tag{4.34}$$

Ici f est distribuée comme un processus gaussien $\mathcal{GP}$ à moyenne $m(\bar{x})$ et à covariance $k(\bar{x}, \bar{x}')$. Comme le montre la section précédente (Sec. 4.2.2.1), la régression génère une seule fonction considérée comme la mieux adaptée à l'ensemble de données. Cependant, il peut y avoir plus d'une fonction qui s'adapte aussi bien aux points de données observés. Les processus gaussiens permettent de définir la distribution de probabilité sur toutes les fonctions possibles, il est alors possible de calculer les moyennes, et les variances pour indiquer la confiance des prédictions. La démarche est la suivante :

1. Les fonctions qui s'adaptent aux points observés permettent de définir les fonctions *a posteriori*,
2. En utilisant les fonctions *a posteriori* actuelles comme *a priori*, les fonctions *a posteriori* se mettent à jour avec les nouvelles observations,
3. La fonction moyenne, calculée par la distribution *a posteriori* des fonctions possibles, est la fonction utilisée pour les prédictions de régression.

Pour simplifier l'explication des processus gaussiens, nous utilisons le modèle de régression linéaire bayésien $f(\bar{x}) = \bar{\phi}(\bar{x})^T \bar{w}$ avec une distribution *a priori* $\bar{w} \sim \mathcal{N}(\bar{0}, \Sigma_p)$ [82],

$$\begin{aligned} \mathbb{E}\left[f(\bar{x})\right] &= \bar{\phi}(\bar{x})^T \mathbb{E}\left[\bar{w}\right] = 0 \\ \mathbb{E}\left[f(\bar{x}) f(\bar{x}')\right] &= \bar{\phi}(\bar{x})^T \mathbb{E}\left[\bar{w}\bar{w}^T\right] \bar{\phi}(\bar{x}') = \bar{\phi}(\bar{x})^T \Sigma_p \bar{\phi}(\bar{x}') \end{aligned} \tag{4.35}$$

$f(x)$ et $f(\bar{x}')$ sont conjointement gaussiennes de moyenne nulle et de covariance donnée par $\bar{\phi}(x)^T \Sigma_p \bar{\phi}(x')$, $\mathbb{E}\left[f(\bar{x})\right]$ définie l'espérance de $f(\bar{x})$. La fonction de covariance, aussi nommée noyau, la plus simple, est celle de l'exponentielle au carré (SE : Square Expenontial), donnée

en fonction de la distance euclidienne, $r = |\bar{x} - \bar{x}'|$ par,

$$k_{\mathrm{SE}}(r) = \sigma_f^2 \exp\left(-\frac{1}{2\ell^2} r^2\right) \tag{4.36}$$

Il existe également la classe de *Matérn* des fonctions de covariance qui dépendent des paramètres ν, ℓ et des fonctions Bessel modifiée K_ν [82].

$$k_{Matérn}(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} (\frac{\sqrt{2\nu r}}{\ell})^\nu K_\nu (\frac{\sqrt{2\nu r}}{\ell}) \tag{4.37}$$

Il est possible que les cas les plus intéressants pour l'apprentissage automatique soient $\nu = 3/2$ et $\nu = 5/2$, pour lesquels les résultats sont également présentés dans la section (Sec. 4.3). Les expressions de la covariance, et donc les noyaux s'écrivent :

$$\begin{aligned} k_{\nu=3/2}(r) &= \sigma_f^2 \left(1 + \frac{\sqrt{3}r}{\ell}\right) \exp\left(-\frac{\sqrt{3}r}{\ell}\right) \\ k_{\nu=5/2}(r) &= \sigma_f^2 \left(1 + \frac{\sqrt{5}r}{\ell} + \frac{5r^2}{3\ell^2}\right) \exp\left(-\frac{\sqrt{5}r}{\ell}\right) \end{aligned} \tag{4.38}$$

où, σ_f^2 est la variance du signal (sans bruit) et ℓ la longueur caractéristique qui définit l'échelle de la variable. Ces deux paramètres sont des hyperparamètres et sont appris au cours de l'algorithme d'apprentissage, détails dans (Sec. B.2.3). La fonction (4.36) de la covariance exponentielle au carré, (SE) aussi connu sous le nom (RBF) *radial basis functions*, correspond à un modèle de régression linéaire bayésien avec un nombre infini de fonctions de base.

Dans un premier temps, considérerons le cas simple de données brutes où les observations sont sans bruit. La distribution *a priori* combinée des sorties d'apprentissage, $\bar{f}$, et des sorties de test $\bar{f}_\star$ est

$$\begin{bmatrix} \bar{f} \\ \bar{f}_\star \end{bmatrix} \sim \mathcal{N}\left(\bar{0}, \begin{bmatrix} K(\bar{\bar{X}}, \bar{\bar{X}}) & K(\bar{\bar{X}}, \bar{\bar{X}}_\star) \\ K(\bar{\bar{X}}_\star, \bar{\bar{X}}) & K(\bar{\bar{X}}_\star, \bar{\bar{X}}_\star) \end{bmatrix}\right) \tag{4.39}$$

En supposant que l'ensemble $\mathcal{D}$ considéré, $\mathcal{D} = \{(\bar{x}_i, y_i) \,|\, i = 1, \ldots, N_B\}$, comprend N_B points d'observations, et $N_{B\star}$ points test, alors $K(\bar{\bar{X}}, \bar{\bar{X}}_\star)$ désigne la matrice $N_B \times N_{B\star}$ des covariances évaluées à toutes les couples de points d'apprentissage et de test, de manière similaire, on peut définir les autres matrices, $K(\bar{\bar{X}}, \bar{\bar{X}})$ matrice $N_B \times N_B$, $K(\bar{\bar{X}}_\star, \bar{\bar{X}}_\star)$ matrice $N_{B\star} \times N_{B\star}$ et $K(\bar{\bar{X}}_\star, \bar{\bar{X}})$ une matrice $N_{B\star} \times N_B$.

La distribution *a posteriori* des fonctions est obtenue en restreignant cette distribution *a priori* conjointe pour qu'elle ne contienne que les fonctions qui correspondent aux points de données observées. Les valeurs de la fonction $\bar{f}_\star$ peuvent être déduites à partir de la distribution, *a posteriori*, conjointe en évaluant la moyenne et la matrice de covariance à

partir de l'équation,

$$\begin{aligned} p(\bar{f}_\star|\bar{\bar{X}}_\star, \bar{\bar{X}}, \bar{f}) &= \mathcal{N}\Big(K(\bar{\bar{X}}_\star, \bar{\bar{X}})K(\bar{\bar{X}}, \bar{\bar{X}})^{-1}\bar{f}, \\ &\qquad K(\bar{\bar{X}}_\star, \bar{\bar{X}}_\star) - K(\bar{\bar{X}}_\star, \bar{\bar{X}})K(\bar{\bar{X}}, \bar{\bar{X}})^{-1}K(\bar{\bar{X}}, \bar{\bar{X}}_\star)\Big) \\ \bar{m}(\bar{\bar{X}}_\star) &= K(\bar{\bar{X}}_\star, \bar{\bar{X}})K(\bar{\bar{X}}, \bar{\bar{X}})^{-1}\bar{f} \\ cov(\bar{\bar{X}}_\star) &= K(\bar{\bar{X}}_\star, \bar{\bar{X}}_\star) - K(\bar{\bar{X}}_\star, \bar{\bar{X}})K(\bar{\bar{X}}, \bar{\bar{X}})^{-1}K(\bar{\bar{X}}, \bar{\bar{X}}_\star) \end{aligned} \tag{4.40}$$

Prédiction avec des observations bruitées

Dans les situations de modélisation avec la dynamique moléculaire, les données sont bruitées du fait que le seul moyen de les obtenir est à partir de calcul de moyennes, ce qui entraîne des incertitudes sur ces données $y = f(x, \bar{w}) + \varepsilon$.

Dans l'hypothèse d'un bruit gaussien additif et uniformément distribué, ε, avec une variance σ_n^2, la probabilité *a priori* d'observations bruitées ainsi que le noyau sont donnés,

$$\begin{aligned} cov(y, y') &= k(\bar{x} - \bar{x}') + \sigma_n^2\delta_{ij} \\ cov(\bar{y}) &= K(\bar{\bar{X}}, \bar{\bar{X}}) + \sigma_n^2\bar{\bar{I}} \end{aligned} \tag{4.41}$$

$$k_{\text{SE}} = \sigma_f^2 \exp(-\frac{1}{2\ell^2}|\bar{x} - \bar{x}'|^2) + \sigma_n^2\bar{\bar{I}} \tag{4.42}$$

où δ_{ij} représente le symbole de Kronecker, l'indice i est affecté à y et j à y', $\delta_{ij} = 0$ si $i = j$. Un terme supplémentaire s'ajoute à l'expression du noyau (4.42) et fait apparaître un troisième hyperparamètre qui doit être appris au cours de l'algorithme d'apprentissage, voir (Sec. B.2.3).

Il est possible d'écrire la distribution combinée des valeurs observées et des valeurs de la fonction aux emplacements de test en fonction de la probabilité *a priori*, comme suit

$$\begin{bmatrix} \bar{y} \\ \bar{f}_\star \end{bmatrix} \sim \mathcal{N}\left(\bar{0}, \begin{bmatrix} K(\bar{\bar{X}}, \bar{\bar{X}}) + \sigma_n^2\bar{\bar{I}} & K(\bar{\bar{X}}, \bar{\bar{X}}_\star) \\ K(\bar{\bar{X}}_\star, \bar{\bar{X}}) & K(\bar{\bar{X}}_\star, \bar{\bar{X}}_\star) \end{bmatrix}\right) \tag{4.43}$$

Pour simplifier, notons $\bar{\bar{K}} = K(\bar{\bar{X}}, \bar{\bar{X}})$ et $\bar{\bar{K}}_\star = K(\bar{\bar{X}}, \bar{\bar{X}}_\star)$, et dans le cas où il n'y a qu'un seul point de test $\bar{x}_\star$, écrivons $\bar{k}(\bar{x}_\star) = \bar{k}_\star$ pour désigner le vecteur des covariances entre le point de test et les N_B points d'apprentissage. En dérivant la distribution conditionnelle correspondant à l'équation (4.40), nous obtenons les équations prédictives principales pour la régression par processus gaussien.
Pour un seul point d'essai $x_\star$, la fonction moyenne de prédiction et la variance $\sigma_N^2(\bar{x}_\star)$, illustrées dans la figure 4.11, peuvent s'écrire,

$$m(\bar{x}_\star) = \bar{f}_\star = \bar{k}_\star^T(\bar{\bar{K}} + \sigma_n^2\bar{\bar{I}})^{-1}\bar{y} \tag{4.44}$$

$$\sigma_N^2(\bar{x}_\star) = \underbrace{k(\bar{x}_\star, \bar{x}_\star)}_{T_1} - \underbrace{\bar{k}_\star^T(\bar{\bar{K}} + \sigma_n^2\bar{\bar{I}})^{-1}\bar{k}_\star}_{T_2} \tag{4.45}$$

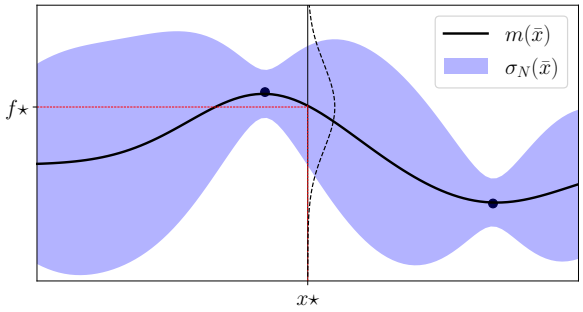


FIGURE 4.11 – Graphique représentant les processus gaussiens avec 2 observations. La ligne noire pleine est la prédiction de la moyenne obtenue par $\mathcal{GP}$ et la zone colorée montre l'intervalle de confiance à 95% (la moyenne ± 1.96 de la variance). La gaussienne superposée correspond à la moyenne $m(.)$ et à l'écart type $\sigma(.)$ de la prédiction au point inconnu $x_\star$

L'équation (4.44) est une combinaison linéaire des observations $\bar{y}$, elle représente alors une prédiction d'une fonction linéaire. Cette fonction peut être écrite autrement sous la forme de combinaison de n fonctions noyaux, chacune centrée sur un point d'apprentissage

$$\bar{f}(\bar{x}_\star) = \sum_{i=1}^{N_B} \alpha_i k_i(\bar{x}_i, \bar{x}_\star) \quad ; \quad N_B \in \mathbb{N},\ \bar{x}_i \in \bar{\bar{X}},\ \alpha_i \in \mathbb{R} \tag{4.46}$$

où $\bar{\alpha} = (\bar{\bar{K}} + \sigma_n^2 \bar{\bar{I}})^{-1} \bar{y}$. L'équation (4.45) est la différence entre deux termes : Le premier T_1 correspond à la covariance *a priori*, le deuxième terme T_2 représente l'information que donnent les observations sur la fonction. Il faut noter que la variance ne dépend pas des cibles observées, mais uniquement des entrées, cette propriété est typique à une distribution gaussienne. La distribution prédictive dans ce cas de données bruitées se déduit à partir de l'équation (4.16) et se résume à

$$p(y_\star | \bar{\bar{X}}, \bar{x}_\star) = \mathcal{N}(m(\bar{x}_\star), \sigma_N^2(\bar{x}_\star)) \tag{4.47}$$

La distribution prédictive pour le modèle $\mathcal{GP}$ donne plus que les barres d'erreur ponctuelles de l'équation (4.45) simplifiée. La figure 4.12 illustre la procédure d'enrichissement utilisée dans notre cas. Initialement, nous disposons que de quelques points (par exemple, deux points). À partir de ces deux points, une fonction moyenne $m(\bar{x})$ est prédite, accompagnée d'une incertitude associée à cette prédiction. À chaque étape, nous identifions le point présentant la plus grande incertitude $\sigma_N(\bar{x})$. Ce dernier est ajouté pour enrichir la base de données. Cette procédure est répétée jusqu'à l'obtention d'une meilleure approximation (comme illustrée dans la troisième figure) où les incertitudes sont minimales.

En résumé : La régression bayésienne et la régression par les processus gaussiens sont deux approches de la régression statistique qui utilisent des hypothèses sur la distribution des paramètres pour prédire une variable cible.
La principale différence entre ces deux approches est que la régression bayésienne utilise une

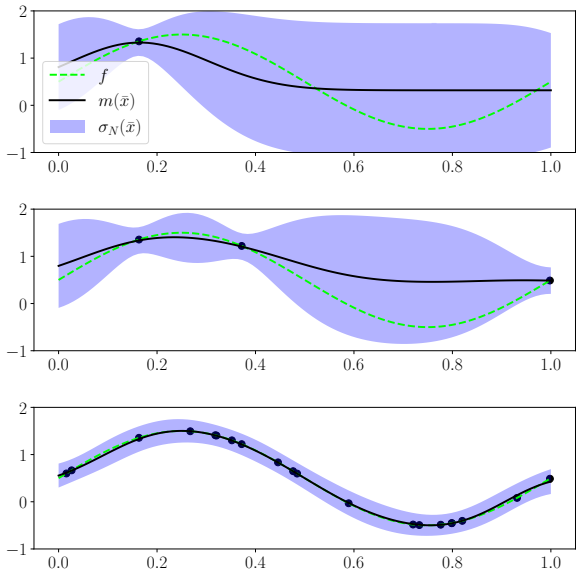


FIGURE 4.12 – Représentation de l'apprentissage par enrichissement au maximum des incertitudes sur la prédiction σ_N. Pour $N \gg$ l'incertitude sur la prédiction $\sigma_N \to \varepsilon$. Les poinst noirs représente les données d'apprentisage.

distribution *a priori* pour les paramètres du modèle de régression, tandis que la régression par les processus gaussiens utilise une distribution *a priori* pour les fonctions de régression.

4.3 Étude comparative des différentes techniques de mise à l'échelle

La normalisation ou la mise à l'échelle des ensembles de données est un élément crucial du pré-traitement des données pour l'apprentissage. Une bonne adaptation de l'échelle est nécessaire pour une prédiction précise avec les algorithmes d'apprentissage. En effet, les ensembles de données constituant la base d'apprentissage, qui ont des échelles très différentes et qui contiennent des valeurs importantes, peuvent entraîner des difficultés de visualisation des données et, surtout, peuvent dégrader les performances prédictives de nombreux algorithmes, par exemple pour des applications en microcanaux, l'échelle de la vitesse varie de 0 à des valeurs maximales de l'ordre de 10^2 m/s, alors que la pression est susceptible de varier sur une échelle de plusieurs bars $p \in [10^4, 10^5]$ Pa. La manipulation numérique de ces données dont les plages de variation sont très différentes peut poser un problème. Effectivement, des données non mises à l'échelle peuvent ralentir ou rendre la convergence impossible.

La plupart des estimateurs sont en effet conçus en supposant que chaque donnée prenne des valeurs proches de zéro ou que toutes les données varient sur des échelles comparables. Notamment, les estimateurs basés sur le gradient supposent souvent des données approximativement normalisées (centrées avec des variances unitaires). Il faut noter que la majorité

des algorithmes d'apprentissage sont plus efficaces lorsque les données sont à une échelle relativement similaire et proches de la distribution normale.

Il existe une variété de méthodes dans le code source `scikit-learn` qui permettent de prétraiter les données pour l'apprentissage [2], parmi elles : `MinMaxScaler`, `StandardScaler`. La méthode requise dépend généralement du type de modèle et des données d'apprentissage. Cette section est dédiée à présenter une étude comparative des deux méthodes de normalisation, afin de faire le bon choix lors de nos simulations.

4.3.1 Outils de normalisation et de mise à l'échelle

Pour rappel, nous désignons par X le vecteur comprenant l'ensemble des valeurs de x_i qui constituent la base de données, $X = \{x_i | i \in [\![1, N_B]\!]\}$. Par conséquent, la matrice $\mathcal{X} = \{X_i | i \in [\![1, N_{var}]\!]\}$, regroupe tous les vecteurs des variables dont dépend la réponse.

4.3.1.1 Méthode de mise à échelle par les valeurs minimales et maximales

Cette méthode permet de mettre à l'échelle les données dans une plage spécifique. Elle permet de transformer alors les données de sorte qu'elles se situent dans un intervalle défini, $[x^\star_m, x^\star_M]$, une fois transformée. Par exemple, un choix classique, l'intervalle entre zéro et un $[0, 1]$.
Pour des données x appartenant initialement à l'intervalle $[x_m, x_M]$, où x_m et x_M sont respectivement les bornes inférieure et supérieure de l'intervalle initial, la transformation s'écrit

$$x^\star = \left(\frac{x - x_m}{x_M - x_m} \right) (x^\star_M - x^\star_m) + x^\star_m \tag{4.48}$$

Cette transformation est le plus souvent utilisée en tant que méthode alternative à la mise à l'échelle qui dépend cette fois-ci de la moyenne décrite ci-dessous.

4.3.1.2 Méthode de normalisation par la moyenne

La méthode de normalisation par la mise en moyenne transforme les données en les centrant autour d'une moyenne et en le mettant à l'échelle en fonction de l'écart-type. Cette méthode suppose que les données suivent une distribution normale.

$$x^\star = \frac{x - \langle x \rangle}{\sigma_x} \tag{4.49}$$

où $\tilde{x}$ la moyenne du vecteur de données dans la base, X, et σ_x l'écart type sur les données.

Il est également important de noter que le choix de la méthode de mise à l'échelle dépend du type de données et des exigences spécifiques de votre problème. Il est recommandé d'effectuer des expérimentations et de comparer les résultats en utilisant différentes méthodes de mise à l'échelle pour déterminer celle qui convient le mieux à votre cas d'utilisation. Les deux méthodes de mise l'échelle `StandardScaler` et `MinMaxScaler` sont toutes deux très sensibles à la présence de valeurs aberrantes.

4.3.2 Choix de la fonction de base

Dans le cas de cette étude, nous avons pris comme exemple une fonction, $f^{(e)}$ qui représente la fonction exacte et s'écrit :

$$f^{(e)}(p,u) = \frac{u}{C_1 + \frac{C2}{p}} \tag{4.50}$$

où u est la vitesse variant dans un intervalle, $u = [0, 100]$ m/s, la pression varie dans une gamme, $p = [5 \times 10^4, 2 \times 10^5]$ Pa, $C_1 = 2.5 \times 10^{-7}$ m, $C_2 = 6.65 \times 10^{-3}$ m Pa. Ces deux dernières constantes $(C_1, C_2/p)$ correspondent respectivement aux valeurs de la hauteur de la couche de Knudsen et le libre parcours moyen des molécules de gaz.
Le choix de la fonction de base utilisée dans la méthode de régression bayésienne repose sur la dépendance de cette pente en p et u. En effet, comme $C_1 > \frac{C2}{p}$, le développement de la fonction (4.50) s'écrit :

$$f^{(e)}(p,u) = \frac{u}{C_1} \sum_{i=0}^{\infty} \left(-\frac{C_2}{C_1} \frac{1}{p} \right)^i \tag{4.51}$$

D'après le développement, la forme idéale de la fonction de base seraient en $(\frac{1}{p}, u)$. Pour des raisons de comparaison et de validation, deux choix de fonction de bases sont utilisés dans ce qui suit : $\phi^{(1)}(x) = [x^{-i} | i \in [\![0, M_1]\!]]$ et $\phi^{(2)}(x) = [x^i | i \in [\![0, M_2]\!]]$, M_1 et M_2 représente l'ordre polynomial.

L'approximation $f^{(p)}(p,u)$ de cette fonction analytique (4.51) est prédite en utilisant les méthodes de régression bayésiennes et les méthodes de régression par processus gaussiens. Les erreurs de prédiction sont calculées par les formules suivantes :

$$\varepsilon_2^{(p)} = \frac{\|f^{(e)}(\bar{x}) - f^{(p)}(\bar{x})\|_2}{\|f^{(e)}(\bar{x})\|_2} = \left[\frac{\displaystyle\int_{\bar{x}_m}^{\bar{x}_M} \left(f^{(e)}(\bar{x}) - f^{(p)}(\bar{x}) \right)^2 d\bar{x}}{\displaystyle\int_{\bar{x}_m}^{\bar{x}_M} \left(f^{(e)}(\bar{x}) \right)^2 d\bar{x}} \right]^{\frac{1}{2}} \tag{4.52}$$

$$\varepsilon_\infty^{(p)} = \frac{\|f^{(e)}(\bar{x}) - f^{(p)}(\bar{x})\|_\infty}{\|f^{(e)}(\bar{x})\|_\infty} = \frac{\max\limits_{\bar{x} \in [\bar{x}_m, \bar{x}_M]} \left| f^{(e)}(\bar{x}) - f^{(p)}(\bar{x}) \right|}{\max\limits_{\bar{x} \in [\bar{x}_m, \bar{x}_M]} \left| f^{(e)}(\bar{x}) \right|} \tag{4.53}$$

où $f^{(e)}(p,u)$ est donné par l'équation (4.51) et $f^{(p)}(p,u)$ est la fonction apprise, dont l'approximation par méthode bayésienne et par la méthode de régression par les processus gaussiens s'écrivent respectivement :

$$\begin{aligned} \text{Cas Bayes} \quad &: f^{(p)}(\bar{x}) = \sum_{i=1}^{C} w_i \bar{\phi}_i(\bar{x}) \quad ; \quad \bar{\phi}_k(\bar{x}) = \prod_{i=1}^{D} \phi_k^{(i)}(x) \\ \text{Cas } \mathcal{GP} \quad &: f^{(p)}(\bar{x}) = \sum_{i=1}^{N_B} \alpha_i k(\bar{x}_i, \bar{x}_\star) \end{aligned} \tag{4.54}$$

avec $\phi^{(1)}(p) = [p^{-i} | i \in [\![0, M_p]\!]]$ et $\phi^{(2)}(u) = [u^i | i \in [\![0, M_u]\!]]$, w_{ij} sont les coefficients

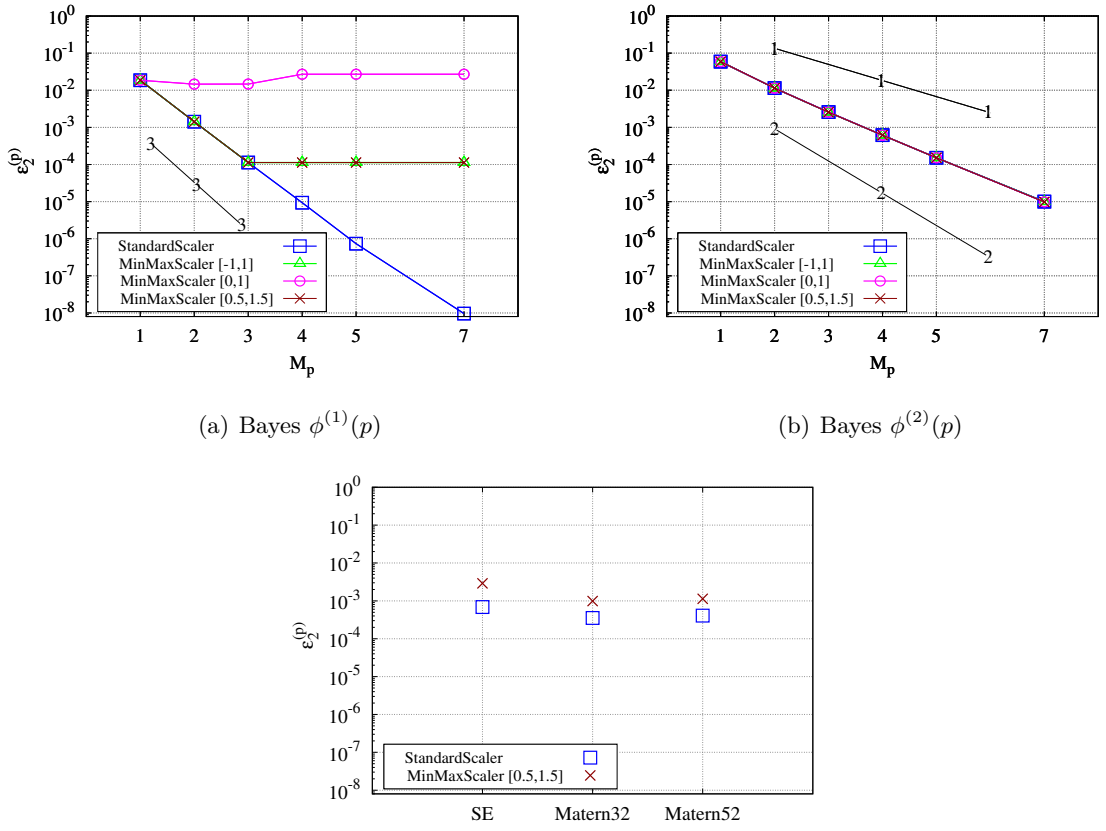
d'interpolation, $k(x_i, x_\star)$ est le noyau utilisé dans la méthode de régression par processus gaussiens.

4.3.3 Résultats de comparaison

Dans les figures représentées dans cette section, le polynôme utilisé pour la vitesse est un polynôme d'ordre 1, $\phi^{(2)}(u) = [1, u]$.

4.3.3.1 Comparaison entre les différentes méthodes de mise à échelle

Sur la figure 4.13, les erreurs relatives, $\varepsilon_2^{(p)}$, obtenues dans le cas d'une remise à l'échelle par la méthode par les valeurs minimales et maximales (`MinMaxScaler`), se révèlent *a priori* peu probantes. En effet, d'une part, l'erreur est plus importante lorsque les fonctions de base



(a) Bayes $\phi^{(1)}(p)$

(b) Bayes $\phi^{(2)}(p)$

(c) Processus gaussiens ($\mathcal{GP}$)

FIGURE 4.13 – Comparaison de la précision des calculs pour les différentes méthodes de mise à l'échelle (4.48) et (4.49), ainsi que les différentes méthodes de régression (Bayes et $\mathcal{GP}$). La différence entre les méthodes SE, MATERN32 et MATERN52 réside dans le noyau utilisé.

sont conçues en inverse de p, alors que ce résultat ne reflète pas le développement de (4.51) dans la section (Sec. 4.3.2). Par ailleurs, il faut porter une attention particulière au choix de l'intervalle de mise à l'échelle pour avoir de meilleurs résultats.

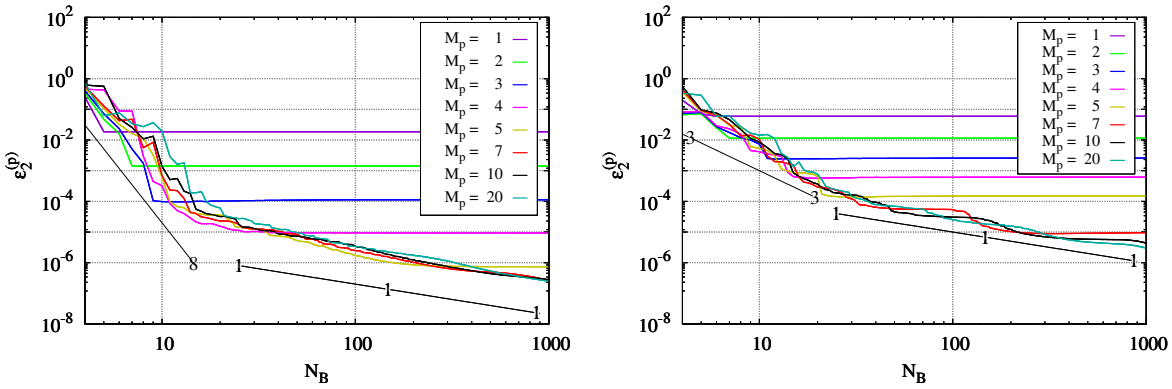
Dans la 4.13(a), la méthode de remise à l'échelle par `MinMaxScaler` sature à une certaine valeur d'erreurs, une fois l'ordre $M_p > 3$. La normalisation par la méthode `StandardScaler` en revanche, donne de meilleurs résultats, les erreurs sont de plus en plus petites en augmentant l'ordre M_p.
La 4.13(b) montre que lorsque la fonction de base utilisée est polynomiale $\phi(x)^{(1)}$, les erreurs décroissent avec l'augmentation de l'ordre polynomial en pression, et ce Quelle que soit la méthode de remise à l'échelle utlilisée (`MinMaxScaler` ou `StandardScaler`).

Les trois différentes méthodes de régression par processus gaussiens ($\mathcal{GP}$) 4.13(c) donnent des résultats moins satisfaisants que la méthode bayésienne pour ce cas particulier. Il faut noter que l'avantage des processus gaussiens réside dans le fait qu'avoir des connaissances *a priori* sur la fonction à approximer (définition des fonctions de bases) n'est pas obligatoire pour bien approximer la fonction. En revanche, dans ce cas précis, le développement montre que la fonction à apprendre dépend explicitement de la vitesse u et de l'inverse de la pression p. Ce développement a permis alors d'avoir la bonne fonction de base et par conséquent la bonne approximation.

En résumé, les méthodes de remise à l'échelle par la méthode `StandardScaler` semblent donner de meilleurs résultats pour les deux méthodes de régression. Dans le travail qui suit, seuls les résultats avec ce type de normalisation seront illustrés.

La figure 4.14 montre les erreurs relatives en norme L_2, $\varepsilon_2^{(p)}$, entre l'approximation (la fonction prédite) et la solution analytique, donnée par l'équation (4.51), pour un nombre de points N_B dans la base de données d'apprentissage en fonction de l'ordre polynomial en pression M_p. Les calculs sont réalisés pour un nombre d'enrichissements de la base $N_B = 10^5$ points, cependant seuls les premiers $N_B = 1000$ points sont représentés.

D'après la figure 4.14(a), la meilleure approximation est obtenue en utilisant les fonctions



(a) $\phi^{(1)}(p) = \left\{p^{-i} | i \in [\![0, M_p]\!]\right\}$ (b) $\phi^{(2)}(p) = \left\{p^{i} | i \in [\![0, M_p]\!]\right\}$.

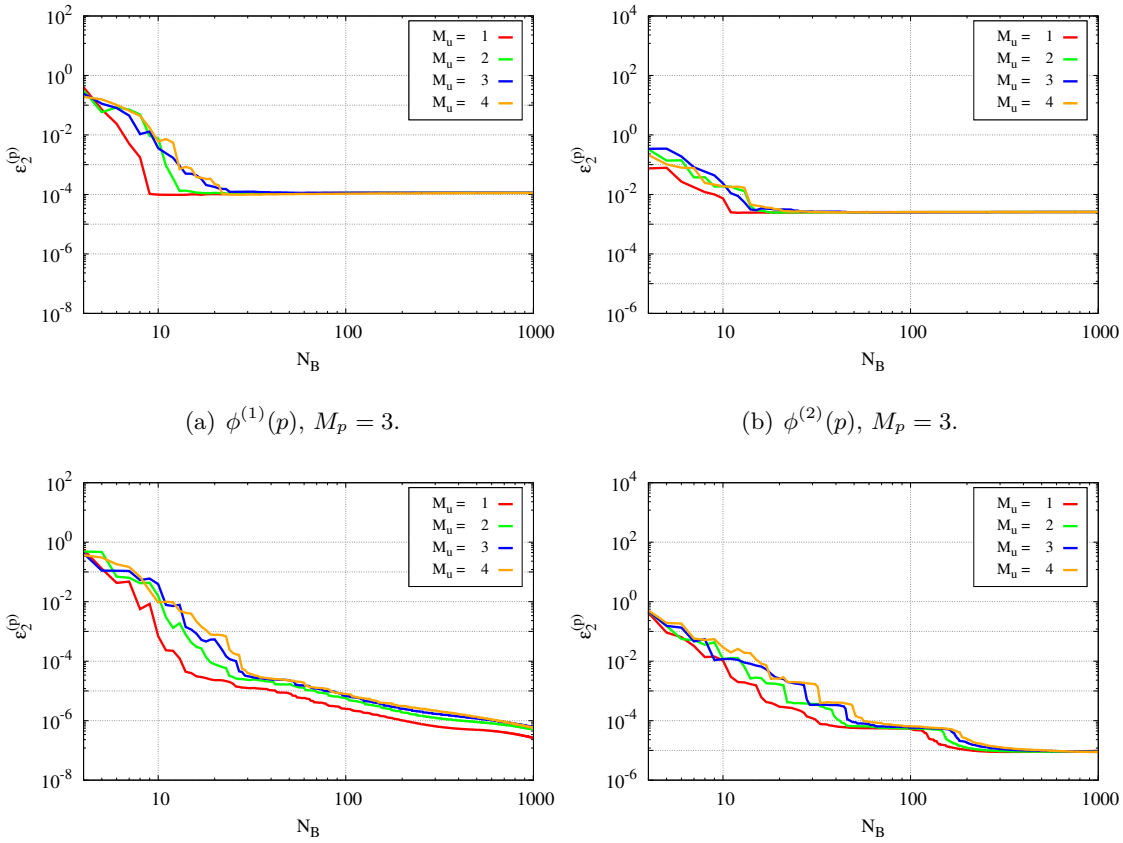
FIGURE 4.14 – Variation des erreurs relatives entre l'approximation (la fonction prédite) et la solution analytique en fonction du nombre de points $N_B = 10^3$ dans la base de données d'apprentissage pour des ordres polynomiaux en pression ($M_p = \{1, 2, 3, 4, 5, 7, 10, 20\}$).

de base en inverse de p, $\phi^{(1)}(p)$, où les erreurs atteignent l'ordre 10^{-7} lorsqu'on augmente l'ordre en pression, M_p. De plus, l'erreur devient constante, voire varie très peu, lorsque le nombre de points correspond au nombre de degrés de liberté de chaque polynôme.

La vitesse de convergence, c'est-à-dire les pentes, changent avec l'augmentation du nombre de points N_B dans la base de données d'apprentissage. La pente au départ est de l'ordre 8, mais une fois le nombre de points coïncide avec le degré de liberté pour chacun des polynômes, la vitesse de convergence diminue et arrive à un ordre d'environ (1.5).

Pour avoir une erreur de $\varepsilon_2^{(p)} = 10^{-5}$, il faut prendre un ordre 5 au minimum lorsqu'il s'agit des fonctions de base polynomiale alors que l'ordre 3 est largement suffisant quand les fonctions de base sont construites en $\frac{1}{p} \rightarrow \phi^{(1)}(p)$.

Les sous-figures de la figure 4.15 montrent que l'ordre en vitesse ne modifie pas les erreurs de prédiction.



(a) $\phi^{(1)}(p)$, $M_p = 3$.

(b) $\phi^{(2)}(p)$, $M_p = 3$.

(c) $\phi^{(1)}(p)$, $M_p = 7$.

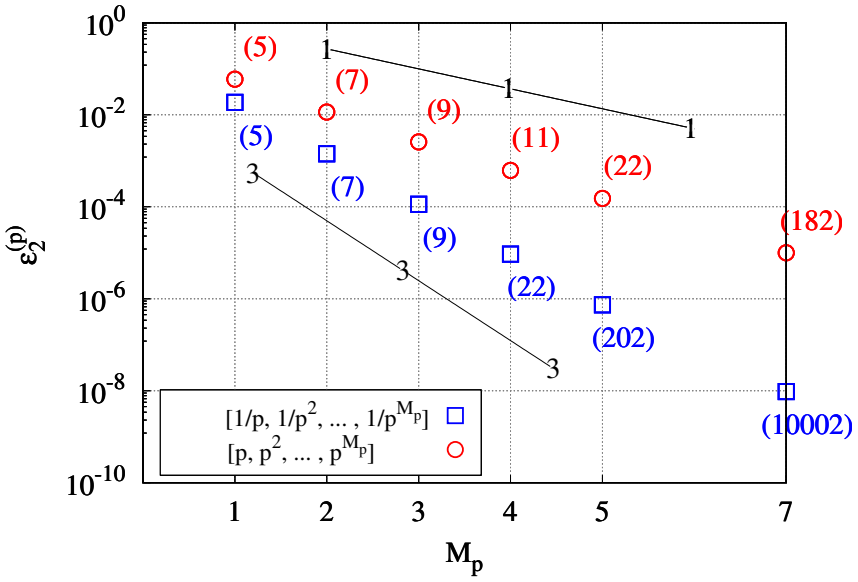
(d) $\phi^{(2)}(p)$, $M_p = 7$.

FIGURE 4.15 – Représentation de l'influence de l'ordre en vitesse sur les erreurs relatives d'approximation entre la fonction prédite et la solution analytique en fonction du nombre de points $N_B = 10^3$ dans la base de données d'apprentissage. Les résultats sont faits sur une gamme des ordres de pression ($M_p = \{1, 2, 3, 4, 5, 7, 10, 20\}$), mais seulement les ordres ($M_p = \{3, 7\}$) sont présentés.

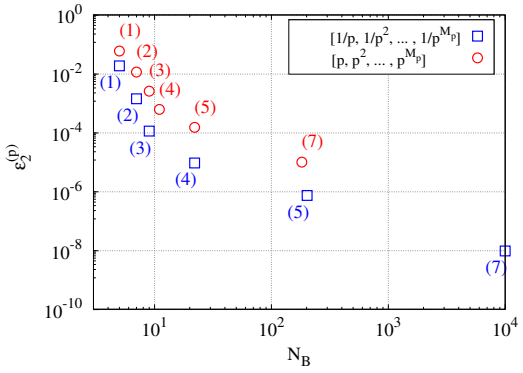
Pour un ordre donné en pression (par exemple $Mp = 3$ dans les figures 4.15(a) et 4.15(b), les

erreurs sont quasiment les mêmes, excepté pour un nombre de points $N_B = 10$ au début, où l'on remarque que l'ordre $M_u = 1$ en vitesse converge mieux). De plus, pour un ordre plus élevé en pression, représenté dans les figures 4.15(c) et 4.15(d), prendre l'ordre 1 en vitesse semble être plus favorable qu'un ordre plus élevé en vitesse. Par conséquent, l'ordre 1 a été utilisé dans la suite de tous les calculs.

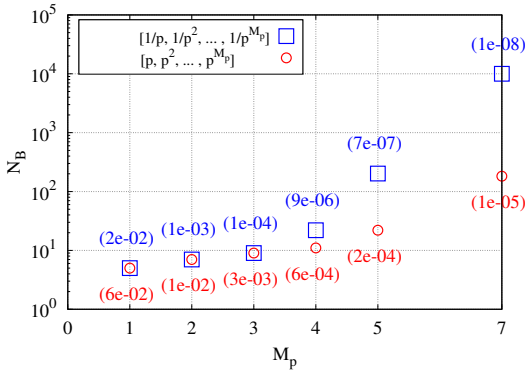
Afin de mieux comprendre le comportement de la convergence, les figures 4.16 illustrent les erreurs en fonction de l'ordre polynomial en p et les erreurs en fonction du nombre de points N_B requis pour que le changement de pente ait lieu, et ce, pour chaque ordre. La figure 4.16(a) confirme ainsi que l'erreur décroît très rapidement lorsque la fonction de base est construite en inverse de pression $\frac{1}{p}$ qu'en p (ordre 3 *Vs* ordre 1). Les erreurs décroissent de façon exponentielle.



(a) Précision de calcul : les nombres inscrit entre parenthèses représentent le nombre de points dans la base lorsque la pente change de valeur.



(b) Entre parenthèses : l'ordre en pression M_p.

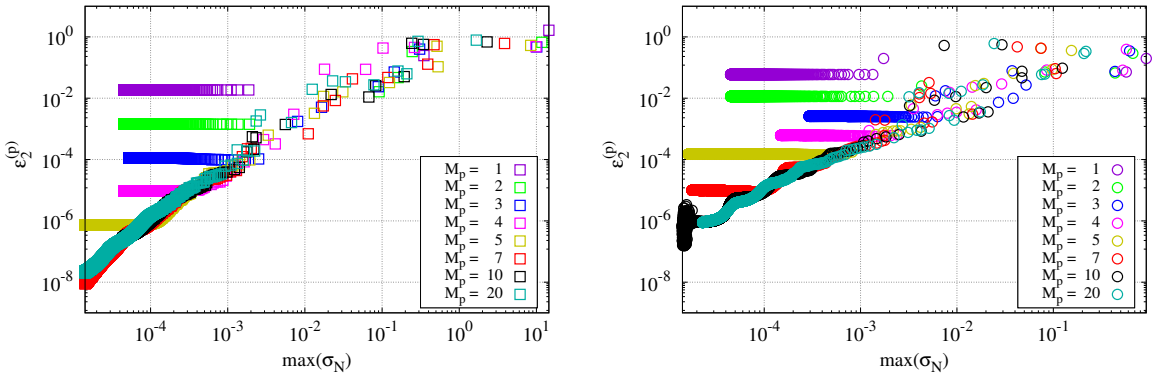


(c) Entre parenthèses : l'erreur $\varepsilon_2^{(p)}$.

FIGURE 4.16 – Comportement de la convergence et précision de calcul. Les symboles (carrés) : représentent les fonctions de base, $\phi^{(1)}(p)$, et les symboles (cercles) : représentent les fonctions de base sont, $\phi^{(2)}(p)$. N_B le nombre de points dans la base de données d'apprentissage et M_p l'ordre en pression.

4.3.3.2 Erreurs de prédiction et incertitudes sur la prédiction

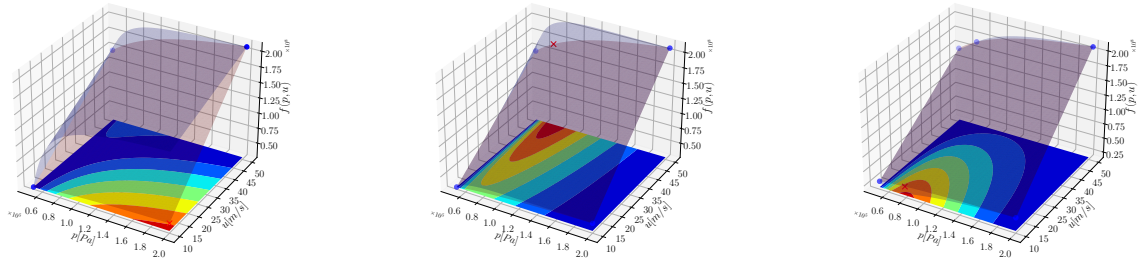
Lorsque la fonction analytique est inconnue, il est intéressant de considérer les incertitudes sur la prédiction afin de confirmer la convergence et par conséquent d'arrêter les calculs. La figure 4.17 présente une comparaison entre la variation des erreurs relatives $\varepsilon_2^{(p)}$, entre l'approximation (la fonction prédite) et la solution analytique, et les incertitudes de prédiction données par σ_N pour une gamme des ordres de pression ($M_p = \{1, 2, 3, 4, 5, 7, 10, 20\}$).



