

**INTRODUCTION D'ORBITALES CORRELÉES DANS LES
MÉTHODES MONTE CARLO QUANTIQUE**

Thomas Bouabça

Laboratoire de Chimie et Physique Quantiques

IRSAMC-CNRS, Université de Toulouse

31062 Toulouse Cedex, France

February 2009

Table des matières

1	Introduction générale	11
2	Présentation des méthodes Monte-Carlo quantique	19
2.1	Bref Historique	19
2.1.1	Mouvement brownien et calcul stochastique	20
2.1.2	Naissance des méthodes de Monte-Carlo	21
2.1.3	Les méthodes de Monte-Carlo appliquées à l'équation de Schrödinger	22
2.2	La méthode "Variational Monte Carlo" (VMC)	23
2.2.1	L'algorithme de Metropolis	25
2.2.2	Déplacement des points	27
2.2.3	La méthode de déplacement basée sur le mouvement Brownien avec dérive	28
2.2.4	La méthode basée sur l'équation de Langevin dans l'es- pace des phases	29
2.2.5	La procédure d'acceptation/rejet	30
2.2.6	Les étapes de l'algorithme de Metropolis	31
2.2.7	Erreurs dans la méthode VMC	32
2.2.8	L'erreur statistique	33
2.2.9	L'erreur systématique	33
2.3	L'équation de Schrödinger comme équation d'évolution d'un processus de diffusion : la méthode Diffusion Monte Carlo (DMC)	34

2.3.1	Principes généraux	35
2.3.2	Simulations numériques	36
2.3.3	Echantillonnage selon l'importance	36
2.3.4	Lien entre VMC et DMC	37
2.3.5	L'erreur statistique	38
2.3.6	L'approximation des nœuds fixés	38
2.3.7	Résultats et limites de la méthode Fixed-Node DMC .	41
2.3.8	Algorithme DMC en pratique	43
2.3.9	DMC dans le cas d'un potentiel nul	45
2.3.10	DMC dans le cas d'un potentiel non nul	46
2.3.11	Résumé d'un pas élémentaire dans l'algorithme DMC .	47
2.3.12	Résumé des étapes principales d'un calcul Diffusion Monte-Carlo	48
3	Optimisation de la fonction d'onde	49
3.1	Echantillonnage corrélé par pondération	52
3.2	Optimisation par la variance	53
3.3	La méthode de Newton	54
3.3.1	Description théorique	54
3.3.2	Implémentation en QMC	55
3.3.3	Les estimateurs	55
3.3.4	Le principe de zéro-variance	56
3.3.5	Schéma de stabilisation de l'algorithme	57
3.3.6	Résumé des étapes d'optimisation pour la méthode de Newton	67
3.4	Méthode linéaire	67
3.4.1	Description théorique	67
3.4.2	Résolution d'un problème aux valeurs propres généralisé	69
3.4.3	Estimateurs	71
3.4.4	Renormalisation	71
3.4.5	Formules utilisées	72
3.4.6	Schéma de stabilisation	73

3.4.7	Résumé des étapes d'optimisation pour l'algorithme linéaire	75
3.4.8	Déroulement pratique d'une optimisation linéaire	76
4	La fonction d'onde d'essai multi-Jastrow	77
4.1	Généralités : Rôle de la fonction d'onde d'essai dans les méthodes Monte Carlo quantique	77
4.1.1	Les conditions de cusp	78
4.2	Fonction d'essai usuelle des simulations Monte Carlo quantique	79
4.2.1	Description d'un déterminant de Slater Ψ_k	80
4.2.2	Le terme de Jastrow	83
4.2.3	Interprétation des coefficients	85
4.3	La nouvelle fonction d'onde proposée : la fonction "multi-Jastrow"	86
4.3.1	Lien avec la fonction d'onde de type "backflow"	88
4.3.2	Quelques remarques importantes sur la forme de la fonction multi-Jastrow	90
5	Calcul des dérivées de la fonction d'onde multi-Jastrow	94
5.1	Calcul des Jastrows et de leurs dérivées	97
5.1.1	Dérivée spatiale simple du Jastrow	99
5.1.2	Dérivée spatiale double du Jastrow	100
5.1.3	Dérivée du Jastrow suivant un paramètre	101
5.1.4	Dérivée du gradient et du Laplacien du Jastrow suivant un paramètre	102
5.1.5	Dérivée doubles	103
5.2	Calcul des déterminants et de leurs dérivées	104
5.2.1	Calcul de Ψ_T	106
5.2.2	Calcul du gradient	107
5.2.3	Dérivées doubles	111
5.2.4	Dérivée partielle de Ψ_T suivant un paramètre c_p du Jastrow	119

5.2.5	Dérivée simple suivant un paramètre du Jastrow de la dérivée simple spatiale de Ψ_T	120
5.2.6	Dérivée du Laplacien de Ψ_T suivant un paramètre du Jastrow	124
5.2.7	Dérivée double suivant les paramètres du Jastrow de la fonction d'onde	130
6	Application de la fonction d'onde multi-Jastrow à quelques systèmes atomiques et moléculaires	134
6.1	Introduction	135
6.2	The multi-Jastrow trial wavefunction	138
6.2.1	The standard form	138
6.2.2	The multiJastrow form	140
6.2.3	Relationship with backflow-type trial wavefunction	143
6.3	Optimization of the trial wavefunction	144
6.4	Evaluating derivatives of the multi-Jastrow trial wavefunction	145
6.4.1	Evaluation of Ψ	146
6.4.2	Evaluation of $\nabla\Psi$	147
6.4.3	Evaluation of $\nabla^2\Psi$	148
6.4.4	Derivative of Ψ with respect to a molecular Jastrow parameter	150
6.4.5	Derivative of $\nabla^2\Psi$ with respect to molecular Jastrow parameter	151
6.5	Results	154
6.5.1	Atoms	154
6.5.2	The HF molecule	156
6.6	Summary and perspectives	160
7	Coût calculatoire de la fonction d'onde multi-Jastrow	176
7.1	La fonction Ψ	177
7.2	Le gradient de Ψ	178
7.2.1	Dans le cas du Jastrow commun	178

7.2.2	Dans le cas du Jastrow orbitalaire	178
7.3	Le Laplacien de Ψ	179
7.3.1	Dans le cas du Jastrow commun	179
7.3.2	Dans le cas du Jastrow orbitalaire	180
7.4	Dérivée simple de Ψ par rapport à un paramètre	182
7.5	Dérivée double de Ψ	182
7.6	Dérivée partielle du Laplacien suivant un paramètre	184
8	Comparaison des temps de calcul pour une fonction d'onde usuelle et une fonction d'onde multi-Jastrow	188
8.1	Atome d'oxygène	189
8.1.1	Simulations VMC	189
8.1.2	Optimisation	193
8.1.3	Optimisation pour un paramètre pour un Jastrow sur deux orbitales	198
8.2	Atome de cuivre	198
8.2.1	Simulation VMC du cuivre pour un Jastrow commun .	199
8.2.2	Simulation VMC du cuivre avec des Jastrows orbitalaires	199
8.2.3	Optimisation pour un Jastrow commun	201
8.2.4	Optimisation pour un Jastrow orbitalaire	202
9	Etude des nœuds des fonctions d'onde CASSCF	205
9.1	Idées principales	205
9.2	Manuscript sur l'acroléine (J. Chem. Phys., sous presse) . . .	209
9.3	Introduction	210
9.4	The acrolein molecule and the $n \rightarrow \pi^*$ transition	215
9.5	Computational details	216
9.6	Results	219
9.6.1	Configuration Interaction	219
9.6.2	Fixed-Node Diffusion Monte Carlo	220
9.7	Summary and Perspectives	224

10 Conclusion générale	234
A Dérivation mathématique de la formule de Feynman-Kac	237
B Existence de la décomposition SVD	239
B.1 Décomposition SVD	239
B.2 De la matrice A à la matrice B bidiagonale : La méthode de Householder	241
B.3 Décomposition de B en SVD par la méthode QR	243
C Usage des comatrices pour le calcul des déterminants	245
D Formules du Jastrow commun et de ses dérivées	247
D.1 Notations	247
D.2 Forme simple	248
D.2.1 Expression du terme J	248
D.2.2 Dérivée spatiale de J suivant une coordonnée	248
D.2.3 Dérivée double spatiale de J suivant une coordonnée	248
D.2.4 Dérivée de J suivant un paramètre c_p	248
D.2.5 Dérivée de $\partial_{x_i} J$ suivant un paramètre c_p	249
D.2.6 Dérivée de $\partial_{x_i}^2 J$ suivant un paramètre c_p	249
D.2.7 Dérivée double de J suivant deux paramètres c_1 et c_2	249
D.3 Forme “générale” pour J	250
D.3.1 Expression de J	250
D.3.2 Dérivée spatiale de J suivant une coordonnée	250
D.3.3 Dérivée double spatiale de J suivant une coordonnée	251
D.4 Dérivées suivant les paramètres	251
D.4.1 Dérivée simple de J suivant un paramètre c_p	251
D.4.2 Dérivée de $\partial_{x_i} J$ suivant un paramètre c_p	252
D.4.3 Dérivée de $\partial_{x_i}^2 J$ suivant un paramètre c_p	253
D.4.4 Dérivée double de J suivant deux paramètres c_1 et c_2	254

E	Formules du Jastrow orbitaire et de ses dérivées	256
E.1	Notations	256
E.2	Forme simple	257
E.2.1	Expression du terme J	257
E.2.2	Dérivée spatiale de J suivant une coordonnée	257
E.2.3	Dérivée double spatiale de J suivant une coordonnée	258
E.2.4	Dérivée de J suivant un paramètre c_p	258
E.2.5	Dérivée de $\partial_{x_i} J$ suivant un paramètre c_p	258
E.2.6	Dérivée simple du Laplacien du jastrow suivant les pa- ramètres	259
E.2.7	Dérivée de J suivant les paramètres c_1 et c_2	259
E.3	Forme “générale”	260
E.3.1	Expression de J	260
E.3.2	Dérivée spatiale de J suivant une coordonnée	261
E.3.3	Dérivée double spatiale de J suivant une coordonnée	261
E.3.4	Dérivée de J suivant un paramètre c_p	262
E.3.5	Dérivée de $\partial_{x_i} J$ suivant un paramètre c_p	263
E.3.6	Dérivée de $\partial_{x_i}^2 J$ suivant un paramètre c_p	265
E.3.7	Dérivée double de J suivant deux paramètres c_1 et c_2	266

Remerciements

Je tiens à remercier chaleureusement Michel Caffarel qui a été un directeur de thèse hors pair. Je souhaite à bon nombre de thésards de tomber sur une personne de cette qualité. Je remercie également énormément Anthony Scemama qui m'a apporté avec gentillesse et patience toute l'aide dont j'avais besoin. Cette thèse n'aurait pas pu voir le jour sans eux.

Merci à Julien Toulouse qui m'a donné les bons conseils pour débloquer mon travail à un moment critique et merci à l'ensemble des membres du jury. Je remercie également le Laboratoire de Chimie et Physique Quantiques pour m'avoir accueilli durant ces 4 années et de m'avoir apporté tout ce dont j'avais besoin : un bureau (très) spacieux, du chauffage, de la lumière...

Je tiens à remercier le LCPQ, le CEA, le CICT et l'IDRIS pour tous les moyens informatiques mis à ma disposition. Pardon à tous ceux qui attendaient dans la queue la fin de mes calculs !

Merci également à David et ses cours d'improvisations qui ont été un exutoire et m'ont permis chaque semaine de décompresser.

Merci enfin infiniment à tous ceux qui dans ma vie professionnelle et personnelle m'ont accompagné, enrichi, aidé, supporté... et parfois causé bien du soucis ! Je ne serais pas là sans eux : mes parents, ma soeur, Françoise, Virginie, Jean-Matthieu, Vinh, Mathias ... et tant d'autres.

Merci à vous.

Chapitre 1

Introduction générale

L'objet de la chimie théorique est d'étudier à l'aide des outils mathématiques et des concepts de la mécanique quantique les propriétés des assemblées moléculaires : de la molécule simple aux molécules les plus complexes de la biochimie, en passant par les agrégats, les solides, etc. Cette activité est au carrefour de nombreuses disciplines : mathématiques, physique, chimie, informatique. L'objectif principal est d'obtenir une connaissance approfondie des systèmes moléculaires : étude de la faisabilité et de la dynamique des réactions chimiques, prédiction des géométries moléculaires (par exemple, repliement des protéines), appariement des électrons, etc. A la base de cette connaissance est la résolution de l'équation fondamentale régissant les phénomènes quantiques, à savoir l'équation de Schrödinger. Cette équation est une équation aux dérivées partielles impliquant $3N$ variables indépendantes (les positions des N électrons) dont il est extraordinairement difficile d'obtenir des solutions approchées de qualité pour un système arbitraire. En pratique, des simulations intensives sur ordinateur utilisant des algorithmes issus de plusieurs dizaines d'années de recherche sont nécessaires.

Actuellement, deux grandes classes de méthodes sont utilisées pour cela : les méthodes dites *ab initio* reposant sur le principe variationnel et les méthodes basées sur la théorie de la fonctionnelle de la densité (méthodes DFT). Les méthodes *ab initio* utilisent un développement explicite de la fonction

d'onde sur une base de produits de fonctions élémentaires et permettent d'obtenir avec une grande précision des propriétés pour des systèmes de petite taille (de l'ordre de quelques dizaines d'électrons). En revanche, pour étudier des systèmes de grande taille seule l'approche DFT exprimant l'énergie totale comme une fonctionnelle de la densité électronique est possible. Malheureusement, il est difficile de contrôler la précision obtenue avec cette méthode : la qualité des résultats dépend beaucoup de la fonctionnelle approchée employée et du système considéré.

A l'heure actuelle, il est important de souligner qu'il n'existe pas de méthode systématique permettant de traiter avec une précision contrôlée des systèmes de taille et de nature arbitraires. Cette précision étant indispensable pour faire de la chimie et de la physique *prédictive* sur des systèmes réalistes (chimie du vivant, nanosciences, etc.), mettre au point cette méthode générale reste un des enjeux majeurs de la recherche dans les sciences de la matière.

Depuis une vingtaine d'années, une troisième voie émerge en chimie quantique. Il s'agit des méthodes Monte-Carlo quantique (quantum Monte Carlo, QMC). S'appuyant sur une approche stochastique très différente de celles utilisées dans les deux autres approches, elles offrent une alternative originale de résolution de l'équation de Schrödinger. Les méthodes QMC sont des méthodes qui sont beaucoup utilisées en physique numérique (étude des systèmes de fermions fortement corrélés [1,2], des systèmes de spins[3], des liquides quantiques [4-19], de la matière nucléaire,[21,22], etc.) et qui sont considérées dans ces domaines comme des méthodes tout à fait matures. En chimie quantique, la situation est différente, principalement à cause des caractéristiques très exigeantes du problème à N -corps en chimie (la précision demandée sur les énergies et les différences d'énergie est beaucoup plus grande).

Les processus stochastiques à la base des méthodes QMC sont des processus qui permettent de décrire l'évolution temporelle de phénomènes aléatoires [23]. Ils sont simulés à l'aide d'équations différentielles ordinaires auxquelles

est ajouté un terme de bruit stochastique. Ils servent en traitement du signal, en mathématiques financières pour l'évaluation des prix, et ici en physique quantique. Insistons bien sur le fait que, dans nos travaux, les processus stochastiques sont utilisés uniquement comme un outil mathématique pour résoudre numériquement l'équation de Schrödinger. En particulier, nous ne faisons aucun lien physique entre les processus stochastiques mis en jeu et les phénomènes probabilistes sous-jacents à la mécanique quantique.

Les méthodes QMC possèdent un énorme avantage sur ses deux méthodes concurrentes, à savoir qu'elles sont en parfaite adéquation avec l'évolution des moyens informatiques à venir. C'est un point très important car tout algorithme de chimie théorique est fortement tributaire des caractéristiques des outils informatiques utilisés pour le réaliser. Aujourd'hui, les méthodes Monte Carlo quantique permettent de traiter des systèmes assez gros (jusqu'à 1000 électrons à ce jour [20], et potentiellement plusieurs centaines d'atomes sont envisageables) grâce à un coût calculatoire qui augmente en $O(N^3)$ où N est le nombre d'électrons traités (même "scaling" que les méthodes DFT). Cette augmentation raisonnable du coût est un point très intéressant des méthodes QMC mais leur véritable force repose sur la possibilité de paralléliser massivement les simulations (parallélisation totale sans communication entre les processeurs). Beaucoup soulignent que si les méthodes QMC devaient un jour supplanter les deux autres approches cela serait très probablement dû à cet avantage décisif. Insistons sur le fait que cette possibilité de paralléliser massivement est en phase avec l'évolution du paysage mondial dans le domaine informatique et l'émergence de ce qu'on appelle le "cloud computing" [24-34]. En effet, depuis quelques années les acteurs mondiaux les plus importants de l'informatique : Microsoft, IBM, Sun, Oracle, Google ... construisent des fermes de serveurs capables de mettre à disposition des capacités de calcul phénoménales pour lancer des calculs sur grille. Il se dessine ainsi un nouvel environnement informatique que nous pouvons par analogie, comparer à la distribution d'énergie électrique. Pour effectuer ses calculs, l'utilisateur délègue à un de ces opérateurs externes la gestion des infra-

structures et des problèmes hardware afférents (climatisation, alimentation électrique, etc.) ou software (compilation, optimisation des codes, etc.). Il peut alors se “brancher” sur le réseau et consommer du temps-calcul comme il consommerait de l’électricité. Cette évolution informatique et économique est en parfaite adéquation avec les besoins des méthodes QMC : un nombre important de processeurs pour un temps relativement limité.

Grâce aux développements méthodologiques récents et à la croissance de la puissance de calcul disponible, les méthodes Monte-Carlo quantique ont permis d’obtenir des résultats intéressants dans plusieurs domaines d’étude :

- Systèmes à base de métaux de transition [35-50]
- Molécules du vivant (porphyrine [51,52], hémocyanine [53], etc.)
- Agrégats [54-63]
- ...

Au-delà de la recherche académique, il est également envisageable d’employer les méthodes QMC pour traiter des problèmes à fort impact technologique :

- Electronique moléculaire : nanotubes de carbone [64-72], molécules de polythiophène utilisées comme capteurs, dans la technologie OLED pour les écrans souples [73], etc.
- Energie : piles, cellules solaires organiques...
- Chimie : problème de catalyse ...
- Pharmacie : dans les départements “drug design” pour l’aide dans le criblage et la définition d’un “lead” [74,75]

Les méthodes Monte-Carlo quantique peuvent ainsi offrir un outil d’aide à la décision en fournissant de précieuses informations aux expérimentateurs pour l’étude et la conception de molécules. L’étude théorique par simulation sur ordinateur permet donc de réduire les coûts et les temps d’étude d’une molécule.

Très bien adaptées aux moyens informatiques et possédant un très bon scaling en fonction du nombre d’électrons, les méthodes QMC se posent donc comme des méthodes de choix pour l’avenir. Cependant, plusieurs problèmes subsistent encore pour que les méthodes Monte-Carlo quantique deviennent

vraiment des méthodes standards dans le domaine du calcul électronique.

Un des problèmes les plus importants qui a motivé les travaux présentés dans cette thèse est que les énergies totales issues des calculs Monte-Carlo quantique sont obtenues avec une certaine erreur systématique dite erreur des nœuds fixés (“Fixed-Node” error). Cette erreur est en général très petite : de l’ordre de quelques pourcents de l’énergie de corrélation totale. Cependant, cette erreur systématique peut avoir malgré tout des répercussions importantes sur le calcul des *différences* d’énergie. Or, il est bien connu que ce sont les différences d’énergie qui sont importantes en chimie (affinité électronique, potentiel d’ionisation, énergie de liaison, barrières de potentiel, etc.) et non pas les énergies totales. Il est important de souligner que dans les calculs traditionnels de la chimie quantique (*ab initio* et DFT) ce problème des différences d’énergie existe également. Cependant, bien que les erreurs sur les énergies totales soient plus grandes que dans le cas des approches QMC, les compensations d’erreur sur les différences se font mieux. Ceci est principalement dû au fait que ces algorithmes déterministes permettent plus facilement les compensations d’erreur : utilisation d’une même méthode, d’une même base et d’un même espace fonctionnel assez restreint.

L’objet principal de ce travail de thèse est de présenter deux nouvelles propositions permettant à la fois de réduire l’erreur des nœuds fixés absolue sur des bases physiques mais aussi de mieux contrôler la compensation des erreurs de nœuds fixés dans les différences d’énergie de la chimie. L’objectif final est de pouvoir réaliser de manière standardisée des calculs de qualité sur des systèmes moléculaires de taille arbitraire. Précisons maintenant les éléments essentiels de ces deux propositions.

Les méthodes Monte-Carlo quantique reposent sur le choix d’une fonction d’essai de qualité représentant la fonction d’onde exacte du système traité. L’erreur des nœuds fixés sera d’autant plus petite que cette fonction d’onde approchée sera “proche” de la fonction d’onde exacte inconnue. Actuellement, la forme typique de la fonction d’essai pour les simulations est de la forme : produit d’un préfacteur dit de Jastrow et d’une partie mono-électronique

représentée par un petit nombre de déterminants de Slater construits à partir d'orbitales moléculaires à un électron. La partie déterminantale est assez classique en chimie quantique. Typiquement issue d'un calcul de type Hartree-Fock, elle correspond à la meilleure représentation possible (minimum de l'énergie) pour la fonction d'onde multi-électronique à l'aide des fonctions à un électron (orbitales moléculaires). La partie dite de Jastrow est plus originale. Elle permet de représenter à peu de frais la nature corrélée des électrons. En pratique, un calcul Monte-Carlo débute par le choix d'une fonction d'onde de cette forme. Pour évaluer des différences d'énergie, les deux fonctions d'onde sont optimisées *indépendamment* par rapport aux paramètres du terme de Jastrow et à ceux des orbitales moléculaires, ceci afin de réduire les erreurs de nœuds fixés sur chaque énergie totale. Cette stratégie est raisonnable dans le cas où il y a peu de paramètres. Mais l'optimisation Monte-Carlo d'une fonction d'essai est une procédure longue et complexe. Elle devient de plus en plus difficile à réaliser quand on augmente la taille du système et le nombre de paramètres à optimiser.

Notre approche repose sur une philosophie différente. Notre souci a été en effet de considérer des fonctions d'onde "simples" composées d'éléments pré-optimisés. Pour cela nous avons envisagé deux stratégies complémentaires :

- la première est l'implémentation d'une nouvelle fonction d'onde modulaire permettant une construction "morceaux par morceaux" grâce à des éléments préoptimisés. Cette approche est prévue pour permettre le traitement des gros systèmes.

- la seconde consiste en l'étude des différences d'énergie par l'usage de deux fonctions d'essai "cohérentes" dont on espère que les erreurs de nœuds fixés se compensent au mieux.

Notre nouvelle fonction d'essai cherche à représenter plus physiquement la partie de la corrélation électronique qui a essentiellement un caractère local. Pour cela, nous avons implémenté une fonction d'essai composé d'un déterminant de Slater dont les orbitales portent chacune leur propre facteur

de Jastrow. Le but étant de créer une bibliothèque de Jastrows locaux préoptimisés et *transférables* répondant chacun à des situations physiques précises : liaison simple σ carbone-carbone, orbitale $1s$ d'un atome, paire libre, orbitales d , groupe fonctionnel, etc... En combinant ces Jastrows locaux avec des orbitales strictement localisées [20,76-80], il serait alors envisageable de construire "brique par brique" (une sorte de lego moléculaire) une fonction d'essai beaucoup plus physique et facile à évaluer (car les matrices associées aux différentes quantités à évaluer sont très creuses).

Dans notre deuxième approche, les fonctions d'essai sont composées d'un jeu d'orbitales moléculaires optimisées à partir d'un calcul *ab initio* "State-Averaged". Les calculs Monte-Carlo s'effectuent sans réoptimisation des orbitales. Naturellement, ces deux approches sont complémentaires. Lors de l'étude des différences d'énergie, nous pouvons utiliser des orbitales moléculaires optimisées dans un schéma *ab initio* puis rajouter sur les groupes fonctionnels impliqués dans cette différence d'énergie un Jastrow local. De cette manière, nous décrivons de manière spécifique les corrélations locales tout en gardant des structures proches pour le reste de la molécule afin de compenser au mieux les différences parasites dues aux erreurs dans le calcul de cette partie.

Le plan que nous avons choisi pour notre présentation est le suivant :

- Dans le chapitre 2 nous exposerons le principe des méthodes Monte-Carlo quantique en mettant en valeur le rôle de la fonction d'essai.
- Dans le chapitre 3, nous présenterons les méthodes d'optimisation que nous avons employées.
- Nous présenterons dans le chapitre 4 notre nouvelle fonction d'onde multi-Jastrow.
- Nous décrivons le calcul de ses diverses dérivées dans le chapitre 5.
- Le chapitre 6 est constitué d'un article présentant cette fonction multi-Jastrow et les premiers résultats obtenus pour quelques atomes et la molécule HF.

- Dans le chapitre 7, nous discuterons les performances théoriques de notre algorithme en terme de temps-calcul.
- Cette discussion sera complétée dans le chapitre 8 par la présentation de calculs qui illustrent les formules théoriques.
- Dans le chapitre 9, nous aborderons la question d'un traitement rationnel des différences d'énergies à l'aide de fonctions d'onde *ab initio* de type CASSCF. Des résultats détaillés seront donnés pour la transition $n \rightarrow \pi^*$ de la molécule acroléine.
- Enfin, le chapitre 10 présentera nos conclusions générales.

Chapitre 2

Présentation des méthodes Monte-Carlo quantique

Cette partie a pour but de présenter les méthodes de Monte-Carlo quantique (ou quantum Monte Carlo, QMC) utilisées en chimie théorique. Les méthodes QMC regroupent sous cette dénomination commune un vaste nombre de variantes aux acronymes différents (GFMC, SRMC, PDMC, etc.). Nous présenterons ici les deux méthodes principales, qui sont de loin les plus utilisées et que nous avons utilisé ici, à savoir les méthodes dites “Variational Monte Carlo” (VMC) et “Diffusion Monte Carlo” (DMC). Si ces deux méthodes reposent toutes deux sur des processus stochastiques, elles restent néanmoins différentes dans leur esprit. A travers la présentation de ces méthodes, nous sensibiliserons tout particulièrement le lecteur au rôle important joué par la fonction d’onde d’essai, Ψ_T . Une présentation complète des méthodes Monte-Carlo quantique se trouve, par exemple, dans les références [81,82,83].

2.1 Bref Historique

Nous présentons ici quelques repères historiques sur les méthodes Monte-Carlo quantique. Dans notre première partie, nous proposons une histoire

non exhaustive du mouvement brownien et sa description qui ont mené à la création du calcul stochastique, outil mathématique indispensable à la base de toute méthode Monte-Carlo. Nous raconterons ensuite l'origine des méthodes de Monte-Carlo et de l'algorithme de Metropolis, Enfin, nous montrerons comment les méthodes de Monte-Carlo classiques ont été appliquées à l'équation de Schrödinger.

2.1.1 Mouvement brownien et calcul stochastique

C'est au XIX^{ème} siècle que le botaniste Robert Brown introduit la notion de mouvement brownien pour décrire le mouvement aléatoire de grains de pollen en suspension dans l'eau [84]. Louis le Bachelier sera le premier à quantifier le concept de mouvement brownien dans sa thèse [85] en 1900. Cependant, il utilise le mouvement brownien pour le domaine de la finance et ne le connecte pas avec les phénomènes physiques de propagation des particules. Ce problème sera redécouvert quelques années plus tard par Albert Einstein. Il résoudra le problème en 1905 [86] en décrivant le mouvement brownien comme un processus qui sera appelé plus tard "processus de Wiener". En 1908, Paul Langevin proposera une autre moyen d'appréhender le mouvement brownien [106] en le considérant du point de vue des trajectoires des particules. Il applique les équations de mouvement de Newton en considérant que la particule de masse m est soumise à deux forces : une force visqueuse $f = -kv$ et un terme de bruit blanc gaussien $\eta(t)$. Il établit ainsi l'équation dite de Langevin, point de départ de la théorie de la diffusion :

$$m \frac{dv(t)}{dt} = -kv(t) + \eta(t) \quad (2.1)$$

Les points de vue d'Einstein et de Langevin seront réunis par Adriaan Fokker [88] qui dans sa thèse de 1913 décrira l'évolution d'une densité de probabilité p . Il établira l'équation de Fokker-Planck qui généralise les deux processus précédents. Par exemple, pour un densité $p(x, t)$ unidimensionnelle, cette équation est de la forme :

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}[D_1(x)f(x, t)] + \frac{\partial^2}{\partial x^2}[D_2(x)f(x, t)] \quad (2.2)$$

D_1 est le vecteur de dérive (drift vector) et D_2 le tenseur de diffusion. Cette équation sera redécouverte par Andreï Kolmogorov qui établit la théorie axiomatique des probabilités en 1933. Elle est ainsi également connue sous le nom d'équation ("backward") de Kolmogorov. La théorie des probabilités sera étendue par Kiyoshi Itô. En 1942, Itô publie sa théorie de l'intégrale stochastique qui permet d'étudier les phénomènes aléatoires à travers le temps. Cette théorie permet d'intégrer le mouvement brownien et de calculer directement des trajectoires stochastiques. Mark Kac, en 1949 établira la formule de Feynman-Kac[89] qui permet de faire le lien entre la résolution d'une équation de Fokker-Planck et la résolution d'une intégrale stochastique (voir dérivation de cette formule dans l'appendice A).

2.1.2 Naissance des méthodes de Monte-Carlo

Nous devons, au comte de Buffon au XVIII^{ème} siècle, la première résolution d'un problème par une approche probabiliste. Il s'agit du problème de l'aiguille de Buffon [90]. En jettant des aiguilles sur le parquet et en comptant le nombre d'aiguilles intersectant avec les rainures de celui-ci, il est possible d'obtenir une estimation du nombre π .

Mais ce n'est qu'à la fin de la seconde guerre mondiale que naîtront les méthodes de Monte-Carlo. Ces méthodes ont été développées au laboratoire de Los Alamos dans le cadre du projet Manhattan visant à mettre au point la bombe atomique. Il était en effet nécessaire de trouver des méthodes numériques permettant de résoudre l'équation de Boltzmann gouvernant la fonction de distribution des neutrons dans du matériel fissible. L'idée originale vient du mathématicien Stanislaw Ulam, grand amateur de jeux de hasard qui introduisit cette méthode en 1946 pour calculer la probabilité d'obtenir une main gagnante au jeu de carte du "solitaire". Il convainquit en 1947, John von Neumann de l'intérêt de cette méthode pour résoudre numériquement les problèmes de diffusion des neutrons ou d'autres problèmes de physique statistique. Il travaille ainsi avec Enrico Fermi, Robert D. Richtmyer et Nicholas Metropolis sur ces idées. En 1949, Ulam et Metropolis signent le premier

article[91] décrivant le principe des méthodes de Monte-Carlo. C'est Metropolis qui nommera cette classe de méthodes de son nom, en référence au casino de la ville de Monte-Carlo. Metropolis voulait ainsi rendre hommage à la passion de Ulam pour le poker[92]. En 1953, Metropolis, Rosenbluth, Teller et Teller proposent leur fameux algorithme [93] (dit maintenant "de Metropolis", ou MRT² si on veut être plus équitable) qui est à la base des simulations de Monte-Carlo classique. L'algorithme de Metropolis sera généralisé en 1970 par Hasting[94] qui permettra l'utilisation de probabilités de transition non symétriques. C'est cet algorithme, communément appelé algorithme de Metropolis-Hasting que nous décrirons dans la partie de cet exposé consacré à la méthode Variational Monte Carlo (VMC), section 2.2.

2.1.3 Les méthodes de Monte-Carlo appliquées à l'équation de Schrödinger

C'est dans un deuxième temps que s'est développée l'utilisation des méthodes Monte Carlo pour résoudre l'équation de Schrödinger, on parle alors de méthodes Monte Carlo *quantique*. Dès 1931, Schrödinger soulignait la possibilité d'associer son équation en temps imaginaire à un processus de diffusion. Ces méthodes sont basées sur l'idée originale de Schrödinger, et par la suite d'autres chercheurs, soulignant le lien formel entre mécanique quantique en temps imaginaire et processus de diffusion. La première application des méthodes de Monte-Carlo quantique pour une molécule a eu lieu en 1975 avec l'étude de He_3^+ par Anderson[95]. En 1982, Ceperley, Reynolds, Alder et Lester [96] présentent la méthode Diffusion Monte-Carlo dans l'approximation des nœuds fixés (Fixed-Node DMC) que nous présenterons plus loin et qui est l'algorithme standard utilisé de nos jours. A partir de ce moment, les méthodes Monte-Carlo quantique pour le calcul des structures électroniques prendront leur essor, permettant des applications variées (par exemple, [97-103]).

2.2 La méthode “Variational Monte Carlo” (VMC)

La méthode variationnelle de Monte-Carlo (VMC) repose sur le principe du calcul d’une intégrale multidimensionnelle par échantillonnage de la densité de probabilité. Considérons en effet l’intégrale

$$\int dx f(x) \Pi(x) \quad (2.3)$$

où $\Pi(x)dx$ est une densité de probabilité, c’est à dire que $\Pi(x) \geq 0$ et $\int \Pi(x)dx = 1$. On peut approcher $\int dx f(x) \Pi(x)$ par :

$$\int dx f(x) \Pi(x) \approx \frac{1}{N} \sum_{i=1, N} f(x_i) \quad (2.4)$$

où x_i est un ensemble fini de N points répartis selon la densité de probabilité $\Pi(x)$. En VMC, une fonction réelle Ψ_T , appelée fonction d’onde d’essai, est proposée et définit la densité de probabilité suivante (probabilité de présence de la mécanique quantique associée à une fonction d’onde)

$$\Pi(R) = \frac{\Psi_T^2(R)}{\int \Psi_T^2 dR} \quad (2.5)$$

à partir de laquelle différentes quantités seront moyennées. R représente ici l’ensemble des coordonnées électroniques, c’est à dire

$$R = (\vec{r}_1, \dots, \vec{r}_{N_e})$$

où N_e est le nombre d’électrons. Remarquons que, contrairement aux approches traditionnelles de la chimie quantique, la fonction d’onde ne dépend pas ici des coordonnées de spin. Ce n’est toutefois pas un problème puisqu’il est bien connu que pour des systèmes non relativistes il est toujours possible d’exprimer les contraintes résultant du principe de Pauli dans un formalisme sans spin (voir, par exemple, l’article de Matsen "Spin-Free quantum chemistry [104]). En pratique, la symétrie de spin est réalisée en imposant que la fonction d’onde soit antisymétrique par rapport à l’échange des paires d’électrons de même spin (antisymétrie par rapport au sous-ensemble des N_α

électrons et par rapport au sous-ensemble des N_β électrons). Ψ_T est choisie pour être une bonne approximation de la fonction d'onde Ψ_0 inconnue de l'état fondamental. L'un des points importants de la méthode VMC est de fournir une approximation variationnelle E_{VMC} de l'énergie E_0 de l'état fondamental.

$$E_0 \equiv \frac{\hat{H}\Psi_0}{\Psi_0} \quad (2.6)$$

où \hat{H} est l'Hamiltonien du système considéré.

Nous verrons par la suite que la méthode VMC sera utilisée suivant le même principe pour obtenir d'autres quantités.

Exprimons maintenant E_{VMC} comme valeur moyenne stochastique. Pour cela partons de l'expression de l'énergie variationnelle

$$E_{VMC} = \int \frac{dR \Psi_T^* \hat{H} \Psi_T}{\int dR \Psi_T \Psi_T^*}. \quad (2.7)$$

Introduisons la quantité

$$E_L = \frac{\hat{H}\Psi_T}{\Psi_T}. \quad (2.8)$$

E_L est appelée "énergie locale" et joue un rôle central dans les méthodes Monte-Carlo quantique. C'est une quantité homogène à une énergie et qui dépend de \vec{R} . Notons que plus la fonction d'onde d'essai est proche de la fonction exacte et moins l'énergie locale varie en fonction de \vec{R} . E_{VMC} se réécrit alors :

$$E_{VMC} = \int E_L \frac{\Psi_T \Psi_T^*}{\int \Psi_T \Psi_T^*} = \langle E_L \rangle_{|\Psi_T|^2} \quad (2.9)$$

E_{VMC} est donc la valeur moyenne de la quantité E_L suivant la densité de probabilité $\frac{\Psi_T^2}{\int \Psi_T^2 dR}$ (Ψ_T étant réelle, $\Psi_T = \Psi_T^*$)

Ainsi comme précédemment E_{VMC} peut s'approcher par :

$$E_{VMC} \approx \frac{1}{N} \sum_{i=1}^N E_L(x_i) \quad (2.10)$$

où x_i est un ensemble fini de N points répartis selon la densité de probabilité normalisée associée à Ψ_T^2 . E_L est donnée par l'expression :

$$E_L = V(R) - \frac{1}{2} \frac{\nabla^2 \Psi_T}{\Psi_T} \quad (2.11)$$

avec $V(R)$ est l'énergie potentielle

Nous allons à présent voir comment réaliser un échantillonnage des points x_i grâce à l'algorithme de Metropolis.

2.2.1 L'algorithme de Metropolis

L'algorithme de Metropolis permet de simuler une densité de probabilité $\Pi(R) = \frac{\Psi_T^2}{\int dR \Psi_T^2}$. Concrètement, on applique itérativement un ensemble de règles de déplacements et d'acceptation/réjection de ces déplacements sur une population de copies du système (appelés "marcheurs" ou "walkers"). Dans la limite d'un grand nombre d'itérations les marcheurs se répartissent en moyenne selon la densité voulue. Dit de manière différente, si l'on se fixe un domaine de l'espace δV et un nombre total de points N , pour M points dans δV on a :

$$\frac{M}{N} \approx \int_{\delta V} \Pi(R) dR \quad (2.12)$$

En termes mathématiques, l'algorithme de Metropolis consiste à engendrer une chaîne de Markov discrète ergodique dont la loi stationnaire est Π . Rappelons qu'une chaîne de Markov discrète est une succession de variables aléatoires X_0, X_1, \dots, X_n dont l'état X_n ne dépend que de l'état X_{n-1} . Si l'on note $P(\cdot)$ la probabilité conditionnelle, on a donc

$$P(X_n = x | X_{n-1}, \dots, X_1, X_0) = P(X_n = x | X_{n-1}). \quad (2.13)$$

La probabilité conditionnelle est également appelée probabilité de transition. C'est elle qui va caractériser notre chaîne de Markov. Nous conserverons ce terme de probabilité de transition dans la suite de l'exposé et nous la noterons sous une forme plus physique $p(X \rightarrow Y)$ avec le sens suivant :

$$p(X \rightarrow Y) = p(Y|X) \quad (2.14)$$

Dans notre cas, la variable X_i aussi appelé “marcheur” est un vecteur constitué de l’ensemble des positions des électrons du système considéré. S’il y a N_e électrons, X_i sera donc un vecteur de dimension $3N_e$ (chaque électron ayant trois coordonnées spatiales). Partant d’une distribution initiale (aléatoire) des électrons en X_0 , l’application de l’algorithme de Metropolis va engendrer lorsque n tend vers $+\infty$ des marcheurs $X_n, X_{n+1} \dots$ dont la position des électrons est répartie suivant la densité Π .

Pour engendrer cette chaîne, l’algorithme de Metropolis va s’appuyer sur deux éléments :

- des règles de déplacement des marcheurs auxquelles est associée une probabilité de transition
- une procédure d’acceptation/rejet

Les marcheurs vont explorer l’espace en suivant la probabilité de transition et la procédure de rejet/acceptation va valider ou non ce choix, créant ainsi un nouvel élément de la chaîne. Il est important de souligner que pour que l’algorithme fonctionne, une condition nécessaire est que la condition d’ergodicité soit vérifiée. L’ergodicité signifie que les marcheurs partant de n’importe quel endroit vont pouvoir parcourir n’importe quelle zone de l’espace avec une probabilité non nulle sans être bloqué quelque part.

En pratique, on réalisera plusieurs chaînes de Markov indépendantes et on calculera les moyennes sur l’ensemble de ces trajectoires. La condition d’ergodicité va imposer l’égalité entre les moyennes spatiales et temporelles. Cela signifie que si l’on simule k chaînes de Markov indépendantes :

$$\begin{array}{l}
 X_0^1, \quad X_1^1, \dots \quad X_n^1, \dots \\
 \vdots \qquad \qquad \qquad \vdots \\
 X_0^k, \quad X_1^k, \dots \quad X_n^k, \dots \} : \text{moyenne temporelle le long d'une trajectoire} \\
 \underbrace{\hspace{10em}} \\
 \text{moyenne spatiale à un instant donné}
 \end{array} \tag{2.15}$$

On a :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n F(X_n^k) = \lim_{k \rightarrow +\infty} \frac{1}{k} \sum_{i=1}^k F(X_n^k) \tag{2.16}$$

pour n assez grand et où F est une fonction quelconque de la configuration. La première somme s'effectue le long d'une trajectoire. L'indice supérieur k décrivant le numéro de la trajectoire est fixé. La limite s'effectue en faisant tendre n vers $+\infty$. La seconde somme s'effectue dans l'espace. L'indice inférieur n décrivant le rang dans la chaîne de Markov est fixé. La somme porte sur l'indice supérieur qui décrit le nombre de trajectoires indépendantes. La limite s'effectue en faisant tendre ce nombre k vers $+\infty$.

La moyenne le long d'une trajectoire est identique à la moyenne de tous les marcheurs des k trajectoires indépendantes à un moment donné suffisamment grand.

Ainsi, lorsque nous moyennons une quantité \widehat{O} , la moyenne porte sur la moyenne temporelle et spatiale :

$$\langle \widehat{O} \rangle \approx \frac{1}{K_{space}} \frac{1}{N_{time}} \sum_{k=1}^{K_{space}} \sum_{n=1}^{N_{time}} \widehat{O}(X_n^k) \quad (2.17)$$

Voyons à présent les deux éléments constitutifs de l'algorithme de Metropolis. Nous commençons par décrire les diverses méthodes de déplacement des marcheurs puis la procédure d'acceptation/rejet.

2.2.2 Déplacement des points

Le déplacement des points peut se faire suivant plusieurs méthodes. Citons les deux méthodes de déplacement employées durant ce travail :

- la méthode basée sur le mouvement Brownien avec dérive[105]
- la méthode basée sur l'équation de Langevin dans l'espace des phases (R, P) [106]

Dans l'article original de Metropolis[93] et dans beaucoup d'applications publiées, le déplacement des points se fait suivant une marche aléatoire libre, c'est à dire avec la règle de déplacement suivante :

$$R^{n+1} = R^n + \tau \eta \quad (2.18)$$

où η est une variable aléatoire de loi uniforme sur l'intervalle $[-\frac{1}{2}, \frac{1}{2}]$. La probabilité de transition associée est donnée par :

$$p(R^n \rightarrow R^{n+1}) = \frac{1}{(2\tau)^{3N_e}} \chi\left(\frac{R^n - R^{n+1}}{\tau}\right) \quad (2.19)$$

où χ est la fonction caractéristique de l'hypercube centré en 0 et de côté τ , N_e étant le nombre d'électrons dans l'espace ordinaire tridimensionnel.

Il est clair que ce choix de probabilité de transition pour visiter l'espace de configuration n'est pas optimal puisque les marcheurs parcourent indifféremment n'importe quelle région de l'espace sans tenir compte des probabilités associées.

2.2.3 La méthode de déplacement basée sur le mouvement Brownien avec dérive

Afin de ne pas parcourir à l'aveugle l'espace de configuration et donc d'échantillonner les régions pertinentes à forte probabilité, il a été proposé [105] une autre méthode qu'on appelle souvent marche aléatoire biaisée. C'est la méthode utilisée habituellement dans les simulations VMC. La dynamique des points est donnée par :

$$dR(t) = \frac{\nabla\Psi}{\Psi}[R(t)]dt + dW(t) \quad (2.20)$$

où $W(t)$ est un processus de Wiener. La discrétisation temporelle donne alors les règles de déplacement suivantes.

$$R^{n+1} = R^n + \frac{\nabla\Psi}{\Psi}[R^n]\tau + \sqrt{\tau}\eta \quad (2.21)$$

où η est un vecteur aléatoire gaussien dont toutes les composantes sont indépendantes. Plus précisément

$$\langle \eta_i \rangle = 0$$

et

$$\langle \eta_i \eta_j \rangle = \delta_{ij}$$

La probabilité de transition du processus est alors :

$$p(R^n \rightarrow R^{n+1}) = \left[\frac{1}{\sqrt{\pi\tau}} \right]^{3N_e} \exp\left[-\frac{(R^{n+1} - R^n - R^n\tau)^2}{2\tau}\right] \quad (2.22)$$

où $b \equiv \frac{\nabla\Psi_T(R^n)}{\Psi_T(R^n)}$ est appelé vecteur de dérive (ou drift) et τ est le pas de temps. La valeur de ce pas de temps dans l'espace de configuration est fixée avant la simulation.

Si τ est trop petit, les déplacements seront très petits. Il faudra faire beaucoup de pas Monte-Carlo pour être sûr d'avoir exploré tout l'espace et ainsi obtenir des moyennes convergées.

Si τ est trop grand, les déplacements des points seront plus importants et la probabilité de tomber dans une région physiquement inadmissible sera beaucoup plus importante. L'événement sera systématiquement refusé et le taux d'acceptation dans la procédure d'acceptation/rejet à suivre sera plus faible.

Un critère pratique est de fixer τ pour obtenir un taux d'acceptation ni trop grand ni trop petit, c'est à dire typiquement dans l'intervalle (40%,70%).

L'introduction du terme de dérive va accélérer la convergence. En effet, le gradient va favoriser les déplacements des régions à faible densité de probabilité vers les régions à plus haute densité de probabilité.

2.2.4 La méthode basée sur l'équation de Langevin dans l'espace des phases

Cependant, le déplacement des points par la méthode de la marche biaisée souffre d'une limite. Il n'arrive pas à décrire avec précision les densités près des surfaces nodales (i.e les zones où la densité tend vers zéro). En effet, lorsque Ψ_T tend vers 0, le drift $\frac{\nabla\Psi_T}{\Psi_T}$ va tendre vers $+\infty$ et crée un déplacement trop brutal. La méthode de Langevin va corriger cet aspect là en faisant intervenir le drift non pas directement dans le déplacement des points, mais indirectement via le moment. La dynamique de Langevin est définie dans l'espace des configurations de la manière suivante :

$$dR(t) = \frac{1}{m}P(t)dt \quad (2.23)$$

$$dP(t) = -\nabla V(R(t)) - \gamma P(t)dt + \sigma dW(t) \quad (2.24)$$

qui se discrétise en

$$\begin{cases} R^{n+1} = R^n + \frac{\tau}{m} P^n e^{-\frac{\gamma\tau}{2}} - (\frac{\tau^2}{2m}) \nabla V(R^n) e^{-\frac{\gamma\tau}{4}} + \eta_R^n \\ P^{n+1} = P^n e^{-\gamma\tau} - (\frac{\tau}{2}) [\nabla V(R^n) + \nabla V(R^{n+1})] e^{-\frac{\gamma\tau}{2}} + \eta_P^n \end{cases} \quad (2.25)$$

Ici, η_R et η_P sont des vecteurs gaussiens corrélés. La probabilité de transition approchée est alors :

$$P[(R^n, P^n) \rightarrow (R^{n+1}, P^{n+1})] = Z^{-1} \exp\left[-\frac{1}{2(1-c_{12}^2)} \left(\left(\frac{|d_1|}{\sigma_1}\right)^2 + \left(\frac{|d_2|}{\sigma_2}\right)^2 - 2c_{12} \frac{d_1}{\sigma_1} \cdot \frac{d_2}{\sigma_2} \right)\right] \quad (2.26)$$

avec

$$d_1 = R^{n+1} - R^n - \tau \frac{P^n}{m} e^{-\frac{\gamma\tau}{2}} + \frac{\tau^2}{2m} \nabla V(R^n) e^{-\frac{\gamma\tau}{4}}$$

et

$$d_2 = P^{n+1} - P^n e^{-\gamma\tau} + \frac{1}{2} \tau [\nabla V(R^n) + \nabla V(R^{n+1})] e^{-\frac{\gamma\tau}{2}}$$

Afin de diminuer le taux de rejet, on effectue une opération supplémentaire : l'inversion du moment. Le nouveau point (R_*^{n+1}, P_*^{n+1}) est obtenu en prenant l'opposé du moment P :

$$(R_*^{n+1}, P_*^{n+1}) = (R^{n+1}, -P^{n+1}) \quad (2.27)$$

τ est fixé pour obtenir un taux d'acceptation de l'ordre de 90%.

2.2.5 La procédure d'acceptation/rejet

Après avoir proposé un nouveau point, la procédure d'acceptation/rejet va sanctionner ce choix. La règle est la suivante. Pour une nouvelle position R^{n+1} à partir d'une ancienne position R^n on pose :

$$A(R^n \rightarrow R^{n+1}) = \text{Min}\left[1, \frac{p(R^{n+1} \rightarrow R^n) \Pi(R^{n+1})}{p(R^n \rightarrow R^{n+1}) \Pi(R^n)}\right] \quad (2.28)$$

On tire alors un nombre aléatoire uniforme η entre 0 et 1.

Si

$$A(R^n \rightarrow R^{n+1}) \geq \eta \quad (2.29)$$

le mouvement est accepté et le nouvel élément dans la chaîne est R^{n+1} . Sinon, le mouvement est rejeté et le nouvel élément de la chaîne de Markov est R^n . Insistons sur le fait que, après un rejet, l'ancien point est considéré comme le nouveau point et est compté une fois de plus dans les moyennes. Ainsi, après k rejets (on fait du “sur place” pendant k tirages) les moyennes contiendront $k+1$ fois la valeur des observables en ce point.

Pour bien comprendre comment fonctionne la procédure, plaçons nous dans le cas simple où la probabilité de transition est symétrique.

$$p(R^n \rightarrow R^{n+1}) = p(R^{n+1} \rightarrow R^n) \quad (2.30)$$

Dans ce cas là,

$$A(R^n \rightarrow R^{n+1}) = \text{Min}\left[1, \frac{\Pi(R^{n+1})}{\Pi(R^n)}\right] \quad (2.31)$$

Lorsque $\Pi(R^{n+1})$ est plus grand que $\Pi(R^n)$, $A(R^n \rightarrow R^{n+1})$ vaut 1. Le nouveau point est donc accepté avec certitude. C'est le cas où les marcheurs se déplacent des zones de basse probabilité vers les zones de haute probabilité. Dans le cas où $\Pi(R^{n+1})$ est plus petit que $\Pi(R^n)$, le nouveau point n'est pas systématiquement refusé, mais l'est d'autant plus souvent que le rapport des deux probabilités est petit. Maintenant, lorsque nous utilisons une probabilité de transition non symétrique comme dans le cas de la marche aléatoire biaisée ou la méthode de Langevin, cette idée générale subsiste, c'est à dire que la procédure d'acceptation/rejet va favoriser les déplacements des zones de faible densité vers les zones de forte densité.