(a) $\phi^{(1)}(p) = \{p^{-i} | i \in [\![0, M_p]\!]\}$ (b) $\phi^{(2)}(p) = \{p^{i} | i \in [\![0, M_p]\!]\}$.

FIGURE 4.17 – Comparaison du comportement de la variation des erreurs relatives $\varepsilon_2^{(p)}$ et les incertitudes de prédiction σ_N en fonction du nombre de points N_B dans la base de données d'apprentissage.

La Figure 4.18 présente quelques itérations ($i = 1, 2, 3$) du processus d'apprentissage, dans laquelle les surfaces de réponses sont exposées. Les surfaces représentent la solution ana-



(a) $i = 1$ (b) $i = 2$ (c) $i = 3$

FIGURE 4.18 – Surfaces de réponses des cas d'apprentissage par la méthode de régression bayésienne polynomiale $M_u = 1$; $M_p = 5$ respectivement par ligne. Les points bleus représentent les données d'apprentissage et les croix rouges pour indiquer le nouveau point ajouté permettant d'enrichir la base de données.

lytique (en rouge) et la surface de réponse prédite (en bleu). Ces surfaces permettent de visualiser les relations entre les variables d'entrée et de sortie. De plus, une carte est utilisée pour indiquer les incertitudes associées à ces prédictions, en déterminant le degré d'incerti-

tude. Les zones présentant une incertitude maximale sont représentées en rouge, tandis que celles présentant une incertitude minimale sont en bleu. Cette représentation colorée permet d'appréhender visuellement les niveaux d'incertitude dans les prédictions.

Les deux figures, 4.21 et 4.22, expose plus d'itérations et illustre également représentent respectivement les surfaces de réponses, les iso-surfaces d'incertitudes sur la prédiction σ_N et les iso-surfaces d'erreurs $\varepsilon_2^{(p)}$ au cours du calcul pour un nombre d'apprentissages $N_B = 1, 2, 3, 4, 5, 7, 10, 15, 20$. L'ensemble des figures, 4.17, 4.21 et 4.22, montrent que les deux quantités, $\varepsilon_2^{(p)}$ et σ_N, sont corrélées. Cela permet d'avoir un critère d'arrêt.

4.3.3.3 Modalités d'enrichissement

Différentes approches permettent d'enrichir la base de données. Nous avons entrepris une analyse comparative de trois méthodes :

- (i) l'ajout de points pour lesquels l'incertitude sur la prédiction est maximale ($\text{Max}(\sigma_N)$) comme illustré dans la figure 4.19(a),
- (ii) enrichissement par ajout aléatoire de points, comme le montre la figure 4.19(b),
- (iii) L'enrichissement par partitionnement du domaine en une grille uniforme, suivant une approche similaire aux plans d'expériences, comme démontré dans (Fig. 4.19(c)).

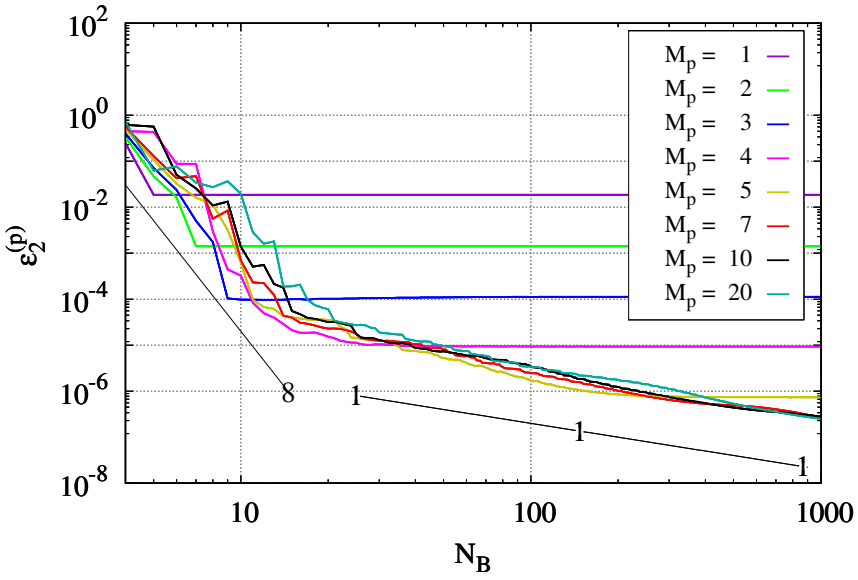
Sur la base de ces trois figures, il est possible d'éliminer l'enrichissement utilisant la grille. Les deux premières méthodes d'enrichissement semblent conduire à des résultats pratiquement similaires. Afin de les distinguer plus clairement et de réaliser une étude comparative, des calculs avec l'introduction de bruit ont été effectués. Les résultats sont présentés dans la figure 4.20.

4.3.3.4 Résultats de calcul sur des données bruitées

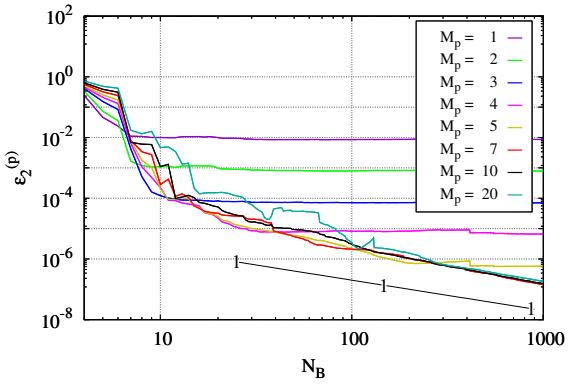
La figure 4.20 illustre les erreurs de prédiction pour différents niveaux de bruit obtenus par le biais de l'outil de calcul de la moyenne.
Bien que le bruit soit présent, on observe une réduction des erreurs de prédiction, qui atteignent des valeurs plus faibles que celles du bruit. Néanmoins, à la différence des cas précédents sans bruit, il convient de noter que l'augmentation de l'ordre polynomial ne conduit pas nécessairement à une meilleure approximation.

4.4 Conclusion

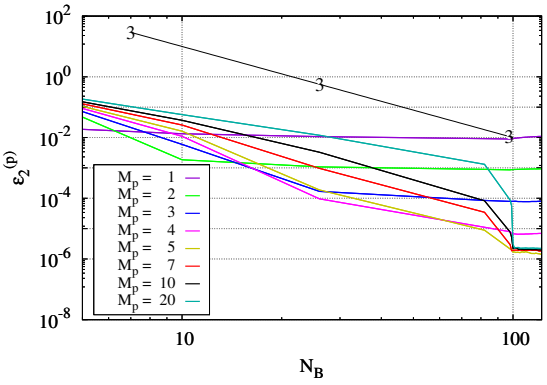
Dans ce chapitre, ont été détaillées les méthodes d'apprentissages choisies, afin d'aborder des problèmes multi-échelle dans des micro conduites. Compte tenu des besoins spécifiques liés au couplage des approches, les méthodes permettant de modéliser et de construire des réponses sous forme de fonctions continues ont été privilégiées. Parmi ces méthodes, les classes d'apprentissage par régression bayésienne et par processus gaussiens ont finalement été retenues.



(a) enrichissement avec Max(σ)



(b) enrichissement par ajout de points aléatoirement



(c) enrichissement avec une grille uniforme comme celle des plans d'expériences

FIGURE 4.19 – Comparaison : enrichissement avec une grille uniforme comme celle des plans d'expériences, soit la grille parcours les deux variables p et u en bleu et pour p

En outre, une étude approfondie de la question de la mise à l'échelle des données d'entrée a été effectuée, en évaluant les différentes options et en choisissant celle qui convenait le mieux à notre contexte. Enfin, notre étude a également porté sur l'impact de l'intensité du bruit lié aux données d'entrées. Bien que cela n'affecte pas de manière significative la convergence associée à la hiérarchie des bases de fonctions choisie, il a été constaté que la précision des prédictions est réduite, en ordre de grandeur, à un niveau similaire à celui du bruit intrinsèque des données.

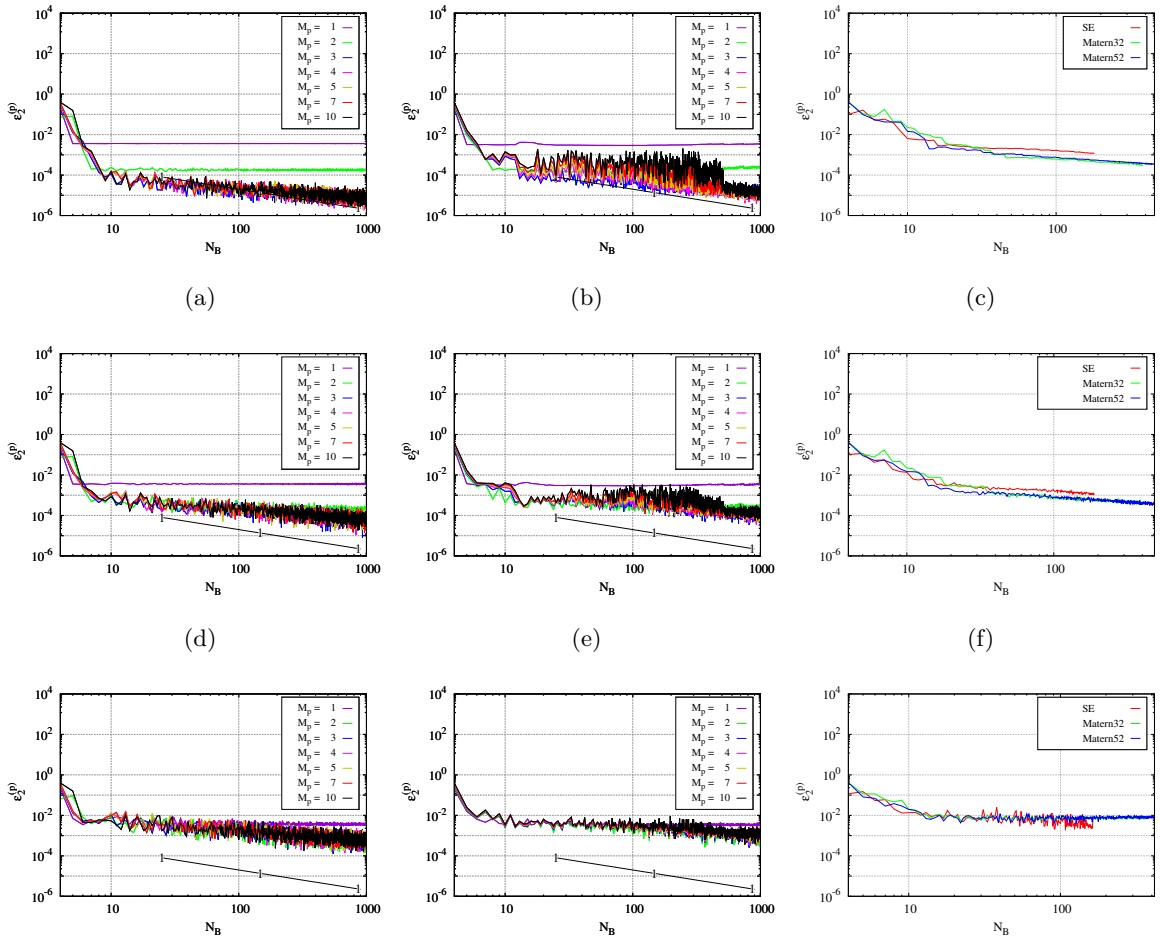
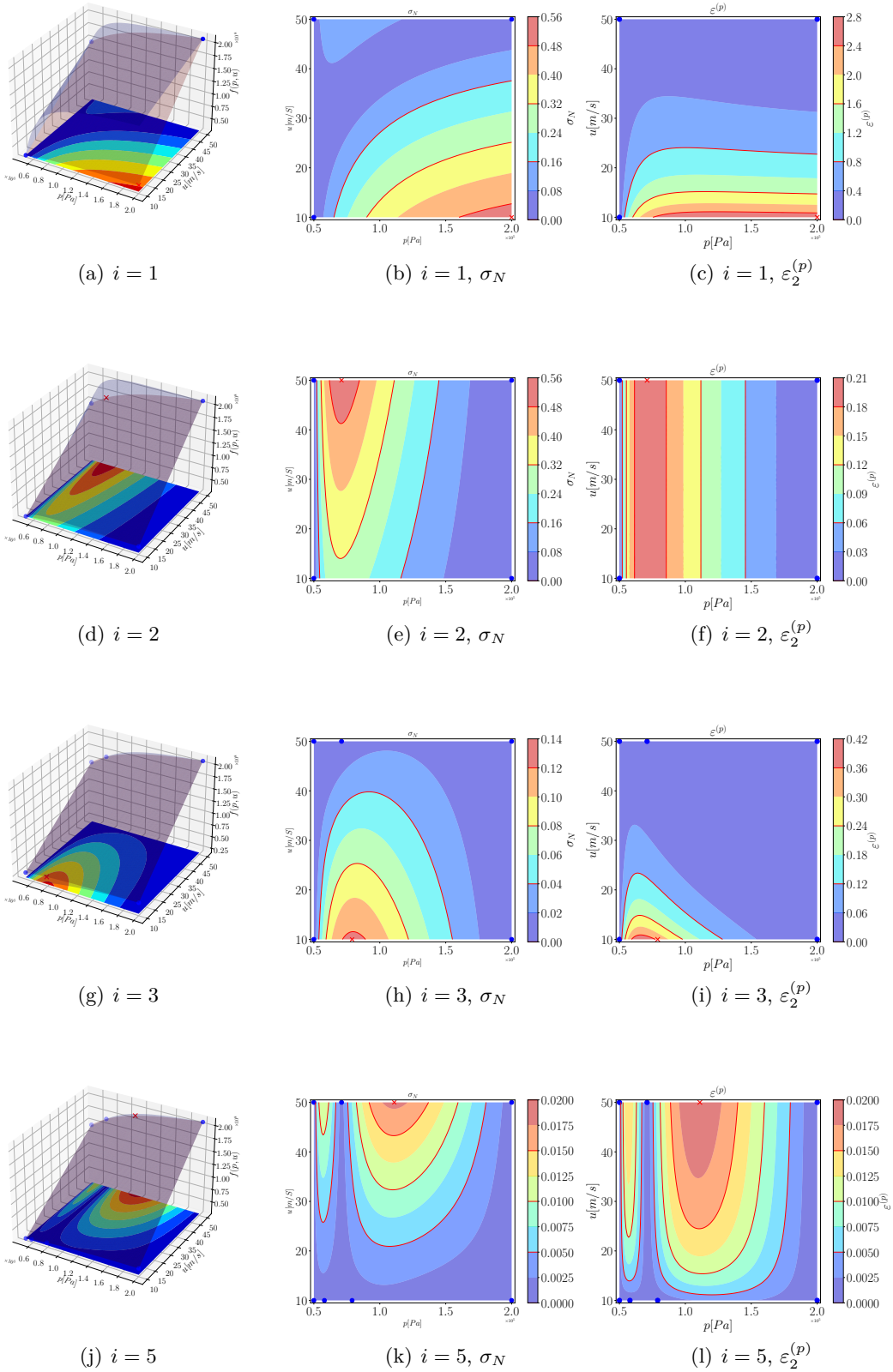


FIGURE 4.20 – Erreurs relatives entre l'approximation (la fonction prédite pour des données bruitées) et la solution analytique pour un nombre de points N_B dans la base de données d'apprentissage en fonction de l'ordre polynomial en (M_p,M_u). La première colonne : enrichissement avec Max(σ) pour la méthode bayésienne, deuxième colonne : enrichissement par ajout de points aléatoirement, troisième colonne : enrichissement avec Max(σ) pour les processus gaussiens. Par lignes, les bruits sont respectivement : $10^{-4}, 10^{-3}, 10^{-2}$.



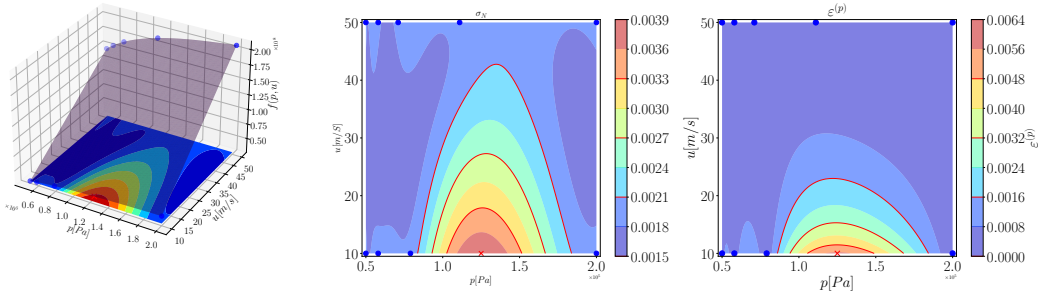
(a) $i = 1$ (b) $i = 1$, σ_N (c) $i = 1$, $\varepsilon_2^{(p)}$

(d) $i = 2$ (e) $i = 2$, σ_N (f) $i = 2$, $\varepsilon_2^{(p)}$

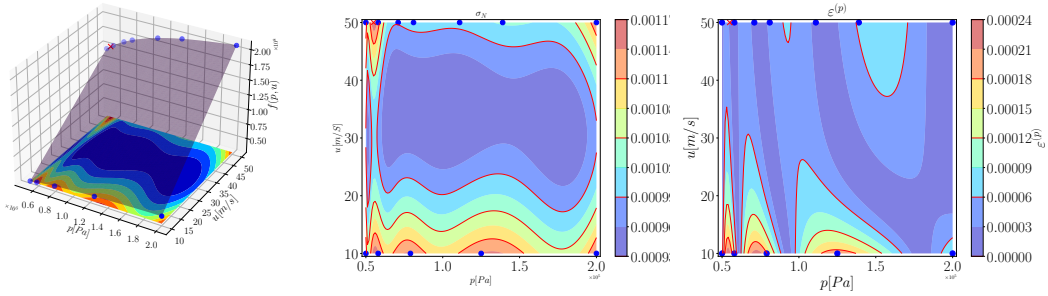
(g) $i = 3$ (h) $i = 3$, σ_N (i) $i = 3$, $\varepsilon_2^{(p)}$

(j) $i = 5$ (k) $i = 5$, σ_N (l) $i = 5$, $\varepsilon_2^{(p)}$

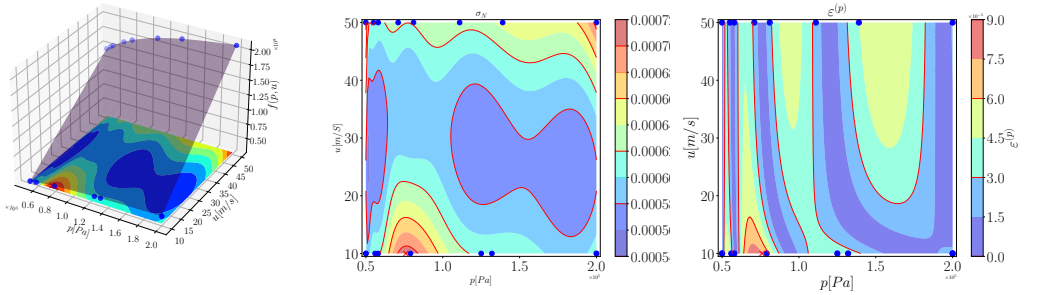
FIGURE 4.21 – Surfaces de réponses des cas d'apprentissage par la méthode de régression bayésienne polynomiale $M_u = 1$; $M_p = 5$ respectivement par ligne.



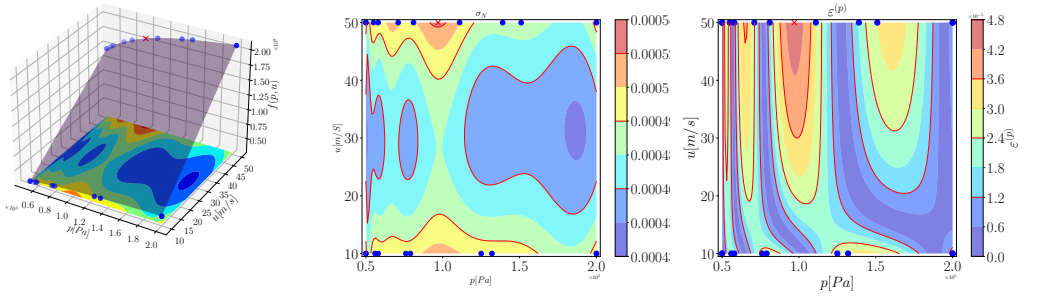
(a) $i = 7$ (b) $i = 7$, σ_N (c) $i = 7$, $\varepsilon_2^{(p)}$



(d) $i = 10$ (e) $i = 10$, σ_N (f) $i = 10$, $\varepsilon_2^{(p)}$



(g) $i = 15$ (h) $i = 15$, σ_N (i) $i = 15$, $\varepsilon_2^{(p)}$



(j) $i = 20$ (k) $i = 20$, σ_N (l) $i = 20$, $\varepsilon_2^{(p)}$

FIGURE 4.22 – Surfaces de réponses des cas d'apprentissage par la méthode de régression bayésienne polynomiale $M_u = 1$; $M_p = 5$ respectivement par ligne.

Le succès est une échelle sur laquelle on ne peut pas monter les mains dans le dos.

Proverbe Américain

Chapitre 5

Méthodes hybrides et multi-échelles pour les écoulements en nano/microcanaux de grande extension

Sommaire

5.1 Introduction 110
5.2 Principe des méthodes de couplage 111
5.3 Couplage fort par la méthode des plots 112
5.3.1 Couplage par plots moléculaires 112
5.3.2 Couplage hybride entre les solveurs continu et discret 113
5.3.3 Inconvénients de la méthode hybride de couplage explicite par plots 115
5.4 Couplage faible par la méthode des plots virtuel 115
5.4.1 Du couplage fort au couplage faible : éléments principaux d'améliorations 115
5.4.2 Couplage hybride faible explicite par la méthode des plots virtuels . 117
5.4.3 Construction de la réponse du plot virtuel 121
5.4.3.1 Algorithme d'apprentissage des fonctions f_{u_τ}, f_{u_w} et f_q . . 121
5.4.3.2 Dynamique moléculaire dans un plot virtuel 124
5.4.4 Validation de la procédure de couplage itératif, performance et coûts de calcul 126
5.4.4.1 Simplification du modèle hybride 126
5.4.4.2 Géométrie et conditions aux limites des problèmes modèles 127
5.4.4.3 Écoulement de Couette 128
5.4.4.4 Écoulement de Poiseuille 131
5.4.4.5 Conclusion 134
5.5 Couplage faible hybride multi-échelle implicite par la méthode des plots virtuels 134
5.5.1 Du couplage explicite au couplage implicite : éléments principaux d'amélioration 134

5.5.2 Couplage hybride faible implicite par la méthode des plots virtuels . 135
5.5.3 Construction de la réponse du plot virtuel 139
5.5.3.1 Algorithme d'apprentissage des fonctions f_{u_τ} et f_{u_w} 139
5.5.3.2 Dynamique moléculaire dans un plot virtuel 141
5.5.4 Validation de la procédure de couplage itératif, performance et coûts de calcul 142
5.5.4.1 Problème 142
5.5.4.2 Apprentissage par des processus gaussiens 144
5.5.4.3 Validation de la solution hybride pour un canal de section variable 150
5.5.4.4 Apprentissage bayésien 151
5.5.5 Apprentissage de la fonction $f_{d_n u_\tau}$ dans une boite de dynamique moléculaire 153
5.5.5.1 Murs stochastiques diffus 153
5.5.5.2 Murs atomiques 154
5.6 Conclusion 155

Préambule

Ce chapitre présente deux méthodes de couplage hybride, développées au cours de cette thèse, pour la résolution d'écoulements multi-échelles. Ces deux méthodes diffèrent principalement par la nature explicite ou implicite du couplage entre les solveurs des deux échelles de la physique. Des validations de chacune des méthodes sont d'abord présentées. Les avantages et les inconvénients sont ensuite discutés.

5.1 Introduction

Comme présenté en introduction de ce mémoire (page 16), certains problèmes physiques mettent en jeu des échelles de temps ou d'espace très différentes. Pour simuler cette physique, il est indispensable d'adopter des modélisations et des outils numériques appropriés pour chacune des échelles : c'est le principe des méthodes hybrides et multi-échelles.
Différents types de méthodes numériques ont été développées au cours de ces dernières années [19, 75, 78, 84, 98, 99, 102, 103] (voir chapitre 1, page 24) afin d'aborder de genre de problématique. Ces méthodes numériques sont dites hybrides, car elles couplent des modélisations numériques différentes adaptées à chacune des échelles résolues. Ainsi, l'approche hybride multi-échelle vise la résolution toutes les échelles de la physique dans une simulation unifiée.

Pour l'application à des nano/microcanaux de grande extension, deux méthodes hybrides et multi-échelles sont développées dans ce travail. Elles se distinguent principalement par la nature explicite ou implicite du couplage entre les solveurs des deux échelles de la physique.

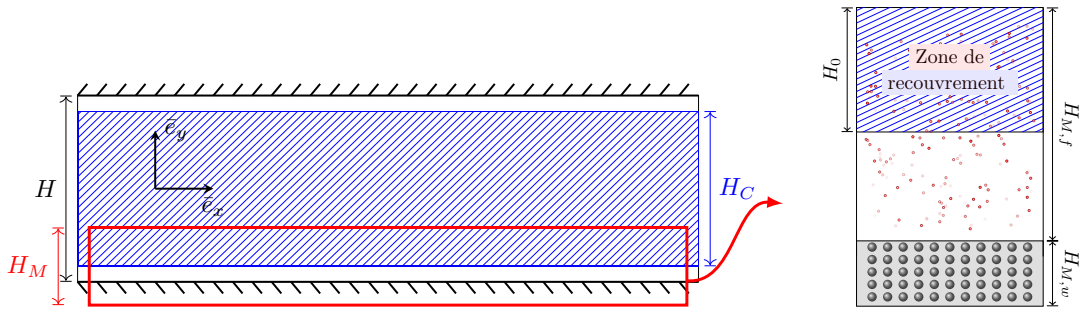
5.2 Principe des méthodes de couplage hybride multi-échelle pour les écoulements en nano/microcanaux de grande extension

Le principe de la méthode hybride multi-échelle pour les écoulements en nano/microcanaux de grande extension consiste à subdiviser le domaine physique en sous-domaines adaptés pour chacune des échelles de la physique, avec une zone de recouvrement pour assurer des communications entre ces échelles. Chacun des domaines est résolu à l'aide d'un solveur continu ou discret.

— La région fluide loin des parois est modélisée par la mécanique des milieux continus : ce modèle est valable à l'extérieur de la couche de Knudsen, c'est-à-dire au-delà d'une distance à la paroi de l'ordre du libre parcours moyen des molécules du fluide. Ce modèle est adapté à la résolution des grandes échelles de l'écoulement.

— La région de proche paroi fluide/solide est décrite par le comportement statistique des molécules et atomes composant le fluide et le mur. Cette région englobe le mur solide et s'étend sur l'épaisseur de la couche de Knudsen. Elle permet de représenter finement la dynamique à petite échelle en prenant en compte toutes les interactions entre les molécules du fluide et les atomes du mur.

La figure 5.1(a) illustre la décomposition de domaines, appliquée à un écoulement en canal



(a) Décomposition du domaine (b) Zoom

FIGURE 5.1 – Représentation multi-échelle d'un micro-canal de hauteur H. (a) : décomposition du domaine en une région bleue, de hauteur H_C, pilotée par les grandes échelles de l'écoulement (mécanique des milieux continus) et une région encadrée en rouge, de hauteur H_M, où les interactions fluide/paroi reposent sur une représentation atomistique de la dynamique. (b) : zoom sur une région de proche paroi. La région discrète est subdivisée en une région de paroi ($H_{M,w}$) et une région fluide ($H_{M,f}$). La zone bleue contenant les points discrets rouges indique la région de recouvrement où cohabitent les représentations continue et atomistique.

entre 2 plaques parallèles distantes de H, en une partie fluide continue de hauteur $H_C \leq H$ (hachurée en bleu) et une partie discrète de hauteur H_M (rectangle rouge) composée des atomes du mur et des molécules du fluide. Une zone de recouvrement assure la communication entre les solveurs continu et discret (Fig. 5.1(b)) par l'échange de variables telles que la vitesse ou la température ou des propriétés physiques.

Des méthodes numériques adaptées sont utilisées pour résoudre les deux échelles de la phy-

sique. Dans ce travail, les équations continues de Navier-Stokes sont discrétisées par la méthode des volumes finis et la dynamique discrète des atomes et molécules résulte de l'intégration temporelle des équations classiques issues de la seconde loi de Newton.
Parmi les méthodes de couplage, nous distinguons le couplage

- fort dans lequel la région moléculaire est simulée, soit totalement, soit partiellement (voir la méthode [98, 99] rappelée à la section 5.3),
- faible dans lequel la région moléculaire n'est pas explicitement simulée, mais un modèle représentant le comportement du fluide à l'extrémité de la couche de Knudsen est construit. Ce modèle s'appuie sur des simulations de dynamiques moléculaires. Nos travaux présentés par la suite reposent sur un couplage de type faible (voir Sec. 5.4 et Sec. 5.5.2).

Selon la façon dont les échanges d'informations s'effectuent entre les échelles, le couplage peut être

- explicite par échange de valeurs numériques, la vitesse par exemple, comme proposé dans [98, 99] ou dans notre méthode présentée en section 5.4 pour des canaux de grandes extensions,
- implicite par la construction et l'utilisation d'une fonction continue qui dépend des grandeurs physiques de l'écoulement (voir Sec. 5.5.2).

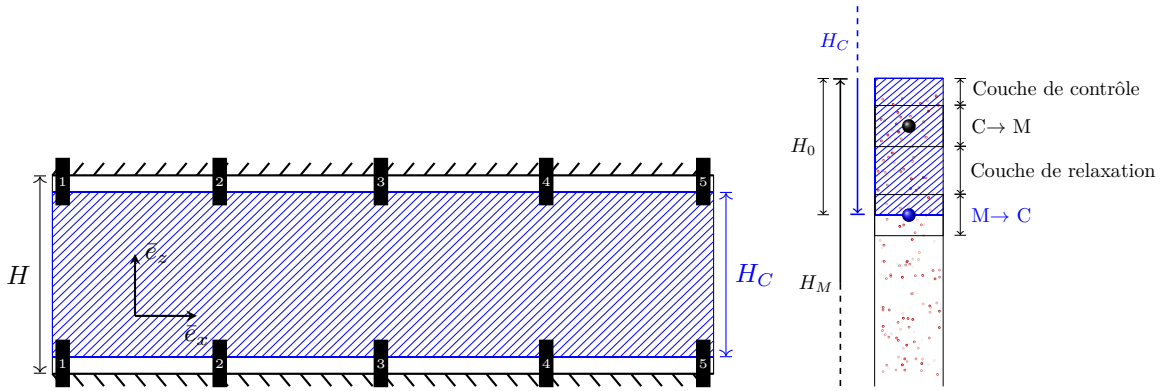
5.3 Couplage fort hybride multi-échelle explicite par la méthode des plots [98, 99]

5.3.1 Couplage par plots moléculaires

La méthode hybride et multi-échelle proposée par Vu et al. [98, 99] repose sur le principe de couplage présenté à la section précédente. C'est un couplage fort, car il résulte d'un échange de valeurs entre les régions moléculaire et continue.

Au lieu de simuler la totalité de la région discrète, ce qui serait excessivement coûteux en temps de calcul, une discrétisation de cette région est réalisée par un petit nombre choisi de micro-sous-domaines, nommés *plots*, intelligemment distribués le long de la conduite (5 plots distribués uniformément sur la Fig. 5.2(a)). Dans chacun de ces plots, le mouvement des atomes/molécules est résolu par dynamique moléculaire avec des conditions de périodicité dans les directions horizontales x et y : l'écoulement moyen est mono-dimensionnel et s'exprime par la vitesse $\bar{v} = u(z)\bar{e}_x$. Le domaine continu est ainsi légèrement décalé par rapport aux parois du canal, de sorte que ses frontières coïncident avec la partie haute des régions moléculaires : la région continue semble alors reposer sur des plots qui ont pour rôle de la connecter aux parois physiques du canal.
La zone de recouvrement dans le plot, qui assure le transfert des informations entre les échelles, est subdivisée en quatre régions distinctes (Fig. 5.2(b)) :

1. La région d'échange du microscopique au continu, notée $M \rightarrow C$. La vitesse macroscopique (et éventuellement la température et la densité) est calculée à partir de la



(a) Couplage par plots (b) Plot

FIGURE 5.2 – Représentation du couplage fort hybride multi-échelle explicite par la Méthode des plots [98, 99]. (a) : décomposition du domaine par en un milieu continu (zone bleue) et en 5 plots de dynamique moléculaire. (b) : région supérieure du plot moléculaire avec la zone de recouvrement H_0 subdivisée en 4 sous-zones permettant les communications entre les régions continu et moléculaire.

dynamique des particules contenues dans cette zone. Cette valeur est ensuite imposée comme conditions aux limites pour le problème continu, à la position moyenne du plot.

2. La région d'échange du milieu continu vers l'échelle microscopique $C \rightarrow M$. Dans cette zone, la dynamique moléculaire est contrainte pour satisfaire la vitesse moyenne (et éventuellement la température et la densité) macroscopique.

3. La couche de relaxation assure la décorrélation de la dynamique des deux zones d'échange, $M \rightarrow C$ et $C \rightarrow M$.

4. La couche de contrôle correspond à la région supérieure de la zone moléculaire. Elle permet de piloter la densité du fluide par l'ajout ou le retrait de molécules du fluide. Une paroi fictive est appliquée à son extrémité supérieure pour maintenir le nombre de particules constant dans le domaine moléculaire.

Comme mentionné ci-dessus, la région d'échange $M \rightarrow C$ permet d'imposer les conditions aux limites du problème continu, uniquement à la position des plots. Entre ces plots, une interpolation (généralement linéaire) est utilisée pour reconstruire les valeurs en frontière (Fig. 5.2(a)).

5.3.2 Couplage hybride entre les solveurs continu et discret

Le couplage entre les solveurs continu et moléculaire repose sur le principe itératif suivant. La procédure est décrite pour l'échange de la vitesse mais s'adapte afin d'éventuellement échanger température et densité.

À partir des valeurs macroscopiques de vitesse calculées dans les domaines $M \rightarrow C$ de tous les plots moléculaires, des interpolations sont utilisées pour reconstruire l'intégralité des conditions aux limites du problème continu, avant sa résolution. De la solution obtenue sont ensuite extraites les valeurs de la vitesse moyenne sur chaque plot, aux ordonnées $z_{C \rightarrow M}$. La dynamique des atomes et molécules est ensuite calculée dans chacun des plots tout en

contrôlant la vitesse moyenne dans le domaine $C \to M$ à l'aide de la dynamique moléculaire sous contrainte.

Une illustration du processus itératif pour un problème de Poiseuille est donnée sur la figure 5.3 où le profil de vitesse est donné à la fois dans un des plots moléculaires et le domaine continu. La solution issue de la dynamique moléculaire est indiquée en rouge avec des symboles circulaires creux alors que la solution issue de la mécanique des milieux continus est indiquée en bleu avec des ronds pleins. La zone d'échange correspond à la région de cohabitation des deux modèles.

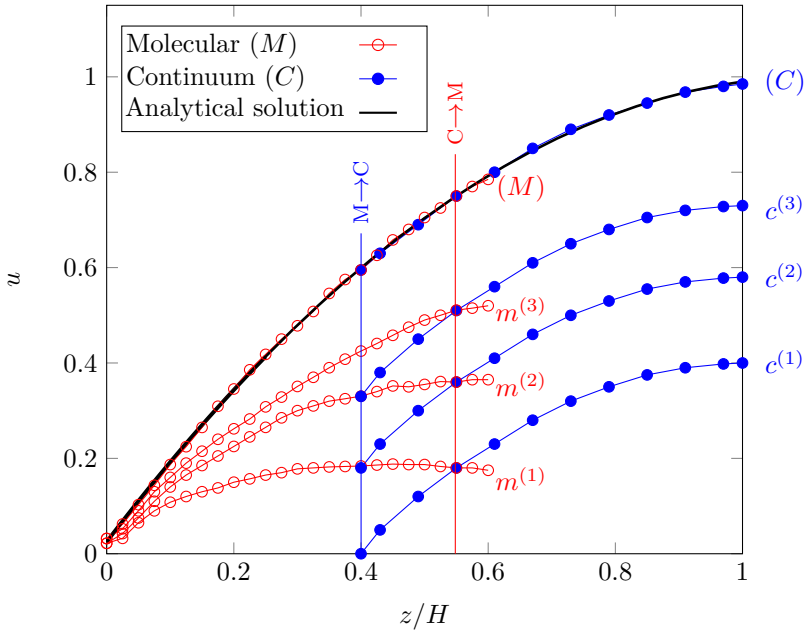


FIGURE 5.3 – Processus itératif de convergence d'un profil de vitesse pour un écoulement de Poiseuille. Les lignes $M \to C$ et $C \to M$ indiquent les hauteurs pour lesquelles les informations sont échangées entre les modèles continu (C) et moléculaire (M), et la flèche $\to$ le sens de la communication. Les solutions continues (**courbes bleues notées** $c^{(i)}$) et moléculaires (**courbes rouges notées** $m^{(i)}$) sont présentées pour les trois premières itérations du couplage. Les courbes notées (C) et (M) correspondent aux solutions convergées, c'est-à-dire à une demi-parabole.

Le principe de couplage est le suivant :

Étape 1 : Initialisation. Le processus est initialisé par un calcul Navier-Stokes avec des conditions aux limites d'adhérence à la paroi ($i = 1$, solution $c^{(1)}$).

Étape 2 : Extraction des informations macroscopiques. La vitesse macroscopique $u(z_{C\to M})$ est extraite à la position $z_{C\to M}$.

Étape 3 : Simulations moléculaires. Des calculs de dynamique moléculaire sont menés en contraignant la vitesse moléculaire moyenne du domaine $C \to M$ à rester égale à $u(z_{C\to M})$.

Étape 4 : Prise de moyenne. Une fois la simulation de dynamique moléculaire convergée (solution moyenne $m^{(i)}$), la vitesse moléculaire moyenne du domaine $M \to C$ est extraite pour l'imposer comme une future condition aux limites du problème continu.

Étape 5 : Calcul d'erreur. Si l'écart relatif entre les anciennes et nouvelles conditions aux limites est inférieur à un critère fixé, le processus itératif est supposé convergé. Dans le cas contraire, il se poursuit.

Étape 6 : Résolution macroscopique. Un nouveau calcul Navier-Stokes est réalisé avec les nouvelles conditions aux limites ($i \leftarrow i+1$, solution $c^{(i)}$), puis le processus itératif recommence à l'**Étape 2**.

5.3.3 Inconvénients de la méthode hybride de couplage explicite par plots

Bien que la méthode de couplage explicite par plots soit infiniment moins coûteuse qu'une simulation directe par dynamique moléculaire complète, ou même réduite à l'ensemble du voisinage des parois, elle présente un certain nombre d'inconvénients.

Nombre et positionnement des plots. Pour limiter le coût des simulations de dynamique moléculaire, le nombre de plots doit rester relativement faible. Cependant, les variables issues des simulations dans les plots sont utilisées pour reconstruire les conditions aux limites du problème continu, ce qui impose de disposer le long du canal de suffisamment de plots positionnés "astucieusement" le long de la paroi. Le nombre et le positionnement des plots sont des paramètres critiques, car la solution convergée issue du processus itératif en est dépendante.

Difficultés des simulations de dynamique moléculaire sous contraintes. L'imposition de contraintes dans la région $C \rightarrow M$ peut s'avérer délicate, particulièrement pour les fluides peu denses [97]. De plus, la coexistence de plusieurs contraintes dans la région $C \rightarrow M$, par exemple la vitesse moyenne, la température et la densité, rend relativement difficile un contrôle précis aux valeurs macroscopiques prescrites. En outre, l'absence d'une paroi matérielle en haut du plot requiert de réinsérer les particules qui quittent le domaine. Cette réinsertion est délicate, car elle doit éviter de trop perturber la dynamique des molécules déjà présentes dans le plot.

Non-conservation du débit massique. Bien que le débit massique soit conservé dans la partie continue du canal, ce n'est plus le cas si les débits dans les plots moléculaires sont pris en compte (pas de conservation du débit massique entre les différents plots). Ainsi, une légère variation de débit existe entre l'entrée et la sortie du canal.

Coût élevé de calcul lié au couplage fort. À chaque itération du processus itératif, la totalité des plots moléculaires doivent être mis à jour, ce qui est évidemment d'autant plus coûteux que le processus itératif convergera lentement.

5.4 Couplage faible hybride multi-échelle explicite par la méthode des plots virtuels

5.4.1 Du couplage fort au couplage faible : éléments principaux d'améliorations

La méthode que nous avons développée de couplage faible explicite, par la méthode des plots virtuels, a pour but de résoudre certains défauts inhérents à la méthode proposée par Vu

et al. [98, 99], méthode dont le principe a été brièvement rappelé dans la section 5.3. Ainsi, notre méthode de couplage hybride multi-échelle est présentée en reprenant les inconvénients mentionnés à la section 5.3.3 et en précisant les approches utilisées pour les limiter ou les éliminer.

Positionnement des plots virtuels sur chaque face des mailles de bord. Au lieu de réaliser un couplage entre les échelles sur un nombre très réduit de plots moléculaires "intelligemment" positionnés le long de la paroi du canal et d'utiliser une interpolation pour reconstruire les conditions aux limites du problème macroscopique, nous introduisons des plots dits virtuels pour l'ensemble des faces volumes finis de frontière du domaine continu.

La nature virtuelle de ces plots moléculaires réside dans le fait que la dynamique moléculaire n'est jamais résolue en leur sein. Leur rôle est de fournir les conditions aux limites $u(z_{M\to C})$ et $v(z_{M\to C})$ en frontière du domaine continu, en réponse à des données issues du modèle continu (Navier-Stokes). Cette réponse résulte d'un apprentissage supervisé à travers un nombre, que l'on souhaite le plus petit, possible de simulations de dynamiques moléculaires. Celles-ci sont réalisées de façon découplée de toutes données issues du modèle continu.

Notons ici que les simulations de dynamique moléculaire restent monodimensionnelles dans la direction normale à la paroi et périodiques dans les autres directions. De ce fait, seule la composante tangentielle u est non nulle et peut être extraite. La composante normale v est reconstruite à partir de la conservation de la masse.

Contrôle de la dynamique moléculaire par un écoulement de type Couette-Poiseuille. Contrairement à la méthode de Vu et al. [98, 99] qui contraint la dynamique du fluide au niveau des plots en $C \to M$, à être identique à celle du milieu continu (couplage fort, Fig. 5.2), la méthode des plots virtuels dissocie complètement les simulations de dynamique moléculaire et du milieu continu.

Comme les simulations moléculaires ont lieu dans la couche de Knudsen, c'est-à-dire au voisinage immédiat de la paroi, il est raisonnable de modéliser cet écoulement comme celui d'un problème de Couette-Poiseuille plan dès lors que la courbure locale de la paroi est négligeable devant le libre parcours moyen des molécules du gaz. Dans le problème de Couette-Poiseuille, la taille du canal, le gradient de pression et la vitesse de la paroi supérieure sont connus. De ces simulations sont extraites des vitesses moyennes et des débits qui viendront enrichir la base d'apprentissage permettant de construire les réponses des plots virtuels. L'utilisation de problème de type Couette-Poiseuille pour la dynamique moléculaire permet également de s'affranchir de la gestion des molécules de gaz qui quittaient le plot par la face supérieure, ce qui résout le problème de leur réinsertion.

Conservation du débit massique. Afin de tenir compte des variations de débit massique entre les plots virtuels adjacents, des conditions aux limites pour la composante verticale de la vitesse, $v(z_{M\to C})$, sont imposées en frontière du problème macroscopique. Cette vitesse est définie sur chaque face des mailles volumes finis de frontière pour assurer la conservation de débit avec les deux plots virtuels adjacents à cette face.

Coût modéré de calcul lié à l'apprentissage de la réponse des simulations de dynamique moléculaire. La nature virtuelle des plots est relative à l'absence de dynamique moléculaire en leur sein. Leur rôle est de mimer la réponse que la dynamique moléculaire fournirait pour une sollicitation macroscopique $u(z_{C\to M})$ à la hauteur $z_{C\to M}$. Cette réponse est déterminée par un apprentissage supervisé dépendant de différents paramètres macroscopiques ($u(z_{C\to M})$, gradient de pression, ...). Elle permet d'imposer les conditions aux limites du problème continu en $z_{M\to C}$. Cet apprentissage peut requérir des simulations supplémentaires de dynamique moléculaire pour améliorer sa précision.

5.4.2 Couplage hybride faible explicite par la méthode des plots virtuels

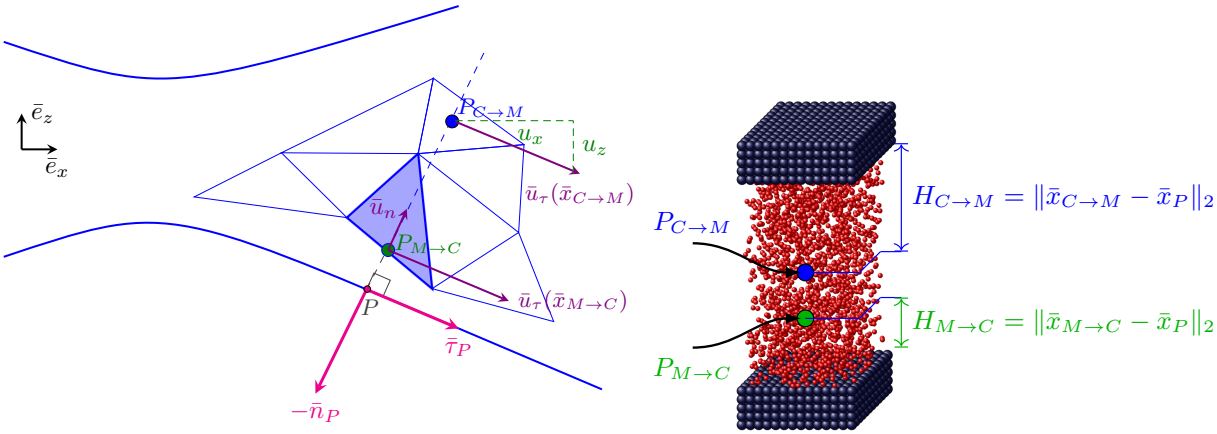
Le canal est subdivisé en deux sous domaines (voir Fig. 5.1). Dans le cœur de l'écoulement, les grandes échelles sont représentées par la mécanique des milieux continus (équations de Navier-Stokes) dont les équations sont discrétisées par la méthode des volumes finis. Au voisinage des parois solides, aucune simulation n'est effectuée. En contrepartie, une méthode d'apprentissage reposant sur la dynamique moléculaire permet de reconstruire les valeurs des conditions aux limites du domaine continu comme si un plot moléculaire reliait chaque maille volumes finis de frontière à la paroi du canal : c'est un plot virtuel.

Afin de décrire la méthode de couplage faible par la méthode des plots virtuels, il est nécessaire d'introduire quelques notations.
Soit le point $P_{M\to C}$, centre d'une face d'un volume de contrôle appartenant à la frontière du domaine continu et adjacent à la zone moléculaire, et $\bar{x}_{M\to C}$ ses coordonnées. On définit le point P comme la projection orthogonale du point $P_{M\to C}$ sur la paroi du canal et $\bar{x}_P$ ses coordonnées. On note $\bar{n}_P$ le vecteur unitaire normal à la paroi au point P et pointant vers le fluide, et $\bar{\tau}_P$ le vecteur tangent tel que $\bar{\tau}_P \wedge \bar{n}_P = \bar{e}_x \wedge \bar{e}_z$. On définit $H_{M\to C} = \|\bar{x}_{M\to C} - \bar{x}_P\|_2$ la distance entre les points $P_{M\to C}$ et P, puis on construit le point $P_{C\to M}$ tel que ses coordonnées $\bar{x}_{C\to M}$ s'écrivent $\bar{x}_{C\to M} = H_{C\to M}\bar{n}_P + \bar{x}_P$ avec $H_{C\to M} > H_{M\to C}$. Comme illustré dans la figure 5.4, les points P, $P_{M\to C}$ et $P_{C\to M}$ sont donc alignés et $P_{M\to C}$ (resp. $P_{C\to M}$) sera le point de communication du domaine moléculaire vers celui continu (resp. continu vers celui moléculaire).
Dans la Figure 5.4(a), une section du maillage est exposée pour mettre en évidence les points de communication, les projections, les normales et tangentes, ainsi que les vecteurs de vitesse. Cette représentation offre une visualisation détaillée des éléments essentiels. De plus, la Figure 5.4(b) présente les différentes hauteurs de communication et décrit la configuration d'un calcul par dynamique moléculaire dans un plot virtuel.

Pour indiquer la $i^{ème}$ itération du processus itératif, nous utiliserons un exposant sous la forme $\bullet^{(i)}$.

Le principe de couplage est le suivant :

Étape 1 : Résolution du problème continu.
Les conditions aux limites du problème continu étant connues à l'itération i pour tout



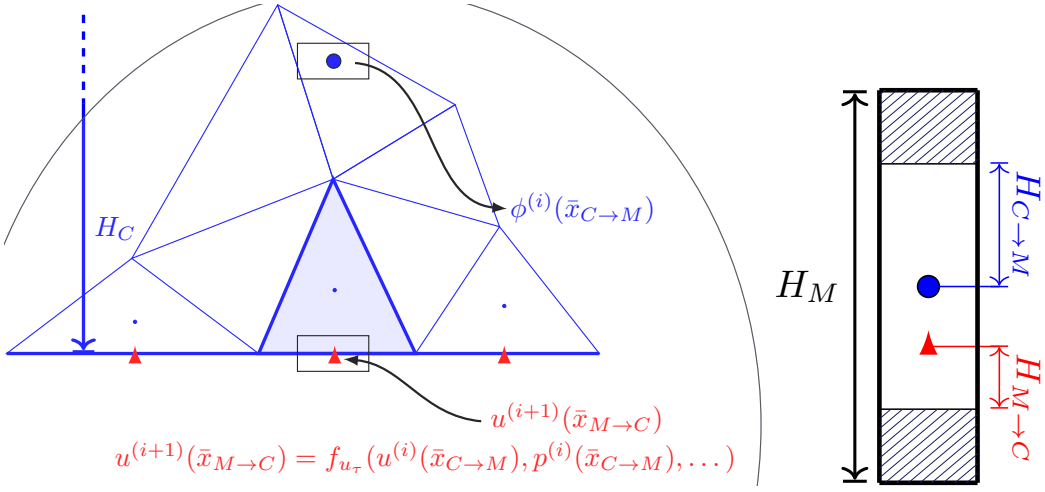
(a) canal $2D$, convergent divergent (b) Dynamique moléculaire dans un plot virtuel

FIGURE 5.4 – Couplage explicite faible, par la méthode des plots virtuels.

point $P_{M\to C}$, c'est-à-dire $u^{(i)}(\bar{x}_{M\to C})$ et $v^{(i)}(\bar{x}_{M\to C})$, les équations de Navier-Stokes discrétisées en volumes finies sont résolues.

Étape 2 : Extraction de grandeurs macroscopiques.

La vitesse tangente $u_\tau^{(i)} = \bar{u}^{(i)} \cdot \bar{\tau}_p$, le gradient de pression tangent $(d_\tau p)^{(i)} := (\partial p/\partial \bar{\tau}_P)^{(i)}$, la pression $p^{(i)}$ et la hauteur $H_{C\to M}$ sont extraits sur l'ensemble des points $P_{C\to M}$ (voir la figure 5.5).



(a) Zoom sur une partie canal $2D$ (b) Dynamique moléculaire

FIGURE 5.5 – Couplage explicite faible, par la méthode des plots virtuels. Les points bleus représentent les centres de mailles à la frontière. Les triangles rouges désignent l'information issue de dynamique moléculaire et imposée au continu comme étant la condition aux limites. La grandeur ϕ peut être la vitesse, pression, température, gradient de pression, etc.

Étape 3 : Construction de la réponse moléculaire dans le plot virtuel

Selon les hauteurs $H_{C\to M}$ et $H_{M\to C}$, les valeurs extraites des pressions $p^{(i)}(\bar{x}_{C\to M})$, des vitesses $u_\tau^{(i)}(\bar{x}_{C\to M})$ et des gradients de pression $(d_\tau p)^{(i)}(\bar{x}_{C\to M})$ sur l'ensemble des points $P_{C\to M}$, une, deux ou trois simulations moléculaires d'un problème de Couette-Poiseuille sont réalisées. Elles servent à enrichir les fonctions apprises $f_{u_\tau}^{(i)}$, $f_{u_w}^{(i)}$ et $f_q^{(i)}$

et ainsi fabriquer les nouvelles fonctions $f_{u_\tau}^{(i+1)}$, $f_{u_w}^{(i+1)}$ et $f_q^{(i+1)}$:

— $f_{u_\tau}^{(i+1)}$ est la fonction apprise à l'itération $i+1$ donnant la vitesse tangentielle moyenne à la hauteur $H_{M\to C}$ du plot virtuel moléculaire soumis à une vitesse moyenne $u_\tau^{(i)}(\bar{x}_{C\to M})$ à la hauteur $H_{C\to M}$, pour une pression moyenne $p^{(i)}(\bar{x}_{C\to M})$ et un gradient de pression moyen $(d_\tau p)^{(i)}(\bar{x}_{C\to M})$:

$$f_{u_\tau}^{(i+1)} := f_{u_\tau}^{(i+1)}(u_\tau^{(i)}(\bar{x}_{C\to M}), p^{(i)}(\bar{x}_{C\to M}), (d_\tau p)^{(i)}(\bar{x}_{C\to M}), H_{C\to M}, H_{M\to C})$$

— $f_{u_w}^{(i+1)}$ est la fonction apprise à l'itération $i+1$ pour effectuer les simulations de Couette-Poiseuille dans une boite de dynamique moléculaire de hauteur fluide $2H_{C\to M}$. Elle fournit la vitesse horizontale moyenne de la paroi solide supérieure de sorte que la vitesse moyenne macroscopique $u_\tau^{(i)}(\bar{x}_{C\to M})$ soit prescrite à la hauteur $H_{C\to M}$, pour une pression moyenne $p^{(i)}(\bar{x}_{C\to M})$ et un gradient de pression moyen $(d_\tau p)^{(i)}(\bar{x}_{C\to M})$:

$$f_{u_w}^{(i+1)}(u_\tau^{(i)}(\bar{x}_{C\to M}), p^{(i)}(\bar{x}_{C\to M}), (d_\tau p)^{(i)}(\bar{x}_{C\to M}), H_{C\to M})$$

— $f_q^{(i+1)}$ est la fonction apprise à l'itération $i+1$ pour représenter le débit massique moyen traversant la hauteur $H_{M\to C}$ du plot virtuel moléculaire lorsque celui-ci est soumis à une vitesse moyenne $u_\tau^{(i)}(\bar{x}_{C\to M})$ à la hauteur $H_{C\to M}$, pour une pression moyenne $p^{(i)}(\bar{x}_{C\to M})$ et un gradient de pression moyen $(d_\tau p)^{(i)}(\bar{x}_{C\to M})$:

$$f_q^{(i+1)}(u_\tau^{(i)}(\bar{x}_{C\to M}), p^{(i)}(\bar{x}_{C\to M}), (d_\tau p)^{(i)}(\bar{x}_{C\to M}), H_{C\to M}, H_{M\to C})$$

Ces trois fonctions apprises $f_{u_\tau}^{(i+1)}$, $f_{u_w}^{(i+1)}$ et $f_q^{(i+1)}$ à l'étape $i+1$ sont reconstruites par des méthodes d'apprentissages supervisés. La construction de la réponse moléculaire, qui couple l'étape d'apprentissage et les simulations de dynamique moléculaire, est décrite plus en détails à la section 5.4.3.

Si aucune simulation de dynamique moléculaire n'est nécessaire, le processus d'apprentissage est achevé et le couplage entre les échelles physiques est résolu.

L'algorithme itératif de la méthode hybride est arrêté.

Étape 4 : Reconstruction des conditions aux limites du problème continu.

Pour chacune des faces de frontière des mailles volumes finis, les nouvelles composantes de la vitesse $\bar{u}$ selon $\bar{\tau}_P$ et $\bar{n}_p$ sont définies par :

$$u_\tau^{(i+1)}(\bar{x}_{M\to C}) = f_{u_\tau}^{(i+1)}(u_\tau^{(i)}(\bar{x}_{C\to M}), p^{(i)}(\bar{x}_{C\to M}), (d_\tau p)^{(i)}(\bar{x}_{C\to M}), H_{C\to M}, H_{M\to C}) \tag{5.1}$$

et

$$u_n^{(i+1)}(\bar{x}_{M\to C}) = \frac{1}{\rho_f}\frac{q_- - q_+}{m(\sigma)} \tag{5.2}$$

avec

$$q_\pm = f_q^{(i+1)}(u_\tau^{(i)}(\bar{x}_{C\to M}^\pm), p^{(i)}(\bar{x}_{C\to M}^\pm), (d_\tau p)^{(i)}(\bar{x}_{C\to M}^\pm), H_{C\to M}^\pm, H_{M\to C}^\pm) \tag{5.3}$$

où $\bar{x}^-_{C\to M}$ (resp. $H^-_{C\to M}$ et $H^-_{M\to C}$) et $\bar{x}^+_{C\to M}$ (resp. $H^+_{C\to M}$ et $H^+_{M\to C}$) désignent les coordonnées du point $P_{C\to M}$ (resp. les hauteurs de communication) pour les plots virtuels amont et aval à $\bar{x}_{C\to M}$, dans le sens de $\bar{\tau}_P$, et $m(\sigma)$ la taille de la face de frontière de centre $\bar{x}_{M\to C}$ du volume de contrôle de la méthode des volumes finis. La composante normale u_n sert à assurer la conservation du débit dans le canal (région continue et plots virtuels).

Les composantes de la vitesse $\bar{u}$ selon $\bar{e}_x$ et $\bar{e}_z$ s'écrivent donc en frontière du domaine continu :

$$u^{(i+1)}(\bar{x}_{M\to C}) = u_\tau^{(i+1)}(\bar{x}_{M\to C})\bar{\tau}_P.\bar{e}_x + u_n^{(i+1)}(\bar{x}_{M\to C})\bar{n}_P.\bar{e}_x$$

et

$$v^{(i+1)}(\bar{x}_{M\to C}) = u_\tau^{(i+1)}(\bar{x}_{M\to C})\bar{\tau}_P.\bar{e}_y + u_n^{(i+1)}(\bar{x}_{M\to C})\bar{n}_P.\bar{e}_y$$

Étape 5 : Incrément de l'itération de couplage i ($i \leftarrow i+1$) et Retour à l'Étape 1.

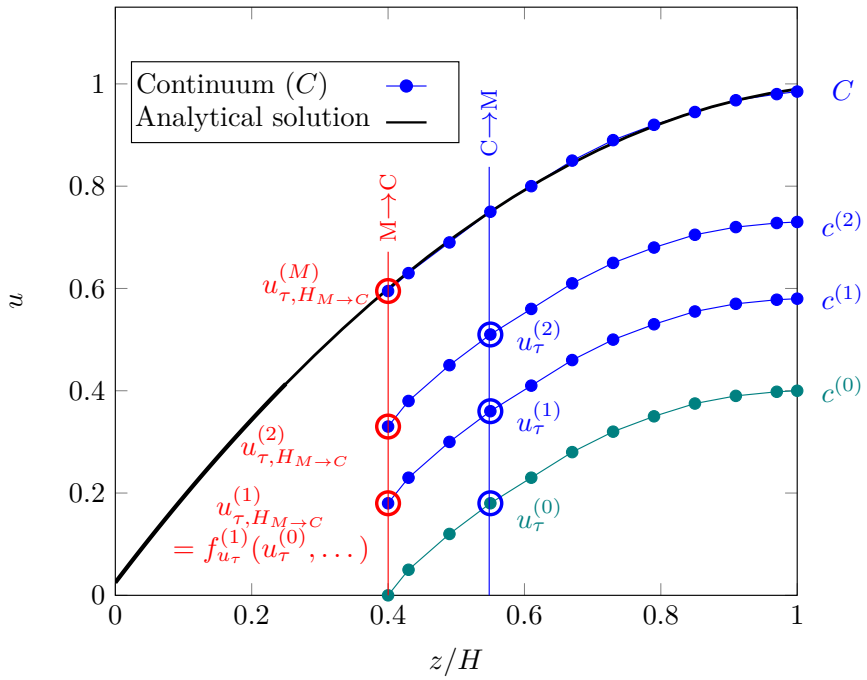


FIGURE 5.6 – Processus itératif de convergence dans le cas de couplage itératif faible avec des plots virtuels. Les lignes $M \to C$ et $C \to M$ indiquent les hauteurs pour lesquelles les informations sont échangées entre le modèle continu (C) et les réponses issues des plots moléculaires virtuels (M) ; la flèche $\to$ souligne le sens de la communication. Les solutions continues (courbes bleues notées $c^{(i)}$) et les valeurs issues des solutions moléculaires (cercles rouges notées $u^{(i+1)}_{\tau,H_{M\to C}} = f^{(i+1)}_{u_\tau}(u^{(i)}_\tau(\bar{x}_{C\to M}), p^{(i)}(\bar{x}_{C\to M}), (d_\tau p)^{(i)}(\bar{x}_{C\to M}), H_{C\to M}, H_{M\to C})$) sont présentées pour les trois premières itérations du couplage pour une coupe verticale donnée.

Pour un écoulement de Poiseuille selon $\bar{e}_x$ dans une conduite plane, le processus itératif entre la solution volumes finis (en bleu) et la sortie du plot moléculaire virtuel (cercles rouges) est illustré sur la Fig. 5.6.

5.4.3 Construction de la réponse du plot virtuel

5.4.3.1 Algorithme d'apprentissage des fonctions f_{u_τ}, f_{u_w} et f_q

L'algorithme d'apprentissage repose sur les méthodes de régression Bayésienne ou Gaussienne présentée au chapitre 4. Les données d'entrée de notre problème d'apprentissage sont la vitesse $u_{C\to M}$ selon $\bar{e}_\tau$, la pression p, le gradient de pression $d_\tau p$ et les hauteurs de communication $H_{C\to M}$ et $H_{M\to C}$. Les grandeurs apprises dans la boite moléculaire du problème de Couette-Poiseuille sont la vitesse moyenne $u_{M\to C}$ selon $\bar{e}_\tau$ à la hauteur $H_{M\to C}$, le débit q sur la hauteur $H_{M\to C}$ et la vitesse de la paroi mobile u_w selon $\bar{e}_\tau$ (à la hauteur $2H_{C\to M}$).

La mise en place de l'algorithme d'apprentissage à l'**Étape 3** de la procédure de couplage hybride entre les solveurs continu et discret (Sec. 5.4.2) requiert une phase d'initialisation précédent l'**Étape 1**. Celle-ci est définie comme l'étape 0 :

Étape 0 : Initialisation de l'algorithme d'apprentissage.

- Données d'entrée :
  - Nombre maximal d'apprentissages supervisés $cpt_{max} > 0$ pour des bornes fixées du plan d'apprentissage.
  - Précisions relatives $\varepsilon_{u_{C\to M}}$, ε_p, $\varepsilon_{d_\tau p}$, $\varepsilon_{H_{C\to M}}$ et $\varepsilon_{H_{M\to C}}$ sur les bornes des grandeurs d'entrée délimitant les intervalles des données d'apprentissage $u_{C\to M}$, p, $d_\tau p$, $H_{C\to M}$ et $H_{M\to C}$.
- Initialisation des variables à l'itération $i = 0$:
  - Définition initiale de la fonction $f_{u_w}^{(0)}$ donnant la vitesse de la paroi u_w pour le problème moléculaire de Couette-Poiseuille. Faute d'information à cette étape, la valeur constante $u_{C\to M}$ est utilisée :

$$f_{u_w}^{(0)}(u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}) = u_{C\to M}$$

  - Initialisation de la base potentielle d'apprentissage B_{pot} à l'ensemble vide $B_{pot} = \varnothing$. La base B_{pot} contient par la suite l'ensemble des octuplets (5 entrées et 3 sorties) de la forme :

$$(5 \text{ entrées} \,;\, 3 \text{ sorties}) = (u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C} \,;\, u_{M\to C}, q, u_w) \quad (5.4)$$

    La dernière variable u_w de l'octuplet n'est utilisée que pour les simulations dans les boites de dynamique moléculaire.
  - Initialisation de deux quintuplets (vitesse, pression, gradient de pression, hauteurs de communication) pour délimiter la zone des données d'apprentissage par valeurs inférieure $X_m^{(new)} = (-\infty, -\infty, -\infty, -\infty, -\infty)$ et supérieure $X_M^{(new)} = (+\infty, +\infty, +\infty, +\infty, +\infty)$.
  - Initialisation du compteur d'apprentissages supervisés $cpt = 0$.

À l'issue de l'**Étape 2** (Sec. 5.4.2), les valeurs de vitesse $u_\tau(\bar{x}_{C\to M})$, de la pression $p(\bar{x}_{C\to M})$ et du gradient de pression $d_\tau p(\bar{x}_{C\to M})$ sont connues pour tous les points $P_{C\to M}$ des faces de frontière. À l'étape i, la procédure d'apprentissage des réponses en vitesse $f_{u_\tau}^{(i+1)}$, en débit $f_q^{(i+1)}$ et en vitesse de paroi $f_{u_w}^{(i+1)}$ est alors la suivante :

Boucle sur le nombre d'apprentissages. Tant que $cpt \leq cpt_{max}$,

1. Définition des bornes des données délimitant la zone d'apprentissage.
   (a) Sauvegarde des anciennes bornes délimitant la zone d'apprentissage $X_m^{(old)} = X_m^{(new)}$, $X_M^{(old)} = X_M^{(new)}$.
   (b) Extraction des nouvelles bornes délimitant les intervalles des données d'apprentissage :

$$X_m^{(new)} = \begin{pmatrix} \min\limits_{P_{C\to M}} u_\tau^{(i)}(\bar{x}_{C\to M}), \\ \min\limits_{P_{C\to M}} p^{(i)}(\bar{x}_{C\to M}), \\ \min\limits_{P_{C\to M}} (d_\tau p)^{(i)}(\bar{x}_{C\to M}), \\ \min\limits_{P_{C\to M}} H_{C\to M}, \\ \min\limits_{P_{C\to M}} H_{M\to C} \end{pmatrix}^T , \; X_M^{(new)} = \begin{pmatrix} \max\limits_{P_{C\to M}} u_\tau^{(i)}(\bar{x}_{C\to M}), \\ \max\limits_{P_{C\to M}} p^{(i)}(\bar{x}_{C\to M}), \\ \max\limits_{P_{C\to M}} (d_\tau p)^{(i)}(\bar{x}_{C\to M}), \\ \max\limits_{P_{C\to M}} H_{C\to M}, \\ \max\limits_{P_{C\to M}} H_{M\to C} \end{pmatrix}^T \tag{5.5}$$

2. Comparaison entre les anciennes $X_{k,j}^{(old)}$ et nouvelles $X_{k,j}^{(new)}$ composantes des bornes délimitant les intervalles des données d'apprentissage, pour

$$k \in \{m, M\} \text{ et } j \in \{u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}\}$$

puis éventuels calculs dans des boites de dynamique moléculaire.

Pour les deux bornes inférieure et supérieure délimitant la zone d'apprentissage, faire pour $k \in \{m, M\}$

- Pour $j \in \{u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}\}$, si $|X_{k,j}^{(new)}/X_{k,j}^{(old)} - 1| > \varepsilon_j$ alors
  (a) Réalisation d'une simulation dans une boite de dynamique moléculaire (voir Sec. 5.4.3.2) :
  — Données d'entrée : $p = X_{k,p}^{(new)}$, $d_\tau p = X_{k,d_\tau p}^{(new)}$, $H_{C\to M} = X_{k,H_{C\to M}}^{(new)}$, $H_{M\to C} = X_{k,H_{M\to C}}^{(new)}$ et la vitesse de la paroi mobile issue de la fonction apprise $f_{u_w}^{(i)}$:

$$u_w = f_{u_w}^{(i)}(X_{k,u_{C\to M}}^{(new)}, p, d_\tau p, H_{C\to M}, H_{M\to C}) \tag{5.6}$$

  — Données de sortie : $u_{C\to M}$, $u_{M\to C}$ et q. Dans l'octuplet (5.4), la sortie u_w est calculée à partir de l'Eq. (5.6).
  (b) Ajout d'un octuplet dans la base potentielle d'apprentissage :

$$B_{pot} = B_{pot} \cup (u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}\,;\, u_{M\to C}, q, u_w)$$

  (c) Les bornes de la zone d'apprentissage ayant été modifiées, le compteur d'apprentissages supervisés est remis à zéro, $cpt = 0$.

3. Apprentissage supervisé :

   (a) Construction de la base réelle d'apprentissage B_{app} par extraction d'octuplets de la base d'apprentissage potentielle B_{pot}. Seules les octuplets dont les données sont dans l'intervalle $[X_{m,j}^{(new)}(1-\varepsilon_j), X_{M,j}^{(new)}(1+\varepsilon_j)]$, pour $j \in \{u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}\}$, sont considérés dans l'étape d'apprentissage.

$$B_{app} = \Big\{ (u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}\,;\, u_{M\to C}, q, u_w) \in B_{pot} \\ \text{tel que } j \in [X_{m,j}^{(new)}(1-\varepsilon_j), X_{M,j}^{(new)}(1+\varepsilon_j)] \\ \text{pour } j \in \{u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}\} \Big\}$$

   (b) Calcul de la nouvelle régression $f_{u_w}^{(i+1)}$ sur la base d'apprentissage B_{app} permettant de définir la vitesse moyenne moléculaire de la paroi mobile en $2H_{C\to M}$. Pour rappel, ce choix permet de contrôler la vitesse moléculaire moyenne $u_{C\to M}$ en $H_{C\to M}$.

   (c) Détermination du quintuplet $(u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C})$ engendrant une vitesse $u_{M\to C}$ la plus incertaine sur la base d'apprentissage B_{app}, c'est-à-dire le quintuplet de données produisant la plus grande variance sur $u_{M\to C}$.

   (d) Réalisation d'une simulation de dynamique moléculaire (voir Sec. 5.4.3.2) :

   — Données d'entrée : p, $d_\tau p$, $H_{C\to M}$, $H_{M\to C}$ ainsi que la vitesse de la paroi mobile issue de la fonction d'apprentissage :

$$u_w = f_{u_w}^{(i+1)}(u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C})$$

   — Données de sortie : $u_{C\to M}$, $u_{M\to C}$ et q.

   (e) Ajout de l'octuplet $(u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}\,;\, u_{M\to C}, q, u_w)$ dans les 2 bases d'apprentissage :

$$B_{pot} = B_{pot} \cup (u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}\,;\, u_{M\to C}, q, u_w)$$
$$B_{app} = B_{app} \cup (u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C}\,;\, u_{M\to C}, q, u_w)$$

   (f) Les bornes de la zone d'apprentissage n'ayant pas été modifiées, le compteur d'apprentissages supervisés est incrémenté,

$$cpt = cpt + 1$$

   (g) Calcul (ou mise à jour) des trois nouvelles régressions sur la base d'apprentissage B_{app} :

   — $f_{u_\tau}^{(i+1)}(u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C})$ la réponse apprise de la vitesse moyenne moléculaire en $H_{M\to C}$.

   — $f_q^{(i+1)}(u_{C\to M}, p, d_\tau p, H_{C\to M}, H_{M\to C})$ la réponse apprise du débit moléculaire.

— $f_{u_w}^{(i+1)}(u_{C\to M}, p, d_\tau p, H_{C\to M})$ la réponse apprise de la vitesse moyenne moléculaire de la paroi mobile en $2H_{C\to M}$.

5.4.3.2 Dynamique moléculaire dans un plot virtuel

Cette section décrit la mise en place d'un calcul par dynamique moléculaire dans un plot virtuel. Les paramètres d'entrée/sortie liés à l'apprentissage sont les suivants :

Données d'entrée :

— la pression moyenne de la boite p,

— le gradient moyen de pression $d_\tau p$,

— la hauteur $H_{C\to M}$ à laquelle la vitesse moyenne doit être contrôlée,

— la hauteur $H_{M\to C}$ permettant de dimensionner la hauteur $H_{M,f} = 2H_{M\to C}$ de la boite de dynamique moléculaire pour la partie fluide,

— la vitesse de la paroi/mur mobile u_w selon $\bar{e}_\tau$ à la hauteur $H_{M,f}$.

Données de sortie :

— la vitesse moyenne macroscopique en $H_{C\to M}$ selon $\bar{e}_\tau$,

— la vitesse moyenne macroscopique en $H_{M\to C}$ selon $\bar{e}_\tau$,

— le débit massique traversant le plot moléculaire entre 0 et $H_{M\to C}$ selon $\bar{e}_\tau$.

Il est à noter que la température moyenne (ou la masse volumique) fait également partie des données d'entrée d'une simulation de dynamique moléculaire puisque pression, température et masse volumique sont liées par la loi d'état. Cependant, dans notre cas, la température n'est pas une donnée d'apprentissage et elle permet de définir l'état thermodynamique du fluide.

Les simulations de dynamique moléculaire sont effectuées dans une boite de section $\ell \times L$ dans les directions $\bar{\tau}_P \times \bar{e}_y$ et de hauteur H_M dans la direction $\bar{n}_P$ (Fig. 5.7). La hauteur $H_M = H_{M,w1} + H_{M,f} + H_{M,w2}$ est subdivisée en 3 régions distinctes : $H_{M,w1}$ et $H_{M,w2}$ désignent les hauteurs des murs atomiques, en partie inférieure et supérieure, et $H_{M,f} = 2H_{C\to M}$ est la hauteur moléculaire fluide. Il faut noter que les dimensions $H_{M,w1}$ et $H_{M,w2}$ peuvent éventuellement être nulles si l'un ou les deux murs sont modélisés par de simples parois stochastiques. Ainsi la vitesse moyenne que nous voulons prescrire en $H_{C\to M}$, en contrôlant par apprentissage la vitesse de la paroi solide supérieur à u_w, est située au milieu de la zone fluide.
La mesure des vitesses moyennes macroscopiques selon $\bar{\tau}_P$, en fonction de la hauteur z, $u(z)$, est réalisée dans de petites tranches horizontales centrées autour de l'ordonnée z et de taille $\ell \times L \times \delta z$, appelées "bins" en anglais. Elles sont obtenues en subdivisant la hauteur fluide en N_b couches de hauteur $\delta z = H_{M,f}/N_b$. Classiquement, N_b est une valeur impaire de sorte que la vitesse au milieu de la zone fluide puisse être aisément calculée. La valeur de N_b est généralement de l'ordre d'une quarantaine, ce qui assure une représentation suffisamment fine du profil de vitesse macroscopique en fonction de $z = (k - 1/2)\delta z$, $k = 1, \ldots, N_b$.

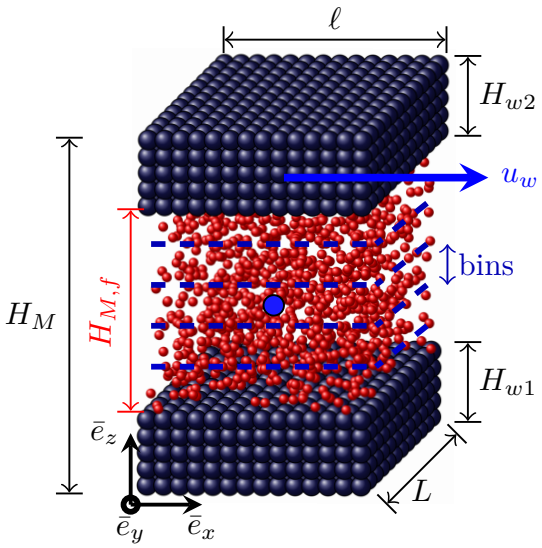


FIGURE 5.7 – Dimensions du plot virtuel de dynamique moléculaire et représentation des boites ("bins") dans lesquelles sont calculées des valeurs moyennes des vitesses selon $\bar{\tau}_P$.

Un écoulement de Couette-Poiseuille isotherme à $300\,K$ est simulé. La température est contrôlée par l'utilisation de thermostats au niveau des murs solides, mais également en ajustant la distribution initiale de la vitesse des molécules du gaz. Ces dernières sont initialisées, en accord avec la densité souhaitée ρ_f, sur un réseau cristallin prédéfini, par exemple selon un réseau CFC (Cubique Face Centré). Les dimensions transverses ℓ et L sont ajustées afin d'obtenir la densité requise. Celle-ci est déduite de la pression moyenne du fluide et de sa température par une loi d'état. Pour l'ensemble des configurations d'écoulements de gaz abordées dans ce travail, une loi d'état correspondant à un fluide de Lennard-Jones [53], de la forme $p = Z(\rho, T)\rho T$, a d'abord été utilisée. Pour les gammes de pression et température parcourues, la valeur du coefficient de compressibilité Z est très voisine de l'unité : la loi de Lennard-Jones a donc finalement été remplacée par la loi des gaz parfaits pour laquelle il est plus aisé d'exprimer la masse volumique en fonction de la pression et de la température. Le gradient de pression $d_\tau p$ est modélisé par une force équivalente $\bar{f} = \rho_f\, d_\tau p\, / m_f\, \bar{\tau}_P$ qui s'applique sur chaque molécule de masse m_f composant le fluide.

La vitesse du mur mobile est engendrée par un terme de forçage dépendant de la nature du mur utilisé. Dans la plupart des simulations de dynamique moléculaire que nous avons étudié, les parois sont modélisées par des murs stochastiques diffus. La distribution de la vitesse correspond à celle d'équilibre des particules à la température du mur. Ce type de mur permet également de le considérer comme une frontière mobile à la vitesse u_w, dans laquelle les atomes sont réfléchis de manière diffuse.

Dans le cas de murs atomiques, la vitesse est imposée pour un groupe d'atomes constituant le mur ; chaque composante de vitesse peut être spécifiée.

Les tranches (bins en anglais) correspondent à une répartition spatiale des atomes de la boite. Chaque groupe d'atomes appartient alors à une tranche dans laquelle les moyennes sont calculées, permettant ainsi de remonter à des grandeurs macroscopiques. Les tranches sont définies par rapport à une origine spécifiée, qui peut être le bord inférieur/supérieur de

la boîte de simulation, ou son point central, ou une valeur spécifiée des coordonnées.

5.4.4 Validation de la procédure de couplage itératif, performance et coûts de calcul

5.4.4.1 Simplification du modèle hybride

Dans la section 5.4.4, nous nous intéressons au coût de calcul de la méthode hybride résultant du couplage itératif entre les solveurs du cœur de l'écoulement et celui moléculaire dédié à de la couche de Knudsen. Les problèmes considérés sont des écoulements de type Couette-Poiseuille glissant en canal, faiblement compressibles, c'est-à-dire pour des nombres de Mach $Ma < 1$. Les solutions de couplage seront comparées à celles issues de la mécanique des milieux continus, pour lesquelles la méthode des volumes finis est utilisée sur l'ensemble du canal. Lorsque cela est possible, des comparaisons à des solutions analytiques sont également proposées. Afin de tester les propriétés de convergence et performances du couplage, les incertitudes relatives aux simulations de dynamiques moléculaires et aux méthodes d'apprentissage sont, dans un premier temps, supprimées. Ainsi, trois simplifications sont effectuées.

1. Les simulations dans les domaines moléculaires sont remplacées par des solutions analytiques de Couette-Poiseuille dans un canal horizontal de hauteur $H_{C\to M}$ et soumis à un gradient de pression $d_\tau p < 0$. La paroi supérieure est en translation dans son propre plan à la vitesse $u_{C\to M}\bar{\tau}_P$ et des conditions de glissement de Navier [71] sont imposées au niveau de la paroi inférieure :

$$u(z=0) - u_p = \frac{2-\sigma}{\sigma}\lambda \left(\bar{\nabla} u.\bar{n}_P\right)\Big|_{z=0} \tag{5.7}$$

où $\bar{n}_P$ est le vecteur normal unitaire pointant vers le fluide, u_p sa vitesse selon $\bar{\tau}_P$, σ est le coefficient d'accommodation caractérisant la nature spéculaire ($\sigma = 0$) ou diffuse ($\sigma = 1$) des chocs des molécules du gaz sur la paroi solide et λ est le libre parcours moyen des molécules du gaz (voir Eq. (1.1), page 19).

La composante horizontale de vitesse est donc donnée par : quel que soit $z \in [0; H_{C\to M}]$,

$$u(z) = \frac{d_\tau p}{2\mu} z^2 + \left(u_{C\to M} - \frac{d_\tau p\, H_{C\to M}^2}{2\mu} - u_p\right) \frac{z + \lambda\sigma_u}{H_{C\to M} + \lambda\sigma_u} + u_p \tag{5.8}$$

avec $\sigma_u = (2-\sigma)/\sigma$.

2. L'apprentissage de la fonction f_{u_w}, qui correspond à la vitesse à imposer sur le mur supérieur de la boite de dynamique moléculaire (Fig. 5.7), n'a plus d'utilité, car la vitesse est directement imposée en $H_{C\to M}$, au point milieu du domaine moléculaire.

3. Les réponses apprises $u_{C\to M}$ et q, pour la construction des fonctions f_{u_τ} et f_q, sont substituées par les sorties exactes des solutions analytiques, c'est-à-dire

$$\begin{aligned} u(H_{M\to C}) &= u(z = H_{M\to C}) \\ &= \frac{d_\tau p}{2\mu} H_{M\to C}^2 + \left(u_{C\to M} - \frac{d_\tau p\, H_{C\to M}^2}{2\mu} - u_p\right) \frac{H_{M\to C} + \lambda\sigma_u}{H_{C\to M} + \lambda\sigma_u} + u_p \end{aligned} \tag{5.9}$$

et

$$q = \int_0^{H_{M\to C}} \rho\, u(z)\, dz \tag{5.10}$$
$$= \rho \left(\frac{d_\tau p}{6\mu} H_{M\to C}^2 + \left(u_{C\to M} - \frac{d_\tau p\, H_{C\to M}^2}{2\mu} - u_p \right) \frac{H_{M\to C} + 2\lambda\sigma_u}{2H_{C\to M} + 2\lambda\sigma_u} + u_p \right) H_{M\to C}$$

Deux critères sont utilisés pour mesurer la convergence de la solution hybride. On introduit

— le vecteur $\bar{u}^{(i)}$, restriction de la solution hybride aux mailles volumes finis du domaine continu pour l'itéré i,

— le vecteur $\bar{u}_{VF}$, restriction au maillage de la solution hybride $\bar{u}^{(i)}$ de la solution volumes finis obtenue sur tout le canal avec des conditions de Navier (5.7) à la paroi.

On définit

— le maximum sur le maillage de la région continue de l'écart relatif de l'itéré hybride $\bar{u}^{(i)}$ à la solution volumes finis $\bar{u}_{VF}$:

$$\varepsilon_\infty^{(c)} = \frac{\|\bar{u}^{(i)} - \bar{u}_{VF}\|_\infty}{\|\bar{u}_{VF}\|_\infty}$$

— le maximum sur tous les points $P_{M\to C}$ de l'écart relatif entre deux itérations successives de la méthode de couplage à la position $\bar{x}_{M\to C}$:

$$\varepsilon_\infty^{(i)} = \frac{\max_{P_{M\to C}} |\bar{u}^{(i)}(\bar{x}_{M\to C}) - \bar{u}^{(i-1)}(\bar{x}_{M\to C})|}{\max_{P_{M\to C}} |\bar{u}^{(i)}(\bar{x}_{M\to C})|}$$

L'écart relatif mesuré entre deux itérations successives, $\varepsilon_\infty^{(i)}$, permet de contrôler la convergence du processus de couplage, mais ne dit rien sur la qualité de la solution hybride. Pour mesurer cette dernière, la solution hybride est comparée à une solution volumes finis $\bar{u}_{VF}$ sur le même maillage, du moins dans la partie continue : cet écart est mesuré par l'écart relatif $\varepsilon_\infty^{(c)}$. Généralement, la solution hybride $\bar{u}^{(i)}$ différera de la solution volumes finis $\bar{u}_{VF}$ dans la mesure où le couplage utilise un modèle particulier dans les plots virtuels (modèle de Couette-Poiseuille analytique) et que les dynamiques dans chacun des plots virtuels ne sont couplées entre elles qu'à travers le domaine continu et jamais entre plots adjacents.

L'initialisation du processus de couplage itératif du modèle hybride s'effectue à l'aide de conditions aux limites d'adhérence en $\bar{x}_{M\to C}$ pour le domaine continu, toujours résolu par la méthode des volumes finis.

Il est à noter que, pour que le couplage entre les échelles soit cohérent, il est nécessaire que la hauteur de communication $H_{M\to C}$ soit choisie supérieure au libre parcours moyen λ, de sorte que le couplage ait lieu à l'extérieur de la couche de Knudsen.

5.4.4.2 Géométrie et conditions aux limites des problèmes modèles

Nous considérons un canal plan de hauteur $2H = 6\,\mu m$ selon $\bar{e}_z$ et de longueur $L = 100H$ selon $\bar{e}_x$ (Fig. 5.8). L'écoulement isotherme à $T_{in} = 300\,K$ est engendré par une différence de

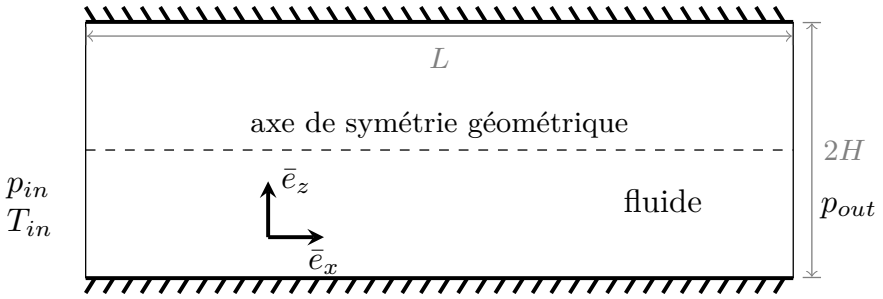


FIGURE 5.8 – Canal plan $2D$

pression entre l'entrée p_{in} et la sortie $p_{out} \leq p_{in}$ et un éventuellement déplacement de la paroi supérieure dans son propre plan à la vitesse u_p. En notant $\bar{v} = u\bar{e}_x + v\bar{e}_z$, les conditions aux limites en vitesse sont en entrée $(\partial u)/(\partial x) = v = 0$ et en sortie $(\partial \bar{v})/(\partial x) = \bar{0}$. Lorsque la surface supérieure du canal est immobile ($u_p = 0$), l'écoulement devient symétrique vis-à-vis de la droite horizontale, $z = H$, ce qui permet de ne résoudre numériquement que la moitié du canal avec les conditions $(\partial u)/(\partial z) = v = 0$ le long de la droite médiane, en $z = H$. Pour le modèle complet résolu par la méthode des volumes finis, des conditions aux limites de type NAVIER [71] (Eq. 5.7) sont utilisées avec $\sigma = 1$ ($\sigma_u = 1$), c'est-à-dire avec des conditions de rebonds diffus des molécules à la paroi isotherme à $300\,K$.