2.2.6 Les étapes de l'algorithme de Metropolis

Nous résumons ici les étapes de l'algorithme de Metropolis

- **Etape 1**

Une règle de déplacement des points est fixée. Mathématiquement, cela revient à choisir une probabilité de transition $p(R^n \rightarrow R^{n+1})$.

- **Etape 2**

On se donne une répartition de points initiale R^0 répartis aléatoirement dans l'espace. A la convergence, les résultats ne dépendent pas de ce choix initial.

- **Etape 3**

Pour chaque marcheur R , on cherche un nouveau point pour la marche. On propose pour cela un nouveau point R^{n+1} suivant la probabilité de transition $p(R^n \rightarrow R^{n+1})$.

- **Etape 4**

La nouvelle position R^{n+1} est acceptée avec une probabilité

$$A(R^n \rightarrow R^{n+1}) = \text{Min}\left[1, \frac{p(R^{n+1} \rightarrow R^n)\Pi(R^{n+1})}{p(R^n \rightarrow R^{n+1})\Pi(R^n)}\right] \quad (2.32)$$

Si le point R^{n+1} est accepté, R^{n+1} devient le nouveau point, sinon R^n devient le point suivant dans la marche.

- **Etape 5**

On recommence les étapes 3 et 4 autant de fois qu'on le désire. Plus on itère, plus les erreurs statistiques seront diminuées. Toutefois la contrepartie de cette précision accrue est un temps de calcul plus important.

Notons au passage que pour simuler notre densité de probabilité $\Pi = \frac{\Psi_T^2}{\int \Psi_T^2 dR}$ il n'est nul besoin de calculer l'intégrale $\int \Psi_T^2 dR$ puisque, comme on le voit dans les expressions, seul le calcul du rapport $\frac{\Psi_T^2(R^n)}{\Psi_T^2(R^{n+1})}$ est nécessaire. Heureusement, car le calcul de cette intégrale de normalisation (équivalent à une fonction de partition) est un problème non-trivial. Remarquons que ne pas avoir à effectuer ce calcul a été une des clefs du succès de l'algorithme de Metropolis lors de sa sortie.

2.2.7 Erreurs dans la méthode VMC

Il existe deux types d'erreur qui apparaissent lorsque l'on effectue un calcul VMC

- une erreur statistique

- une erreur systématique (biais)

2.2.8 L'erreur statistique

En pratique, le nombre N de tirages aléatoires d'une simulation Monte-Carlo est toujours en nombre fini, ce qui implique une erreur statistique $\epsilon(N)$ sur le calcul des propriétés. Ainsi, pour une observable O quelconque :

$$\langle O \rangle_{exact} = \langle O \rangle_{fini} + \epsilon(N) \quad (2.33)$$

avec

$$\epsilon(N) = \frac{\sigma(O)}{\sqrt{N_{eff}}} \quad (2.34)$$

où σ est l'écart-type défini comme la racine carrée de la variance de l'observable O :

$$\sigma^2(O) = \langle (O - \langle O \rangle)^2 \rangle_{\Pi} \quad (2.35)$$

$$N_{eff} = kN \quad (2.36)$$

avec $0 < k < 1$. N_{eff} peut être interprété comme le nombre de points réellement indépendants de la chaîne de Markov.

L'erreur statistique est en $\approx \frac{1}{\sqrt{N_{eff}}}$. Autrement dit, pour diminuer la barre d'erreur par 2, il faut multiplier le nombre de pas (et donc le temps de calcul) par 4. Nous pouvons faire cela en multipliant par 4 le temps de calcul sur chacun des processeurs utilisés ou bien en multipliant le nombre de processeurs par 4. Cette dernière remarque repose sur une des propriétés remarquables des simulations Monte-Carlo, à savoir d'être parfaitement parallélisables. En effet, la méthode VMC consiste à simuler des marches aléatoires indépendantes les unes des autres. Il suffit donc de lancer des marches aléatoires indépendantes sur différents processeurs indépendants (sans communications) puis de moyenner les résultats finaux.

2.2.9 L'erreur systématique

Cette erreur est liée à la qualité du choix de la fonction d'essai considérée comme une approximation de la fonction d'onde exacte inconnue. Dans le

calcul des énergies par la méthode VMC, l'enjeu consiste à choisir une “bonne” fonction d’onde, c’est à dire une fonction d’essai qui donnera l’énergie la plus basse possible. Dans le chapitre suivant nous préciserons plus en détail ce que nous entendons par “bonne”. Notons dès à présent l’importance de ce choix puisque c’est la fonction d’essai qui détermine directement la valeur de l’énergie variationnelle. Nous verrons plus loin le rôle joué par la fonction d’essai dans la méthode Diffusion Monte-Carlo (DMC), en particulier en rapport avec l’erreur systématique dite erreur des “nœuds fixés”. Insistons sur le fait que le choix d’une fonction d’essai est un problème crucial dans les méthodes Monte-Carlo quantique et qu’un des points forts de cette thèse est de présenter une nouvelle forme de fonction d’essai.

L’énergie variationnelle E_{VMC} obéit au principe variationnel :

$$E_{VMC} \geq E_0 \quad (2.37)$$

Le cas limite $E_{VMC} = E_0$ est atteint lorsque $\Psi_T = \Psi_0$. Dans ce cas là, la variance de l’énergie locale est nulle.

$$E_L \equiv \frac{\hat{H}\Psi_T}{\Psi_T} \equiv \frac{\hat{H}\Psi_0}{\Psi_0} \equiv E_0 \quad (2.38)$$

2.3 L’équation de Schrödinger comme équation d’évolution d’un processus de diffusion : la méthode Diffusion Monte Carlo (DMC)

La méthode de diffusion de Monte-Carlo (DMC) permet d’obtenir la solution de l’équation de Schrödinger stationnaire. Pour cela, on assimile l’équation de Schrödinger en temps imaginaire ($t \rightarrow -it$) à une équation d’évolution de la densité d’un processus de diffusion qu’il sera alors possible de simuler sur machine en utilisant le point de vue de Langevin des processus (équation différentielle stochastique associé à un processus). Mathématiquement, il s’agit de résoudre une équation de Focker-Planck à l’aide de la formule de Feynman-Kac. Concrètement, la simulation mettra en jeu trois processus élémentaires :

- un processus de diffusion libre (mouvement Brownien libre)
- un terme de dérive déterministe construit à partir du vecteur-dérive (drift)
- un processus de branchement (branching) ou processus de mort/naissance (birth/death process)

La partie qui suit a pour but de présenter les aspects théoriques de la méthode DMC[96]

2.3.1 Principes généraux

Considérons l'équation de Schrödinger en temps imaginaire :

$$\frac{\partial \Phi(R, t)}{\partial t} = [D\nabla^2 - (V(R) - E_T)]\Phi(R, t) \quad (2.39)$$

Dans cette équation $D = \frac{\hbar}{2m_e}$, $V(R)$ est le potentiel coulombien d'interaction entre les différentes charges (en unités atomiques) :

$$V(R) = \sum_{i>j} \frac{1}{r_{ij}} - \sum_{i,\alpha} \frac{Z_\alpha}{r_{i,\alpha}} + \sum_{\alpha>\beta} \frac{Z_\alpha Z_\beta}{r_{\alpha\beta}} \quad (2.40)$$

où les indices latins correspondent aux électrons et les indices grecs aux noyaux, et E_T est une énergie de référence arbitraire.

En résolvant l'équation 2.39, on obtient en décomposant sur la base des fonctions propres de l'Hamiltonien :

$$\Phi(R, t) = \sum_i N_i e^{-(E_i - E_T)t} \Phi_i(R) \quad (2.41)$$

où

$$N_i = \langle \Phi(R, 0) | \Phi_i \rangle \quad (2.42)$$

Les E_i sont les valeurs propres de l'Hamiltonien et les fonctions Φ_i les états propres associés. Pour un temps suffisamment long, seul le terme de plus basse énergie subsiste :

$$\Phi(R, t) = N_0 e^{-(E_0 - E_T)t} \Phi_0(R) \quad (2.43)$$

En prenant $E_T = E_0$, on obtient alors un état proportionnel à l'état fondamental Φ_0

2.3.2 Simulations numériques

Reprenons l'équation de Schrödinger en temps imaginaire (Eq.2.39). Nous pouvons décomposer le membre de droite en deux termes :

- $D\nabla^2$ qui, s'il était seul, permettrait d'interpréter l'équation 2.39 comme un processus de diffusion libre qui pourrait être simulé par une marche aléatoire libre de particules indépendantes à travers l'espace.
- $-[V(R) - E_T]\Phi(R)$ qui, s'il était seul, permettrait d'interpréter l'équation comme l'équation d'évolution d'un processus de mort/naissance

L'équation 2.39 peut donc être simulée numériquement par une combinaison d'un processus de diffusion et de mort/naissance. Remarquons dès à présent que pour permettre une telle analogie entre la fonction d'onde et la densité d'un tel processus, il est nécessaire que la fonction d'onde soit de signe constant (positive ou négative à un changement de phase près). Cette contrainte est vérifiée pour les systèmes bosoniques mais ne l'est plus pour les systèmes fermioniques. Nous reviendrons sur cette question importante par la suite.

Le processus de mort/naissance va diminuer la densité de probabilité dans les régions où $V(R)$ est plus grand que l'énergie de référence E_T en tuant avec grande probabilité des marcheurs se trouvant dans cette région et l'augmenter dans les régions où $V(R)$ est plus petit que E_T en créant des copies des marcheurs à ces endroits là. Cependant, il est important de remarquer que le potentiel coulombien $V(R)$ n'étant pas borné, l'utilisation d'un tel processus de mort/naissance peut devenir très inefficace à cause de trop grandes fluctuations de population. Pour mieux guider le processus de diffusion et limiter l'explosion de la population, on utilise la technique dit de l'"échantillonnage selon l'importance" (importance sampling).

2.3.3 Echantillonnage selon l'importance

L'échantillonnage selon l'importance est une méthode permettant d'étudier les propriétés d'une distribution en utilisant un échantillonnage associé à une distribution alternative. Elle permet de diminuer les fluctuations sta-

tistiques des quantités à calculer en décroissant les variations spatiales des estimateurs à moyenner. Concrètement, l'équation 2.39 est multipliée par une fonction d'essai $\Psi_T(R)$ et on cherche l'équation vérifiée par la nouvelle distribution donnée par $f(R, t) = \Phi(R, t)\Psi_T(R)$. A partir de l'équation 2.39 il est facile de vérifier que l'on a :

$$\frac{\partial f(R, t)}{\partial t} = [D\nabla^2 f - D\nabla(F_Q(R)f) - (E_L(R) - E_T)f] \quad (2.44)$$

avec $E_L = \frac{H\Psi_T}{\Psi_T}$ (énergie locale) et $F_Q = \nabla \ln|\Psi_T(R)|^2 = 2\frac{\nabla\Psi_T}{\Psi_T}$. L'équation décrit à présent un processus de diffusion pour la densité de probabilité $f(R, t)$ et un processus de branchement associé à $(E_L(R) - E_T)$. Au contraire de ce qu'on avait précédemment, pour un choix de Ψ_T limitant les variations de E_L et pour un choix de E_T adéquat, le terme de branchement peut être contrôlé. Un autre terme fait son apparition dans l'équation, à savoir le terme de dérive associé à F_Q . Dans les régions où $|\Psi_T|^2$ est faible, le vecteur F_Q est de grand module et joue le rôle d'une force de rappel. Ce terme va donc pousser les marcheurs des régions de faible densité vers les régions de forte densité.

2.3.4 Lien entre VMC et DMC

Quoique reposant toutes deux sur des processus stochastiques, du point de vue théorique, les méthodes VMC et DMC sont distinctes. La méthode VMC repose sur un échantillonnage d'une densité de probabilité Ψ_T^2 donnée *a priori* tandis que la méthode DMC repose sur la projection de la fonction d'onde exacte Φ_0 sur la fonction d'onde d'essai Ψ_T ou, alternativement comme l'échantillonnage de la densité de probabilité $\Psi_T\Phi_0$.

Sur un plan algorithmique, lorsque nous supprimons l'étape de branching de la méthode DMC, nous retombons sur une méthode VMC. C'est pourquoi la première étape d'un calcul DMC pour une fonction Ψ_T donnée consiste à échantillonner la densité de probabilité $|\Psi_T|^2$. Commencer un calcul DMC à partir de la densité de probabilité VMC, au lieu d'une densité de probabilité uniforme, permet de réduire les fluctuations et de faire converger le calcul plus rapidement.

Nous allons voir à présent les types d'erreurs présentes dans un calcul DMC.

2.3.5 L'erreur statistique

Tout comme dans la méthode VMC, l'erreur statistique $\epsilon(N)$ associée au calcul de la moyenne d'une observable O est due au caractère fini de la simulation.

$$\epsilon(N) = \frac{\sigma(O)}{\sqrt{N_{eff}}} \quad (2.45)$$

Cette erreur diminue et tend vers zéro lorsque le nombre de pas Monte Carlo tend vers $+\infty$.

2.3.6 L'approximation des nœuds fixés

Comme nous l'avons vu précédemment, la méthode DMC consiste à interpréter la quantité $\Psi_T \Phi_0$ comme une densité de probabilité. Or, pour pouvoir procéder de la sorte, il est nécessaire que le produit de la fonction d'essai et de la fonction exacte soit de signe constant. Quand nous considérons un système bosonique, il n'y a pas de problème : Φ_0 est de signe constant et en imposant à Ψ_T de ne pas changer de signe, notre produit peut bien être interprété comme une densité. En revanche, lorsque le système est fermionique comme ceux que nous étudions ici, Φ_0 change de signe. Comment gérer efficacement ce changement de signe des fonctions d'onde fermionique est connu dans la littérature sous le terme de "problème du signe". C'est un des problèmes les plus importants de la physique numérique actuelle. Pour imposer à $\Psi_T \Phi_0$ de garder le même signe, on se place alors dans l'approximation dite des nœuds fixés (fixed-node approximation). Pour cela, on impose aux zéros (ou nœuds) de la fonction Φ_0 d'être identiques à ceux de la fonction Ψ_T . Φ_0 et Ψ_T vont donc changer de signe en même temps et le produit $\Psi_T \Phi_0$ va donc bien garder le même signe. Concrètement, l'espace est partitionné suivant les nœuds de la fonction d'essai. Cette partition se maintient grâce au drift $\frac{\nabla \Psi_T}{\Psi_T}$ qui agit au voisinage des nœuds comme une force de rappel fortement répulsive. Lorsque Ψ_T tend vers zéro, le drift tend en module vers $+\infty$. Les

marcheurs restent donc confinés dans des cellules où la fonction d'essai a le même signe. En imposant les zéros à la fonction Φ_0 , nous n'évaluons donc plus exactement $\Psi_T\Phi_0$ mais $\Psi_T\tilde{\Phi}_0$. La méthode DMC dans l'approximation des nœuds fixés ne va donc pas permettre de retrouver l'état fondamental exact, Φ_0 , mais un état approché, $\tilde{\Phi}_0$, solution de l'équation de Schrödinger dont les nœuds correspondent à ceux de la fonction d'essai Ψ_T . L'énergie obtenue n'est donc pas exactement l'énergie de l'état fondamental. En ayant éliminé les erreurs statistiques, il reste donc une petite erreur due à la nature des nœuds de la fonction Ψ_T . C'est l'erreur des nœuds fixés (ou erreur "fixed-node"). On peut montrer [96] que cette erreur est variationnelle. Cela signifie que l'énergie "fixed-node" est toujours supérieure ou égale à l'énergie exacte.

L'erreur des nœuds fixés reste très petite, de l'ordre de quelques pourcents de l'énergie de corrélation. Néanmoins, et c'est un point fondamental que nous serons amenés à discuter dans la suite, c'est dans ces quelques pourcents que se joue une bonne partie de la chimie.

La fonction Ψ_T en imposant la structure nodale à la solution va donc conditionner le résultat ; par exemple, lorsque l'on change de fonctions de base pour exprimer les orbitales moléculaires ou lorsque la fonction d'onde représentant le système est issue de méthodes différentes (fonctions Hartree-Fock, CASSCF, CISD...). Nous étudierons cet aspect dans le chapitre 9. Remarquons également que la nouvelle fonction d'onde multi-Jastrow que nous proposons permet de changer la structure nodale. Nous reviendrons sur cette remarque importante lorsque nous la présenterons plus en détail.

Si les nœuds de Ψ_T sont exactement ceux de la fonction exacte inconnue Φ_0 alors $\tilde{\Phi}_0 = \Phi_0$ et l'erreur des nœuds fixés est nulle. Plus la fonction d'essai sera proche de la fonction de l'état fondamental, plus l'erreur des nœuds fixés aura tendance à diminuer. C'est pourquoi nous cherchons à mettre toutes les informations physiques dans la fonction d'essai pour réduire cette erreur. Nous insistons néanmoins sur le fait qu'il s'agit d'un principe heuristique. En réalité, très peu de choses sont connues sur la topologie des hypersur-

faces nodales. Il n’y a pas à ce jour de stratégie systématique de réduction de l’erreur des nœuds fixés. Des stratégies ont été mises en œuvre pour diminuer cette erreur mais sans donner de résultats totalement probants. Les algorithmes proposés ne sont pas stables. Citons par exemple les méthodes de relâchement des nœuds[107] qui marche pour des très petits systèmes ou pour des systèmes très homogènes (gaz électronique uniforme [108]) ou la méthode Fermion Monte-Carlo[109] dont les résultats ont été contestés[110].

A chaque calcul, nous devons proposer une fonction d’onde. Ce n’est qu’à l’issu de la simulation DMC que nous comparons les résultats obtenus à d’autres (par exemple, expérimentaux ou issus de calculs par d’autres méthodes) pour conclure sur l’erreur des nœuds fixés. Mais nous ne pouvons rien prévoir *a priori*.

Une stratégie classique consiste à optimiser les paramètres de la fonction d’onde (paramètres du Jastrow, orbitales moléculaires et poids des déterminants de Slater) en VMC pour espérer abaisser l’énergie DMC. Cette stratégie donne de bons résultats[111] lorsqu’il s’agit de trouver la meilleure énergie totale. Cependant, cette stratégie tourne court dès lors que nous étudions des différences d’énergies, $E_1 - E_2$. Dans ce cas là, les erreurs des nœuds fixés des deux calculs peuvent s’ajouter et être du même ordre que la différence d’énergie. Or ce sont les différences d’énergies qui comptent en chimie. La méthode “brute force” consistant à optimiser le mieux possible un état pour réduire l’erreur des nœuds fixés puis optimiser séparément le deuxième état avant d’effectuer la différence donne des résultats incertains. Nous avons pour notre part choisi deux stratégies différentes. En effet, se reposer uniquement sur l’optimisation de l’énergie VMC pour améliorer les nœuds lorsque nous calculons des différences d’énergies n’est pas une démarche contrôlée. Nous avons ainsi constaté qu’il était possible d’obtenir des résultats DMC bien différents, alors que les énergies VMC sont très proches. Ceci n’est pas étonnant. En effet, ce qui définit l’erreur des nœuds fixés est la structure nodale seulement. Une fonction avec une mauvaise énergie VMC mais à la structure nodale exacte donnera une énergie DMC exacte. A l’inverse, une

fonction d'onde avec une bien meilleure énergie VMC mais à la structure nodale différente de celle de la fonction exacte donnera une plus mauvaise énergie DMC.

Il est donc nécessaire d'avoir une approche différente en tenant compte bien plus finement de la cohérence des structures nodales entre les différents états. Nous discuterons de cette question plus en détail lorsque nous présenterons le cœur de nos travaux :

- sur l'acroléine où nous proposons une approche permettant de garder la cohérence des nœuds lorsque nous étudions des différences d'énergie entre états
- dans l'introduction d'une nouvelle forme de fonction d'onde qui nous permet de modifier les nœuds de manière plus physique.

Répétons le en guise de conclusion de cette partie : la structure nodale est encore très largement méconnue (voir [112] pour une discussion intéressante sur la structure nodale) et il n'y a pas à ce jour de stratégie systématique de réduction de l'erreur des nœuds fixés. L'erreur des nœuds fixés reste l'un des problèmes majeurs dans le domaine des simulations QMC.

2.3.7 Résultats et limites de la méthode Fixed-Node DMC

Nous présentons ici deux tableaux provenant de [113] présentant des résultats obtenus par la méthode FN-DMC et qui résument notre problématique. Le premier tableau, Table 9.1, compare des résultats obtenus par la méthode Fixed-Node DMC avec nœuds Hartree-Fock face à ceux obtenus par la méthode CCSD(T), une des méthodes de la chimie théorique les plus précises à ce jour quand les effets multi-configurationnels (corrélation non-dynamique) ne sont pas trop marqués. Nous pouvons constater que les énergies totales obtenues par la méthode DMC sont systématiquement plus basses que celles obtenues par la méthode CCSD(T). Bien que la méthode CCSD(T) ne soit pas variationnelle, on peut raisonnablement conclure que les calculs FN-DMC sont plus précis.

Cependant, lorsque l'on regarde le deuxième tableau, Table 9.2, qui s'inté-

resse à des *différences* d'énergies électroniques lors de réactions élémentaires, nous pouvons constater que la méthode CCSD(T) avec la base cc-pVTZ donne des résultats globalement plus précis. On voit donc que les compensations d'erreur opèrent mieux dans le cas de l'approche *ab initio* que dans l'approche QMC.

TAB. 2.1 – Calculs de Manten et Luchöw, Ref.113 : Energies totales calculées avec la méthode Fixed-Node DMC. Les noeuds sont issus d'une fonction Hartree-Fock exprimées dans la base cc-pVTZ. Résultats comparées à CCSD(T)/cc-pVXZ avec X=D,T. Toutes les énergies sont en unités atomiques.

Molecule	FN-DMC	CCSD(T)/cc-pVTZ	CCSD(T)/cc-pVDZ
H ₂	-1.1739(1)	-1.1723	-1.1634
CH ₂	-39.1165(3)	-39.0614	-39.0220
CH ₄	-40.5005(3)	-40.4381	-40.3868
NH ₃	-56.5485(4)	-56.4732	-56.4020
H ₂ O	-76.4207(2)	-76.3322	-76.2410
C ₂ H ₂	-77.3110(4)	-77.1876	-77.1092
C ₂ H ₄	-78.5644(3)	-78.4388	-78.3544
HNC	-93.3987(4)	-93.2751	-93.1884
HF	-100.4466(3)	-100.3379	-100.2275
N ₂	-109.5046(3)	-109.3739	-109.2753
N ₂ H ₂	-110.6054(5)	-110.4780	-110.3670
CO	-113.2877	-113.1555	-113.0544
H ₂ CO	-114.4739(4)	-114.3338	-114.2183
HNO	-130.4371(4)	-130.2984	-130.1710
H ₂ O ₂	-151.5213(3)	-151.3586	-151.1937
HOF	-175.5120(4)	-175.3343	-175.1519
CO ₂	-188.5429(4)	-188.3271	-188.1475
F ₂	-199.4841(4)	-199.2961	-199.0975
O ₂	-225.3410(4)	-225.1326	-224.9091

2.3.8 Algorithme DMC en pratique

Après avoir posé les bases théoriques de la méthode DMC, nous allons présenter l'algorithme concrètement utilisé. Nous allons montrer comment

TAB. 2.2 – Calculs de Manten et Luchöw, Ref.113 : Contribution électronique, ΔH_e , aux réactions d'enthalpies (en kJ mol^{-1}). FN-DMC correspond à des calculs Fixed-Node DMC avec les nœuds d'une fonction d'onde Hartree-Fock en base cc-pVTZ. CC correspond à des calculs CCSD(T)/cc-pVXZ avec X=D (CC-DZ) ou X=T (CC-TZ). Les données expérimentales sont également reportées. Dans chaque compartiment du tableau Δ représente la différence avec les résultats FN-DMC.

Réaction	Expt	FN-DMC	Δ	CC-TZ	Δ	CC-DZ	Δ
$\text{CH}_2^a + \text{H}_2 \rightarrow \text{CH}_4$	-544(2)	-552(2)	-8.	-537.	7.	-529.	15.
$\text{C}_2\text{H}_2 + \text{H}_2 \rightarrow \text{C}_2\text{H}_4$	-203(2)	-209(2)	-6	-207	-4	-215	-12
$\text{C}_2\text{H}_2 + 3\text{H}_2 \rightarrow 2 \text{CH}_4$	-446(2)	-442(2)	4	-451	-5	-457	-11
$\text{N}_2\text{H}_2 \rightarrow \text{N}_2 + \text{H}_2$	-174 ^b	-192(2)	-18	-179	-5	-188	-14
$\text{CO} + \text{H}_2 \rightarrow \text{H}_2\text{CO}$	-21(1)	-32(2)	-11	-16	5	-1	20
$\text{N}_2 + 3\text{H}_2 \rightarrow 2\text{NH}_3$	-164(1)	-186(2)	-22	-146	18	-101	63
$\text{F}_2 + \text{H}_2 \rightarrow 2\text{HF}$	-563(1)	-618(2)	-55	-545	18	-510	53
$\text{O}_3 + 3 \text{H}_2 \rightarrow 3\text{H}_2\text{O}$	-933(2)	-1050(2)	-117	-912	21	-850	83
$\text{H}_2\text{CO} + 2\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-251(1)	-262(2)	-11	-241	10	-217	34
$\text{H}_2\text{O}_2 + \text{H}_2 \rightarrow 2\text{H}_2\text{O}$	-365(2)	-384(2)	-19	-351	14	-328	37
$\text{CO} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-272(1)	-294(2)	-22	-257	15	-218	54
$\text{HCN} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{NH}_3$	-320(3)	-338(2)	-18	-313	7	-289	31
$\text{HNO} + 2\text{H}_2 \rightarrow \text{H}_2\text{O} + \text{NH}_3$	-444(1)	-485(2)	-41	-427	17	-381	63
$\text{HNC} \rightarrow \text{HCN}$	-64 ^b	-66(2)	-2	-63	1	-66	-2
$\text{H}_2\text{O} + \text{F}_2 \rightarrow \text{HOF} + \text{HF}$	-129(4)	-141(2)	-12	-115	14	-107	22
$\text{CO}_2 + 4\text{H}_2 \rightarrow 2\text{H}_2\text{O} + \text{CH}_4$	-244(1)	-272(2)	-28	-226	18	-178	66
$2\text{CH}_2^a \rightarrow \text{C}_2\text{H}_4$	-844(3)	-871(2)	-27	-830	14	-815	29

^aEtat singulet du méthylène, $\tilde{a}^1 A_1$.

^b Les énergies CCSD(T)-R12 ont été utilisées au lieu des énergies expérimentales, Ref.114.

implémenter numériquement les concepts de diffusion et de branching.

Reprenons l'équation de Schrödinger (temps réel, unités atomiques) :

$$i \frac{\partial \phi(R, t)}{\partial t} = (\hat{H} - E_T) \phi(R, t) \quad (2.46)$$

avec \hat{H} l'hamiltonien du système et E_T une énergie de référence arbitraire. En passant en temps imaginaire, c'est à dire $t \rightarrow it$ il vient :

$$\frac{\partial \phi(R, t)}{\partial t} = -(\hat{H} - E_T) \phi(R, t) \quad (2.47)$$

qui peut se réécrire sous la forme :

$$\phi(R, t + \tau) = \int G(R \rightarrow R', \tau) \phi(R', t) dR' \quad (2.48)$$

où $G(R \rightarrow R', \tau)$ est une fonction de Green qui obéit à la même équation :

$$\frac{\partial G(R \rightarrow R', t)}{\partial t} = -(\hat{H}(R) - E_T) G(R \rightarrow R', t) \quad (2.49)$$

avec comme condition initiale :

$$G(R \rightarrow R', 0) = \delta(R - R') \quad (2.50)$$

Cette fonction est de la forme :

$$G(R \rightarrow R', \tau) = \langle R | \exp[-\tau(\hat{H}(R) - E_T)] | R' \rangle \quad (2.51)$$

En utilisant la décomposition spectrale de l'opérateur $\exp[-\tau\hat{H}]$,

$$\exp(-\tau\hat{H}) = \sum_i |\psi_i\rangle \exp(-\tau E_i) \langle \psi_i | \quad (2.52)$$

la fonction de Green se réécrit :

$$G(R \rightarrow R', t) = \sum_i \exp[-\tau(E_i - E_T)] \psi_i(R) \psi_i^*(R') \quad (2.53)$$

2.3.9 DMC dans le cas d'un potentiel nul

Lorsque l'équation de Schrödinger se réduit à l'équation d'évolution du processus stochastique libre (équation de la diffusion) :

$$\frac{\partial \phi(R, \tau)}{\partial \tau} = \frac{1}{2} \sum_i^N \nabla_i^2 \phi(R, \tau) \quad (2.54)$$

la fonction de Green s'écrit :

$$G(R \rightarrow R', \tau) = \left[\frac{1}{\sqrt{2\pi\tau}} \right]^{3N_e} \exp\left[-\frac{|R' - R|^2}{2\tau} \right] \quad (2.55)$$

Elle peut être interprétée comme une densité de probabilité de transition et permet de simuler le processus de diffusion via l'algorithme de Metropolis.

2.3.10 DMC dans le cas d'un potentiel non nul

Lorsque l'on rajoute un potentiel V au terme cinétique T de l'Hamiltonien, $\hat{H} = \hat{T} + \hat{V}$ nous utilisons alors la formule de Trotter-Suzuki [115,116,117] :

$$e^{-\tau(\hat{A} + \hat{B})} = e^{-\frac{\tau\hat{B}}{2}} e^{-\tau\hat{A}} e^{-\frac{\tau\hat{B}}{2}} + O(\tau^3) \quad (2.56)$$

Elle permet de réécrire la fonction de Green

$$\begin{aligned} G(R \rightarrow R', \tau) &= \langle R | \exp[-\tau(\hat{T} + \hat{V} - E_T)] | R' \rangle \approx \\ &e^{-\frac{\tau(\hat{V}(R) - E_T)}{2}} \langle R | \exp[-\tau\hat{T}] | R' \rangle e^{\frac{-\tau(\hat{V}(R') - E_T)}{2}} + O(\tau^3) \end{aligned} \quad (2.57)$$

La forme approchée de la fonction de Green est alors :

$$G(R \rightarrow R', \tau) \approx \left[\frac{1}{\sqrt{2\pi\tau}} \right]^{3N_e} \exp\left[-\frac{|R' - R|^2}{2\tau} \right] \exp\left[-\frac{\tau}{2}(V(R) + V(R') - 2E_T) \right] \quad (2.58)$$

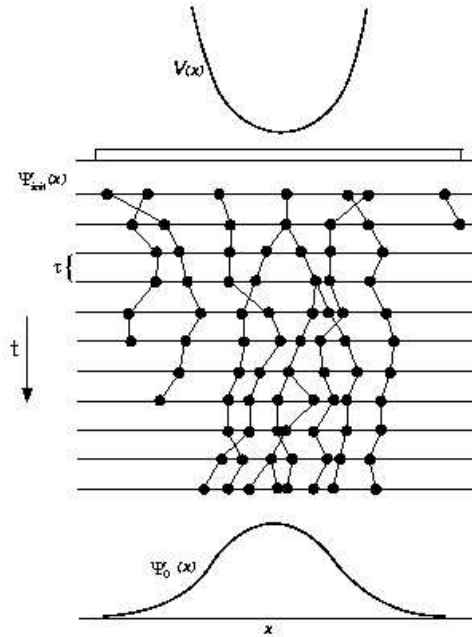
Nous retrouvons la probabilité de transition qui nous servira à simuler la diffusion dans l'algorithme de Metropolis. Le facteur $P = \exp[-\frac{\tau}{2}(V(R) + V(R') - 2E_T)]$ quant à lui permet de simuler le processus de mort/naissance. Le nombre de marcheurs n'est plus fixe. A chaque pas, P détermine le nombre de marcheurs qui survivent. On peut par exemple coder cela en introduisant :

$$M = E(P + \eta) \quad (2.59)$$

avec $E()$ désignant la partie entière et η un nombre aléatoire uniforme entre 0 et 1. M représente ici le nombre de marcheurs à l'étape suivante. Cette dernière étape est justifiée par le fait que :

$$\int_0^1 du E(x+u) = x \quad (2.60)$$

Nous voyons que lorsque le potentiel V est important, P est petit et les marcheurs tendent à disparaître. Au contraire, lorsque $V \ll E_T$, de nouveaux marcheurs ont tendance à naître dans cette région. Le dessin ci-dessous décrit cette situation pour un Hamiltonien unidimensionnel.



2.3.11 Résumé d'un pas élémentaire dans l'algorithme DMC

Nous résumons ici les étapes d'un seul pas DMC. Rappelons que pour accélérer les calculs les points de départ ne sont pas uniformément répartis comme dans la méthode VMC mais répartis suivant la densité $\Pi = \frac{\Psi_T^2}{\int \Psi_T^2 dR}$

- **Etape 1**

Déplacement des points : $R^{n+1} = R^n + D\tau F_Q(R^n) + \sqrt{\tau}\eta$

avec η un vecteur gaussien normal.

- **Etape 2**

Acceptation du nouveau point avec une probabilité de :

$$A(R^n \rightarrow R^{n+1}) = \text{Min}(1, \frac{G(R^{n+1} \rightarrow R^n)\Psi_T^2(R^{n+1})}{G(R^n \rightarrow R^{n+1})\Psi_T^2(R^n)}) \quad (2.61)$$

- **Etape 3**

Processus de mort/naissance :

Le nombre de marcheurs est $M=E(P+\eta)$ où η est un nombre aléatoire uniforme entre 0 et 1.

2.3.12 Résumé des étapes principales d'un calcul Diffusion Monte-Carlo

Un calcul DMC dans sa globalité va donc se dérouler suivant les étapes suivantes

- Choix d'une fonction d'essai Ψ_T
- Obtention de cette fonction à l'issu d'un calcul *ab initio* ou/et optimisation des paramètres
- Calcul VMC pour obtenir une densité $\Pi = \frac{\Psi_T^2}{\int \Psi_T^2 dR}$
- Calcul DMC

Chapitre 3

Optimisation de la fonction d'onde

L'optimisation de la fonction d'onde d'essai Ψ_T est une étape importante des simulations Monte Carlo quantique. Quelque soit la forme de fonction d'onde choisie, il est très souvent avantageux d'optimiser les paramètres de cette fonction. En général, il n'y a pas de jeu de paramètres par défaut à utiliser pour n'importe quel système. Bien au contraire, la sensibilité au choix des paramètres est souvent très grande. Idéalement, il serait souhaitable d'optimiser la fonction d'onde d'essai de façon à minimiser l'erreur des nœuds fixés (seule erreur résiduelle quand le temps de simulation augmente indéfiniment). Malheureusement, une telle méthode n'a pas encore été mise au point. En pratique, les optimisations se font au niveau variationnel par la méthode VMC et les quantités qui sont minimisées sont soit la variance de l'énergie, soit l'énergie totale elle-même. Ces trois dernières années des progrès importants ont été faits concernant le minimisation de l'énergie. Dans ces approches, on détermine les paramètres de la fonction d'onde qui minimisent l'énergie VMC puis ces paramètres sont utilisés pour calculer les énergies fixed-node DMC. Insistons sur le fait que la démarche consistant à minimiser l'énergie VMC pour espérer abaisser l'énergie DMC est une démarche naturelle mais qui n'a pas de raison de fonctionner dans le

cas général. En effet, la minimisation de l'énergie VMC peut parfois conduire à des énergies DMC plus hautes. La question de l'optimisation est donc au cœur de notre réflexion sur l'amélioration de l'erreur fixed-node et c'est pour cela qu'il s'agit d'une question fondamentale en QMC.

L'optimisation de la fonction d'onde d'essai par rapport à un grand nombre de paramètres est un problème hautement non-trivial comme tout problème d'optimisation en mathématique. Il s'agit d'optimiser de très nombreux paramètres (une centaine, voire plus) d'une fonction d'onde dont les moyennes sont bruitées parce qu'issues d'une simulation stochastique. Nous présentons ici les différentes méthodes d'optimisation que nous avons implémentées.

La première méthode d'optimisation, qui était la méthode la plus répandue au début de ce travail de thèse, repose sur la minimisation de la variance de l'énergie locale à partir d'une méthode d'échantillonnage corrélé [118]. L'avantage de minimiser la variance est que cette quantité est bornée inférieurement (par zéro) pour toute simulation stochastique *finie*. Comme conséquence, toute réduction de la variance de l'énergie locale pour un nombre d'événements Monte Carlo beaucoup plus grand que le nombre de paramètres à optimiser est la plupart du temps une vraie amélioration des paramètres. En revanche, ce n'est pas du tout le cas pour l'énergie totale qui n'est pas bornée inférieurement pour une simulation *finie* (avec des potentiels coulombiens, il est possible d'avoir des énergies locales infiniment basses). En abaissant la variance, on est censé "améliorer" la fonction d'onde et donc on espère abaisser l'énergie variationnelle. L'expérience montre que cela fonctionne dans la plupart des cas, mais évidemment ce résultat n'est pas assuré dans le cas général. Malheureusement, la méthode de l'optimisation par réduction de la variance n'est pas assez précise lorsque l'on souhaite travailler sur le niveau d'optimisation aussi fin de l'énergie qui sera requis dans notre travail.

Pendant le déroulement de cette thèse sont apparues successivement deux autres méthodes d'optimisation qui travaillent directement sur la minimisa-

tion de l'énergie variationnelle, qui sont bien plus efficaces et que j'ai donc mises en œuvre :

- Une méthode de Newton[119] (2005)
- Une méthode linéaire[120],[121] (2006)

La méthode de Newton résout le problème de la difficulté de considérer l'énergie totale comme quantité à minimiser en introduisant un principe de zéro-variance qui permet un bien meilleur contrôle des fluctuations de l'énergie. La méthode linéaire, quant à elle, s'appuie sur d'autres idées que nous développerons dans ce chapitre.

Ces deux nouvelles méthodes ont été testées dans les articles originaux [119,120,121] sur des fonctions d'onde de forme "standard" (produits d'une somme de déterminants par un préfacteur de Jastrow). Nous les avons implémentées dans le code QMC=Chem[122] développé dans notre groupe pour permettre d'optimiser notre nouvelle fonction d'onde que nous présenterons plus loin. Après avoir testé ces deux méthodes sur nos fonctions d'onde nous avons abandonné la méthode de Newton au profit de la méthode linéaire. En effet, nous nous sommes aperçus lors de nos essais pour optimiser les paramètres des Jastrows que cette étape s'effectuait avec des difficultés pratiques. Des problèmes de stabilité ne rendaient pas l'optimisation aussi facile que ce que l'article original le laissait entendre et de nombreux essais/erreurs étaient nécessaires. Ces difficultés ont disparu avec l'approche linéaire qui permet, même avec nos fonctions d'onde originales une optimisation facile et quasi-automatique. Outre l'aspect pratique pour effectuer les travaux de cette thèse, cette automaticité est un pré-requis important pour espérer une diffusion des méthodes Monte Carlo quantique auprès d'un public non-spécialisé dans l'optimisation des fonctions d'onde. Une autre contrainte qui a rendu nécessaire l'abandon de la méthode de Newton au profit de la méthode linéaire est une contrainte de temps de calcul. En effet, la méthode de Newton nécessite, en plus des quantités nécessaires à la méthode linéaire, la dérivée double de la fonction d'onde suivant tous les paramètres. Lorsque le Jastrow est en préfacteur comme dans les fonctions d'onde standards du QMC, ce

problème ne se présente pas car la dérivée double représente seulement le calcul de quelques coefficients supplémentaires. Mais lorsqu'il se trouve dans les déterminants comme c'est le cas dans notre nouvelle fonction d'onde, la dérivée double se traduit par le calcul de nombreux déterminants supplémentaires. Nous développerons cet aspect-là plus précisément dans le chapitre 5 qui décrit les détails techniques de l'implémentation de notre algorithme.

3.1 Echantillonnage corrélé par pondération

Avant de présenter la méthode d'optimisation par la variance, nous devons au préalable présenter le principe de la pondération qui est la façon la plus naturelle d'effectuer un échantillonnage corrélé.

Nous avons vu qu'une moyenne $\langle O_1 \rangle$ obtenue par un calcul Monte Carlo est donnée avec une erreur statistique ϵ . Cette erreur provient du nombre fini de tirages Monte Carlo (points ou marcheurs) à partir desquels s'effectue la moyenne.

$$\langle O_1 \rangle = \langle O_1 \rangle_{exact} + \epsilon_1$$

Ainsi, pour comparer deux moyennes $\langle O_1 \rangle$ et $\langle O_2 \rangle$ calculées suivant deux densités de probabilité différentes, nous devons nous assurer que les barres d'erreurs ϵ_1 et ϵ_2 soient (très) inférieures à la différence $\langle O_1 \rangle - \langle O_2 \rangle$. Une autre manière d'estimer ces différences et de réécrire ces quantités à l'aide d'une densité commune. Plaçons-nous dans le cadre d'une approche VMC avec utilisation de deux fonctions d'onde d'essai, Ψ_1 et Ψ_2 . Nous avons

$$\begin{aligned} \Delta O &\equiv \langle O_1 \rangle - \langle O_2 \rangle \\ &= \frac{\int O_1 |\Psi_1|^2 dR}{\int |\Psi_1|^2 dR} - \frac{\int O_2 |\Psi_2|^2 dR}{\int |\Psi_2|^2 dR} \\ &= \frac{\int O_1 |\Psi_0|^2 (\frac{\Psi_1}{\Psi_0})^2 dR}{\int |\Psi_0|^2 (\frac{\Psi_1}{\Psi_0})^2 dR} - \frac{\int O_2 |\Psi_0|^2 (\frac{\Psi_2}{\Psi_0})^2 dR}{\int |\Psi_0|^2 (\frac{\Psi_2}{\Psi_0})^2 dR} \end{aligned}$$

En définissant un poids de la manière suivante :

$$w_i = \left(\frac{\Psi_i}{\Psi_0} \right)^2$$

nous estimons la différence de moyenne pour une trajectoire finie de la manière suivante :

$$\Delta O \approx \sum_i \left[\frac{w_1(R_i)O_1(R_i)}{\sum_j w_1(R_j)} - \frac{w_2(R_i)O_2(R_i)}{\sum_j w_2(R_j)} \right] \quad (3.1)$$

où les R_i sont des configurations prises le long de trajectoires VMC obtenues à partir de la densité de référence : $|\Psi_0|^2$. Ainsi, les quantités O_1 et O_2 sont évaluées en prenant les *mêmes points* d'une trajectoire commune (celle-ci peut d'ailleurs être issue d'une simulation VMC à partir de la densité $|\Psi_1|^2$ ou $|\Psi_2|^2$). La pondération est très efficace quand les quantités Ψ_1 et Ψ_2 sont "proches" de Ψ_0 . En effet, dans ce cas-là, les poids w_i ($i = 1, 2$) sont très proches de 1 et chaque contribution de la somme dans l'Eq.3.1 est toujours petite.

Pour quantifier cette notion de "petitesse", nous avons introduit la quantité suivante,

$$\eta = \frac{[\sum_i w(R_i)]^2}{\sum_i w^2(R_i)}$$

qui est une mesure de la fluctuation du poids w le long de la trajectoire. Empiriquement, on considère l'échantillonnage acceptable pour

$$\eta \gtrsim 0.9$$

3.2 Optimisation par la variance

Nous minimisons la variance σ^2 de l'énergie locale $E_L = \frac{H\Psi_T}{\Psi_T}$.

$$\sigma^2 \equiv \langle (E_L - \bar{E})^2 \rangle_{|\Psi_T|^2}$$

où \bar{E} est la valeur moyenne de l'énergie. La variance est bornée inférieurement par zéro. Cette limite est atteinte dans le cas où $\Psi_T = \Psi_0$ (fonction d'onde exacte). On a alors :

$$E_L = \frac{H\Psi_T}{\Psi_T} \equiv E_0$$

où E_0 est l'énergie exacte. La méthode d'optimisation par la variance consiste à minimiser la quantité :

$$\sigma^2(p) = \frac{\sum_{i=1}^{N_{MC}} w_{i,p} \left(\frac{H\Psi_T(i,p)}{\Psi_T(i,p)} - \bar{E} \right)^2}{\sum_{i=1}^{N_{MC}} w_{i,p}} \quad (3.2)$$

avec

$$w_{i,p} = \frac{\Psi_T^2(i,p)}{\Psi_0^2(i)}.$$

Ici, p désigne le jeu de paramètres à optimiser et N_{MC} le nombre de configurations Monte Carlo. L'introduction d'un poids nous permet d'utiliser le principe de pondération présenté plus haut : pour rechercher le minimum de $\sigma^2(p)$ nous avons besoin de différentes valeurs $\sigma^2(p_1), \sigma^2(p_2), \dots$ et toutes ces quantités sont calculées le long d'une seule trajectoire correspondant à une densité commune $|\Psi_0|^2$ et indépendante de p .

3.3 La méthode de Newton

3.3.1 Description théorique

La méthode de Newton utilisée est une méthode de Newton classique. Elle se base sur un développement de Taylor de l'énergie à l'ordre 2 suivant les paramètres.

$$E^{(2)}(p) = E(p^0) + \sum_i \frac{\partial E(p^0)}{\partial p_i} \Delta p_i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 E(p^0)}{\partial p_i \partial p_j} \Delta p_i \Delta p_j$$

où on a posé $\Delta p = p - p^0$. La méthode de Newton est une méthode itérative qui permet sous certaines conditions de converger vers un minimum local. A l'étape n de l'itération, les nouveaux paramètres $p^{(n+1)}$ sont donnés par la formule :

$$p^{(n+1)} = p^{(n)} - H^{(n)-1} G^{(n)} \quad (3.3)$$

où $H^{(n)}$ est la matrice Hessienne à l'itération n . La matrice Hessienne étant définie de la manière suivante :

$$H_{ij}^{(n)} \equiv \frac{\partial^2 E}{\partial p_i \partial p_j} [p^{(n)}] \quad (3.4)$$

et $G^{(n)}$ le gradient à l'itération n

$$G_i^{(n)} = \frac{\partial E}{\partial p_i} [p^{(n)}]. \quad (3.5)$$

3.3.2 Implémentation en QMC

Les éléments des matrices H et G sont obtenus par un calcul VMC. Cependant, toute la difficulté en QMC (qui n'existe pas dans les méthodes déterministes) est de trouver un bon estimateur de ces quantités afin de contrôler le bruit statistique. En effet, pour un temps de calcul infini, nous sommes assurés d'obtenir les quantités voulues, mais pour un temps de simulation fini, ces quantités sont obtenues avec une barre d'erreur. Le problème revient donc à implémenter le schéma 3.3 avec les matrices bruitées suivantes :

$$H = \begin{pmatrix} h_{11} + \epsilon_{11} & h_{12} + \epsilon_{12} & \dots \\ h_{21} + \epsilon_{21} & h_{22} + \epsilon_{22} & \dots \\ \vdots & \vdots & \vdots \end{pmatrix} \quad (3.6)$$

$$G = \begin{pmatrix} g_1 + \epsilon_1 \\ g_2 + \epsilon_2 \\ \dots \end{pmatrix} \quad (3.7)$$

C'est la présence des fluctuations ϵ qui rend le problème d'optimisation en QMC très délicat.

3.3.3 Les estimateurs

Les estimateurs des éléments de matrice de H et G utilisés dans les calculs VMC sont donnés par les formules :

$$G_i = 2 \langle \frac{\Psi_i}{\Psi} (E_L - \bar{E}) \rangle \quad (3.8)$$

et

$$H_{ij} = 2 \left[\langle \left(\frac{\Psi_{ij}}{\Psi} + \frac{\Psi_i \Psi_j}{\Psi^2} \right) (E_L - \bar{E}) \rangle - \langle \frac{\Psi_i}{\Psi} \rangle \bar{E}_j - \langle \frac{\Psi_j}{\Psi} \rangle \bar{E}_i \right] \quad (3.9)$$

$$+ \langle \frac{\Psi_i}{\Psi} E_{L,j} \rangle - \langle \frac{\Psi_i}{\Psi} \rangle \langle E_{L,j} \rangle + \langle \frac{\Psi_j}{\Psi} E_{L,i} \rangle - \langle \frac{\Psi_j}{\Psi} \rangle \langle E_{L,i} \rangle$$

Dans ces deux formules nous avons introduit les notations suivantes :

$$\Psi_i \equiv \frac{\partial \Psi}{\partial c_i}$$

$$\begin{aligned}\Psi_{ij} &\equiv \frac{\partial^2 \Psi}{\partial c_i \partial c_j} \\ E_{L,i} &\equiv \frac{\partial E_L}{\partial c_i}\end{aligned}\tag{3.10}$$

L'expression du gradient G_i est facilement obtenue en dérivant l'énergie variationnelle $\frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle}$ suivant le paramètre c_i .

En dérivant le gradient G_i suivant le paramètre c_j , nous obtenons la dérivée seconde de l'énergie :

$$\begin{aligned}\frac{\partial^2 \langle \Psi | H | \Psi \rangle}{\partial c_i \partial c_j} &= 2 \left[\left\langle \left(\frac{\Psi_{ij}}{\Psi} + \frac{\Psi_i \Psi_j}{\Psi^2} \right) (E_L - \bar{E}) \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \bar{E}_j - \left\langle \frac{\Psi_j}{\Psi} \right\rangle \bar{E}_i \right] \\ &\quad + \left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle + \left\langle \frac{\Psi_j}{\Psi} E_{L,i} \right\rangle\end{aligned}\tag{3.11}$$

Le point important à constater est que les deux estimateurs du Hessien données par les formules 3.9 et 3.11 sont différents. En effet, dans la formule proposée en (3.9) deux termes supplémentaires ont été ajoutés, à savoir :

- $-\left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_{L,j} \rangle$
- $-\left\langle \frac{\Psi_j}{\Psi} \right\rangle \langle E_{L,i} \rangle$

Cependant, comme nous allons le voir, ces quantités sont de moyenne nulle. Les deux estimateurs ont bien la même moyenne égale à l'élément de matrice du Hessien recherché. En revanche, introduire ces termes supplémentaires est très utile car l'estimateur (3.9) présente des fluctuations statistiques bien plus petites que l'estimateur (3.11). Une telle approche consiste à invoquer un principe dit de "zéro-variance" qui permet de diminuer les fluctuations statistiques sans changer les moyennes. Utilisant un tel principe l'algorithme de Newton peut alors être appliqué avec une efficacité bien plus grande.

3.3.4 Le principe de zéro-variance

Le principe de zéro-variance consiste à remplacer un estimateur O par un estimateur \tilde{O} ayant la même moyenne mais une variance plus petite. Le principe de zéro-variance pour les méthodes Monte Carlo en général a été discuté dans [123] puis généralisé aux approches Monte Carlo quantique et au calcul des observables dans [124,125].

Ici sera utilisé le principe présenté dans [126] qui consiste à remplacer systématiquement la moyenne $\langle AB \rangle$ par :

$$\langle AB \rangle - \langle A \rangle \langle B \rangle$$

quand $\langle A \rangle = 0$ ce qui conduit à réduire fortement les fluctuations du produit AB quand les fluctuations de B sont importantes :

$$\sigma^2(\langle AB \rangle - \langle A \rangle \langle B \rangle) \ll \sigma^2(\langle AB \rangle)$$

Dans notre cas est appliqué ce principe en notant que :

$$\langle E_{L,i} \rangle = 0$$

En effet :

$$\langle E_{L,j} \rangle = \left\langle \frac{\partial}{\partial c_j} \left(\frac{H\Psi}{\Psi} \right) \right\rangle \quad (3.12)$$

$$= \left\langle \frac{H\Psi_j}{\Psi} \right\rangle - \left\langle \frac{\Psi_j H\Psi}{\Psi^2} \right\rangle \quad (3.13)$$

$$= \left\langle \frac{H\Psi_j}{\Psi} \right\rangle - \left\langle \frac{H\Psi_j}{\Psi} \right\rangle \quad (3.14)$$

$$= 0. \quad (3.15)$$

Le passage de la deuxième à la troisième égalité utilise le fait que H est un opérateur hermitique et donc : $\langle \Psi_j H\Psi \rangle = \langle \Psi H\Psi_j \rangle$.

3.3.5 Schéma de stabilisation de l'algorithme

Malgré le principe de zéro-variance qui améliore beaucoup le problème des fluctuations, il est encore nécessaire de conditionner au mieux l'algorithme vis à vis des fluctuations statistiques.

Inversion de la matrice par décomposition en valeurs singulières

Pour réduire l'impact des fluctuations sur l'étape d'inversion de la matrice Hessienne, nous avons proposé d'inverser cette matrice par une méthode de décomposition en valeurs singulières (Singular Value Decomposition ou SVD) qui nous permet d'éliminer les composantes parasites dues au bruit.

Voyons plus en détail cette méthode : La SVD permet de décomposer une matrice A de taille $m \times n$ en trois matrices :

$$A = QDT^t \quad (3.16)$$

- Q est une matrice aux colonnes orthogonales de taille $m \times n$ (i.e $Q^tQ = Id$)
- T est une matrice orthogonale de taille $n \times n$ (i.e $T^tT = TT^t = Id$)
- D est une matrice diagonale de taille $n \times n$ dont les éléments sont positifs ou nuls :

$$D = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix}$$

Ici la matrice H pour laquelle nous appliquons la SVD est carrée de taille $n \times n$:

$$H = Q \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{pmatrix} T^t \quad (3.17)$$

L'inversion s'effectue donc en posant :

$$H^{-1} = Q^t \begin{pmatrix} \frac{1}{d_1} & & & \\ & \frac{1}{d_2} & & \\ & & \ddots & \\ & & & \frac{1}{d_n} \end{pmatrix} T \quad (3.18)$$

Cependant l'un ou plusieurs des éléments d_i peut, à cause du bruit statistique, être nul ou très petit (de l'ordre de 10^{-12} ou moins). Dans ce cas-là, on peut montrer qu'en remplaçant $\frac{1}{d_i}$ par 0 la partie physiquement importante de la matrice est conservée. Nous démontrons l'existence de la décomposition en valeurs singulières d'une matrice dans l'annexe B. Nous avons implémenté dans le code QMC=Chem la subroutine SVDCMP. SVDCMP est une

routine écrite en Fortran dont le code source est décrit dans l'ouvrage Numerical Recipes [127]. La subroutine SVDCMP décompose en valeur singulière une matrice A de taille quelconque $n \times m$ suivant l'algorithme de Golub-Reinsch[128]. Cet algorithme décompose la matrice A en valeur singulière en 2 étapes :

- une décomposition de la matrice A en une matrice bidiagonale B par la méthode de Householder[129]
- la décomposition de la matrice B en valeur singulière par la méthode QR[130],[131],[132]

Une matrice bidiagonale est une matrice de la forme :

$$B = \begin{pmatrix} b_{11} & b_{12} & 0 & 0 \dots & 0 \\ 0 & b_{22} & b_{23} & 0 \dots & 0 \\ \vdots & 0 & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & & & & \end{pmatrix}$$

Pour une description complète de la méthode de Golub and Reinsch, voir l'annexe B.

Stabilisation de l'optimisation

Une optimisation de Newton n'est efficace qu'au voisinage du minimum dans une espace convexe. Or, nous avons à optimiser en une fois de très nombreux paramètres (une soixantaine et plus) non-linéaires pour une fonction bruitée. Afin de stabiliser le schéma, il a été proposé dans [121] d'ajouter un paramètre λ à la diagonale de la matrice Hessienne :

$$\tilde{H} = \begin{pmatrix} H_{11} + \lambda & H_{12} & \dots & \dots \\ H_{21} & H_{12} + \lambda & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ H_{N_{opt}1} & \dots & \dots & H_{N_{opt}N_{opt}} + \lambda \end{pmatrix} \quad (3.19)$$

Regardons le rôle de λ .

Dans le cas-limite où $\lambda = 0$, la méthode est une méthode de Newton "pure".

Lorsque $\lambda \gg H_{ij} \quad \forall i, j$, on a :

$$\tilde{H} \sim \begin{pmatrix} \lambda & 0 & \dots & \dots \\ 0 & \lambda & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda \end{pmatrix} \quad (3.20)$$

et l'algorithme de Newton est devenu un algorithme de "plus grande pente" (steepest descent) :

$$\Delta p = -\frac{1}{\lambda} G \quad (3.21)$$

Plus λ sera grand, plus $\frac{1}{\lambda}$ sera petit. Augmenter λ pour un gradient G donné va donc avoir pour conséquence de limiter la variation de Δp et permettre que les nouveaux paramètres restent dans des domaines de valeurs physiquement admissibles.

Afin de trouver la meilleure valeur de λ possible voici, toujours en suivant [119], comment nous procédons :

Le point de départ est un jeu de paramètres p_0 qui donne une fonction d'onde Ψ_0 . A l'issu d'une simulation VMC permettant d'obtenir les matrices G et H échantillonnées suivant la densité de probabilité $\frac{|\Psi_0|^2}{\int |\Psi_0|^2}$, nous proposons une valeur λ_0 de λ . Cette valeur est arbitraire. Empiriquement, nous la choisissons à la première itération égale à 0.01. Nous résolvons alors l'équation 3.3 pour trois Hessiens \tilde{H} différents : \tilde{H}_1 , \tilde{H}_2 et \tilde{H}_3 , obtenus en ajoutant à la diagonale de la matrice H respectivement $\lambda_1 = \frac{\lambda}{10}$, $\lambda_2 = \lambda$, $\lambda_3 = 10\lambda$. Nous obtenons alors trois jeux de paramètres p_1, p_2, p_3 . Ces trois jeux de paramètres permettent d'obtenir trois nouvelles fonctions d'onde $\Psi(\lambda_1)$, $\Psi(\lambda_2)$, $\Psi(\lambda_3)$ qui donneront trois énergies : E_1, E_2, E_3 . La comparaison de ces trois énergies permettra d'établir une énergie minimale pour laquelle correspondra un λ_{opt} optimal. Pour obtenir ces énergies, nous réalisons une simulation VMC "courte". Par courte, nous entendons 10 à 50 fois plus courte que la précédente. Cela nous permet de produire un échantillonnage corrélé.

Echantillonnage corrélé par pondération

Cette simulation est effectuée suivant la densité de probabilité, $\frac{|\Psi_0|^2}{\int |\Psi_0|^2}$. Nous calculons les trois quantités :

$$E_i = \frac{\langle w_i E_L \rangle}{\langle w_i \rangle} \quad (3.22)$$

avec

$$E_{Li} = \frac{H\Psi(\lambda_i)}{\Psi(\lambda_i)}$$

et

$$w_i = \frac{\Psi(\lambda_i)^2}{\Psi_0^2}$$

pour $i = 1, 3$. Nous utilisons un échantillonnage corrélé car nous nous intéressons aux différences relatives entre E_1 , E_2 , et E_3 . En effet, comme nous l'avons vu plus haut, estimer une différence entre énergies dans les approches stochastiques est difficile à cause de la présence des erreurs statistiques respectives qui peuvent s'ajouter. Sans échantillonnage corrélé et pondération il faudrait donc réaliser de très longues simulations pour réduire ces barres d'erreurs.

L'échantillonnage corrélé par pondération n'est possible que quand on échantillonne des quantités selon des densités de probabilité assez proches de la densité de probabilité de référence : $\frac{|\Psi_0|^2}{\int |\Psi_0|^2}$. Pour mesurer cette différence qui doit toujours rester petite, nous introduisons comme dans la section 3.1 les quantités

$$\eta_i = \frac{\langle w_i \rangle^2}{\langle w_i^2 \rangle}$$

pour $i = 1, 3$ et nous définissons comme critère d'acceptabilité de nos quantités

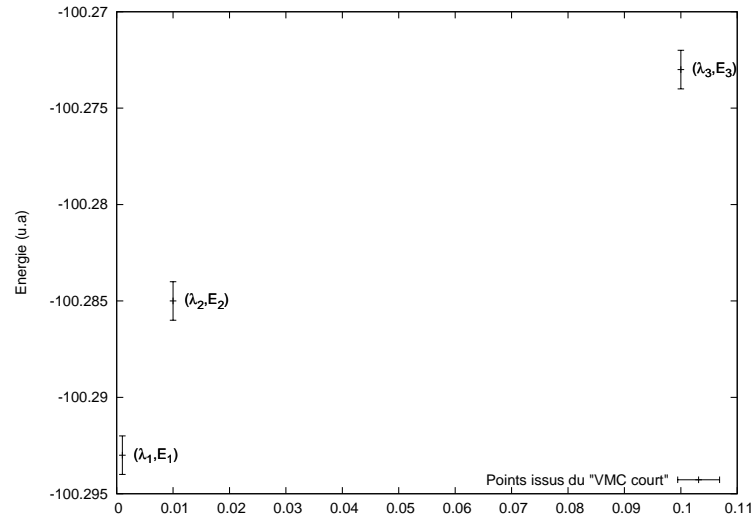
$$\eta_i \gtrsim 0.9$$

Notons que η est toujours compris dans l'intervalle $[0,1]$. Il reste maintenant à comparer E_1 , E_2 et E_3

Comparaison des énergies

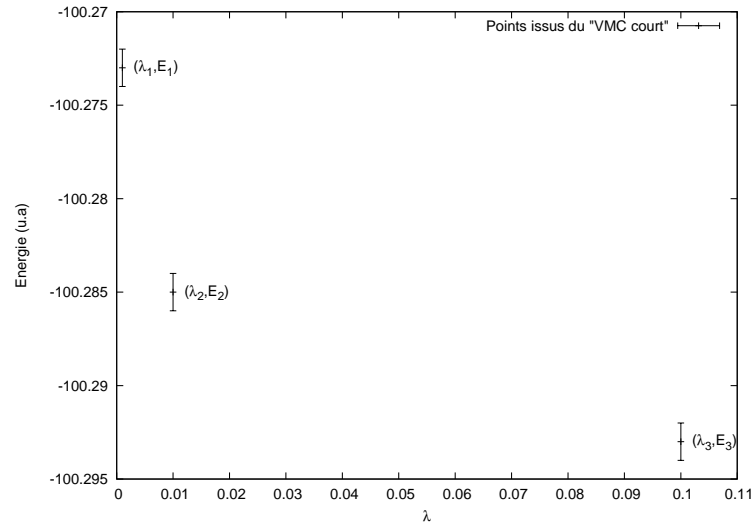
Il existe quatre cas possibles :

- Cas 1 : $E_1 < E_2 < E_3$



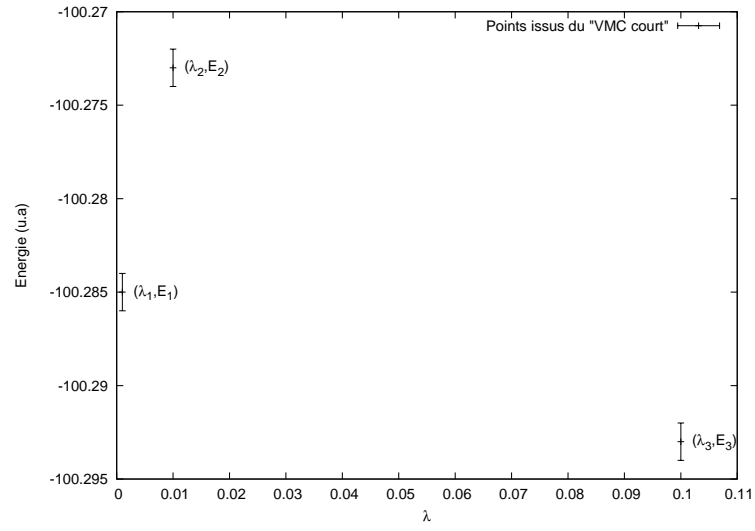
Nous prenons $\lambda_{opt} = \lambda_1$ et pour l'itération suivante, $\lambda_{old} = \lambda_{opt}$.

- Cas 2 : $E_1 > E_2 > E_3$



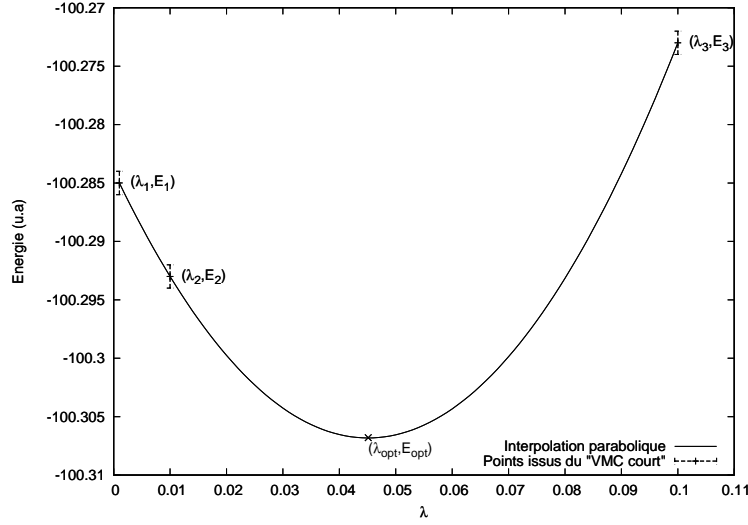
Nous prenons $\lambda_{opt} = \lambda_3$ et pour l'itération suivante, $\lambda_{old} = \lambda_{opt}$.

- Cas 3 : $E_1 < E_2$ et $E_2 > E_3$



Nous prenons par sécurité $\lambda_{opt} = \lambda_3$ et pour l'itération suivante, $\lambda_{old} = \lambda_{opt}$.

- Cas 4 : $E_1 > E_2$ et $E_2 < E_3$



Nous utilisons une interpolation parabolique afin d'obtenir λ_{opt} . Nous construisons par interpolation de Lagrange le polynôme du deuxième degré $P(\lambda)$ passant par les trois points (λ_1, E_1) , (λ_2, E_2) , et (λ_3, E_3) .

$$\begin{aligned}
 P(\lambda) = & \\
 & E_1 \frac{(\lambda - \lambda_2)(\lambda - \lambda_3)}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)} \\
 & + E_2 \frac{(\lambda - \lambda_1)(\lambda - \lambda_3)}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)} \\
 & + E_3 \frac{(\lambda - \lambda_1)(\lambda - \lambda_2)}{(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2)} \quad (3.23)
 \end{aligned}$$

λ_{opt} est choisi comme le minimum du polynôme, c'est à dire tel que

$$P'(\lambda_{opt}) = 0$$

où P' est la dérivée du polynôme suivant λ . Nous conservons pour l'itération suivante $\lambda_0 = \lambda_{opt}$.

Finalement, le choix de λ_{opt} étant établi, nous posons

$$\tilde{H}_{opt} = H + \lambda_{opt} Id$$

et nous résolvons l'équation 3.3 pour \tilde{H}_{opt} afin d'obtenir à l'issu de cette itération un nouveau jeu de paramètre p_{new} .

3.3.6 Résumé des étapes d'optimisation pour la méthode de Newton

Pour une optimisation :

- Choix d'une fonction d'onde Ψ_0 avec un jeu de paramètres p_{old}
- Calcul VMC pour obtenir les quantités nécessaires à la méthode de Newton
- Obtention de la matrice Hessienne H_0 et du gradient G_0
- Ajout des coefficients $\frac{\lambda}{10}$, λ , 10λ (avec $\lambda = 0.01$). à la diagonale de la matrice H_0 pour obtenir trois matrices : H_1 , H_2 , H_3
- Résolution du problème $p_{new} = p_{old} - H_i^{-1}G_0$ pour $i = 1, 2, 3$
- Obtention de trois jeux de paramètres p_1 , p_2 , p_3 issus des matrices H_1 , H_2 , H_3
- Calcul VMC "court" et calcul des quantités corrélées E_1 , E_2 , et E_3
- Interpolation parabolique et obtention de λ_{opt}
- Résolution du problème $p_{new} = p_{old} - H_{\lambda_{opt}}^{-1}G_0$
- Obtention des paramètres p_{new}
- $p_{new} = p_{old}$ et on recommence l'itération.

3.4 Méthode linéaire

3.4.1 Description théorique

La méthode d'optimisation linéaire repose sur un développement de Taylor à l'ordre 1 de la fonction d'onde Ψ_0 par rapport aux différents paramètres p à optimiser puis à la recherche des paramètres optimaux qui minimisent l'énergie variationnelle :

$$E(\Psi) = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (3.24)$$

Au voisinage des paramètres p_0 , la fonction d'onde Ψ est développée linéairement sous la forme :

$$\Psi_{lin} = \Psi_0 + \sum_{i=1}^{N_{opt}} \Delta p_i \Psi_i \quad (3.25)$$

avec $\Delta p_i = p_i - (p_0)_i$ [$(p_0)_i$ est le $i^{\text{ème}}$ paramètre du jeu p_0] $\Psi_i = \frac{\partial \Psi}{\partial c_i}$ et N_{opt} est le nombre de paramètres à optimiser. Optimiser la fonction d'onde suivant l'énergie revient donc à chercher les composantes $(\Delta p_i)_{i=1, N_{opt}}$ de Ψ dans la base (Ψ_0, Ψ_1, \dots) qui minimise E dans l'expression :

$$H\Psi = E\Psi \quad (3.26)$$

qui se réécrit également :

$$H(\Psi_0 + \sum_i \Delta p_i \Psi_i) = E(\Psi_0 + \sum_i \Delta p_i \Psi_i) \quad (3.27)$$

En projetant sur les composantes de la base (Ψ_0, Ψ_1, \dots) , nous obtenons le système :

$$\left\{ \begin{array}{l} \langle \Psi_0 H \Psi_0 \rangle + \sum_i \Delta p_i \langle \Psi_0 H \Psi_i \rangle = E(\langle \Psi_0 \Psi_0 \rangle + \sum_i \Delta p_i \langle \Psi_0 \Psi_i \rangle) \\ \langle \Psi_1 H \Psi_0 \rangle + \sum_i \Delta p_i \langle \Psi_1 H \Psi_i \rangle = E(\langle \Psi_1 \Psi_0 \rangle + \sum_i \Delta p_i \langle \Psi_1 \Psi_i \rangle) \\ \dots \\ \langle \Psi_{N_{opt}} H \Psi_0 \rangle + \sum_i \Delta p_i \langle \Psi_{N_{opt}} H \Psi_i \rangle = E(\langle \Psi_{N_{opt}} \Psi_0 \rangle + \sum_i \Delta p_i \langle \Psi_{N_{opt}} \Psi_i \rangle) \end{array} \right. \quad (3.28)$$

Soit, en réécrivant les éléments de matrice sous une forme adaptée aux simulations VMC :

$$\left\{ \begin{array}{l} \langle \frac{\Psi_0}{\Psi_0} \frac{H\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_0}{\Psi_0} \frac{H\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} = E \langle \frac{\Psi_0}{\Psi_0} \frac{\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_0}{\Psi_0} \frac{\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \langle \frac{\Psi_1}{\Psi_0} \frac{H\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_1}{\Psi_0} \frac{H\Psi_i}{\Psi_i} \rangle_{|\Psi_0|^2} = E \langle \frac{\Psi_1}{\Psi_0} \frac{\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_1}{\Psi_0} \frac{\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \dots \\ \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{H\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{H\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} = E \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} \end{array} \right. \quad (3.29)$$

Posons

$$H = \begin{pmatrix} \langle \frac{\Psi_0}{\Psi_0} \frac{H\Psi_0}{\Psi_0} \rangle_{|\Psi_0|^2} & \langle \frac{\Psi_0}{\Psi_0} \frac{H\Psi_1}{\Psi_0} \rangle_{|\Psi_0|^2} & \dots & \langle \frac{\Psi_0}{\Psi_0} \frac{H\Psi_{N_{opt}}}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \langle \frac{\Psi_1}{\Psi_0} \frac{H\Psi_0}{\Psi_0} \rangle_{|\Psi_0|^2} & \langle \frac{\Psi_1}{\Psi_0} \frac{H\Psi_1}{\Psi_0} \rangle_{|\Psi_0|^2} & \dots & \langle \frac{\Psi_1}{\Psi_0} \frac{H\Psi_{N_{opt}}}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \vdots & \vdots & \vdots & \vdots \\ \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{H\Psi_0}{\Psi_0} \rangle_{|\Psi_0|^2} & \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{H\Psi_0}{\Psi_0} \rangle_{|\Psi_0|^2} & \dots & \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{H\Psi_{N_{opt}}}{\Psi_0} \rangle_{|\Psi_0|^2} \end{pmatrix} \quad (3.30)$$

$$S = \begin{pmatrix} \langle \frac{\Psi_0}{\Psi_0} \frac{\Psi_0}{\Psi_0} \rangle_{|\Psi_0|^2} & \langle \frac{\Psi_0}{\Psi_0} \frac{\Psi_1}{\Psi_0} \rangle_{|\Psi_0|^2} & \cdots & \langle \frac{\Psi_0}{\Psi_0} \frac{\Psi_{N_{opt}}}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \langle \frac{\Psi_1}{\Psi_0} \frac{\Psi_0}{\Psi_0} \rangle_{|\Psi_0|^2} & \langle \frac{\Psi_1}{\Psi_0} \frac{\Psi_1}{\Psi_0} \rangle_{|\Psi_0|^2} & \cdots & \langle \frac{\Psi_1}{\Psi_0} \frac{\Psi_{N_{opt}}}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \vdots & \vdots & \vdots & \vdots \\ \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{\Psi_0}{\Psi_0} \rangle_{|\Psi_0|^2} & \langle \frac{\Psi_{N_{opt}}}{\Psi_1} \frac{\Psi_0}{\Psi_0} \rangle_{|\Psi_0|^2} & \cdots & \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{\Psi_{N_{opt}}}{\Psi_0} \rangle_{|\Psi_0|^2} \end{pmatrix} \quad (3.31)$$

$$\Delta p = \begin{pmatrix} 1 \\ \Delta p_1 \\ \vdots \\ \Delta p_{N_{opt}} \end{pmatrix} \quad (3.32)$$

Résoudre ce problème d'optimisation revient donc à résoudre le problème aux valeurs propres généralisé suivant :

$$H\Delta p = ES\Delta p \quad (3.33)$$

3.4.2 Résolution d'un problème aux valeurs propres généralisé

Etant données deux matrices H et S , résoudre le problème aux valeurs propres signifie trouver les valeurs propres λ et les vecteurs propres associés u_λ qui satisfassent à l'équation matricielle :

$$Hu_\lambda = \lambda Su_\lambda \quad (3.34)$$

Dans notre cas, nous cherchons plus exactement la plus petite valeur propre réelle et le vecteur propre associé. Notons que la matrice de recouvrement S , Eq.(3.31), est symétrique mais pas la matrice Hamiltonienne H , Eq.(3.30), qui pour un temps de simulation VMC *fini* ne vérifie pas $H_{ij} = H_{ji}$ à cause du bruit statistique (cette égalité est obtenue uniquement pour un temps de simulation infini). Nous sommes donc conduit à résoudre un problème aux valeurs propres généralisé et non-symétrique. Pour se faire nous avons utilisé la routine DGGEV (ainsi que les subroutine appelées par DGGEV) de la bibliothèque Lapack[133]. L'algorithme de résolution se base sur la décomposition généralisée de Schur [134] :

Etant données deux matrices carrées H et S de taille $n \times n$, il existe deux matrices unitaires Q et Z et deux matrices triangulaires supérieures T et U telles que :

$$H = QTZ^T \quad (3.35)$$

et

$$S = QUZ^T. \quad (3.36)$$

Résoudre le problème aux valeurs propres pour (H, S) :

$$QTZ^T u_\lambda = \lambda QUZ^T u_\lambda \quad (3.37)$$

est équivalent à résoudre le problème au valeurs propres pour (T, U) :

$$Tu_\lambda = \lambda Uu_\lambda \quad (3.38)$$

Plus précisément, les valeurs de λ sont données par les n rapports :

$$\frac{T_{ii}}{U_{ii}}$$

La réduction généralisée de Schur sur H et S s'effectue en deux étapes :

- Une décomposition supérieure de Hessenberg des matrices H et S :

$$H = VAW^t \quad (3.39)$$

$$S = VBW^t \quad (3.40)$$

où V et W sont des matrices orthogonales et A et B sont des matrices de Hessenberg supérieures

- une décomposition généralisée de Schur sur les matrices A et B . Cette décomposition est facilitée par le fait que A et B sont des matrices de Hessenberg grâce à la première étape

Une matrice de Hessenberg est une matrice B “presque” triangulaire supé-

rieure. Elle est plus exactement de la forme :

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} & \dots & b_{1(n-1)} & b_{1n} \\ b_{21} & b_{22} & b_{23} & \dots & b_{2(n-1)} & b_{2n} \\ 0 & b_{32} & b_{33} & \dots & b_{3(n-1)} & b_{3n} \\ 0 & 0 & b_{43} & \dots & b_{4(n-1)} & b_{4n} \\ 0 & 0 & 0 & \dots & b_{2(n-1)} & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & b_{(n-1)(n-1)} & b_{(n-1)n} \\ 0 & 0 & 0 & \dots & b_{n(n-1)} & b_{2n} \end{pmatrix}$$

3.4.3 Estimateurs

Nous avons donné dans la partie précédente les estimateurs :

- $S_{ij} = \langle \frac{\Psi_i \Psi_j}{\Psi_0 \Psi_0} \rangle$
- $H_{ij} = \langle \frac{\Psi_i}{\Psi_0} \frac{H \Psi_j}{\Psi_0} \rangle$

car nous avons développé Ψ suivant la base (Ψ_0, Ψ_1, \dots) . Mais nous aurions très bien pu effectuer une rotation de la partie (Ψ_1, \dots) pour une obtenir une nouvelle base sur laquelle développer Ψ , c'est ce que nous ferons un peu plus loin.

Comme nous venons de le noter plus haut, la matrice H_{ij} n'est pas symétrique et le devient uniquement dans la limite du nombre de pas Monte Carlo tendant vers l'infini. Il est donc tentant de vouloir symétriser cette matrice, en remplaçant par exemple les éléments H_{ij} par $\frac{H_{ij} + H_{ji}}{2}$. Mais cela serait une erreur car comme nous allons le voir la non-symétrie de la matrice permet de réaliser un principe de zéro-variance "fort" [135].

3.4.4 Renormalisation

La valeur des paramètres Δp dépend de la renormalisation que l'on impose à Ψ_{lin} . Pour améliorer encore l'efficacité de l'algorithme, il a été proposé [121] que les paramètres non-linéaires soient renormalisés :

$$\Delta p_{\text{renorm}} = \frac{\Delta p}{1 - \sum_{i=1}^{N_{opt}} N_i \Delta p_i} \quad (3.41)$$

avec

$$N_i = S_{i0} \quad (3.42)$$

pour les paramètres linéaires (c'est à dire ici les coefficients des déterminants)

et

$$N_i = -\frac{(1-\zeta)\sum_{j=1}^{N_{nonlin}} \Delta p_j S_{ij}}{(1-\zeta) + \zeta \sqrt{(1 + \sum_{j,k, \text{param.nonlin}} \Delta p_j \Delta p_k S_{jk})}} \quad (3.43)$$

pour les paramètres non-linéaires de la fonction d'onde. Cette dernière expression vient de la condition imposée aux dérivées de la fonction d'onde par rapport au paramètre nonlinéaires Ψ_i . Nous imposons en effet que ces dérivées soient orthogonales à une combinaison linéaire de Ψ_0 et Ψ_{lin} soit :

$$\langle \Psi_i | \zeta \Psi_0 + (1-\zeta) \Psi_{lin} \rangle = 0 \quad (3.44)$$

qui mène à la relation ci-dessus. Il y a plusieurs choix pour ζ , nous avons respecté le choix proposé dans [121] de $\zeta = \frac{1}{2}$ qui impose $\|\Psi_{lin}\| = \|\Psi_0\|$

3.4.5 Formules utilisées

Afin d'abaisser les fluctuations, la renormalisation est effectuée *a priori*. Les matrices H et S subissent une rotation. Les éléments de la matrice de saut deviennent donc :

$$S_{00} = 1$$

Pour $i, j > 0$

$$S_{i0} = S_{0j} = 0$$

et

$$S_{ij} = \left\langle \frac{\Psi_i \Psi_j}{\Psi_0 \Psi_0} \right\rangle - \left\langle \frac{\Psi_i}{\Psi_0} \right\rangle \left\langle \frac{\Psi_j}{\Psi_0} \right\rangle$$

Les éléments de l'Hamiltonien deviennent :

$$H_{00} = \langle E_L \rangle$$

Pour $i, j > 0$:

$$H_{i0} = \left\langle \frac{\Psi_i}{\Psi_0} E_L \right\rangle - \left\langle \frac{\Psi_i}{\Psi_0} \Psi_0 \right\rangle \langle E_L \rangle$$

$$H_{0j} = [\langle \frac{\Psi_j}{\Psi_0} E_L \rangle - \langle \frac{\Psi_j}{\Psi_0} \Psi_0 \rangle \langle E_L \rangle] + \langle E_{L,j} \rangle$$

et

$$\begin{aligned} H_{ij} &= [\langle \frac{\Psi_i}{\Psi_0} \frac{\Psi_j}{\Psi_0} E_L \rangle - \langle \frac{\Psi_i}{\Psi_0} \rangle \langle \frac{\Psi_j}{\Psi_0} E_L \rangle - \langle \frac{\Psi_j}{\Psi_0} \rangle \langle \frac{\Psi_i}{\Psi_0} E_L \rangle \\ &+ \langle \frac{\Psi_i}{\Psi_0} \rangle \langle \frac{\Psi_j}{\Psi_0} \rangle \langle E_L \rangle] + [\langle \frac{\Psi_i}{\Psi_0} E_{L,j} \rangle - \langle \frac{\Psi_i}{\Psi_0} \rangle \langle E_{L,j} \rangle] \end{aligned}$$

On retrouve le même principe qui consiste à soustraire $\langle a \rangle \langle b \rangle$ à $\langle ab \rangle$ afin de diminuer la variance.

3.4.6 Schéma de stabilisation

A des fins de stabilisation, l'algorithme introduit à nouveau un paramètre λ ajouté sur la diagonale de H à l'exception du premier élément correspondant la fonction d'onde Ψ_0 .

Pour $i > 0$

$$H_{ii} \rightarrow H_{ii} + \lambda$$

A la différence de l'algorithme de Newton, le choix de λ est important mais s'avère moins crucial pour la stabilisation de l'algorithme. Différentes stratégies peuvent être employées pour trouver un paramètre λ satisfaisant. Voici celle que nous avons proposée de notre côté.

Reprenons le système d'équations qui décrit le problème. En additionnant les colonnes les unes avec les autres nous obtenons le système :

$$\left\{ \begin{array}{l} \langle \frac{\Psi_0}{\Psi_0} \frac{H\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_0}{\Psi_0} \frac{H\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} = E \langle \frac{\Psi_0}{\Psi_0} \frac{\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_0}{\Psi_0} \frac{\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \langle \frac{\Psi_1}{\Psi_0} \frac{H\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_1}{\Psi_0} \frac{H\Psi_i}{\Psi_i} \rangle_{|\Psi_0|^2} = E \langle \frac{\Psi_1}{\Psi_0} \frac{\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_1}{\Psi_0} \frac{\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \dots \\ \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{H\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{H\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} = E \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} \end{array} \right. \quad (3.45)$$

En rajoutant le paramètre λ nous obtenons :

$$\left\{ \begin{array}{l} \langle \frac{\Psi_0}{\Psi_0} \frac{H\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_0}{\Psi_0} \frac{H\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} = E \langle \frac{\Psi_0}{\Psi_0} \frac{\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_0}{\Psi_0} \frac{\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \langle \frac{\Psi_1}{\Psi_0} \frac{H\Psi_0}{\Psi_0} + \sum_i (\Delta p_i \frac{\Psi_1}{\Psi_0} \frac{H\Psi_i}{\Psi_i}) + \lambda \Delta p_1 \frac{\Psi_1}{\Psi_0} \frac{\Psi_1}{\Psi_0} \rangle_{|\Psi_0|^2} = E \langle \frac{\Psi_1}{\Psi_0} \frac{\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_1}{\Psi_0} \frac{\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} \\ \dots \\ \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{H\Psi_0}{\Psi_0} + \sum_i (\Delta p_i \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{H\Psi_i}{\Psi_i}) + \lambda \Delta p_{N_{opt}} \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{\Psi_{N_{opt}}}{\Psi_0} \rangle_{|\Psi_0|^2} = E \langle \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{\Psi_0}{\Psi_0} + \sum_i \Delta p_i \frac{\Psi_{N_{opt}}}{\Psi_0} \frac{\Psi_i}{\Psi_0} \rangle_{|\Psi_0|^2} \end{array} \right. \quad (3.46)$$

Nous multiplions à présent chaque ligne L_i par Δp_i .

En posant

$$\Psi_{lin} = \Psi_0 + \sum_{i=1, N_{opt}} \Delta p_i \Psi_i$$

et en sommant toutes ces lignes nous obtenons :

$$\langle \frac{\Psi_{lin}^2}{\Psi_0^2} \frac{H\Psi_{lin}}{\Psi_{lin}} \rangle_{|\Psi_0|^2} = E(\lambda) - \lambda \frac{\sum_{i=1, N_{opt}} (\Delta p_i)^2 \langle \frac{\Psi_i^2}{\Psi_0^2} \rangle_{|\Psi_0|^2}}{|\Psi_{lin}|^2} \quad (3.47)$$

Nous sommes donc amenés à trouver le paramètre λ qui minimise la quantité :

$$E(\lambda) - \lambda \frac{\sum_{i=1, N_{opt}} (\Delta p_i)^2 \langle \frac{\Psi_i^2}{\Psi_0^2} \rangle_{|\Psi_0|^2}}{|\Psi_{lin}|^2}$$

Nous constatons par ailleurs que cette quantité permet de déterminer :

$$\langle \frac{\Psi_{lin}^2}{\Psi_0^2} \frac{H\Psi_{lin}}{\Psi_{lin}} \rangle_{|\Psi_0|^2}$$

c'est à dire la quantité

$$\frac{H\Psi_{lin}}{\Psi_{lin}}$$

obtenue par échantillonnage pondéré suivant la trajectoire construite à partir

de la densité $|\Psi_0|^2$ avec un poids :

$$w^2 = \frac{\Psi_{lin}^2}{\Psi_0^2}$$

Nous calculons donc la quantité

$$\eta = \frac{\langle w^2 \rangle_{|\Psi_0|^2}^2}{\langle w^4 \rangle_{|\Psi_0|^2}}$$

pour évaluer la confiance en notre échantillonnage. En posant ainsi le problème, nous voyons que nous sommes amenés à trouver les paramètres Δp_i qui minimisent $\langle \frac{H\Psi_{lin}}{\Psi_{lin}} \rangle$. L'échantillonnage ne pouvant bien évidemment s'effectuer selon les trajectoires associées à $|\Psi_{lin}|^2$ puisque celle-ci est inconnue *a priori*, elle s'effectue donc par échantillonnage pondéré via la trajectoire connue de $|\Psi_0|^2$. Toutefois nous raisonnons sur Ψ_{lin} qui reste une approximation linéarisée de notre fonction d'onde (les paramètres à optimiser étant majoritairement dans l'exponentielle du ou des termes de Jastrow). Cela permet de comprendre que malgré une minimisation de la quantité $\langle \frac{H\Psi_{lin}}{\Psi_{lin}} \rangle$, les nouveaux paramètres peuvent ne pas correspondre dans le cas où l'approximation linéaire n'est plus valable. Néanmoins, en pratique nous n'utiliserons pas cette formule. Le calcul de η nécessite en effet le stockage des quantités $\Psi_i\Psi_j\Psi_k\Psi_l$ qui peut occuper trop d'espace. Nous nous contenterons de donner une valeur arbitraire à λ comme développé ci-dessous. La robustesse de cet algorithme par rapport à l'algorithme de Newton permet de faire de ce choix de λ une question moins importante. C'est bien cette question de stabilité de l'algorithme qui nous a fait préférer l'algorithme linéaire à l'algorithme de Newton.

3.4.7 Résumé des étapes d'optimisation pour l'algorithme linéaire

Afin de clarifier les idées synthétisons ici les étapes de l'algorithme linéaire.

- Calcul des éléments des matrices de H et S à l'aide d'une simulation VMC.
- Résolution du problème aux valeurs propres généralisé.
- Renormalisation des paramètres.

3.4.8 Déroutement pratique d'une optimisation linéaire

Nous consignons ici les règles purement empiriques que nous avons dégagées lors de nos différentes optimisations. En effet, malgré l'apparent caractère "automatique" des algorithmes développés ci-dessus, l'optimisation reste une affaire de doigté et de stratégies qui permettent de faciliter cette optimisation. Nous jouons ainsi sur :

- la valeur du paramètre λ
- le nombre de pas Monte Carlo
- le choix et l'ordre des paramètres

Nous regardons comme critères qui nous permettent d'évaluer la justesse de notre optimisation :

- les variations de l'énergie
- les variations des paramètres
- la valeur des gradients

Ainsi, nous concluons à la justesse de notre optimisation lorsque l'énergie et les paramètres ne varient plus et les gradients sont nuls (à l'intérieur des barres d'erreurs). Evidemment, tout ceci revêt un aspect "recette de cuisine" mais c'est le propre de toutes les stratégies d'optimisation d'un grand nombre de paramètres pour les systèmes complexes (de manière très générale, l'optimisation n'est un problème polynomial et on ne peut pas espérer définir un algorithme universel qui marcherait à tous les coups sans réfléchir). Cette procédure n'est bien sûr pas la seule possible. Elle a été mise au point suite à l'expérience acquise lors de différents essais et erreurs lors des optimisations.

Chapitre 4

La fonction d'onde d'essai multi-Jastrow

4.1 Généralités : Rôle de la fonction d'onde d'essai dans les méthodes Monte Carlo quantique

Comme nous l'avons vu précédemment, la fonction d'essai Ψ_T est la quantité centrale des simulations VMC et DMC. Dans les simulations VMC, la fonction d'onde représente la fonction à échantillonner (son module au carré plus exactement) et on cherche à ce qu'elle soit une approximation la plus fidèle possible de la fonction d'onde exacte de l'état fondamental.

Dans la méthode DMC, la fonction d'onde peut être considérée comme un auxiliaire très important facilitant la convergence des simulations (à travers l'échantillonnage selon l'importance, "importance sampling"). De plus, en imposant ses zéros à la fonction d'onde-solution, la fonction d'essai va déterminer l'importance de l'erreur dite des nœuds fixés. Pour réduire celle-ci, il faut trouver une fonction d'essai dont les zéros coïncident ou sont les plus proches possibles de ceux de la fonction d'onde exacte. Rappelons qu'il n'y a aucune stratégie pour trouver *a priori* les nœuds de la fonction d'onde exacte. Concrètement, nous choisissons des fonctions d'essai et après avoir effectué les simulations DMC, nous estimons l'erreur des nœuds fixés (en

la comparant avec des calculs de référence). Néanmoins, un principe heuristique suggère que, plus on injecte d'informations physiques pertinentes dans la fonction d'essai, plus on s'attend à une réduction importante de l'erreur des nœuds fixés.

La fonction d'essai joue donc un rôle crucial dans les méthodes Monte-Carlo et tout milite pour qu'elle soit choisie la plus proche possible de la fonction d'onde exacte inconnue. Parmi les informations physiques pertinentes, on peut citer :

- l'antisymétrie due au principe d'exclusion de Pauli.
- les conditions de cusp entre les particules.
- les éventuelles symétries spatiales.
- la corrélation électronique longue distance (effets de quasi-dégénérescence) et courte distance (trou de Coulomb)
- etc.

4.1.1 Les conditions de cusp

Les conditions de cusp [136] décrivent les comportements de la fonction d'onde exacte quand deux particules coalescent (électron-électron ou électron-noyau).

Le cusp électron-noyau

Considérons le cas simple de l'équation de Schrödinger pour un noyau et un électron (unités atomiques) :

$$\left[-\frac{1}{2m_n}\nabla_n^2 - \frac{1}{2}\nabla_e^2 - \frac{Z}{r}\right]\Psi = E\Psi \quad (4.1)$$

où m_n est la masse du noyau et Z sa charge. Dans l'approximation de la masse du noyau infiniment grande par rapport à celle de l'électron, l'expression se réduit à :

$$\left[-\frac{1}{2}\nabla_e^2 - \frac{Z}{r}\right]\Psi = E\Psi \quad (4.2)$$

Lorsque r tend vers 0, le potentiel en $-\frac{Z}{r}$ tend vers $-\infty$. Pour que le terme de gauche ne diverge pas, la fonction d'onde Ψ doit donc avoir une forme

particulière pour r petit afin de faire disparaître la singularité. En réécrivant l'équation en coordonnées sphériques nous avons pour la partie radiale ($l = 0$) :

$$-\frac{1}{2}\Psi'' - \frac{1}{r}(Z\Psi + \Psi') = E\Psi \quad (4.3)$$

où les dérivées sont définies par rapport à r . Pour que le terme en $-\frac{Z}{r}$ disparaisse nous devons donc avoir :

$$\frac{\Psi'}{\Psi} = -Z \quad (4.4)$$

Lorsque la fonction d'onde est par exemple de la forme $\Psi = e^{-cr}$, nous devons avoir $c = Z$

Le cusp électron-électron

Dans le cas de deux électrons très proches, nous avons localement l'équation de Schrödinger suivante :

$$\left[-\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 + \frac{1}{r_{12}}\right]\Psi = E\Psi \quad (4.5)$$

Dans le cas de deux électrons de spins différents, la condition de cusp pour une fonction de la forme $\Psi = e^{cr_{12}}$ est donc :

$$c = \frac{1}{2} \quad (4.6)$$

Lorsque les électrons sont de même spin, ils sont dans un état relatif du moment angulaire p . Le cusp est réduit de moitié

$$c = \frac{1}{4} \quad (4.7)$$

4.2 Fonction d'essai usuelle des simulations Monte Carlo quantique

Les fonctions d'essai typiquement utilisées sont de la forme

$$\Psi_T = e^J \sum_k c_k \Psi_k \quad (4.8)$$

On y distingue deux parties :

- une partie constituée d'une combinaison linéaire de déterminants de Slater
- un préfacteur dit de Jastrow[137] qui décrit une partie de la corrélation électronique (corrélation dynamique essentiellement).

4.2.1 Description d'un déterminant de Slater Ψ_k

Ψ_k a la forme d'un déterminant de Slater :

$$\Psi_k = \begin{vmatrix} \Phi_1(\vec{r}_1) & \Phi_1(\vec{r}_2) & \dots & \dots & \Phi_1(\vec{r}_{N_\uparrow}) \\ \Phi_2(\vec{r}_1) & \Phi_2(\vec{r}_2) & \dots & \dots & \Phi_2(\vec{r}_{N_\uparrow}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \times \begin{vmatrix} \Phi_1(\vec{r}_{N_\uparrow+1}) & \Phi_1(\vec{r}_{N_\uparrow+2}) & \dots & \dots & \Phi_1(\vec{r}_{N_e}) \\ \Phi_2(\vec{r}_{N_\uparrow+1}) & \Phi_2(\vec{r}_{N_\uparrow+2}) & \dots & \dots & \Phi_2(\vec{r}_{N_e}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \quad (4.9)$$

N_\uparrow est le nombre d'électrons de spins up, N_e le nombre total d'électrons. Chaque ligne i comporte la même orbitale Φ_i tandis que chaque colonne j comporte le même électron \vec{r}_j .

Notons que nous travaillons dans un formalisme sans-spin ("spin-free"[104]). Dans ce formalisme les électrons de spin up et les électrons de spin down sont considérés comme indépendants et les fonctions d'onde peuvent s'écrire comme un produit de deux déterminants. Le premier déterminant comporte tous les électrons de spin up et le second tous les électrons de spin down.

Les déterminants de Slater sont issus d'un calcul *ab initio* préalable effectué avec un des programmes standards de la chimie quantique (par exemple, Gaussian[138], Gamess[139], Molpro[140] ou Molcas[141] parmi les plus connus). Ce déterminant est monoélectronique comme tout déterminant issu d'un calcul *ab initio* : chaque colonne ne comporte qu'un électron et aucune corrélation explicite entre eux n'est introduite. Chaque électron

ne voit les autres que par le biais d'un champ moyen. Les éléments de matrice sont constitués d'une combinaison linéaire sur une base de fonctions monoélectroniques. Ainsi, l'élément ij du déterminant est de la forme :

$$\Phi_i(\vec{r}_j) = \sum_{k=1}^{N_{basis}} c_{ik} \chi_k(\vec{r}_j) \quad (4.10)$$

Dans les calculs *ab initio*, la base $(\chi_k)_{k=1, N_{basis}}$ est constituée en général de fonctions gaussiennes centrées sur chacun des noyaux, c'est à dire des fonctions de la forme

$$P[(x - R_x), (y - R_y), (z - R_z)] e^{-\alpha(\vec{r} - \vec{R})^2} \quad (4.11)$$

où \vec{R} est la position du noyau considéré, α l'exposant de la gaussienne et $P(x, y, z)$ un polynôme en x, y, z . Dans nos calculs, nous avons utilisé des gaussiennes de type s, p ou d définies de la manière suivante :

Type	Formule
s	$e^{-\alpha r^2}$
p	
p_x	$x e^{-\alpha r^2}$
p_y	$y e^{-\alpha r^2}$
p_z	$z e^{-\alpha r^2}$
d	
x^2	$x^2 e^{-\alpha r^2}$
y^2	$y^2 e^{-\alpha r^2}$
z^2	$z^2 e^{-\alpha r^2}$
xy	$xy e^{-\alpha r^2}$
xz	$xz e^{-\alpha r^2}$
yz	$yz e^{-\alpha r^2}$

Pour une base donnée, les calculs *ab initio* nous permettent d'obtenir les coefficients optimisés c_{ik} des orbitales moléculaires constituant les déterminants de Slater. Dans tout le reste de ce travail, nous n'optimiserons pas ces

derniers. Les fonctions gaussiennes sont utilisées car elles offrent un bon compromis entre une description physique convenable et une facilité d'utilisation pour les calculs. Néanmoins elles décrivent très mal les conditions de cusp électron-noyau. A proximité des noyaux, la divergence créée par les électrons crée de très fortes fluctuations qui empêchent la convergence du calcul. En effet, en dérivant $e^{-\alpha r^2}$ suivant r nous obtenons une fonction du type $r e^{-\alpha r^2}$. Nous ne pouvons de cette manière respecter la condition de cusp 4.4

$$\frac{\Psi'}{\Psi} = -Z$$

Pour remédier à cette difficulté, nous remplaçons les orbitales $1s$ gaussiennes par des fonctions de Slater, c'est à dire des fonctions de la forme $e^{-\alpha r}$. Pour les méthodes QMC, le calcul de fonctions gaussiennes ou Slater est indifférent en terme de coût contrairement aux méthodes *ab initio*.

Pour les autres orbitales, nous "coupons" la fonction à courte distance et la remplaçons à cet endroit par une autre fonction qui décrit convenablement le cusp. La fonction que nous utilisons a la forme suivante :

$$c(1 + ar^2 + br^3)e^{-\gamma r} \tag{4.12}$$

Les paramètres sont ajustés pour satisfaire la continuité avec la fonction coupée et sa dérivée et pour décrire convenablement le cusp à l'origine.

Toutes ces manipulations sur les fonctions sont possibles car les méthodes de Monte-Carlo permettent une grande flexibilité dans la forme que prend la fonction d'onde. L'obligation d'utiliser des déterminants monoélectroniques est une contrainte des calculs *ab initio*. Mais en ce qui concerne les méthodes Monte Carlo quantique, cette condition n'est pas nécessaire. Il n'y pas de contraintes mathématiques sur la structure de la fonction d'onde. Néanmoins comme nous cherchons une fonction d'onde la plus proche possible de la fonction d'onde exacte, il y a des contraintes physiques qui commandent la forme.

4.2.2 Le terme de Jastrow

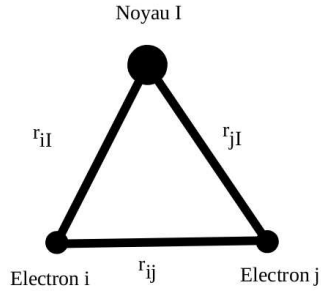
Utiliser un terme de Jastrow est une des originalités des méthodes QMC. Il permet de modéliser à peu de frais la corrélation électronique dynamique, c'est à dire la répulsion des électrons. Dans les méthodes *ab initio* dites "post-Hartree-Fock" (MCSCF, CI[142]) la corrélation dynamique est récupérée en introduisant des sommes de déterminants monoélectroniques. Le terme de Jastrow permet d'éviter l'utilisation de ces déterminants en grand nombre. Cela a pour conséquence avantageuse de rendre les fonctions d'essai plus compactes et plus "controlâbles" : une partie importante de la corrélation dynamique est présente dans le terme de Jastrow et c'est en modifiant ce dernier que l'on agit dessus.