5.4.4.3 Écoulement de Couette

Dans cette section, les équations résolues dans le domaine continu sont celles de Navier-Stokes incompressibles à coefficients constants.

Considérons d'abord un écoulement de Couette glissant conformément à la section 5.4.4.2, avec $p_{in} = p_{out} = 5 \times 10^4$ et $u_p = 100\,m/s$ sur le mur supérieur en $z = 2H$. La pression étant constante, le libre parcours moyen est donc identique tout le long de la paroi du canal, $\lambda/H = 1.1\%$. Le champ de vitesse est alors invariant dans la direction $\bar{e}_x$: $\bar{u} = u(z)\bar{e}_x$. Le couplage hybride n'est effectué que pour la paroi inférieure immobile. Le fluide le long de la paroi supérieure subit un glissement donné par la loi de Navier (5.7).
Le schéma numérique volumes finis étant d'ordre deux, la solution volumes finis est donc exacte. Il est donc attendu que la solution hybride qui couple itérativement une solution volumes finis dans le cœur de l'écoulement et la solution analytique dans les plots virtuels converge également vers la solution exacte d'un écoulement de Couette glissant, quelles que soient les hauteurs de communication $H_{C\to M}$ et $H_{M\to C}$.
Pour une hauteur de communication fixée $H_{M\to C}/H = 10\%$, l'écart relatif $\varepsilon_\infty^{(c)}$ commis entre la solution hybride à l'itération i et la solution exacte (ou volumes finis), tend effectivement vers zéro lorsque les itérés relatifs $\varepsilon_\infty^{(i)}$ tendent vers zéro (Fig. 5.9). Cette convergence est d'ordre 1, quelle que soit la distance entre les deux zones de communication $(H_{C\to M} - H_{M\to C})/H$. Nous notons cependant que pour une valeur donnée $\varepsilon_\infty^{(i)}$, la solution hybride sera d'autant meilleure que la distance entre les régions de communication augmente. Cette amélioration est de l'ordre d'une décade lorsque $(H_{C\to M} - H_{M\to C})$ augmente également d'une décade (de 1.25% à 10%). Cette meilleure précision est liée à une sensibilité accrue de

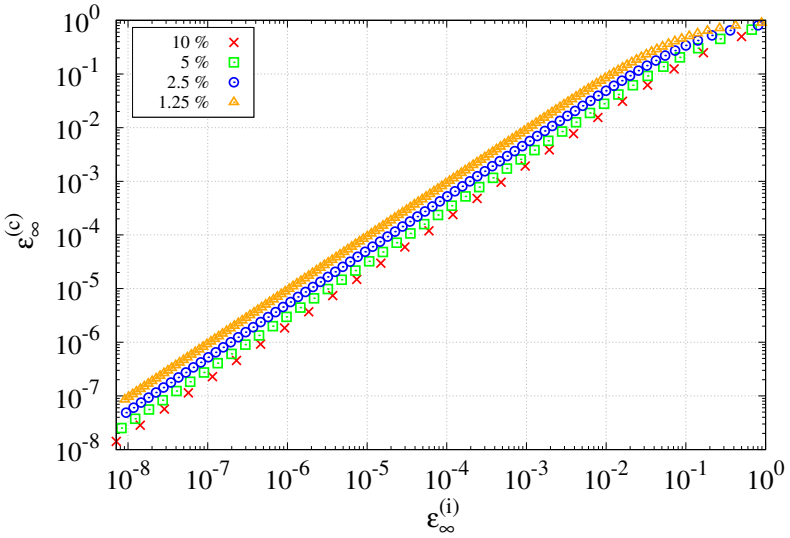


FIGURE 5.9 – Erreur relative de la solution hybride $\varepsilon_\infty^{(c)}$ en fonction de l'écart relatif entre deux itérés successifs $\varepsilon_\infty^{(i)}$ pour un cas écoulement de Couette incompressible, isotherme à $T_{in} = 300K$ avec $H_{M\to C}/H = 10\%$ et différentes valeurs de $(H_{C\to M} - H_{M\to C})/H$ allant de $1,25\%$ à 10%, avec 80 mailles dans la section de hauteur $H - H_{M\to C}$.

la solution hybride vis-à-vis des itérations de couplage lorsque les régions de communication s'éloignent.

La convergence à l'ordre 1 de la solution hybride vers la solution exacte est également retrouvée pour une distance de communication $(H_{C\to M} - H_{M\to C})/H$ fixée et une hauteur de communication du plot virtuel vers la région continue $H_{M\to C}$ variable (Fig. 5.10). De façon

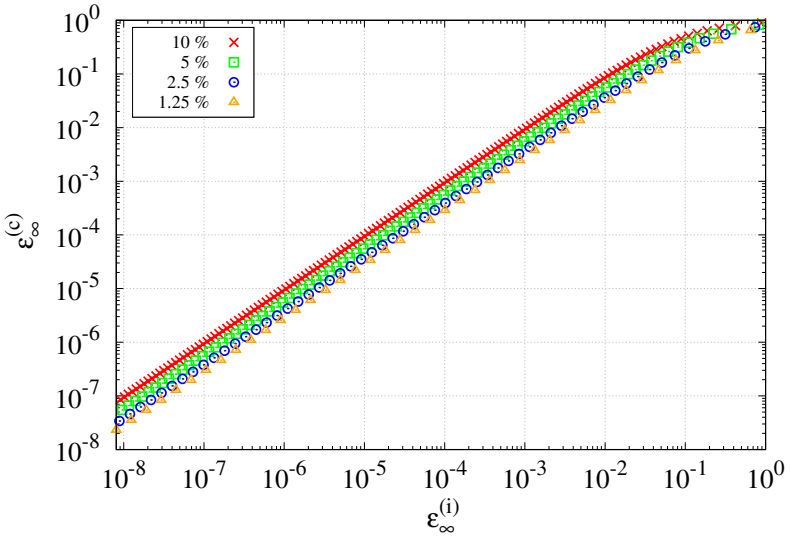


FIGURE 5.10 – Erreur relative de la solution hybride $\varepsilon_\infty^{(c)}$ en fonction de l'écart relatif entre deux itérés successifs $\varepsilon_\infty^{(i)}$ pour, $H_{M\to C}/H = 1.25\%, 2.5\%, 5\%, 10\%$, cas d'un écoulement de Couette incompressible, isotherme à $T_{in} = 300K$, avec 80 mailles dans la section de hauteur $H - H_{M\to C}$.

similaire à l'analyse précédente, la réduction de la hauteur $H_{M\to C}/H$ d'une décade, de 10%

à 1.25%, améliore d'autant la précision de la solution hybride.
Si l'écart relatif à la solution volumes finis $\varepsilon_\infty^{(c)}$ est examiné en fonction du nombre d'itérations i, la vitesse de convergence décroît alors dramatiquement lorsque la distance entre les deux régions de communication $H_{C\to M}$ et $H_{M\to C}$ diminue, pour $H_{M\to C}$ fixé (Fig. 5.11). Cette

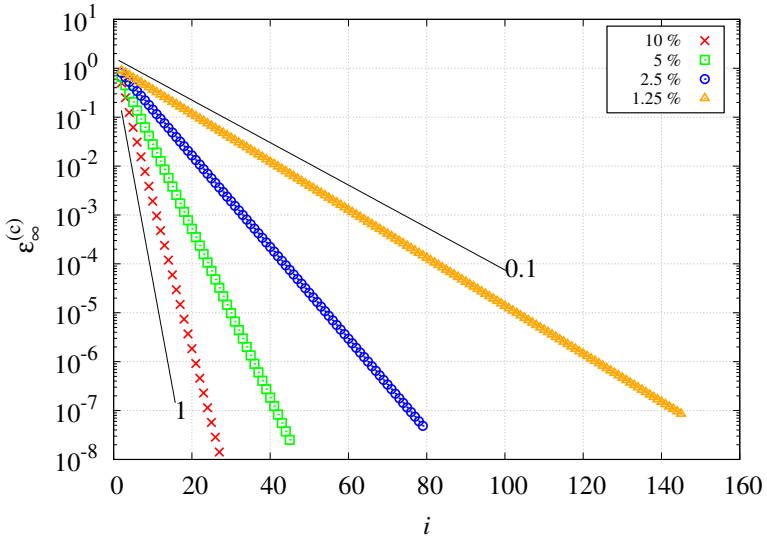


FIGURE 5.11 – Erreur relative à la solution exacte $\varepsilon_\infty^{(c)}$ en fonction des itérations de couplage i et différentes valeurs $(H_{C\to M}-H_{M\to C})/2H$. Écoulement de Couette avec $H_{M\to C}/2H = 5\%$

dégradation de la vitesse de convergence est également observée lorsque $H_{M\to C}$ augmente pour une valeur $H_{C\to M} - H_{M\to C}$ fixée (figure non présentée).
Pour une erreur relative entre deux itérations $\varepsilon_\infty^{(i)} = 10^{-8}$, la figure 5.12 synthétise les dé-

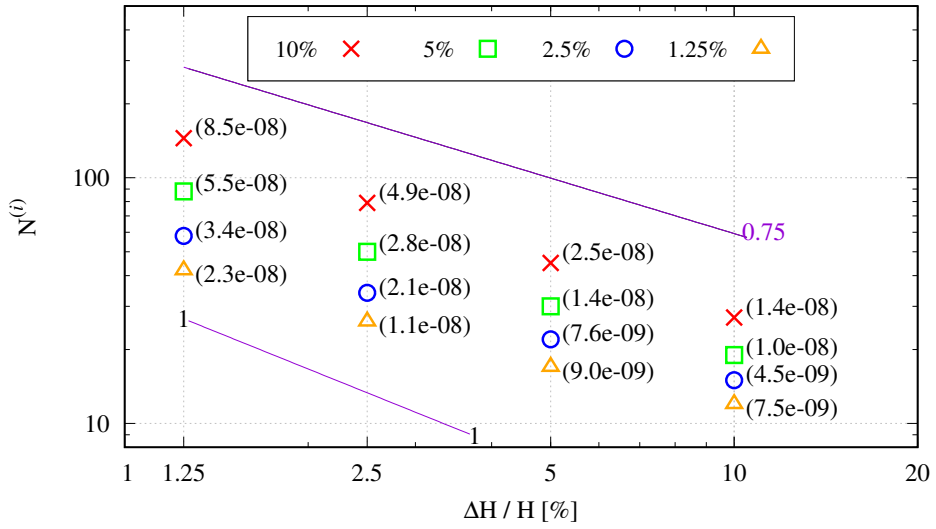


FIGURE 5.12 – Nombre d'itérations de couplage $N^{(i)}$ pour atteindre un écart relatif entre deux itérations $\varepsilon_\infty^{(i)} = 10^{-8}$. Les valeurs entre parenthèses indiquent l'erreur relative à la solution exacte $\varepsilon_\infty^{(c)}$. Écoulement de Couette avec $H_{M\to C}/2H = 5\%$.

pendances du nombre total d'itérations $N^{(i)}$ (ordonnée) et la précision relative atteinte $\varepsilon_\infty^{(c)}$ (valeurs entre parenthèses) en fonction de la distance entre les régions de communication

$\Delta H = H_{C\to M} - H_{M\to C}$ (abscisse) et la hauteur de communication du plot virtuel vers le milieu continu $H_{M\to C}$ (différents symboles). Conformément aux observations précédentes, à $H_{M\to C}$ fixé (symbole donné), le nombre d'itérations $N^{(i)}$ augmente fortement avec la diminution de ΔH selon une loi $N^{(i)} \simeq (\Delta H)^{-0.75}$, tout en dégradant la qualité de la solution hybride $\varepsilon_\infty^{(c)}$ (valeurs en parenthèses). Si maintenant la distance de communication ΔH est fixée (abscisse fixée), la réduction de $H_{C\to M}$ diminue à la fois le nombre d'itérations pour converger, mais également améliore la précision de la solution hybride $\varepsilon_\infty^{(c)}$ (valeurs en parenthèses). Ainsi, les paramètres des simulations qui minimisent le nombre d'itérations $N^{(i)}$ tout en assurant la meilleure précision possible $\varepsilon_\infty^{(c)}$ correspondent à la plus petite hauteur communication du plot virtuel vers le milieu continu $H_{C\to M}$ ($H_{C\to M}/H = 1.25\%$) et la plus grande distance entre les régions de communication $\Delta H = H_{C\to M} - H_{M\to C}$ ($\Delta H/H = 10\%$).

5.4.4.4 Écoulement de Poiseuille

Nous nous intéressons maintenant à un écoulement de Poiseuille dans le canal présenté à la Sec. 5.4.4.2 avec $p_{in} = 2\times10^5\,Pa$, $p_{out} = 5\times10^4\,Pa$ et $u_p = 0$, et nous conservons des conditions de glissement de Navier (Eq. 5.7) à la paroi inférieure. Compte tenu de la géométrie et des conditions aux limites, nous travaillerons maintenant que sur un demi-canal en imposant des conditions de symétries en $z = H$.

Soit un écoulement incompressible avec un glissement constant le long de la paroi du canal, égal à celui obtenu à la pression moyenne $(p_{in}+p_{out})/2$ ($\lambda/H = ?\%$). Comme pour le problème de Couette présenté auparavant, la vitesse $u(z)\bar{e}_x$ devient donc invariante dans la direction de l'écoulement, avec une solution qui s'exprime sous la forme de l'équation (5.8) où $H_{C\to M} = H$ et $u_p = 0$ (demi-canal). La figure 5.13 illustre le processus itératif de convergence de la partie

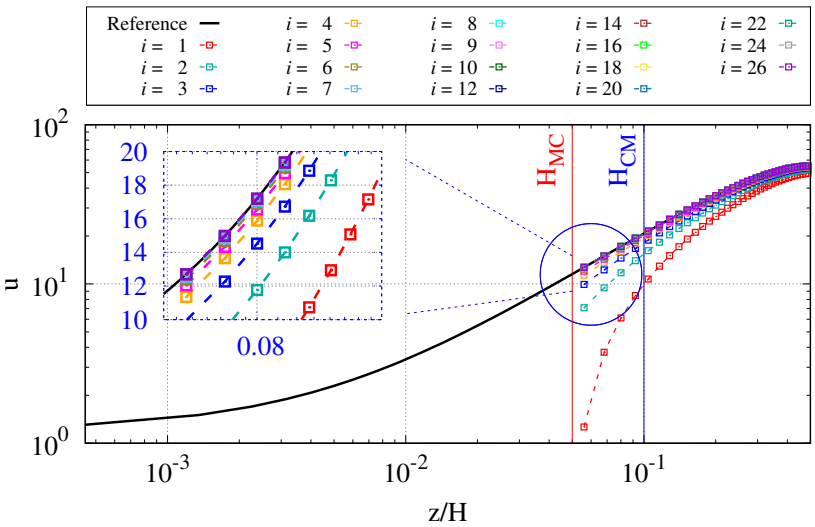


FIGURE 5.13 – Convergence itérative du couplage hybride vers la solution exacte de Poiseuille pour $p_{in} = 2.10^5$, $p_{out} = 5.10^4$, $H_{M\to C}/H = 10\%$ et $(H_{C\to M} - H_{M\to C})/H = 10\%$, pour un cas incompressible et un glissement constant, avec 80 mailles dans la section de hauteur $H - H_{M\to C}$. La valeur i indique le numéro de l'itération.

continue vers la solution exacte, pour $H_{M\to C}/H = 10\%$ et $(H_{C\to M} - H_{M\to C})/H = 10\%$.

Le schéma volumes finis étant d'ordre 2, la solution de référence volumes finis $\bar{u}_{VF}$ pour calculer l'écart $\varepsilon_\infty^{(c)}$ n'est donc plus exacte. La table 5.1 indique l'écart relatif à la solution

maillage (H)	5	10	20	40
$\varepsilon_\infty^{(c)}$	2.60×10^{-2}	2.24×10^{-2}	5.28×10^{-3}	1.05×10^{-3}
maillage	80	160	320	640
$\varepsilon_\infty^{(c)}$	1.82×10^{-4}	5.85×10^{-5}	1.16×10^{-5}	3.04×10^{-6}

TABLE 5.1 – Erreur de couplage $\varepsilon_\infty^{(c)}$ à la solution exacte en fonction du maillage de la zone continue, pour une erreur $\varepsilon_\infty^{(i)} = 10^{-8}$, $H_{M\to C}/H = 10\%$ et $(H_{C\to M} - H_{M\to C})/H = 5\%$. Cas incompressible et glissement constant.

volumes finis $\varepsilon_\infty^{(c)}$ pour différents nombres de mailles de la partie continue du modèle hybride et un critère d'arrêt $\varepsilon_\infty^{(i)} = 10^{-8}$. Quel que soit le maillage de la région continue, le nombre d'itérations est toujours le même, $N^{(i)} = 27$. À l'exclusion du maillage le plus grossier, la convergence de la solution hybride est globalement d'ordre deux. À noter que la substitution de la solution de référence volumes finis par la solution de Poiseuille analytique donne des erreurs tout à fait similaires.

Pour une erreur relative entre deux itérations $\varepsilon_\infty^{(i)} = 10^{-8}$, la figure 5.14 présente le nombre

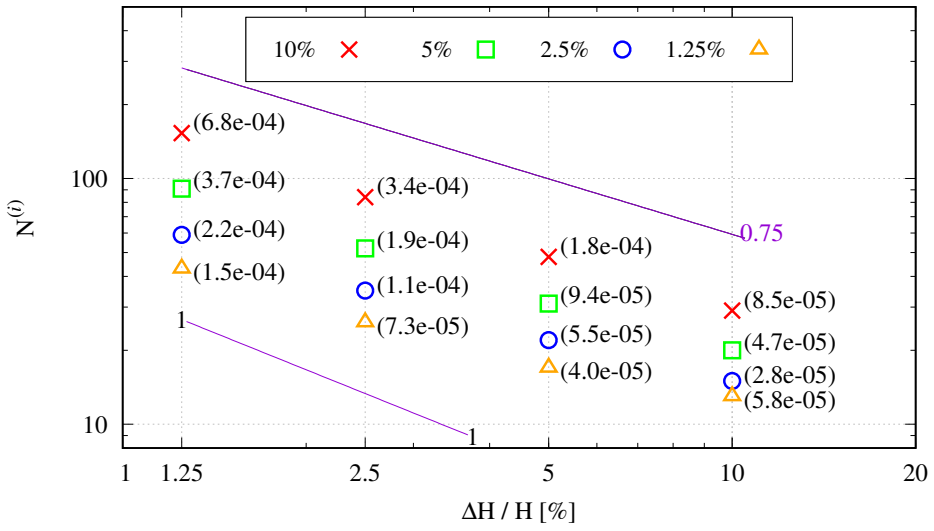


FIGURE 5.14 – Nombre d'itérations de couplage $N^{(i)}$ pour atteindre un écart relatif entre deux itérations successives $\varepsilon_\infty^{(i)} = 10^{-8}$ en fonction de la différence entre les hauteurs de communication $\Delta H/H = (H_{C\to M} - H_{C\to M})/H$. Les valeurs entre parenthèses indiquent l'erreur relative à la solution exacte $\varepsilon_\infty^{(c)}$. Écoulement de Poiseuille incompressible et glissement constant avec 80 mailles sur une hauteur $H - H_{M\to C}$ (conditions de symétrie).

d'itérations $N^{(i)}$ et l'écart relatif à la solution exacte $\varepsilon_\infty^{(c)}$ pour différentes valeurs de la distance entre les zones de communication $\Delta H = H_{C\to M} - H_{M\to C}$ et de la hauteur de communication du domaine "moléculaire" à celui continu $H_{M\to C}$. Globalement, les mêmes comportements que ceux analysés pour l'écoulement de Couette sont retrouvés : $N^{(i)} \approx (\Delta H)^{-0.75}$, dégradation de la qualité de la solution hybride avec la diminution de ΔH et l'augmentation de $H_{M\to C}$.

Pour $H_{M\to C}/H = 5\%$, l'écart à la référence $\varepsilon_\infty^{(c)}$ décroit puis atteint un palier avec la diminution de $\varepsilon_\infty^{(i)}$ et celui-ci est d'autant plus faible que $H_{C\to M} - H_{M\to C}$ est grand (Fig. 5.15). De même, pour $(H_{C\to M} - H_{M\to C})/H = 1.25$, la diminution de $H_{M\to C}$ se traduit par une

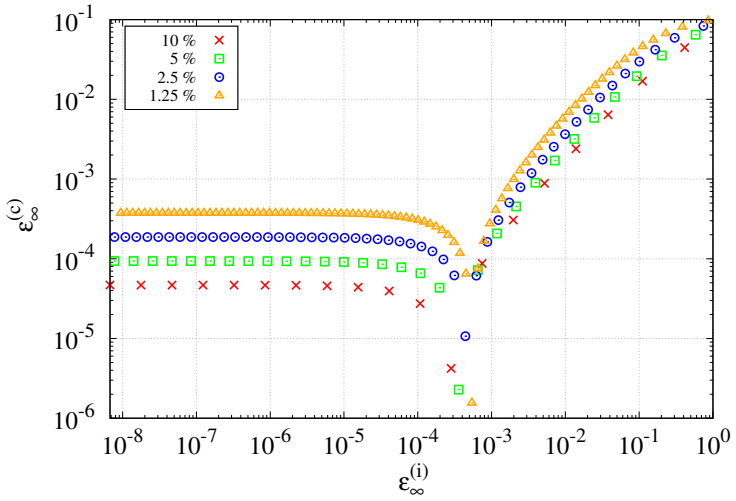


FIGURE 5.15 – Erreur de convergence $\varepsilon_\infty^{(c)}$ en fonction de $\varepsilon_\infty^{(i)}$, pour $(H_{C\to M} - H_{M\to C})/H = 1.25\%, 2.5\%, 5\%, 10\%$ et $H_{M\to C}/H = 5\%$. Écoulement de Poiseuille incompressible et glissement constant avec 80 mailles sur une hauteur $H - H_{M\to C}$ (conditions de symétrie).

réduction de l'écart relatif $\varepsilon_\infty^{(c)}$ avant qu'un plateau soit atteint dont le niveau décroît avec $H_{M\to C}$ (Fig. 5.16). Le niveau de saturation obtenu pour les courbes des deux graphes pré-

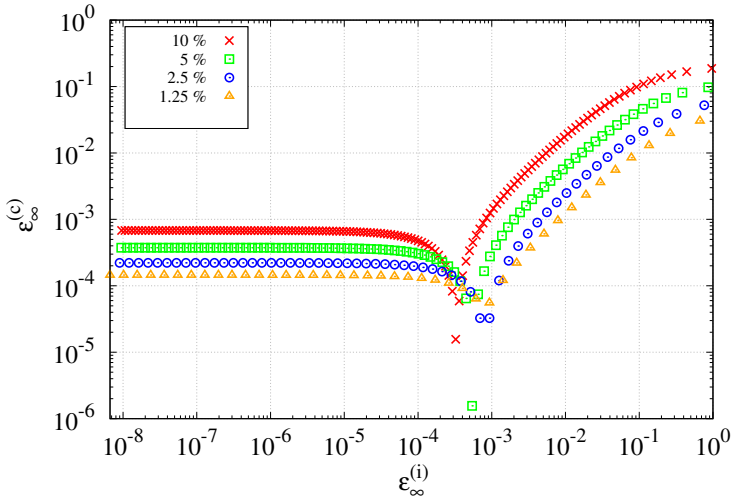


FIGURE 5.16 – Erreur de convergence $\varepsilon_\infty^{(c)}$ en fonction de $\varepsilon_\infty^{(i)}$, pour $H_{M\to C}/H = 1.25\%, 2.5\%, 5\%, 10\%$ et $(H_{C\to M} - H_{M\to C})/H = 1.25\%$. Écoulement de Poiseuille incompressible et glissement constant avec 80 mailles sur une hauteur $H - H_{M\to C}$ (conditions de symétrie).

cédents dépend naturellement du maillage de la solution hybride dans la région continue : les plateaux seront d'autant plus bas que les maillages seront fins, car la solution volumes finis dans le cœur de l'écoulement tendra alors vers la solution analytique utilisée dans le plot “moléculaire” virtuel.

5.4.4.5 Conclusion

En l'absence des processus d'apprentissage et des simulations de dynamique moléculaire qui peuvent tous deux introduire des niveaux d'approximation supplémentaires et donc réduire la précision de la méthode numérique hybride, la qualité des solutions hybrides et les coûts de calcul ont été évalués à travers des études de convergence vers des solutions élémentaires de Couette-Poiseuille et les nombres d'itérations requis pour les atteindre.
L'adoption de simulations de type Couette-Poiseuille dans la couche de Knudsen implique nécessairement que les régions de connexion aux hauteurs $H_{C\to M}$ et $H_{M\to C}$ doivent rester faibles, de l'ordre du libre parcours moyen des molécules du gaz. Or, d'après les situations élémentaires étudiées, la réduction de la distance $H_{C\to M} - H_{M\to C}$ dégrade notablement la qualité de la solution, et ce, pour un nombre d'itérations $N^{(i)}$ d'autant plus élevé que cette distance se réduit. L'approche explicite se révèle donc coûteuse et peu précise, même en l'absence de l'apprentissage et de la dynamique moléculaire.

Pour remédier à ces deux inconvénients, une méthode hybride implicite avec des plots virtuels a été mise en œuvre et fait l'objet de la section suivante.

5.5 Couplage faible hybride multi-échelle implicite par la méthode des plots virtuels

5.5.1 Du couplage explicite au couplage implicite : éléments principaux d'amélioration

La méthode de couplage faible implicite par la méthode des plots virtuels vise à améliorer l'efficacité du couplage entre les échelles et ainsi réduire le nombre d'itérations à réaliser entre les plots virtuels et le domaine continu pour converger vers la solution hybride. Les principaux points d'amélioration entre les approches explicite et implicite sont les suivants :

De l'explicite vers l'implicite. Le couplage explicite entre les échelles moléculaires et continues est aisé à mettre en œuvre, car les communications s'effectuent itérativement via la transmission de simples valeurs numériques entre les codes de calcul (vitesses, pression, ...). L'inconvénient de ce processus est que la convergence est lente, d'autant plus lente que les régions d'échange sont voisines. De plus, la qualité de la solution hybride ainsi obtenue se dégrade avec le rapprochement des régions de communication.

La mise en place de la version implicite substitue la transmission de valeurs par celle d'une fonction au modèle continu. Cette implicitation permet ainsi d'adapter automatiquement la réponse issue du plot virtuel à la sollicitation transmise du milieu continu vers les petites échelles lors de la résolution du solveur volumes finis.

De l'échange entre deux points distants au couplage ponctuel. L'échange de données dans la méthode explicite requière que les points de communication $P_{C\to M}$ et $P_{M\to C}$ soient proches, à la fois l'un de l'autre et de la paroi, pour que le modèle de Couette-Poiseuille soit pertinent. Outre l'augmentation importante du nombre d'itérations $N^{(i)}$ avec le rapprochement des régions d'échange, cela engendre également

des difficultés dans les évaluations des vitesses macroscopiques issues de la dynamique moléculaire dont les valeurs deviennent très voisines et donc fortement sensibles aux incertitudes liées à l'évaluation de ces moyennes.

La version implicite ne repose plus sur la communication de vitesses entre deux points proches, mais sur la construction d'une relation entre la vitesse et le taux de cisaillement au même point. Ainsi, à partir des simulations de dynamique moléculaire, le cisaillement macroscopique est extrait et vient enrichir la fonction apprise qui permet l'évaluation du cisaillement local en fonction de la vitesse locale et du nombre de Knudsen de maille.

Dynamique moléculaire : du problème de Couette-Poiseuille à celui de Couette. Lors du couplage avec le schéma explicite, le gradient de pression est pris en compte dans le modèle car les régions d'échange peuvent être éloignées de la paroi solide. Ainsi, un problème de Couette-Poiseuille doit être considéré lors des simulations de dynamique moléculaire.

Dans la version implicite, le couplage se réalise entre la maille volumes finis adjacente à la paroi et la paroi : la contribution de courbure du profil de vitesse liée au gradient de pression devient donc négligeable. Ainsi, les simulations de dynamique moléculaire ne sont réalisées que pour un problème de Couette.

Du couplage de débit entre les domaines à sa conservation implicite. La subdivision du canal en deux régions, l'une continue et l'autre moléculaire dans des plots virtuels découplés entre eux, requière que la frontière du domaine continu soit perméable pour que la différence de débits entre des plots adjacents puisse être compensée par un apport ou un prélèvement de débit du domaine continu.

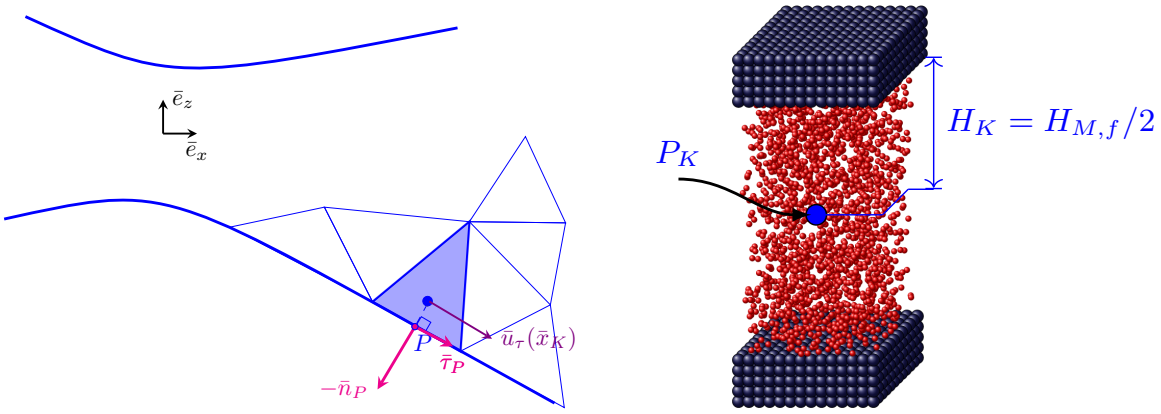
Dans le modèle implicite, le domaine continu recouvre la totalité du canal de sorte que la conservation du débit est simplement assurée par la résolution des équations de la mécanique des fluides : aucune vitesse normale à la paroi n'est reconstruite.

5.5.2 Couplage hybride faible implicite par la méthode des plots virtuels

Pour le couplage implicite, le domaine continu occupe la totalité du canal physique et les plots virtuels recouvrent la région continue dans le voisinage proche des parois, sur l'épaisseur de la couche de Knudsen (Fig. 5.17). L'échange d'informations entre la grande et la petite échelle s'effectue au point P_K, à la coordonnée $\bar{x}_K$, correspondant au centre des mailles K ayant une arête commune avec la frontière solide. Soit P la projection orthogonale du point P_K sur la paroi du canal. On note $\bar{n}_P$ le vecteur unitaire normal à la paroi au point P et pointant vers le fluide, et $\bar{\tau}_P$ le vecteur tangent tel que $\bar{\tau}_P \wedge \bar{n}_P = \bar{e}_x \wedge \bar{e}_z$. On définit les composantes de vitesse $u_\tau = \bar{u}.\bar{\tau}_P$ et $u_n = \bar{u}.\bar{n}_P$ et la fonction

$$f_{d_n u_\tau}(u_\tau(\bar{x}_K), \mathrm{Kn}_K)$$

modélisant la réponse apprise $d_n u_\tau := \nabla u_\tau.\bar{n}_p$ à l'abscisse $\bar{x}_K$ du plot virtuel soumis à une sollicitation de vitesse $u_\tau(\bar{x}_K)$, pour un nombre de Knudsen de maille $\mathrm{Kn}_K = \lambda/\|\bar{x}_K - \bar{x}_P\|$. Contrairement à la méthode explicite où la pression p et deux longueurs $H_{C\to M}$ et $H_{M\to C}$ étaient utilisées parmi les variables d'entrée des fonctions d'apprentissage, la version implicite



(a) Canal de section variable (b) Dynamique moléculaire dans un plot virtuel

FIGURE 5.17 – Couplage explicite faible, par la méthode des plots virtuels.

s'appuie sur la variable adimensionnelle Kn_K qui donne le rapport entre le libre parcours moyen λ des molécules de gaz, lequel s'exprime en fonction de la pression, et la distance entre le centre de la maille P_K et le point P de la paroi. L'utilisation de ce nombre de Knudsen de maille permet de réduire le nombre de variables des fonctions apprises. Il faut noter que la valeur du nombre de Knudsen de maille doit vérifier l'inégalité $\mathrm{Kn}_K < 1$ pour assurer que le point P_K se trouve à l'extérieur de la couche de Knudsen.
Le principe du couplage implicite est très similaire à celui explicite et est illustré dans la figure 5.18 :

Étape 1 : Résolution du problème continu.

L'expression des conditions aux limites du problème continu étant connues à l'itération i de couplage pour tout point P de la frontière, c'est-à-dire à travers la fonction apprise $f^{(i)}_{d_n u_\tau}(u^{(i)}_\tau, \mathrm{Kn}^{(i)}_K)$, les équations de Navier-Stokes discrétisées en volumes finies sont résolues pour donner la nouvelle solution à l'itération i.

Étape 2 : Extraction de grandeurs macroscopiques.

La vitesse tangente $u^{(i)}_\tau = \bar{u}^{(i)} \cdot \bar{\tau}_p$, et le nombre de Knudsen de maille $\mathrm{Kn}^{(i)}_K$ (fonction de la pression $p^{(i)}$ et la hauteur $\|\bar{x}_K - \bar{x}_P\|$) sont extraits sur l'ensemble des points P_K des mailles adjacentes K à la paroi.

Étape 3 : Construction de la réponse moléculaire dans le plot virtuel

Selon les vitesses $u^{(i)}_\tau(\bar{x}_K)$ et les nombres de Knudsen de maille $\mathrm{Kn}^{(i)}_K$ sur l'ensemble des points P_K, une, deux ou trois simulations dans une boite de dynamique moléculaire d'un problème de Couette sont réalisées pour enrichir les fonctions apprises $f^{(i)}_{d_n u_\tau}$ et $f^{(i)}_{u_w}$, et ainsi fabriquer les nouvelles fonctions $f^{(i+1)}_{d_n u_\tau}$ et $f^{(i+1)}_{u_w}$:

— $f^{(i+1)}_{d_n u_\tau}$ est la fonction apprise à l'itération $i+1$ reliant le cisaillement moyen au point P_K du plot virtuel moléculaire à la vitesse moyenne $u^{(i+1)}_\tau(\bar{x}_K)$ et au nombre de Knudsen de maille $\mathrm{Kn}^{(i+1)}_K$:

$$f^{(i+1)}_{d_n u_\tau}(u^{(i+1)}_\tau(\bar{x}_K), \mathrm{Kn}^{(i+1)}_K)$$

— $f^{(i+1)}_{u_w}$ est la fonction apprise à l'itération $i+1$ pour effectuer les simulations

de Couette dans une boite de dynamique moléculaire de hauteur fluide $H_{M,f}$ (choix arbitraire, mais généralement choisi égal à $2\|\bar{x}_K - \bar{x}_P\|$). Elle fournit la vitesse horizontale moyenne de la paroi solide supérieure de sorte que la vitesse moyenne macroscopique $u_\tau(\bar{x}_K)$ soit prescrite à la hauteur $H_{M,f}/2$, pour une pression moyenne p qui corresponde exactement au nombre de Knudsen de maille Kn_K sur la longueur de référence $H_{M,f}/2$:

$$f_{u_w}^{(i+1)}(u_\tau(\bar{x}_K), \mathrm{Kn}_K)$$

Ces deux fonctions apprises $f_{d_n u_\tau}^{(i+1)}$ et $f_{u_w}^{(i+1)}$ sont reconstruites par des méthodes d'apprentissages supervisés. La construction de la réponse moléculaire, qui couple les étapes d'apprentissage et les simulations de dynamique moléculaire, est décrite plus en détails à la section 5.5.3

Si aucune simulation de dynamique moléculaire n'est nécessaire, le processus d'apprentissage est achevé et le couplage entre les échelles physiques est résolu. **L'algorithme itératif de la méthode hybride est arrêté**.

Étape 4 : Reconstruction des conditions aux limites du problème continu.

Pour chacune des faces de frontière des mailles volumes finis, les composantes de la vitesse $\bar{u}^{(i)}$ à l'itération $i+1$ doivent être spécifiées. Dans le repère local à la paroi, la composante normale est nulle, $u_n^{(i+1)}(\bar{x}_P) = 0$. La nouvelle composante tangentielle de la vitesse $\bar{u}^{(i+1)}$ selon $\bar{\tau}_P$ est définie par :

$$u_\tau^{(i+1)}(\bar{x}_P) = u_\tau^{(i+1)}(\bar{x}_K) - f_{d_n u_\tau}^{(i+1)}(u_\tau^{(i+1)}(\bar{x}_K), \mathrm{Kn}_K^{(i+1)}) \times \|\bar{x}_P - \bar{x}_K\| \tag{5.11}$$

Cette relation permet de reconstruire une vitesse de glissement $u_\tau^{(i+1)}(\bar{x}_P)$ en utilisant le cisaillement déterminé sur le point P_K. Dans cette relation de glissement, $u_\tau^{(i+1)}(\bar{x}_K)$ et $\mathrm{Kn}_K^{(i+1)}$ (c'est-à-dire la pression $p^{(i+1)}(\bar{x}_K)$) sont inconnus et doivent être calculés simultanément avec les équations de Navier-Stokes discrétisées par la méthode des volumes finis. Ramenées dans le repère $(\bar{e}_x, \bar{e}_z)$, les composantes $u^{(i+1)}$ et $v^{(i+1)}$ de la vitesse $\bar{u}^{(i+1)}$ s'écrivent donc en frontière du domaine continu :

$$u^{(i+1)}(\bar{x}_P) = u_\tau^{(i+1)}(\bar{x}_P)\bar{\tau}_P.\bar{e}_x$$

et

$$v^{(i+1)}(\bar{x}_P) = u_\tau^{(i+1)}(\bar{x}_P)\bar{\tau}_P.\bar{e}_y$$

Étape 5 : Incrément de l'itération de couplage i ($i \leftarrow i+1$) et Retour à l'Étape 1.

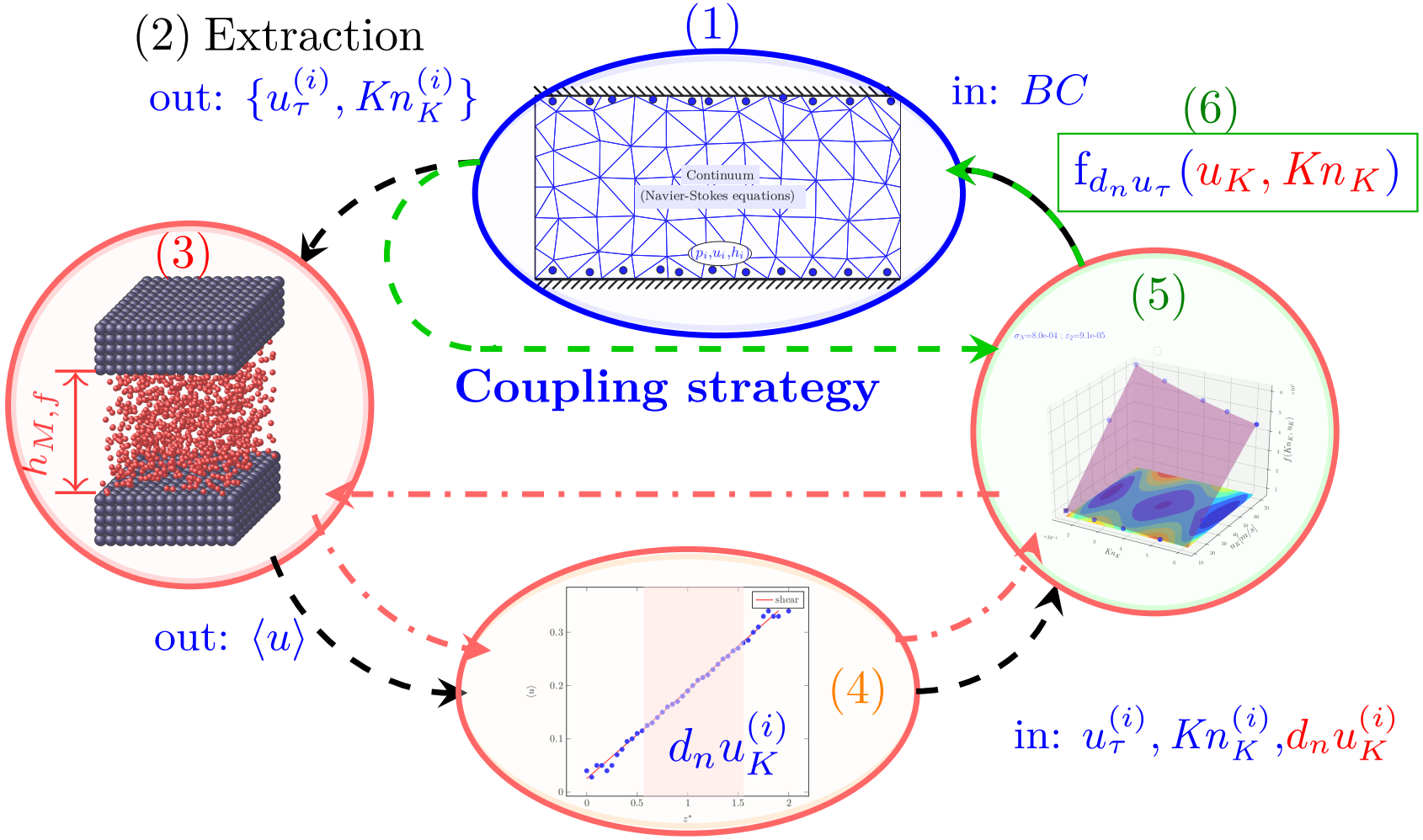


FIGURE 5.18 – Représentation graphique du couplage faible hybride implicite par la méthode des plots virtuels. Les labels (1), (2) correspondent aux étapes **1** et **2** de l'algorithme, page 136. Les labels (3), (4), (5) (encadré en rouge) schématisent la construction de la réponse moléculaire dans le plot virtuel (**Étape 3**). Les labels (3) et 4 font référence à la dynamique moléculaire ((Sec. 5.5.3.2), page 141) et le label (5) se rapporte à l'algorithme d'apprentissage ((Sec. 5.5.3.1)). Le cadre vert (label (6)) se réfère à l'**Étape 4** de l'algorithme hybride implicite. Les flèches intermédiaires montrent les différentes possibilités qui existent : ainsi, lorsque les bornes [min, max] des valeurs extraites ne varient pas, les labels (2), (3) et (4) sont ignorés. Le processus d'apprentissage ("label (5)") est directement mis en œuvre (la flèche verte au milieu). Á chaque passage dans l'algorithme d'apprentissage (label (5)) un nouveau point est appris, il est alors opportun de définir (calculer) la cible d_{nu_K} par l'intermédiaire de la dynamique moléculaire (flèche rouges).

5.5.3 Construction de la réponse du plot virtuel

5.5.3.1 Algorithme d'apprentissage des fonctions f_{u_τ} et f_{u_w}

L'algorithme d'apprentissage repose toujours sur les méthodes de régression Bayésienne ou Gaussienne présentée au chapitre 4. Outre quelques ajustements, il est très similaire à celui présenté pour le couplage explicite (Sec. 5.4.3.1, page 121).