De nombreuses formes ont été proposées ces dernières années[137][83] pour le terme de Jastrow. Nous avons choisi de prendre deux types de Jastrow :

- **Forme minimale pour le terme de Jastrow :**

$$J = \sum_{i < j} \frac{a_{\sigma_{ij}} r_{ij}}{1 + b_{\sigma_{ij}} r_{ij}} - \sum_I \sum_i p_{Ii} r_{iI} \quad (4.13)$$

La somme s'effectue sur les paires d'électrons (i, j) et sur les noyaux I . r_{ij} est la distance entre les électrons i et j et r_{iI} la distance entre l'électron i et le noyau I .



La notation σ_{ij} est introduite pour faire la distinction entre les électrons de même spin et les électrons de spins opposés. Il y a en effet un coefficient a_{like} pour une paire d'électrons (i, j) de même spin et un coefficient a_{unlike} si les électrons i et j sont de spins différents. Il en est de même pour le coefficient b . C'est pourquoi afin d'éviter de faire la distinction à chaque fois, nous regrouperons les coefficients c (ici a et b mais d'autres dans la suite de l'exposé) sous la notation : $c_{\sigma_{ij}}$ pour signifier qu'ils prennent des valeurs différentes suivant que les électrons soient de même spin ou non.

A courte distance électron-électron, le rapport $\frac{a_{\sigma_{ij}} r_{ij}}{1+b_{\sigma_{ij}} r_{ij}}$ s'approche de $a_{\sigma_{ij}} r_{ij}$. Le coefficient $a_{\sigma_{ij}}$ va permettre de décrire ce type d'interaction. La valeur théorique de a_{like} pour des électrons de même spin est $1/4$, et $a_{unlike} = 1/2$. C'est la condition de cusp électron-électron. A longue distance, le rapport $\frac{a_{\sigma_{ij}} r_{ij}}{1+b_{\sigma_{ij}} r_{ij}}$ s'approche de $\frac{a_{\sigma_{ij}}}{b_{\sigma_{ij}}}$, le coefficient

$a_{\sigma_{ij}}$ déterminant exclusivement la corrélation à courte distance, c'est donc le coefficient $b_{\sigma_{ij}}$ qui va décrire ce type d'interaction.

- **Forme plus générale pour le terme de Jastrow :**

$$J = \sum_I \sum_{i < j} [s_{ij}(X_{ij}) - p_I(X_{iI}) - p_I(X_{jI}) + g_1^I X_{iI}^2 X_{jI}^2 + g_2^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2] \quad (4.14)$$

avec

$$X_{ij} = \frac{r_{ij}}{1 + b_{\sigma_{ij}} r_{ij}} \quad (4.15)$$

$$X_{iI} = \frac{r_{iI}}{1 + b_I r_{iI}} \quad (4.16)$$

et

$$s_{ij}(X) = \frac{1}{N_I} (e_{1\sigma_{ij}} X + e_{2\sigma_{ij}} X^2 + e_{3\sigma_{ij}} X^3 + e_{4\sigma_{ij}} X^4) \quad (4.17)$$

$$p_I(X) = \frac{1}{N_e - 1} (f_{1I} X + f_{2I} X^2 + f_{3I} X^3 + f_{4I} X^4) \quad (4.18)$$

où N_I est le nombre de noyaux, N_e le nombre d'électrons. Les e_i se distinguent suivant que les électrons sont de spins opposés ou non. De même pour le coefficient b nucléaire. Les coefficients f_{iI} et b_I ne dépendent que des noyaux I et non du spin.

4.2.3 Interprétation des coefficients

La forme simple (ou minimale) du Jastrow est un cas particulier de la forme complexe. Un Jastrow simple est un Jastrow complexe qui à l'exception des coefficients $e_{1,\sigma_{ij}}$, f_1 et $b_{\sigma_{ij}}$ a tous les autres coefficients nuls. L'utilisation de ces seuls coefficients permet une description minimale suffisante dans une première approche. Les coefficients $(e_k)_{k=1,4}$ permettent de décrire le comportement des électrons entre eux puisqu'ils ne font intervenir que des termes en r_{ij} (distances entre électrons). Les coefficients $(f_k)_{k=1,4}$ permettent la description radiale du comportement d'un électron i avec un noyaux I : ils ne font intervenir que les distances entre un électron et un

noyau. Les coefficients g_1^I et g_2^I mettent en jeu les distances r_{ij} , r_{iI} et r_{jI} . Ils permettent de décrire le comportement radiale des électrons i et j autour du noyau I . Le terme de Jastrow permet de décrire la corrélation à courte distance par le biais des paramètres e_k , f_k , g et à longue distance grâce aux paramètres b_{e-e} et b_{e-n} .

4.3 La nouvelle fonction d’onde proposée : la fonction “multi-Jastrow”

Le terme de Jastrow présenté ci-dessus est placé en facteur de la partie déterminantale et agit sur les électrons de manière tout à fait indépendante de l’orbitale moléculaire qu’ils occupent principalement. Or, physiquement, il est clair que l’interaction électron-électron n’est pas un phénomène localement uniforme. Pour illustrer cela, citons quelques situations contrastées :

- Dans l’orbitale atomique $1s$ les deux électrons sont très proches du noyau et donc très proches l’un de l’autre. L’absence d’écrantage par les autres électrons conduit à un comportement de la fonction d’onde très différent de celui correspondant à la région de valence.
- Dans une liaison chimique locale (par exemple dans une liaison localisée) les électrons ne se voient pas de la même manière que dans un système π délocalisé.
- Dans les métaux de transition, la corrélation électronique dans la couche $3d$ ou $4d$ est très différente de celle intervenant dans les couches s ou p .
- etc.

Afin de combler cette insuffisance nous proposons une nouvelle forme de fonction d’onde basée sur le remplacement du Jastrow global par des termes de Jastrow attachés à tout ou partie des orbitales moléculaires ou atomiques.

Plus précisément, la fonction d'onde proposée a la forme suivante :

$$\Psi_T = \begin{vmatrix} e^{J_1(\vec{r}_1|\vec{r}_2,\dots)}\Phi_1(\vec{r}_1) & e^{J_1(\vec{r}_2|\vec{r}_1,\dots)}\Phi_1(\vec{r}_2) & \dots & \dots & e^{J_1(\vec{r}_{N_\uparrow}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_\uparrow}) \\ e^{J_2(\vec{r}_1|\vec{r}_2,\dots)}\Phi_2(\vec{r}_1) & e^{J_2(\vec{r}_2|\vec{r}_1,\dots)}\Phi_2(\vec{r}_2) & \dots & \dots & e^{J_2(\vec{r}_{N_\uparrow}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_\uparrow}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \\ \times \begin{vmatrix} e^{J_1(\vec{r}_{N_\uparrow+1}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_\uparrow+1}) & e^{J_1(\vec{r}_{N_\uparrow+2}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_\uparrow+2}) & \dots & \dots & e^{J_1(\vec{r}_{N_e}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_e}) \\ e^{J_2(\vec{r}_{N_\uparrow+1}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_\uparrow+1}) & e^{J_2(\vec{r}_{N_\uparrow+2}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_\uparrow+2}) & \dots & \dots & e^{J_2(\vec{r}_{N_e}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_e}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \quad (4.19)$$

Nous donnons ici la forme mono-configurationnelle pour Ψ_T , la généralisation à un ensemble quelconque de configurations étant triviale. Notons que les termes de Jastrow J_k apparaissent maintenant à l'intérieur des déterminants et sont attachés aux différentes orbitales moléculaires Φ_k . Les termes de Jastrow sont écrits de la manière suivante :

$$e^{J_k(\vec{r}_i|\vec{r}_1,\dots,\vec{r}_{N_e})} \quad (4.20)$$

en supposant ici que i est différent de 1 et de N_e . Le premier vecteur-position à la gauche de la barre joue un rôle particulier : il s'agit du vecteur-position correspondant aux coordonnées de l'électron occupant l'orbitale moléculaire associée. A la droite de la barre apparaissent toutes les autres coordonnées ($N_e - 1$ vecteurs) des électrons autres que i (électrons dits secondaires). Afin de préserver l'antisymétrie globale de la fonction d'onde, on vérifie facilement que les J_k doivent être des fonctions complètement symétriques par rapport à l'échange des coordonnées des électrons secondaires :

$$\forall k(\forall i \neq 1)(\forall j \neq 1) \quad e^{J_k(\vec{r}_1|\dots,\vec{r}_i,\dots,\vec{r}_j,\dots)} = e^{J_k(\vec{r}_1|\dots,\vec{r}_j,\dots,\vec{r}_i,\dots)} \quad (4.21)$$

On propose, pour chaque orbitale k , par similtude avec le Jastrow global, des termes de Jastrow du type e^J . Soulignons que cette structure de Jastrow est commune à tous les termes rattachés aux orbitales mais que chaque orbitale

peut avoir un jeu de paramètres différents. Nous avons choisi de considérer des Jastrows J de la forme :

- **Jastrow simple :**

$$J_k(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} \frac{a_k r_{ij}}{1 + b_k r_{ij}} - \sum_I p_{Ik} r_{iI} \quad (4.22)$$

L'électron principal i est fixé et la somme s'effectue sur les électrons secondaires et les noyaux. Le terme $e^{J_k(\vec{r}_i | \dots)}$ va donc représenter l'interaction de l'électron i placé dans l'orbitale k avec l'ensemble des autres électrons et noyaux. Notons que pour conserver la symétrie des J_k , nous ne pouvons plus distinguer entre spins de même type ou non. Le coefficient $\frac{1}{2}$ a été quant à lui introduit pour permettre de retrouver exactement la forme du Jastrow global dans le cas-limite où tous les Jastrows orbitaux sont choisis identiques.

- **Jastrow complexe :**

$$J_k(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} [s_{ij}(X_{ij}) - \sum_I p_{kI}(X_{iI}) - p_I(X_{jI}) + \frac{1}{2} \sum_I \sum_{j \neq i} g_{1ij}^I X_{iI}^2 X_{jI}^2 + g_{2ij}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2] \quad (4.23)$$

avec

$$X_{ij} = \frac{r_{ij}}{1 + b_k r_{ij}} \quad (4.24)$$

$$X_{iI} = \frac{r_{iI}}{1 + b_{kI} r_{iI}} \quad (4.25)$$

et

$$s_{ij}(X) = (e_{1k} X + e_{2k} X^2 + e_{3k} X^3 + e_{4k} X^4) \quad (4.26)$$

$$p_I(X) = (f_{1I_k} X + f_{2I_k} X^2 + f_{3I_k} X^3 + f_{4I_k} X^4) \quad (4.27)$$

4.3.1 Lien avec la fonction d'onde de type "backflow"

Le type de fonction d'onde employée dans la littérature qui se rapproche le plus de notre nouvelle fonction d'onde est la fonction d'onde dite "avec backflow" (Backflow trial wave function). L'idée d'une telle fonction d'onde

est due à Feynman[143] et a été employée assez systématiquement ces dernières années dans plusieurs études numériques sur les propriétés microscopiques des fluides de fermions (gaz électronique homogène, diagramme de phase de He³, diagramme de phase de l’hydrogène, etc.). Une référence parmi d’autres pour ce type de travaux est [144]. L’idée physique derrière la notion de backflow est de décrire la perturbation provoquée par le déplacement d’une particule au sein d’un ensemble de particules. L’approche employée consiste à définir pour chaque particule de nouvelles coordonnées “renormalisées” par l’ensemble des autres particules de la manière suivante :

$$\vec{x}_i = \vec{r}_i + \sum_{j \neq i} \eta(r_{ij})(\vec{r}_i - \vec{r}_j) \quad (4.28)$$

où la somme s’effectue sur l’ensemble de toutes les autres particules et $\eta(r)$ est une fonction de la distance interelectronique à optimiser éventuellement. Une fois définies ces variables, la forme usuelle de la fonction d’essai construite sur des ondes planes est utilisée :

$$\Psi_T = e^J \det_{\uparrow}(e^{k_j x_i}) \det_{\downarrow}(e^{k_j x_i}) \quad (4.29)$$

Notons que, dans le cas de l’absence d’un Jastrow global ($J = 0$) cette fonction d’onde avec backflow est un cas particulier de notre fonction multi-Jastrow avec comme orbitales les ondes planes $e^{k_j \vec{r}}$ et comme Jastrow orbitales les termes complémentaires. Cependant, il est clair que cette fonction n’a pas du tout la grande plasticité que nous offre notre fonction d’onde. En particulier, il n’y a pas dans cette fonction d’onde de Jastrows orbitales qui dépendent de l’orbitale, ce qui est le point fort de notre forme. La fonction ci-dessus est bien adaptée à des systèmes relativement homogènes, comme par exemple le gaz d’électron homogène ou les systèmes apparentés. Dans le cas des systèmes moléculaires complexes, la densité électronique et les propriétés de corrélation sont très inhomogènes. Avec notre fonction d’onde nous pouvons ajouter ou enlever des Jastrows à loisir sur chaque orbitale. Cette modularité est le point fort de l’approche multi-Jastrow.

4.3.2 Quelques remarques importantes sur la forme de la fonction multi-Jastrow

- Une première remarque est que la fonction multi-Jastrow peut être vue comme une généralisation de la forme usuelle qui n'a qu'un seul Jastrow en préfacteur. En effet, il est facile de vérifier que nous retombons sur la forme usuelle lorsque tous les Jastrows orbitaux sont choisis identiques. Notons cependant que nous retrouvons dans cette limite un Jastrow commun un peu particulier pour lequel les coefficients qui dépendaient éventuellement du spin relatif des paires d'électrons n'en dépendent plus maintenant (autrement dit, les coefficients "like" et "unlike" sont égaux). Ceci est dû au fait que les Jastrows individuels ne distinguent pas les électrons up et les électrons down. Si nous voulions introduire une telle dépendance, imposer la contrainte d'antisymétrie globale de la fonction d'onde multi-Jastrow serait beaucoup plus difficile (on tomberait alors dans un problème de type $N!$).
- Grâce à la présence des Jastrows orbitaux, il est désormais possible d'individualiser les comportements et de répondre aux problèmes physiques posés en préambule de cette partie. Un électron occupant une orbitale "voit" les autres électrons d'une certaine manière et cela peut changer suivant l'orbitale où se trouve l'électron. Toutefois, l'individualisation se fait sur l'orbitale même et non sur les électrons. Le Jastrow orbitalaire n'attribue pas une orbitale à chaque électron. Lorsqu'un électron i se trouve dans une orbitale définie, il voit tous les électrons de la même manière. Ces derniers ne sont pas définis par l'orbitale où ils se trouvent. L'interaction avec l'électron i s'effectue par le biais des distances (entre électrons et électrons-noyaux) et à travers les coefficients du Jastrow propre à l'orbitale. L'électron i ne fait pas la différence explicite entre des électrons qui se situeraient dans une même orbitale ou dans des orbitales différentes. Cependant, la distinction se fait indirectement notamment grâce aux coefficients du Jastrow qui permettent

de distinguer entre interactions à courte et à longue distance. Nous parlons de distinctions indirectes car les notions d'appartenance à une orbitale et de distance ne se recouvrent pas totalement. Mais la présence dans le Jastrow orbitalaire des distances électron-noyau en plus des distances électron-électron et leur couplage via les termes g_1 et g_2 donnent un bon comportement suivant la répartition spatiale.

- Nous pourrions bien sûr réfléchir ultérieurement à d'autres formes pour les termes de Jastrow pour capturer encore d'avantage ces subtilités. La forme du Jastrow que nous avons prise est une forme couramment utilisée que nous avons adapté à notre contrainte. Il est tout à fait possible par la suite de modifier cette forme tout en gardant les deux idées clés de la démarche, à savoir :

i.) Un Jastrow orbitalaire introduit à l'intérieur des déterminants de Slater

ii.) Une contrainte sur la symétrie des électrons secondaires dans le facteur de Jastrow pour rendre la structure compatible avec l'antisymétrie globale de la fonction d'onde.

Soulignons tout de même que les Jastrows orbitalaires tels que nous les avons implémentés donnent déjà de très bons résultats et décrivent très bien les situations physiques que nous avons étudiées.

- Une autre remarque extrêmement importante est que l'introduction de ces Jastrows permet de modifier les nœuds de la fonction d'onde. Le Jastrow global étant un préfacteur positif, il n'avait aucune influence sur les zéros de la fonction d'essai. Au contraire, les Jastrows orbitalaires étant introduits directement à l'intérieur de la matrice de Slater vont en changer le déterminant et les zéros dépendront des valeurs des paramètres des différents Jastrows. Les Jastrows orbitalaires offrent donc un moyen de modifier les nœuds et ainsi d'influer sur l'importance de l'erreur des nœuds fixés. C'est un des aspects importants que

nous discuterons dans la partie résultat.

- Une crainte naturelle face à notre fonction d’onde multi-Jastrow est que le coût du calcul de la fonction d’onde et de ses dérivées soit trop lourd. En admettant que l’introduction des Jastrows orbitales soient un plus, s’ils requièrent des calculs beaucoup plus longs que la forme standard, cet avantage pourrait disparaître. Ceci d’autant plus que les paramètres des différents Jastrows orbitales demandent également à être optimisés, procédure elle-même complexe et coûteuse en temps quand le nombre de paramètres est grand. Le problème général de l’optimisation a été présenté dans le chapitre 3. Afin d’éviter cette difficulté nous avons effectué un travail important pour optimiser les différents aspects de l’implémentation pratique de notre proposition dans le code Monte Carlo quantique développé au laboratoire (code QMC=Chem pour “Quantum Monte Carlo for Chemistry”, [122]). Nous développerons plus en détail cette implémentation dans le chapitre 5. L’un des résultats importants au terme de cette implémentation et que l’usage des Jastrows orbitales n’augmente les temps de calculs que modérément. Des précisions sur cet aspect seront données dans l’étude numérique comparative présentée dans le chapitre 8. Les temps de calcul ne sont donc pas un facteur limitatif pour utiliser nos Jastrows orbitales.
- Une critique qui pourrait être faite à notre fonction multi-Jastrow est que les Jastrows orbitales complexifient beaucoup trop le problème d’un point de vue physique. L’élégance et la simplicité de la forme classique laisserait place à une jungle de paramètres qui n’auraient plus de sens physique. En fait, il n’en est rien. Les Jastrow orbitales permettent une très grande flexibilité. Il est tout à fait possible de ne pas mettre de Jastrow sur une orbitale ; ou de mettre le même Jastrow sur plusieurs orbitales à la fois pour décrire de grands ensembles. C’est ce

que nous ferons. En pratique, il y a bien moins de Jastrows différents que d'orbitales moléculaires. Ainsi, par exemple, la description d'un atome léger nécessitera l'usage de deux Jastrows : un pour décrire l'orbitale $1s$ la plus profonde et un pour décrire la région de valence.

- Généralisant cette dernière remarque, il est important de rappeler qu'à plus long terme le but qui motive l'introduction de ces Jastrow orbitales est de construire une bibliothèque de Jastrows préalablement optimisés et définis une fois pour toute et qui permettrait de construire des fonctions d'onde de qualité pour les "grosses" molécules (plusieurs centaines d'atomes) par simple combinaison de ces briques de base. L'idée est d'utiliser ces Jastrows préalablement optimisés pour décrire des phénomènes locaux sur de telles molécules. Par exemple, nous pourrions mettre des Jastrows sur les orbitales $1s$ des atomes (le même Jastrow pour chaque type d'atome), mettre un même Jastrow orbitaire sur chaque liaison double carbone-carbone, mettre un Jastrow spécial pour décrire un groupe fonctionnel particulier, etc. Les Jastrows orbitales offrent donc des perspectives extrêmement prometteuses. L'un des aspects importants de ce travail de thèse a été de régler toutes les difficultés techniques liées à cette introduction. C'est l'objet des deux prochains chapitres que de décrire le travail qui a permis d'avoir désormais à disposition dans le code Qmc=Chem des Jastrows orbitales que nous pouvons optimiser. Cette contrainte technique étant levée, il est à présent possible de se concentrer uniquement sur l'aspect physique du problème pour développer tout le potentiel qu'offrent ces Jastrows.

Chapitre 5

Calcul des dérivées de la fonction d'onde multi-Jastrow

Nous allons dans cette partie présenter en détail comment calculer les différentes quantités nécessaires aussi bien pour les simulations Monte Carlo quantique proprement dites (méthodes VMC et DMC) que pour les algorithmes d'optimisation des paramètres de la fonction d'onde (voir chapitre 3). Toutes ces quantités ont été implémentées dans le code QMC=Chem de notre groupe.[122] Clairement, l'introduction des Jastrows par orbitale entraîne une complexification importante du calcul des différentes dérivées et ceci est censé accroître sensiblement les temps de calcul. L'efficacité que peut apporter l'utilisation de plusieurs Jastrows peut alors se trouver contrebalancer par les temps de calcul nécessaires à leur mise en oeuvre. C'est l'existence d'un compromis acceptable qui permet de déterminer leur utilité pratique. Fort heureusement, comme nous le verrons dans l'étude comparative du chapitre 8, les techniques développées pour mettre en oeuvre nos idées physiques dans le code QMC=Chem permettent de régler cette question : les temps de calcul par rapport à l'utilisation d'un seul Jastrow global ne sont en effet pas drastiquement allongés. Ce résultat important nous permet par la suite dans nos discussions de ne pas considérer la question du temps de calcul comme un point important. Seule la pertinence d'un point de vue physique

sera prise en compte.

Finalement, soulignons que les quantités intervenant dans l'étape d'optimisation de nos fonctions d'onde ont du être calculées avec une assez grande précision par des calculs VMC. Etant donné la précision requise il a fallu paralléliser les méthodes pour pouvoir optimiser de "grands" systèmes. Par exemple, cette parallélisation s'est révélée indispensable pour pouvoir optimiser complètement la fonction d'onde multi-Jastrow de l'atome de cuivre. Je tiens à remercier A. Scemama (Ingénieur de Recherche au laboratoire) qui a mis en place dans le code QMC=Chem la procédure de parallélisation automatique que j'ai beaucoup utilisée.

Rappelons que les quantités nécessaires pour effectuer une simulation Monte-Carlo quantique (VMC ou DMC) sont :

- la fonction d'onde Ψ_T
- le vecteur-dérive (drift) $\frac{\vec{\nabla}\Psi_T}{\Psi_T}$
- l'énergie locale $E_L = \frac{H\Psi_T}{\Psi_T} = -\frac{1}{2}\frac{\Delta\Psi_T}{\Psi_T} + V$

Nous avons donc besoin des quantités élémentaires suivantes :

- la fonction d'onde Ψ_T
- le gradient $\vec{\nabla}\Psi_T$
- le Laplacien $\Delta\Psi_T$

Le gradient $\vec{\nabla}\Psi_T$ est le vecteur constitué des dérivées simples de Ψ_T suivant les trois coordonnées spatiales de chacun des électrons.

Le Laplacien a pour formule :

$$\Delta\Psi_T = \sum_{i=1}^{N_e} \frac{\partial^2\Psi_T}{\partial x_i^2} + \frac{\partial^2\Psi_T}{\partial y_i^2} + \frac{\partial^2\Psi_T}{\partial z_i^2} \quad (5.1)$$

et nécessite donc le calcul des dérivées doubles spatiales de Ψ_T suivant chacune des $3N_e$ coordonnées.

Pour l'optimisation, nous avons besoin des quantités suivantes :

- $\frac{\Psi_{T,i}}{\Psi_T}$ ($\Psi_{T,i}$ est la dérivée partielle de Ψ_T suivant le paramètre d'indice i)

- $\frac{\Psi_{T,ij}}{\Psi_T}$: uniquement pour la méthode de Newton (voir, chapitre 3), $\Psi_{T,ij}$ est la dérivée double de Ψ_T suivant les paramètres d'indices i et j .
- la dérivée de l'énergie locale par rapport à un paramètre d'indice i :

$$E_{L,i} = -\frac{1}{2} \left(\frac{\Delta \Psi_{T,i}}{\Psi_T} - \frac{\Psi_{T,i}}{\Psi_T} \frac{\Delta \Psi_T}{\Psi_T} \right)$$

Ce qui donne les quantités élémentaires suivantes :

- la dérivée partielle de Ψ_T suivant le paramètre d'indice i : $\Psi_{T,i}$
- la dérivée double de Ψ_T suivant les paramètres d'indices i et j : $\Psi_{T,ij}$ (uniquement pour la méthode de Newton)
- la dérivée partielle du Laplacien de la fonction d'onde, $\Delta \Psi_T$, suivant le paramètre d'indice i : $\Delta \Psi_{T,i}$

Voyons maintenant comment calculer efficacement toutes ces quantités. Rappelons l'expression de la fonction d'onde multi-Jastrow :

$$\Psi_T = \begin{vmatrix} e^{J_1(\vec{r}_1|\vec{r}_2,\dots)}\Phi_1(\vec{r}_1) & e^{J_1(\vec{r}_2|\vec{r}_1,\dots)}\Phi_1(\vec{r}_2) & \dots & \dots & e^{J_1(\vec{r}_{N_\uparrow}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_\uparrow}) \\ e^{J_2(\vec{r}_1|\vec{r}_2,\dots)}\Phi_2(\vec{r}_1) & e^{J_2(\vec{r}_2|\vec{r}_1,\dots)}\Phi_2(\vec{r}_2) & \dots & \dots & e^{J_2(\vec{r}_{N_\uparrow}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_\uparrow}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \\ \times \begin{vmatrix} e^{J_1(\vec{r}_{N_\uparrow+1}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_\uparrow+1}) & e^{J_1(\vec{r}_{N_\uparrow+2}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_\uparrow+2}) & \dots & \dots & e^{J_1(\vec{r}_{N_e}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_e}) \\ e^{J_2(\vec{r}_{N_\uparrow+1}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_\uparrow+1}) & e^{J_2(\vec{r}_{N_\uparrow+2}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_\uparrow+2}) & \dots & \dots & e^{J_2(\vec{r}_{N_e}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_e}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \quad (5.2)$$

Les différents déterminants associés à la fonction d'onde et à ses dérivées comportent désormais des éléments formés à partir :

- des orbitales moléculaires Φ_i ainsi que de leurs dérivées,
- des Jastrows orbitales J_k ainsi que de leurs dérivées.

Les quantités Φ_i , $\vec{\nabla}\Phi_i$ et $\Delta\Phi_i$ sont les mêmes que l'on ait un Jastrow unique en commun ou que l'on utilise des Jastrows orbitales. Ainsi, nous n'expliquerons pas ici leur calcul et dans le programme QMC=Chem nous avons réutilisé les sous-programmes de calcul qui étaient déjà implémentés.

Dans tout le travail qui suit, nous n’optimiserons que les paramètres des Jastrows et nous ne toucherons pas aux orbitales Φ_i . Nous ne considérerons donc pas les dérivées des Φ_i suivant un paramètre.

Passons maintenant au calcul des Jastrows orbitales ainsi que de leurs dérivées (spatiales ou suivant un paramètre). Nous verrons que les temps de calcul de ces quantités croissent en $O(N_{Jast})$ où N_{Jast} est le nombre de Jastrows différents. Notons que, dans la pratique, ce nombre est inférieur au nombre d’orbitales que comportent le déterminant (voir le chapitre 6 pour une discussion physique du choix des Jastrows orbitales). Par exemple, nous avons choisi de mettre sur l’atome de cuivre qui comporte 15 orbitales, 3 Jastrows différents ; sur la molécule HF qui comporte 6 orbitales différentes, 1 seul Jastrow de valence ; sur l’oxygène avec 5 orbitales, deux Jastrows.

5.1 Calcul des Jastrows et de leurs dérivées

Nous allons maintenant et dans toute la suite de cette partie nous fixer sur un Jastrow orbitaire précis :

$$J_k(\vec{r}|\dots) \quad (5.3)$$

Nous allons calculer les N_e quantités suivantes :

$$J_k(\vec{r}_1|\dots), J_k(\vec{r}_2|\dots), \dots, J_k(\vec{r}_{N_e}|\dots) \quad (5.4)$$

Nous allons montrer que le temps mis pour calculer chacune de ces N_e quantités est du même ordre de grandeur que celui de l’unique Jastrow commun. Ainsi, chaque Jastrow orbitaire distinct sera calculé en un temps similaire à celui mis pour calculer le Jastrow commun. C’est ce qui explique le scaling en $O(N_{Jast})$. Cette linéarité s’étendra également au calcul des dérivées.

Réécrivons pour la clarté de notre exposé les formules du Jastrow commun, e^{J_c} avec :

$$J_c = \sum_{i < j} P(i, j) + \sum_i Q(i) \quad (5.5)$$

et du Jastrow orbitalaire :

$$J_k(\vec{r}_{i_0} | \dots) = \sum_{j \neq i_0} P(i_0, j) + Q(i_0) \quad (5.6)$$

Dans les deux cas P et Q ont la même structure :

$$P(i, j) = s_{ij}(X_{ij}) + g_1^I X_{iI}^2 X_{jI}^2 + g_2^I [(X_{iI}^2 + X_{jI}^2) X_{ij}^2] \quad (5.7)$$

et

$$Q(i) = -p_I(X_{iI}) \quad (5.8)$$

avec

$$X_{ij} = \frac{r_{ij}}{1 + b_{e-e} r_{ij}} \quad (5.9)$$

$$X_{iI} = \frac{r_{iI}}{1 + b_I r_{iI}} \quad (5.10)$$

et

$$s_{ij}(X) = \frac{1}{N_I} (e_1 X + e_2 X^2 + e_3 X^3 + e_4 X^4) \quad (5.11)$$

$$p_I(X) = \frac{1}{N_e - 1} \sum_I (f_{1iI} X + f_{2iI} X^2 + f_{3iI} X^3 + f_{4iI} X^4) \quad (5.12)$$

N_e , nombre total d'électrons et N_I nombre total de noyaux.

Dans le cas des Jastrows orbitalaires tous les coefficients sont propres au Jastrow lui-même. Dans le cas du Jastrow commun, les coefficients e_i , g_1^I , g_2^I , et b_{e-e} peuvent être différents selon que les électrons sont de même spin ou non. Ces différences ne changeront pas notre raisonnement.

Maintenant, constatons que P est symétrique.

$$P(i, j) = P(j, i) \quad (5.13)$$

Pour calculer les N_e quantités : $J_k(\vec{r}_1 | \dots)$, $J_k(\vec{r}_2 | \dots)$, \dots , $J_k(\vec{r}_{N_e} | \dots)$, il est donc nécessaire et suffisant de calculer uniquement les paires $P(i, j)$ pour $i < j$, ce qui revient à calculer exactement les mêmes paires $P(i, j)$ que dans le cas du Jastrow commun. Le nombre de paires à calculer est :

$$N_P = \frac{(N_e - 1)(N_e - 2)}{2} \quad (5.14)$$

Les quantités $Q(i)$ sont exactement les mêmes dans le cas du Jastrow commun et des Jastrows orbitales. Il y en a :

$$N_Q = N_e \quad (5.15)$$

Dans le cas du Jastrow commun, pour évaluer le Jastrow nous additionnons les N_Q quantités $Q(i)$ et les N_P paires $P(i, j)$. Dans le cas des Jastrows orbitales, pour reconstituer les N_e quantités, nous devons effectuer N_Q additions (chaque élément $Q(i)$ est ajouté à $J_k(\vec{r}_i | \dots)$) et $2N_P$ additions (chaque paire $P(i, j)$ est ajoutée au Jastrow $J_k(\vec{r}_i | \dots)$ et $J_k(\vec{r}_j | \dots)$). Finalement, il reste à évaluer les exponentielles, une seule dans le cas du Jastrow commun, N_e pour les N_e Jastrows orbitales. Les temps de calcul du Jastrow commun et des N_e Jastrows orbitales sont donc du même ordre.

Passons maintenant au calcul des différentes dérivées des Jastrows. Nous donnerons ici les éléments principaux et nous renvoyons à l'annexe E pour les expressions complètes qui sont assez lourdes.

5.1.1 Dérivée spatiale simple du Jastrow

Posons $\partial_i^l T$ la dérivée suivant la $l^{\text{ème}}$ coordonnée de l'électron i d'une quantité T . Pour le Jastrow commun, nous avons :

$$\partial_i^l J_c = \sum_{j < i} \partial_i^l P(i, j) + \sum_k \partial_i^l Q(k) \quad (5.16)$$

La somme des éléments P s'effectue sur toutes les paires (i, j) d'électrons différents.

Pour les Jastrows orbitales il y a deux expressions différentes :

- si la dérivée se fait suivant l'électron i_0 principal :

$$\partial_{i_0}^l J_k(\vec{r}_{i_0} | \dots) = \sum_{j \neq i_0} \partial_{i_0}^l P(i_0, j) + \partial_{i_0}^l Q(i_0) \quad (5.17)$$

L'électron i_0 est fixé. La somme s'effectue sur tous les autres électrons j .

- si la dérivée se fait suivant les coordonnées d'un des électrons secondaires i_1 :

$$\partial_{i_1}^l J_k(\vec{r}_{i_0} | \dots) = \partial_{i_1}^l P(i_0, i_1) \quad (5.18)$$

Il n'y pas dans ce cas de dérivées suivant $Q(i_0)$ car ce dernier terme ne se retrouve dans le déterminant que dans la colonne de l'électron i_0 .

Nous avons comme précédemment une symétrie :

$$\partial_i^l P(i, j) = \partial_i^l P(j, i) \quad (5.19)$$

En remplaçant $P(i, j)$ et $Q(i)$ par $\partial_i^l P(i, j)$, $\partial_i^l Q(i)$, nous avons une analogie dans la structure avec la partie précédente. Nous pouvons donc tenir le même raisonnement. Les temps pour calculer $\partial_i^l J_c$ et les $N_e \partial_i^l J_k$ sont donc du même ordre. Nous profitons par ailleurs d'une antisymétrie de l'expression $s()$ pour éviter de recalculer des éléments identiques :

$$\partial_i^l s_{ij}(X_{ij}) = -\partial_j^l s_{ij}(X_{ij}) \quad (5.20)$$

Ainsi, le nombre initial de paires $\nabla_l P(i, j)$ est $2 \times 3 \times (N_e - 1)^{N_e}$. Pour chaque paire i, j , il y a deux gradients non nuls : ∂_i^l et ∂_j^l qui ont chacun trois coordonnées. Le nombre total de paires à calculer est donc

$$2 \times 3 \times \frac{(N_e - 1)(N_e - 2)}{2} \quad (5.21)$$

soit

$$3 \times (N_e - 1)(N_e - 2) \quad (5.22)$$

5.1.2 Dérivée spatiale double du Jastrow

Posons $\partial_{x_i^l}^2 R$ la dérivée double suivant la $l^{\text{ème}}$ coordonnée de l'électron i de la quantité R .

Pour le Jastrow commun l'expression est :

$$\partial_{2,i}^l J_c = \sum_{j \neq i} \partial_{2,i}^l P(i, j) + \sum_k \partial_{2,i}^l Q(k) \quad (5.23)$$

Pour le Jastrow orbitalaire :

- si la dérivée se fait suivant l'électron i_0 qui est la même que le premier argument du Jastrow :

$$\partial_{2,i_0}^l J_k(\vec{r}_{i_0} | \dots) = \sum_{j \neq i_0} \partial_{2,i_0}^l P(i_0, j) + \partial_{2,i_0}^l Q(i_0) \quad (5.24)$$

- si la dérivée se fait suivant l'électron i_1 différent du premier argument i_0 :

$$\partial_{2,i}^i J_k(\vec{r}_{i_0} | \dots) = \partial_{2,i_0}^l P(i_0, i_1) \quad (5.25)$$

Nous avons :

$$\frac{\partial^2 P(i, j)}{\partial x^2} = \frac{\partial^2 P(j, i)}{\partial x^2} \quad (5.26)$$

Nous retrouvons une symétrie. La structure identique à celles précédentes qui nous permet de tenir le même raisonnement que ci-dessus. Les temps de calcul de $\partial_{2,i}^l J_c$ et des $N_e \partial_{2,i}^l J_k$ sont donc du même ordre.

5.1.3 Dérivée du Jastrow suivant un paramètre

Notons J_{x,c_p} la dérivée du Jastrow (x = commun ou orbitalaire) suivant le paramètre c_p et T_{c_p} la dérivée d'une quantité T .

Pour le Jastrow commun nous avons :

$$J_{c,c_p} = \sum_{i < j} P_{c_p}(i, j) + \sum_i Q_{c_p}(i) \quad (5.27)$$

Pour le Jastrow orbitalaire nous avons :

$$J_{k,c_p}(r_{i_0} | \dots) = \sum_{j \neq i_0} P_{c_p}(i_0, j) + Q_{c_p}(i_0) \quad (5.28)$$

Il y a encore une symétrie. Comme dans la partie consacrée au Jastrow et en remplaçant, $P(i, j)$ et $Q(i)$ par leurs dérivées $P_{c_p}(ij)$ et $Q_{c_p}(i)$ suivant un paramètre c_p , nous montrons que les calculs pour déterminer la dérivée du Jastrow commun suivant un paramètre et la dérivée des N_e Jastrows orbitaux conduisent aux mêmes opérations de base.

5.1.4 Dérivée du gradient et du Laplacien du Jastrow suivant un paramètre

Pour le gradient nous pouvons effectuer exactement le même raisonnement pour obtenir les formules.

Pour le Jastrow commun :

$$\partial_i^l J_{c,c_p} = \sum_{j \neq i} \partial_i^l P_{c_p}(i, j) + \sum_k \partial_i^l Q_{c_p}(k) \quad (5.29)$$

Pour le Jastrow orbitalaire :

- si la dérivée se fait suivant l'électron i_0 de référence :

$$\partial_{i_0}^l J_{k,c_p}(\vec{r}_{i_0} | \dots) = \sum_{j \neq i_0} \partial_{i_0}^l P_{c_p}(i_0, j) + \partial_{i_0}^l Q_{c_p}(i_0) \quad (5.30)$$

- si la dérivée se fait suivant l'un des électrons secondaires i_1 :

$$\vec{\nabla}_{i_1} J_{k,c_p}(\vec{r}_{i_0} | \dots) = \partial_{i_0}^l P_{c_p}(i_0, i_1) \quad (5.31)$$

et pour le Laplacien :

Dans le cas du Jastrow commun :

$$\partial_{x_i}^2 J_{c,c_p} = \sum_{j \neq i} \partial_{x_i}^2 P_{c_p}(i, j) + \sum_i \partial_{x_i}^2 Q_{c_p}(i) \quad (5.32)$$

Dans le cas du Jastrow orbitalaire :

- si la dérivée se fait suivant l'électron de référence i_0 :

$$\partial_{x_{i_0}}^2 J_{k,c_p}(\vec{r}_{i_0} | \dots) = \sum_{j \neq i_0} \partial_{x_{i_0}}^2 P_{c_p}(i_0, j) + \partial_{x_{i_0}}^2 Q_{c_p}(i_0) \quad (5.33)$$

- si la dérivée se fait suivant l'un des électrons secondaires i_1 :

$$\Delta_{i_1} J_{k,c_p}(\vec{r}_{i_0} | \dots) = \partial_{x_{i_0}}^2 P_{c_p}(i_0, i_1) \quad (5.34)$$

Les dérivées des Jastrows suivant les paramètres $e_1, e_2, e_3, e_4, f_1, f_2, f_3, f_4, g_1, g_2$ reposent comme le montrent les formules présentées dans l'annexe E sur des sommes d'éléments déjà calculés pour construire les Jastrows. Cela permet de limiter le nombre d'opérations à effectuer puisque ces quantités ont été stockées au préalable. Seules les dérivées suivant les paramètres b_{e-e} et b_{e-n} demandent de calculer de nouvelles quantités.

5.1.5 Dérivée doubles

Nous renvoyons à l'annexe E pour les formules exactes.

Là encore, comme dans toutes les parties précédentes, la comparaison des structures entre Jastrow commun et Jastrow orbitalaire montre que le nombre d'opérations pour calculer ces deux types de Jastrows est similaire. Cependant, beaucoup de dérivées doubles sont nulles. Parmi celle qui ne le sont pas, nous avons les dérivées doubles suivant les paramètres :

- b_{e-e} et e_i pour $i=1$ à 4
- b_{e-n} et f_i pour $i=1$ à 4
- b_{e-e} et g_i pour $i=1$ à 2
- b_{e-n} et g_i pour $i=1$ à 2

elles sont constituées de sommes de quantités déjà calculées au préalable pour le calcul du Jastrow même. Ces quantités peuvent ainsi être réutilisées à ce propos pour éviter des calculs supplémentaires. Ainsi, seules les dérivées doubles suivant les paramètres :

- b_{e-e} et b_{e-e}
- b_{e-n} et b_{e-n}
- b_{e-e} et b_{e-n}

demandent des calculs supplémentaires.

Nous venons de voir que pour un type de Jastrow orbitalaire, calculer les N_e Jastrows et toutes les dérivées nécessaires ne représente que peu d'opérations supplémentaires par rapport à son équivalent sous la forme d'un Jastrow commun. Bien évidemment, cela n'est vrai que pour le cas d'un seul Jastrow orbitalaire. Comme nous en employons plusieurs, le temps de calcul est ainsi multiplié en conséquence.

Le calcul des éléments de la matrice de Slater représente la partie la plus coûteuse en temps-calcul de la méthode QMC dans le cas où on utilise un Jastrow commun. Dans le cas d'une utilisation de plusieurs Jastrows orbitaires, ce n'est plus tout le temps vrai. Nous développerons ces remarques dans la partie consacrée à l'étude comparative du coût des calculs (chapitre

8). La multiplication des déterminants, notamment dans l'optimisation linéaire rend leur part dans le temps de calcul non négligeable.

5.2 Calcul des déterminants et de leurs dérivées

Nous allons maintenant décrire les techniques mises en œuvre pour calculer les déterminants permettant de construire les quantités nécessaires aux simulations. Rappelons ici que pour permettre une simulation VMC ou DMC, les quantités requises sont :

- la fonction d'onde Ψ_T
- le gradient $\vec{\nabla}\Psi_T$
- l'énergie locale E_L construite grâce essentiellement au Laplacien de Ψ_T .

$$E_L = -\frac{1}{2} \frac{\Delta\Psi_T}{\Psi_T} + V \quad (5.35)$$

L'optimisation linéaire requiert en plus de ces quantités, les dérivées de ces dernières suivant les paramètres du Jastrow. La méthode de Newton nécessite une quantité supplémentaire, à savoir la dérivée double de Ψ_T suivant les paramètres à optimiser. La fonction d'essai Ψ_T étant le produit de deux déterminants (dans le cas mono-configurationnel, le cas multi-configurationnel étant une généralisation triviale) les dérivées de cette fonction, spatiales ou suivant un paramètre, s'effectuent par dérivations successives des colonnes ou des lignes de la matrice correspondant au déterminant. D'un point de vue des techniques algorithmiques, calculer la dérivée d'un déterminant d'une matrice suivant une colonne ou une ligne revient à se poser la question élémentaire du calcul d'un déterminant dans lequel une ou plusieurs lignes ou une ou plusieurs colonnes ont été remplacées par rapport à un déterminant de référence ($\Psi_{T\uparrow}$ ou $\Psi_{T\downarrow}$). Le calcul d'un déterminant s'effectuant en N^3 (N taille du déterminant), il représente une part importante du temps de calcul. Il a donc été nécessaire de trouver des techniques minimisant ce temps. Dans le cas du Jastrow commun le temps de calcul des déterminants est peu important face au temps de calcul des éléments de matrice. Dans le cas de

la fonction multi-Jastrow, le nombre de déterminants est bien plus grand. Pour les dérivées spatiales, cela est dû au fait que chaque électron \vec{r}_i est présent dans toutes les colonnes des matrices via les Jastrows orbitaux. Pour les dérivées suivant les paramètres du Jastrow, cela provient du fait que les Jastrows sont à l'intérieur de la matrice. Les temps de calcul de ces déterminants deviennent une partie prépondérante de l'algorithme. Nous renvoyons au chapitre 7 une analyse détaillée du nombre de déterminants à calculer et nous présenterons une comparaison des temps de calcul sur des exemples concrets dans le chapitre 8. Il est donc indispensable de trouver des solutions pratiques pour minimiser ces calculs. Un premier algorithme a été proposé en se basant sur les propriétés des comatrices. Cependant, il a dû être abandonné car il nécessitait encore trop de déterminants à calculer. Nous renvoyons le lecteur à l'annexe C pour une description de celui-ci.

Voyons maintenant l'algorithme que nous avons finalement choisi et implémenté dans le code QMC=Chem pour calculer tous ces déterminants. Le cœur en est les deux formules suivantes :

$$\frac{\det(M + \delta M)}{\det(M)} = \frac{\det(M^{-1})}{\det(M^{-1})} \frac{\det(M + \delta M)}{\det(M)} = \det(Id + M^{-1}\delta M) \quad (5.36)$$

et

$$\frac{\det(M + \delta M)}{\det(M)} = \frac{M^{-1}}{M^{-1}} \frac{\det(M + \delta M)}{\det(M)} = \det(Id + \delta M M^{-1}) \quad (5.37)$$

où M et δM sont deux matrices quelconques. Ici, la matrice M est supposée inversible. Dans la première formule nous considérons le produit $M^{-1}\delta M$. Dans la seconde formule, nous considérons le produit $\delta M M^{-1}$. Ces formules permettent tout en connaissant le déterminant de la matrice M d'obtenir le déterminant de la matrice légèrement modifiée $M + \delta M$ grâce au calcul du déterminant de $Id + \delta M M^{-1}$ (resp. $Id + M^{-1}\delta M$). Dans tous les cas que nous rencontrerons, la nouvelle matrice $Id + \delta M M^{-1}$ (resp. $Id + M^{-1}\delta M$) est d'une forme extrêmement pratique pour le calcul des déterminants en un nombre limité d'opérations. Nous utiliserons cette formule pour calculer des déterminants de matrices pour lesquelles nous aurons interverti soit :

- une ligne

- deux lignes
- une colonne
- deux colonnes

5.2.1 Calcul de Ψ_T

Dans toute la suite de notre exposé nous considérerons le cas d'une fonction d'onde composée d'un seul déterminant de Slater (en fait, produit d'un déterminant \uparrow et d'un déterminant \downarrow). Pour le cas d'une somme de déterminants :

$$\Psi_T = \sum_n c_n \Psi_n \quad (5.38)$$

il suffit d'appliquer ce qui suit à chaque déterminant Ψ_n .

Rappelons l'expression de Ψ_T

$$\Psi_T = \begin{vmatrix} e^{J_1(\vec{r}_1|\vec{r}_2,\dots)}\Phi_1(\vec{r}_1) & e^{J_1(\vec{r}_2|\vec{r}_1,\dots)}\Phi_1(\vec{r}_2) & \dots & \dots & e^{J_1(\vec{r}_{N_\uparrow}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_\uparrow}) \\ e^{J_2(\vec{r}_1|\vec{r}_2,\dots)}\Phi_2(\vec{r}_1) & e^{J_2(\vec{r}_2|\vec{r}_1,\dots)}\Phi_2(\vec{r}_2) & \dots & \dots & e^{J_2(\vec{r}_{N_\uparrow}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_\uparrow}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \\ \times \begin{vmatrix} e^{J_1(\vec{r}_{N_\uparrow+1}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_\uparrow+1}) & e^{J_1(\vec{r}_{N_\uparrow+2}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_\uparrow+2}) & \dots & \dots & e^{J_1(\vec{r}_{N_e}|\vec{r}_1,\dots)}\Phi_1(\vec{r}_{N_e}) \\ e^{J_2(\vec{r}_{N_\uparrow+1}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_\uparrow+1}) & e^{J_2(\vec{r}_{N_\uparrow+2}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_\uparrow+2}) & \dots & \dots & e^{J_2(\vec{r}_{N_e}|\vec{r}_1,\dots)}\Phi_2(\vec{r}_{N_e}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \quad (5.39)$$

Posons alors

$$\Psi_T = |\Psi_\uparrow||\Psi_\downarrow| \quad (5.40)$$

où $|M|$ désigne le déterminant de la matrice M et où

$$\Psi_\uparrow = \begin{pmatrix} \Psi_{11}^\uparrow & \dots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \Psi_{N_\uparrow 1}^\uparrow & \dots & \Psi_{N_\uparrow N_\uparrow}^\uparrow \end{pmatrix} \quad (5.41)$$

$$\Psi_{\downarrow} = \begin{pmatrix} \Psi_{1(N_{\uparrow}+1)}^{\downarrow} & \cdots & \Psi_{1N_e}^{\downarrow} \\ \vdots & \ddots & \vdots \\ \Psi_{N_{\downarrow}N_{\uparrow}+1}^{\downarrow} & \cdots & \Psi_{N_{\downarrow}N_e}^{\downarrow} \end{pmatrix} \quad (5.42)$$

avec

$$\Psi_{ij}^{\uparrow} \equiv \Phi_i(\vec{r}_j) e^{J_i(\vec{r}_j)|\vec{r}_1, \dots)} \quad (5.43)$$

et

$$\Psi_{ij}^{\downarrow} \equiv \Phi_i(\vec{r}_{j+N_{\uparrow}}) e^{J_i(\vec{r}_{j+N_{\uparrow}})|\vec{r}_1, \dots)} \quad (5.44)$$

Ici, N_e est le nombre total d'électrons, N_{\uparrow} le nombre d'électrons de spins up, N_{\downarrow} le nombre d'électrons de spins down, Ψ_{\uparrow} la matrice des électrons de spins up, et Ψ_{\downarrow} la matrice des électrons de spins down.

Le calcul de Ψ_T s'effectue par un calcul des déterminants up $|\Psi_{\uparrow}|$ et down $|\Psi_{\downarrow}|$ par un pivot de Gauss en $O(N^3)$ où N est la taille de la matrice carrée considérée. Les matrices doivent également être inversées. L'inversion et le calcul des déterminants s'effectuent par l'application de l'algorithme du pivot de Gauss à ces deux matrices. Le déterminant et l'inverse sont obtenus par le même pivot.

Aucune hypothèse n'est faite sur la structure de ces matrices. Celles-ci sont de forme quelconque ce qui explique l'usage de cet algorithme-ci et non d'autres plus performants adaptés à des types de matrices spécifiques (triangulaires, creuses, etc.). Les déterminants $|\Psi_{\uparrow}|$ et $|\Psi_{\downarrow}|$ sont parmi tous les déterminants calculés, les seuls déterminants obtenus de cette manière. Par comparaison avec le cas du Jastrow commun, il y a autant de déterminants à calculer.

5.2.2 Calcul du gradient

Après avoir calculé nos déterminants de référence $|\Psi_{\uparrow}|$ et $|\Psi_{\downarrow}|$ que nous utiliserons dans toute la suite, nous allons maintenant calculer les dérivées partielles de Ψ_T pour constituer le gradient $\vec{\nabla} \Psi_T$.

Notons $\partial_{x_i^m}$ la dérivée partielle suivant la coordonnée m de l'électron i . Nous noterons également :

- $x_i^1 = x_i$
- $x_i^2 = y_i$
- $x_i^3 = z_i$

les coordonnées (x, y, z) de l'électron i .

Posons

$$\partial_{x_i^m}(\Psi_{ij}^\uparrow) = \frac{\partial \Psi_{ij}^\uparrow}{\partial x_i^m} \quad (5.45)$$

$$\partial_{x_i^m}(\Psi_{ij}^\downarrow) = \frac{\partial \Psi_{ij}^\downarrow}{\partial x_i^m} \quad (5.46)$$

La formule de la dérivée partielle de Ψ_T suivant la coordonnée x_i^m est :

$$\begin{aligned} \partial_i^m \Psi_T = & \sum_{k=1}^{N_\uparrow} \begin{vmatrix} \Psi_{11}^\uparrow & \dots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \dots & \partial_{x_i^m}(\Psi_{jk}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix} |\Psi_\downarrow| \\ + & |\Psi_\uparrow| \sum_{k=1, N_\downarrow} \begin{vmatrix} \Psi_{1(N_\uparrow+1)}^\downarrow & \dots & \partial_{x_i^m}(\Psi_{1k}^\downarrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j(N_\uparrow+1)}^\downarrow & \dots & \partial_{x_i^m}(\Psi_{jk}^\downarrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix} \end{aligned} \quad (5.47)$$

Dans la première matrice, nous dérivons une colonne après l'autre. La somme porte ainsi sur les N_\uparrow colonnes que nous dérivons. Il en est de même pour la seconde matrice. La somme porte sur les N_\downarrow colonnes. Sachant que nous dérivons la fonction d'onde suivant une coordonnée de l'électron i , l'expression des dérivées des coefficients de la matrice est différente suivant que la colonne k dérivée est la $i^{\text{ème}}$ ou non. Cela provient du fait que dans la $i^{\text{ème}}$ colonne, nous dérivons également la fonction $\Phi_l(\vec{r}_i)$. Pour n'importe quelle ligne l nous avons donc :

- si $i \neq k$

$$\partial_{x_i^m}(\Psi_{lk}^\uparrow) = \partial_{x_i^m}(e^{J_l(\vec{r}_k|\dots)})\Phi_l(\vec{r}_k) \quad (5.48)$$

$$\partial_{x_i^m}(\Psi_{lk}^\downarrow) = \partial_i^m(e^{J_l(\vec{r}_{N_\uparrow+k}|\dots)})\Phi_i(\vec{r}_{N_\uparrow+k}) \quad (5.49)$$

- si $i = k$

$$\partial_i^m(\Psi_{lk}^\uparrow) = \partial_i^m(e^{J_l(\vec{r}_k|\dots)}\Phi_l(\vec{r}_k)) + e^{J_l(\vec{r}_k|\dots)}\partial_i^m\Phi_l(\vec{r}_k) \quad (5.50)$$

$$\partial_{x_i^m}(\Psi_{ik}^\downarrow) = \partial_{x_i^m}(e^{J_l(\vec{r}_{N_\uparrow+k}|\dots)}\Phi_l(\vec{r}_{N_\uparrow+k})) + e^{J_l(\vec{r}_k|\dots)}\partial_{x_i^m}\Phi_l(\vec{r}_k) \quad (5.51)$$

Voyons maintenant comment calculer ces déterminants avec la formule 5.2.2. Prenons pour cela le cas de la matrice Ψ^\uparrow pour laquelle la colonne k a été dérivée :

$$\begin{pmatrix} \Psi_{11}^\uparrow & \dots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \dots & \partial_{x_i^m}(\Psi_{jk}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad (5.52)$$

Rappelons que cette dérivée est différente suivant que j est ou non également la valeur de l'électron suivant lequel la fonction Ψ_T est dérivée pour obtenir le gradient. La méthode est strictement la même pour tous les autres déterminants.

Nous utiliserons la formule :

$$\frac{\det(M + \delta M)}{\det(M)} = \det(Id + M^{-1}\delta M) \quad (5.53)$$

avec

$$M = \Psi^\uparrow \quad (5.54)$$

et

$$\delta M = \begin{pmatrix} 0 & 0 & \dots & -\Psi_{1k}^\uparrow + \partial_{x_i^m}\Psi_{1k}^\uparrow & \dots & 0 \\ 0 & 0 & \dots & -\Psi_{2k}^\uparrow + \partial_{x_i^m}\Psi_{2k}^\uparrow & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & -\Psi_{N_\uparrow k}^\uparrow + \partial_{x_i^m}\Psi_{N_\uparrow k}^\uparrow & \dots & 0 \end{pmatrix} \quad (5.55)$$

Regardons comment s'effectuent les simplifications du calcul. Les éléments de δM n'étant différents de zéro que sur une colonne, le produit $M^{-1}\delta M$ est

de la forme :

$$P = \begin{pmatrix} 0 & 0 & \dots & P_{1k} & \dots & 0 \\ 0 & 0 & \dots & P_{2k} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & P_{N_{\uparrow}k} & \dots & 0 \end{pmatrix} \quad (5.56)$$

La matrice $Q = Id + P$ devient donc :

$$Q = \begin{pmatrix} 1 & 0 & \dots & P_{1k} & \dots & 0 \\ 0 & 1 & \dots & P_{2k} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & P_{kk} + 1 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_{N_{\uparrow}k} & \dots & 1 \end{pmatrix} \quad (5.57)$$

et le déterminant $\det(Q)$:

$$\det(Q) = P_{kk} + 1 \quad (5.58)$$

Autrement dit, le calcul de $\det(Q)$ et donc de $\det(M + \delta M)$ se résume au calcul de la quantité P_{kk}

$$P_{kk} = \sum_{m=1, N_{\uparrow}} \Psi_{km}^{\uparrow, -1} \delta M_{mk} \quad (5.59)$$

Le point clé qui nous a incité à utiliser la formule est la forme simple de la matrice P qui permet de calculer très facilement le déterminant. Cette forme simple est possible parce que nous avons multiplié une matrice δM avec une seule colonne non nulle à droite de la matrice M^{-1} . Nous verrons plus tard pour le calcul des dérivées suivant un paramètre que nous obtiendrons également une matrice à la forme simple en multipliant à gauche de M^{-1} une matrice avec une seule ligne non nulle.

5.2.3 Dérivées doubles

Nous devons dériver deux fois le produit des déterminants des deux matrices :

$$\Psi_{\uparrow}\Psi_{\downarrow}$$

L'expression peut se décomposer en trois parties distinctes :

$$\frac{\partial^2 \Psi_T}{\partial x_i^2} = S_1 + S_2 + S_3 \quad (5.60)$$

- S_1 est le produit des déterminants où chaque matrice a une colonne dérivée une fois
- S_2 est le produit des déterminants d'une matrice dont deux colonnes distinctes sont dérivées une fois par la matrice non altérée de spin opposé
- S_3 est le produit des déterminants d'une matrice dont une même colonne est dérivée deux fois de suite par la matrice non altérée de spin opposé

Calcul de S_1

L'expression de S_1 est la suivante :

$$S_1 = 2 \sum_{k=1, N_{\uparrow}} \sum_{l=1, N_{\downarrow}} \begin{vmatrix} \Psi_{11}^{\uparrow} & \cdots & \partial_{x_i^m}(\Psi_{1k}^{\uparrow}) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^{\uparrow} & \cdots & \partial_{x_i^m}(\Psi_{jk}^{\uparrow}) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix} \begin{vmatrix} \Psi_{1(N_{\uparrow}+1)}^{\downarrow} & \cdots & \partial_{x_i^m}(\Psi_{1l}^{\downarrow}) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j(N_{\uparrow}+1)}^{\downarrow} & \cdots & \partial_{x_i^m}(\Psi_{jl}^{\downarrow}) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix} \quad (5.61)$$

Il s'agit du produit des dérivées simples des déterminants up et down. Chaque déterminant de spin up comportant une dérivée dans une colonne est multiplié par chaque déterminant de spin down comportant une dérivée. Ces quantités ont déjà été calculées ci-dessus pour le gradient. Nous attirons le lecteur sur le facteur deux devant les sommes.

Pour calculer S_1 , il nous faut donc obtenir les produits :

$$\left| \begin{array}{cccc} \Psi_{11}^\uparrow & \cdots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \cdots & \partial_{x_i^m}(\Psi_{jk}^\uparrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right| \left| \begin{array}{cccc} \Psi_{1(N_\uparrow+1)}^\downarrow & \cdots & \partial_{x_i^m}(\Psi_{1l}^\downarrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j(N_\uparrow+1)}^\downarrow & \cdots & \partial_{x_i^m}(\Psi_{jl}^\downarrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right|$$

Nous partons donc des éléments de la dérivée simple. Lorsque nous considérons, par exemple, les termes

$$\left| \begin{array}{cccc} \Psi_{11}^\uparrow & \cdots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \cdots & \partial_{x_i^m}(\Psi_{jk}^\uparrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right| \left| \begin{array}{ccc} \Psi_{1(N_\uparrow+1)}^\downarrow & \cdots & \Psi_{1N_e}^\downarrow \\ \vdots & \ddots & \vdots \\ \Psi_{N_\uparrow 1}^\downarrow & \cdots & \Psi_{N_\uparrow N_e}^\downarrow \end{array} \right|$$

$$+ \left| \begin{array}{ccc} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \Psi_{N_\uparrow 1}^\uparrow & \cdots & \Psi_{N_\uparrow N_\uparrow}^\uparrow \end{array} \right| \left| \begin{array}{cccc} \Psi_{1(N_\uparrow+1)}^\downarrow & \cdots & \partial_{x_i^m}(\Psi_{1l}^\downarrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j(N_\uparrow+1)}^\downarrow & \cdots & \partial_{x_i^m}(\Psi_{jl}^\downarrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right|$$

Nous dérivons alors dans le premier terme la $l^{\text{ème}}$ colonne de la matrice de spin down et dans le second terme, la $k^{\text{ème}}$ colonne. Nous obtenons bien deux fois le produit voulu.

Calcul de S_2

L'expression de S_2 est :

$$\begin{aligned}
 S_2 = & 2 \sum_{k < l} \begin{vmatrix} \dots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \dots & \partial_{x_i^m}(\Psi_{1l}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \partial_{x_i^m}(\Psi_{jk}^\uparrow) & \dots & \partial_{x_i^m}(\Psi_{jl}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} & |\Psi_{\downarrow l}| \\
 + & 2|\Psi_{\uparrow}| \sum_{k < l} \begin{vmatrix} \dots & \partial_{x_i^m}(\Psi_{1k}^\downarrow) & \dots & \partial_{x_i^m}(\Psi_{1l}^\downarrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \partial_{x_i^m}(\Psi_{jk}^\downarrow) & \dots & \partial_{x_i^m}(\Psi_{jl}^\downarrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \\
 & (5.62)
 \end{aligned}$$

Cette fois-ci deux colonnes distinctes ont été dérivées dans la même matrice.

Pour les mêmes raisons que pour le calcul de S_1 , nous avons également un facteur 2 devant les sommes car les sommes portent sur des colonnes k et l tel que $k < l$. En effet, pour obtenir :

$$\begin{vmatrix} \dots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \dots & \partial_{x_i^m}(\Psi_{1l}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \partial_{x_i^m}(\Psi_{jk}^\uparrow) & \dots & \partial_{x_i^m}(\Psi_{jl}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

Nous pouvons soit dériver la colonne l de la matrice :

$$\begin{vmatrix} \dots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \dots & \Psi_{1l}^\uparrow & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \partial_{x_i^m}(\Psi_{jk}^\uparrow) & \dots & \Psi_{jl}^\uparrow & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

soit dériver la colonne k de la matrice :

$$\begin{vmatrix} \dots & \Psi_{1k}^\uparrow & \dots & \partial_{x_i^m}(\Psi_{1l}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \Psi_{jk}^\uparrow & \dots & \partial_{x_i^m}(\Psi_{jl}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

Voyons maintenant comment nous calculons ces déterminants. Prenons pour cela le cas de la matrice Ψ^\uparrow pour laquelle les colonnes k et l ont été dérivées :

$$\begin{pmatrix} \dots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \dots & \partial_{x_i^m}(\Psi_{1l}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \partial_{x_i^m}(\Psi_{jk}^\uparrow) & \dots & \partial_{x_i^m}(\Psi_{jl}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad (5.63)$$

Nous utiliserons la formule 5.2.2 :

$$\frac{\det(M + \delta M)}{\det(M)} = \det(Id + M^{-1}\delta M)$$

avec

$$M = \Psi^\uparrow \quad (5.64)$$

et

$$\delta M = \begin{pmatrix} 0 & \dots & -\Psi_{1k}^\uparrow + \partial_{x_i^m}\Psi_{1k}^\uparrow & \dots & 0 & \dots & -\Psi_{1l}^\uparrow + \partial_{x_i^m}\Psi_{1l}^\uparrow & \dots & 0 \\ 0 & \dots & -\Psi_{2k}^\uparrow + \partial_{x_i^m}\Psi_{2k}^\uparrow & \dots & 0 & \dots & -\Psi_{2l}^\uparrow + \partial_{x_i^m}\Psi_{2l}^\uparrow & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & -\Psi_{N_1k}^\uparrow + \partial_{x_i^m}\Psi_{N_1k}^\uparrow & \dots & 0 & \dots & -\Psi_{N_1l}^\uparrow + \partial_{x_i^m}\Psi_{N_1l}^\uparrow & \dots & 0 \end{pmatrix} \quad (5.65)$$

Cette fois-ci, les éléments de δM sont différents de zéro sur deux colonnes,

le produit $M^{-1}\delta M$ est alors de la forme :

$$P = \begin{pmatrix} 0 & 0 & \dots & P_{1k} & \dots & 0 & \dots & P_{1l} & \dots & 0 \\ 0 & 0 & \dots & P_{2k} & \dots & 0 & \dots & P_{2l} & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & P_{N_{\uparrow}k} & \dots & 0 & \dots & P_{N_{\uparrow}l} & \dots & 0 \end{pmatrix} \quad (5.66)$$

La matrice $Q = Id + P$ devient donc :

$$Q = \begin{pmatrix} 1 & 0 & \dots & P_{1k} & 0 & P_{1l} & \dots & 0 \\ 0 & 1 & \dots & P_{2k} & 0 & P_{2l} & \dots & 0 \\ \vdots & & 1 & \vdots & 0 & \vdots & & \vdots \\ & \vdots & & P_{kk} + 1 & 0 & P_{kl} & & \vdots \\ \vdots & \vdots & \vdots & & 1 & \vdots & & \vdots \\ \vdots & \vdots & & P_{lk} & 0 & P_{ll} + 1 & & \vdots \\ \vdots & \vdots & \vdots & & 0 & \vdots & 1 & 0 \\ 0 & 0 & \dots & P_{N_{\uparrow}k} & 0 & P_{N_{\uparrow}l} & \dots & 1 \end{pmatrix} \quad (5.67)$$

et le déterminant $\det(Q)$:

$$\det(Q) = (P_{kk} + 1)(P_{ll} + 1) - P_{lk}P_{kl} \quad (5.68)$$

Le calcul de $\det(Q)$ et donc de $\det(M + \delta M)$ nécessite alors quatre quantités $P_{kl}, P_{lk}, P_{ll}, P_{kk}$. Ces quantités sont nécessaires pour le calcul de la dérivée double suivant les colonnes l et k . Cependant, nous cherchons à calculer les dérivées doubles pour toutes les paires (l, k) avec $1 \leq l < k \leq N_{\uparrow}$ (respectivement N_{\downarrow}). Il est donc inutile de recalculer à chaque fois le produit d'une colonne l avec la matrice M .

Notons

$$R = \begin{pmatrix} -\Psi_{11}^\uparrow + \partial_{x_i^m} \Psi_{11}^\uparrow & \dots & -\Psi_{1k}^\uparrow + \partial_{x_i^m} \Psi_{1k}^\uparrow & \dots & 0 & \dots & -\Psi_{1l}^\uparrow + \partial_{x_i^m} \Psi_{1l}^\uparrow & \dots \\ -\Psi_{21}^\uparrow + \partial_{x_i^m} \Psi_{21}^\uparrow & \dots & -\Psi_{2k}^\uparrow + \partial_{x_i^m} \Psi_{2k}^\uparrow & \dots & 0 & \dots & -\Psi_{2l}^\uparrow + \partial_{x_i^m} \Psi_{2l}^\uparrow & \dots \\ \vdots & & \vdots & & \vdots & & \vdots & \\ \vdots & & \vdots & & \vdots & & \vdots & \\ \vdots & & \vdots & & \vdots & & \vdots & \\ -\Psi_{N_\uparrow 1}^\uparrow + \partial_{x_i^m} \Psi_{N_\uparrow 1}^\uparrow & \dots & -\Psi_{N_\uparrow k}^\uparrow + \partial_{x_i^m} \Psi_{N_\uparrow k}^\uparrow & \dots & 0 & \dots & -\Psi_{N_\uparrow l}^\uparrow + \partial_{x_i^m} \Psi_{N_\uparrow l}^\uparrow & \dots \end{pmatrix} \quad (5.69)$$

Pour obtenir la totalité de ces déterminants, les quantités nécessaires sont donc l'ensemble des éléments de R . Remarquons que les éléments sur la diagonale de P sont les éléments permettant de calculer les éléments du gradient comme montré ci-dessus. En pratique, pour calculer les éléments de cette partie 2 du Laplacien ainsi que les éléments du gradient (et donc également les éléments de la partie 1), nous procédons de la sorte :

- nous calculons la matrice R
- nous calculons le produit $D = MR$
- nous manipulons les coefficients de la matrice D pour obtenir le gradient et les parties 1 et 2 du Laplacien. Quelque soient les colonnes k et l de la matrice, le déterminant de Q est donné par les formules suivantes. Pour les dérivées simples de la colonne k :

$$\det(Q) = (D_{kk} + 1) \quad (5.70)$$

et pour les dérivées doubles des colonnes k et l

$$\det(Q) = (D_{kk} + 1)(D_{ll} + 1) - D_{lk}D_{kl} \quad (5.71)$$

Calcul de S_3

Rappelons la notation $\partial_{x_i^m}^2$ qui désigne la dérivée double suivant la coordonnée m de l'électron i .

La formule pour S_3 est :

$$\begin{aligned}
S_3 = & \sum_{k=1, N_\uparrow} \left| \begin{array}{cccc} \Psi_{11}^\uparrow & \dots & \partial_{x_i^m}^2(\Psi_{1k}^\uparrow) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \dots & \partial_{x_i^m}^2(\Psi_{jk}^\uparrow) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right| \times |\Psi_\downarrow| \\
+ & \sum_{k=1, N_\downarrow} |\Psi_\uparrow| \times \left| \begin{array}{cccc} \Psi_{1(N_\uparrow+1)}^\downarrow & \dots & \partial_{x_i^m}^2(\Psi_{1k}^\downarrow) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{j(N_\uparrow+1)}^\downarrow & \dots & \partial_{x_i^m}^2(\Psi_{jk}^\downarrow) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right|
\end{aligned} \tag{5.72}$$

Une colonne est remplacée par une dérivée double. Nous reprenons le même principe que dans la partie consacrée au gradient pour calculer la matrice δM du déterminant up pour laquelle une colonne k a été dérivée deux fois :

$$\left(\begin{array}{cccc} \Psi_{11}^\uparrow & \dots & \partial_{x_i^m}^2(\Psi_{1k}^\uparrow) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \dots & \partial_{x_i^m}^2(\Psi_{jk}^\uparrow) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right) \tag{5.73}$$

On applique la formule 5.2.2 :

$$\frac{\det(M + \delta M)}{\det(M)} = \det(Id + M^{-1}\delta M)$$

avec

$$M = \Psi^\uparrow \tag{5.74}$$

et

$$\delta M = \left(\begin{array}{cccc} 0 & 0 & \dots & -\Psi_{1j}^\uparrow + \partial_{x_i^m}^2 \Psi_{1j}^\uparrow & \dots & 0 \\ 0 & 0 & \dots & -\Psi_{2j}^\uparrow + \partial_{x_i^m}^2 \Psi_{2j}^\uparrow & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & -\Psi_{N_\uparrow j}^\uparrow + \partial_{x_i^m}^2 \Psi_{N_\uparrow j}^\uparrow & \dots & 0 \end{array} \right) \tag{5.75}$$

Comme précédemment, le produit $M^{-1}\delta M$ est de la forme :

$$P = \begin{pmatrix} 0 & 0 & \dots & P_{1k} & \dots & 0 \\ 0 & 0 & \dots & P_{2k} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & P_{N_{\uparrow}k} & \dots & 0 \end{pmatrix} \quad (5.76)$$

La matrice $Q = Id + P$ devient donc :

$$Q = \begin{pmatrix} 1 & 0 & \dots & P_{1k} & \dots & 0 \\ 0 & 1 & \dots & P_{2k} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & P_{kk} + 1 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_{N_{\uparrow}k} & \dots & 1 \end{pmatrix} \quad (5.77)$$

et le déterminant $\det(Q)$:

$$\det(Q) = P_{kk} + 1 \quad (5.78)$$

Pour chaque dérivée double, seul le coefficient P_{kk} est nécessaire. C'est pourquoi nous ne calculerons que les élément P_{kk} de la matrice $P = M^{-1}\delta M$

$$P_{jj} = \sum_{k=1}^{N_{\uparrow}} \Psi_{jk}^{\uparrow,-1} \delta M_{kj} \quad (5.79)$$

5.2.4 Dérivée partielle de Ψ_T suivant un paramètre c_p du Jastrow

Notons Ψ_{c_p} cette dérivée. On a la formule :

$$\Psi_{c_p} = \sum_{j=1, N_{\uparrow}} \left| \begin{array}{ccc} \Psi_{11}^{\uparrow} & \cdots & \Psi_{1N_{\uparrow}}^{\uparrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^{\uparrow}}{\partial c_p} & \cdots & \frac{\partial \Psi_{jN_{\uparrow}}^{\uparrow}}{\partial c_p} \\ \vdots & \ddots & \vdots \end{array} \right| |\Psi_{\downarrow}| \quad (5.80)$$

$$|\Psi_{\uparrow}| \sum_{j=1, N_{\downarrow}} \left| \begin{array}{ccc} \Psi_{1(N_{\uparrow}+1)}^{\downarrow} & \cdots & \Psi_{1N_e}^{\downarrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{l(N_{\uparrow}+1)}^{\downarrow}}{\partial c_p} & \cdots & \frac{\partial \Psi_{lN_e}^{\downarrow}}{\partial c_p} \\ \vdots & \ddots & \vdots \\ \Psi_{N_{\downarrow}1}^{\downarrow} & \cdots & \Psi_{N_{\downarrow}N_e}^{\downarrow} \end{array} \right|$$

Un paramètre peut être présent sur plusieurs lignes de la matrice. Il y a alors plusieurs déterminants à calculer. Isolons en un et calculons le. Nous poserons que la $j^{\text{ème}}$ ligne est dérivée suivant le paramètre c_p . Nous appliquons la formule 5.37. Cette fois nous intervertissons une ligne, δM est multiplié à gauche :

$$\frac{\det(M + \delta M)}{\det(M)} = \det(Id + \delta M M^{-1})$$

Dans ce cas-là, nous nous intéressons au produit $\delta M M^{-1}$ (et non $M^{-1} \delta M$ comme pour l'interversion d'une colonne). Nous appliquons la formule de base avec :

$$M = \Psi^{\uparrow} \quad (5.81)$$

et

$$\delta M = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ -\Psi_{1j}^{\uparrow} + \frac{\partial \Psi_{j1}^{\uparrow}}{\partial c_p} & \cdots & -\Psi_{N_{\uparrow}j}^{\uparrow} + \frac{\partial \Psi_{jN_{\uparrow}}^{\uparrow}}{\partial c_p} \\ \vdots & \ddots & \vdots \end{pmatrix} \quad (5.82)$$

Les éléments de δM n'étant différents de zéro que sur une ligne, le produit $\delta M M^{-1}$ est de la forme :

$$P = \begin{pmatrix} 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ P_{1j} & P_{2j} & \vdots & \vdots & \vdots & P_{N_\uparrow j} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & \dots & 0 \end{pmatrix} \quad (5.83)$$

La matrice $Q = Id + P$ devient donc :

$$Q = \begin{pmatrix} 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ P_{1j} & P_{2j} & \vdots & P_{jj} + 1 & \vdots & P_{N_\uparrow j} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \quad (5.84)$$

et le déterminant $\det(Q)$:

$$\det(Q) = P_{jj} + 1 \quad (5.85)$$

Autrement dit, le calcul de $\det(Q)$ et donc de $\det(M + \delta M)$ se résume au calcul de la quantité P_{jj}

$$P_{jj} = \sum_{k=1}^{N_\uparrow} \delta M_{jk} \Psi_{kj}^{\uparrow, -1} \quad (5.86)$$

5.2.5 Dérivée simple suivant un paramètre du Jastrow de la dérivée simple spatiale de Ψ_T

Nous n'avons pas besoin de la dérivée du gradient suivant un paramètre. Néanmoins, nous aurons besoin par la suite de la dérivée partielle suivant un

paramètre du Jastrow de la dérivée spatiale simple :

$$\begin{pmatrix} \Psi_{11}^\uparrow & \cdots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \cdots & \frac{\partial \partial_{x_i^m}(\Psi_{jk}^\uparrow)}{\partial c_p} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \cdots & \partial_{x_i^m}(\Psi_{jk}^\uparrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad (5.87)$$

Nous partons de la fonction Ψ_{c_p} et nous procédons comme dans la partie ci-dessus en intervertissant les colonnes pour obtenir les quantités voulues. Cette fois-ci, notre fonction de référence n'est plus Ψ_\uparrow mais $\Psi_{c_p}^\uparrow$ avec

$$\Psi_{c_p}^\uparrow = \begin{pmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \cdots & \frac{\partial \Psi_{jN_\uparrow}^\uparrow}{\partial c_p} \\ \vdots & \ddots & \vdots \end{pmatrix}$$

Nous avons ici effectué un abus de langage et d'écriture pour simplifier l'exposé. En effet, nous parlons de $|\Psi_{c_p}^\uparrow|$ comme le déterminant de la matrice des électrons de spin up dont nous aurions dérivé une ligne suivant le paramètre c_p . Il faudra bien sûr appliquer cela à toutes les lignes qui comportent ce paramètre c_p .

Nous appliquons la formule 5.2.2.

$$\frac{\det(M + \delta M)}{\det(M)} = \det(Id + M^{-1} \cdot \delta M)$$

avec

$$M = \begin{pmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \cdots & \frac{\partial \Psi_{jN_\uparrow}^\uparrow}{\partial c_p} \\ \vdots & \ddots & \vdots \end{pmatrix} \quad (5.88)$$

et

$$\delta M = \begin{pmatrix} 0 & \dots & \partial_{x_i^m}(\Psi_{1k}^\dagger) & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \frac{\partial \partial_{x_i^m}(\Psi_{jk}^\dagger)}{\partial c_p} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \partial_{x_i^m}(\Psi_{jk}^\dagger) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad (5.89)$$

Le déterminant $\det(M)$ est calculé dans la partie précédente. Mais pour appliquer la formule, nous avons également besoin d'inverser les matrices $\Psi_{c_p}^\dagger$ et $\Psi_{c_p}^\dagger$. Afin de ne pas utiliser une inversion via le pivot de Gauss en N_\uparrow^3 (resp. N_\downarrow^3) coûteuse en temps, nous avons choisi pour économiser du temps de calcul d'utiliser la formule de Sherman-Morrison[145]. Ce choix est indispensable au vu du nombre important de déterminants à calculer. Si l'inversion par pivot de Gauss pèse peu dans les temps de calcul dans le cas du Jastrow commun, le nombre important de déterminants à calculer dans le cas de la fonction multi-Jastrow rend son usage problématique.

L'inversion par la formule Sherman-Morrison

Posons A une matrice carrée, u et v deux vecteurs. La formule de Sherman-Morrison[145] permet d'inverser la matrice $A + u \otimes v$ en connaissant l'inverse de A . Rappelons que $\delta A = u \otimes v$ est la matrice dont l'élément δA_{ij} est le produit $u_i v_j$.

Voyons comment nous obtenons notre expression :

$$\begin{aligned} (A + u \otimes v)^{-1} &= (1 + A^{-1}.u \otimes v)^{-1}.A^{-1} \\ &= (1 - A^{-1}.u \otimes v + A^{-1}.u \otimes v.A^{-1}.u \otimes v - \dots).A^{-1} \\ &= A^{-1} - A^{-1} \cdot \otimes c.A - -1(1 - \lambda + \lambda^2 - \dots) \\ &= A^{-1} - \frac{(A^{-1}.u) \otimes (v.A^{-1})}{1+\lambda} \end{aligned}$$

avec

$$\lambda \equiv v.A^{-1}.u \quad (5.90)$$

Ainsi, en posant

$$z \equiv A^{-1}.u \quad (5.91)$$

$$w \equiv (A^{-1})^T.v \quad (5.92)$$

$$\lambda = v.z \quad (5.93)$$

On obtient :

$$(A + u \otimes v)^{-1} = A^{-1} - \frac{z \otimes w}{1 + \lambda} \quad (5.94)$$

Application de la formule de Sherman-Morrison à notre problème

Nous souhaitons calculer l'inverse de la matrice :

$$M = \begin{pmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{l1}^\uparrow}{\partial c_p} & \cdots & \frac{\partial \Psi_{lN_\uparrow}^\uparrow}{\partial c_p} \\ \vdots & \ddots & \vdots \end{pmatrix} \quad (5.95)$$

à partir de $\Psi^{\uparrow,-1}$. Nous avons échangé la $l^{\text{ème}}$ ligne. Nous appliquons donc la formule de Sherman-Morrison avec :

$$M = \Psi^\uparrow \quad (5.96)$$

$$u = \begin{pmatrix} -\Psi_{l1}^\uparrow + \frac{\partial \Psi_{l1}^\uparrow}{\partial c_p} \\ -\Psi_{l2}^\uparrow + \frac{\partial \Psi_{l2}^\uparrow}{\partial c_p} \\ \vdots \\ \vdots \end{pmatrix} \quad (5.97)$$

$$v = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \rightarrow l^{\text{ème}} \text{ ligne} \\ 0 \\ \vdots \end{pmatrix} \quad (5.98)$$

5.2.6 Dérivée du Laplacien de Ψ_T suivant un paramètre du Jastrow

Nous utiliserons la notation $\partial_{x_i^m}^2$ pour désigner la dérivée suivant le paramètre c_p de la dérivée double suivant la coordonnées m de l'électron i . La dérivées suivant un paramètre du Jastrow du Laplacien est la somme sur les électrons i de la dérivée suivant le paramètre c_p des dérivées doubles suivant les trois coordonnées des électrons.

$$\Delta_l \Psi_T = \sum_{i=1, N_{el}} \partial_{x_i^1}^2 \Psi_T + \partial_{x_i^2}^2 \Psi_T + \partial_{x_i^3}^2 \Psi_T \quad (5.99)$$

La dérivée suivant le paramètre c_p de la dérivée double suivant la coordonnée m de l'électron i a pour formule :

$$\begin{aligned} & \partial_{x_i^m}^2 \Psi \\ = & \sum_{j=1, N_\uparrow} \sum_{k=1, N_\uparrow} \left| \begin{array}{cccc} \Psi_{11}^\uparrow & \cdots & \partial_{x_i^m}^2(\Psi_{1k}^\uparrow) & \cdots \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \cdots & \frac{\partial \partial_{x_i^m}^2(\Psi_{jk}^\uparrow)}{\partial c_p} & \cdots \cdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right| |\Psi_\downarrow| \quad (5.100) \end{aligned}$$

$$+ \sum_{j=1, N_\downarrow} \sum_{k=1, N_\downarrow} |\Psi_\uparrow| \left| \begin{array}{cccc} \Psi_{1(N_\uparrow+1)}^\downarrow & \cdots & \partial_{x_i^m}^2(\Psi_{1k}^\downarrow) & \cdots \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j(N_\uparrow+1)}^\downarrow}{\partial c_p} & \cdots & \frac{\partial \partial_{x_i^m}^2(\Psi_{jk}^\downarrow)}{\partial c_p} & \cdots \cdots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right| \quad (5.101)$$

$$+ \sum_{j=1, N_\uparrow} \sum_{k=1, N_\downarrow} \left| \begin{array}{ccc|ccc} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow & \Psi_{1(N_\uparrow+1)}^\downarrow & \cdots & \partial_{x_i^m}^2(\Psi_{1k}^\downarrow) & \cdots \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \cdots & \frac{\partial \Psi_{jN_\uparrow}^\uparrow}{\partial c_p} & \Psi_{j(N_\uparrow+1)}^\downarrow & \cdots & \partial_{x_i^m}^2(\Psi_{jk}^\downarrow) & \cdots \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right| \quad (5.102)$$

$$+ \sum_{j=1, N_\downarrow} \sum_{k=1, N_\uparrow} \left| \begin{array}{ccc|ccc} \Psi_{11}^\uparrow & \cdots & \partial_{x_i^m}^2(\Psi_{1k}^\uparrow) & \cdots \cdots & \Psi_{1(N_\uparrow+1)}^\downarrow & \cdots & \Psi_{1N_e}^\downarrow \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \Psi_{j1}^\uparrow & \cdots & \partial_{x_i^m}^2(\Psi_{jk}^\uparrow) & \cdots \cdots & \frac{\partial \Psi_{j(N_\uparrow+1)}^\downarrow}{\partial c_p} & \cdots & \frac{\partial \Psi_{jN_e}^\downarrow}{\partial c_p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \end{array} \right| \quad (5.103)$$

$$+ 2 \sum_{j=1, N_\uparrow} \sum_{1 \leq k < l \leq N_\uparrow} \left| \begin{array}{cccc} \cdots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \cdots & \partial_{x_i^m}(\Psi_{1l}^\uparrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \frac{\partial \partial_{x_i^m}(\Psi_{jk}^\uparrow)}{\partial c_p} & \cdots & \frac{\partial \partial_{x_i^m}(\Psi_{jl}^\uparrow)}{\partial c_p} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right| |\Psi_\downarrow| \quad (5.104)$$

$$+ 2 \sum_{j=1, N_\downarrow} \sum_{1 \leq k < l \leq N_\downarrow} |\Psi_\uparrow| \left| \begin{array}{cccc} \cdots & \partial_{x_i^m}(\Psi_{1k}^\downarrow) & \cdots & \partial_{x_i^m}(\Psi_{1l}^\downarrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \frac{\partial \partial_{x_i^m}(\Psi_{jk}^\downarrow)}{\partial c_p} & \cdots & \frac{\partial \partial_{x_i^m}(\Psi_{jl}^\downarrow)}{\partial c_p} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right| \quad (5.105)$$

$$+2 \sum_{j=1, N_{\uparrow}} \sum_{1 \leq k < l \leq N_{\downarrow}} \left| \begin{array}{ccc|ccc} \Psi_{11}^{\uparrow} & \cdots & \Psi_{1N_{\uparrow}}^{\uparrow} & \cdots & \partial_{x_i^m}(\Psi_{1k}^{\downarrow}) & \cdots & \partial_{x_i^m}(\Psi_{1l}^{\downarrow}) & \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j1}^{\uparrow}}{\partial c_p} & \cdots & \frac{\partial \Psi_{jN_{\uparrow}}^{\uparrow}}{\partial c_p} & \cdots & \partial_{x_i^m}(\Psi_{jk}^{\downarrow}) & \cdots & \partial_{x_i^m}(\Psi_{jl}^{\downarrow}) & \cdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right| \quad (5.106)$$

$$+2 \sum_{j=1, N_{\downarrow}} \sum_{1 \leq k < l \leq N_{\uparrow}} \left| \begin{array}{ccc|ccc} \cdots & \partial_{x_i^m}(\Psi_{1k}^{\uparrow}) & \cdots & \partial_{x_i^m}(\Psi_{1l}^{\uparrow}) & \cdots & \Psi_{1(N_{\uparrow}+1)}^{\downarrow} & \cdots & \Psi_{1k}^{\downarrow} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \partial_{x_i^m}(\Psi_{jk}^{\uparrow}) & \cdots & \partial_{x_i^m}(\Psi_{jl}^{\uparrow}) & \cdots & \frac{\partial \Psi_{j(N_{\uparrow}+1)}^{\downarrow}}{\partial c_p} & \cdots & \frac{\partial \Psi_{jk}^{\downarrow}}{\partial c_p} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right| \quad (5.107)$$

$$+ \sum_{j=1, N_{\uparrow}} \sum_{k=1, N_{\uparrow}} \sum_{l=1, N_{\downarrow}} \left| \begin{array}{ccc|ccc} \Psi_{1(N_{\downarrow}+1)}^{\uparrow} & \cdots & \partial_{x_i^m}(\Psi_{1k}^{\uparrow}) & \cdots & \Psi_{11}^{\downarrow} & \cdots & \partial_{x_i^m}(\Psi_{1l}^{\downarrow}) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j(N_{\downarrow}+1)}^{\uparrow}}{\partial c_p} & \cdots & \frac{\partial \partial_{x_i^m}(\Psi_{jk}^{\uparrow})}{\partial c_p} & \cdots & \Psi_{j1}^{\downarrow} & \cdots & \partial_{x_i^m}(\Psi_{jl}^{\downarrow}) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right| \quad (5.108)$$

$$+ \sum_{j=1, N_{\downarrow}} \sum_{k=1, N_{\downarrow}} \sum_{l=1, N_{\uparrow}} \left| \begin{array}{ccc|ccc} \Psi_{11}^{\uparrow} & \cdots & \partial_{x_i^m}(\Psi_{1l}^{\uparrow}) & \cdots & \Psi_{1(N_{\uparrow}+1)}^{\downarrow} & \cdots & \partial_{x_i^m}(\Psi_{1k}^{\downarrow}) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^{\uparrow} & \cdots & \partial_{x_i^m}(\Psi_{jl}^{\uparrow}) & \cdots & \frac{\partial \Psi_{j(N_{\uparrow}+1)}^{\downarrow}}{\partial c_p} & \cdots & \frac{\partial \partial_{x_i^m}(\Psi_{jk}^{\downarrow})}{\partial c_p} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right| \quad (5.109)$$

$$(5.110)$$

Explicitons cette formule. Les sommes sont groupées par paire. Les paires (5.100,5.101), (5.102,5.103), (5.104,5.105), (5.106,5.107), (5.108,5.109). Au sein d'une même paire, nous avons les mêmes types de dérivées spatiales ou suivant un paramètre mais symétriques suivant le spin. Si pour une paire (i, j) , il y a dans le terme deux dérivées spatiales dans la matrice des spins up et une dérivée spatiale dans la matrice des spins down, alors dans le terme j , il y aura une dérivée suivant le paramètre dans la matrice des spins up et deux dérivées spatiales dans la matrice de spins down.

Sur la première ligne dans l'expression (5.100), la première somme porte sur le nombre de lignes de la matrice de spins up. Chaque ligne j est dérivée suivant le paramètre c_p . Bien sûr, les lignes ne contenant pas le paramètre c_p auront une dérivée nulle. La seconde somme porte sur le nombre k de colonnes de la matrice de spins up. Chaque colonne sera dérivée spatialement une fois. Nous attirons l'attention sur la forme du coefficient jk de la matrice de spin up, à l'intersection de la dérivée suivant le paramètre et de la dérivée spatiale. Celui-ci est dérivée suivant le paramètre c_p et suivant la dérivée spatiale. Ce type de déterminants n'a pas été encore calculé.

Pour la seconde ligne (5.101), la description est identique en remplaçant la matrice up par la matrice down.

Dans la troisième ligne (5.102), la première somme porte sur les lignes de la matrice de spins up. Chaque ligne est dérivée suivant le paramètre c_p . La seconde somme porte sur les colonnes de la matrice de spins down. Chaque colonne est dérivée spatialement deux fois.

Dans la quatrième ligne (5.103), la première somme porte sur les lignes de la matrice de spins down. Chaque ligne est dérivée suivant le paramètre c_p . La seconde somme porte sur les colonnes de la matrice de spins up. Chaque colonne est dérivée spatialement deux fois.

Dans la cinquième ligne (5.104), la première somme porte sur les lignes de la matrice de spins up. Chaque ligne j est dérivée suivant le paramètre c_p . La seconde somme porte sur les colonnes de la matrice de spins up. Deux colonnes distinctes k et l sont simultanément dérivées une fois. Nous attirons l'attention sur le facteur 2 devant les sommes en accord avec l'inégalité stricte $k < l$. Comme pour le cas du Laplacien, il y a deux manières de dériver les colonnes k et l : dériver la colonne k après avoir dérivé la colonne j ou dériver la colonne j après avoir dérivé la colonne k . Il est donc nécessaire de compter chaque matrice avec deux colonnes dérivées deux fois. Comme dans la ligne (5.100), les coefficients jk et jl , à l'intersection de la dérivée suivant la lignes et des dérivées suivant les colonnes, sont les dérivées des coefficients de la matrice de spins down suivant le paramètre c_p et la coordonnée spatiale. Ce

type de déterminants n'a pas encore été calculé.

La description de la sixième ligne (5.105) est identique à celle de la ligne précédente en intervertissant les matrices up et down.

Dans la septième ligne (5.106), la première somme porte sur les lignes de la matrice de spins up. Chaque ligne j est dérivée suivant le paramètre c_p . La seconde somme porte sur les colonnes de la matrice de spins up. Deux colonnes distinctes k et l sont simultanément dérivées une fois. Pour les mêmes raisons qu'auparavant, un facteur 2 se trouve devant les sommes.

La description de la huitième ligne (5.107) est identique à celle de la ligne précédente en intervertissant les matrices up et down.

Dans l'avant-dernière ligne (5.108), la première somme porte sur les lignes de la matrice de spins up. Chaque ligne j est dérivée suivant le paramètre c_p . La seconde somme porte sur les colonnes de la matrice de spins down. Chaque colonne est dérivée deux fois.

La description de la dernière ligne (5.109) est identique à la ligne précédente en intervertissant les matrices up et down.

A l'exception des déterminants explicitement indiqués ci-dessus, tous les autres déterminants ont déjà été calculés précédemment. Il en reste donc encore deux types à calculer :

- Les déterminants de la forme :

$$\begin{vmatrix} \Psi_{11}^\dagger & \cdots & \partial_{x_i^m}^2(\Psi_{1k}^\dagger) & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j1}^\dagger}{\partial c_p} & \cdots & \frac{\partial \partial_{x_i^m}^2(\Psi_{jk}^\dagger)}{\partial c_p} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \quad (5.111)$$

qui se calculent en appliquant la formule 5.2.2

$$\frac{\det(M + \delta M)}{\det(M)} = \det(Id + M^{-1} \cdot \delta M)$$

avec :

$$M = \begin{pmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \cdots & \frac{\partial \Psi_{jN_\uparrow}^\uparrow}{\partial c_p} \\ \vdots & \ddots & \vdots \end{pmatrix} \quad (5.112)$$

$$\delta M = \begin{pmatrix} 0 & \cdots & \Psi^\uparrow - \Psi_{1k}^\uparrow + \partial_{x_i^m}^2(\Psi_{1k}^\uparrow) & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & -\frac{\partial \Psi_{jk}^\uparrow}{\partial c_p} + \frac{\partial \partial_{x_i^m}^2(\Psi_{jk}^\uparrow)}{\partial c_p} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad (5.113)$$

• Les déterminants de la forme :

$$\begin{vmatrix} \cdots & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & \cdots & \partial_{x_i^m}(\Psi_{1l}^\uparrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \frac{\partial \partial_{x_i^m}(\Psi_{jk}^\uparrow)}{\partial c_p} & \cdots & \frac{\partial \partial_{x_i^m}(\Psi_{jl}^\uparrow)}{\partial c_p} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \quad (5.114)$$

qui se calculent en appliquant la formule 5.2.2

$$\frac{\det(M + \delta M)}{\det(M)} = \det(Id + M^{-1}.\delta M)$$

avec :

$$M = \begin{pmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \cdots & \frac{\partial \Psi_{jN_\uparrow}^\uparrow}{\partial c_p} \\ \vdots & \ddots & \vdots \end{pmatrix} \quad (5.115)$$

$$\delta M = \begin{pmatrix} 0 & \partial_{x_i^m}(\Psi_{1k}^\uparrow) & 0 & \partial_{x_i^m}(\Psi_{1l}^\uparrow) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & -\frac{\partial \Psi_{jk}^\uparrow}{\partial c_p} + \frac{\partial \partial_{x_i^m}(\Psi_{jk}^\uparrow)}{\partial c_p} & \cdots & -\frac{\partial \Psi_{jl}^\uparrow}{\partial c_p} + \frac{\partial \partial_{x_i^m}(\Psi_{jl}^\uparrow)}{\partial c_p} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \quad (5.116)$$

5.2.7 Dérivée double suivant les paramètres du Jastrow de la fonction d'onde

Nous dérivons la fonction d'onde suivant les paramètres c_1 et c_2 . Cette dérivée a exactement la même structure que la dérivée double spatiale de la fonction d'onde. Mais cette fois-ci, au lieu d'invertir des colonnes, nous intervertissons des lignes. Suivant la remarque que nous avons faite précédemment, c'est donc la multiplication de δM à gauche de M dans la formule 5.2.2 :

$$\frac{\det(M + \delta M)}{\det(M)} = \det(\text{Id} + \delta M M^{-1})$$

qui permet de calculer simplement le déterminant $\det(M + \delta M)$.

La dérivée double peut se décomposer en 3 parties :

$$\Psi_{c_1, c_2} = D_{c_1, c_2}^1 + D_{c_1, c_2}^2 + D_{c_1, c_2}^3 \quad (5.117)$$

- D_{c_1, c_2}^1 est composé des déterminants des produits de matrices ayant chacune une ligne dérivée suivant l'un des deux paramètres
- D_{c_1, c_2}^2 est composé des déterminants des produits d'une matrice ayant deux lignes distinctes dérivées suivant les paramètres par la matrice de spin opposé non altérée
- D_{c_1, c_2}^3 est composé des déterminants des produits d'une matrice ayant une seule ligne dérivée deux fois suivant les paramètres par la matrice de spin opposé non altérée

Calcul de D_{c_1, c_2}^1 :

L'expression de D_{c_1, c_2}^1 est :

$$\begin{aligned}
 D_{c_1, c_2}^1 = & \sum_{j=1, N_{\uparrow}} \sum_{k=1, N_{\downarrow}} \left| \begin{array}{ccc} \Psi_{11}^{\uparrow} & \cdots & \Psi_{1N_{\uparrow}}^{\uparrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^{\uparrow}}{\partial c_1} & \cdots & \frac{\partial \Psi_{jN_{\uparrow}}^{\uparrow}}{\partial c_1} \\ \vdots & \ddots & \vdots \\ \Psi_{11}^{\uparrow} & \cdots & \Psi_{1N_{\uparrow}}^{\uparrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^{\uparrow}}{\partial c_2} & \cdots & \frac{\partial \Psi_{jN_{\uparrow}}^{\uparrow}}{\partial c_2} \\ \vdots & \ddots & \vdots \end{array} \right| \left| \begin{array}{ccc} \Psi_{1(N_{\uparrow}+1)}^{\downarrow} & \cdots & \Psi_{1N_e}^{\downarrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{k(N_{\uparrow}+1)}^{\downarrow}}{\partial c_2} & \cdots & \frac{\partial \Psi_{kN_e}^{\downarrow}}{\partial c_2} \\ \vdots & \ddots & \vdots \\ \Psi_{1(N_{\uparrow}+1)}^{\downarrow} & \cdots & \Psi_{1N_e}^{\downarrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{k(N_{\uparrow}+1)}^{\downarrow}}{\partial c_1} & \cdots & \frac{\partial \Psi_{kN_e}^{\downarrow}}{\partial c_1} \\ \vdots & \ddots & \vdots \end{array} \right| \\
 & + \sum_{j=1, N_{\uparrow}} \sum_{k=1, N_{\downarrow}} \left| \begin{array}{ccc} \Psi_{11}^{\uparrow} & \cdots & \Psi_{1N_{\uparrow}}^{\uparrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^{\uparrow}}{\partial c_2} & \cdots & \frac{\partial \Psi_{jN_{\uparrow}}^{\uparrow}}{\partial c_2} \\ \vdots & \ddots & \vdots \\ \Psi_{11}^{\uparrow} & \cdots & \Psi_{1N_{\uparrow}}^{\uparrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^{\uparrow}}{\partial c_1} & \cdots & \frac{\partial \Psi_{jN_{\uparrow}}^{\uparrow}}{\partial c_1} \\ \vdots & \ddots & \vdots \end{array} \right| \left| \begin{array}{ccc} \Psi_{1(N_{\uparrow}+1)}^{\downarrow} & \cdots & \Psi_{1N_e}^{\downarrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{k(N_{\uparrow}+1)}^{\downarrow}}{\partial c_1} & \cdots & \frac{\partial \Psi_{kN_e}^{\downarrow}}{\partial c_1} \\ \vdots & \ddots & \vdots \\ \Psi_{1(N_{\uparrow}+1)}^{\downarrow} & \cdots & \Psi_{1N_e}^{\downarrow} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{k(N_{\uparrow}+1)}^{\downarrow}}{\partial c_2} & \cdots & \frac{\partial \Psi_{kN_e}^{\downarrow}}{\partial c_2} \\ \vdots & \ddots & \vdots \end{array} \right|
 \end{aligned}
 \tag{5.118}$$

Les sommes portent sur les lignes j et k que l'on dérive suivant un paramètre. Ces dérivées peuvent éventuellement être nulles. Ces déterminants ont déjà été calculés ci-dessus.