Les données d'entrée de notre problème d'apprentissage sont la vitesse u_K selon $\bar{e}_\tau$ à la hauteur $H_{M,f}/2$ et le nombre de Knudsen local Kn_K. Les grandeurs apprises dans la boite moléculaire, soumise à des contraintes générant un problème de Couette, sont le cisaillement moyen $d_n u_K$ à la hauteur $H_{M,f}/2$ et la vitesse de la paroi mobile u_w selon $\bar{e}_\tau$ à la hauteur $H_{M,f}$.

La mise en place de l'algorithme d'apprentissage à l'**Étape 3** de la procédure de couplage hybride entre les solveurs continu et discret (Sec. 5.4.2) requiert une phase d'initialisation précédent l'**Étape 1**. Comme pour la version explicite, elle est définie comme l'**étape 0** :

Étape 0 : Initialisation de l'algorithme d'apprentissage.

- Données d'entrée :
  - Nombre maximal d'apprentissages supervisés $cpt_{max} > 0$ pour des bornes fixées du plan d'apprentissage.
  - Précisions relatives ε_{u_K} et $\varepsilon_{\mathrm{Kn}_K}$ sur les bornes des grandeurs d'entrée délimitant les intervalles des données d'apprentissage u_K et Kn_K.
- Initialisation des variables à l'itération $i = 0$:
  - Définition initiale de la fonction $f_{u_w}^{(0)}$ donnant la vitesse de la paroi u_w pour le problème moléculaire de Couette. Faute d'information à cette étape, la valeur constante u_K est utilisée :

$$f_{u_w}^{(0)}(u_K, \mathrm{Kn}_K) = u_K$$

  - Initialisation de la base potentielle d'apprentissage B_{pot} à l'ensemble vide $B_{pot} = \varnothing$. La base B_{pot} contient par la suite l'ensemble des quadruplets (2 entrées et 2 sorties) de la forme :

$$(2 \text{ entrées} \,;\, 2 \text{ sorties}) = (u_K, \mathrm{Kn}_K \,;\, d_n u_K, u_w)$$

    La dernière variable u_w du quadruplet n'est utilisée que pour les simulations dans les boites de dynamique moléculaire.
  - Initialisation de deux couples (vitesse, Knudsen de maille) pour délimiter la zone des données d'apprentissage par valeurs inférieure $X_m^{(new)} = (-\infty, -\infty)$ et supérieure $X_M^{(new)} = (+\infty, +\infty)$.
  - Initialisation du compteur d'apprentissages supervisés $cpt = 0$.

À l'issue de l'**Étape 2** (Sec. 5.5.2), les valeurs de vitesse $u_\tau^{(i)}(\bar{x}_K)$ et du nombre de Knudsen de maille $\mathrm{Kn}_K^{(i)}$ sont connues pour tous les points P_K des mailles adjacentes aux parois. La procédure d'apprentissage des réponses en cisaillement $f_{d_n u_\tau}^{(i+1)}$ et en vitesse de paroi $f_{u_w}^{(i+1)}$ est alors la suivante :

Boucle sur le nombre d'apprentissages. Tant que $cpt \leq cpt_{max}$,

1. Définition des bornes des données délimitant la zone d'apprentissage.
   (a) Sauvegarde des anciennes bornes délimitant la zone d'apprentissage $X_m^{(old)} = X_m^{(new)}$, $X_M^{(old)} = X_M^{(new)}$.
   (b) Extraction des nouvelles bornes délimitant les intervalles des données d'apprentissage :

$$X_m^{(new)} = \begin{pmatrix} \min\limits_{P_K} u_\tau^{(i)}(\bar{x}_K), \\ \min\limits_{P_K} \mathrm{Kn}_K^{(i)} \end{pmatrix}^T, \qquad X_M^{(new)} = \begin{pmatrix} \max\limits_{P_K} u_\tau^{(i)}(\bar{x}_K), \\ \max\limits_{P_K} \mathrm{Kn}_K^{(i)} \end{pmatrix}^T$$

2. Comparaison entre les anciennes $X_{k,j}^{(old)}$ et nouvelles $X_{k,j}^{(new)}$ composantes des bornes délimitant les intervalles des données d'apprentissage, pour

$$k \in \{m, M\} \text{ et } j \in \{u_K, \mathrm{Kn}_K\}$$

   puis éventuels calculs dans des boites de dynamique moléculaire.

   Pour les deux bornes inférieure et supérieure délimitant la zone d'apprentissage, faire pour $k \in \{m, M\}$

   - Pour $j \in \{u_K, \mathrm{Kn}_K\}$, si $|X_{k,j}^{(new)}/X_{k,j}^{(old)} - 1| > \varepsilon_j$ alors
     (a) Réalisation d'une simulation dans une boite de dynamique moléculaire (voir Sec. 5.5.3.2) :
     — Données d'entrée : $\mathrm{Kn}_K = X_{k,\mathrm{Kn}_K}^{(new)}$ et la vitesse de la paroi mobile issue de la fonction apprise f_{u_w} :

$$u_w = f_{u_w}^{(i)}(X_{k,u_K}^{(new)}, \mathrm{Kn}_K)$$

     — Données de sortie : u_K et $d_n u_K$.
     (b) Ajout d'un quadruplet dans la base potentielle d'apprentissage :

$$B_{pot} = B_{pot} \cup (u_K, \mathrm{Kn}_K \,;\, d_n u_K, u_w)$$

     (c) Les bornes de la zone d'apprentissage ayant été modifiées, le compteur d'apprentissages supervisés est remis à zéro, $cpt = 0$.

3. Apprentissage supervisé :
   (a) Construction de la base réelle d'apprentissage B_{app} par extraction de quadruplets de la base d'apprentissage potentielle B_{pot}. Seuls les quadruplets dont

les données sont dans l'intervalle $[X_{m,j}^{(new)}, X_{M,j}^{(new)}]$, pour $j \in \{u_K, \mathrm{Kn}_K\}$, sont considérés dans l'étape d'apprentissage.

$$B_{app} = \Big\{ (u_K, \mathrm{Kn}_K \,;\, d_n u_K, u_w) \in B_{pot} \quad \text{tel que } j \in [X_{m,j}^{(new)}(1-\varepsilon_j), X_{M,j}^{(new)}(1+\varepsilon_j)] \quad \text{pour } j \in \{u_K, \mathrm{Kn}_K\} \Big\} \tag{5.12}$$

(b) Calcul de la nouvelle régression $f_{u_w}^{(i+1)}$ sur la base d'apprentissage B_{app} permettant de définir la vitesse moyenne moléculaire de la paroi mobile en $H_{M,f}$. Pour rappel, ce choix permet de contrôler la vitesse moléculaire moyenne u_K en $H_{M,f}/2$.

(c) Détermination du couple (u_K, Kn_K) engendrant un cisaillement $d_n u_K$ le plus incertain sur la base d'apprentissage B_{app}, c'est-à-dire le couple de données produisant la plus grande variance sur $d_n u_K$.

(d) Réalisation d'une simulation de dynamique moléculaire (voir Sec. 5.5.3.2) :

— Données d'entrée : Kn_K et la vitesse de la paroi mobile issue de la fonction d'apprentissage :

$$u_w = f_{u_w}^{(i+1)}(u_K, \mathrm{Kn}_K)$$

— Données de sortie : u_K et $d_n u_K$.

(e) Ajout du quadruplet $(u_K, \mathrm{Kn}_K \,;\, d_n u_K, u_w)$ dans les 2 bases d'apprentissage :

$$\begin{aligned} B_{pot} &= B_{pot} \cup (u_K, \mathrm{Kn}_K \,;\, d_n u_K, u_w) \\ B_{app} &= B_{app} \cup (u_K, \mathrm{Kn}_K \,;\, d_n u_K, u_w) \end{aligned} \tag{5.13}$$

(f) Les bornes de la zone d'apprentissage n'ayant pas été modifiées, le compteur d'apprentissages supervisés est incrémenté,

$$cpt = cpt + 1$$

(g) Calcul (ou mise à jour) des deux nouvelles régressions sur la base d'apprentissage B_{app} :

— $f_{d_n u_\tau}^{(i+1)}(u_K, \mathrm{Kn}_K)$ la réponse apprise de la vitesse moyenne moléculaire en $H_{M,f}/2$.

— $f_{u_w}^{(i+1)}(u_K, \mathrm{Kn}_K)$ la réponse apprise de la vitesse moyenne moléculaire de la paroi mobile en $H_{M,f}$.

5.5.3.2 Dynamique moléculaire dans un plot virtuel

La mise en place, la simulation et l'exploitation des résultats sont très similaires à ce qui a été décrit pour la méthode explicite à la section 5.4.3.2, page 124. Nous ne présenterons donc dans cette section que les aspects spécifiques de notre dynamique moléculaire lié au

caractère implicite du couplage. En particulier, l'écoulement étudié ici s'effectue en l'absence d'un gradient de pression : c'est un écoulement de Couette plan.
Les paramètres d'entrée/sortie liés à l'apprentissage sont les suivants :

Données d'entrée :

- le Nombre de Knudsen de maille $\mathrm{Kn}_K = 2\lambda/H_{M,f}$ permet de reconstruire la densité moyenne du gaz dans la boite dont la hauteur fluide $H_{M,f}$ est choisie par l'utilisateur ($2\|\bar{x}_K - \bar{x}_P\|$ généralement),
- la vitesse de la paroi/mur mobile u_w selon $\bar{e}_\tau$ à la hauteur $H_{M,f}$.

Données de sortie :

- la vitesse moyenne macroscopique en $H_{M,f}/2$ selon $\bar{e}_\tau$,
- le cisaillement moyen macroscopique en $H_{M,f}/2$.

L'évaluation du cisaillement moyen macroscopique à la hauteur $H_{M,f}/2$ est obtenue en deux étapes :

1. Les vitesses moyennes macroscopiques sont d'abord évaluées dans chacune des tranches (bins) de hauteur δz,
2. Une régression linéaire basée sur les vitesses moyennes et centrée en $H_{M,f}/2$ est réalisée sur l'intervalle $[0.3H_{M,f}; 0.7H_{M,f}]$ pour extraire le cisaillement de la vitesse tangentielle à la paroi loin des frontières solides de la boite moléculaire.

5.5.4 Validation de la procédure de couplage itératif, performance et coûts de calcul

5.5.4.1 Problème

Nous considérons à nouveau le canal plan de hauteur $2H = 6\,\mu m$ selon $\bar{e}_z$ et de longueur $L = 100H$ selon $\bar{e}_x$ (Fig. 5.8, page 128). L'écoulement de Poiseuille, compressible, glissant et isotherme à $T_{in} = 300\,K$ est engendré par une différence de pression entre l'entrée $p_{in} = 2 \times 10^5\,Pa$ et la sortie $p_{out} = 5 \times 10^4\,Pa$. Des conditions de symétrie sont imposées en $z = H$ (voir Sec. 5.4.4.2, page 127).

Comme pour le couplage faible explicite, nous allons dans un premier temps nous affranchir des simulations dans les boites de dynamique moléculaire et les remplacer par un modèle analytique moins coûteux à résoudre. Dans ce qui suit, les longueurs seront adimensionnées par la hauteur $H_K = \|\bar{x}_K - \bar{x}_P\|$ pour faire apparaitre explicitement le nombre de Knudsen de maille Kn_K.
Nous considérerons dans le plot virtuel un écoulement de Couette glissant pour lequel la vitesse de la paroi inférieure est nulle et celle de la paroi supérieure en $z = 2$ est ajustée à $u_w = f_{u_w}(u_K, \mathrm{Kn}_K)$, de sorte que la vitesse $u_\tau(z = 1)$ soit prescrite à la valeur u_K :

$$u_\tau(z) = \frac{f_{u_w}(u_K, \mathrm{Kn}_K)}{2 + 2\sigma_u \mathrm{Kn}_K}\,(z + \sigma_u \mathrm{Kn}_K) \tag{5.14}$$

avec $\sigma_u = 1$.

Fonction f_{u_w}. Dans la relation (5.14), f_{u_w} est une fonction apprise par la méthode Bayésienne ou Gaussienne. La base des données d'apprentissage B_{app} (voir Eqs. 5.12-5.13, page 141) utilise donc les triplets $(u_K, \mathrm{Kn}_K, u_w)$, avec u_K redéfini par la relation ;

$$u_K = u_\tau(z = 1) \tag{5.15}$$

Cette relation détermine la vraie vitesse à mi-hauteur de l'écoulement de Couette glissant (Eq. 5.14). C'est la valeur issue de la relation (5.15) qui doit être utilisée dans la base d'apprentissage pour f_{u_w} et $f_{d_n u_\tau}$, et non la valeur u_K utilisée dans la fonction f_{u_w} de l'équation (5.14).

Fonction $f_{d_n u_\tau}$. Pour l'apprentissage de la fonction $f_{d_n u_\tau}$, nous utilisons également une méthode Bayésienne ou Gaussienne. La base des données d'apprentissage B_{app} (voir Eqs. 5.12-5.13, page 141) pour la fonction $f_{d_n u_\tau}$ utilise les triplets $(u_K, \mathrm{Kn}_K, d_n u_K)$ où

$$d_n u_K := \left.\frac{du_\tau}{dz}\right|_{z=1} = \frac{u_K}{1 + \sigma_u \mathrm{Kn}_K} \tag{5.16}$$

et u_K est défini par la relation (5.15).

Il faut noter que si nous n'utilisons pas les fonctions apprises f_{u_w} et $f_{d_n u_\tau}$, mais directement les expressions mathématiques exactes, la convergence de la méthode hybride implicite s'effectue en deux itérations et la solution est exacte ($\varepsilon_\infty^{(c)} = 0$, Eq. 5.17) : la première correspond à l'initialisation avec des conditions d'adhérence et la seconde utilise le modèle exact de glissement à la paroi.

Afin de mesurer à la fois la qualité de l'apprentissage et de la solution hybride, et le comportement itératif de la solution de couplage, nous introduisons

- le vecteur $\bar{u}^{(i)}$, solution hybride exprimée sur les mailles volumes finis du domaine continu pour l'itéré i,
- le vecteur $\bar{u}_{VF}$, solution volumes finis obtenue sur tout le canal avec des conditions de Navier (5.7) à la paroi (même maillage que la solution hybride),
- la fonction exacte $f^{(e)}_{d_n u_\tau}(u_K, \mathrm{Kn}_K) = u_K/(1 + \sigma_u \mathrm{Kn}_K)$ à laquelle la fonction apprise $f_{d_n u_\tau}$ sera comparée.

Deux critères sont utilisés pour mesurer la qualité des solutions. On définit pour la norme l_∞

- l'écart relatif de l'itéré hybride $\bar{u}^{(i)}$ à la solution volumes finis $\bar{u}_{VF}$:

$$\varepsilon_\infty^{(c)} = \frac{\|\bar{u}^{(i)} - \bar{u}_{VF}\|_\infty}{\|\bar{u}_{VF}\|_\infty} \tag{5.17}$$

- l'écart relatif pour la fonction d'apprentissage $f_{d_n u_\tau}$:

$$\varepsilon_\infty^{(p)} = \frac{\max\limits_{(u_K, \mathrm{Kn}_K) \in \mathcal{I}} |f_{d_n u_\tau}(u_K, \mathrm{Kn}_K) - f^{(e)}_{d_n u_\tau}(u_K, \mathrm{Kn}_K)|}{\max\limits_{(u_K, \mathrm{Kn}_K) \in \mathcal{I}} |f^{(e)}_{d_n u_\tau}(u_K, \mathrm{Kn}_K)|}$$

avec $\mathcal{I} = [u_K^{(min)} : u_K^{(max)}] \times [\mathrm{Kn}_K^{(min)} : \mathrm{Kn}_K^{(max)}]$, avec $\bullet^{min,max}$ désignant les minima et maxima des variables sur B_{app}, la base courante des données d'apprentissage.

Les paramètres numériques du modèle hybride sont les suivants :

Le maillage est régulier et constitué de 3000×10 volumes de contrôle rectangulaires dans les directions $\bar{e}_x$ et $\bar{e}_z$.

L'initialisation du processus itératif est la solution volumes finis obtenue avec des conditions aux limites d'adhérence aux parois.

Les précisions relatives sur les bornes d'apprentissage sont fixées à $\varepsilon_{u_K} = \varepsilon_{\mathrm{Kn}_K} = 10\%$ (voir l'étape 0 intitulée "Initialisation de l'algorithme d'apprentissage", page 139).

5.5.4.2 Apprentissage par des processus gaussiens

Les noyaux utilisés dans la méthode d'apprentissage avec des processus gaussiens sont les fonctions Square Exponential (SE), Matern32 et Matern52 (voir Sec. 4.2.3, page 91). Ces fonctions sont utilisées pour apprendre la vitesse u_w de la paroi du problème de Couette ainsi que le cisaillement $d_n u_\tau$.

La figure 5.19 illustre l'écart relatif $\varepsilon_\infty^{(c)}$ de la méthode hybride à la solution de référence vo-

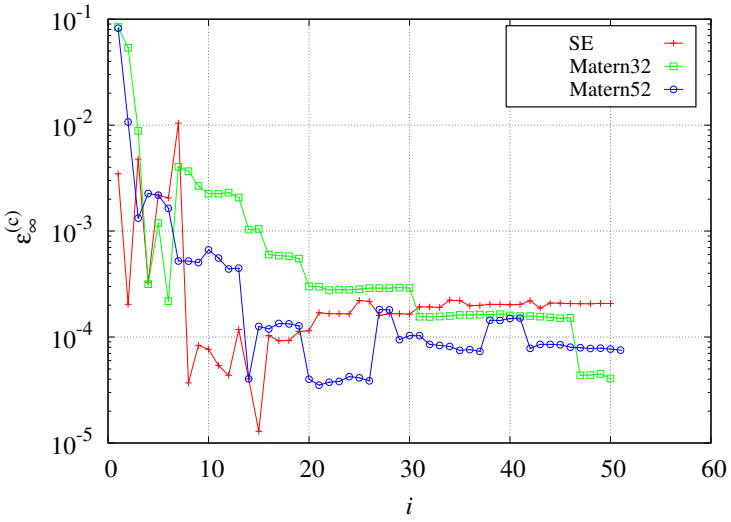
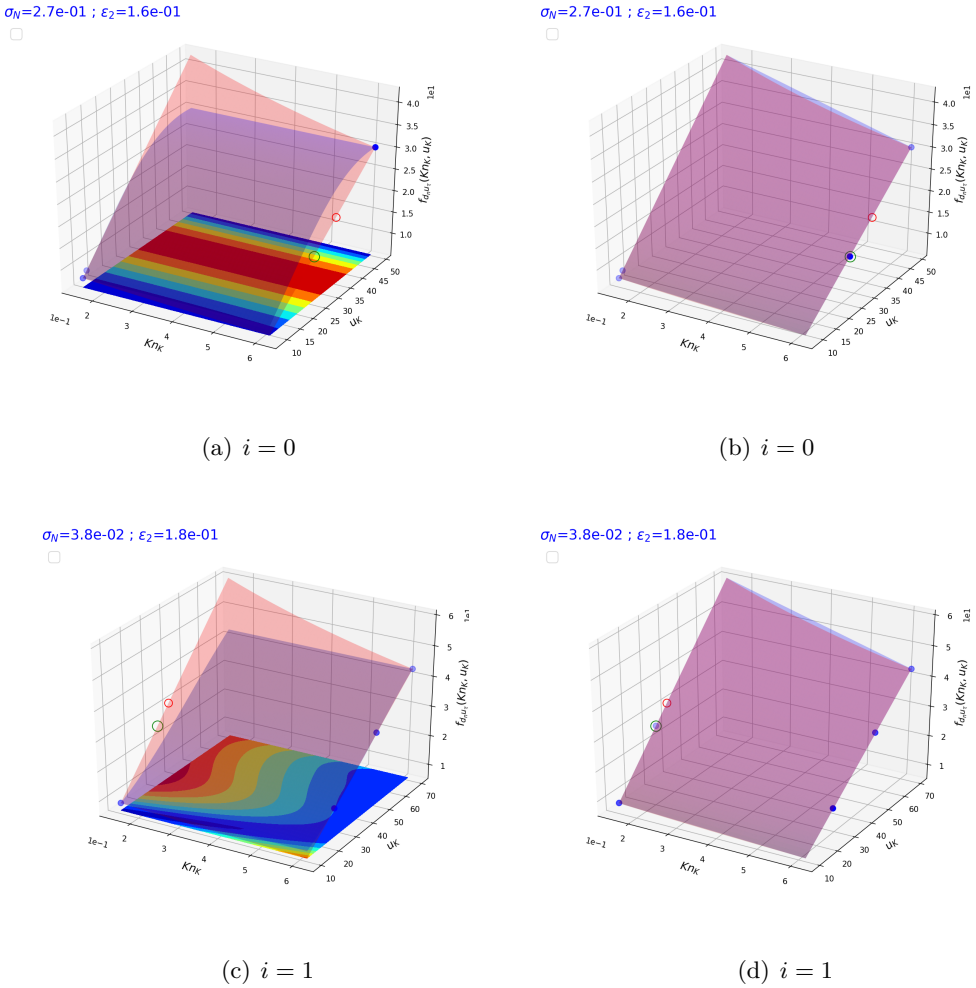


Figure 5.19 – Écart relatif de la solution hybride à la solution volumes finis en fonction du nombre d'itérations de couplage pour l'apprentissage des fonctions f_{u_w} et $f_{d_n u\tau}$ par des processus Gaussiens, avec les noyaux SE, Matern32 et Matern52

lumes finis en fonction du nombre d'itérations de couplage et des noyaux d'apprentissage. Un écart relatif de l'ordre de 10^{-4} est globalement obtenu pour l'ensemble des noyaux. Hormis une convergence plus lente pour le noyau Matern32, les écarts à la solution volumes finis se comportent de manières similaires et il est difficile d'isoler une hiérarchie ou un noyau préférentiel à utiliser pour le cas traité. De l'ordre d'une quinzaine d'itérations de couplage sont nécessaires pour obtenir une précision relative inférieure à $\varepsilon = 0.1\%$ pour Matern32 et de l'ordre de $\varepsilon = 0.01\%$ pour SE et Matern52. Pour réaliser cette quinzaine d'itérations,

le nombre de simulations dans les plots virtuels est de l'ordre d'une petite vingtaine.

Sur la figure 5.20 sont illustrées les premières étapes de construction des données d'appren-



(a) $i = 0$ (b) $i = 0$

(c) $i = 1$ (d) $i = 1$

FIGURE 5.20 – Données d'apprentissage et surfaces de réponse pour les deux premières itérations de couplage réalisées avec le processus Gaussien SE, $i = 0$ et $i = 1$. Figs. 5.20(a) et 5.20(c) : Les surfaces rouge et bleue correspondent à la solution exacte $d_n u_\tau$ et celle apprise $f_{d_n u_K}$. Les points bleus représentent les données de la base d'apprentissage. La surface colorée dans le plan (u_K, Kn_K) donne la distribution des incertitudes σ_N relatives à la fonction d'apprentissage et le cercle rouge localise la position où le maximum est atteint, $(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})$. Le cercle vert correspond à la donnée (u_K, Kn_K) où u_K résulte d'une simulation de Couette dans le plot virtuel avec une vitesse de la paroi supérieure $f_{u_w}(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})$ $(\mathrm{Kn}_K = \widetilde{\mathrm{Kn}_K})$. Figs. 5.20(b) et 5.20(d) : Le cercle de couleur verte est inséré dans la base d'apprentissage (point bleu) et une nouvelle surface de réponse est apprise (surface rouge), laquelle sera transmise à la méthode des volumes finis pour reconstruire la condition de glissement à la paroi.

tissage pour la fonction $f_{d_n u_\tau}(u_K, \mathrm{Kn}_K)$ dans le cas de la fonction noyau SE. Les définitions et notations suivantes sont introduites :

— la surface rouge est la solution exacte $d_n u_K(u_K, \mathrm{Kn}_K)$ (Eq. 5.16),

— la surface bleue est la solution apprise $f_{d_n u_\tau}(u_K, \mathrm{Kn}_K)$,

— les points bleus sont les données qui correspondent aux triplets $(u_K, \mathrm{Kn}_K, d_n u_K)$ de la base d'apprentissage B_{app},

— le champ de couleur dans le plan (u_K, Kn_K) donne l'incertitude σ_N prédite par le modèle (voir Eq. 4.28c, page 88) en utilisant les données de la base d'apprentissage (points bleus),

— le cercle rouge correspond au couple de données $(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})$ qui maximise l'incertitude σ_N ; l'ordonnée est la réponse exacte,

— le cercle vert désigne la position du nouveau point d'apprentissage issu de l'incertitude maximale (cercle rouge) et défini par (u_K, Kn_K), avec (Eqs. 5.14 et 5.15)

$$u_K = \frac{f_{u_w}^{(i)}(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})}{2}, \qquad \widetilde{\mathrm{Kn}_K} = \mathrm{Kn}_K \tag{5.18}$$

et $f_{u_w}^{(i)}$ la fonction d'apprentissage à l'itération i.

Les valeurs numériques pour construire la base potentielle d'apprentissage B_{pot} et la base d'apprentissage B_{app} sont reportées dans le tableau 5.2. Avec les notations présentées ci-

i	j	NS/App	$\widetilde{\mathrm{Kn}_K} = \mathrm{Kn}_K$	$\widetilde{u_K}$	$f_{u_w}(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})$	u_K	$d_n u_K$
$i=0$	1	$\mathrm{NS}^{(min,1)}$	0.15	10	$f_{u_w}^{(0)}$ =10	5.1	4.4
	2	$\mathrm{NS}^{(max,2)}$	0.61	49	$f_{u_w}^{(0)}$ =49	24	15
	3	$\mathrm{App}^{(3)}$	0.15	12	$f_{u_w\,[1;2]}^{(1)}$ =17	8.8	1.9
	4	$\mathrm{App}^{(4)}$	0.61	33	$f_{u_w\,[1-3]}^{(1)}$ =50	25	4.0
			– –	– –	$f_{u_w\,[1-4]}^{(1)}$	– –	$d_n u_\tau = f_{d_n u_\tau\,[1-4]}^{(1)}$
$i=1$	5	$\mathrm{NS}^{(min,5)}$	0.15	10	$f_{u_w\,[1-4]}^{(1)}$ =20	10	4.4
	6	$\mathrm{NS}^{(max,6)}$	0.61	70	$f_{u_w\,[1-4]}^{(1)}$ =140	70	43
	1	–	0.15	10	$f_{u_w}^{(0)}$ =10	5.1	4.4
	2	–	0.61	49	$f_{u_w}^{(0)}$ =49	24	15
	3	–	0.15	12	$f_{u_w\,[1;2]}^{(1)}$ =17	8.8	1.9
	4	–	0.61	33	$f_{u_w\,[1-3]}^{(1)}$ =50	25	4.0
	7	$\mathrm{App}^{(7)}$	0.15	36	$f_{u_w\,[2;4-6]}^{(2)}$ =60	30	21
			– –	– –	$f_{u_w\,[2;4-7]}^{(2)}$	– –	$d_n u_\tau = f_{d_n u_\tau\,[2;4-7]}^{(2)}$

TABLE 5.2 – Tableau de données constituant la base d'apprentissage potentielle B_{pot} aux itérations de couplage $i = 0$ et $i = 1$. Pour une valeur de i donnée, l'entier j numérote les données dans la base d'apprentissage potentielle B_{pot} ; seules les lignes écrites en noir appartiennent à la base d'apprentissage B_{app}, celles en gris en sont exclues. La notation NS souligne que les données sont issues de Navier-Stokes (modèle continu) et App indique l'apprentissage du quadruplet $(u_K, \mathrm{Kn}_K, u_w = f_{u_w\,[j_0;\cdots;j_1]}(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K}), d_n u_K)$ pour lequel l'incertitude du modèle/données sur la base d'apprentissage contenant les lignes $j_0, \cdots$ et j_1 est maximale pour $(u_K, \mathrm{Kn}_K) \in [\mathrm{NS}^{(min,k)}, \mathrm{NS}^{(max;k)}]$.

dessus, les deux premières itérations de la méthode de couplage hybride implicite sont maintenant décrites :

1. Figures 5.20(a), 5.20(b) et Tableau 5.2 pour $i = 0$.

(a) Tab. 5.2 : $\widetilde{u_K}$ et $\widetilde{\mathrm{Kn}_K}$ en $j=1$ et $j=2$.
D'une première simulation volumes finis, avec des conditions d'adhérence à la paroi, sont extraits deux couples : les valeurs minimales $\widetilde{u_K}^{(min,1)}$ et $\widetilde{\mathrm{Kn}_K}^{(min,1)}$, et maximales $\widetilde{u_K}^{(max,2)}$ et $\widetilde{\mathrm{Kn}_K}^{(max,2)}$ de la vitesse tangentielle u_τ et du nombre de Knudsen de maille Kn_K pour tous les points du maillage P_K.

(b) Tab. 5.2 : $f_{u_w}^{(0)}(\widetilde{u_K},\widetilde{\mathrm{Kn}_K})$, u_K, Kn_K et $d_n u_K$ en $j=1$ et $j=2$.
A l'aide de la fonction initialement définie par $f_{u_w}^{(0)}(\widetilde{u_K},\widetilde{\mathrm{Kn}_K})=\widetilde{u_K}$ et des relations (5.18) puis (5.16), deux triplets de données sont calculés : $(u_K^{(min,1)},\mathrm{Kn}_K{}^{(min,1)},(d_n u_K)^{(min,1)})$ et $(u_K^{(max,2)},\mathrm{Kn}_K{}^{(max,2)},(d_n u_K)^{(max,2)})$. Ces deux triplets sont représentés par deux points bleus situés aux extrémités du graphe 5.20(a).

(c) Tab. 5.2 : apprentissage de $f_{u_w\,[1;2]}^{(1)}$ en $j=3$ avec les données des lignes $j=1$ et $j=2$.
La fonction $f_{u_w}^{(1)}$ est apprise avec les triplets de données :

$$\begin{pmatrix} u_K^{(min,1)}, \\ \mathrm{Kn}_K{}^{(min,1)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(min,1)},\widetilde{\mathrm{Kn}_K}^{(min,1)}) \end{pmatrix}^T, \quad \begin{pmatrix} u_K^{(max,2)}, \\ \mathrm{Kn}_K{}^{(max,2)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(max,2)},\widetilde{\mathrm{Kn}_K}^{(max,2)}) \end{pmatrix}^T$$

(d) Tab. 5.2 : $\widetilde{u_K}$, $\widetilde{\mathrm{Kn}_K}$, $f_{u_w\,[1;2]}^{(1)}(\widetilde{u_K},\widetilde{\mathrm{Kn}_K})$, u_K, Kn_K et $d_n u_K$ en j=3.
Un troisième point $(\widetilde{u_K}^{(3)},\widetilde{\mathrm{Kn}_K}^{(3)})$ est ajouté à la position où l'incertitude $\sigma_{f_{d_n u_\tau}}(u_K,\mathrm{Kn}_K)$ est maximale sur l'intervalle $(u_K,\mathrm{Kn}_K)\in\mathcal{I}$ avec $\mathcal{I}=[u_K^{(min,1)},u_K^{(max,2)}]\times[\mathrm{Kn}_K{}^{(min,1)},\mathrm{Kn}_K{}^{(max,2)}]$, pour les données $(u_K^{(min,1)},\mathrm{Kn}_K{}^{(min,1)})$ et $(u_K^{(max,2)},\mathrm{Kn}_K{}^{(max,2)})$. Les valeurs $(u_K^{(3)},\mathrm{Kn}_K^{(3)},(d_n u_K)^{(3)})$ utilisées dans la base d'apprentissage sont issues de la relation (5.18) puis (5.16) avec $f_{u_w\,[1;2]}^{(1)}$, ce qui engendre le troisième point bleu du graphe 5.20(a).

(e) Tab. 5.2 : apprentissage de $f_{u_w\,[1-3]}^{(1)}$ en $j=4$ avec les données des lignes $j=1$ à $j=3$.
La fonction $f_{u_w}^{(1)}$ est mise à jour avec les triplets de données :

$$\begin{pmatrix} u_K^{(min,1)}, \\ \mathrm{Kn}_K{}^{(min,1)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(min,1)},\widetilde{\mathrm{Kn}_K}^{(min,1)}) \end{pmatrix}^T, \begin{pmatrix} u_K^{(max,2)}, \\ \mathrm{Kn}_K{}^{(max,2)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(max,2)},\widetilde{\mathrm{Kn}_K}^{(max,2)}) \end{pmatrix}^T, \begin{pmatrix} u_K^{(3)}, \\ \mathrm{Kn}_K{}^{(3)}, \\ f_{u_w\,[1;2]}^{(1)}(\widetilde{u_K}^{(3)},\widetilde{\mathrm{Kn}_K}^{(3)}) \end{pmatrix}^T$$

(f) Tab. 5.2 : $\widetilde{u_K}$, $\widetilde{\mathrm{Kn}_K}$, $f_{u_w\,[1-3]}^{(1)}(\widetilde{u_K},\widetilde{\mathrm{Kn}_K})$, u_K, Kn_K et $d_n u_K$ en j=4.
Un quatrième point $(\widetilde{u_K}^{(4)},\widetilde{\mathrm{Kn}_K}^{(4)})$ est ajouté à la position où l'incertitude $\sigma_{f_{d_n u_\tau}}$ est maximale sur l'intervalle $(u_K,\mathrm{Kn}_K)\in\mathcal{I}$, pour les données $(u_K^{(min,1)},\mathrm{Kn}_K{}^{(min,1)})$, $(u_K^{(max,2)},\mathrm{Kn}_K{}^{(max,2)})$ et $(u_K^{(3)},\mathrm{Kn}_K^{(3)})$. La fonction $\sigma_{f_{d_n u_\tau}}(u_K,\mathrm{Kn}_K)$ est représentée par des aplats de couleurs et le lieu où la fonction atteint son maximum par un cercle rouge (l'ordonnée est placée sur la vraie surface de couleur rouge, Fig. 5.20(a)). Les valeurs $(u_K^{(4)},\mathrm{Kn}_K^{(4)},(d_n u_K)^{(4)})$ utilisées dans la base d'apprentissage sont issues de la relation (5.18) puis (5.16) avec $f_{u_w\,[1-3]}^{(1)}$, ce qui engendre le cercle noir sur le graphe 5.20(a).

(g) Tab. 5.2 : apprentissage de $f_{d_n u_\tau\,[1-4]}^{(1)}$ avec les données des lignes $j=1$ à $j=4$.

Avec les quatre triplets de la base d'apprentissage $(u_K^{(min,1)}, \mathrm{Kn}_K{}^{(min,1)}, (d_n u_K)^{(min,1)})$, $(u_K^{(max,2)}, \mathrm{Kn}_K{}^{(max,2)}, (d_n u_K)^{(max,2)})$, $(u_K^{(3)}, \mathrm{Kn}_K^{(3)}, (d_n u_K)^{(3)})$ et $(u_K^{(4)}, \mathrm{Kn}_K^{(4)}, (d_n u_K)^{(4)})$, la fonction $f^{(1)}_{d_n u_\tau}(u_K, \mathrm{Kn}_K)$ est apprise (surface bleue sur la figure 5.20(b) ; la fonction tracée en rouge est la solution exacte). La fonction $f^{(1)}_{d_n u_\tau}(u_K, \mathrm{Kn}_K)$ est ensuite transmise à la méthode des volumes finis pour construire la condition aux limites de glissement.

(h) Tab. 5.2 : apprentissage de $f^{(1)}_{u_w\,[1-4]}$ avec les données des lignes $j = 1$ à $j = 4$. La fonction $f^{(1)}_{u_w}$ est mise à jour par l'utilisation des triplets de données

$$\begin{pmatrix} u_K^{(min,1)}, \\ \mathrm{Kn}_K{}^{(min,1)}, \\ f^{(0)}_{u_w}(\widetilde{u_K}^{(min,1)}, \widetilde{\mathrm{Kn}_K}^{(min,1)}) \end{pmatrix}^T, \begin{pmatrix} u_K^{(max,2)}, \\ \mathrm{Kn}_K{}^{(max,2)}, \\ f^{(0)}_{u_w}(\widetilde{u_K}^{(max,2)}, \widetilde{\mathrm{Kn}_K}^{(max,2)}) \end{pmatrix}^T,$$

$$\begin{pmatrix} u_K^{(3)}, \\ \mathrm{Kn}_K{}^{(3)}, \\ f^{(1)}_{u_w\,[1;2]}(\widetilde{u_K}^{(3)}, \widetilde{\mathrm{Kn}_K}^{(3)}) \end{pmatrix}^T, \begin{pmatrix} u_K^{(4)}, \\ \mathrm{Kn}_K{}^{(4)}, \\ f^{(1)}_{u_w\,[1-3]}(\widetilde{u_K}^{(4)}, \widetilde{\mathrm{Kn}_K}^{(4)}) \end{pmatrix}^T$$

2. Figures 5.20(c), 5.20(d) et Tableau 5.2 pour $i = 1$.

(a) Tab. 5.2 : $\widetilde{u_K}$ et $\widetilde{\mathrm{Kn}_K}$ en $j = 5$ et $j = 6$.
D'une nouvelle simulation volumes finis, avec les conditions de glissement à la paroi (5.11) utilisant la fonction $f^{(1)}_{d_n u_\tau\,[1-4]}$, sont extraits deux couples : les valeurs minimales $\widetilde{u_K}^{(min,5)}$ et $\widetilde{\mathrm{Kn}_K}^{(min,5)}$, et maximales $\widetilde{u_K}^{(max,6)}$ et $\widetilde{\mathrm{Kn}_K}^{(max,6)}$ de la vitesse tangentielle u_τ et du nombre de Knudsen de maille Kn_K pour tous les points du maillage P_K.

(b) Tab. 5.2 : $f^{(1)}_{u_w\,[1-4]}(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})$, u_K, Kn_K et $d_n u_K$ en $j = 5$ et $j = 6$.
Les bornes minimales et maximales ayant changé, ces deux nouveaux points doivent être intégrés à la base potentielle.
A l'aide de la fonction $f^{(1)}_{u_w\,[1-4]}(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})$ et des relations (5.18) puis (5.16), deux triplets de données sont calculés : $(u_K^{(min,5)}, \mathrm{Kn}_K{}^{(min,5)}, (d_n u_K)^{(min,5)})$ et $(u_K^{(max,6)}, \mathrm{Kn}_K{}^{(max,6)}, (d_n u_K)^{(max,6)})$. Ces deux triplets sont représentés par deux points bleus situés aux extrémités du graphe 5.20(c).

(c) Tab. 5.2 : construction de la base d'apprentissage B_{app} à partir de la base potentielle B_{pot} constituée de $j = 1$ à $j = 6$.
Seuls les données de la base d'apprentissage dont les couples $(u_K^{(j)}, \mathrm{Kn}_K^{(j)}) \in \mathcal{I}$, avec $\mathcal{I} = [u_K^{(min,5)}, u_K^{(max,6)}] \times [\mathrm{Kn}_K^{(min,5)}, \mathrm{Kn}_K^{(max,6)}]$, sont copiées dans la base d'apprentissage ($j = 5$, 6, 2, 4). Les troisième et quatrième points bleus correspondent aux données des lignes $j = 2$ et $j = 4$, déjà apprises.

(d) Tab. 5.2 : détermination de $f^{(2)}_{u_w\,[2;4-6]}$ en $j = 7$ avec les données de la base d'apprentissage B_{app} ($j = 2$ et $j = 4$ à $j = 6$).

La fonction $f_{u_w}^{(2)}$ est apprise avec les triplets de données :

$$\begin{pmatrix} u_K^{(min,5)}, \\ \mathrm{Kn}_K{}^{(min,5)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(min,5)}, \widetilde{\mathrm{Kn}_K}^{(min,5)}) \end{pmatrix}^T, \begin{pmatrix} u_K^{(max,6)}, \\ \mathrm{Kn}_K{}^{(max,6)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(max,6)}, \widetilde{\mathrm{Kn}_K}^{(max,6)}) \end{pmatrix}^T$$

$$\begin{pmatrix} u_K^{(2)}, \\ \mathrm{Kn}_K{}^{(2)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(2)}, \widetilde{\mathrm{Kn}_K}^{(2)}) \end{pmatrix}^T, \begin{pmatrix} u_K^{(4)}, \\ \mathrm{Kn}_K{}^{(4)}, \\ f_{u_w\,[1-3]}^{(1)}(\widetilde{u_K}^{(4)}, \widetilde{\mathrm{Kn}_K}^{(4)}) \end{pmatrix}^T$$

(e) Tab. 5.2 : $\widetilde{u_K}$, $\widetilde{\mathrm{Kn}_K}$, $f_{u_w\,[2;4-6]}^{(2)}(\widetilde{u_K}, \widetilde{\mathrm{Kn}_K})$, u_K, Kn_K et $d_n u_K$ en j=7.
Un septième point $(\widetilde{u_K}^{(7)}, \widetilde{\mathrm{Kn}_K}^{(7)})$ est ajouté à la position où l'incertitude $\sigma_{f_{d_n u_\tau}}$ est maximale sur l'intervalle $(u_K, \mathrm{Kn}_K) \in \mathcal{I}$. La fonction $\sigma_{f_{d_n u_\tau}}(u_K, \mathrm{Kn}_K)$ est représentée par des aplats de couleurs et le lieu où la fonction atteint son maximum par un cercle rouge (l'ordonnée est placée sur la vraie surface de couleur rouge, Fig. 5.20(c)). Les valeurs $(u_K^{(7)}, \mathrm{Kn}_K^{(7)}, (d_n u_K)^{(7)})$ utilisées dans la base d'apprentissage sont issues de la relation (5.18) puis (5.16) avec $f_{u_w\,[2;4-4]}^{(2)}$, ce qui engendre le cercle noir sur le graphe 5.20(c).

(f) Tab. 5.2 : apprentissage de $f_{d_n u_\tau\,[2;4-7]}^{(2)}$ avec les données des lignes $j = 2$, $j = 5$ à 7.
Avec les quatre triplets de la base d'apprentissage $(u_K^{(min,5)}, \mathrm{Kn}_K{}^{(min,5)}, (d_n u_K)^{(min,5)})$, $(u_K^{(max,6)}, \mathrm{Kn}_K{}^{(max,6)}, (d_n u_K)^{(max,6)})$, $(u_K^{(2)}, \mathrm{Kn}_K^{(2)}, (d_n u_K)^{(2)})$, $(u_K^{(4)}, \mathrm{Kn}_K^{(4)}, (d_n u_K)^{(4)})$ et $(u_K^{(7)}, \mathrm{Kn}_K^{(7)}, (d_n u_K)^{(7)})$, la fonction $f_{d_n u_\tau}^{(2)}(u_K, \mathrm{Kn}_K)$ est apprise (surface bleue sur la figure 5.20(d) ; la fonction tracée en rouge est la solution exacte). La fonction $f_{d_n u_\tau}^{(2)}(u_K, \mathrm{Kn}_K)$ est ensuite transmise à la méthode des volumes finis pour construire la condition aux limites de glissement.