Calcul de D_{c_1, c_2}^2 :

L'expression de D_{c_1, c_2}^2 est :

$$\begin{aligned}
D_{c_1, c_2}^2 = & 2 \sum_{j=1, N_\uparrow} \sum_{k=1, N_\uparrow, j_2 \neq j_1} \begin{vmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_1} & \cdots & \frac{\partial \Psi_{jN_\uparrow}^\uparrow}{\partial c_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{k1}^\uparrow}{\partial c_2} & \cdots & \frac{\partial \Psi_{kN_\uparrow}^\uparrow}{\partial c_2} \\ \vdots & \ddots & \vdots \end{vmatrix} |\Psi_\downarrow| \\
& + 2 \sum_{j=1, N_\downarrow} \sum_{k=1, N_\downarrow, j_2 \neq j_1} |\Psi_\uparrow| \begin{vmatrix} \Psi_{1(N_\uparrow+1)}^\downarrow & \cdots & \Psi_{1N_e}^\downarrow \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j(N_\uparrow+1)}^\downarrow}{\partial c_1} & \cdots & \frac{\partial \Psi_{jN_e}^\downarrow}{\partial c_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{k(N_\uparrow+1)}^\downarrow}{\partial c_2} & \cdots & \frac{\partial \Psi_{kN_e}^\downarrow}{\partial c_2} \\ \vdots & \ddots & \vdots \end{vmatrix}
\end{aligned} \tag{5.119}$$

Dans une matrice, on dérive deux lignes distinctes j et k une fois suivant les paramètres. Nous avons pour les même raisons que précédemment un facteur 2 devant les sommes. Pour calculer D_{c_1, c_2}^2 , on applique la formule 5.2.2 :

$$\frac{\det(M + \delta M)}{\det(M)} = \det(Id + \delta M.M^{-1})$$

avec :

$$M = \Psi^\uparrow \tag{5.120}$$

$$\delta M = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ -\Psi_{j1}^\uparrow + \frac{\partial \Psi_{j1}^\uparrow}{\partial c_1} & \cdots & -\Psi_{1N_\uparrow}^\uparrow + \frac{\partial \Psi_{jN_\uparrow}^\uparrow}{\partial c_1} \\ \vdots & \ddots & \vdots \\ \Psi_{k1}^\uparrow + \frac{\partial \Psi_{k1}^\uparrow}{\partial c_2} & \cdots & -\Psi_{kN_\uparrow}^\uparrow + \frac{\partial \Psi_{kN_\uparrow}^\uparrow}{\partial c_2} \\ \vdots & \ddots & \vdots \end{pmatrix} \tag{5.121}$$

Calcul de D_{c_1, c_2}^3 :

L'expression de D_{c_1, c_2}^3 est :

$$\begin{aligned}
 D_{c_1, c_2}^3 = & \sum_{j=1, N_\uparrow} \left| \begin{array}{ccc} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \Psi_{j1}^\uparrow}{\partial c_1 c_2} & \cdots & \frac{\partial^2 \Psi_{jN_\uparrow}^\uparrow}{\partial c_1 c_2} \\ \vdots & \ddots & \vdots \end{array} \right| |\Psi_\downarrow| \\
 & + \sum_{j=1, N_\downarrow} |\Psi_\uparrow| \left| \begin{array}{ccc} \Psi_{1(N_\uparrow+1)}^\downarrow & \cdots & \Psi_{1N_e}^\downarrow \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \Psi_{j(N_\uparrow+1)}^\downarrow}{\partial c_1 c_2} & \cdots & \frac{\partial^2 \Psi_{jN_e}^\downarrow}{\partial c_1 c_2} \\ \vdots & \ddots & \vdots \end{array} \right|
 \end{aligned} \tag{5.122}$$

Chaque ligne de matrice est dérivée deux fois suivant les paramètres. Pour calculer D_{c_1, c_2}^3 , on applique la formule 5.2.2 :

$$\frac{\det(M + \delta M)}{\det(M)} = \det(Id + \delta M.M^{-1})$$

avec :

$$M = \Psi^\uparrow \tag{5.123}$$

$$\delta M = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ -\Psi_{j1}^\uparrow + \frac{\partial^2 \Psi_{j1}^\uparrow}{\partial c_1 c_2} & \cdots & -\Psi_{1N_\uparrow}^\uparrow + \frac{\partial^2 \Psi_{jN_\uparrow}^\uparrow}{\partial c_1 c_2} \\ \vdots & \ddots & \vdots \end{pmatrix} \tag{5.124}$$

Chapitre 6

Application de la fonction d'onde multi-Jastrow à quelques systèmes atomiques et moléculaires

**Multi-Jastrow trial wavefunctions for electronic structure
calculations with quantum Monte Carlo**

Thomas Bouabça and Michel Caffarel

Laboratoire de Chimie et Physique Quantiques, CNRS-IRSAMC Université
de Toulouse, France.

Abstract

A new type of electronic trial wavefunction suitable for quantum Monte Carlo calculations of molecular systems is presented. In contrast with the standard Jastrow-Slater form built with a unique global Jastrow term, it is proposed to introduce multiple Jastrow factors attached to individual molecular orbitals. Such a form allows to describe *differently* the local electronic correlations

occurring in various environments (1s-core orbitals, 3d-magnetic orbitals, localized two-center σ -orbitals, delocalized π -orbitals, atomic lone pairs, etc.). Two remarkable points associated with the use of a multi-Jastrow form are the possibility of better describing the nodal structure of the wavefunction and the possibility of defining and pre-optimizing local and *transferable* correlated units for building complex trial wavefunctions from simple parts. The practical aspects associated with the computation of the intricate derivatives of the multi-Jastrow trial function are presented in detail. Some illustrative applications for a light and a more heavy atom (O and Cu), and for the potential energy curve and spectroscopic constants of the HF molecule are discussed. Our numerical calculations show that the computational overhead associated with the use of the multi-Jastrow form can be kept under control and that this new wavefunction contains indeed some interesting new physics.

6.1 Introduction

At the heart of quantum Monte Carlo (QMC) simulations of electronic ground-state properties is the use of a trial wavefunction, Ψ_T , which approximates the unknown wavefunction, Ψ_0 . The role played by Ψ_T is many-fold. First, it is used to implement the “Importance Sampling” strategy common to any *efficient* Monte Carlo scheme (either classical or quantum). As well-known, apart for very simple systems, Monte Carlo estimators can converge only if the random sampling of the configuration space is made efficiently (*i.e.*, only regions associated with a significant density probability are visited). Second, the trial wavefunction is used for reducing the statistical fluctuations of the total energy estimator. This is based on the existence of a zero-variance principle which states that the closer the trial wavefunction is from the exact wave function, the smallest the statistical errors are. In the limit, $\Psi_T = \Psi_0$ the statistical error entirely vanishes (zero-variance). Note that this zero-variance principle valid for the energy has been recently extended to arbitrary properties[6]. Finally, for fermionic systems the trial

wavefunction plays another central role related to the existence of the Pauli exclusion principle. As well-known, there does not exist a stable and exact QMC algorithm for fermionic systems (because of the famous “sign problem”, for molecules see, *e.g.*, Ref.[2]). In practice, this difficulty is circumvented by using a stable algorithm consisting in solving the Schrödinger equation with the additional constraint that the solution has the same sign as the trial wavefunction. Unfortunately, this constraint which implies to impose to Ψ_0 the same nodal hypersurfaces as Ψ_T , introduces a small bias, the so-called “fixed-node” (FN) error. This error is in general a very small fraction of the total energy; nevertheless, it can lead to undesirable physical consequences when considering the very small energy differences we are interested in (electron affinities, ionization potentials, binding energies, etc.). Note also that it can be shown that the fixed-node energy is an upper bound of the exact energy (variational principle, [3]).

Now, in view of the practical importance of all these points, an intense activity has been developed towards devising better trial wavefunctions associated with smaller statistical fluctuations (to decrease the computational effort and/or to treat larger systems) and smaller fixed-node errors (better control of the quality of the results). The general guiding idea to improve trial wavefunctions is to incorporate into them as much physics as possible. In addition, the form of the trial wavefunctions must absolutely be in practice as compact as possible (the function, its first and second derivatives are to be calculated at each Monte Carlo step). Finally, on top of that, it is in general advantageous to optimize by minimization of the energy or the variance all or some part of the parameters appearing in the trial wavefunction.

The most standard form used presently for the trial wavefunction, known as the Slater-Jastrow form, is written as

$$\Psi_T = e^{J(\mathbf{r}_1, \dots, \mathbf{r}_N)} \sum_{k=1}^{N_{det}} c_k Det_k(\{\Phi_i^\uparrow\}) Det_k(\{\Phi_i^\downarrow\}) \quad (6.1)$$

N being the number of electrons, N_{det} the number of terms in the multi-determinant expansion, and $\{\Phi_i^\sigma\}$ ($\sigma = \uparrow, \downarrow$) a set of molecular orbitals. In

this expression the prefactor e^J is known as the Jastrow factor. Its role is to impose the exact behavior of the wavefunction in the $[r_{ij} \rightarrow 0]$ -limit (electron-electron cusp condition) and, also, to incorporate some two-body (electron-electron and electron-nucleus) and three-body (electron-electron-nucleus) correlations (to describe the best as possible the shape of the Coulomb hole,[4]). The determinantal part is in general issued from some *ab-initio*-wavefunction or Density Functional Theory (DFT) calculations. The multi-determinant expansion is introduced for describing situations where the exact wavefunction is expected to be strongly multi-configurational (bond-breaking, excited-states, etc.). Because molecular orbitals are optimized within a pure *ab initio* framework without the presence of a Jastrow term, it is in general important to re-optimize most of the parameters of the trial wavefunction[5] (Jastrow parameters, molecular orbitals, and expansion coefficients c_k).

Other forms for the trial wavefunction have been introduced in the literature but, so far, they have remained of marginal use. We can cite, for example, the geminal wavefunction of Sorella and coworkers,[6] the Pfaffian wavefunction of Mitáš and collaborators,[7], and the backflow trial wavefunction of Rios et al.[8].

Here, we propose a new trial wavefunction valid for arbitrary molecular systems. This wavefunction is a generalization of the standard form, Eq.(137), where the common Jastrow factor is replaced by several Jastrow factors J_i , each of them being attached to a particular molecular orbital, Φ_i . In what follows, this trial wavefunction will be called the multi-Jastrow wavefunction and the J_i 's, the *molecular* Jastrow factors. The expression of the new wave function is written as follows

$$\Psi_T = \sum_{k=1}^{N_{det}} c_k Det_k(\{e^{J_i} \Phi_i^\uparrow\}) Det_k(\{e^{J_i} \Phi_i^\downarrow\}). \quad (6.2)$$

The physical idea behind this form is to allow an *orbital-dependent* description of the structure of the correlation effects. Indeed, it is quite legitimate not to place on the same footing the correlation effects between electrons

occupying a $1s$ -core orbital, a $3d$ -magnetic orbital, a localized two-center σ -orbital, a delocalized π -orbital, etc. As we shall see in our applications, the multi-Jastrow trial wavefunction appears to be a physical improvement with respect to the standard form, Eq.(137). Not only the variational energy is lowered but the change of the nodal structure associated with the introduction of molecular Jastrows into determinants can also lead to better results. Finally, employing a multi-Jastrow form opens perspectives toward devising systematic procedures for constructing good-quality wavefunctions for electronic systems of arbitrary size.

The organization of this paper is as follows. In Section 6.2 we present the multi-Jastrow trial wavefunction. In Section 6.3 the algorithms for optimizing the trial wavefunction are presented. In Section 6.4 the most difficult technical part of the approach, namely evaluating efficiently the derivatives of the multi-Jastrow trial wavefunction with respect to spatial coordinates and parameters is presented in detail. In Section 6.5, illustrative applications for the oxygen and copper atoms and the HF molecule are presented. Finally, we present some summary and perspectives in the last section 6.6.

6.2 The multi-Jastrow trial wavefunction

6.2.1 The standard form

As briefly recalled in the introduction the standard form used for the trial wavefunction is as follows

$$\Psi_T = e^{J(\mathbf{r}_1, \dots, \mathbf{r}_N)} \sum_{k=1}^{N_{det}} c_k \begin{vmatrix} \phi_{k_1^\uparrow}(\mathbf{r}_1) & \dots & \phi_{k_1^\uparrow}(\mathbf{r}_{N_\uparrow}) \\ \vdots & \vdots & \vdots \\ \phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_1) & \dots & \phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_{N_\uparrow}) \end{vmatrix} \begin{vmatrix} \phi_{k_1^\downarrow}(\mathbf{r}_{N_\uparrow+1}) & \dots & \phi_{k_1^\downarrow}(\mathbf{r}_N) \\ \vdots & \vdots & \vdots \\ \phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_{N_\uparrow+1}) & \dots & \phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_N) \end{vmatrix} \quad (6.3)$$

In this expression the prefactor e^J is the Jastrow factor. Many different forms for J have been proposed in the literature. A typical “minimal” form is given

as follows

$$J(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i < j} \frac{a_{\sigma_{ij}} r_{ij}}{1 + b r_{ij}} - \sum_{\alpha} c_{\alpha} \sum_i r_{i\alpha} \quad (6.4)$$

where the latin indexes i and j refer to electrons and the greek index α to nuclei. Parameters $a_{\sigma_{ij}}$ are introduced to impose the electron-electron cusp condition, namely $a_{\sigma_{ij}} = 1/2$ for spin-unlike electrons and $a_{\sigma_{ij}} = 1/4$ for spin-like electrons. The parameter b is introduced to take into account the screening of the electron-electron interaction at large inter-electronic distances, and the parameters c_{α} are introduced to allow some re-adjustment of the electronic density once the r_{ij} terms have been introduced into the wavefunction. More sophisticated forms can be written and are generally expressed under the generic form

$$J(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i < j} v_{e-e}(r_{ij}) + \sum_{\alpha} v_{e-n}(r_{i\alpha}) + \sum_{i < j} \sum_{\alpha} v_{e-e-n}(r_{ij}, r_{i\alpha}, r_{j\alpha}) \quad (6.5)$$

where the functions, v_{e-e} , v_{e-n} , and v_{e-e-n} describe the two-body electron-electron and electron-nucleus correlations and the three-body electron-electron-nucleus correlations, respectively. Each of these functions are usually expanded in terms of simple elementary functions (polynomials, Padé, sums of exponentials, etc.).

In the expression, Eq.(6.3), of the trial wavefunction the determinantal part is a finite sum of N_{det} Slater determinants of single-particle molecular orbitals, c_k being the various amplitude coefficients. N_{\uparrow} , N_{\downarrow} are the number of \uparrow (or α) and \downarrow (or β) electrons, respectively. This part is obtained using standard DFT or *ab initio* wavefunction-based method : Hartree-Fock (HF), CASSCF, small-Configuration Interaction, etc. The various molecular orbitals $\phi_{k_i^{\sigma}}$, ($\sigma = \uparrow, \downarrow$) are extracted from a set of active orbitals $\phi_1(\mathbf{r}), \dots, \phi_{N_{act}}(\mathbf{r})$. Note that the formalism used in quantum Monte Carlo calculations is a spin-free formalism [9]. The trial wavefunction depends only on space coordinates, \mathbf{r}_i and, the molecular orbitals employed are spatial orbitals depending only on the ordinary 3D-coordinates. To get a wavefunction obeying the Pauli principle, the antisymmetry of the wavefunction under the

exchange of the spatial coordinates of any pair of electrons having the same spin has to be imposed.[9] Here, this property is fulfilled because the global Jastrow term is by construction fully symmetric under the exchange of any pair of electrons and because the product of two Slater determinants for $\sigma = \uparrow$ and $\sigma = \downarrow$ realizes the antisymmetry.

6.2.2 The multiJastrow form

The multiJastrow form is similar to the standard form, except that the global Jastrow term is no longer attached to the determinantal part as a whole like in Eq.(6.3), but is now split into multiple *molecular* Jastrow terms J_i attached to each molecular orbital $\phi_i(\mathbf{r})$. To be more precise, the wavefunction proposed here is given as

$$\Psi_T = \sum_{k=1}^{N_{det}} c_k \begin{vmatrix} e^{J_{k_1^\uparrow}(\mathbf{r}_1|\mathbf{r}_i \neq \mathbf{r}_1)} \phi_{k_1^\uparrow}(\mathbf{r}_1) & \dots & e^{J_{k_1^\uparrow}(\mathbf{r}_{N_\uparrow}|\mathbf{r}_i \neq \mathbf{r}_{N_\uparrow})} \phi_{k_1^\uparrow}(\mathbf{r}_{N_\uparrow}) \\ \vdots & \vdots & \vdots \\ e^{J_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_1|\mathbf{r}_i \neq \mathbf{r}_1)} \phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_1) & \dots & e^{J_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_{N_\uparrow}|\mathbf{r}_i \neq \mathbf{r}_{N_\uparrow})} \phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_{N_\uparrow}) \end{vmatrix} \\ \times \begin{vmatrix} e^{J_{k_1^\downarrow}(\mathbf{r}_{N_\uparrow+1}|\mathbf{r}_i \neq \mathbf{r}_{N_\uparrow+1})} \phi_{k_1^\downarrow}(\mathbf{r}_{N_\uparrow+1}) & \dots & e^{J_{k_1^\downarrow}(\mathbf{r}_N|\mathbf{r}_i \neq \mathbf{r}_N)} \phi_{k_1^\downarrow}(\mathbf{r}_N) \\ \vdots & \vdots & \vdots \\ e^{J_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_{N_\uparrow+1}|\mathbf{r}_i \neq \mathbf{r}_{N_\uparrow+1})} \phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_{N_\uparrow+1}) & \dots & e^{J_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_N|\mathbf{r}_i \neq \mathbf{r}_N)} \phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_N) \end{vmatrix}. \quad (6.6)$$

In this expression the set of orbitals $\{\phi_{k_i^\sigma}\}$, $\sigma = \uparrow, \downarrow$ of a given component of the sum is a subset of size N_σ of the full set of the active molecular orbitals used. The molecular Jastrow functions $J_k(\mathbf{r}_j|\mathbf{r}_i \neq \mathbf{r}_j)$ are arbitrary (smooth) functions of *all* electronic coordinates $(\mathbf{r}_1, \dots, \mathbf{r}_N)$. For convenience, we have adopted the following notation

$$J_k(\mathbf{r}_j|\mathbf{r}_i \neq \mathbf{r}_j) \equiv J_k(\mathbf{r}_j|\mathbf{r}_1, \dots, \mathbf{r}_i \neq \mathbf{r}_j, \dots, \mathbf{r}_N) \quad (6.7)$$

Remark that in the list of arguments of the molecular Jastrows the “reference” electron (here, labelled with j), which is the electron occupying the molecular orbital attached to the Jastrow, has been particularized from the $(N - 1)$

remaining “external” electrons. This distinction is introduced because we impose to all molecular Jastrows J_k to be symmetric under the exchange of any pair of *external* electrons. This property can be expressed as follows

$$J_k(\mathbf{r}_j|\mathbf{r}_1, \dots, \mathbf{r}_l, \dots, \mathbf{r}_m, \dots, \mathbf{r}_\mathbf{N}) = J_k(\mathbf{r}_j|\mathbf{r}_1, \dots, \mathbf{r}_m, \dots, \mathbf{r}_l, \dots, \mathbf{r}_\mathbf{N}) \quad \forall l \neq j \text{ and } \forall m \neq j \quad (6.8)$$

Now, the elementary but most important point to remark is that our proposed form (6.6) still obeys the Pauli exclusion principle. To verify this we just need to check that the wavefunction changes its sign when two electrons having the same spin are exchanged. Without loss of generality, let us consider that the two electrons to be exchanged are the electrons 1 and 2 of spin \uparrow . When doing the exchange, the first and second columns of all \uparrow -determinants are exchanged and the signs of these determinants are changed. In all other columns, (columns of \uparrow -determinants except columns 1 and 2 and all columns of \downarrow -determinants), electrons 1 and 2 have the status of external electrons. Now, because of the symmetry property, Eq.(6.8), this exchange leaves the wavefunction unchanged. Finally, as desired, the overall sign of the wavefunction is changed when spin-like electrons are exchanged.

Various forms for the molecular Jastrow terms can be used. In this work we have employed simple polynomial forms of the renormalized interelectronic distances, $\tilde{r}_{ij}^{(k)}$ and electron-nucleus distances $\tilde{r}_{i\alpha}^{(k)}$

$$\tilde{r}_{ij}^{(k)} \equiv \frac{r_{ij}}{1 + b_k r_{ij}}$$

$$\tilde{r}_{i\alpha}^{(k)} \equiv \frac{r_{i\alpha}}{1 + b_{\alpha k} r_{i\alpha}}$$

where latin indices i, j are used for electrons and the Greek index α is used for nuclei.

The molecular Jastrow is written as

$$J_k(\mathbf{r}_i|\mathbf{r}_j \neq \mathbf{r}_i) = \frac{1}{2} \sum_{j \neq i} \sum_{\alpha} [s_k(\tilde{r}_{ij}) - p_{k\alpha}(\tilde{r}_{i\alpha}) - p_{k\alpha}(\tilde{r}_{j\alpha}) + g_{k\alpha}^{(1)} \tilde{r}_{i\alpha}^2 \tilde{r}_{j\alpha}^2 + g_{k\alpha}^{(2)} (\tilde{r}_{i\alpha}^2 + \tilde{r}_{j\alpha}^2) \tilde{r}_{ij}^2] \quad (6.9)$$

where

$$s_k(r) = \frac{1}{N_{nucl}} (e_k^{(1)} r + e_k^{(2)} r^2 + e_k^{(3)} r^3 + e_k^{(4)} r^4)$$

$$p_{k\alpha}(r) = \frac{1}{N_{nucl}}(f_{k\alpha}^{(1)}r + f_{k\alpha}^{(2)}r^2 + f_{k\alpha}^{(3)}r^3 + f_{k\alpha}^{(4)}r^4)$$

where N_{nucl} denotes the number of nuclei. In these expressions, the quantities $\{b_k, b_{\alpha k}, e_k^{(l)}, f_{k\alpha}^{(l)}, g_{k\alpha}^{(l)}\}$ play the role of parameters. For a given molecular Jastrow, there are $7 + 5N_{nucl}$ such parameters.

Now, a few important remarks are in order. First, as already briefly discussed in the introduction, the physical role played by the molecular Jastrow's is rather clear. Each reference electron occupying a given molecular orbital interacts with the complete set of the $(N - 1)$ remaining external electrons. The nature of this interaction and its representation via a Jastrow term is clearly very dependent on the type of molecular orbital occupied by the reference electron. For example, $1s$ electrons close to some nucleus (large- Z attraction) do not experience the same local coulombic interaction as $3d$ electrons moving in a small region of space crowded by many electrons (large electron-electron repulsion). This is also the case when considering lone-pair electrons which are of atomic nature or π -electrons involved in a multiple bond which can be mainly involved between two atoms or circulating over a ring (aromatic compounds) and so on. A second remark is that to build a properly antisymmetrized wavefunction constructed from fully correlated elementary blocks (here, the product $J_i\phi_i$ depending explicitly on *all* electronic coordinates) is in general not tractable from a computational point of view due to the $N!$ problem coming from the \uparrow - and \downarrow -antisymmetrizers. Here, the particular form of our trial wavefunction and, more specifically, the symmetry among the external electrons allows to express the antisymmetrizers as standard Slater determinants like in the usual case where the elementary block is a product of one-particle functions. Associated with this remark, let us note that it would have been more physical to consider a spin-dependent interaction between the reference electron of a given spin and external electrons of the same and opposite spin. Unfortunately, introducing such a structure would have led to untractable form for the trial wavefunction. Finally, a last remark concerns the fact that allowing the orbitals to depend on the positions of the other electrons as done here is not new. A

first : geminals, pfaffian wavefunctions. Backflow trial wavefunctions.

6.2.3 Relationship with backflow-type trial wavefunction

A multideterminantal backflow (BF) trial wavefunction can be obtained from the standard Slater-Jastrow form, Eq.(6.3), by adding to the coordinates \mathbf{r}_i of the particles in the Slater a backflow displacement depending on the coordinates of all particles

$$\mathbf{r}_i \rightarrow \mathbf{r}_i + \xi_i(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad (6.10)$$

Physically, this backflow displacement is supposed to reproduce the perturbation of the flow of particles around a given particle. Using the same notations as above, the BF wavefunction can be written as follows

$$\Psi_T = e^{J(\mathbf{r}_1, \dots, \mathbf{r}_N)} \sum_{k=1}^{N_{det}} c_k \left| \begin{array}{ccc} \phi_{k_1^\uparrow}(\mathbf{r}_1 + \xi_1(\mathbf{r}_1, \dots, \mathbf{r}_N)) & \dots & \phi_{k_1^\uparrow}(\mathbf{r}_{N_\uparrow} + \xi_{N_\uparrow}(\mathbf{r}_1, \dots, \mathbf{r}_N)) \\ \vdots & & \vdots \\ \phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_1 + \xi_1(\mathbf{r}_1, \dots, \mathbf{r}_N)) & \dots & \phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_{N_\uparrow} + \xi_{N_\uparrow}(\mathbf{r}_1, \dots, \mathbf{r}_N)) \end{array} \right|$$

$$\times \left| \begin{array}{ccc} \phi_{k_1^\downarrow}(\mathbf{r}_{N_\uparrow+1} + \xi_{N_\uparrow+1}(\mathbf{r}_1, \dots, \mathbf{r}_N)) & \dots & \phi_{k_1^\downarrow}(\mathbf{r}_N + \xi_N(\mathbf{r}_1, \dots, \mathbf{r}_N)) \\ \vdots & & \vdots \\ \phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_{N_\uparrow+1} + \xi_{N_\uparrow+1}(\mathbf{r}_1, \dots, \mathbf{r}_N)) & \dots & \phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_N + \xi_N(\mathbf{r}_1, \dots, \mathbf{r}_N)) \end{array} \right| \quad (6.11)$$

Comparing Eqs.(6.6) and (144) allows to see the differences between how the electronic correlations are introduced. In the backflow case, the local displacement ξ for a given electron is independent on the orbital occupied. In contrast, in the multi-Jastrow case the coupling with other electrons is strongly dependent on the orbital occupied, a point which is central in our approach. Another point is that in the multi-Jastrow case the partitioning of the 3D-space between different regions : atomic regions, bonding regions, etc. is possible by using localized orbitals and localized molecular Jastrows. Furthermore, it is not necessary to introduce a Jastrow term for each molecular orbitals, a point which is not possible with the BF form.

6.3 Optimization of the trial wavefunction

In this section we present the optimization scheme employed in this work. As known, to define an efficient procedure for optimizing the parameters entering quantum Monte Carlo wavefunctions is in general non-trivial. For example, we have found that the standard variance-minimization algorithm using a correlated approach of Umrigar *et al.* [10] was insufficient to get reasonable parameters for our multi-Jastrow wavefunctions. We have therefore turned our attention to the very recently proposed method of Umrigar and collaborators[11] based on the use of a linear expansion of the trial wavefunction within the space spanned by the wavefunction and its first derivatives with respect to all parameters

$$\psi_T^{lin}(\mathbf{p}, \mathbf{R}) = \psi_T(\mathbf{p}_0, \mathbf{R}) + \sum_{i=1}^{N_{opt}} (\mathbf{p} - \mathbf{p}_0)_i \frac{\partial \psi_T}{\partial p_i}(\mathbf{p}_0, \mathbf{R}) \quad (6.12)$$

where $\mathbf{R} \equiv (\mathbf{r}_1, \dots, \mathbf{r}_N)$ and \mathbf{p} is the set of the N_{opt} parameters to be optimized. The optimal parameters (lowest total energy) are obtained by diagonalizing the Hamiltonian within the linear functional space consisting of $\Psi_0(\mathbf{R}) \equiv \psi_T(\mathbf{p}_0, \mathbf{R})$ and the N_{opt} functions $\psi_i(\mathbf{R}) \equiv \frac{\partial \psi_T}{\partial p_i}(\mathbf{p}_0, \mathbf{R})$. The $(N_{opt} + 1)$ -dimensional generalized eigenvalue problem is thus written as

$$\mathbf{H}\Delta\mathbf{p} = E\mathbf{S}\Delta\mathbf{p} \quad (6.13)$$

where $\Delta\mathbf{p}_i = (\mathbf{p} - \mathbf{p}_0)_i$ for $i = 1$ to N_{opt} and $\Delta\mathbf{p}_0 = 1$. The matrix elements are given by $H_{ij} = \langle \psi_i | H | \psi_j \rangle$ and $S_{ij} = \langle \psi_i | \psi_j \rangle$. In practice, this approach has been implemented by using the same strategy as in the original work[11], namely :

i.) Calculation of the matrix elements H_{ij} and S_{ij} using the Variational Monte Carlo (VMC) approach with $\psi_T^2(\mathbf{p}_0, \mathbf{R})$ as probability density, the estimators of the matrix elements obeying a “strong” zero-variance principle as introduced in [12].

ii.) Renormalization of the components of $\Delta\mathbf{p}$ corresponding to nonlinear parameters by using a rescaling factor.

iii.) Use of a stabilization procedure to minimize the effect of the statistical noise.

The preceding steps are iterated until the change of parameters, $\Delta\mathbf{p}$'s, is negligible. In this procedure only the last step has been modified and it has been done as follows. As proposed in Ref.[11] a positive constant, λ , is added to the diagonal part of the Hamiltonian matrix except for the first element

$$H_{ij}(\lambda) \equiv H_{ij} + \lambda\delta_{ij}(1 - \delta_{i0}) \quad (6.14)$$

Physically, λ can be viewed as a repulsive potential making the components $\Delta p_i(\lambda)$ as small as desired when the magnitude of λ is increased. Now, let us denote as $\psi_T^{lin}(\lambda)$ and $E(\lambda)$ the ground-state eigensolutions of the eigenproblem associated with $H(\lambda)$. Here, we propose to use the wavefunction $\psi_T^{lin}(\lambda)$ as a trial wavefunction for the true Hamiltonian H (corresponding to $\lambda = 0$). Doing that, it is easy to verify that the corresponding variational energy, $E^{lin}(\lambda)$, is given by

$$E^{lin}(\lambda) = \frac{\langle \psi_T^{lin}(\lambda) | H | \psi_T^{lin}(\lambda) \rangle}{\langle \psi_T^{lin}(\lambda) | \psi_T^{lin}(\lambda) \rangle} = E(\lambda) - \lambda \frac{\sum_{i=1}^{N_{opt}} \Delta p_i^2(\lambda)}{1 + \sum_{i=1}^{N_{opt}} \Delta p_i^2(\lambda)} \quad (6.15)$$

To get the optimal value of λ the variational energy is minimized. In practice, we evaluate the left-hand side of Eq.(6.15) for different values of λ with a correlated approach using the initial density as reference density

$$E^{lin}(\lambda) = \frac{\langle \frac{H\Psi_T^{lin}}{\Psi_T^{lin}} (\frac{\Psi_T^{lin}}{\Psi_0})^2 \rangle_{\Psi_0^2}}{\langle (\frac{\Psi_T^{lin}}{\Psi_0})^2 \rangle_{\Psi_0^2}}. \quad (6.16)$$

6.4 Evaluating derivatives of the multi-Jastrow trial wavefunction

The quantities to be calculated at each Monte Carlo step of a Variational Monte Carlo (VMC) or Fixed-Node Diffusion Monte Carlo (FN-DMC) simulation (see, *e.g.* [21]) are the values of the trial wavefunction ψ_T , its first derivatives $\nabla\psi_T$, and its Laplacian $\nabla^2\psi_T$ for a given electronic configuration $\mathbf{R} \equiv (\mathbf{r}_1, \dots, \mathbf{r}_{N_{el}})$ (or a set of configurations when a population of walkers is

considered). When optimizing the trial wavefunction as described in Sec.6.3 the derivatives of ψ_T with respect to each parameter, $\frac{\partial \psi_T}{\partial c_p}$, and the Laplacian of these derivatives, $\nabla^2 \frac{\partial \psi_T}{\partial c_p}$ must also be computed. Note that all these quantities being linear in the products of determinants we only need to consider their calculation for one product of determinants, the generalization to an arbitrary number of determinants being trivial. To facilitate the derivations to follow, let us introduce some notations.

$$\Psi_{ij}^\sigma \equiv e^{J_i(\mathbf{r}_j | \mathbf{r}_k \neq \mathbf{r}_j)} \phi_i(\mathbf{r}_j)$$

where $\sigma = \uparrow, \downarrow$. The matrices Ψ_{ij}^σ will be called Slater matrices and their determinants will be denoted as

$$\Psi^\sigma \equiv \det(\Psi_{ij}^\sigma)$$

For the sake of clarity we will suppose that the occupied up and down orbitals are the same (“restricted-type” calculations but using different orbitals is not a problem), so that the generic component whose various derivatives are to be computed can be written as follows

$$\Psi = \Psi^\uparrow \Psi^\downarrow \tag{6.17}$$

Before expliciting the various derivatives, let us present an elementary but fundamental formula we will use in the following to write down most of the expressions. Let M be a finite inversible matrix and δM a general matrix (in practice, δM will be a “small” perturbation of M resulting from the change of one or two columns or lines) it is trivial to verify that

$$\det(M + \delta M) = \det(M) \det(Id + M^{-1} \delta M) \tag{6.18}$$

where $Id : \text{Papier_multijastrow.tex, v1.62009/07/0418 : 06 : 43bouabcaExp}$ is the identity matrix.

6.4.1 Evaluation of Ψ

The computation of Ψ is done as usual way by using standard linear algebra techniques to compute the determinant of the matrices Ψ_{ij}^σ .

6.4.2 Evaluation of $\nabla\Psi$

Let us note

$$\partial_{i,l} \equiv \frac{\partial}{\partial x_i^l}$$

where i denotes the electron index running from $i = 1$ to N_σ , and l the space index, $x^1 \equiv x, x^2 \equiv y, x^3 \equiv z$. Using these notations we have

$$\begin{aligned} \partial_{i,l}\Psi = & \sum_{k=1}^{N_\uparrow} \begin{vmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1k-1}^\uparrow & \partial_{i,l}\Psi_{1k}^\uparrow & \Psi_{1k+1}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \cdots & \Psi_{jk-1}^\uparrow & \partial_{i,l}\Psi_{jk}^\uparrow & \Psi_{jk+1}^\uparrow & \cdots & \Psi_{jN_\uparrow}^\uparrow \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{N_\uparrow 1}^\uparrow & \cdots & \Psi_{N_\uparrow k-1}^\uparrow & \partial_{i,l}\Psi_{N_\uparrow k}^\uparrow & \Psi_{N_\uparrow k+1}^\uparrow & \cdots & \Psi_{N_\uparrow N_\uparrow}^\uparrow \end{vmatrix} \Psi^\downarrow \\ & + \Psi^\uparrow \sum_{k=1}^{N_\downarrow} \begin{vmatrix} \Psi_{11}^\downarrow & \cdots & \Psi_{1k-1}^\downarrow & \partial_{i,l}\Psi_{1k}^\downarrow & \Psi_{1k+1}^\downarrow & \cdots & \Psi_{1N_\downarrow}^\downarrow \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\downarrow & \cdots & \Psi_{jk-1}^\downarrow & \partial_{i,l}\Psi_{jk}^\downarrow & \Psi_{jk+1}^\downarrow & \cdots & \Psi_{jN_\downarrow}^\downarrow \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{N_\downarrow 1}^\downarrow & \cdots & \Psi_{N_\downarrow k-1}^\downarrow & \partial_{i,l}\Psi_{N_\downarrow k}^\downarrow & \Psi_{N_\downarrow k+1}^\downarrow & \cdots & \Psi_{N_\downarrow N_\downarrow}^\downarrow \end{vmatrix} \end{aligned} \quad (6.19)$$

To evaluate the determinants we consider the following decomposition of the matrices corresponding to the determinants \uparrow and \downarrow

$$M^\sigma + \delta M_k^\sigma = \begin{pmatrix} \Psi_{11}^\sigma & \cdots & \Psi_{1k-1}^\sigma & \partial_{i,l}\Psi_{1k}^\sigma & \Psi_{1k+1}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\sigma & \cdots & \Psi_{jk-1}^\sigma & \partial_{i,l}\Psi_{jk}^\sigma & \Psi_{jk+1}^\sigma & \cdots & \Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{N_\sigma 1}^\sigma & \cdots & \Psi_{N_\sigma k-1}^\sigma & \partial_{i,l}\Psi_{N_\sigma k}^\sigma & \Psi_{N_\sigma k+1}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma \end{pmatrix} \quad (6.20)$$

where M^σ is the Slater matrix, Ψ_{ij}^σ , and δM_k^σ a ‘‘perturbing’’ matrix having only the k -th column different from zero :

$$(\delta M_k^\sigma)_j \equiv -\Psi_{jk}^\sigma + \partial_{i,l}\Psi_{jk}^\sigma \quad j = 1, N_\sigma. \quad (6.21)$$

The matrix $Id + [M^\sigma]^{-1}\delta M_k^\sigma$ has thus the following structure

$$Id + [M^\sigma]^{-1}\delta M_k^\sigma = \begin{pmatrix} 1 & 0 & \dots & (P_k^\sigma)_1 & \dots & 0 & 0 \\ 0 & 1 & \dots & (P_k^\sigma)_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & 1 + (P_k^\sigma)_k & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & (P_k^\sigma)_{N_\sigma-1} & \dots & 1 & 0 \\ 0 & 0 & \dots & (P_k^\sigma)_{N_\sigma} & \dots & 0 & 1 \end{pmatrix} \quad (6.22)$$

where the vector (P_k^σ) is defined as

$$(P_k^\sigma)_m = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{mj}^{-1} (\delta M_k^\sigma)_j \quad m = 1, N_\sigma. \quad (6.23)$$

Now, applying formula Eq.(6.18) we get

$$\det(M^\sigma + \delta M_k^\sigma) = \Psi^\sigma [1 + (P_k^\sigma)_k] \quad (6.24)$$

and, finally, the expression of the derivative of Ψ is given by

$$\partial_{i,l}\Psi = \Psi \left\{ \sum_{k=1}^{N_\uparrow} [1 + (P_k^\uparrow)_k] + \sum_{k=1}^{N_\downarrow} [1 + (P_k^\downarrow)_k] \right\}. \quad (6.25)$$

As seen this expression requires the computation of the Slater matrix elements and their derivatives, and also the inverse of the Slater matrices which are computed in standard implementation of QMC algorithms (see, *e.g.* [21]).

6.4.3 Evaluation of $\nabla^2\Psi$

The second derivatives of Ψ can be written as

$$\partial_{i,l}^2\Psi = (\partial_{i,l}^2\Psi^\uparrow)\Psi^\downarrow + \Psi^\uparrow\partial_{i,l}^2\Psi^\downarrow + 2\partial_{i,l}\Psi^\uparrow\partial_{i,l}\Psi^\downarrow \quad (6.26)$$

The first derivatives, $\partial_{i,l}\Psi^\sigma$ are computed as presented in the preceding section. The new terms, $\partial_{i,l}^2\Psi^\sigma$, can be written as

$$\partial_{i,l}^2\Psi^\sigma = \sum_{k=1}^{N_\sigma} \sum_{m=1}^{N_\sigma} \begin{vmatrix} \Psi_{11}^\sigma & \dots & \partial_{i,l}\Psi_{1k}^\sigma & \dots & \partial_{i,l}\Psi_{1m}^\sigma & \dots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots & & \vdots \\ \Psi_{j1}^\sigma & \dots & \partial_{i,l}\Psi_{jk}^\sigma & \dots & \partial_{i,l}\Psi_{jm}^\sigma & \dots & \Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots & & \vdots \\ \Psi_{N_\sigma 1}^\sigma & \dots & \partial_{i,l}\Psi_{N_\sigma k}^\sigma & \dots & \partial_{i,l}\Psi_{N_\sigma m}^\sigma & \dots & \Psi_{N_\sigma N_\sigma}^\sigma \end{vmatrix} \quad (6.27)$$

where the derivatives appear only in columns k and m . Let us distinguish the cases corresponding to $k = m$ and $k \neq m$.

- Case $k = m$

In this case where the column $k = m$ is given by the vector $\partial_{i,l}^2\Psi_{jk}^\sigma$ with $j = 1, N_\sigma$ we are in a situation absolutely similar to what we have treated before for the first derivative. The perturbing matrix δM^σ to consider is the matrix having as unique nonzero column the following vector

$$(\delta M_k^\sigma)_j \equiv -\Psi_{jk}^\sigma + \partial_{i,l}^2\Psi_{jk}^\sigma \quad j = 1, N_\sigma. \quad (6.28)$$

and the value of the second derivative is given by

$$\partial_{i,l}^2\Psi^\sigma = \Psi^\sigma \sum_{k=1}^{N_\sigma} [1 + (P_k^\sigma)_k] \quad (6.29)$$

where

$$(P_k^\sigma)_k = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{kj}^{-1} (\delta M_k^\sigma)_j \quad (6.30)$$

- Case $k \neq m$

This case leads to a new situation where the perturbing matrix δM_{km}^σ consists now of two nonvanishing columns k and m represented by the vectors $-\Psi_{jk}^\sigma + \partial_{i,l}\Psi_{jk}^\sigma$ and $-\Psi_{jm}^\sigma + \partial_{i,l}\Psi_{jm}^\sigma$. Defining as before the vectors P_k^σ and P_m^σ as follows

$$(P_k^\sigma)_n = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{nj}^{-1} (\delta M_k^\sigma)_j$$

$$(P_m^\sigma)_n = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{nj}^{-1} (\delta M_m^\sigma)_j \quad (6.31)$$

the following structure for the matrix $Id + [M^\sigma]^{-1} \delta M_{km}^\sigma$ is obtained

$$Id + [M^\sigma]^{-1} \delta M_{km}^\sigma = \begin{pmatrix} 1 & \dots & (P_k^\sigma)_1 & \dots & 0 & \dots & (P_m^\sigma)_1 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 + (P_k^\sigma)_k & \dots & 0 & \dots & (P_m^\sigma)_k & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & (P_k^\sigma)_m & \dots & 0 & \dots & 1 + (P_m^\sigma)_m & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & (P_k^\sigma)_{N_\sigma} & \dots & 0 & \dots & (P_m^\sigma)_{N_\sigma} & \dots & 1 \end{pmatrix} \quad (6.32)$$

The determinant of this matrix is given by $[(1 + P_k^\sigma)_k][(1 + P_m^\sigma)_m] - (P_m^\sigma)_k (P_k^\sigma)_m$ and, finally, the value of the second derivative can be written as

$$\partial_{i,l}^2 \Psi^\sigma = \sum_{k=1}^{N_\sigma} \sum_{m=1}^{N_\sigma} \Psi^\sigma \{ [1 + (P_k^\sigma)_k](1 + (P_m^\sigma)_m) - (P_m^\sigma)_k (P_k^\sigma)_m \} \quad (6.33)$$

6.4.4 Derivative of Ψ with respect to a molecular Jastrow parameter

In this section we evaluate the derivative of Ψ with respect to a parameter c_p appearing in the expression of a molecular Jastrow term, J_i . Here, we shall only consider the case where this parameter belongs only to a single Jastrow (no common parameter for distinct Jastrows). Note that, common Jastrows can be attached to several different molecular orbitals. Denoting ∂_{c_p} the derivative with respect to c_p we have

$$\partial_{c_i} \Psi = \sum_{k=1}^{N_\uparrow} \begin{vmatrix} \Psi_{11}^\uparrow & \dots & \Psi_{1j}^\uparrow & \dots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & & \vdots & & \vdots \\ \partial_{c_p} \Psi_{k1}^\uparrow & \dots & \partial_{c_p} \Psi_{kj}^\uparrow & \dots & \partial_{c_p} \Psi_{kN_\uparrow}^\uparrow \\ \vdots & & \vdots & & \vdots \\ \Psi_{N_\uparrow 1}^\uparrow & \dots & \Psi_{N_\uparrow j}^\uparrow & \dots & \Psi_{N_\uparrow N_\uparrow}^\uparrow \end{vmatrix} \Psi^\downarrow$$

$$+ \Psi^\uparrow \sum_{k=1}^{N_\downarrow} \begin{vmatrix} \Psi_{11}^\downarrow & \dots & \Psi_{1j}^\downarrow & \dots & \Psi_{1N_\downarrow}^\downarrow \\ \vdots & & \vdots & & \vdots \\ \partial_{c_p} \Psi_{k1}^\downarrow & \dots & \partial_{c_p} \Psi_{kj}^\downarrow & \dots & \partial_{c_i} \Psi_{kN_\downarrow}^\downarrow \\ \vdots & & \vdots & & \vdots \\ \Psi_{N_\downarrow 1}^\downarrow & \dots & \Psi_{N_\downarrow j}^\downarrow & \dots & \Psi_{N_\downarrow N_\downarrow}^\downarrow \end{vmatrix} \quad (6.34)$$

As seen, this expression is similar to that obtained for the gradient of Ψ , Eq.(6.25), except that the derivatives occur on one line and not on a column. It is thus simple to verify that the same derivation applies, we get

$$\partial_{c_p} \Psi = \Psi \left\{ \sum_{k=1}^{N_\uparrow} [1 + (P_k^\uparrow)_k] + \sum_{k=1}^{N_\downarrow} [1 + (P_k^\downarrow)_k] \right\}. \quad (6.35)$$

with

$$(P_k^\sigma)_k = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{kj}^{-1} (\delta M_k^\sigma)_j \quad (6.36)$$

and

$$(\delta M_k^\sigma)_j = -\Psi_{jk}^\sigma + \partial_{c_p} \Psi_{jk}^\sigma \quad j = 1, N_\sigma \quad (6.37)$$

6.4.5 Derivative of $\nabla^2 \Psi$ with respect to molecular Jastrow parameter

The derivative of $\nabla^2(\Psi^\uparrow \Psi^\downarrow)$ with respect to a Jastrow's parameter c_p can be decomposed in a series of elementary terms. Most of these terms have been already derived in the previous sections. The remaining terms to be computed are the gradient and Laplacian of the derivative of Ψ^σ , namely $\nabla \frac{\partial \Psi^\sigma}{\partial c_p}$ and $\nabla^2 \frac{\partial \Psi^\sigma}{\partial c_p}$.

Evaluation of $\nabla \frac{\partial \Psi^\sigma}{\partial c_p}$

Using the notations introduced above the gradient of Ψ^σ is written as

$$\partial_{i,l} \Psi^\sigma = \sum_{k=1}^{N_\sigma} \begin{vmatrix} \Psi_{11}^\sigma & \cdots & \Psi_{1k-1}^\sigma & \partial_{i,l} \Psi_{1k}^\sigma & \Psi_{1k+1}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\sigma & \cdots & \Psi_{jk-1}^\sigma & \partial_{i,l} \Psi_{jk}^\sigma & \Psi_{jk+1}^\sigma & \cdots & \Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{N_\sigma 1}^\sigma & \cdots & \Psi_{N_\sigma k-1}^\sigma & \partial_{i,l} \Psi_{N_\sigma k}^\sigma & \Psi_{N_\sigma k+1}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma \end{vmatrix} \quad (6.38)$$

and the derivative with respect to the Jastrow's parameter c_p as

$$\partial_{c_p} \partial_{i,l} \Psi^\sigma = \sum_{j=1}^{N_\sigma} \sum_{k=1}^{N_\sigma} \begin{vmatrix} \Psi_{11}^\sigma & \cdots & \partial_{i,l} \Psi_{1k}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \partial_{c_p} \Psi_{j1}^\sigma & \cdots & \partial_{c_p} \partial_{i,l} \Psi_{jk}^\sigma & \cdots & \partial_{c_p} \Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \Psi_{N_\sigma 1}^\sigma & \cdots & \partial_{i,l} \Psi_{jk}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma \end{vmatrix} \quad (6.39)$$

To compute the new determinants appearing in this expression, the fundamental formula, Eq.(6.18), is used with

$$M^\sigma = \begin{pmatrix} \Psi_{11}^\sigma & \cdots & \Psi_{1k}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \partial_{c_p} \Psi_{j1}^\sigma & \cdots & \partial_{c_p} \Psi_{jk}^\sigma & \cdots & \Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \Psi_{N_\sigma 1}^\sigma & \cdots & \Psi_{jk}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma \end{pmatrix} \quad (6.40)$$

and

$$\delta M^\sigma = \begin{pmatrix} 0 & \cdots & -\Psi_{1k}^\sigma + \partial_{i,l} \Psi_{1k}^\sigma & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & -\partial_{c_p} \Psi_{jk}^\sigma + \partial_{c_p} \partial_{i,l} \Psi_{jk}^\sigma & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & -\Psi_{jk}^\sigma + \partial_{i,l} \Psi_{jk}^\sigma & \cdots & 0 \end{pmatrix} \quad (6.41)$$

How to calculate the determinant of M has already been presented. Now, to employ Eq.(6.18) we also need to evaluate the inverse of M . To avoid costly

N_σ^3 evaluations using standard methods we propose to employ the Sherman-Morison formula.[14] Let us denote A an inversible matrix and u and v two vectors, this formula is written as

$$(A + u \otimes v)^{-1} = A^{-1} - \frac{z \otimes w}{1 + \lambda} \quad (6.42)$$

where

$$\lambda \equiv (v, A^{-1}u), \quad z \equiv A^{-1}u, \quad \text{and} \quad w \equiv (A^{-1})^T v$$

Here, to calculate the inverse of matrix M we apply the Sherman-Morison formula with the two vectors

$$u_i = -\Psi_{li}^\sigma + \frac{\partial \Psi_{li}^\sigma}{c_l} \quad (6.43)$$

and

$$v_i = \delta_{il} \quad (6.44)$$

Evaluation of $\nabla_i^2 \frac{\partial \Psi^\sigma}{\partial c_p}$

To compute $\nabla_i^2 \frac{\partial \Psi^\sigma}{\partial c_p}$ it is easy to check that we need to evaluate the two types of determinant

$$\begin{vmatrix} \Psi_{11}^\sigma & \dots & \Delta_i(\Psi_{1k}^\sigma) & \dots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \partial_{c_p} \Psi_{j1}^\sigma & \dots & \partial_{c_p} \Delta_i(\Psi_{jk}^\sigma) & \dots & \partial_{c_p} \Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \Psi_{N_\sigma 1}^\sigma & \dots & \Delta_i(\Psi_{jk}^\sigma) & \dots & \Psi_{N_\sigma N_\sigma}^\sigma \end{vmatrix} \quad (6.45)$$

and

$$\begin{vmatrix} \Psi_{11}^\sigma & \dots & \partial_{i,l} \Psi_{1k}^\sigma & \dots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \partial_{c_p} \Psi_{j1}^\sigma & \dots & \partial_{c_p} \partial_{i,l} \Psi_{jk}^\sigma & \dots & \partial_{c_p} \Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \Psi_{N_\sigma 1}^\sigma & \dots & \partial_{i,l} \Psi_{jk}^\sigma & \dots & \Psi_{N_\sigma N_\sigma}^\sigma \end{vmatrix} \quad (6.46)$$

and we make use here also of our fundamental formula combined with the Sherman-Morison formula.

6.5 Results

6.5.1 Atoms

As a first application of the multi-Jastrow wavefunction we have considered atomic systems. Numerous studies on atoms using QMC techniques have been presented in the last twenty years (see, for example, [15]). Here, we shall consider two systems, namely the 3P ground-state oxygen atom ($1s^2 2s^2 2p^4$, 8 electrons) and the 2S ground-state copper atom ($[Ar]3d^{10}4s$, 29 electrons).

The oxygen atom

Table 9.1 presents our results for the oxygen atom. The total energies obtained at the VMC and FN-DMC levels with different trial wavefunctions are presented. Two different criteria to quantify the quality of the trial wavefunctions are employed : the variance of the energy, $\sigma^2(Energy)$, and the amount of correlation energy recovered. The variance of the energy is defined as the variance of the local energy during the simulation

$$\sigma^2(Energy) \equiv \langle E_L^2 \rangle - \langle E_L \rangle^2 \quad (6.47)$$

where the symbol $\langle \dots \rangle$ denotes the Monte Carlo average (here, either VMC or FN-DMC) and E_L is the local energy defined as $E_L = \frac{H\Psi_T}{\Psi_T}$. Note that the variance of the energy is a natural quantity to introduce in a QMC framework since the statistical error is directly proportional to it. In particular, the lower the variance of the local energy is, the “closer” the trial wavefunction is from the exact wavefunction. The second criterium is standard in quantum chemistry : The correlation energy is defined as the difference between the exact non-relativistic Born-Oppenheimer energy and the Hartree-Fock energy obtained with an infinite one-particle basis set (complete basis set limit or “Hartree-Fock limit”).

At the VMC level, three different trial wavefunctions have been employed. The first one is a standard Hartree-Fock wavefunction built from the atomic orbitals expressed over Slater-Type Orbitals (STO) as given by Bunge et

al.[16]. As it should be, the VMC result recovers within statistical uncertainties the analytical variational energy. The second trial wavefunction has a standard form consisting of a Slater determinant multiplied by a Jastrow factor. All 19 parameters of the Jastrow term have been optimized, while the atomic orbitals have been kept unchanged. The optimization leads to a gain of 83.6 % of the correlation energy, a typical result for light atoms (see, *e.g.* Ref.[15] where Schmidt and Moskowitz have recovered about 80. % of the correlation energy for the oxygen atom). The optimized parameters for the global Jastrow term are listed in Table 9.2. Finally, the last trial wavefunction is our multi-Jastrow wavefunction. Using a trial-and-error approach it has been found that only two different atomic Jastrows were sufficient to represent the atom. The first one attached to the innermost $1s$ orbital is called the “ $1s$ -Jastrow” while the second one attached to the $2s$ and $2p$ atomic orbitals will be referred to as the valence Jastrow. After full optimization of the two atomic Jastrow terms (12 parameters each) the VMC energy obtained is only slightly lower than the energy resulting from the use of a common Jastrow. The various parameters obtained for the atomic Jastrow terms are listed in Table 9.3. This result is of course disappointing but illustrates that, at least for the oxygen atom, making the difference between the core and valence regions for the Jastrow correlations does not lead to significant improvements.

This conclusion also extends to the fixed-node DMC results. Indeed, as seen in Table 9.1 the FN-DMC results obtained with Hartree-Fock or multi-Jastrow nodes are identical within statistical uncertainties. This result seems to indicate that for this simple atomic system the nodal properties of the trial wavefunction are not improved when introducing different atomic Jastrow terms.

The copper atom

The case of the copper atom is much more interesting since for a transition metal atom with many d electrons occupying a small region of space,

it is expected that electronic correlations may differ drastically from one shell to the other. The VMC and FN-DMC results obtained are presented in Table 6.4. At the VMC level with the standard Jastrow-Slater form for the trial wavefunction, about 44% of the correlation energy is recovered. In sharp contrast, our multi-Jastrow form allows to recover much more, namely about 75% of the correlation energy. This result is clearly very satisfactory and gives strong support for the validity of distinguishing Jastrow correlations according to the different shells in a structured system such as the copper atom. To get this result, three different atomic Jastrow terms have been considered : one for the $1s$ core orbital, another for the $2s2p3s3p$ orbitals and the last one for the $3d$ and $4s$ orbitals. Let us emphasize that, to the best of our knowledge, the total variational energy obtained here for the copper atom, $E = -1640.173(7)$ is very likely the best energy ever obtained for this atom with an explicit trial wavefunction consisting of *only one* determinant.

Unfortunately, like for the oxygen atom, the fixed-node results presented in Table 6.4 do not show any improvement of the nodes, despite the important improvement regarding the variational energy.

6.5.2 The HF molecule

As a first molecular example, the calculation of the potential energy curve of the HF molecule has been considered. In particular, the three basic spectroscopic constants describing the overall shape of the curve : the equilibrium distance r_e , the harmonic frequency ω_e , and the dissociation energy D_e are computed.

Minimal CASSCF wavefunction

To describe the one-particle Slater part of the trial wavefunction a minimal form leading to the correct dissociation of the molecule into its neutral atomic fragments has been considered. More precisely, the Slater part is given

by the following two-determinant form

$$\Psi_{Tdet} = [\sigma_{1s}^2 \sigma_{2s}^2 \pi_x^2 \pi_y^2] \sigma^2 + [\sigma_{1s}^2 \sigma_{2s}^2 \pi_x^2 \pi_y^2] \sigma^{*2} \quad (6.48)$$

where the molecular orbitals σ_{1s} , σ_{2s} , π_x , and π_y have a strong atomic character corresponding to the fluor atom and (σ, σ^*) are the bonding/anti-bonding valence orbitals built for the most part from the $1s^H$ and $2p_z^F$ atomic orbitals.

In Figure 6.1 the various CASSCF curves obtained with this trial wavefunction and basis sets of increasing size are presented (correlation-consistent basis sets of Dunning and coworkers, $cc - pVnZ$ with $n = 2$ to $n = 5$, [17]) and compared to the estimate exact non-relativistic curve.[18] As seen, the size of the basis set has a strong impact on the overall shape of the energy curves. In our example, results seem to be converged only when basis sets corresponding to $n = 4$ and beyond are considered. To have a more quantitative assessment of the quality of these curves, the spectroscopic constants corresponding to each curve are presented in Table 6.7. For completeness, we also give the SCF results corresponding to the mono-determinantal configuration $[\sigma_{1s}^2 \sigma_{2s}^2 \pi_x^2 \pi_y^2] \sigma^2$.

As well-known, SCF results lead to too short equilibrium distances and to too large harmonic frequencies. Here, for large basis sets r_e is too small by about 2% and ω_e is too large by about 16%. The reason for that is related to the wrong curvature of the SCF potential energy curve due to the incorrect dissociation of the SCF representation (energy curve too “stiff” at large distances because of the undesirable high-energy ionic components of the SCF wavefunction). Using now a minimal CASSCF wavefunction, as in Eq.6.48, this problem is easily avoided and both equilibrium distance and harmonic frequency are greatly improved with an error of 0.25% (2% in the SCF case) and 1% (16% in the SCF case), respectively. Regarding dissociation energies, the CASSCF results as a function of the basis set is found to converge to about 115 kcal/mol, a quantity much too small by about 26 kcal/mol. Note that the SCF dissociation energy cannot not defined due to the wrong dissociation property.

At this point, it should be emphasized that using trial wave functions

having the correct dissociation property is an important point for our study. Indeed, to get a clear assessment of the quality of our new proposed form it is essential to properly distinguish between the various types of electron correlations at work. In particular, the quality of the nodal pattern of the multi-Jastrow wave function is expected to be intimately related to the ability of such a wavefunction to correctly describe both non-dynamical and dynamical correlation effects, although it is clearly difficult to make such a statement quantitative. By introducing various molecular Jastrow factors to describe the instantaneous correlation effects in molecular orbitals, the focus here is essentially on the dynamical correlation effects and very little on the static ones (related to near-degeneracy effects). Accordingly, introducing the major part of these latter effects via appropriate CASSCF representation is essential to get a well-balanced description of the wavefunction.

Fixed-Node DMC results

Let us begin by discussing the FN-DMC results obtained with trial wavefunctions having the nodal pattern of the CASSCF wavefunctions, Eq.6.48. Such calculations can be considered as standard since they are based on the widely used strategy consisting in using trial wavefunctions of the Jastrow-Slater type. Let us look at the impact of the fixed-node error on such a molecular system. To see the effect of the basis set on the CASSCF nodal pattern we present in Fig.6.2 the various curves obtained using the different basis sets introduced above. The corresponding spectroscopic constants are presented in Table 6.8. Note that the extraction of such constants from noisy (Monte Carlo) potential energy curves requires some care. Here, to calculate them the potential energy curves have been systematically fitted using a five-parameter generalized Morse potential form as follows

$$E(r) = p_1 + (p_2 - p_1)e^{-p_3x} + p_5(e^{-2p_3x} - e^{-p_3x}) \quad (6.49)$$

with

$$x = \frac{r - p_4}{p_4}. \quad (6.50)$$

In this formula each parameter p_i is directly related to one of the main features of the energy curve. In this way, the fit is “physically” meaningful and is neither redundant nor insufficient. The parameter p_1 is directly related to $E(r \rightarrow +\infty)$, p_2 to $E(r_e)$, p_3 to the curvature at the minimum, p_4 to r_e , and p_5 to the way the curve goes up at small r (or, equivalently, to the asymmetry of the curve around the minimum).

A first remark concerning the fixed-node energy curves for HF is that the DMC results are rather accurate. For example, at the minimum, the total correlation energy recovered with the largest basis set is about 94%. As seen in Fig.6.2 the fixed-node curves have a clear dependence on the dimension of the basis set. At the DZ level (smallest basis set), the fixed-node error is much larger than with the other basis sets and we conclude that such a basis is not appropriate. In the neighborhood of the minimum and for the three largest basis sets (TZ, QZ, and 5Z) the dependence is much weaker and similar results are obtained. However, for larger distances ($R \gtrsim 2$.) it is no longer the case and the nodal pattern appears not to be fully stabilized as a function of the basis set. To get a more quantitative view, let us now turn our attention to the spectroscopic constants presented in Table 6.8. As a function of the basis set the dissociation energy is found to converge around -138.kcal/mol. This value is greatly improved with respect to the uncorrelated value of -115 kcal/mol presented in Table 6.8 but is still in error by a contribution of about 4 kcal/mol. Now, the equilibrium distances and the harmonic frequencies are acceptable but are found not to be really stabilized as a function of the basis set used. This fact is a direct consequence of the lack of stabilization of the DMC energies at large distances already pointed out.

Finally, let us look at the results obtained with the multi-Jastrow wavefunction. In order to take into account the correlation effects occurring in the valence region, the wavefunction has been built by using only one molecular Jastrow term attached to the valence molecular orbitals. The various curves obtained using various basis sets are presented in Fig.6.3 and the correspon-

ding spectroscopic values are presented in Table 6.8. The global shape of the curves are similar to the curves obtained previously. However, some important differences can be observed. First, the dependence of the curves on the size of the basis set is different, particularly at large interatomic distances. In contrast with the standard case, the FN-DMC results for the multi-Jastrow case are found to be converged as a function of the basis set at all distances. Second, the value of the dissociation energy is much improved. The value obtained with the 5Z basis set, $D_e = -140.5(4)$ is very good and is only about 1kcal/mol different from the experimental one. Regarding the values of r_e and ω_e the values are reasonable but do not seem really improved with the standard fixed-node case. Clearly, to make a more accurate evaluation of these quantities we would need more FN-DMC data around the minimum and larger statistics to decrease the noise; however, such a study is out of the scope of the present work.

6.6 Summary and perspectives

In this work we have presented a new trial wavefunction for QMC simulations of molecular systems. This wavefunction is similar to the standard Slater-Jastrow wavefunction except that the common global Jastrow factor is replaced by several Jastrow factors attached to individual molecular orbitals. The physical motivation behind this “multi-Jastrow” form is to allow an orbital-dependent description of the correlation effects in atoms and molecules. Applications to atoms have illustrated that the optimized multi-Jastrow form can become particularly good when the number of electrons is large. For example, in the case of the copper atom (the heavier atom considered here), using three different atomic Jastrow terms (for the 1s, (2s-2p-3s-3p), and the (3d-4s) orbitals) leads to a total energy corresponding to about 75% of the correlation energy. Such a high-quality result obtained with a wavefunction comprising only one determinant is remarkable and represents, to the best of our knowledge, the best variational energy ever obtained for the copper atom with a simple trial wavefunction. Regarding fixed-node re-

sults, it has been found for the atoms considered that the multi-Jastrow wavefunction does not change the fixed-node energies with respect to standard calculations using Hartree-Fock nodes. Application to the HF molecule has shown that, in contrast with atoms, the nodal pattern can be changed and that a very good dissociation energy can be obtained. Regarding now computational aspects, it has been shown that by using the suitable strategy the computational overhead associated with the use of multi-Jastrow wavefunctions is limited and can be kept under control.

Finally, regarding perspectives it is important to note that since the major part of (dynamical) correlations effects are local, it seems natural to put together the use of our molecular Jastrow factors with the use of localized molecular orbitals. Doing that, we can imagine to define and pre-optimize once for all some pairs of local Jastrow term plus localized MO's for various chemical situations. By introducing such local elementary pieces, we can hope to be able of building in a "molecular Lego way" trial wavefunctions of good quality for arbitrary complex molecular systems. Work along this direction is presently under way.

Acknowledgments M.C. would like to acknowledge a preliminary discussion with B.Braïda (Univ. of Paris VI) about the possibility of using a multi-Jastrow form for molecules. The authors would like to thank IDRIS (CNRS, Orsay), CCRT (CEA/DAM, Ile-de-France), and CALMIP (Université de Toulouse) for generous computational support.

Bibliographie

- [1] R. Assaraf and M. Caffarel, J. Chem. Phys. **119**, 10536 (2003).
- [2] D. M. Ceperley and B. J. Alder, J. Chem. Phys. **64**, 5833 (1984).
- [3] P.J. Reynolds, D.M. Ceperley, B.J. Alder, and W.A. Lester Jr., J. Chem. Phys. **77**, 5593 (1982).
- [4] C. A. Coulson and A. H. Neilson, Proc. Phys. Soc. **78**, 831 (1961).
- [5] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **128**, 174101 (2008).
- [6] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004).
- [7] L. Mitáš, Phys. Rev. Lett. **96**, 240402 (2006).
- [8] P. Lopez Rios, A. Ma, N.D. Drummond, M.D. Towler, and R.J. Needs Phys. Rev. E **74**, 066701 (2006).
- [9] F. A. Matsen, "Spin-free quantum chemistry," Adv. Quantum Chem. , 59 (1964).
- [10] C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).
- [11] C.J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R.G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007).
- [12] M. P. Nightingale and V. Melik-Alaverdian, Phys. Rev. Lett. **87**, 043401 (2001).
- [13] B.L. Hammond, W.A. Lester Jr., and P.J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, World Scientific Lecture and Course Notes in Chemistry Vol.1 (1994).

- [14] See, *e.g.*, in William H. Press, Brian P. Flannery, Saul A. Teukolsky, William T. Vetterling, Numerical Recipes in C : The Art of Scientific Computing, pp.73+. Cambridge University Press (1992).
- [15] K. E. Schmidt and J. W. Moskowitz, *J. Chem. Phys.* **93**, 4172 (1990).
- [16] C. F. Bunge, J. A. Barrientos, A. Viver Bunge, and J. A. Cogordan, *Phys. Rev. A* **46**, 3691 (1992).
- [17] T.H. Dunning, Jr. *J. Chem. Phys.* **90**, 1007 (1989).
- [18] H. Müller, R. Franke, S. Vogtner, R. Jaquet, and W. Kutzelnigg, *Theor. Chem. Acc.* **100**, 85 (1998).
- [19] S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia, and C. Froese Fischer, *Phys. Rev. A* **47**, 3649 (1993).

TAB. 6.1 – Variational Monte Carlo (VMC) and Fixed-Node Diffusion Monte Carlo (FN-DMC) total ground-state energies for the 3P oxygen atom using different trial wavefunctions and nodal structures.

Method	Energy (a.u.)	$\sigma^2(\text{Energy})$	Correlation energy
Hartree-Fock[16]	-74.8094	-	0%
“Exact” non-relativistic[19]	-75.0673	-	100%
VMC :			
Hartree-Fock (no Jastrow of any type)	-74.808(2)	15.9(3)	0.% \pm 0.8%
Hartree-Fock + optimized common Jastrow	-75.025(1)	1.178(7)	83.6% \pm 0.4%
Hartree-Fock + optimized molecular Jastrows	-75.0275(4)	1.133(8)	84.6% \pm 0.2%
FN-DMC :			
Hartree-Fock nodes	-75.052(1)		94.1% \pm 0.4%
Molecular Jastrow nodes	-75.052(1)		94.1% \pm 0.4%

TAB. 6.2 – Optimized parameters of the global Jastrow term for the 3P oxygen atom. The corresponding variational energy is -75.025(1).

Parameters	
$(e_{\sigma\sigma}^{(1)}, e_{\sigma\bar{\sigma}}^{(1)})$	(0.23746, 0.52828)
$(b_{\sigma\sigma}^{(e-e)}, b_{\sigma\bar{\sigma}}^{(e-e)})$	(0.94149, 1.03020)
$f^{(1)}$	0.04367
$b^{(e-n)}$	0.78642
$(e_{\sigma\sigma}^{(2)}, e_{\sigma\bar{\sigma}}^{(2)})$	(-0.34807, -0.63983)
$(e_{\sigma\sigma}^{(3)}, e_{\sigma\bar{\sigma}}^{(3)})$	(0.72185, 0.91421)
$(e_{\sigma\sigma}^{(4)}, e_{\sigma\bar{\sigma}}^{(4)})$	(-0.82900, -1.53998)
$f^{(2)}$	-0.13863
$f^{(3)}$	0.11238
$f^{(4)}$	2.11583
$(g_{\sigma\sigma}^{(1)}, g_{\sigma\bar{\sigma}}^{(1)})$	(-0.84541, -1.20295)
$(g_{\sigma\sigma}^{(2)}, g_{\sigma\bar{\sigma}}^{(2)})$	(0.61555, 1.33749)

TAB. 6.3 – Optimized parameters for the two atomic Jastrow terms for the 3P oxygen atom. The corresponding variational energy is -75.0275(4).

Parameters	1s-Jastrow	Valence Jastrow
$e_1^{(1)}$	0.56220	0.26483
$b_1^{(e-e)}$	1.09867	1.61627
$f_1^{(1)}$	0.02286	0.04952
$b_1^{(e-n),1}$	1.77005	0.84338
$e_1^{(2)}$	-0.54624	0.47433
$e_1^{(3)}$	0.08169	1.59447
$e_1^{(4)}$	-1.20430	-1.73644
$f_1^{(2)}$	0.57051	1.03889
$f_1^{(3)}$	-0.37920	-0.29200
$f_1^{(4)}$	1.77559	1.37422
$g_1^{(1)}$	-1.74206	-0.92854
$g_1^{(2)}$	0.91483	3.82277

TAB. 6.4 – Variational Monte Carlo (VMC) and Fixed-Node Diffusion Monte Carlo (FN-DMC) total ground-state energies for the 2S copper atom using different trial wavefunctions and nodal structures.

Method	Energy (a.u.)	$\sigma^2(\text{Energy})$	Correlation energy
Hartree-Fock[16]	-1638.9637	-	0%
“Exact” non-relativistic[19]	-1640.5677	-	100%
VMC :			
Hartree-Fock (no Jastrow of any type)	-1638.970(6)	423(7)	0.% \pm 0.8%
Hartree-Fock + optimized common Jastrow	-1639.672(5)	587(14)	44.2% \pm 0.3%
Hartree-Fock + optimized molecular Jastrows	-1640.1725(74)	131(3)	75.4% \pm 0.5%
FN-DMC :			
Hartree-Fock nodes	-1640.401(9)		89.6% \pm 0.6%
Molecular Jastrow nodes	-1640.407(8)		90.0% \pm 0.5%

TAB. 6.5 – Optimized parameters of the global Jastrow term for the 2S copper atom. The corresponding variational energy is -1639.672(5).

Parameters	
$(e_{\sigma\sigma}^{(1)}, e_{\sigma\bar{\sigma}}^{(1)})$	(0.21339,0.44788)
$(b_{\sigma\sigma}^{(e-e)}, b_{\sigma\bar{\sigma}}^{(e-e)})$	(2.75707,1.10234)
$f^{(1)}$	0.65641
$b^{(e-n)}$	0.05134
$(e_{\sigma\sigma}^{(2)}, e_{\sigma\bar{\sigma}}^{(2)})$	(1.77743,-0.93810)
$(e_{\sigma\sigma}^{(3)}, e_{\sigma\bar{\sigma}}^{(3)})$	(-0.00195,1.02403)
$(e_{\sigma\sigma}^{(4)}, e_{\sigma\bar{\sigma}}^{(4)})$	(0.000036,0.00277)
$f^{(2)}$	0.03779
$f^{(3)}$	-0.02917
$f^{(4)}$	0.00308
$(g_{\sigma\sigma}^{(1)}, g_{\sigma\bar{\sigma}}^{(1)})$	(0.000117,-0.000591)
$(g_{\sigma\sigma}^{(2)}, g_{\sigma\bar{\sigma}}^{(2)})$	(-0.01993,0.000028)

TAB. 6.6 – Optimized parameters for the three atomic Jastrow terms for the 2S copper atom. The corresponding variational energy is -1640.1725(74).

Parameters	Core Jastrow	2s-2p-3s Jastrow	3d-4s Jastrow
$e_1^{(1)}$	0.85507	0.64044	0.60019
$b_1^{(e-e)}$	4.70071	2.7002	0.95653
$f_1^{(1)}$	0.08045	0.4055	1.13849
$b_1^{(e-n)}$	0.40395	0.4763	0.58689
$e_1^{(2)}$	-0.17918	-0.2813	-0.65535
$e_1^{(3)}$	-0.02402	-0.1085	0.43481
$e_1^{(4)}$	-0.00238	-0.0245	-0.05078
$f_1^{(2)}$	0.66866	0.1698	0.34402
$f_1^{(3)}$	0.08836	-0.4218	-0.19808
$f_1^{(4)}$	0.00666	-0.2467	-0.26311
$g_1^{(1)}$	-0.15821	0.00019	-0.17380
$g_1^{(2)}$	0.21130	-0.8553	-0.00218

TAB. 6.7 – Spectroscopic constants of the HF molecule at the SCF and minimal CASSCF levels as a function of the basis set.

Method	R_e (a.u.)	ω_e (cm^{-1})	ΔE (kcal/mol)
SCF (cc-pVDZ basis)	1.7042	4727.	-
SCF (cc-pVTZ basis)	1.6971	4829.	-
SCF (cc-pVQZ basis)	1.6956	4806.	-
SCF (cc-pV5Z basis)	1.6953	4805.	-
2-det CASSCF (cc-pVDZ basis)	1.7412	3960.	109.27
2-det CASSCF (cc-pVTZ basis)	1.7309	4083.	113.88
2-det CASSCF (cc-pVQZ basis)	1.7283	4088.	114.80
2-det CASSCF (cc-pV5Z basis)	1.7283	4096.	115.04
Exact	1.7326	4137	141.95

TAB. 6.8 – Spectroscopic constants of the HF molecule with fixed-node Diffusion Monte Carlo (DMC) using various nodal patterns. DZ,TZ,QZ, and 5Z refer to the use of the cc-pVnZ basis set with n=2,3,4, and 5, respectively.

Method	R_e (a.u.)	ω_e (cm ⁻¹)	ΔE (kcal/mol)
DMC with 2-det CASSCF (DZ basis) nodes	1.7042	4508.	131.9(3)
DMC with 2-det CASSCF (TZ basis) nodes	1.7186	4378.	137.7(3)
DMC with 2-det CASSCF (QZ basis) nodes	1.7368	3915.	138.0(4)
DMC with 2-det CASSCF (5Z basis) nodes	1.7257	3782.	138.2(4)
DMC with 2-det CASSCF (DZ basis) + valence Jastrow nodes	1.7209	3927.	135.3(3)
DMC with 2-det CASSCF (TZ basis) + valence Jastrow nodes	1.7146	4706.	139.4(4)
DMC with 2-det CASSCF (QZ basis) + valence Jastrow nodes	1.7248	4304.	140.3(4)
DMC with 2-det CASSCF (5Z basis) + valence Jastrow nodes	1.7174	4160	140.5(4)
Exact	1.7326	4137	141.95

FIGURE CAPTION

- Fig. 6.1 Potential energy curves of the HF molecule calculated at the 2-det CASSCF level using various basis sets : cc-pVnZ, n=2,3,4, and 5. The “exact” non-relativistic curve at the bottom of the figure is obtained from [18].
- Fig. 6.2 Potential energy curves of the HF molecule calculated at the Fixed-Node DMC level using as trial wavefunction 2-det CASSCF wavefunctions obtained from various basis sets : cc-pVnZ, n=2,3,4, and 5. The “exact” non-relativistic curve at the bottom of the figure is obtained from [18].
- Fig. 6.3 Potential energy curves of the HF molecule calculated at the Fixed-Node DMC level using as trial wavefunction the multi-Jastrow wavefunction built from 2-det CASSCF wavefunction with various basis sets : cc-pVnZ, n=2,3,4, and 5. The “exact” non-relativistic curve at the bottom of the figure is obtained from [18].

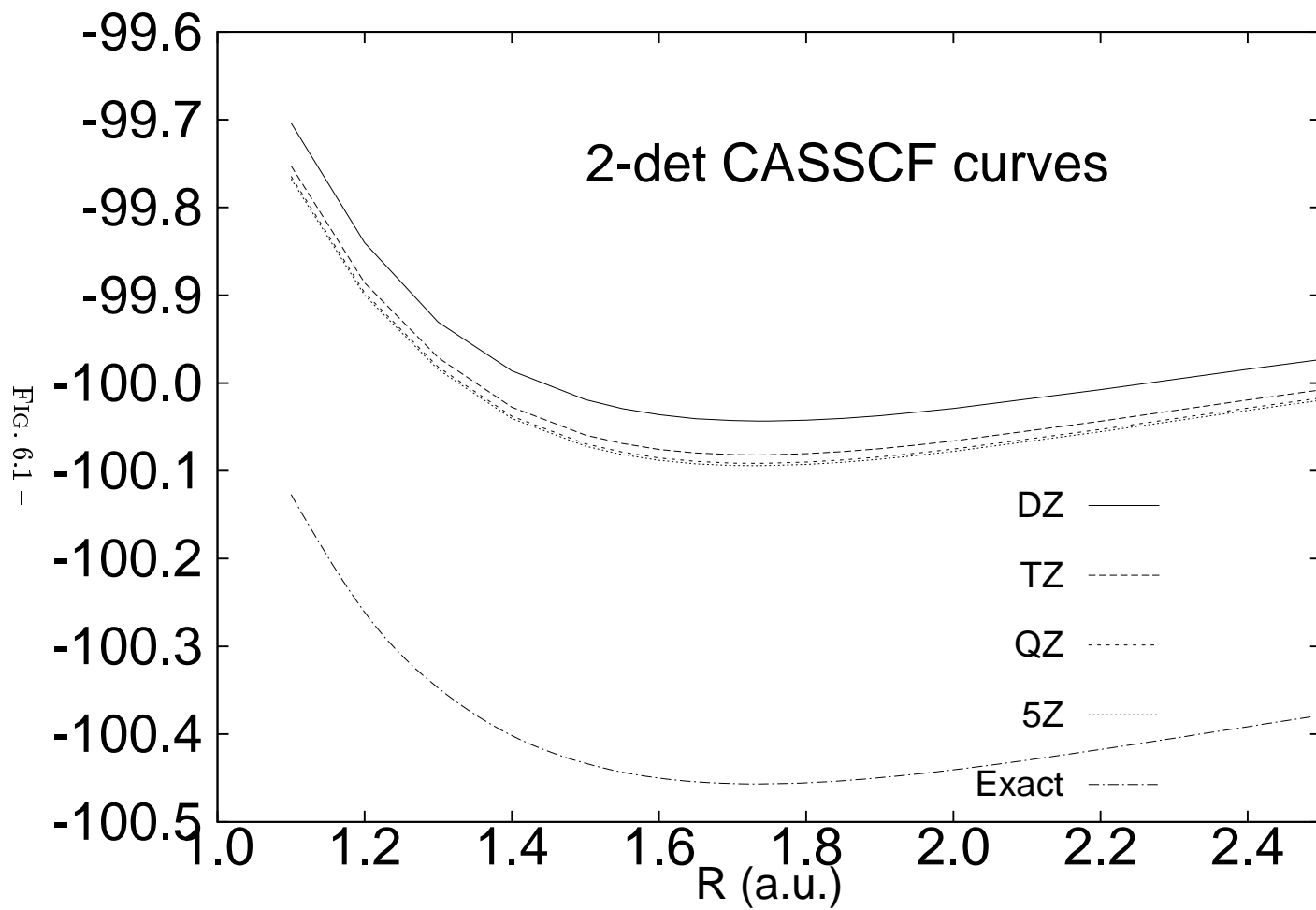


FIG. 6.1 —

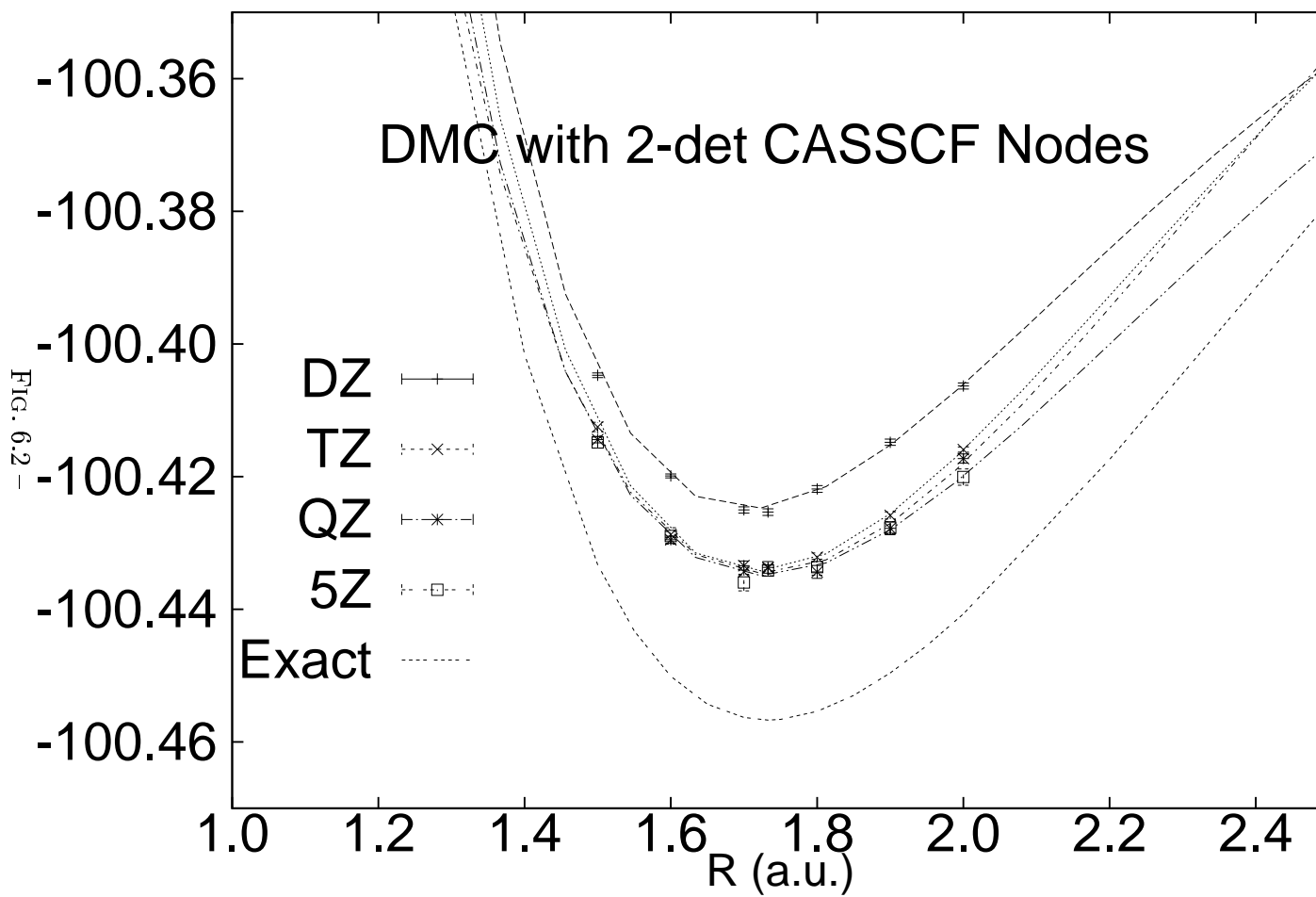


FIG. 6.2 -

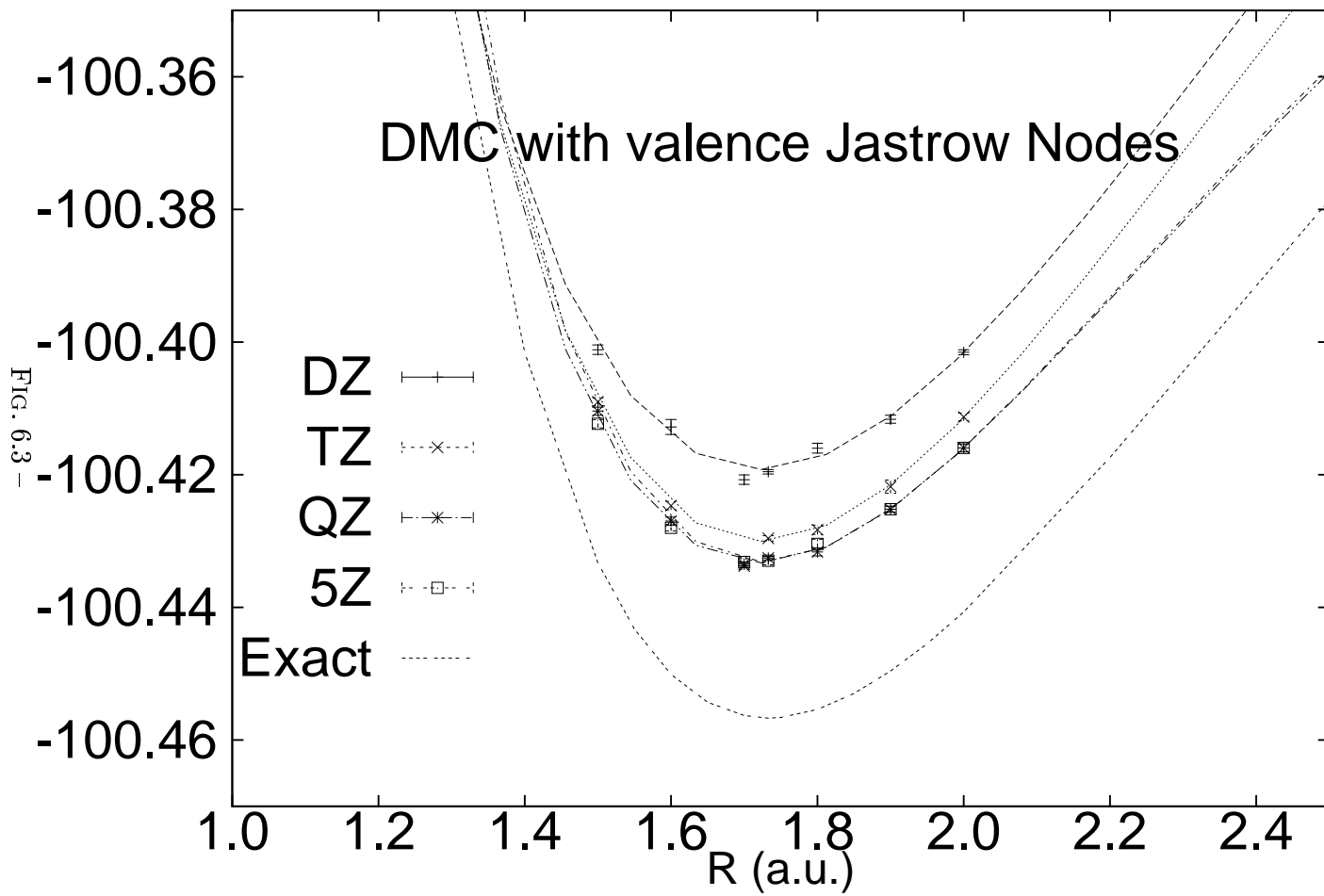


FIG. 6.3

Chapitre 7

Coût calculatoire de la fonction d'onde multi-Jastrow

Comme nous l'avons vu précédemment, la fonction d'onde multi-Jastrow a un contenu physique beaucoup plus pertinent que la fonction d'onde usuelle. Malheureusement, le prix à payer est un surcoût numérique lié à l'évaluation de nombreux déterminants. Ce chapitre a pour but d'évaluer précisément ce surcoût. Dans le chapitre suivant les performances pratiques seront illustrées sur des exemples concrets. Nous allons maintenant montrer que, malgré le nombre important de déterminants à calculer, il est possible de mettre en place un algorithme d'évaluation très performant. Pour cela, nous allons comparer le nombre total de déterminants différents qu'il faut calculer dans le cas d'un Jastrow commun et dans le cas de Jastrow orbitales. Les parties suivent l'ordre chronologique de l'implémentation dans notre code Monte Carlo quantique. Cela signifie que lorsqu'une quantité requièrera des déterminants déjà calculés précédemment, nous ne les compterons pas comme des nouveaux déterminants. C'est d'ailleurs la stratégie que nous avons adoptée dans le code.

Rappelons que dans le cas d'un Jastrow commun, les coordonnées d'un électron i ne sont présentes que dans la colonne i . Ainsi, pour dériver suivant une des coordonnées de \mathbf{r}_i , il faudra dériver la matrice selon une colonne.

Dans le cas d'un Jastrow orbitalaire, les coordonnées d'un électron i sont présentes dans toutes les colonnes puisqu'elles apparaissent dans les Jastrows. Pour dériver suivant une des coordonnées, il faudra donc dériver la matrice suivant toutes les colonnes. De même, dans le cas du Jastrow commun, les paramètres du Jastrow par rapport auxquels on veut dériver n'apparaissent qu'à l'extérieur de la matrice dans le facteur e^J . Dans le cas de l'utilisation de Jastrow orbitaires, les paramètres d'un Jastrow apparaissent dans toutes les lignes où le Jastrow est présent.

Par simplicité, dans les dérivations qui vont suivre nous nous sommes limités à des fonctions d'onde mono-configurationnelles constituées d'un seul produit de deux déterminants associés aux électrons \uparrow (α) et \downarrow (β). Pour des fonctions multiconfigurationnelles du type

$$\Psi = \sum_{k=1, N_{det}} c_k \Psi_k$$

il faudra multiplier les résultats par N_{det} , c'est à dire le nombre de termes de la somme. Remarquons néanmoins qu'en pratique ce facteur est très pessimiste puisqu'en présence de plusieurs déterminants il est possible de calculer une seule fois un certain nombre de parties communes, mais nous n'entrons pas ici dans ces détails techniques. Dans ce qui suit, nous utiliserons les notations suivantes :

- N_e nombre total d'électrons
 - N_\uparrow nombre d'électrons de spin up
 - N_\downarrow nombre d'électrons de spin down
- avec $N_e = N_\uparrow + N_\downarrow$.

7.1 La fonction Ψ

Le calcul de la fonction Ψ s'effectue dans le cas usuel et dans le cas multi-Jastrow de la même manière, c'est à dire par l'évaluation directe des déterminants de spin up et down. Même si les éléments de matrice dans les deux cas ne sont pas identiques, le calcul reste similaire (les tailles des

matrices sont les mêmes pour les deux situations).

Le calcul de la fonction d'onde requiert donc le calcul de deux déterminants pour chacun des deux cas.

7.2 Le gradient de Ψ

7.2.1 Dans le cas du Jastrow commun

La formule pour la dérivée suivant la $i^{\text{ème}}$ coordonnée de l'électron j est :

$$\frac{\partial \Psi}{\partial x_j^i} = \begin{vmatrix} \Phi_1(\vec{r}_1) & \dots & \frac{\partial \Phi_1(\vec{r}_j)}{\partial x_j^i} & \dots & \Phi_1(\vec{r}_{N_\uparrow}) \\ \Phi_2(\vec{r}_1) & \dots & \frac{\partial \Phi_2(\vec{r}_j)}{\partial x_j^i} & \dots & \Phi_2(\vec{r}_{N_\uparrow}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \times \begin{vmatrix} \Phi_1(\vec{r}_{N_\uparrow+1}) & \Phi_1(\vec{r}_{N_\uparrow+2}) & \dots & \dots & \Phi_1(\vec{r}_{N_e}) \\ \Phi_2(\vec{r}_{N_\uparrow+1}) & \Phi_2(\vec{r}_{N_\uparrow+2}) & \dots & \dots & \Phi_2(\vec{r}_{N_e}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

La dérivée d'une coordonnée d'un électron s'effectue en dérivant la colonne correspondante. Chacune de ces dérivées nécessite le calcul d'un déterminant supplémentaire. Il y a N_e électrons avec trois coordonnées spatiales.

Le calcul du gradient de la fonction d'onde requiert donc le calcul de $3N_e$ déterminants supplémentaires.

7.2.2 Dans le cas du Jastrow orbitalaire

Dans le cas du Jastrow orbitalaire, nous devons calculer les déterminants supplémentaires correspondant à :

$$\begin{vmatrix} \Psi_{11}^\sigma & \dots & \partial_j^i(\Psi_{1k}^\sigma) & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{l1}^\sigma & \dots & \partial_j^i(\Psi_{lk}^\sigma) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

avec $k = 1, N_\sigma$ et $\sigma = \uparrow$ ou \downarrow . La dérivée par rapport à une coordonnée d'un électron s'effectue en dérivant successivement toutes les colonnes des deux matrices up et down (N_e en tout). Chaque dérivée nécessite donc le calcul de N_e déterminants supplémentaires. Il y a N_e électrons avec trois coordonnées spatiales.

Le calcul du gradient de la fonction d'onde requiert donc le calcul de

$$3N_e^2$$

déterminants supplémentaires.

7.3 Le Laplacien de Ψ

7.3.1 Dans le cas du Jastrow commun

La formule pour la dérivée double suivant la $i^{\text{ème}}$ coordonnée de l'électron j est (j est choisi ici arbitrairement comme électron \uparrow) :

$$\frac{\partial^2 \Psi}{\partial x_j^{i2}} = \begin{vmatrix} \Phi_1(\vec{r}_1) & \dots & \frac{\partial^2 \Phi_1(\vec{r}_j)}{\partial x_j^{i2}} & \dots & \Phi_1(\vec{r}_{N_\uparrow}) \\ \Phi_2(\vec{r}_1) & \dots & \frac{\partial^2 \Phi_2(\vec{r}_j)}{\partial x_j^{i2}} & \dots & \Phi_2(\vec{r}_{N_\uparrow}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix} \begin{vmatrix} \Phi_1(\vec{r}_{N_\uparrow+1}) & \dots & \dots & \Phi_1(\vec{r}_{N_e}) \\ \Phi_2(\vec{r}_{N_\uparrow+1}) & \dots & \dots & \Phi_2(\vec{r}_{N_e}) \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

Chacune de ces dérivées doubles nécessite le calcul d'un déterminant supplémentaire. Il y a N_e électrons avec chacun trois coordonnées spatiales.

En l'absence de simplifications, le calcul du Laplacien de la fonction d'onde requiert donc le calcul de $3N_e$ déterminants supplémentaires.

7.3.2 Dans le cas du Jastrow orbitalaire

Dans le cas du Jastrow orbitalaire, calculons la dérivée pour une coordonnée d'un électron. Nous avons vu que nous pouvions décomposer le calcul en trois parties :

- Le produit des matrices de spin up et down dont chacune possède une seule colonne constituée de dérivées premières (partie 1).

$$\left| \begin{array}{cccc} \Psi_{11}^\uparrow & \dots & \partial_i^m(\Psi_{1k}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \dots & \partial_i^m(\Psi_{jk}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right| \left| \begin{array}{cccc} \Psi_{1(N_\uparrow+1)}^\downarrow & \dots & \partial_i^m(\Psi_{1l}^\downarrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j(N_\uparrow+1)}^\downarrow & \dots & \partial_i^m(\Psi_{jl}^\downarrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{array} \right|$$

Il n'y a pas de déterminants supplémentaires à calculer puisque ceux-ci l'ont déjà été précédemment.

- À l'intérieur d'une même matrice deux colonnes différentes sont dérivées une seule fois (partie 2).

$$\left| \begin{array}{cccc} \dots & \partial_i^m(\Psi_{1k}^\uparrow) & \dots & \partial_i^m(\Psi_{1l}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dots & \partial_i^m(\Psi_{jk}^\uparrow) & \dots & \partial_i^m(\Psi_{jl}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right|$$

Le nombre de déterminants supplémentaires est $C_{N_\uparrow}^2 + C_{N_\downarrow}^2$ (C_N^p est le coefficient du binôme) qui décrit le nombre de possibilités de dériver deux colonnes parmi d'une part les N_\uparrow colonnes de la matrice de spin up et parmi les N_\downarrow colonnes de la matrice de spin down soit $\frac{N_\uparrow(N_\uparrow-1)}{2} + \frac{N_\downarrow(N_\downarrow-1)}{2}$. Cela n'est valable que pour une coordonnée d'un électron. Il faut faire de même pour la totalité de $3N_e$ coordonnées. Le nombre de déterminants supplémentaires est donc :

$$3N_e \left[\frac{N_\uparrow(N_\uparrow-1)}{2} + \frac{N_\downarrow(N_\downarrow-1)}{2} \right]$$

- Chaque colonne est dérivée deux fois (partie 3).

$$\begin{vmatrix} \Psi_{11}^\uparrow & \dots & \partial_i^{2m}(\Psi_{1k}^\uparrow) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \dots & \partial_i^{2m}(\Psi_{jk}^\uparrow) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

Il y a donc N_e déterminants supplémentaires.

Le calcul du Laplacien de la fonction d'onde requiert donc le calcul de

$$3N_e \left[N_e + \frac{N_\uparrow(N_\uparrow - 1)}{2} + \frac{N_\downarrow(N_\downarrow - 1)}{2} \right]$$

déterminants supplémentaires soit de l'ordre de $\approx N_e^3$.

Dérivées par rapport aux paramètres des Jastrows orbitales :

Quelques notations

Nous allons à présent déterminer le nombre de déterminants nécessaires pour calculer les dérivées suivant les paramètres des Jastrows orbitales. Pour le cas du Jastrow commun, remarquons qu'aucune évaluation de déterminant n'est nécessaire puisque le Jastrow est en préfacteur de la partie déterminantale. Nous noterons :

- N_{opt} le nombre de paramètres distincts
- $N(c_p)$ le nombre total de lignes (dans les matrices de spin up et down) comportant le paramètre c_p
- $N(c_p, \uparrow)$ (respectivement, $N(c_p, \downarrow)$) le nombre de lignes dans les matrices de spin up (resp. down) comportant le paramètre c_p .
- $N_{line} = \sum_{p=1}^{N_{opt}} N(c_p)$
- $N_{line}(\uparrow) = \sum_{p=1}^{N_{opt}} N(c_p, \uparrow)$
- $N_{line}(\downarrow) = \sum_{p=1}^{N_{opt}} N(c_p, \downarrow)$

Noter que $N(c)$ peut être supérieur à 1 car un même Jastrow orbitale peut être attaché à plusieurs orbitales moléculaires à la fois. N_{line} est la somme

de tous les nombres $N(c_p)$. Il représente le nombre total de lignes à dériver si l'on souhaite dériver partiellement Ψ ou le Laplacien de Ψ suivant tous les paramètres. N_{line} est au minimum égal à N_{opt} mais il peut être bien plus important. Par exemple, pour l'étude de la molécule HF présentée plus loin, nous avons attaché un Jastrow orbitalaire dit de valence comportant 17 paramètres sur les quatre orbitales moléculaires de valence (pour les électrons de spins up et les électrons de spin down). Nous avons donc :

- $N_{opt}=17$
- pour n'importe quel paramètre $N(c_p) = 8$ (4 orbitales de la matrice de spin up et 4 orbitales de la matrice de spin down)
- $N(c_p, \uparrow) = 4$
- $N(c_p, \downarrow) = 4$
- $N_{line} = 17 \times 8 = 136$

7.4 Dérivée simple de Ψ par rapport à un paramètre

Pour chaque paramètre c_p il faut dériver toute les lignes qui le contiennent. Ce sont des déterminants du type :

$$\begin{vmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \cdots & \frac{\partial \Psi_{jN_\uparrow}^\uparrow}{\partial c_p} \\ \vdots & \ddots & \vdots \end{vmatrix}$$

Il y a donc N_{line} déterminants supplémentaires à calculer.

7.5 Dérivée double de Ψ

Nous dérivons Ψ suivant les paramètres c_{p_1} et c_{p_2} .

Les dérivées doubles ont exactement la même structure que ceux qui a été fait pour le Laplacien (dérivées doubles par rapport aux coordonnées). Mais

cette fois-ci nous dérivons les lignes à la place des colonnes. Il y a donc également trois types de déterminants :

- Les déterminants où une seule ligne est dérivée qui servent dans le produit. Ces déterminants ont déjà été calculés pour la dérivée partielle de la fonction d'onde Ψ .
- Les déterminants où deux lignes distinctes d'une même matrice sont dérivées

$$\begin{vmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_{p_1}} & \cdots & \frac{\partial \Psi_{jN_\uparrow}^\uparrow}{\partial c_{p_1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \Psi_{1k1}^\uparrow}{\partial c_{p_2}} & \cdots & \frac{\partial \Psi_{kN_\uparrow}^\uparrow}{\partial c_{p_2}} \\ \vdots & \ddots & \vdots \end{vmatrix}$$

- Les déterminants où une ligne d'une matrice est dérivée deux fois de suite.

$$\begin{vmatrix} \Psi_{11}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \Psi_{j1}^\uparrow}{\partial c_{p_1} \partial c_{p_2}} & \cdots & \frac{\partial^2 \Psi_{jN_\uparrow}^\uparrow}{\partial c_{p_1} \partial c_{p_2}} \\ \vdots & \ddots & \vdots \end{vmatrix}$$

Attention, ces derniers déterminants n'existent que lorsque les deux paramètres que nous considérons se trouvent sur une même ligne (c'est à dire, appartiennent au même Jastrow orbitalaire). Nous pouvons réunir ces deux derniers types de déterminants et se passer de la distinction entre paramètres appartenant au même Jastrow ou non. Pour cela nous pouvons les décrire comme des déterminants dont les lignes pas nécessairement distinctes sont dérivées deux fois. Grâce à cela, nous pouvons en déduire le nombre de déterminants supplémentaires nécessaires.

$$N(c_{p_1}, \uparrow)N(c_{p_2}, \uparrow) + N(c_{p_1}, \downarrow)N(c_{p_2}, \downarrow)$$

7.6 Dérivée partielle du Laplacien suivant un paramètre

Calculons la dérivée pour une coordonnée d'un électron et un paramètre c_p . Nous avons vu que nous pouvions décomposer le calcul en trois parties, chaque déterminant sera alors dérivé suivant les lignes :

- Les déterminants des matrices de spin up et down dont chacune à une seule colonne dérivée

$$\begin{vmatrix} \Psi_{11}^\uparrow & \dots & \partial_i^m(\Psi_{1k}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \dots & \frac{\partial \partial_i^m(\Psi_{jk}^\uparrow)}{\partial c_p} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_{j1}^\uparrow & \dots & \partial_i^m(\Psi_{jk}^\uparrow) & \dots \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

Il y a une colonne à dériver, soit $N_\uparrow + N_\downarrow$. Puis, il y a pour la matrice up (resp. down) $N(c_p, \uparrow)$ (resp. $N(c_p, \downarrow)$) lignes à dériver. Cela fait un total de $N(c_p, \uparrow)N_\uparrow + N(c_p, \downarrow)N_\downarrow$ déterminants supplémentaires à calculer.