(g) Tab. 5.2 : apprentissage de $f_{u_w\,[2:4-7]}^{(2)}$ avec les données des lignes $j = 2$ et $j = 4$ à 7.
La fonction $f_{u_w}^{(2)}$ est mise à jour par l'utilisation des triplets de données

$$\begin{pmatrix} u_K^{(min,5)}, \\ \mathrm{Kn}_K{}^{(min,5)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(min,5)}, \widetilde{\mathrm{Kn}_K}^{(min,5)}) \end{pmatrix}^T, \begin{pmatrix} u_K^{(max,6)}, \\ \mathrm{Kn}_K{}^{(max,6)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(max,6)}, \widetilde{\mathrm{Kn}_K}^{(max,6)}) \end{pmatrix}^T, \begin{pmatrix} u_K^{(2)}, \\ \mathrm{Kn}_K{}^{(2)}, \\ f_{u_w}^{(0)}(\widetilde{u_K}^{(2)}, \widetilde{\mathrm{Kn}_K}^{(2)}) \end{pmatrix}^T,$$

$$\begin{pmatrix} u_K^{(4)}, \\ \mathrm{Kn}_K{}^{(4)}, \\ f_{u_w\,[1-3]}^{(1)}(\widetilde{u_K}^{(4)}, \widetilde{\mathrm{Kn}_K}^{(4)}) \end{pmatrix}^T \begin{pmatrix} u_K^{(7)}, \\ \mathrm{Kn}_K{}^{(7)}, \\ f_{u_w\,[2;4-6]}^{(2)}(\widetilde{u_K}^{(7)}, \widetilde{\mathrm{Kn}_K}^{(7)}) \end{pmatrix}^T$$

La figure 5.21 illustre pour les noyaux SE, Matern32 et Matern52 la dépendance entre $\varepsilon_\infty^{(c)}$, l'écart relatif de la solution hybride à la solution de référence volumes finis, en fonction de l'erreur d'apprentissage $\varepsilon_\infty^{(p)}$, écart relatif entre la solution apprise et la fonction exacte. Il apparaît que la solution hybride sera d'autant plus précise que la fonction apprise est bonne.

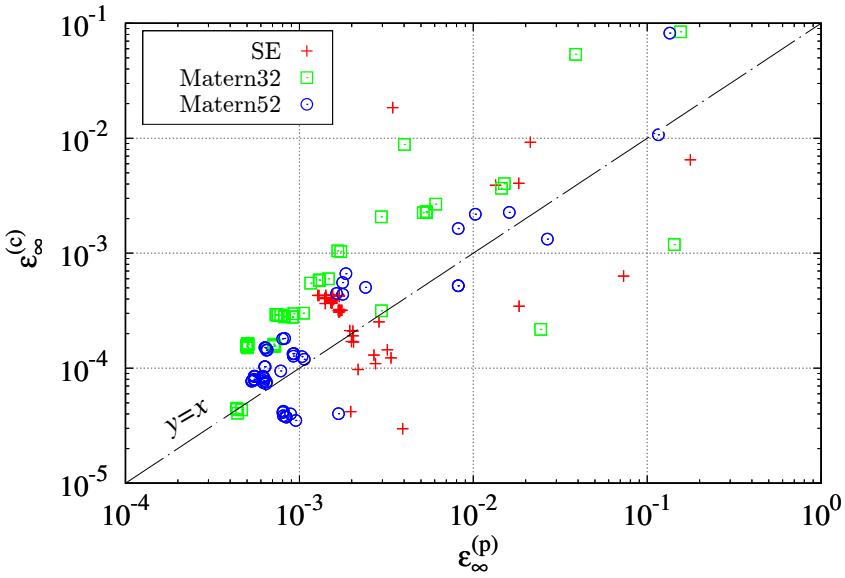


FIGURE 5.21 – Écart relatif de la solution hybride à la solution volumes finis, $\varepsilon_\infty^{(c)}$, en fonction de l'erreur d'apprentissage du cisaillement $\varepsilon_\infty^{(p)}$.

5.5.4.3 Validation de la solution hybride pour un canal de section variable

Pour illustrer le fonctionnement de la méthode hybride implicite sur des maillages triangulaires et des parois d'inclinaison variable, des simulations d'un écoulement isotherme, glissant et compressible ont été menées dans un canal de section variable, discrétisé par des volumes de contrôle triangulaires (Fig. 5.22). Les conditions aux limites sont identiques à celles utili-

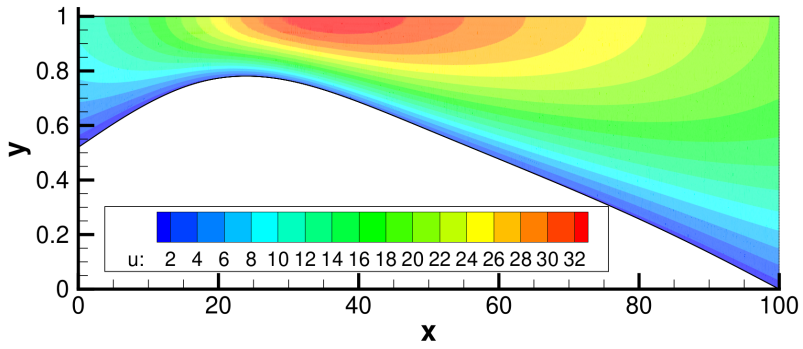


FIGURE 5.22 – Solution hybride obtenue dans un canal de section variable.

sées pour le canal rectiligne : $T_{in} = 300\,K$, $p_{in} = 2\times10^5\,Pa$ et $p_{out} = 5\times10^4\,Pa$. La longueur du canal au niveau de l'axe de symétrie vaut $100H$ avec $H = 3\,\mu m$. La demi-hauteur du canal à l'entrée est H et en sortie $2H$.

La solution hybride, obtenue avec le noyau SE, diffère de la solution de référence volumes finis d'environ $\varepsilon_\infty^{(c)} = 0.1\%$ au bout d'une quinzaine d'itérations de couplage (Fig. 5.23). Bien que cet écart relatif soit faible, il est environ dix fois plus grand que celui obtenu pour la simulation hybride dans le simple canal rectiligne (Fig. 5.19, page 144). Cette relative détérioration de la précision est probablement liée à l'utilisation du maillage triangulaire

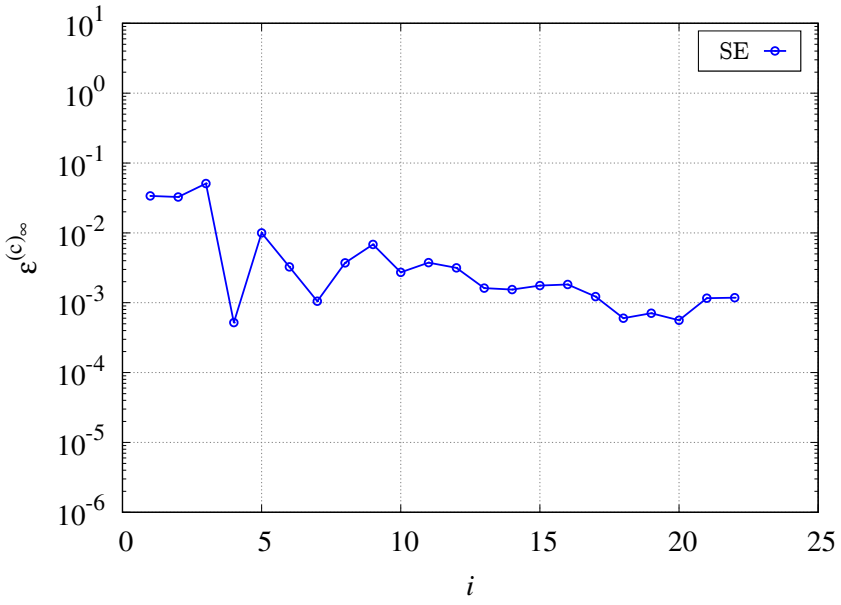


FIGURE 5.23 – Écart relatif de la solution hybride à la solution volumes finis pour l'écoulement dans un canal de section variable. Le noyau du processus Gaussien est SE.

qui doit vérifier la contrainte que le centre des triangles reste en dehors de la couche de Knudsen ($\mathrm{Kn}_K \leq 1$), ce qui impose de relativement grands triangles à proximité des parois, principalement dans la région de plus faible section du canal.

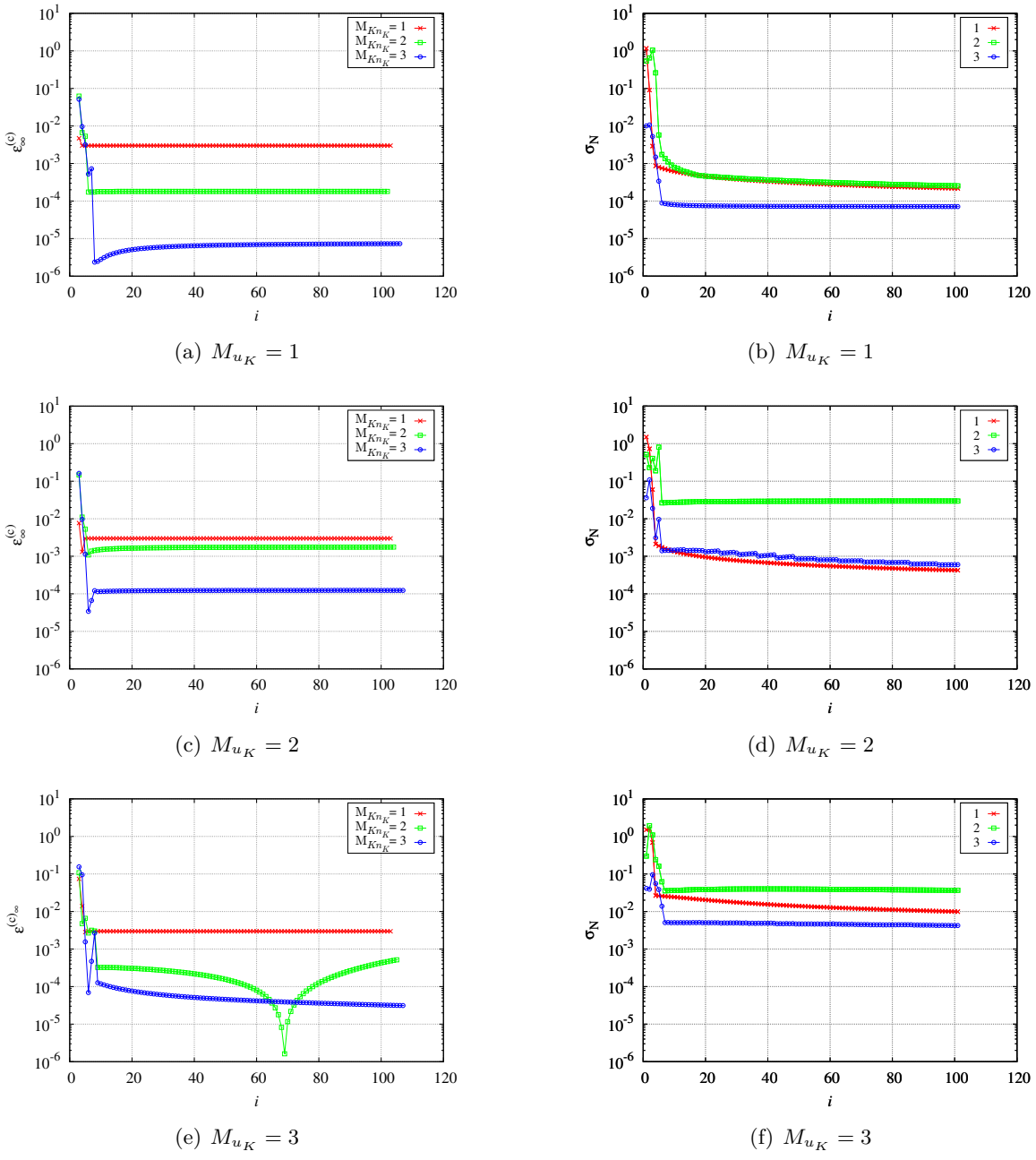
5.5.4.4 Apprentissage bayésien

Les bases d'apprentissage pour les fonctions f_{u_w} et $f_{d_n u_\tau}$ sont choisies identiques, sous la forme

$$\{1, ((u_K)^m(\mathrm{Kn}_K)^n, (m,n) \in [\![1, M_{u_K}]\!] \times [\![1, M_{\mathrm{Kn}_K}]\!])\}$$

où M_{u_K} M_{Kn_K} sont deux paramètres entiers dont l'influence sur la qualité de l'approximation et la vitesse de convergence de la solution hybride est étudiée. L'ajout de points dans la base d'apprentissage est uniquement basé sur l'incertitude de prédiction sur le cisaillement $f_{d_n u_\tau}$, et aucunement sur celle portant sur f_{u_w}.

La figure 5.24 présente la variation de l'erreur relative de la solution hybride $\varepsilon_{\infty}^{(c)}$ et l'incertitude σ_N sur la prédiction du cisaillement appris $f_{d_n u_\tau}$, en fonction du nombre d'itérations de couplage, pour M_{u_K} et M_{Kn_K} compris entre 1 et 3.
D'après l'expression analytique du cisaillement (voir Eq. 5.16, page 143), il est évident qu'une base particulièrement pertinente pour apprendre $f_{d_n u_\tau}$ est de choisir $M_{u_K} = 1$ (relation linéaire de du/dz en fonction de u_K). Pour le nombre de Knudsen de maille ($\mathrm{Kn}_K < 1$), la solution exacte peut être développée en une série polynomiale en Kn_K : ainsi, plus la valeur de M_{Kn_K} sera grande, meilleure sera l'approximation $f_{d_n u_\tau}$ (Fig. 5.24(a)).
L'augmentation de la base polynomiale pour u_K, $M_{u_K} = 2$ (Fig. 5.24(c)) et $M_{u_K} = 3$ (Fig. 5.24(e)), semble dégrader l'approximation du fait de l'introduction de degrés de liberté inutiles (base sur-dimensionnée). Cependant, malgré une augmentation de l'écart relatif,



(a) $M_{u_K} = 1$ (b) $M_{u_K} = 1$

(c) $M_{u_K} = 2$ (d) $M_{u_K} = 2$

(e) $M_{u_K} = 3$ (f) $M_{u_K} = 3$

FIGURE 5.24 – Écarts relatifs de la solution hybride à la solution volumes finis, $\varepsilon_\infty^{(c)}$ (gauche), et incertitude σ_N sur la prédiction du cisaillement appris $f_{d_n u_\tau}$ (droite) en fonction du nombre d'itérations et du nombre M_{u_K} et M_{Kn_K} de polynômes dans la base d'apprentissage.

celui-ci reste faible, au plus de l'ordre de quelque pour mille. Globalement, pour une valeur M_{u_K} donnée, la valeur $\varepsilon_\infty^{(c)}$ décroît avec l'augmentation du degré polynomial en Knudsen de maille.

Les figures 5.24(b), (d) et (f) montrent la variation de l'incertitude σ_N sur la prédiction du cisaillement appris $f_{d_n u_\tau}$ au cours des itérations de couplage. Aucune hiérarchie claire n'est observée quant aux évolutions de σ_N pour différents M_{Kn_K}. En revanche comme pour $\varepsilon_\infty^{(c)}$, l'augmentation de l'ordre polynomial en u_K semble dégrader l'incertitude sur la prédiction.

On peut également noter que l'évolution de σ_N se fait majoritairement au cours des 10 premières itérations de couplage (variation sur trois ordres de grandeur) et que le gain est marginal au cours des itérations suivantes. Cette évolution est assez bien corrélée à l'évolution de l'erreur à la solution exacte dont la variation se fait également principalement sur les premières itérations du couplage.
En pratique, le couplage itératif pour les méthodes bayésiennes pourra dans la suite être limité à une vingtaine d'itérations.

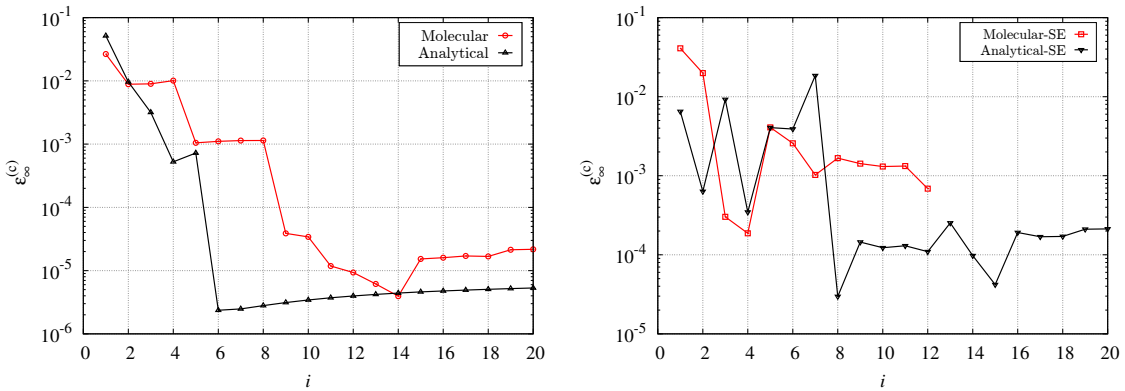
5.5.5 Apprentissage de la fonction $f_{d_n u_\tau}$ dans une boite de dynamique moléculaire

Dans cette section, les solutions analytiques sont remplacées par des simulations dans des boites de dynamique moléculaire et les parois seront représentées soit comme des murs stochastiques diffus, soit par un modèle atomique.

Nous considérons toujours le même canal plan de hauteur $2H = 6\,\mu m$ selon $\bar{e}_z$ et de longueur $L = 100H$ selon $\bar{e}_x$. L'écoulement de Poiseuille, compressible, glissant et isotherme à $T_{in} = 300\,K$ est engendré par une différence de pression entre l'entrée $p_{in} = 2 \times 10^5\,Pa$ et la sortie $p_{out} = 5 \times 10^4\,Pa$. Des conditions de symétrie sont imposées en $z = H$ (voir Sec. 5.4.4.2, page 127). Le maillage est régulier et est constitué de 3000×10 volumes de contrôle rectangulaires dans les directions $\bar{e}_x$ et $\bar{e}_z$. Le processus itératif est initialisé avec une condition aux limites d'adhérence à la paroi.

5.5.5.1 Murs stochastiques diffus

La figure 5.25 montre les convergences de la solution hybride obtenue, d'une part avec une



(a) bayésien (b) processus gaussien

FIGURE 5.25 – Écarts relatifs entre deux types de solutions hybrides. L'un est obtenu avec une boite moléculaire munie de parois stochastiques diffuses et le second s'appuie sur une solution analytique glissante. Gauche : modèle bayésien pour $M_{\mathrm{Kn}_K} = 3$ et $M_{u_K} = 1$. Droite : modèle gaussien pour le noyau SE.

boite moléculaire et des murs stochastiques diffus et d'autre part avec le modèle analytique (voir Sec. 5.5.4, page 142) pour un processus gaussien de noyau SE et une méthode bayésienne

de base $\{1, (u_K(\mathrm{Kn}_K)^n,\ n \in [\![1,3]\!])\}$. Quelle que soit la méthode d'apprentissage, la solution hybride construite avec les boites moléculaires se révèle être très proche de la solution de référence en volumes finis avec des conditions de Navier de type diffus à la paroi : pour le modèle bayésien (resp. gaussien), $\varepsilon_\infty^{(c)} \approx 10^{-5}$ (resp. $\varepsilon_\infty^{(c)} \approx 10^{-3}$). Même si les écarts relatifs sont faibles, ils restent néanmoins plus grands que ceux relevés lorsque la boite de dynamique moléculaire est substituée par la solution analytique.

5.5.5.2 Murs atomiques

Les parois sont maintenant modélisées par des atomes de platine organisés selon un réseau cubique face centrée (voir Sec. 3.3, page 58 pour les potentiels d'interaction).
Deux types de murs sont étudiés. Le premier est un mur lisse tel que présenté sur le dessin de la Fig. 5.18, page 138. Le second possède des rugosités d'épaisseur constante (10% la hauteur fluide $H_{M,f}$) et réparties périodiquement dans le sens de l'écoulement avec une longueur d'onde égale à la longueur de la boite $l = 5\,nm$ (Fig. 5.26).

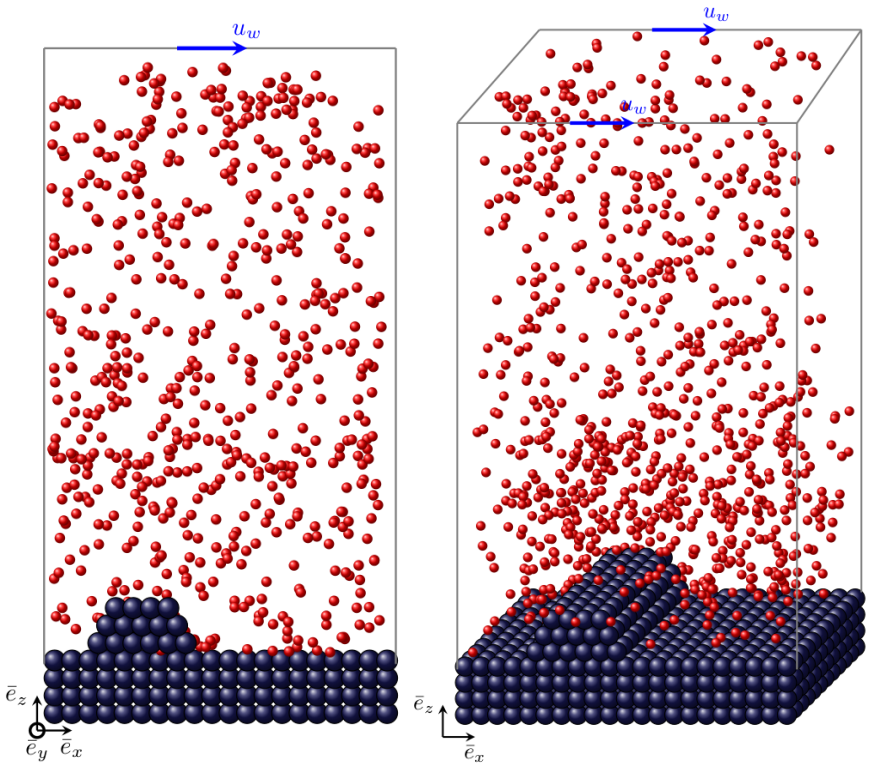


FIGURE 5.26 – Boite moléculaire avec des murs atomiques rugueux.

La figure 5.27 montre les écarts relatifs de la solution hybride à celle volumes finis utilisant un modèle de Navier pour modéliser des parois diffuses. Pour le mur simple, sans rugosité, l'écart $\varepsilon_\infty^{(c)}$ atteint 10^{-4} au bout de vingt itérations de couplage. Bien que cet écart soit faible, il est quasiment un ordre de grandeur plus grand que celui obtenu pour une solution avec un mur stochastique diffus. L'utilisation du mur atomique semble donc légèrement modifier la nature des interactions entre le mur et le gaz.

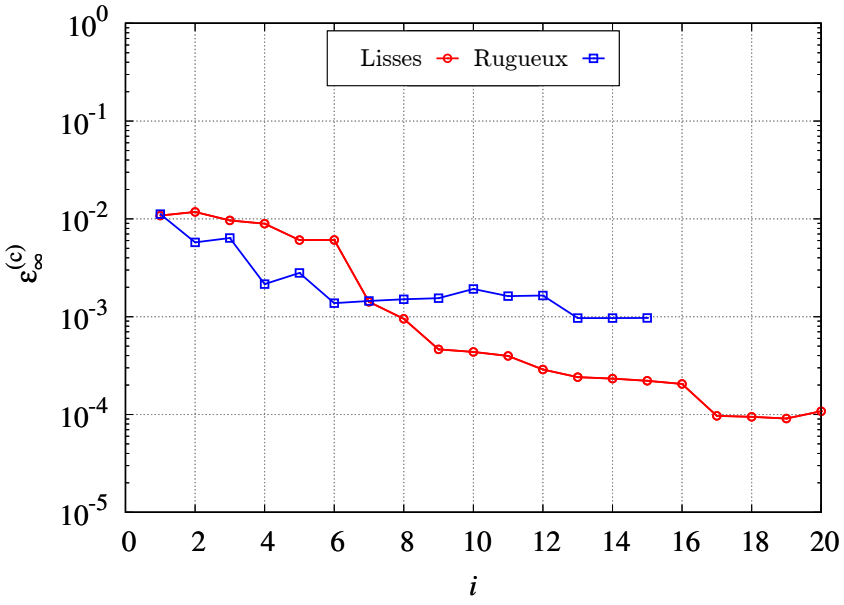


FIGURE 5.27 – Écarts relatifs de la solution hybride à celle volumes finis avec des conditions de glissement. La comparaison se fait en utilisant les régressions bayésiennes polynomiales $M_{\mathrm{Kn}_K} = 3$ et $M_{u_K} = 1$ avec des murs atomiques lisses et rugueux.

L'introduction de rugosités sur la paroi semble montrer que la solution hybride s'éloigne encore davantage de celle donnée par le modèle de Navier diffus (solution de référence).

5.6 Conclusion

Dans ce chapitre, nous avons présenté deux stratégies de couplage faible par la méthode des plots virtuels pour résoudre un écoulement isotherme, compressible et glissant circulant dans un canal plan sous une différence de pression.

La méthode la plus simple est basée sur un couplage explicite et une transmission d'informations par valeurs entre les grandes et petites échelles. Cette méthode, testée en s'affranchissant des erreurs provenant de l'apprentissage et de la résolution de la petite échelle par la dynamique moléculaire, s'avère très coûteuse en nombre d'itérations de couplage lorsque les zones de communication s'approchent l'une de l'autre, ce qui dégrade également la précision des solutions.

Pour palier ces défauts, une méthode implicite a été développée fonctionnant non pas par transmission de valeurs, mais par envoi d'une fonction apprise. La connaissance de cette fonction permet de reconstruire une vitesse de glissement dont la valeur est implicitement couplée avec les inconnues du modèle continu. Ce modèle a d'abord été étudié en remplaçant la simulation dans la boite moléculaire par une solution analytique, ce qui a permis d'examiner la précision et le nombre d'itérations requis pour obtenir une solution satisfaisante (une vingtaine) pour des méthodes d'apprentissage gaussiennes et bayésiennes. Par la suite, les simulations des petites échelles ont été introduites par l'utilisation de boites de simulation moléculaire avec différents modèles de parois : des murs stochastiques diffus, atomiques lisses et atomiques rugueux. Les résultats semblent indiquer que les solutions hybrides s'éloignent

de plus en plus de celle obtenue avec un modèle de Navier glissant et diffus ($\sigma_u = 1$) lorsque les modèles de murs se complexifient.

If I have seen further it is by standing on the shoulders of Giants.

Isaac Newton

Chapitre 6

Conclusion générale et perspectives

Ce travail a principalement porté sur le développement d'une méthode hybride efficace pour résoudre des écoulements de gaz dans les micro-canaux à grandes extensions. En effet, dans des conduites de fort rapport surface/volume, les interactions entre le fluide et la paroi ont un impact significatif sur la dynamique des fluides et les transferts thermiques. Les nombreux modèles macroscopiques existants dans la littérature ne permettent pas nécessairement de prendre pleinement en compte la diversité des phénomènes qui peuvent avoir lieu à l'échelle atomique (adsorption, condensation, modification des propriétés et caractéristiques de la paroi par des couches très minces, donc des films, etc) à proximité de l'interface.

Dans ce cadre, ce mémoire propose une méthode hybride multi-échelle permettant de simuler les écoulements de fluides dans de longues conduites micro ou nanométriques en gardant une représentation fine, à l'échelle moléculaire, du comportement du fluide à proximité de la paroi. Ces micro-simulations sont connectées à la région fluide principale modélisée par une approche adaptée aux grandes échelles. Ainsi, en pratique, les équations continues de Navier-Stokes et de l'énergie sont couplées à des simulations par dynamique moléculaire, permettant de prendre en compte de manière précise les interactions moléculaires au niveau de l'interface fluide/paroi.
Parmi les originalités du travail, figure l'utilisation de méthodes d'apprentissage supervisé, par régressions bayésienne ou gaussienne, afin de coupler via des fonctions continues les deux approches micro et macroscopique. L'utilisation de fonctions de couplage issues de l'apprentissage permet de réduire considérablement les coûts de calcul, tout en garantissant une convergence satisfaisante entre la méthode hybride et les solutions de référence.

Basées sur des "plots virtuels", c'est-à-dire sur des micro simulations non nécessairement couplées spatialement au domaine continue, deux déclinaisons du couplage, explicite et implicite, ont été imaginées.
D'un point de vue qualitatif, les résultats obtenus sont intéressants dans la mesure où la convergence entre la méthode hybride et multi-échelle (couplage) et la solution considérée comme référence obtenue à partir d'une résolution des équations de Navier-Stokes avec des modèles de glissement [64, 71] est observée.
Quantitativement, nous avons également de bons résultats, car nous observons une convergence d'ordre L_∞ de l'erreur entre la méthode hybride et multi-échelle (couplage) et la solution de référence avec le modèle de glissement. Nous avons apporté une attention toute

particulière à démontrer que cette nouvelle stratégie permet de limiter le coût numérique tout en conservant une erreur numérique acceptable. Pour ce faire, nous avons établi un indicateur d'erreur se basant sur les incertitudes et erreurs sur la prédiction $\sigma_N, \varepsilon^{(p)}$, tout en contrôlant également les erreurs de couplage $\varepsilon^{(c)}$.

Les tests réalisés ont démontré la capacité de la méthode hybride à reproduire avec précision les résultats obtenus par des modèles de glissement dynamique, tout en fournissant des informations supplémentaires sur les phénomènes microscopiques se produisant près de l'interface fluide-paroi. De plus, l'évaluation de la méthode a révélé l'importance de maintenir une distance optimale entre les régions de connexion dans la couche de Knudsen, soulignant ainsi l'impact crucial de la géométrie et de la configuration des micro-canaux sur la qualité des solutions.

Le principal apport de ce travail réside dans l'adoption d'une approche implicite avec des plots virtuels, permettant ainsi à la méthode hybride d'offrir une solution prometteuse pour surmonter les limitations de l'approche explicite, en améliorant à la fois la précision des résultats et l'efficacité des calculs. En résumé, cette méthode de couplage hybride et multi-échelle constitue une avancée significative dans la modélisation des écoulements dans les micro-canaux, ouvrant de nouvelles perspectives pour la conception et l'optimisation de dispositifs et systèmes microfluidiques, ainsi que pour la compréhension approfondie des phénomènes de transport à l'échelle atomique.

Bien que le principal objectif ait été atteint, il convient toutefois d'examiner certains aspects de la méthode proposée afin d'y apporter à terme certaines améliorations. Pour l'approche implicite retenue, les principales perspectives sont listées ici

- Un premier point réside dans la prise en compte de parois et/ou de phénomène proche paroi plus complexes. En effet, dans ce mémoire, la méthode a été validée vis-à-vis de cas pour lesquels les conditions de Maxwell s'appliquent. La prise en compte de parois plus complexes, par exemple rugueuses, pour lesquelles le coefficient d'accommodation peut prendre une forme tensorielle, pourrait permettre de montrer un intérêt plus probant de la mise en place de notre méthode de couplage.
- En l'état, l'approche développée permet un couplage entre des micro simulations, ou plots virtuels, et équations de Navier-Stokes discrétisés en volumes finis uniquement sur les volumes de contrôle dont les faces coïncident avec les parois de la conduite. Ce point représente une limitation puisque la distance caractéristique sur laquelle s'effectue le couplage doit être de l'ordre du nombre de Knudsen local. Ce paramètre physique contraint donc la taille des mailles en proches parois, qui augmente lorsque le gaz se raréfie le long de la conduite. Un travail est à mener afin de permettre la communication entre domaines moléculaire et continu sur des distances plus importantes. Dans ce cas, les maille comprises entre les points de communications ne représentent pas une physique à petite échelle, mais une solution de raccord. Elles servent à améliorer le couplage par le calcul de variations (gradient) plus précises.
- Comme annoncé dans l'introduction, l'approche proposée repose sur une modélisation

plane des parois. Des parois courbes à grande échelle comme le convergent-divergent peuvent être traités, mais une attention particulière devra être prise lorsque la courbure est de l'ordre du libre parcours moyen.

- Enfin, un dernier point d'amélioration serait le passage en 3D au niveau continue. Néanmoins, ce point n'est pas prioritaire, puisqu'il ne présenterait pas de difficulté particulière autre que le passage au 3D du code Navier-Stokes.

Annexe A

Méthode de dynamique moléculaire

A.1 Propriétés physiques

Il est possible de calculer les propriétés physiques, à savoir, la conductivité thermique, la viscosité, etc, en utilisant la méthode de dynamique moléculaire par `LAMMPS` par exemple. Nombreuses sont les techniques utilisées dans ce logiciel. Dans cette section, certaines techniques et les résultats obtenus à partir des simulations de dynamique moléculaire sont exposés.

A.1.1 Conductivité thermique

Le terme conductivité thermique, κ, désigne la mesure de la capacité d'un matériau à transmettre l'énergie thermique de manière diffuse, conformément à la loi de Fourier,

$$\vec{J} = -\kappa \nabla T \tag{A.1}$$

où J est le flux thermique en unités d'énergie par surface par temps et ∇T est le gradient spatial de température. La conductivité thermique est souvent approximée comme une quantité isotrope, c'est-à-dire une quantité salaire, et a des unités d'énergie par distance, par temps, par température en K.

De manière générale, pendant la simulation, il faut mesurer la chaleur qui est transférée à travers le système, en utilisant la commande `fix ave/heat` par exemple, qui permet de mesurer la chaleur qui est transférée à travers un bord spécifique du système à chaque pas de temps de la simulation. La conductivité thermique est obtenue à partir de la formule (Eq. A.2).

Il existe plusieurs méthodes implémentées dans `LAMMPS` qui permettent d'obtenir la conductivité thermique. Dans l'ensemble, les méthodes consistent à configurer deux régions aux extrémités opposées d'une boîte de simulation, ou une au milieu et une à l'extrémité d'une boîte périodique, et les maintenir à des températures différentes, soit (i) à l'aide des thermostats (*lg*), soit (ii) en utilisant des commandes `fix heat` (*ht*)... L'énergie fournie à la région chaude est équivalente à l'énergie perdue dans la région froide et est proportionnelle au flux de chaleur entre les régions (voir les articles d'Ikeshoji et Hafskjold [46] ainsi que celui de Wirnsberger et al. [100]). Il est à souligner que, dans `LAMMPS`, les correctifs de

thermostatisation tels que `fix nvt`, `fix langevin`, et `fix temp/rescale` stockent l'énergie cumulative qu'ils apportent/soustraient.

De plus, deux autres méthodes peuvent être également citées : (i) la méthode impliquant l'algorithme de Muller-Plathe (*mp*) [69] hors équilibre, dans laquelle l'énergie cinétique est échangée entre les atomes dans deux couches différentes de la boîte de simulation, ce qui induit un gradient de température, voir (Fig. A.1) ; et (ii) la méthode basée sur la formule Green-Kubo (*gk*) [34]. Dans cette dernière, le flux de chaleur est obtenu à partir des fluctuations des énergies potentielle et cinétique par atome et du tenseur des contraintes par atome dans une simulation à l'équilibre en régime permanent.

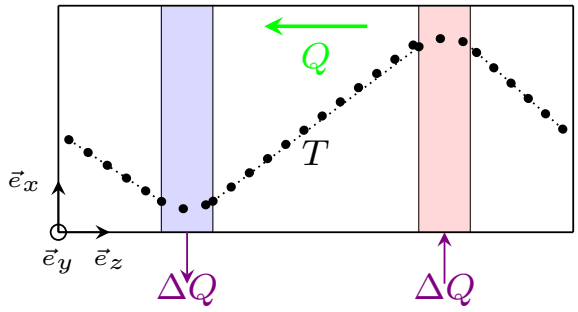


FIGURE A.1 – Principe de calcul de conductivité thermique dans `LAMMPS` par la méthode de MÜLLER-PLATHE [69],(*mp*).

Dans la suite de cet article, nous allons comparer les valeurs de conductivité trouvées par `LAMMPS` aux valeurs dans la littérature EVANS [26] et MEIER et al. [66]. Les propriétés du fluide, $\rho = 0.6$, $T = 1.35$, et $r_c = 2.5\sigma$, sont choisies de sorte qu'elles soient en accord avec l'article de EVANS [26], dans lequel il a calculé la conductivité thermique d'un petit système de 108 atomes en utilisant une méthode à l'aide du thermostat. Le système simulé comprend 4000 atomes, et est une boîte de dimensions $\ell_x = \ell_y = \frac{1}{2}\ell_z$, la direction la plus longue est celle de non-équilibre. Pour obtenir la conductivité thermique, κ, et l'extraire à partir de la sortie du fichier journal (log lammps) pour les méthodes hors équilibre NEMD, la formule (Eq. A.2) est utilisée, où ΔQ, est le flux d'énergie, et $\frac{\Delta T}{\Delta z}$ le gradient de température. Les résultats des différentes simulations sont présentés dans le tableau (Tab. A.1).

$$\kappa = \frac{\Delta Q}{2A\Delta t}\frac{\Delta z}{\Delta T} \tag{A.2}$$

ρ	T	κ_{ref}	κ_{mp}	κ_{gk}	κ_{lg}	κ_{ht}	ε_∞ %
0.6	1.35	3.360	3.426	3.438	3.415	3.435	1.96 %
0.5	1.35	2.513	2.608	2.499	–	–	1.40 %
0.4	1.35	2.001	1.970	2.008	–	–	1.54 %
0.3	2.50	0.409	–	0.441	–	–	1.54 %

TABLE A.1 – Valeurs de conductivité κ, pour différentes densités ρ et températures T. Les simulations ont été effectuées sur 2×10^6 d'itérations.

A.1.2 Viscosité dynamique

La viscosité dynamique d'un fluide μ peut être obtenue en utilisant le logiciel de simulation de dynamique moléculaire `LAMMPS` (Large-scale Atomic/Molecular Massively Parallel Simulator). Il existe de nombreuses techniques utilisées dans ce logiciel permettant de mesurer la viscosité dynamique d'un fluide. Dans cette section, certaines techniques sont brièvement présentées et les résultats obtenus à partir des simulations de dynamique moléculaire sont exposés.

La viscosité dynamique d'un fluide, μ, désigne une mesure de sa résistance au mouvement relatif de couches de fluide adjacentes, au cisaillement. Une mesure de la propension d'un fluide à transmettre la quantité de mouvement dans une direction perpendiculaire à la direction du mouvement. Elle est obtenue à partir de,

$$J = -\mu \nabla u_x \tag{A.3}$$

avec J le flux de quantité de mouvement, exprimé en unités de quantité de mouvement par surface et par temps, et ∇u_x désigne le gradient spatial de la vitesse du fluide se mouvant perpendiculairement à la direction du mouvement.

Afin d'obtenir la viscosité dynamique, plusieurs méthodes peuvent être citées : (i) utiliser plusieurs couches mobiles afin de cisailler le fluide entre elles, sans oublier de thermaliser les parois via des thermostats, (ii) utiliser des commandes `fix deform` pour cisailler le fluide et `fix nvt/sllod` pour thermostater le fluide via les équations de mouvement SLLOD [95].

En outre, deux autres méthodes sont utilisées pour comparer les résultats : l'une consiste à effectuer une simulation hors équilibre rNEMD inverse en utilisant la commande `fix viscosity` et l'autre simulation entièrement équilibrée ; (i) la méthode impliquant l'algorithme de Muller-Plathe (*mp*) [69] hors équilibre, dans laquelle l'énergie cinétique dans une dimension est échangée entre les atomes dans deux couches différentes de la boîte de simulation dans une dimension différente, ce qui induit un gradient de vitesse, voir (Fig. A.2) ; et (ii) la méthode basée sur la formule GREEN-KUBO (*gk*) [34], qui relie la moyenne d'ensemble de l'auto corrélation du tenseur de contrainte/pression à la viscosité.

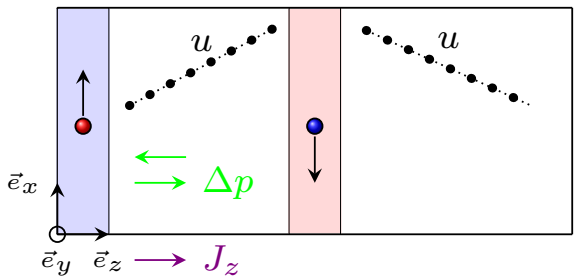


FIGURE A.2 – Principe de calcul de viscosité dynamique du fluide dans `LAMMPS` par la méthode de MÜLLER-PLATHE [69],(*mp*).

Dans la suite de cette section, les valeurs de viscosité trouvées par `LAMMPS` sont comparées aux valeurs dans la littérature : Meier et al. [66] (μ_{Mei}), et Galliéro et al. [34] (μ_{Gal}).

Comparaison des méthodes pour différentes propriétés

Le tableau (Tab. A.2) présente les valeurs de viscosité obtenues par les diverses méthodes et pour plusieurs densités ρ et températures T. Ce tableau montre que les simulations ont permis d'obtenir des résultats qui se rapprochent aux données de référence, avec une erreur d'environ 10% au maximum.

ρ	T	μ_{Mei}	μ_{Gal}	μ_{mp}	ε_{Mei} %
0.5	2.0	0.625	0.622	0.683	9.28 %
0.6	1.5	0.841	0.835	0.901	7.13 %
0.4	1.5	0.438	0.414	0.484	10.5 %
0.3	1.5	0.327	0.321	0.341	4.28 %
0.3	2.5	0.409	0.402	0.441	7.8 %
0.35	2.5	0.472	0.466	0.497	5.29 %
0.5	2.5	0.648	0.671	0.757	4.28 %

Table A.2 – Valeurs de la viscosité μ, pour différentes densités ρ et températures T. Les simulations ont été effectuées sur 5×10^7 d'itérations.

Comparaison de temps de calcul par méthode

Les propriétés du fluide, $\rho = 0.6$, $T = 1.5$, et $r_c = 2.5\sigma$, sont choisies de sorte qu'elles soient en accord avec l'article de Meier et al. [66]. Le tableau (Tab. A.3) représente les résultats des simulations effectuées en variant le nombre de processus sur lequel les calculs sont lancés. Ce tableau montre qu'à partir de 8 processus, le temps τ_{mp} requis pour obtenir la valeur de la viscosité dynamique est optimal, c'est-à-dire lorsque les échanges sont plus présents, la méthode de Müller-Plathe [69] présente des résultats plus satisfaisants et se rapproche plus aux valeurs de la référence.

Comparaison pour différents flux d'échange

Pour les mêmes propriétés du fluide, $\rho = 0.6$, $T = 1.5$, et $r_c = 2.5\sigma$, dans cette partie, les valeurs de la viscosité obtenues sont comparées pour différents débits d'échanges. Les résultats présentés dans le tableau (Tab. A.4) sont réalisés sur un nombre fixe de couches à cisailler, N_l, pour pouvoir remonter à la viscosité. Ce tableau montre que l'erreur diminue en augmentant le nombre d'échanges, c'est-à-dire lorsque les échanges sont plus présents, la méthode de Müller-Plathe [69] présente des résultats plus satisfaisants et se rapproche plus aux valeurs de la référence.
Une fois le nombre d'échanges optimum est obtenu, le calcul de la viscosité dynamique pour différentes densités ρ et températures T est effectué pour ce nombre d'échanges, $N_{Ex} = 200$, voir (Tab. A.5).

N_{proc}	τ_{mp}	τ_{gk}	μ_{mp}	μ_{gk}
1	04 :38 :56	15 :51 :02	0.937	0.8917
2	03 :20 :54	11 :27 :50	0.941	0.9228
4	01 :05 :41	04 :04 :27	0.904	0.9013
8	00 :46 :10	02 :54 :02	0.912	0.9219
16	00 :37 :21	02 :16 :52	0.901	0.9200

TABLE A.3 – Valeurs de la viscosité μ, pour différents nombres de processus N_{proc}. Les simulations ont été effectuées sur 5×10^6 d'itérations.

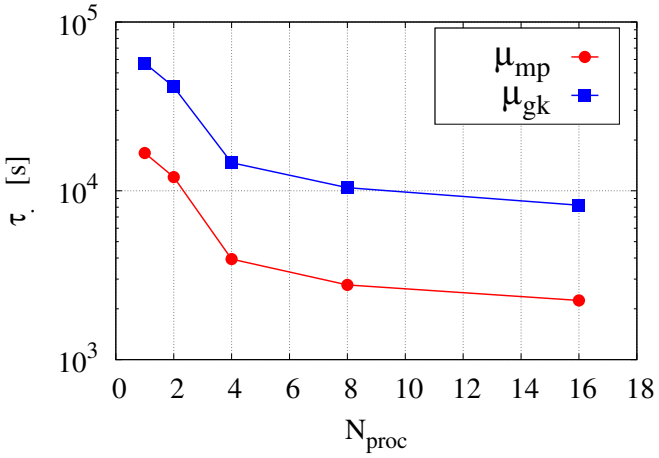


FIGURE A.3 – Variation de temps de calcul τ de la viscosité μ, en fonction du nombre de processus N_{proc} dans `LAMMPS` pour la méthode de MÜLLER-PLATHE [69](mp) et GREEN-KUBO (gk) [34].

A.2 Calcul du flux par la dynamique moléculaire

D'un point de vue milieux continus et comme discuté dans [73], le flux de chaleur total $\bar{q}_{tot}^{(c)}$ peut être divisé en trois contributions :

$$\bar{q}_{tot}^{(c)} = \bar{q}_k + \bar{q}_\mu + \bar{q}_{C_p} = -k\bar{\nabla}T - (\bar{\tau} \cdot \bar{v}) + \rho C_p \bar{v} T \tag{A.4}$$

où $\bar{q}_k = -k\bar{\nabla}T$ étant le flux de chaleur de conduction, $\bar{q}_\mu = -(\bar{\tau} \cdot \bar{v})$ est le flux associé au effets visqueux et $\bar{q}_{C_p} = \rho C_p \bar{v} T$ est le flux de convection. En dynamique moléculaire, le flux de chaleur, $\bar{q}_{tot}^{(m)}$, issue de transport interne entre les molécules est obtenue à l'équilibre par [27, 96] :

$$\begin{aligned} \bar{q}_{tot}^{(m)} &\equiv \frac{1}{V_c}\left(\frac{d}{dt}\sum_i \bar{r}_i\, U_i^{int}\right) \\ &= \frac{1}{V_c}\left(\underbrace{\sum_i \dot{\bar{r}}_i U_i^{int}}_{[T_1]} + \underbrace{\sum_i \bar{r}_i \frac{d}{dt}\, U_i^{int}}_{[T_2]}\right) \end{aligned} \tag{A.5}$$

où U_i^{int} est l'énergie interne de la particule i correspondant à l'énergie totale, $\bar{r}_i$ est la position de la particule i et $\dot{\bar{r}}_i$ sa vitesse. Le premier terme $[T_1]$ du membre de droite de l'équation

N_{Ex}	τ_{mp}	μ_{mp}	ε_{Mei} %
1	02 :32 :57	0.8971	5.77 %
5	01 :10 :14	0.8885	5.77 %
10	00 :38 :53	0.8878	5.52 %
50	00 :37 :48	0.8745	3.98 %
100	00 :39 :19	0.8540	1.54 %
200	00 :35 :55	0.8342	0.80 %
1000	04 :38 :56	0.8308	1.21 %

TABLE A.4 – Valeurs de la viscosité μ, pour différentes fréquences d'échanges de molécules (N_{Ex}).

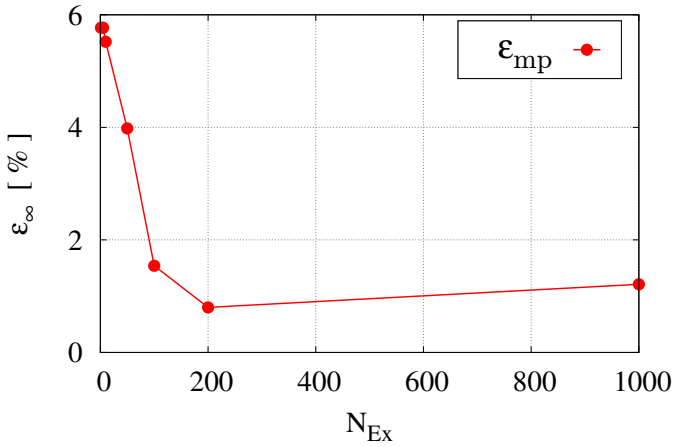


FIGURE A.4 – Erreurs max sur la viscosité μ, pour différentes fréquences d'échanges de molécules (N_{Ex}).

ρ	T	τ_{mp}	μ_{Mei}	μ_{Gal}	μ_{mp}	ε_{Mei} %	ε_{gal} %
0.6	1.5	01 :51 :02	0.841	0.835	0.834	0.83 %	0.12 %
0.5	1.5	01 :04 :46	0.581	0.586	0.696	19.7 %	18.7 %
0.4	1.5	00 :58 :27	0.438	0.414	0.522	19.1 %	26.0 %
0.3	1.5	00 :52 :30	0.327	0.321	0.351	7.33 %	9.34 %
0.2	1.5	00 :24 :07	0.252	0.244	0.257	1.98 %	5.32 %
0.1	1.5	00 :18 :21	0.199	–	0.196	1.50 %	−− %

TABLE A.5 – Valeurs de la viscosité μ, pour différentes densités ρ et températures T. Les simulations ont été effectuées sur 5×10^7 d'itérations ; nombre de molécules échangées $N_{Ex} = 200$; nombre de couches cisaillées $N_l = 20$.

précédente (A.5) peut se développer de la façon suivante :

$$
\begin{aligned}
[T_1] = \sum_i \bar{\dot{r}}_i \; U_i^{int} &= \sum_i \bar{\dot{r}}_i \left(\frac{1}{2} m_i \bar{\dot{r}}_i^2 + U_i^{pot} \right) \\
&= \sum_i \bar{\dot{r}}_i \left(\frac{1}{2} m_i \bar{\dot{r}}_i^2 + (U_i^{(1)} + U_i^{(2)}) \right) \\
&= \frac{1}{2} \sum_i m_i \bar{\dot{r}}_i^2 \bar{\dot{r}}_i + \frac{1}{2} \sum_i \sum_{j \neq i} U_{ij} \bar{\dot{r}}_i + \sum_i U_i \bar{\dot{r}}_i
\end{aligned}
\tag{A.6}
$$

où l'énergie potentielle U_i^{pot} se limite au potentiel volumique $U_i^{(1)} = U_i$ appliqué sur la particule i et au potentiel de paire entre les particules i et j : $U_{ij} = U_{ji} = U_{ij}(r_{ij})$. Le

potentiel de paire sur la particule i est $U_i^{(2)} = \sum_{j\neq i} U_{ij}$. Le facteur $\frac{1}{2}$ intervient, car la double sommation sur i et j prend deux fois en compte la contribution pour chaque paire. Ce terme $[T_1]$ est le plus dominant pour les gaz car le mécanisme est principalement la convection de particules, il est aussi appelé 'terme convectif' [27]. Le second terme $[T_2]$ peut également être développé sous la forme :

$$\begin{aligned}
[T_2] = \sum_i \bar{r}_i \frac{d}{dt}\, U_i^{int} &= \sum_i \bar{r}_i \frac{d}{dt}\left(\frac{1}{2} m_i \dot{\bar{r}}_i^2 + U_i^{pot}\right) \\
&= \sum_i \bar{r}_i \frac{d}{dt}\left(\frac{1}{2} m_i \dot{\bar{r}}_i^2 + U_i^{(1)} + U_i^{(2)}\right) \\
&= \underbrace{\sum_i \bar{r}_i \frac{d}{dt}\left(\frac{1}{2} m_i \dot{\bar{r}}_i^2\right)}_{[T_3]} + \underbrace{\sum_i \bar{r}_i \frac{d}{dt} U_i^{(2)}}_{[T_4]} + \underbrace{\sum_i \bar{r}_i \frac{d}{dt} U_i}_{[T_5]}
\end{aligned} \tag{A.7}$$

Le terme $[T_3]$ du membre de droite de l'Eq. (A.7) est obtenu en utilisant le théorème de l'énergie cinétique. À partir de la loi de Newton :

$$\begin{aligned}
m_i \frac{d\dot{\bar{r}}_i}{dt} &= \bar{F}_i \\
m_i \frac{d\dot{\bar{r}}_i}{dt} \cdot \dot{\bar{r}}_i &= \bar{F}_i \cdot \dot{\bar{r}}_i \\
m_i \frac{d\dot{\bar{r}}_i}{dt} \cdot \dot{\bar{r}}_i &= \left(\sum_{j\neq i} \bar{f}_{ij} + \bar{f}_i\right) \cdot \dot{\bar{r}}_i \\
\frac{d}{dt}\left(\frac{1}{2} m_i \dot{\bar{r}}_i^2\right) &= \sum_{j\neq i} \bar{f}_{ij} \cdot \dot{\bar{r}}_i + \bar{f}_i \cdot \dot{\bar{r}}_i \\
[T_3] = \sum_i \bar{r}_i \frac{d}{dt}\left(\frac{1}{2} m_i \dot{\bar{r}}_i^2\right) &= \sum_i \sum_{j\neq i} \bar{r}_i \left(\bar{f}_{ij} \cdot \dot{\bar{r}}_i\right) + \sum_i \bar{r}_i \left(\bar{f}_i \cdot \dot{\bar{r}}_i\right)
\end{aligned} \tag{A.8}$$

Comme la force d'interaction peut s'écrire :

$$\begin{aligned}
\bar{f}_{ij} &= -\frac{\partial U_i^{(2)}(r_{ij})}{\partial \bar{r}_i} \\
&= -\frac{\partial U_i^{(2)}(r_{ij})}{\partial \bar{r}_{ij}} \frac{\partial \bar{r}_{ij}}{\partial \bar{r}_i} \\
\bar{f}_{ij} &= \frac{\partial U_i^{(2)}(r_{ij})}{\partial \bar{r}_{ij}}
\end{aligned} \tag{A.9}$$

avec $\frac{\partial \bar{r}_{ij}}{\partial \bar{r}_i} = \frac{\partial(\bar{r}_j - \bar{r}_i)}{\partial \bar{r}_i} = -1$. Le terme $[T_4]$ de l'Eq. (A.7) peut se développer :

$$\begin{aligned}
[T_4] = \sum_i \bar{r}_i \frac{d}{dt} U_i^{(2)} &= \frac{1}{2} \sum_i \sum_{j \neq i} \bar{r}_i \frac{d}{dt} U_{ij} \\
&= \frac{1}{2} \sum_i \sum_{j \neq i} \bar{r}_i \frac{dU_{ij}}{d\bar{r}_{ij}} \frac{d\bar{r}_{ij}}{dt} \\
&= \frac{1}{2} \sum_i \sum_{j \neq i} \bar{r}_i \left(\bar{f}_{ij} \cdot \left[\frac{d}{dt} (\bar{r}_j - \bar{r}_i) \right] \right) \\
&= \frac{1}{2} \sum_i \sum_{j \neq i} \bar{r}_i \left(\bar{f}_{ij} \cdot (\dot{\bar{r}}_j - \dot{\bar{r}}_i) \right)
\end{aligned} \tag{A.10}$$

Le terme $[T_5]$ s'écrit :

$$\begin{aligned}
[T_5] = \sum_i \bar{r}_i \frac{d}{dt} U_i = \sum_i \bar{r}_i \frac{dU_i}{d\bar{r}_i} \frac{\bar{r}_i}{dt} \\
= \sum_i \bar{r}_i \left(-\bar{f}_i \cdot \dot{\bar{r}}_i \right)
\end{aligned} \tag{A.11}$$

L'expression finale du second terme $[T_2]$ sera alors :

$$\begin{aligned}
[T_2] &= \sum_i \bar{r}_i \frac{d}{dt} \, U_i^{int} \\
&= \sum_i \sum_{j \neq i} \bar{r}_i \left(\bar{f}_{ij} \cdot \dot{\bar{r}}_i \right) + \sum_i \bar{r}_i \left(\bar{f}_i \cdot \dot{\bar{r}}_i \right) - \sum_i \bar{r}_i \left(\bar{f}_i \cdot \dot{\bar{r}}_i \right) + \frac{1}{2} \sum_i \sum_{j \neq i} \bar{r}_i \left[\bar{f}_{ij} \cdot (\dot{\bar{r}}_j - \dot{\bar{r}}_i) \right] \\
&= \frac{1}{2} \sum_i \sum_{j \neq i} \bar{r}_i \left[\bar{f}_{ij} \cdot (\dot{\bar{r}}_j + \dot{\bar{r}}_i) \right]
\end{aligned} \tag{A.12}$$

Par permutation d'indices :

$$\begin{aligned}
[T_2] &= \frac{1}{2} \sum_j \sum_{i \neq j} \bar{r}_j \left[\bar{f}_{ji} \cdot (\dot{\bar{r}}_j + \dot{\bar{r}}_i) \right] \\
&= -\frac{1}{2} \sum_j \sum_{i \neq j} \bar{r}_j \left[\bar{f}_{ij} \cdot (\dot{\bar{r}}_i + \dot{\bar{r}}_j) \right]
\end{aligned} \tag{A.13}$$

par sommation :

$$\begin{aligned}
2 \times [T_2] &= \frac{1}{2} \sum_j \sum_{i \neq j} \underbrace{(\bar{r}_i - \bar{r}_j)}_{-\bar{r}_{ij}} \left[\bar{f}_{ij} \cdot (\dot{\bar{r}}_i + \dot{\bar{r}}_j) \right] \\
[T_2] &= -\frac{1}{4} \sum_i \sum_{j \neq i} \bar{r}_{ij} \left[\bar{f}_{ij} \cdot (\dot{\bar{r}}_i + \dot{\bar{r}}_j) \right]
\end{aligned} \tag{A.14}$$

En distribuant puis en inversant les indices :

$$\begin{aligned}
[T_2] &= -\frac{1}{4}\sum_i\sum_{j\neq i}\bar{r}_{ij}\left[\bar{f}_{ij}\cdot(\dot{\bar{r}}_i+\dot{\bar{r}}_j)\right]\\
&= -\frac{1}{4}\sum_i\sum_{j\neq i}\bar{r}_{ij}\left[\bar{f}_{ij}\cdot\dot{\bar{r}}_i\right]\underbrace{-\frac{1}{4}\sum_i\sum_{j\neq i}\bar{r}_{ij}\left[\bar{f}_{ij}\cdot\dot{\bar{r}}_j\right]}_{\text{en inversant les indices}}\\
&= -\frac{1}{4}\sum_i\sum_{j\neq i}\bar{r}_{ij}\left[\bar{f}_{ij}\cdot\dot{\bar{r}}_i\right]-\frac{1}{4}\sum_j\sum_{i\neq j}\underbrace{\bar{r}_{ji}}_{-\bar{r}_{ij}}\left[\underbrace{\bar{f}_{ji}}_{-\bar{f}_{ij}}\cdot\dot{\bar{r}}_i\right]\\
&= 2\times\left(-\frac{1}{4}\sum_i\sum_{j\neq i}\bar{r}_{ij}\left[\bar{f}_{ij}\cdot\dot{\bar{r}}_i\right]\right)
\end{aligned} \tag{A.15}$$

L'expression finale du second terme $[T_2]$ est alors :

$$[T_2] = \sum_i \bar{r}_i\frac{d}{dt}\,U_i^{int} = -\frac{1}{2}\sum_i\sum_{j\neq i}\bar{r}_{ij}\left[\bar{f}_{ij}\cdot\dot{\bar{r}}_i\right] \tag{A.16}$$

À partir des deux équations (A.6) et (A.16) substituées dans (Eq. (A.5)) nous retrouvons alors l'expression finale du flux [62, 96] :

$$\bar{q}_{tot}^{(m)} = \frac{1}{2V_c}\left(\sum_i m_i\dot{\bar{r}}_i^2\dot{\bar{r}}_i+\sum_i\sum_{j\neq i}U_{ij}\dot{\bar{r}}_i-\sum_i\sum_{j\neq i}\bar{r}_{ij}\left[\bar{f}_{ij}\cdot\dot{\bar{r}}_i\right]+2\sum_i U_i\dot{\bar{r}}_i\right) \tag{A.17}$$

Si la force que l'on veut imposer est une force $\bar{f} = (f_x, f_y, f_z)^\intercal$ issue d'un potentiel, le potentiel dont elle dérive est :

$$\bar{f}_i = -\frac{\partial U_i}{\partial\bar{r}_i} \quad\Rightarrow\quad U_i = -(\bar{f}\cdot\bar{r}_i) \tag{A.18}$$

avec I_3 la matrice identité. L'expression du flux, comme pour Irving-Kirkwood, devient :

$$\bar{q}_{tot}^{(m)} = \frac{1}{2V_c}\left(\sum_i m_i\dot{\bar{r}}_i^2\dot{\bar{r}}_i+\sum_i\sum_{j\neq i}U_{ij}\dot{\bar{r}}_i-\sum_i\sum_{j\neq i}\bar{r}_{ij}\left[\bar{f}_{ij}\cdot\dot{\bar{r}}_i\right]-2\sum_i\left(\bar{f}\cdot\bar{r}_i\right)\dot{\bar{r}}_i\right) \tag{A.19}$$

A.3 Description du maillage co-localisé

La figure A.5 montre les détails sur ce maillage et sur la discrétisation effectuée pour la résolution des équations de Navier-Stokes et de l'énergie, de (2.14) à (2.17). Soient

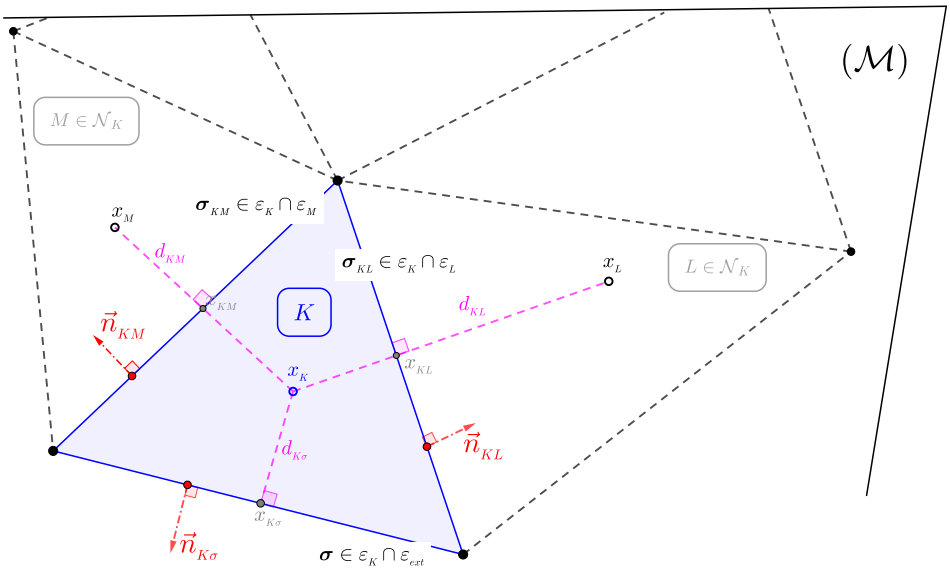


FIGURE A.5 – Illustration bidimensionnelle du maillage de Delaunay – Notations principales utilisées pour la discrétisation

+ $\mathcal{M}$ l'ensemble des mailles fluides ou solides, en fonction de la position dans le domaine de calcul,
+ $\mathcal{E}_{int}$ et $\mathcal{E}_{ext}$ les ensembles des faces des volumes de contrôle fluide, intérieures et extérieures respectivement au domaine Ω.

Pour chaque volume de contrôle $K \in \mathcal{M}$, nous définissons :

* x_K un point caractéristique, souvent appelé milieu du volume de contrôle, de coordonnées $\bar{x}_K$,
* $m(K)$ son aire (2D) ou volume (3D),
* $\mathcal{N}_K$ l'ensemble des volumes de contrôle voisins par les arêtes (2D) ou faces (3D),
* $\mathcal{E}_K$, l'ensemble des arêtes (2D) ou faces (3D) de la maille K,
  ** Si $\sigma \in \mathcal{E}_K \cap \mathcal{E}_{ext}$, nous désignons par
     — $m(\sigma)$ la longueur (2D) ou surface (3D) de l'interface σ,
     — $d_{K\sigma}$ la distance entre x_σ et x_K,
     — $\bar{n}_{K\sigma}$ le vecteur unitaire orthogonal (la normale) à $\sigma \in \mathcal{E}_K$, sortant de K,
     — $x_{K\sigma}$ le centre de gravité du bord $\sigma \in \mathcal{E}_{K,ext}$ et $\bar{x}_{K\sigma}$ ses coordonnées.
  ** Si $L \in \mathcal{N}_K$, $\sigma_{KL} \in \mathcal{E}_K \cap \mathcal{E}_L$ est la face commune entre les mailles K et L, nous définissons :
     — $m(\sigma_{KL})$ sa longueur (2D) ou surface (3D) de l'interface σ_{KL},

— d_{KL} la distance entre x_K et x_L,

— $\bar{n}_{KL}$ le vecteur unitaire orthogonal à σ_{KL}, sortant de K,

— x_{KL} la projection orthogonale de x_K sur le bord $\sigma_{KL} \in \mathcal{E}_{K,int}$, avec $L \in \mathcal{N}_K$, et $\bar{x}_{KL}$ ses coordonnées.

Afin de garantir que la solution diffusive soit monotone, il faut s'assurer que le flux de diffusion est toujours dirigé vers l'intérieur de la zone dans laquelle la concentration est la plus élevée. Pour cela, les maillages sont construits de manière à ce qu'ils satisfassent des conditions qui permettent à la fois de garantir la monotonie et de rendre également possible la construction du flux normal à la frontière commune entre les mailles K et L et ce, en utilisant simplement les inconnues de ces deux volumes de contrôle. Ces deux conditions sont les suivantes : $\forall K \in \mathcal{M}$, $\forall L \in \mathcal{N}_K$,

$$\sigma_{KL} \perp (x_K x_L) \tag{A.20a}$$

$$(\bar{x}_L - \bar{x}_K) \cdot \bar{n}_{KL} > 0 \tag{A.20b}$$

Pour pouvoir décrire de manière compendieuse les flux de masse par $\sigma_{KL} \in \mathcal{E}_K \cap \mathcal{E}_L$, il faudra en outre imposer que x_{KL} se trouve sur la droite $(x_K x_L)$, ou en d'autres termes que la direction de la droite $(x_K x_L)$ soit $\bar{n}_{KL}$:

$$(\bar{x}_{KL} - \bar{x}_K) \wedge \bar{n}_{KL} = \bar{0} \tag{A.21}$$

Annexe B

Contexte sur l'apprentissage de hyperparamètres

B.1 Fonction de vraisemblance, priori, posteriori, production productive

B.1.1 Fonction de vraisemblance

Avec un apprentissage de N paire de données : $\bar{y} = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, et en regardant simultanément les N points d'entraînement, de sorte que le vecteur $\bar{y}$ représente tous les points d'apprentissage individuels $y_{(i)}$, il est possible d'obtenir l'expression de la fonction de vraisemblance, qui est une fonction des paramètres ajustables : les coefficients $\bar{w}$ et l'inverse de la variance β qui représente la précision sur les données. En substituant σ^2 par β^{-1}, cette fonction de vraisemblance s'écrit sous la forme (Eq. B.1) :

$$\begin{aligned} p\left(\bar{f}|\bar{x}, \bar{w}\right) &= \prod_{i=1}^{N} p\left(f_{(i)}|\bar{w}^T \bar{\phi}\left(x_i\right), \beta^{-1}\right) \\ &= \prod_{i=1}^{N} \left(2\pi\beta^{-1}\right)^{-\frac{1}{2}} \exp\left\{-\frac{\beta}{2}\left(y_{(i)} - \bar{w}^T \bar{\phi}\left(x_i\right)\right)^2\right\} \\ &= \mathcal{N}\left(f_{(i)}|\bar{w}^T \bar{\phi}\left(x_i\right), \bar{\bar{S_N}}\right) \end{aligned} \tag{B.1}$$

où $\bar{\phi}\left(x_i\right)$ est le vecteur à la i^{e} ligne de la matrice de conception.Il faut toutefois noter qu'il n'est pas indispensable que ces fonctions aient une signification probabiliste, et en particulier que le coefficient de normalisation n'a pas d'importance dans la mesure où ces fonctions de base seront multipliées par des paramètres adaptatifs w_j. Pour rendre plus explicite le lien entre la régression linéaire et les gaussiens, le modèle peut être réécrit sous la forme :

$$\begin{aligned} p\left(y|x, \bar{w}\right) &\sim \mathcal{N}\left(\bar{w}^T \bar{\phi}(x), \sigma^2\right) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left(y - \bar{w}^T \bar{\phi}\left(x\right)\right)^2\right\} \end{aligned} \tag{B.2}$$

Ainsi, la moyenne de la distribution est $\bar{w}^T \bar{\phi}(x)$, qui est située au point dont la probabilité gaussienne est maximale, sachant que la distribution gaussienne a une valeur maximale dans

l'axe des moyennes. Par conséquent, en maximisant la probabilité, cela permet à la prédiction de s'adapter aux données. Il convient de souligner que la moyenne $\bar{w}^T\bar{\phi}(x)$ est une variable dépendante de x_i, ce qui implique que celle-ci prendra sa propre valeur pour différentes entrées x_i. Plus loin, lorsque le terme "distribution prédictive" sera évoqué et introduit, il s'avérera que $\bar{w}^T\bar{\phi}(x)$ est en réalité la prédiction de la sortie. La variance σ^2 correspond à la variance du biais dans les données.

Pour résumer, la probabilité $y_i = \mathcal{N}(\bar{w}^T\bar{\phi}(x), \sigma^2)$ est une distribution gaussienne univariée qui a une distribution différente en chaque point de l'axe x, où $\bar{w}^T\bar{\phi}(x)$ est la prédiction de y_i, et est une fonction de x.

L'estimation MLE (estimation maximale de la vraisemblance ou Maximal Likelihood Estimation en anglais) consiste à maximiser la vraisemblance. Pour cela, il est opportun de maximiser le logarithme de la fonction de vraisemblance. En substituant la forme de la gaussienne, donnée par (Eq. B.1). La fonction de vraisemblance logarithmique s'écrit alors sous la forme :

$$\begin{aligned}
\ln p\left(\bar{y}|x_i, \bar{w}, \beta\right) &= \ln\left[\left(2\pi\beta^{-1}\right)^{-\frac{N}{2}} \exp\left\{-\frac{\beta}{2}\sum_{i=1}^{N}\left(y_i - \bar{w}^T\bar{\phi}\left(x_i\right)\right)^2\right\}\right] \\
&= \ln\left[\left(2\pi\beta^{-1}\right)^{-\frac{N}{2}}\right] + \ln\left[\exp\left\{-\frac{\beta}{2}\sum_{i=1}^{N}\left(y_i - \bar{w}^T\bar{\phi}\left(x_i\right)\right)^2\right\}\right] \\
\ln p\left(\bar{y}|x_i, \bar{w}, \beta\right) &= \frac{N}{2}\ln\beta - \frac{N}{2}\ln(2\pi) - \frac{\beta}{2}\sum_{i=1}^{N}\left\{f\left(x_i, \bar{w}\right) - y_i\right\}^2
\end{aligned} \tag{B.3}$$

La solution qui maximisera la vraisemblance correspond au paramètre de régression w qui annule la première dérivée de l'équation (Eq. B.3) par rapport à $\bar{w}$.

$$\begin{aligned}
\nabla_{\bar{w}} \ln p(\bar{y}|x_i, \bar{w}, \beta) &= \nabla_{\bar{w}}\left[\frac{N}{2}\ln\beta - \frac{N}{2}\ln(2\pi) - \frac{\beta}{2}\sum_{i=1}^{N}\left\{f\left(x_i, \bar{w}\right) - y_i\right\}^2\right] = 0 \\
&= \left[\cancel{\nabla_{\bar{w}}(\frac{N}{2}\ln\beta)} - \cancel{\nabla_{\bar{w}}(\frac{N}{2}\ln(2\pi))} - \nabla_{\bar{w}}\left(\frac{\beta}{2}\sum_{i=1}^{N}\left\{f\left(x_i, \bar{w}\right) - y_i\right\}^2\right)\right] = 0 \\
\nabla_{\bar{w}} \ln p(\bar{y}|x_i, \bar{w}, \beta) &= \nabla_{\bar{w}}\left[\frac{1}{2}\sum_{i=1}^{N}\left\{f\left(x_i, \bar{w}\right) - y_i\right\}^2\right] = 0
\end{aligned} \tag{B.4}$$

la dernière forme de l'équation (Eq. B.4) est similaire à la fonction d'erreur $J(\bar{w})$, l'équation (Eq. 4.13) à maximiser dans la régression linéaire par les moindres carrés (LSE). En maximisant la fonction d'erreur (Eq. 4.13), la dérivée première sera égale à zéro $\nabla_{\bar{w}}\varepsilon^{(p)}(\bar{w}) = 0$. Ainsi, il est possible d'affirmer que la régression linéaire par moindres carrés (LSE) correspond à la MLE (estimation maximale de la vraisemblance ou Maximal Likelihood Estimation en anglais) dans la méthode bayésienne.

$$\bar{w}_{MLE} = \bar{w}^{\star} \tag{B.5}$$

B.1.2 Distribution *a priori*

Afin d'apprendre la relation entre x et y, nous cherchons à limiter la complexité et donc la croissance des coefficients (poids $\bar{w}$). Cela se fait en définissant une distribution de probabilité *a priori* explicite sur $\bar{w}$. Cette probabilité *a priori* est sous une forme particulière afin de simplifier le traitement. Des fonctions plus lisses et donc moins complexes sont privilégiées. Notons également que la fonction de vraisemblance $p(\bar{y}|\bar{w})$ définie par (Eq. B.1) est l'exponentielle d'une fonction quadratique de $\bar{w}$. Il s'agit alors encore une fois d'une gaussienne univariée. La distribution *a priori* conjuguée correspondante est alors donnée par une distribution gaussienne. Plus précisément, nous considérons une gaussienne multivariée de forme isotrope à moyenne $\bar{m}_0$ nulle régie par un paramètre de précision α (matrice covariance $\bar{\bar{S}}_0 =$ scalaire $*$ matrice d'identité) de sorte que :

$$\begin{aligned} p(\bar{w}|\alpha) &= \mathcal{N}\left(\bar{w}|\bar{m}_0, \bar{\bar{S}}_0\right) \\ &= \mathcal{N}\left(\bar{w}|\bar{0}, \alpha^{-1}\bar{\bar{I}}\right) \\ &= \left(\frac{\alpha}{2\pi}\right)^{\frac{M}{2}} \exp\left\{-\frac{\alpha}{2}\bar{w}^{\mathrm{T}}\bar{w}\right\} \end{aligned} \tag{B.6}$$

où α décrit l'inverse de la variance et donc la précision sur la distribution sur les coefficients $\bar{w}$ et M est le nombre total d'éléments du vecteur $\bar{w}$ pour un polynôme d'ordre $(M-1)$. Le scalaire de précision α ou matrice de précision $\bar{\bar{S}}_0$ est utilisé afin de faciliter la dérivation. Cette variable α, qui contrôle la distribution des paramètres du modèle, est appelée hyperparamètre.

B.1.3 Distribution *a posteriori*

La distribution *a posteriori* sur tous les paramètres inconnus, compte tenu des données, est proportionnelle au produit de la fonction de vraisemblance et de la distribution *a priori* (Eq. B.7).

$$\begin{aligned} \text{distribution } a \textit{ posteriori} \;&\propto\; \text{vraisemblance} \;\times \text{distribution } a \textit{ priori} \\ p(\bar{w}|x, \bar{y}, \alpha, \beta) &\propto\; p(\bar{y}|x, \bar{w}, \beta)\;\; p(\bar{w}|\alpha) \end{aligned} \tag{B.7}$$

En conséquence du choix d'une distribution *a priori* gaussienne conjuguée, la distribution *a posteriori* sera également gaussienne exprimée par l'équation (Eq. B.12). L'enjeu est de trouver les coefficients $\bar{w}$, α et β qui maximisent cette probabilité *a posteriori*.

$$p(\bar{w}|x, \bar{y}, \alpha, \beta) \propto \underbrace{\left[\left(2\pi\beta^{-1}\right)^{-\frac{N}{2}} \exp\left\{-\frac{\beta}{2}\sum_{i=1}^{N}\left(y_i - \bar{w}^T\bar{\phi}\left(x_i\right)\right)^2\right\}\right]}_{\text{vraisemblance}} \underbrace{\left[\left(\frac{\alpha}{2\pi}\right)^{\frac{M}{2}} \exp\left\{-\frac{\alpha}{2}\bar{w}^{\mathrm{T}}\bar{w}\right\}\right]}_{\text{distribution } a \textit{ priori}} \tag{B.8}$$

Le logarithmique de la distribution *a posteriori* s'écrit :

$$\begin{aligned}
\ln p(\bar{w}|x,\bar{y},\alpha,\beta) &\propto \ln(\text{ vraisemblance }) + \ln(\text{distribution } a \text{ priori }) \\
\ln p(\bar{w}|x,\bar{y},\alpha,\beta) &\propto \ln\left[\left(2\pi\beta^{-1}\right)^{-\frac{N}{2}} \exp\left\{-\frac{\beta}{2}\sum_{i=1}^{N}\left(y_i - \bar{w}^T\bar{\phi}\left(x_i\right)\right)^2\right\}\right] \\
&\quad + \ln\left[\left(\frac{\alpha}{2\pi}\right)^{\frac{M}{2}} \exp\left\{-\frac{\alpha}{2}\bar{w}^{\mathrm{T}}\bar{w}\right\}\right] \\
\ln p(\bar{w}|x,\bar{y},\alpha,\beta) &\propto \left[\frac{N}{2}\ln\beta - \frac{N}{2}\ln(2\pi) - \frac{\beta}{2}\sum_{i=1}^{N}\left\{f\left(x_i,\bar{w}\right) - y_i\right\}^2\right] \\
&\quad + \left[\frac{M}{2}\ln\alpha - \frac{M}{2}\ln(2\pi) - \frac{\alpha}{2}\bar{w}^{\mathrm{T}}\bar{w}\right]
\end{aligned} \tag{B.9}$$

Encore une fois, la première dérivée permettra d'obtenir la solution.

$$\begin{aligned}
\nabla_{\bar{w}}\ln p(\bar{w}|x,\bar{y},\alpha,\beta) &\propto \nabla_{\bar{w}}\left[\frac{N}{2}\ln\beta - \frac{N}{2}\ln(2\pi) - \frac{\beta}{2}\sum_{i=1}^{N}\left\{f\left(x_i,\bar{w}\right) - y_i\right\}^2\right] \\
&\quad + \nabla_{\bar{w}}\left[\frac{M}{2}\ln\alpha - \frac{M}{2}\ln(2\pi) - \frac{\alpha}{2}\bar{w}^{\mathrm{T}}\bar{w}\right] = 0 \\
\nabla_{\bar{w}}\ln p(\bar{w}|x,\bar{y},\alpha,\beta) &= \nabla_{\bar{w}}\left[\frac{\beta}{2}\sum_{i=1}^{N}\left\{f\left(x_i,\bar{w}\right) - y_i\right\}^2\right] - \nabla_{\bar{w}}\left[\frac{\alpha}{2}\bar{w}^{\mathrm{T}}\bar{w}\right] = 0
\end{aligned} \tag{B.10}$$

L'équation (Eq. B.10) est similaire à la dérivée de la fonction d'erreur de la technique de régularisation (Eq. 4.15) $\nabla_{\bar{w}}\varepsilon^{(p)}(\bar{w}) = 0$:

$$\begin{aligned}
\nabla_{\bar{w}}\varepsilon^{(p)}(\bar{w}) &= \nabla_{\bar{w}}\frac{1}{2}\left[\sum_{i=1}^{N}\left\{f\left(x_i,\bar{w}\right) - y_i\right\}^2 + \lambda\sum_{j=1}^{M-1}w_j^2\right] \\
&= \nabla_{\bar{w}}\frac{1}{2}\left[-\beta\sum_{i=1}^{N}\left\{f\left(x_i,\bar{w}\right) - y_i\right\}^2 - \alpha\ \bar{w}^{\mathrm{T}}\bar{w}\right] = 0 \quad \text{en divisant } /(-\beta) \\
&= \nabla_{\bar{w}}\frac{1}{2}\left[\sum_{i=1}^{N}\left\{f\left(x_i,\bar{w}\right) - y_i\right\}^2 + \frac{\alpha}{\beta}\ \bar{w}^{\mathrm{T}}\bar{w}\right] = 0
\end{aligned} \tag{B.11}$$

En examinant de près les deux expressions, $\frac{\alpha}{\beta}$ correspond à λ, et $\bar{w}^{\mathrm{T}}\bar{w}$ n'est que $\sum_{j=1}^{M-1}w_j^2$. Finalement, la régression par méthode des moindres carrés (LSE) avec régularisation est égale à MAP (Maximum A Posteriori) de la méthode de régression bayésienne. La seule différence est que les hyperparamètres α et β, sont à leur tour appris et mis à jour dans l'algorithme de la méthode bayésienne.

Dans un premier temps, l'approche consiste à utiliser les valeurs de α et β connus *a priori* et calculer la distribution *a posteriori* des coefficients, l'apprentissage ne se fait alors que sur les coefficients. Pour N paire de données d'entrée, la distribution *a posteriori* est donnée par :

$$p(\bar{w}|\bar{f},\alpha,\beta) = \mathcal{N}\left(\bar{w}|\bar{m}_N,\bar{\bar{S}}_N\right) \tag{B.12}$$

avec

$$\begin{cases} \bar{\bar{S}}_N^{-1} = & \alpha \bar{\bar{I}} + \beta \bar{\bar{\Phi}}^{\mathrm{T}} \bar{\bar{\Phi}} \\ \bar{m}_N = & \beta \bar{\bar{S}}_N \bar{\bar{\Phi}}^{\mathrm{T}} \bar{y} \end{cases} \tag{B.13}$$

La distribution *a posteriori* étant gaussienne, son mode et sa moyenne sont identiques [6]. Par conséquent, le vecteur de poids maximal *a posteriori* est obtenu par $\bar{w}_{MAP} = \bar{m}_N$. En considérant une distribution *a priori* infiniment large $\bar{\bar{S}}_0 = \alpha^{-1}\bar{\bar{I}}$, avec α tend vers 0, la moyenne de la distribution *a posteriori* $\bar{m}_N = \bar{w}_{ML}$. Cette moyenne se ramène donc à la valeur maximale de vraisemblance obtenue en mettant le gradient du log de la vraisemblance à zéro et en résolvant par rapport à $\bar{w}$ (Eq. B.4).

Dans un cadre de LSE, le log négatif de la distribution *a priori* est parfois considéré comme un terme de pénalité, et le résultat de la MAP est connu comme l'estimation pénalisée du maximum de vraisemblance des poids, ce qui peut entraîner une certaine confusion entre les deux approches. Cependant, il convient de savoir que dans le cadre bayésien, l'estimation MAP ne joue aucun rôle particulier. La procédure de maximum de vraisemblance pénalisée est connue dans ce cas sous le nom de régression ridge [41] en raison de l'effet du terme de pénalité quadratique de la priorité logarithmique.

B.1.4 Distribution prédictive

L'enjeu principal de l'apprentissage est de prédire l'évolution de la cible $f_\star$ pour de nouvelles données $\bar{x}$, sans se soucier pour autant de la valeur de $\bar{w}$. La distribution prédictive permet d'arriver à cet objectif, elle est définie par [6], où $\bar{y}$ représente le vecteur des valeurs cible de l'ensemble des données, $\sigma_N^2(x)$ représente la confiance ou l'incertitude sur la prédiction donnée par (Eq. B.15).

$$\begin{aligned} p(f|\bar{f}, \alpha, \beta) &= \int p(f|\bar{w}, \beta)\ p(\bar{w}|\bar{f}, \alpha, \beta) \mathrm{d}\bar{w} \\ &= \mathcal{N}\left(f|\bar{m}_N^{\mathrm{T}} \bar{\phi}(x), \sigma_N^2(x)\right) \end{aligned} \tag{B.14}$$

La distribution prédictive est obtenue en combinant les deux distributions : conditionnelle et *a posteriori*. Le premier terme qui est donné par l'Eq. (Eq. B.2) correspond à la distribution conditionnelle et le second terme représentant la distribution *a posteriori* des paramètres est donné par l'Eq. (Eq. B.12).

$$\sigma_N^2(x) = \frac{1}{\beta} + \bar{\phi}(x)^{\mathrm{T}} \bar{\bar{S}}_N \phi(x) \tag{B.15}$$

où l'inverse de β représente la variance et correspond à l'incertitude sur les données, donc au bruit imposé, et le terme $\bar{\phi}(x)^{\mathrm{T}} \bar{S}_N \phi(x)$ correspond à l'incertitude sur les coefficients ou paramètres $\bar{w}$. Du fait que le bruit et la distribution des paramètres soient des distributions gaussiennes indépendantes, leurs variances sont alors additives. Il est également important de souligner que, avec l'augmentation de la base de données, c'est-à-dire après plusieurs apprentissages, la distribution *a posteriori* devient plus mince. Il est alors possible de montrer que l'incertitude diminue au fur et à mesure que les points de données sont observés

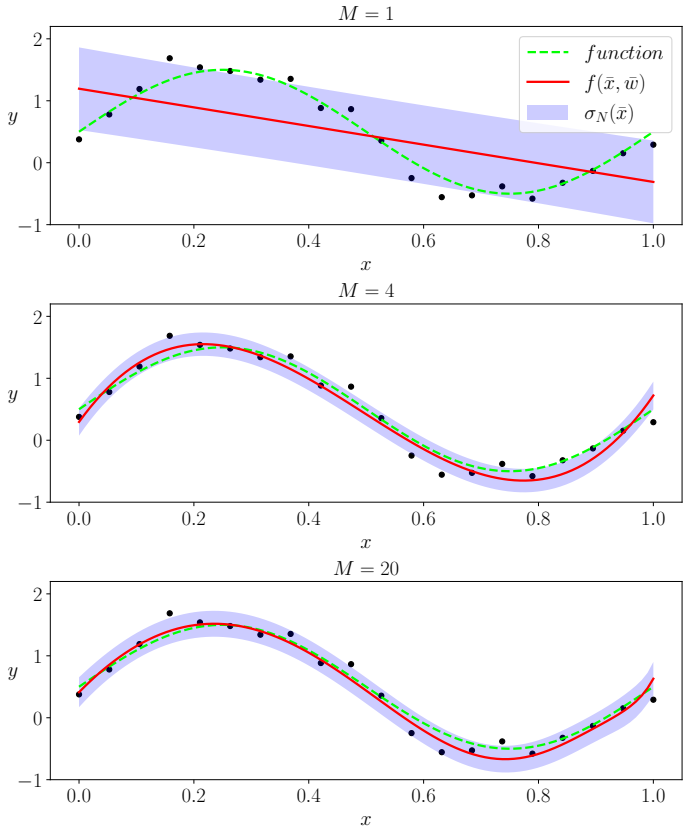


FIGURE B.1 – Représentation des incertitudes sur la prédiction σ_N pour différents ordres polynomiaux $M = [1, \ldots, 8]$, où le bruit aléatoire inclus dans les valeurs de la variable cible $f_\star$ est fixé à $\varepsilon = 1 \times 10^{-2}$. Pour $N >>$ l'incertitude sur la prédiction $\sigma_N \to \varepsilon$.

$(\sigma_{N+1}(x) \leq \sigma_N(x))$ démontré par [80]. Par conséquent, dans la limite où $N \to \infty$, le second terme de l'équation (Eq. B.15) tend vers zéro $(\bar{\phi}(x)^{\mathrm{T}} \bar{S}_N \phi(x) \to 0)$. Finalement, l'incertitude sur la prédiction sera uniquement due au bruit et sa valeur sera égale à l'inverse de β (voir (Fig. 4.10)).

B.2 Apprentissage des hyperparamètres : Bayésien/Gaussien

Étant donné que l'approche bayésienne doit faire des hypothèses *a priori* sur la forme du modèle, sur les hyperparamètres α et β, et si celles-ci ne sont pas correctes, les résultats peuvent être erronés. Notamment, l'évidence du modèle peut être sensible au choix *a priori*. Il est alors judicieux de faire des prédictions en marginalisant par rapport à ces hyperparamètres ainsi que par rapport aux paramètres $\bar{w}$. Cette procédure est appelée approximation de l'évidence ([38] ; [61]). Dans ce qui suit, les développements selon les deux auteurs (Bishop [6] et Tipping [94]) sont détaillés afin de comprendre l'apprentissage des hyperparamètres :

B.2.1 Développement pour le cas de regression

Si l'on introduit des hyperprioris sur α et β, la distribution prédictive est obtenue en marginalisant sur $\bar{w}$, α et β. La fonction de vraisemblance marginale est obtenue en intégrant sur les paramètres de poids $\bar{w}$, de sorte que :

$$p(\bar{y}|\alpha,\beta) = \int p(\bar{y}|\bar{w},\beta)p(\bar{w}|\alpha)\mathrm{d}\bar{w} \tag{B.16}$$

Les paramètres sont estimés en maximisant la vraisemblance logarithmique marginale s'écrivant sous la forme (Eq. B.17), où avec $\bar{\bar{S}}_N^{-1}$ est donné par la relation (Eq. B.13) et $E\left(\bar{m}_N\right)$ par (Eq. B.18) :

$$\ln p(\bar{y}|\alpha,\beta) = \frac{M}{2}\ln\alpha + \frac{N}{2}\ln\beta - E\left(\bar{m}_N\right) - \frac{1}{2}\ln|\bar{\bar{S}}_N^{-1}| - \frac{N}{2}\ln(2\pi) \tag{B.17}$$

$$E\left(\bar{m}_N\right) = \frac{\beta}{2}\left\|\bar{y} - \bar{\bar{\Phi}}\bar{m}_N\right\|^2 + \frac{\alpha}{2}\bar{m}_N^{\mathrm{T}}\bar{m}_N \tag{B.18}$$

Deux étapes permettront d'obtenir les hyperparamètres : Considérons d'abord la maximisation de $p(\bar{t}|\alpha,\beta)$ en fonction de α. En définissant l'équation du vecteur des valeurs propres (Eq. B.19), et en calculant la dérivée de la vraisemblance logarithmique marginale, la quantité γ peut être facilement obtenue. Cela permettra de remonter à la valeur de α et β souhaitée. Commençons par définir l'équation du vecteur des valeurs propres :

$$\left(\beta\bar{\bar{\Phi}}^{\mathrm{T}}\bar{\bar{\Phi}}\right)\bar{u}_i = \lambda_i\bar{u}_i \tag{B.19}$$

Maximiser la vraisemblance marginale revient à dériver la logarithmique de cette fonction et l'égaler à zéro :

$$\begin{aligned}\nabla_\alpha \ln p(\bar{y}|\alpha,\beta) &= \nabla_\alpha\left[\frac{M}{2}\ln\alpha + \frac{N}{2}\ln\beta - E\left(\bar{m}_N\right) - \frac{1}{2}\ln|\bar{\bar{S}}_N^{-1}| - \frac{N}{2}\ln(2\pi)\right] = 0\\ 0 &= \frac{M}{2\alpha} + 0 - \frac{1}{2}\bar{m}_N^{\mathrm{T}}\bar{m}_N - \frac{1}{2}\nabla_\alpha\ln|\bar{\bar{S}}_N^{-1}| - 0\end{aligned} \tag{B.20}$$

À partir de la relation (Eq. B.13) et l'équation du vecteur des valeurs propres définie, la matrice $\bar{\bar{S}}_N^{-1}$ a comme valeurs propres $\alpha + \lambda_i$, donc la dérivée de $\ln|\bar{\bar{S}}_N^{-1}|$ par rapport à α s'écrit sous la forme :

$$\nabla_\alpha \ln|\bar{\bar{S}}_N^{-1}| = \nabla_\alpha \ln\prod_i(\alpha+\lambda_i) = \nabla_\alpha\sum_i\ln(\alpha+\lambda_i) = \sum_i\frac{1}{\alpha+\lambda_i} \tag{B.21}$$

En combinant les deux relations (Eq. B.20) et (Eq. B.21), la quantité γ peut s'écrire :

$$\alpha\bar{m}_N^{\mathrm{T}}\bar{m}_N = M - \alpha\sum_i\frac{1}{\lambda_i+\alpha} = \gamma \tag{B.22}$$

Finalement, la valeur de l'hyperparamètre α qui maximise la vraisemblance obtenue à partir de l'équation (Eq. B.22) est :

$$\alpha = \frac{\gamma}{\bar{m}_N^{\mathrm{T}}\bar{m}_N} \tag{B.23}$$

où $\bar{m}_N$ est donné par l'équation (Eq. B.13), et en considérant qu'il y a M termes dans la somme sur i.

$$\gamma = \sum_i \frac{\lambda_i}{\alpha + \lambda_i} \tag{B.24}$$

Une fois la valeur de *alpha* est connue, la maximisation de $p(\bar{t}|\alpha, \beta)$ en fonction de β permettra d'obtenir l'autre hyperparamètre β. Cette fois, la dérivée sera donc faite par rapport à β :

$$\nabla_\beta \ln |\bar{\bar{S}}_N^{-1}| = \nabla_\beta \sum_i \ln(\lambda_i + \alpha) = \frac{1}{\beta} \sum_i \frac{\lambda_i}{\lambda_i + \alpha} = \frac{\gamma}{\beta} \quad \text{sachant que} \quad d\lambda_i/d\beta = \lambda_i/\beta \tag{B.25}$$

En utilisant les équations (Eq. B.3), (Eq. B.18) et (Eq. B.25), les valeurs de la précision sur les données β sont obtenues par :

$$\frac{1}{\beta} = \frac{1}{N - \gamma} \sum_{i=1}^{N} \left\{ y_i - \bar{m}_N^{\mathrm{T}} \bar{\phi}(x_i) \right\}^2 \tag{B.26}$$

L'algorithme en question permettant l'apprentissage sur ces hyperparamètres est :

Listing B.1 – Algorithme d'apprentissage sur les hyperparamètres

```
def Hyper(Phi, y, alpha_0=1e-10, beta_0=1e-10, max_iter=2000, rtol=1e-16,
    verbose=True):
# ------------------------------------
# Cette fonction permet de déduire les statistiques a posteriori
# et les valeurs optimales de alpha et beta en maximisant la probabilité marginale
    logarithmique
# Arguments:
# Phi: matrice (N x M)
# y: vecteur de valeurs target (f(x)) (N x 1)
# alpha0: valeur initiale de alpha
# beta0: valeur initiale de beta
# maxiter: nombre Maximum d'iterations'
# rtol: critere de convergence (relatif)
# ------------------------------------
    N, M = Phi.shape
    print(N)
    val_propre_0 = np.linalg.eigvalsh(Phi.T.dot(Phi))
    beta = beta_0
    alpha = alpha_0
    for i in range(max_iter):
        beta_prev = beta
        alpha_prev = alpha
        val_propre = val_propre_0 * beta
        # ------------------------------------
        S_N_inv = alpha * np.eye(Phi.shape[1]) + beta * Phi.T.dot(Phi) #
```

```
            eq.(Eq. B.13)(a)
        S_N = np.linalg.inv(S_N_inv)
        m_N = beta * S_N.dot(Phi.T).dot(t)                                  #
            eq.(Eq. B.13)(b)
        # ----------------------------------
        gamma = np.sum(val_propre / (val_propre + alpha))                   # eq.(Eq. B.24)
        alpha = gamma / np.sum(m_N ** 2)                                    # eq.(Eq. B.23)
        # ----------------------------------
        beta_inv = 1 / (N - gamma) * np.sum((y - Phi.dot(m_N)) ** 2) # eq.(Eq. B.26)
            → α
        beta = 1 / beta_inv                                                 # β
        # ----------------------------------
        # check convergence
        if np.isclose(alpha_prev, alpha, rtol=rtol) and np.isclose(beta_prev, beta,
            rtol=rtol):
            if verbose:
                print(f'Convergence after {i + 1} iterations.')
            return alpha, beta, m_N, S_N
    if verbose:
        print(f'Learning : stopped after {max_iter} iterations.')
    # final hyperparameters after convergence
    return alpha, beta, m_N, S_N
# ----------------------------------
```

B.2.2 Développement pour le cas de classification

Selon Tipping généralement pour le cas de classification, la probabilité *a posteriori* sur tous les paramètres inconnus, en tenant compte des données, est exprimée par $p\left(\bar{w}, \bar{\alpha}, \beta | \bar{y}\right)$. Nous cherchons à trouver les paramètres $\bar{w}, \bar{\alpha}$ et β qui maximisent cette probabilité *a posteriori*. Nous pouvons décomposer le modèle *a posteriori* :

$$p\left(\bar{w}, \bar{\alpha}, \beta | \bar{y}\right) = \underbrace{p\left(\bar{w} | \bar{y}, \bar{\alpha}, \beta\right)}_{T_1} \underbrace{p\left(\bar{\alpha}, \beta | \bar{y}\right)}_{T_2} \tag{B.27}$$

Le premier terme de droite (T_1) a été développé précédemment (Eqs. (Eq. B.12)-(Eq. B.13)). Dans le cas du second terme (T_2), il faut déterminer les hyperparamètres α et β qui maximise la seconde partie avant de trouver $\bar{m}_N$ et $\bar{S}_N$, ce terme se décompose lui même :

$$T_2 : p\left(\bar{\alpha}, \beta | \bar{y}\right) \propto p\left(\bar{y} | \bar{\alpha}, \beta\right) p(\bar{\alpha})\ p\left(\beta\right) \tag{B.28}$$

Supposerons des hyperprioris uniformes et ignorerons donc $p(\alpha)$ et $p(\beta)$. Le problème est alors de **maximiser l'évidence** :

$$p\left(\bar{y} | \bar{\alpha}, \beta\right) \equiv p(\bar{y} | \bar{\alpha}, \beta) = \int \underbrace{p(\bar{y} | \bar{w}, \beta)}_{T_3} \underbrace{p(\bar{w} | \bar{\alpha}) d\bar{w}}_{T_4} \tag{B.29}$$

Le premier terme à droite s'écrit :

$$\begin{aligned} T_3 : p(\bar{y}|\bar{w},\beta) &= \prod_{i=1}^{N} \mathcal{N}\left(\bar{y}, \beta^{-1}\right) \\ &= \left(\frac{2\pi}{\beta}\right)^{-\frac{N}{2}} \exp\left\{-\frac{\beta}{2}\|\bar{y} - \bar{\bar{\Phi}}\bar{w}\|^2\right\} \end{aligned} \tag{B.30}$$

Puis passant au deuxième terme de droite (T_4), où M est de même dimension que $\bar{x}$:

$$\begin{aligned} T_4 : p(\bar{w}|\bar{\alpha}) &= \prod_{i=1}^{M} \mathcal{N}\left(0, \alpha_i^{-1}\right) \\ &= \prod_{i=1}^{M} \left(2\pi\alpha_i^{-1}\right)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\alpha_i \bar{w}^2\right\} \\ &= (2\pi)^{-\frac{M}{2}} \prod_{i=1}^{M} \alpha_i^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\bar{w}^T \bar{A} \bar{w}\right\} \end{aligned} \tag{B.31}$$

avec $\bar{A}$ représente la matrice à diagonale α_i. En substituant les deux équations (Eq. B.30) et (Eq. B.31) comme présenté dans (Eq. B.29), le terme T_2 s'écrit alors :

$$\begin{aligned} p(\bar{y}|\bar{\alpha},\beta) &= \int \left(\frac{2\pi}{\beta}\right)^{-\frac{N}{2}} \exp\left\{-\frac{\beta}{2}\|\bar{y} - \bar{\bar{\Phi}}\bar{w}\|^2\right\} (2\pi)^{-\frac{M}{2}} \prod_{i=1}^{M} \alpha_i^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\bar{w}^T \bar{A} \bar{w}\right\} d\bar{w} \\ &= \left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}} \left(\frac{1}{2\pi}\right)^{\frac{M}{2}} \prod_{i=1}^{M} \alpha_i^{\frac{1}{2}} \int \exp - \left\{\frac{\beta}{2}\|\bar{y} - \bar{\bar{\Phi}}\bar{w}\|^2 + \frac{1}{2}\bar{w}^T \bar{A} \bar{w}\right\} d\bar{w} \end{aligned} \tag{B.32}$$

Procédure d'approximation de la vraisemblance

Afin de maximiser notre vraisemblance marginale logarithmique, nous commençons par prendre les dérivées de (Eq. B.17) par rapport à $\bar{\alpha}_i$ et les mettre à zéro :

$$\begin{aligned} \frac{d}{d\alpha_i} \ln p(\bar{y}|\bar{\alpha},\beta) &= \frac{1}{2\alpha_i} - \frac{1}{2}\Sigma_{ii} - \frac{1}{2}m_i^2 = 0 \\ \alpha_i &= \frac{1 - \alpha_i \Sigma_{ii}}{m_i^2} = \frac{\gamma_i}{m_i^2} \end{aligned} \tag{B.33}$$

avec $\gamma_i = 1 - \alpha_i \Sigma_{ii}$

Il faut maintenant dériver (Eq. B.17) par rapport à β et les mettre à zéro :

$$\frac{d}{d\beta} \ln p(\bar{y}|\bar{\alpha},\beta) = \frac{1}{2}\left(\frac{N}{\beta} - \|\bar{y} - \bar{\bar{\Phi}}\bar{m}\|^2 - \mathrm{Tr}\left[\bar{\Sigma}\bar{\bar{\Phi}}^T\bar{\bar{\Phi}}\right]\right) = 0$$

$$\frac{1}{\beta} = \frac{1}{N - \sum_i \gamma_i}\|\bar{y} - \bar{\bar{\Phi}}\bar{m}\|^2 \tag{B.34}$$

Les valeurs de $\bar{\alpha}_i$ et β qui maximisent la vraisemblance marginale sont ensuite calculées itérativement en fixant α et β à des valeurs initiales, et en utilisant les valeurs de $\bar{m}$ et $\bar{\bar{S}}_N$ trouvés à partir de (Eq. B.13), en les utilisant pour calculer de nouvelles estimations pour $\bar{\alpha}$ et β et en répétant ce processus jusqu'à ce qu'un critère de convergence soit satisfait.

Une fois ces valeurs de $\bar{\alpha}_i$ et β obtenues, Il ne subsiste alors qu'à les utiliser pour évaluer la

distribution prédictive sur $f_\star$ pour une nouvelle entrée $x^\star$:

$$\begin{aligned} p\left(f_\star|\bar{x}_\star,\bar{\alpha},\beta\right) &= \int p(t|\bar{w},\beta)p(\bar{w}\mid,\bar{\alpha},\beta)d\bar{w} \\ &= \mathcal{N}\left(\bar{m}^T\bar{\phi}\left(x^\star\right),\sigma^2\left(x^\star\right)\right) \end{aligned} \tag{B.35}$$

Ce qui signifie que la variable cible $f_\star$ est la moyenne de la distribution ci-dessus (Eq. B.35) $\bar{m}^T\bar{\phi}\left(x^\star\right)$

La confiance dans la prédiction est déterminée par la variance de cette distribution $\sigma^2(x^\star)$ qui est donnée par :

$$\sigma_N^2\left(x^\star\right) = \beta^{-1} + \bar{\phi}\left(x^\star\right)^T\bar{\bar{S}}_N\bar{\phi}\left(x^\star\right) \tag{B.36}$$

B.2.3 Hyperparamètres de la méthode de régression par processus gaussiens

Dans la section (Sec. 4.2.3), nous avons vu que la régression par processus gaussien calcule une moyenne pondérée des observations $m_\star(\bar{x}_\star)$ donnée par l'équation (Eq. 4.44). Les valeurs moyennes prédites $\bar{m}$ aux points d'apprentissage sont données par

$$\bar{m} = \bar{\bar{K}}(\bar{\bar{K}} + \sigma_n^2\bar{\bar{I}})^{-1}\bar{y} \tag{B.37}$$

De la même manière que l'apprentissage des paramètres par la méthode bayésienne, expliquée dans la section (Sec. B.2), l'apprentissage des hyperparamètres des processus gaussiens se fait en utilisant les valeurs propres voir [82].

B.3 Calcul de dérivées des Noyaux des Processus Gaussiens

1. **SE :**

$$\begin{aligned} k_{\mathrm{SE}}(r) &= \exp\left(-\frac{1}{2}\left|\left|r\right|\right|_\ell^2\right) \\ \frac{\partial}{\partial x_i}k_{\mathrm{SE}}(r) &= \frac{\partial}{\partial x_i}\exp\left(-\frac{1}{2}\left|\left|r\right|\right|_\ell^2\right) \\ &= \frac{\partial}{\partial x_i}\left(-\frac{r^2}{2\ell^2}\right)\exp\left(-\frac{r^2}{2\ell^2}\right) \\ &= -\frac{r}{\ell_i^2}\exp\left(-\frac{r^2}{2\ell^2}\right)\frac{\partial}{\partial x_i}r \\ &= -\frac{r}{\ell_i^2}k_{\mathrm{SE}}(r) \end{aligned} \tag{B.38}$$

2. **Matern32 :**

$$
\begin{aligned}
k_{\nu=3/2}(r) =& \left(1+\sqrt{3}\,||r||_\ell\right)\exp\left(-\sqrt{3}\,||r||_\ell\right)\\
\frac{\partial}{\partial x_i}k_{\nu=3/2}(r) =& \frac{\partial}{\partial x_i}[(1+\sqrt{3}r)\exp(-\sqrt{3}r)]\\
=& \sqrt{3}\frac{1}{\ell_i}\exp(-\sqrt{3}r) + (1+\sqrt{3}r)\left(-\sqrt{3}\frac{1}{\ell_i}\right)\exp(-\sqrt{3}r)\\
=& \left(-3r\frac{1}{\ell_i}\right)\exp(-\sqrt{3}r)\\
=& \frac{\left(-3r\frac{1}{\ell_i}\right)}{\left(1+\sqrt{3}r\right)}\left(1+\sqrt{3}r\right)\exp(-\sqrt{3}r)\\
=& \frac{\left(-3r\frac{1}{\ell_i}\right)}{\left(1+\sqrt{3}r\right)}k_{\nu=3/2}(r)
\end{aligned}
\tag{B.39}
$$

3. **Matern52 :**

$$
\begin{aligned}
k_{\nu=5/2}(r) =& \left(1+\sqrt{5}\,||r||_\ell + \frac{5}{3}\,||r||_\ell^2\right)\exp\left(-\sqrt{5}\,||r||_\ell\right)\\
\frac{\partial}{\partial x_i}k_{\nu=5/2}(r) =& \frac{\partial}{\partial x_i}[\left(1+\sqrt{5}r+\frac{5}{3}r^2\right)\exp(-\sqrt{5}r)]\\
=& \left(\sqrt{5}\frac{1}{\ell_i^2}+\frac{5}{3}\frac{2x_i}{\ell_i^2}r\right)\exp(-\sqrt{5}r)\\
&+\left(1+\sqrt{5}r+\frac{5}{3}r\right)\left(-\sqrt{5}\frac{x_i}{\ell_i^2}\right)\exp(-\sqrt{5}r)\\
=& \left(\sqrt{5}\frac{x_i}{\ell_i^2}+\frac{5}{3}\frac{2x_i}{\ell_i^2}r+\left(1+\sqrt{5}r+\frac{5}{3}r\right)\left(-\sqrt{5}\frac{x_i}{\ell_i^2}\right)\right)\exp(-\sqrt{5}r)\\
=& \left(\sqrt{5}\frac{x_i}{\ell_i^2}+\frac{5}{3}\frac{2x_i}{\ell_i^2}r-\sqrt{5}\frac{x_i}{\ell_i^2}-5\frac{x_i}{\ell_i^2}r-\frac{5}{3}\sqrt{5}\frac{x_i}{\ell_i^2}r\right)\exp(-\sqrt{5}r)\\
=& \left(\frac{10}{3}\sqrt{5}\frac{x_i}{\ell_i^2}r-5\frac{x_i}{\ell_i^2}r\right)\exp(-\sqrt{5}r)
\end{aligned}
\tag{B.40}
$$

Bibliographie

[1] A. Agrawal. « A Comprehensive Review on Gas Flow in Microchannels ». In : *International Journal of Micro-Nano Scale Transport* 2.1 (2011), p. 1-40 (cf. p. 19).

[2] M. M. Ahsan, M. A. P. Mahmud, P. K. Saha, K. D. Gupta et Z. Siddique. « Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance ». In : *Technologies* 9.3 (2021) (cf. p. 97).

[3] M. Allen et D. Tildesley. *Computer Simulation of Liquids*. Oxford Science Publications, 1987 (cf. p. 22, 23, 33, 36).

[4] N. S. Altman. « An introduction to kernel and nearest-neighbor nonparametric regression ». In : *The American Statistician* 46.3 (1992) (cf. p. 78).

[5] G. Anciaux. « Simulation multi-échelles des solides par une approche couplée dynamique moléculaire/éléments finisDe la modélisation à la simulation haute performance ». 2007BOR13399. Thèse de doct. 2007, 1 vol. (x-159 p.) (Cf. p. 24, 25).

[6] C. Bishop. *Pattern Recognition and Machine Learning*. Springer, jan. 2006 (cf. p. 10, 75-77, 81, 86-90, 92, 175, 176).

[7] M. K. Borg, D. A. Lockerby et J. M. Reese. « A hybrid molecular–continuum method for unsteady compressible multiscale flows ». In : *Journal of Fluid Mechanics* 768 (2015), p. 388-414 (cf. p. 23-25, 29).

[8] M. K. Borg, D. A. Lockerby et J. M. Reese. « A Multiscale Method for Micro/-Nano Flows of High Aspect Ratio ». In : *J. Comput. Phys.* 233 (2013), p. 400-413 (cf. p. 25, 28, 29).

[9] J. Q. Broughton, F. F. Abraham, N. Bernstein et E. Kaxiras. « Concurrent coupling of length scales : methodology and application ». In : *Physical review B* 60.4 (1999), p. 2391 (cf. p. 23).

[10] M. Bugel. « Couplage entre la dynamique moléculaire et la mécanique des milieux continus ». Thèse de doctorat. Université Bordeaux 1, oct. 2009 (cf. p. 24, 26, 27).

[11] M. Bugel et G. Galliéro. « Thermal conductivity of the Lennard-Jones fluid : An empirical correlation ». In : *Chemical Physics* 352.1 (2008) (cf. p. 47, 59-61, 70, 71).

[12] C. J. Burges. « A Tutorial on Support Vector Machines for Pattern Recognition ». In : *Data Mining and Knowledge Discovery* 2 (1998) (cf. p. 10, 75).

[13] C. Cercignani et M. Lampis. « Kinetic models for gas-surface interactions ». In : *Transport Theory and Statistical Physics* 1 (1971) (cf. p. 48).

[14] C. Cercignani. *Slow rarefied flows : Theory and application to Micro-Electro-Mechanical systems*. T. 41. Springer Science & Business Media, 2006 (cf. p. 20, 61).

[15] C. Cercignani. *Mathematical methods in kinetic theory / Carlo Cercignani*. English. 2nd ed. Plenum Press New York, 1990, xi, 252 p. (Cf. p. 19).

[16] E. Chénier, R. Eymard, R. Herbin et O. Touazi. « Collocated finite volume schemes for the simulation of natural convective flows on unstructured meshes ». In : *International Journal for Numerical Methods in Fluids* 56.11 (2008), p. 2045-2068 (cf. p. 54).

[17] D. Chibouti, B. Trouette et E. Chénier. « Jump in the conduction heat flux at the gas/solid interface in micro-channels ». In : *International Journal of Thermal Sciences* 159 (2021), p. 106599 (cf. p. 20, 22, 42).

[18] S. Colin et L. Baldas. *Gas Flows in Microsystems*. MDPI, oct. 2019 (cf. p. 18-20, 23).

[19] I. A. Cosden et J. R. Lukes. « A Hybrid Atomistic–Continuum Model for Fluid Flow Using LAMMPS and OpenFOAM ». In : *Computer Physics Communications* (1er août 2013) (cf. p. 24, 26, 110).

[20] R. Delgado-Buscalioni et P. V. Coveney. « Continuum-particle Hybrid Coupling for Mass, Momentum and Energy Transfers in Unsteady Flow ». In : *Phys. Rev. E* 67.046704 (2003), p. 1-13 (cf. p. 27).

[21] S. H. P. Dharmapalan, M. Borg, J. E. Sprittles et R. Enright. « Molecular Physics of Jumping Nanodroplets ». In : *Nanoscale* (20 juill. 2020) (cf. p. 31).

[22] S. Y. Docherty, M. K. Borg, D. A. Lockerby et J. M. Reese. « Coupling heterogeneous continuum-particle fields to simulate non-isothermal microscale gas flows ». In : *International Journal of Heat and Mass Transfer* 98 (2016), p. 712-727 (cf. p. 30).

[23] S. Y. Docherty, M. K. Borg, D. A. Lockerby et J. M. Reese. « Multiscale simulation of heat transfer in a rarefied gas ». In : *International Journal of Heat and Fluid Flow* 50 (2014), p. 114-125 (cf. p. 29).

[24] J.-M. Dufour. *Histoire de l'analyse des séries chronologiques*. 2003 (cf. p. 92).

[25] F. Ejeian, S. Azadi, A. Razmjou, Y. Orooji, A. Kottapalli, M. Ebrahimi Warkiani et M. Asadnia. « Design and applications of MEMS flow sensors : A review ». In : *Sensors and Actuators A : Physical* 295 (2019), p. 483-502 (cf. p. 17).

[26] D. J. Evans. « Thermal conductivity of the Lennard-Jones fluid ». In : *Phys. Rev. A* 34 (2 1986), p. 1449-1453 (cf. p. 161).

[27] Z. Fan, L. F. C. Pereira, H.-Q. Wang, J.-C. Zheng, D. Donadio et A. Harju. « Force and heat current formulas for many-body potentials in molecular dynamics simulations with applications to thermal conductivity calculations ». In : *Physical Review B* 92 (9 2015), p. 094301 (cf. p. 40, 59, 164, 166).

[28] R. P. Feynman, R. B. Leighton et M. Sands. *The Feynman lectures on physics. Vol. 1*. Addison-Wesley, 1963 (cf. p. 43, 44).

[29] E. G. Flekkøy, G. Wagner et J. Feder. « Hybrid Model for Combined Particle and Continuum Dynamics ». In : *Europhys Lett.* 52 (2000), p. 271-276 (cf. p. 27).

[30] H. Flyvbjerg et H. G. Petersen. « Error estimates on averages of correlated data ». In : *The Journal of Chemical Physics* 91.1 (1989) (cf. p. 49, 60).

[31] D. Frenkel et B. Smit. « Appendix D - Statistical Errors ». In : *Understanding Molecular Simulation (Second Edition)*. Second Edition. San Diego : Academic Press, 2002 (cf. p. 36, 39, 49, 51).

[32] D. Frenkel et B. Smit. *Understanding Molecular Simulation*. Second Edition. San Diego : Academic Press, 2002 (cf. p. 40, 60).

[33] M. Gad-el-Hak. *The MEMS Handbook*. CRC Press, déc. 2005 (cf. p. 17, 20, 21, 23).

[34] G. Galliéro, C. Boned et A. Baylaucq. « Molecular Dynamics Study of the Lennard-Jones Fluid Viscosity : Application to Real Fluids ». In : *Industrial & Engineering Chemistry Research* 44.17 (2005) (cf. p. 47, 59-61, 70, 71, 161-164).

[35] C. Geuzaine et J.-F. Remacle. « Gmsh : A 3-D finite element mesh generator with built-in pre- and post-processing facilities ». In : *International Journal for Numerical Methods in Engineering* 79 (11 2009), p. 1309-1331 (cf. p. 54).

[36] G. Giupponi, G. D. Fabritiis et P. V. Covney. « A Hybrid Method Coupling Fluctuating Hydrodynamics and Molecular Dynamics for the Simulation of Macromolecules ». In : *J. Chem. Phys.* 126 (2007), p. 154903-1-154903-8 (cf. p. 27).

[37] A. Gojan. *Advanced Modeling of Pellet-Cladding Interaction*. 2016 (cf. p. 8, 18).

[38] S. Gull et J. Skilling. « Maximum entropy and bayesian methods ». In : *J. Skilling* (1989) (cf. p. 90, 176).

[39] N. G. Hadjiconstantinou. « Hybrid Atomistic-Continuum Formulations and the Moving Contact-Line Problem ». In : *J. Comput. Phys.* 154 (1999), p. 245-265 (cf. p. 24, 27).

[40] N. G. Hadjiconstantinou, A. L. Garcia, M. Z. Bazant et G. He. « Statistical Error in Particle Simulations of Hydrodynamic Phenomena ». In : *J. Comput. Phys.* 187 (2003), p. 274-297 (cf. p. 27).

[41] A. E. Hoerl et R. W. Kennard. « Ridge Regression : Biased Estimation for Nonorthogonal Problems ». In : *Technometrics* 12.1 (1970), p. 55-67 (cf. p. 81, 175).

[42] N. Houlfort et F.-A. Laurent. « LES RÉGRESSIONS LINÉAIRES ET LOGISTIQUES ». In : *Méthodes qualitatives, quantitatives et mixtes, 2e édition : Dans la recherche en sciences humaines, sociales et de la santé* (2020) (cf. p. 75).

[43] C. Huang, C. Li, P. Y. Choi, K. Nandakumar et L. W. Kostiuk. « Effect of cut-off distance used in molecular dynamics simulations on fluid properties ». In : *Molecular Simulation* 36.11 (2010), p. 856-864 (cf. p. 46).

[44] P. H. Hünenberger. « Thermostat Algorithms for Molecular Dynamics Simulations ». In : *Advanced Computer Simulation : Approaches for Soft Matter Sciences I*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2005 (cf. p. 48, 51, 53).

[45] P. H. Hünenberger. « Thermostat Algorithms for Molecular Dynamics Simulations ». In : *Advanced Computer Simulation : Approaches for Soft Matter Sciences I*. Sous la dir. de C. Dr. Holm et K. Prof. Dr. Kremer. Berlin, Heidelberg : Springer Berlin Heidelberg, 2005, p. 105-149 (cf. p. 60).

[46] T. Ikeshoji et B. Hafskjold. « Non-equilibrium molecular dynamics calculation of heat conduction in liquid and through liquid-gas interface ». In : *Molecular Physics* 81.2 (1994), p. 251-261 (cf. p. 160).

[47] J. H. Irving et J. G. Kirkwood. « The Statistical Mechanical Theory of Transport Processes. IV. The Equations of Hydrodynamics ». In : *The Journal of Chemical Physics* 18.6 (1950), p. 817-829 (cf. p. 41).

[48] S. G. Kandlikar, S. Colin, Y. Peles, S. Garimella, R. F. Pease, J. J. Brandner et D. B. Tuckerman. « Heat Transfer in Microchannels—2012 Status and Research Needs ». In : *Journal of Heat Transfer* 135.9 (26 juill. 2013) (cf. p. 21, 42).

[49] P. Kapitza. « E10 - THE STUDY OF HEAT TRANSFER IN HELIUM II**J. Phys. U.S.S.R. 4, 181 (1941). » In : *Helium 4*. Sous la dir. de Z. M. GALASIEWICZ. Pergamon, 1971, p. 114-153 (cf. p. 21, 42).

[50] G. E. Karniadakis, A. Beskok et N. Aluru. *Microflows and Nanoflows*. Springer-Verlag, 2005 (cf. p. 17, 20-23).

[51] B. H. Kim. *Molecular dynamics simulations of heat transfer in nanoscale liquid films*. Texas A&M University, 2009 (cf. p. 52).

[52] S. Kohlhoff, P. Gumbsch et H. F. Fischmeister. « Crack propagation in b.c.c. crystals studied with a combined finite-element and atomistic model ». In : *Philosophical Magazine A* 64.4 (1991), p. 851-878 (cf. p. 23-25).

[53] J. Kolafa et I. Nezbeda. « The Lennard-Jones fluid : An accurate analytic and theoretically-based equation of state ». In : *Fluid Phase Equilibria* 100 (1994), p. 1-34 (cf. p. 60, 71, 125).

[54] J. Koplik et J. R. Banavar. « Continuum Deductions from Molecular Hydrodynamics ». In : *Annual Review of Fluid Mechanics* (1995) (cf. p. 23).

[55] U. Lācis, M. Pellegrino, J. Sundin, G. Amberg, S. Zaleski, B. Hess et S. Bagheri. « Nanoscale sheared droplet : volume-of-fluid, phase-field and no-slip molecular dynamics ». In : *Journal of Fluid Mechanics* 940 (2022), A10 (cf. p. 24, 31, 32).

[56] LAMMPS. http://lammps.sandia.gov (cf. p. 37).

[57] D. Levesque et L. Verlet. « Molecular dynamics and time reversibility ». In : *Journal of Statistical Physics* 72.3-4 (1993) (cf. p. 43).

[58] H. Liu, J. Zhang, P. Capobianchi, M. K. Borg, Y. Zhang et D. Wen. « A multiscale volume of fluid method with self-consistent boundary conditions derived from molecular dynamics ». In : *Physics of Fluids* 33.6 (2021), p. 062004 (cf. p. 31, 32).

[59] J. Liu, S. Chen, X. Nie et M. O. Robbins. « A Continuum-Atomistic Multi-Timescale Algorithm for Micro/Nano Flows ». In : *Commun. Comp. Phys.* 4.5 (2008), p. 1279-1291 (cf. p. 27).

[60] D. A. Lockerby, A. Patronis, M. K. Borg et J. M. Reese. « Asynchronous coupling of hybrid models for efficient simulation of multiscale systems ». In : *Journal of Computational Physics* 284 (2015), p. 261-272 (cf. p. 24).

[61] D. J. MacKay. « Bayesian interpolation ». In : *Neural computation* 4.3 (1992) (cf. p. 90, 176).

[62] S. Maruyama. « Molecular dynamics methods in microscale heat transfer ». In : *Advances in Numerical Heat Transfer* 2.6 (2002) (cf. p. 40, 53, 60, 168).

[63] S. H. Maslen. « On Heat Transfer in Slip Flow ». In : *Journal of the Aerospace Sciences* 25.6 (1958) (cf. p. 21, 22, 42, 56-58, 71).

[64] J. C. Maxwell. « IV. On the dynamical theory of gases ». In : *Philosophical Transactions of the Royal Society of London* 157 (1867) (cf. p. 157).

[65] J. C. Maxwell. « VII. On stresses in rarified gases arising from inequalities of temperature ». In : *Philosophical Transactions of the Royal Society of London* 170 (1879), p. 231-256 (cf. p. 21, 42, 48, 57).

[66] K. Meier, A. Laesecke et S. Kabelac. « Transport coefficients of the Lennard-Jones model fluid. I. Viscosity ». In : *The Journal of Chemical Physics* 121.8 (2004), p. 3671-3687 (cf. p. 161, 163).

[67] B. Michel, J. Sercombe et G. Thouvenin. « A new phenomenological criterion for pellet–cladding interaction rupture ». In : *Nuclear Engineering and Design* 238.7 (2008), p. 1612-1628 (cf. p. 8, 18).

[68] K. M. Mohamed et A. A. Mohamad. « A Review of the Development of Hybrid Atomistic–Continuum Methods for Dense Fluids ». In : *Microfluidics and Nanofluidics* (1er mars 2010) (cf. p. 22, 24).

[69] F. Müller-Plathe. « A simple nonequilibrium molecular dynamics method for calculating the thermal conductivity ». In : *The Journal of Chemical Physics* 106.14 (1997), p. 6082-6085 (cf. p. 14, 161-164).

[70] K. P. Murphy. *Machine learning : a probabilistic perspective*. Cambridge, Mass. [u.a.] : MIT Press, 2013 (cf. p. 84, 88).

[71] C. Navier. « Mémoires de l'Académie Royale des Sciences de l'Institut de France ». In : t. 6. 1823. Chap. Mémoire sur les lois du mouvement des fluides (cf. p. 21, 32, 42, 57, 126, 128, 157).

[72] S. Nedea, A. van Steenhoven, A. Markvoort, P. Spijker et D. Giordano. « Gas-surface interactions using accommodation coefficients for a dilute and a dense gas in a micro-or nanochannel : Heat flux predictions using combined molecular dynamics and Monte Carlo techniques ». In : *Physical Review E* 89.5 (2014) (cf. p. 64).

[73] X. Nicolas, E. Chénier et G. Lauriat. « Thermal boundary conditions for convective heat transfer of dilute gases in slip flow regime ». In : *International Journal of Thermal Sciences* 135 (jan. 2019) (cf. p. 20, 21, 42, 57, 58, 71, 164).

[74] X. Nicolas, E. Chénier, C. Tchekiken et G. Lauriat. « Revisited analysis of gas convection and heat transfer in micro channels : Influence of viscous stress power at wall on Nusselt number ». In : *International Journal of Thermal Sciences* 134 (2018) (cf. p. 56-58).

[75] X. B. Nie, S. Y. Chen, W. N. E et M. O. Robbins. « A Continuum and Molecular Dynamics Hybrid Method for Micro- and Nano-Fluid Flow ». In : *J. Fluid Mech.* 500 (2004), p. 55-64 (cf. p. 26, 27, 110).

[76] M. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015 (cf. p. 76, 77).

[77] S. T. O'Connell et P. A. Thompson. « Molecular Dynamics Continuum Hybrid Computations : a Tool for Studying Complex Fuid Flows ». In : *Phys. Rev. E* 52 (1995), R5792-R5795 (cf. p. 27).

[78] S. T. O'Connell et P. A. Thompson. « Molecular Dynamics–Continuum Hybrid Computations : A Tool for Studying Complex Fluid Flows ». In : *Physical Review E* (1er déc. 1995) (cf. p. 24-26, 30, 110).

[79] S. Patankar. *Numerical Heat Transfer and Fluid Flow*. Series in computational methods in mechanics and thermal sciences. Routledge, 1980 (cf. p. 33).

[80] C. S. Qazaz, C. K. Williams et C. M. Bishop. « An upper bound on the Bayesian error bars for generalized linear regression ». In : *Mathematics of Neural Networks*. Springer, 1997 (cf. p. 89, 176).

[81] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. 2^{e} éd. Cambridge University Press, 2004 (cf. p. 22, 23, 33, 36, 38, 43, 44, 59, 60).

[82] C. E. Rasmussen et C. K. I. Williams. *Gaussian processes for machine learning*. Adaptive computation and machine learning. MIT Press, 2006, p. I-XVIII, 1-248 (cf. p. 75, 78, 86, 88, 90, 92, 93, 181).

[83] W. Ren et W. E. « Heterogeneous Multiscale Method for the Modeling of Complex Fluids and Micro-fluidics ». In : *J. Comput. Phys.* 204 (2005), p. 1-26 (cf. p. 24, 25, 27, 28, 31).

[84] K. Ritos, M. K. Borg, D. A. Lockerby, D. Emerson et J. Reese. « Hybrid molecular-continuum simulations of water flow through carbon nanotube membranes of realistic thickness ». In : *Microfluidics and Nanofluidics* 19.5 (2015), p. 997-1010 (cf. p. 24, 26, 29, 110).

[85] R. E. Rudd. « The Atomic Limit of Finite Element Modeling in MEMS : Coupling of Length Scales ». In : *Analog Integrated Circuits and Signal Processing* 29.1 (2001), p. 17-26 (cf. p. 24, 25).

[86] R. E. Rudd et J. Q. Broughton. « Coarse-grained molecular dynamics and the atomic limit of finite elements ». In : *Phys. Rev. B* 58 (10 1998), R5893-R5896 (cf. p. 24, 25).

[87] R. E. Rudd et J. Q. Broughton. « Coarse-grained molecular dynamics : Nonlinear finite elements and finite temperature ». In : *Physical Review B* 72.14 (2005) (cf. p. 24-26).

[88] R. B. Schoch, J. Han et P. Renaud. « Transport phenomena in nanofluidics ». In : *Rev. Mod. Phys.* 80 (3 2008), p. 839-883 (cf. p. 22).

[89] J. E. Sprittles. « Air entrainment in dynamic wetting : Knudsen effects and the influence of ambient air pressure ». In : *Journal of Fluid Mechanics* 769 (2015), p. 444-481 (cf. p. 20).

[90] D. W. Stops. « The mean free path of gas molecules in the transition régime ». In : *Journal of Physics D : Applied Physics* 3.5 (1970), p. 685 (cf. p. 19).

[91] J. Sun, Y.-L. He et W.-Q. Tao. « Molecular Dynamics–Continuum Hybrid Simulation for Condensation of Gas Flow in a Microchannel ». In : *Microfluidics and Nanofluidics* (8 jan. 2009) (cf. p. 24, 48, 51, 53).

[92] W. C. Swope, H. C. Andersen, P. H. Berens et K. R. Wilson. « A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules : Application to small water clusters ». In : *The Journal of Chemical Physics* 76.1 (1982) (cf. p. 37, 44).

[93] P. K. T. Werder J. H. Walther. « Hybrid Atomistic-Continuum Method for the Simulation of Dense Fluid Flows ». In : *J. Comput. Phys.* 205 (2005), p. 373-390 (cf. p. 27).

[94] M. E. Tipping. « Sparse Bayesian learning and the relevance vector machine ». In : *Journal of machine learning research* 1.Jun (2001), p. 211-244 (cf. p. 90, 176).

[95] B. D. Todd et P. J. Daivis. *Nonequilibrium Molecular Dynamics : Theory, Algorithms and Applications*. Cambridge University Press, 2017 (cf. p. 162).

[96] R. Vogelsang, C. Hoheisel et G. Ciccotti. « Thermal conductivity of the Lennard-Jones liquid by molecular dynamics calculations ». In : *The Journal of Chemical Physics* 86.11 (1987) (cf. p. 40, 164, 168).

[97] V. H. Vu. « Hybrid and multi-scale modeling for the simulation of fluid flows and heat transfer in microchannels ». Thèse de doctorat. Université Paris-Est, déc. 2016 (cf. p. 24, 45, 115).

[98] V. H. Vu, B. Trouette, Q.-D. To et E. Chénier. « Hybrid atomistic-continuum multiscale method for fluid flow with density variation in microchannels ». In : *International Journal of Numerical Methods for Heat and Fluid Flow* (fév. 2018) (cf. p. 22-24, 26, 27, 30, 31, 48, 53, 110, 112, 113, 115, 116).

[99] V. H. Vu, B. Trouette, Q.-D. To et E. Chénier. « Multi-scale modelling and hybrid atomistic-continuum simulation of non-isothermal flows in microchannels ». In : *Microfluidics and Nanofluidics* (2016) (cf. p. 26, 30-32, 110, 112, 113, 115, 116).

[100] P. Wirnsberger, D. Frenkel et C. Dellago. « An enhanced version of the heat exchange algorithm with excellent energy conservation properties ». In : *The Journal of Chemical Physics* 143.12 (2015), p. 124104 (cf. p. 160).

[101] J. Xie, M. K. Borg, L. Gibelli, O. Henrich, D. A. Lockerby et J. M. Reese. « Effective mean free path and viscosity of confined gases ». In : *Physics of Fluids* 31.7 (2019) (cf. p. 20).

[102] T. H. Yen, C. Y. Soong et P. Y. Tzeng. « Hybrid Molecular Dynamics-Continuum Simulation for Nano/Mesoscale Channel Flow ». In : *Microfluid Nanofluid* 3 (2007), p. 665-675 (cf. p. 26, 27, 110).

[103] J. Zhang, M. K. Borg et J. M. Reese. « Multiscale simulation of dynamic wetting ». In : *International Journal of Heat and Mass Transfer* 115 (2017), p. 886-896 (cf. p. 23, 24, 26, 31, 32, 110).

[104] Z. Zhang et X. Xu. « Effective boundary conditions for dynamic contact angle hysteresis on chemically inhomogeneous surfaces ». In : *Journal of Fluid Mechanics* 935 (2022), A34 (cf. p. 23, 31).