- À l'intérieur d'une même matrice deux colonnes sont dérivées une fois (partie 2) :

$$\begin{vmatrix} \Psi_{11}^\uparrow & \dots & \partial_i^{2m}(\Psi_{1k}^\uparrow) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j1}^\uparrow}{\partial c_p} & \dots & \frac{\partial \partial_i^{2m}(\Psi_{jk}^\uparrow)}{\partial c_p} & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

Le nombre de déterminants supplémentaires est $C_{N_\uparrow}^2 + C_{N_\downarrow}^2$ qui décrit le nombre de possibilités de dériver 2 colonnes parmi d'une part les N_\uparrow colonnes de la matrice de spin up et parmi les N_\downarrow colonnes de la matrice de spin down soit $\frac{N_\uparrow(N_\uparrow-1)}{2} + \frac{N_\downarrow(N_\downarrow-1)}{2}$. Puis, il y a pour la matrice up

(resp. down) $N(c_p, \uparrow)$ (resp. $N(c_p, \downarrow)$) lignes à dériver. Le nombre de déterminants supplémentaires est donc :

$$N(c_p, \uparrow) \frac{N_{\uparrow}(N_{\uparrow} - 1)}{2} + N(c_p, \downarrow) \frac{N_{\downarrow}(N_{\downarrow} - 1)}{2}$$

- Chaque colonne est dérivée deux fois (partie 3).

$$\begin{vmatrix} \Psi_{11}^{\uparrow} & \dots & \partial_i^{2m}(\Psi_{1k}^{\uparrow}) & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_{j1}^{\uparrow}}{\partial c_p} & \dots & \frac{\partial \partial_i^{2m}(\Psi_{jk}^{\uparrow})}{\partial c_p} & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

Il y a une colonne à dériver, soit $N_{\uparrow} + N_{\downarrow}$. Puis, il y a pour la matrice up (resp. down) $N(c_p, \uparrow)$ (resp. $N(c_p, \downarrow)$) lignes à dériver. Cela fait un total de $N(c_p, \uparrow)N_{\uparrow} + N(c_p, \downarrow)N_{\downarrow}$ déterminants supplémentaires à calculer.

Lorsque nous dérivons la fonction d'onde deux fois suivant une coordonnée spatiale puis une fois suivant un paramètre, nous devons calculer le nombre de déterminants suivant :

$$N(c_p, \uparrow) \frac{N_{\uparrow}(N_{\uparrow} - 1)}{2} + N(c_p, \downarrow) \frac{N_{\downarrow}(N_{\downarrow} - 1)}{2} + N(c_p, \uparrow)N_{\uparrow} + N(c_p, \downarrow)N_{\downarrow}$$

ce qui donne pour le Laplacien complet ($3N_e$ composantes) :

$$3N_e(N(c_p, \uparrow) \frac{N_{\uparrow}(N_{\uparrow} - 1)}{2} + N(c_p, \downarrow) \frac{N_{\downarrow}(N_{\downarrow} - 1)}{2} + N(c_p, \uparrow)N_{\uparrow} + N(c_p, \downarrow)N_{\downarrow})$$

et pour l'ensemble des paramètres :

$$\sum_{p=1}^{N_{opt}} [3N_e(N(c_p, \uparrow) \frac{N_{\uparrow}(N_{\uparrow} - 1)}{2} + N(c_p, \downarrow) \frac{N_{\downarrow}(N_{\downarrow} - 1)}{2} + N(c_p, \uparrow)N_{\uparrow} + N(c_p, \downarrow)N_{\downarrow})]$$

qui peut se réécrire en sortant de la somme les termes ne dépendant pas de p et en utilisant les notations :

- $N_{line}(\uparrow) = \sum_{p=1}^{N_{opt}} N(c_p, \uparrow)$
- $N_{line}(\downarrow) = \sum_{p=1}^{N_{opt}} N(c_p, \downarrow)$

$$3N_e \left[\frac{N_\uparrow(N_\uparrow - 1)}{2} + N_\uparrow \right] N_{line}(\uparrow) + 3N_e \left[\frac{N_\downarrow(N_\downarrow - 1)}{2} + N_\downarrow \right] N_{line}(\downarrow)$$

Voici donc résumé dans le tableau suivant 7.1 tous les résultats que nous venons d'énoncer.

TAB. 7.1 – Nombre de déterminants à calculer pour un pas élémentaire de l'algorithme Monte Carlo et pour un seul marcheur.

	Jastrow commun	Jastrow orbitalaire
	Nombre de déts nécessaires pour un calcul VMC ou DMC	
Ψ_T	2	2
$\nabla\Psi_T$	$3N_e$	$3N_e^2$
$\Delta\Psi_T$	$3N_e$	$3N_e \left[N_e + \frac{N_\uparrow(N_\uparrow - 1)}{2} + \frac{N_\downarrow(N_\downarrow - 1)}{2} \right]$
	Nombre de déts supplémentaires pour optimisations linéaires ou de Newton	
$\Psi_{T,l}$	0	N_{line}
$\Delta\Psi_{T,l}$	0	$3N_e \left[\frac{N_\uparrow(N_\uparrow - 1)}{2} + N_\uparrow \right] N_{line}(\uparrow) + 3N_e \left[\frac{N_\downarrow(N_\downarrow - 1)}{2} + N_\downarrow \right] N_{line}(\downarrow)$
	Nombre de déts supplémentaires pour optimisation de Newton	
$\Psi_{T,l,m}$	0	$N(c_1, \uparrow)N(c_2, \uparrow) + N(c_1, \downarrow)N(c_2, \downarrow)$

Le nombre de déterminants qui est ici consigné est valable pour un seul marcheur et pour un unique pas Monte-Carlo. La colonne “Jastrow Commun” indique le nombre de déterminants nécessaires pour le cas de du Jastrow commun. La colonne “Jastrow orbitalaire” indique le nombre de déterminants pour le cas de Jastrows orbitales. Trois sous-parties figurent dans ce tableau :

- Nombre de déterminants nécessaires dans une simulation VMC ou DMC. Il s’agit du nombre de déterminants pour calculer les quantités nécessaires aux simulations.
- Nombre de déterminants supplémentaires pour les optimisations linéaires ou de Newton. Il s’agit du nombre de déterminants *supplémentaires* à calculer, propre aux nouvelles quantités introduites par l’optimisation. Cette sous-partie comporte l’ensemble des déterminants supplémentaires nécessaires pour l’optimisation linéaire. Ces déterminants sont aussi nécessaires pour l’optimisation suivant la méthode de Newton mais pour obtenir l’ensemble des déterminants supplémentaires nécessaires à l’optimisation de Newton, il faut rajouter la sous-partie suivante.
- Nombre de déterminants supplémentaires pour l’optimisation suivant la méthode de Newton.

Ψ_T donne le nombre de déterminants nécessaire pour le calcul de la fonction d’onde ; $\nabla\Psi_T$ de son gradient et $\Delta\Psi_T$ de son Laplacien. $\Psi_{T,l}$ donne le nombre de déterminants pour les dérivées simple de la fonction d’onde suivant tous les paramètres ; $\Delta\Psi_T$ le nombre de déterminants pour les dérivées simples du Laplacien de la fonction d’onde suivant tous les paramètres. Enfin, $\Psi_{T,l,m}$ donne le nombre de déterminants pour le calcul des dérivées doubles de la fonction d’onde suivant tous les paramètres.

Chapitre 8

Comparaison des temps de calcul pour une fonction d'onde usuelle et une fonction d'onde multi-Jastrow

Nous allons maintenant présenter les résultats d'une étude numérique comparative illustrant l'efficacité de notre implémentation des fonctions d'onde multi-Jastrow. Plus précisément, nous présentons les temps de calcul obtenus pour une simulation VMC et pour une itération de l'optimisation linéaire sur les atomes d'oxygène et de cuivre. L'étude de ces deux atomes ne se fait ici que sous l'angle des temps de calcul. L'intérêt "physique" des Jastrows orbitales a déjà été discuté dans la partie consacrée aux résultats[6]. La partie déterminantale de la fonction d'onde des atomes d'oxygène et de cuivre utilisée ici est celle obtenue par un calcul Hartree-Fock en base de Slater (nous avons utilisé les fonctions d'onde HF de Bunge *et al.*[146]). Pour l'atome d'oxygène, nous avons considéré l'état fondamental triplet ce qui donne une fonction d'onde écrite comme un produit d'un déterminant de taille 5 (électrons de spin up) et d'un déterminant de taille 3 (électrons de spin down). La fonction d'onde du cuivre est le produit d'un déterminant de taille 15

(spins up) et d'un déterminant de taille 14 (spins down). Dans les deux cas, les Jastrows communs complexes auront 19 paramètres. Les Jastrows orbitaux auront chacun 12 paramètres (il y a moins de paramètres dans ce cas-là parce qu'avec les Jastrows orbitaux il n'y a plus de distinction entre les paires de même spin ou de spins opposés).

Les simulations ont été effectuées pour 100 marcheurs et 40 blocs de 100 pas chacun, soit 400000 pas Monte-Carlo. Le processeur sur lequel tous les calculs ont été lancés est un processeur Intel Xeon 5355 de 2,66Ghz sur les machines de calcul de Laboratoire de Chimie et Physique Quantique de Toulouse. Les points des simulations VMC sont déplacés suivant un mouvement Brownien avec dérive, voir section 2.2.3. Les résultats présentés sous le nom "Optimisation" correspondent à la réalisation d'une itération d'un calcul d'optimisation par méthode linéaire (voir section 3.4). Tous les temps exprimés ci-dessous seront en secondes.

8.1 Atome d'oxygène

8.1.1 Simulations VMC

Nous rappelons que lors d'une simulation VMC il est nécessaire d'évaluer à chaque pas Monte Carlo et pour chaque marcheur la valeur de la fonction d'onde d'essai Ψ_T , ainsi que les valeurs de son gradient, $\nabla\Psi_T$, et de son Laplacien, $\Delta\Psi_T$.

Pour l'atome d'oxygène le nombre de déterminants à calculer est le suivant (application des formules présentées dans le tableau 7.1) :

Simulations VMC avec Jastrow commun Le tableau suivant, 8.1.1, servira de référence. Il s'agit d'une simulation VMC sur l'atome d'oxygène pour le cas du Jastrow commun. La ligne "Jastrow" donne le temps de calcul nécessaire pour évaluer les Jastrows communs lors de la simulation. La ligne "Elements de matrice" donne le temps de calcul nécessaire pour calculer les orbitales atomiques et leurs dérivées. La ligne "Diagonalisation" donne le

	Jastrow commun	Jastrow orbitalaire
Ψ_T	2	2
$\nabla\Psi_T$	24	192
$\Delta\Psi_T$	24	504
Total	50	698

TAB. 8.1 – Nombre de déterminants à calculer pour un pas Monte Carlo et un marcheur lors d’une simulation VMC de l’atome d’oxygène

Calcul	Temps(s)
Jastrow	29.50
Éléments de matrice	18.63
Diagonalisation	5.52
Temps total	77.29

TAB. 8.2 – Temps de calcul d’une simulation VMC pour l’atome d’oxygène avec un seul Jastrow en commun

temps de calcul pour toutes les inversions de matrice. Nous allons à présent voir le cas où des Jastrows orbitaux sont utilisés.

Simulations VMC avec un ou plusieurs Jastrows orbitaux La première simulation a été faite en ne plaçant qu’un seul Jastrow sur une seule orbitale atomique. Plus précisément, comme les matrices de spins up et down ont les mêmes orbitales, le Jastrow est sur une ligne de la matrice des spins up et sur une ligne de la matrice des spins down.

Avant de regarder les différentes contributions aux temps de calcul, analysons tout de suite les temps totaux. Pour une fonction d’onde avec Jastrow commun, le temps total était de 77.29 secondes. Pour le Jastrow orbitalaire, ce temps est passé à 175.50 secondes soit à peine le double de temps. Au vu du grand nombre d’opérations à effectuer et du nombre bien plus grand de déterminants à traiter nous pouvons considérer que l’algorithme que nous avons proposé est très efficace. Au final, l’usage d’un Jastrow orbitalaire augmente relativement peu le temps de calcul. Le compromis efficacité/temps de calcul

Calcul	Temps(s)
Jastrow	30.21
Eléments de matrice	19.17
Calcul des déterminants et inversion par pivot	3.92
Assemblage des matrices	42.39
Déterminants	52.17
Temps total	175.50

TAB. 8.3 – Temps de calcul d’une simulation VMC de l’oxygène pour un Jastrow orbitalaire sur une seule orbitale

que nous mentionnions en présentant notre nouvelle fonction d’onde penche très clairement en sa faveur. Comme nous le soulignons précédemment, il n’y pas d’obstacles calculatoires à l’usage de ces Jastrows orbitalaires.

Détaillons à présent les temps des divers éléments du calcul. La ligne “Calcul des déterminants et inversion par pivot” désigne le temps mis pour inverser et calculer le déterminant par le pivot de Gauss sur les deux matrices up et down. Nous rappelons que nous ne calculons les déterminants et les inverses par pivot de Gauss que sur ces deux matrices. La ligne “Déterminants” désigne le temps mis pour calculer tous les autres déterminants nécessaires à la construction du gradient et du Laplacien de la fonction d’onde. La ligne “Assemblage des matrices” désigne le temps mis pour construire toutes les matrices nécessaires à partir des Jastrows, de leurs dérivées, des orbitales χ et de leurs dérivées. La première remarque importante porte sur le temps de calcul des Jastrows. Nous pouvons constater, comme nous l’avions affirmé dans le chapitre sur la fonction d’onde d’essai, que les temps de calcul du Jastrow orbitalaire et du Jastrow commun sont similaires.

Nous remarquons également que la ligne “Eléments de matrices” donne le même temps que dans le cas du Jastrow commun. Cela est parfaitement normal car nous calculons ces éléments de la même manière. Le temps de calcul des déterminants ne devient plus négligeable. Cependant, nous rappelons que nous avons 696 déterminants à calculer de cette manière face à 48 dans

le cas du Jastrow commun (en omettant les 2 déterminants de la fonction d'onde identique dans les deux cas). Le temps de calcul des deux matrices par pivot de Gauss est de 3.92 secondes. Si nous appliquions un calcul d'un pivot de Gauss pour calculer les 292 autres déterminants, ce calcul prendrait à peu près $\frac{696}{2} \times \frac{3.92}{2} = 1364.16$ secondes. Le calcul des déterminants suivant notre méthode ne prend que 52.17 secondes. Notre algorithme est donc largement plus efficace. Enfin, nous remarquons également un résultat inattendu. Le temps mis pour assembler les matrices n'est pas négligeable. L'assemblage des matrices consiste en des opérations élémentaires sur des éléments de tableaux. Nous suggérons que cette partie du code pourrait être améliorée ultérieurement par une meilleure manipulation des tableaux.

Nous avons ensuite placé le même Jastrow sur les 5 orbitales (soit, sur les 8 lignes des 2 matrices : 5 lignes sur la matrice de spins up, 3 lignes sur la matrice de spins down) Nous constatons que les résultats sont sensiblement

Calcul	Temps(s)
Jastrow	32.53
Eléments de matrice	18.97
Calcul des déterminants et inversion par le pivot	4.15
Assemblage des matrices	42.71
Déterminants	50.41
Temps total	177.56

TAB. 8.4 – Temps de calcul d'une simulation VMC de l'oxygène pour un Jastrow sur 5 orbitales

les mêmes que précédemment. Il y a en effet autant d'éléments de matrices et de déterminants à calculer.

Cette fois-ci nous plaçons un Jastrow différent sur chacune des 5 orbitales. Il y a donc 5 Jastrows à calculer. En comparant les temps de calcul des Jastrows avec le cas où il n'y a qu'un Jastrow orbitalaire, nous pouvons voir la dépendance dans le nombre de Jastrows. Les autres lignes sont peu modifiées. Le temps de calcul augmente uniquement à cause du temps de

Calcul	Temps(s)
Jastrow	105.20
Eléments de matrice	19.82
Calcul des déterminants et inversion par le pivot	4.35
Assemblage des matrices	42.94
Déterminants	53.80
Temps total	254.87

TAB. 8.5 – Temps de calcul d’une simulation VMC de l’oxygène pour 5 Jastrow orbitales sur 5 orbitales

calcul des Jastrows.

8.1.2 Optimisation

Nous présentons à présent une comparaison des temps calcul quand on réalise une optimisation avec la méthode linéaire pour une fonction d’onde ayant un seul Jastrow en commun ou plusieurs Jastrows orbitales. Rappelons que, par rapport à une simulation VMC qui nécessite le calcul de Ψ_T , de son gradient et de son Laplacien par rapport aux coordonnées spatiales, lors de l’étape d’optimisation nous devons aussi calculer les dérivées de la fonction d’essai Ψ_T par rapport aux différents paramètres, $\Psi_{T,l}$ (l indice des paramètres) ainsi que les Laplaciens par rapport aux positions des électrons, $\Delta\Psi_{T,l}$.

Optimisation avec un Jastrow commun Nous avons effectué deux calculs. Un premier calcul (voir le tableau 8.1.2) où un seul paramètre est optimisé. Un second calcul (tableau 8.1.2) où l’ensemble des 19 paramètres est optimisé (nombre maximal de paramètres que comportent nos Jastrows communs).

Les résultats pour le Jastrow commun sont les suivants :

Nous constatons bien évidemment une augmentation du temps de calcul pour la partie “Jastrow”.

Calcul	Temps(s)
Jastrow	59.14
Elements de matrice	19.14
Diagonalisation	4.95
Temps total	95.64

TAB. 8.6 – Temps de calcul d’une itération de l’algorithme d’optimisation de la fonction d’onde de l’atome d’oxygène comprenant un Jastrow commun. Un seul paramètre est optimisé

Calcul	Temps(s)
Jastrow	91.96
Elements de matrice	19.36
Diagonalisation	5.09
Temps total	167.78

TAB. 8.7 – Temps de calcul d’une itération de l’algorithme d’optimisation de la fonction d’onde de l’atome d’oxygène comprenant un Jastrow commun. 19 paramètres sont optimisés.

Nous passons maintenant aux Jastrows orbitales.

Optimisation d’un paramètre d’un Jastrow attaché à une seule orbitale Le nombre de déterminants supplémentaires à calculer dans ce cas est : Les résultats pour les temps calcul sont les suivants :

$\Psi_{T,l}$	2
$\Delta\Psi_{T,l}$	504

TAB. 8.8 – Nombre de déterminants à calculer dans le cas de l’atome d’oxygène pour l’optimisation d’un seul paramètre d’un Jastrow attaché à une seule orbitale

Nous pouvons constater que le temps de calcul des Jastrows s’est allongé par rapport à la simulation VMC. En plus du Jastrow et de ses dérivées spa-

Calcul	Temps(s)
Jastrow	43.46
Eléments de matrice	19.26
Calcul des déterminants et inversion par le pivot	4.09
Assemblage des matrices	119.91
Déterminants	51.57
Déterminants de l'optimisation	46.86
Temps total	334.98

TAB. 8.9 – Temps de calcul d'une itération de l'optimisation dans le cas de l'atome d'oxygène. Un seul Jastrow sur une seule orbitale. Un seul paramètre

tiales, il faut calculer également la dérivée du Jastrow suivant le paramètre. Nous avons ajouté une nouvelle ligne dans ce tableau : “Déterminants de l'optimisation”. Il s'agit du temps de calcul des déterminants calculés spécifiquement pour l'optimisation. Ce temps est de l'ordre de grandeur du temps de calcul pour évaluer les déterminants nécessaires à la simulation VMC. Cela est parfaitement normal car, dans le déroulement de l'algorithme pour obtenir ces nouveaux déterminants propres à l'optimisation, nous appliquons les mêmes procédures que pour le cas de la simulation VMC à la fonction Ψ_l , dérivée de Ψ par rapport au paramètre l .

Optimisation des 12 paramètres d'un Jastrow attaché à une seule orbitale Nous étudions à présent le cas où on optimise complètement le Jastrow attaché à une seule orbitale. Ici, nous avons 12 paramètres. Le nombre de déterminants supplémentaires à calculer est donc :

$\Psi_{T,l}$	24
$\Delta\Psi_{T,l}$	6048

TAB. 8.10 – Nombre de déterminants à calculer pour l'optimisation dans le cas de l'atome d'oxygène. Un seul Jastrow sur une seule orbitale. 12 paramètres

Les temps calcul obtenus sont les suivants :

Calcul	Temps(s)
Jastrow	126.010
Eléments de matrice	19.66
Calcul des déterminants et inversion par le pivot	4.68
Assemblage des matrices	635.63
Déterminants	53.75
Déterminants de l'optimisation	557.31
Temps total	1467.60

TAB. 8.11 – Temps de calcul d'une itération de l'optimisation dans le cas de l'atome d'oxygène. Un seul Jastrow sur une seule orbitale. 12 paramètres.

Le temps de calcul des Jastrows augmente en conséquence. Le temps de calcul des déterminants de l'optimisation augmente lui linéairement avec le nombre de paramètres par rapport auxquels nous dérivons. Plus exactement, il augmente avec le nombre de paramètres par ligne que nous dérivons. Précisons cette notion.

Optimisation de deux paramètres pour deux Jastrows orbitales sur deux orbitales

Nous allons à présent optimiser deux paramètres de différentes manières. Nous prenons dans un premier temps deux Jastrows placés chacun sur une orbitale différente et nous optimisons un paramètre de chaque Jastrow. Le nombre de déterminants supplémentaires est donc :

$\Psi_{T,l}$	2
$\Delta\Psi_{T,l}$	504

TAB. 8.12 – Nombre de déterminants à calculer pour l'optimisation de l'oxygène pour deux Jastrow orbitales - deux paramètres sur un Jastrow

Nous optimisons par la suite un paramètre sur chacun des deux Jastrows. Il y a toujours deux paramètres à optimiser. Le nombre de déterminants

Calcul	Temps(s)
Jastrow	79.26
Eléments de matrice	19.12
Calcul des déterminants et inversion par le pivot	4.50
Assemblage des matrices	167.91
Déterminants	51.25
Déterminants de l'optimisation	93.02
Temps total	466.24

TAB. 8.13 – Temps de calcul d'une itération de l'optimisation de l'oxygène pour deux Jastrows orbitales - deux paramètres sur un Jastrow

supplémentaires à calculer est :

$\Psi_{T,l}$	2
$\Delta\Psi_{T,l}$	504

TAB. 8.14 – Nombre de déterminants à calculer pour l'optimisation de l'oxygène pour deux Jastrow orbitales - 1 paramètre par Jastrow (2 paramètres)

Calcul	Temps(s)
Jastrow	79.13
Eléments de matrice	19.29
Calcul des déterminants et inversion par le pivot	4.50
Assemblage des matrices	169.21
Déterminants	50.83
Déterminants de l'optimisation	94.19
Temps total	466.94

TAB. 8.15 – Temps de calcul d'une itération de l'optimisation de l'oxygène pour deux Jastrows orbitales - 1 paramètre par Jastrow (2 paramètres)

Nous constatons que les temps de calculs sont similaires.

8.1.3 Optimisation pour un paramètre pour un Jastrow sur deux orbitales

Nous présentons enfin un dernier cas très important. Nous mettons un Jastrow sur deux orbitales et nous optimisons un paramètre de ce Jastrow. Le nombre de déterminants supplémentaires à calculer est :

$\Psi_{T,l}$	2
$\Delta\Psi_{T,l}$	504

TAB. 8.16 – Nombre de déterminants à calculer pour l’optimisation de l’oxygène pour un Jastrow sur 2 orbitales - 1 paramètre

Calcul	Temps(s)
Jastrow	43.94
Eléments de matrice	19.05
Calcul des déterminants et inversion par le pivot	4.41
Assemblage des matrices	163.95
Déterminants	51.95
Déterminants de l’optimisation	92.38
Temps total	423.58

TAB. 8.17 – Temps de calcul d’une itération de l’optimisation de l’oxygène pour un Jastrow sur 2 orbitales - 1 paramètre

Le temps de calcul des Jastrows correspond bien à un Jastrow. Mais le temps de calcul des déterminants de l’optimisation est similaire au temps de calcul avec deux paramètres. En effet, ce qui compte c’est le produit du nombre de paramètres à optimiser par le nombre de lignes où le Jastrow contenant ce paramètre est présent.

8.2 Atome de cuivre

Passons maintenant à l’étude de l’atome de cuivre. Le nombre de déterminants à calculer devient :

Il y a donc 400 fois plus de déterminants à calculer. Voyons dans la suite

	Jastrow commun	Jastrow orbitalaire
Ψ_T	2	2
$\nabla\Psi_T$	87	2523
$\Delta\Psi_T$	87	19575
Total	60	22100

TAB. 8.18 – Nombre de déterminants à calculer pour une simulations VMC du cuivre

l'impact de ce nombre.

Toutes les remarques que nous avons faites auparavant pour l'oxygène se retrouveront sur les exemples ci-dessous.

Notre discussion portera donc plus précisément sur la dépendance de l'effort numérique en fonction du nombre d'électrons en comparant le cuivre et l'oxygène.

8.2.1 Simulation VMC du cuivre pour un Jastrow commun

Voici les résultats pour le Jastrow commun.

Calcul	Temps(s)
Jastrow	429.54
Eléments de matrice	335.09
Diagonalisation	97.25
Temps total	933.26

TAB. 8.19 – Temps de calcul d'une simulation VMC du cuivre pour un Jastrow commun

8.2.2 Simulation VMC du cuivre avec des Jastrows orbitales

Nous présentons plusieurs configurations.

- un Jastrow sur l'orbitale 1s
 - deux Jastrows : un Jastrow sur l'orbitale 1s et un Jastrow sur les quatre orbitales (2s, 2px, 2py, 2pz)
 - trois Jastrows : un Jastrow sur l'orbitale 1s, un Jastrow sur les orbitales 2s, 2px, 2py, 2pz, un Jastrow sur les dix autres orbitales (3s,3p's,3d's,4s)
- C'est le système le plus important en terme de temps de calcul que nous avons étudié durant ce travail.

Calcul	Temps(s)
Jastrow	262.23
Eléments de matrice	337.45
Calcul des déterminants et inversion par le pivot	77.11
Assemblage des matrices	512.89
Déterminants	1605.53
Temps total	2986.91

TAB. 8.20 – Temps de calcul d'une simulation VMC du cuivre pour un Jastrow sur une seule orbitale

Calcul	Temps(s)
Jastrow	472.69
Eléments de matrice	337.52
Calcul des déterminants et inversion par le pivot	75.09
Assemblage des matrices	519.15
Déterminants	1602.07
Temps total	3197.66

TAB. 8.21 – Temps de calcul d'une simulation VMC du cuivre pour deux Jastrows orbitales sur cinq orbitales

En examinant les trois tableaux 8.2.2,8.2.2 et 8.2.2, nous pouvons à nouveau noter la linéarité dans le calcul des Jastrows et la constance du temps de calcul des déterminants. Le temps de calcul pour tous les déterminants est de 1603.91 secondes (pour un total de 22098 déterminants) Rappelons que

Calcul	Temps(s)
Jastrow	677.00
Eléments de matrice	337.59
Calcul des déterminants et inversion par le pivot	75.89
Assemblage des matrices	519.15
Déterminants	1603.91
Temps total	3404.27

TAB. 8.22 – Temps de calcul d’une simulation VMC du cuivre pour trois Jastrows orbitales sur quinze orbitales

pour l’atome d’oxygène, les temps de calcul étaient de 53 secondes (pour 696 déterminants). Il y a donc linéarité du temps-calcul en fonction du nombre de déterminants : $\frac{1603.91}{22098} = 0.073$ et $\frac{52.17}{696} = 0.075$. Le calcul des deux matrices par pivot de Gauss prend lui 75.89 secondes soit 37.94 secondes par déterminant. Notre approche est donc largement supérieure à celle d’un simple calcul de tous les déterminants par pivot de Gauss.

8.2.3 Optimisation pour un Jastrow commun

Nous présentons à présent les résultats pour l’optimisation. Les résultats pour le Jastrow commun sont les suivants :

Calcul	Temps(s)
Jastrow	1011.44
Eléments de matrice	336.54
Diagonalisation	100.48
Temps total	1517.84

TAB. 8.23 – Temps de calcul d’une itération de l’optimisation du cuivre pour un Jastrow commun - 1 paramètre

Le temps-calcul est essentiellement dû au calcul du Jastrow.

Calcul	Temps(s)
Jastrow	1323.18
Eléments de matrice	338.12
Diagonalisation	105.58
Temps total	1923.47

TAB. 8.24 – Temps de calcul d’une itération de l’optimisation du cuivre pour un Jastrow commun - 19 paramètre

8.2.4 Optimisation pour un Jastrow orbitalaire

Nous présentons deux optimisations :

- un paramètre à optimiser sur un unique Jastrow sur une orbitale
- douze paramètres à optimiser sur un unique Jastrow sur une orbitale

Le nombre de déterminants supplémentaires à calculer est donc : Les temps

$\Psi_{T,l}$	2
$\Delta\Psi_{T,l}$	19575

TAB. 8.25 – Nombre de déterminants à calculer pour l’optimisation du cuivre pour un Jastrow orbitalaire - 1 paramètre

$\Psi_{T,l}$	24
$\Delta\Psi_{T,l}$	234900

TAB. 8.26 – Nombre de déterminants à calculer pour l’optimisation du cuivre pour un Jastrow orbitalaire - 12 paramètres

de calcul sont :

Résumons les résultats que nous avons obtenus. Pour un calcul VMC ou DMC, nous avons constaté que pour l’oxygène possédant 8 électrons, l’usage des Jastrows orbitalaires augmentait le temps de calcul d’un facteur 2. Pour l’atome de cuivre à 29 électrons (soit à peu près 4 fois plus d’électrons), ce temps de calcul était augmenté d’un facteur 3-4. A un nombre d’élec-

Calcul	Temps(s)
Jastrow	411.53
Eléments de matrice	339.83
Calcul des déterminants et inversion par le pivot	75.25
Assemblage des matrices	2029.89
Déterminants	1614.68
Déterminants de l'optimisation	1405.41
Temps total	5266.66

TAB. 8.27 – Temps de calcul d'une itération de l'optimisation du cuivre pour un Jastrow orbitalaire - 1 paramètre

Calcul	Temps(s)
Jastrow	3552.86
Eléments de matrice	340.45
Calcul des déterminants et inversion par le pivot	92.58
Assemblage des matrices	18036.39
Déterminants	1624.17
Déterminants de l'optimisation	16980.90
Temps total	31587.18

TAB. 8.28 – Temps de calcul d'une itération de l'optimisation du cuivre pour un Jastrow orbitalaire - 12 paramètres

trons constant, le temps de calcul pour le cas du Jastrow orbitalaire dépend donc du nombre de Jastrow distincts calculé. Soulignons qu'il ne dépend pas du nombre d'orbitales possédant des Jastrows orbitaires. Dans la méthode d'optimisation linéaire, le temps de calcul augmente en fonction du nombre de paramètres optimisés *multiplié* par le nombre d'orbitales possédant ces paramètres. Si nous notons T_{ref} le temps de référence pour calculer un seul paramètre sur une seule orbitale, chaque nouveau paramètre sur une orbitale donnée augmente à peu près le temps de calcul de la valeur T_{ref} . T_{ref} est légèrement supérieur au temps mis pour calculer le même atome avec la même

fonction d'onde en VMC. Malgré l'efficacité de notre algorithme, le temps de calcul pour l'optimisation d'un système peut tout de même paraître important. Faisons plusieurs remarques à ce propos. Premièrement, l'optimisation est basée sur des quantités obtenues par simulation VMC. Nous avons donc parallélisé le code pour pouvoir calculer ces quantités efficacement, c'est à dire rapidement. Deuxièmement, rappelons que le but de ce travail est de permettre d'obtenir une bibliothèque de Jastrows orbitales préoptimisés. L'optimisation serait donc une étape effectuée une fois pour toute et il n'y aurait pas besoin de réoptimiser les coefficients à chaque calcul. Enfin, et c'est un point que nous discuterons dans la conclusion et les perspectives, l'usage de Jastrows orbitales permet d'envisager l'usage d'orbitales localisées pour inscrire notre démarche de construction "morçeau par morçeau" des fonctions d'onde des molécules complexes dans une optique de "linear scaling". Les matrices deviendraient des matrices creuses et permettraient des calculs bien plus rapides.

Chapitre 9

Etude des nœuds des fonctions d'onde CASSCF

9.1 Idées principales

Nous présentons ici la seconde partie de notre travail qui propose une stratégie pour l'étude des différences d'énergie à l'aide de fonctions d'onde CASSCF.

Les méthodes Monte-Carlo quantique représentent un moyen très efficace pour calculer l'énergie totale d'un état fondamental et rivalisent sur ce sujet avec les méthodes les plus précises de la chimie quantique (par exemple, CCSD(T) ou MRCI). Cependant, si elles permettent des calculs de grande précision pour les énergies totales, elles sont beaucoup moins performantes pour les différences d'énergie. Or, il est bien connu qu'en chimie c'est bien la connaissance de ces différences d'énergies qui compte en pratique : courbes de potentiel, affinités électroniques, potentiels d'ionisation, énergies de liaison, etc. Rappelons que dans un calcul fixed-node DMC, le résultat est obtenu avec une erreur systématique (biais) : l'erreur des nœuds fixés dépendant de la structure nodale de la fonction d'onde d'essai utilisée. Dans l'absolu cette erreur est très petite, de l'ordre de quelques pourcents de l'énergie de corrélation, et c'est ce qui permet d'obtenir de très bonnes énergies totales.

Malheureusement, lors de l'évaluation de la *différence* de deux énergies totales calculées par QMC, ces deux erreurs peuvent s'ajouter et l'erreur totale peut aller jusqu'à représenter de l'ordre de la différence d'énergie recherchée ! Notons que le problème de l'évaluation des différences d'énergies n'est pas propre aux méthodes QMC mais se retrouve dans les méthodes *ab initio* et DFT. Cependant, les algorithmes déterministes de ces méthodes permettent une meilleure compensation des erreurs : utilisation d'une même méthode, d'une même base, d'un même espace fonctionnel restreint. A l'heure actuelle, il n'existe pas de méthode standardisée en QMC permettant de contrôler les différences d'énergie pour un système arbitraire. Deux stratégies sont possibles :

- Réduire individuellement chacune des deux erreurs des nœuds fixés pour obtenir des énergies totales les plus précises possibles.
- Corréler le plus possible les erreurs liées aux nœuds pour maximiser la compensation d'erreur dans le calcul de la différence d'énergie.

Actuellement, la stratégie la plus employées est la première. Des progrès intéressants ont été faits dans ce domaine pour réduire le plus possible l'erreur fixed-node ([98], [147], [148], [149]). La stratégie consiste grâce à une simulation VMC à optimiser indépendamment les paramètres des fonctions d'essai (paramètres du Jastrow, des orbitales et des coefficients déterminantaux) pour diminuer séparément les erreurs des nœuds fixés des deux états considérés. Mais, optimiser en QMC les paramètres de la fonction d'essai est une procédure longue et coûteuse. De plus, il n'est pas du tout évident qu'une telle stratégie puissent être réalisable pour des systèmes de taille importante (plusieurs centaines d'atomes) avec des milliers de paramètres.

Notre stratégie relève de la deuxième approche. Nous souhaitons faire en sorte que les erreurs se compensent sans pour autant chercher nécessairement à réduire les erreurs des nœuds fixés individuelles. Pour cela, nous proposons d'utiliser une fonction d'onde de la forme CASSCF construite sur un espace actif pertinent et dont le jeu d'orbitales pour les deux états visés est commun. Ces orbitales communes sont optimisées à l'aide d'un calcul CASSCF usuel

avec une approche “State-Averaged” (SA). L’idée est de compenser les erreurs en utilisant des structures proches donnant des nœuds cohérents. Insistons sur le fait que nous utilisons une fonction d’onde construite à partir d’orbitales issues d’un calcul *ab initio*. Nous ne rajoutons pas de facteur de Jastrow et surtout nous ne réoptimisons pas les paramètres de la fonction d’onde. Cette remarque est très importante car obtenir une fonction d’essai à faible coût sans passer par une réoptimisation des paramètres est ce qui permettra à notre avis de traiter de gros systèmes.

Remarquons que l’utilisation d’orbitales “State-Averaged” en QMC n’est pas nouvelle. Dans la littérature, les premiers à utiliser de telles orbitales pour un calcul QMC sont Schautz et collaborateurs [98], [147]. Cependant, leur motivation concernant leur usage est différente de la nôtre. Dans les articles [98], [147], les orbitales SA sont utilisées pour des états de même symétrie pour éviter l’effondrement des deux niveaux d’énergies. Dans cette configuration, il est en effet possible qu’en optimisant séparément les deux états, que l’énergie de l’état excité devienne plus basse que l’énergie du fondamental. L’usage d’orbitales SA prévient ce risque. Ici, nous utilisons pour la première fois des orbitales SA pour deux états orthogonaux par symétrie. Le problème de l’effondrement des énergies ne se présente pas. Nous choisissons d’utiliser des orbitales SA pour garder la cohérence des nœuds. C’est d’ailleurs une des raisons pour lesquelles nous ne réoptimisons pas les coefficients. La réoptimisation pourrait détruire la cohérence.

Nous avons choisi de tester notre hypothèse sur une géométrie modèle de la molécule d’acroléine $\text{CH}_2=\text{CH}-\text{CHO}$ pour laquelle nous disposons de données issues de calculs *ab initio* avec lesquelles comparer nos résultats. Nous avons étudié plus précisément la transition $n \rightarrow \pi^*$ localisée principalement sur le groupe carbonyle. Un électron d’une paire libre n de l’oxygène est promu dans l’orbitale antiliante π^* de la liaison CO. Nous avons effectué nos calculs avec plusieurs bases et nous avons consigné nos résultats dans le tableau 9.3.

Sans surprise, en l’absence de réoptimisation, les énergies totales fixed-

node varient beaucoup, de l'ordre de la différence d'énergie étudiée. Cependant, avec l'utilisation d'orbitales State-Averaged les différences d'énergie sont beaucoup plus constantes quelque soit la base utilisée. En particulier, nous obtenons un gap convenable avec la base minimale. Le meilleur gap est obtenu avec un CAS(6,5) qui permet de décrire au mieux le problème chimique. Notre approche permet donc d'envisager l'étude des différences d'énergie de manière simple et à peu de frais (sans optimisation QMC). Bien évidemment notre approche est complémentaire avec notre travail précédent sur les fonctions d'onde avec Jastrows orbitales. Nous pouvons envisager de construire des fonctions d'onde à partir d'orbitales State-Averaged puis placer un Jastrow local sur le groupe fonctionnel où opère principalement la transition électronique. De cette manière, nous pouvons décrire localement la corrélation tout en gardant le reste de la molécule inchangé. C'est précisément la démarche que nous allons utiliser pour l'étude prochaine de la molécule de $C_{14}H_{13}O$. Nous employerons une fonction CASSCF avec un jeu d'orbitales State-Averaged et nous placerons un Jastrow local sur le groupe carbonyle.

9.2 Manuscript sur l’acroléine (J. Chem. Phys., sous presse)

A study of the fixed-node error in quantum Monte Carlo calculations of electronic transitions : The case of the singlet $n \rightarrow \pi^*$ (CO) transition of the acrolein

Thomas Bouabça, Nadia Ben Amor, Daniel Maynau, and Michel Caffarel

Laboratoire de Chimie et Physique Quantiques, CNRS-IRSAMC Université de Toulouse, France.

ABSTRACT

We report Fixed-Node Diffusion Monte Carlo (FN-DMC) calculations of the singlet $n \rightarrow \pi^*$ (CO) vertical transition of acrolein. The impact of the fixed-node approximation on the excitation energy is investigated. To do that, trial wavefunctions corresponding to various nodal patterns are used. They are constructed either by using a minimal Complete-Active-Space SCF (CASSCF) calculation involving an oxygen lone pair n and the $\pi^*(\text{CO})$ molecular orbitals or a more complete set involving all the molecular orbitals expected to play a significant role in the excitation process. Calculations of both states have been performed with molecular orbitals optimized separately for each state via standard “State Specific”(SS)-CASSCF calculations or by using a *common* set of optimized orbitals [“State Averaged”(SA)-CASSCF calculations] whose effect is to introduce some important correlation between the nodal patterns of the two electronic states. To investigate the role of the basis set three different basis of increasing size have been employed. The comparative study based on the use of all possible combinations of basis sets, active spaces, and type of optimized molecular orbitals shows that the

nodal error on the difference of energies is small when chemically relevant active space and state-averaged-type CASSCF wavefunctions are used, although the fixed-node error on the individual total energies involved can vary substantially. This remarkable result obtained for the acrolein suggests that FN-DMC calculations based on a simple strategy (use of standard *ab initio* wavefunctions and no Monte Carlo optimization of molecular orbital parameters) could be a working computational tool for computing electronic transition energies for more general systems.

9.3 Introduction

Quantum Monte Carlo (QMC) methods are known to be the most powerful approaches for computing total ground-state energies of many-electron systems (see, *e.g.* Ref.[1]). In short, QMC is based on two key ingredients : i.) The use of an approximate trial wave function, Ψ_T . ii.) The introduction of stochastic rules to sample the $3N$ -dimensional electronic configuration space according to some target probability density [proportional to Ψ_T^2 in Variational Monte Carlo (VMC) and $\Psi_T\Phi_0$ in Diffusion Monte Carlo (DMC), Φ_0 being the exact ground-state].

The approximate trial wavefunction is chosen to be the best compromise available between compactness and accuracy. A common choice for ψ_T consists of a product of two terms. The first term has a standard determinantal form describing the one-particle shell-structure of molecules. It is obtained from a preliminary Hartree-Fock (HF), Complete-Active-Space Self-Consistent-Field (CASSCF), Configuration Interaction (CI) or Density Functional Theory (DFT) calculation and is expressed as one (or a combination of a few) determinant(s) of single-particle orbitals. The second term is more original and is introduced to impose the electron-electron CUSP condition of the exact wave function and, also, to incorporate some explicit coupling between electron-nucleus and electron-electron coordinates (see, Ref.[2])

Regarding the choice of the stochastic rules, many variants have been

proposed. This diversity is at the origin of the great variety of acronyms found in the QMC literature : VMC, DMC, PDMC, GFMC, Projector MC, Reptation MC, etc... However, from a theoretical point of view, all these methods actually differ very little. In practice, the most widely used approach is the so-called Fixed-Node Diffusion Monte Carlo (FN-DMC) method and this will be the variant used in what follows. The stochastic rules of FN-DMC can be summarized as follows.

i.) A stochastic diffusion step corresponding to a free Brownian motion for the electrons.

ii.) A deterministic move done along the drift vector given as

$$\mathbf{b} = \nabla \Psi_T / \Psi_T. \quad (9.1)$$

iii.) A branching step simulated by the death or the duplication (birth) of the electronic configurations according to the magnitude of the local energy defined as

$$E_L \equiv H \Psi_T / \Psi_T. \quad (9.2)$$

The quantum Monte Carlo simulation is performed by applying these rules a large number of times over a large enough population of electronic configurations (or walkers). At equilibrium, the exact ground-state energy is obtained by simply averaging the local energy over the walkers.

In a FN-DMC calculation there exist essentially two types of error. The first one is the usual statistical error resulting from a finite simulation time. This error, present in any Monte Carlo scheme (either classical or quantum), behaves as $\frac{1}{\sqrt{K}}$ where K is the number of Monte Carlo steps. In principle, by using a *brute force* approach (more and more CPU time, *i.e.* $K \rightarrow \infty$), it is always possible to decrease the statistical down to the accuracy desired. The origin of the second type of error which will be the focus of the present study is less obvious and is related to the nodal properties of the trial wavefunction. To understand its origin, let us remark that at the nodes of Ψ_T (nodes are defined as configurations verifying $\Psi_T = 0$) the drift vector, Eq.(9.1), diverges and the nodal surfaces act as infinitely repulsive barriers for the

walkers. Consequently, each diffusing walker is trapped for ever within the nodal hyperdomain where it starts and the Schrödinger equation is solved stochastically *separately* within each nodal domain. Now, because the $(3N - 1)$ -dimensional nodes of an antisymmetric trial wavefunction are in general different from the exact ones, the FN-DMC simulation introduces a finite error related to the wrong location of these nodes. This error is referred to as the “fixed-node error” and it can be shown that the fixed-node energy is an upper bound of the exact energy (variational property).[3]

Now, the central issue is the importance of the fixed-node error. In other words, can the FN-DMC algorithm be used with success for realistic chemical problems or, stated differently, can QMC compete with the most accurate methods of quantum chemistry, *e.g.* CCSD(T) (Coupled-Cluster with Single and Double and Perturbative Triple excitations) or MRCI (MultiReference Configuration Interaction) methods? To answer this question it is necessary to distinguish between calculations of total energies and calculations of *differences* of energies. In the first case, it is true that QMC methods are presently the most powerful approaches for computing ground-state energies [and, more generally, ground-state energies in each symmetry sector (space and spin)]. For the various molecular systems treated up to now, it has been found that about 90 to 100% of the correlation energy can be recovered, the magnitude of the “small” residual fixed-node error depending on the type of systems treated and on the nodal pattern of the trial wave function employed. For small systems, such an accuracy corresponds typically to a CCSD(T) calculation with a large basis set (see, Ref.[4]). For larger systems such high-quality *ab initio* calculation are no longer possible [the formal scaling of CCSD(T) with number of electrons N is N^7] but FN-DMC simulations are still possible (to date, the largest calculation has involved about one thousand electrons, Ref.[5]) and there is no reason not to believe that a large fraction of the total correlation energy for these large systems is recovered. On the other hand, it is also clear that total ground-state energies by themselves are not of great interest for chemistry. Indeed, roughly speaking,

it can be said that all chemistry is about *differences* of energies. Chemists are interested in comparing the energies of different structures (reactants, products, etc.), potential energy profiles, height of barriers, electronic transitions, electron affinities, etc. Even the properties other than the energy can be reformulated as *infinitesimal* differences of energies using the Hellmann-Feynman theorem.[6]

At this point, a most important point consists in remarking that these differences of energies so important in chemistry are typically of the same order of magnitude as the fixed-node error itself on each component of the energy difference. Accordingly, we are in a tricky situation where some important and controlled compensation of errors must absolutely be at work if we want the FN-DMC method to be useful. However, it is most important to emphasize that this situation is not unique QMC methods. In any electronic structure calculations (either DFT or *ab initio* wavefunction-based) we are faced with the very same problem. In a DFT calculation, each energy component of a small difference of energies must absolutely be computed with the same method (LDA, B3LYP, GGA, etc.) and the same basis set. In *ab initio* approaches, the same method and basis sets must also be used and there is a long history about devising clever schemes for avoiding calculating large and expensive contributions which are supposed to cancel out when differences are computed [*e.g.*, perturbational approaches such as Symmetry-Adapted-Perturbation Theories (SAPT)[7]).

In this work we address this problem by reporting a detailed analysis of the role of the fixed-node approximation for a prototypical electronic transition occurring in an organic molecule. We have chosen the well-studied singlet $n \rightarrow \pi^*$ vertical transition on the carbonyl part of the acrolein (or propenal) molecule, C_3H_4O . In this transition a lone pair electron of the oxygen atom is promoted to the empty π^* antibonding orbital of the CO bond. The two total energies computed using FN-DMC correspond to two ground-state energies in their respective symmetry space (even/odd parity with respect to the molecular plane). For each state, various nodal patterns associated

with various trial wavefunctions have been employed. Despite the fact that, as expected, the fixed-node errors on each total energy may vary a lot (of the order of the transition energy) we have found that it is possible to propose a simple strategy for getting accurate and controlled transition energies.

Note that the use of FN-DMC for computing transition energies in organic molecules is not new. We can cite for example the work of Grossman *et al.*[8] about excitation energies for the silane and methane molecules, the work of Aspuru-Guzik *et al.*[9] on electronic transitions in free-base porphyrin, the work of Drummond *et al.*[10] on excitation energies and ionization energies of diamondoids, and very recently the study of Tiago *et al.*[11] on the low-energy excitations of several carbon fullerenes. In all these works, comparisons between QMC results and previous high-quality *ab initio* or DFT calculations and, also, experimental results, when available, have illustrated the excellent performances of QMC. However, no *systematic* study of the impact of the nodal structure on the results is in general done, even if, in some cases, some particular aspects are investigated : dependence on the type of orbitals (Hartree-Fock, Kohn-Sham, natural), sensitivity to the number of determinants used, etc. Now, for the reasons discussed above we would like to emphasize that, despite the good results obtained in these works, the problem of obtaining *systematic* good FN-DMC results for the small energy differences of chemistry is not at all solved. To give a precise example, in a recent work on the O₄ molecule, Caffarel and collaborators[12] have found that the height of the dissociation barrier (an energy difference of about 12 kcal/mol) is extremely dependent on the nature of the nodes.

In a number of recent works, Schautz, Filippi, and collaborators [13],[14],[15],[16] have made interesting contributions toward defining a more systematic strategy to obtain controlled energy differences. In these works, a full optimization of the trial wavefunctions at the VMC level is performed. The optimization involves not only the complete set of the parameters of the Jastrow part but also the parameters of all molecular orbitals involved, and the coefficients of the determinantal expansion. By doing this, the hope is to improve nodes

(although it is not certain[17]) and to get better differences of energy (not certain also). Here, we shall follow a quite different route. As we shall see, no optimization will be done and we shall strongly rely on the use of the *fixed* nodal patterns associated with simple and standard trial wavefunctions of quantum chemistry. Our general and long-term objective is to devise a simple (use of standard *ab initio* wavefunctions) and robust (no system-dependent optimization) strategy for computing electronic transitions, which is feasible not only for the small molecules but also for the large systems for which full optimizations appear difficult to perform. In this work, we have found that such a strategy leads to very good results in the case of the acrolein molecule.

The organization of this paper is as follows. In Section 2 the acrolein molecule and the electronic transition considered are presented. In Section 3 the computational details (methods, trial wavefunctions, and basis sets) are given. In Section 4 the FN-DMC results are presented and some comparisons with *ab initio* Configuration Interaction calculations are made. Finally, we present some summary and perspectives in the last section.

9.4 The acrolein molecule and the $n \rightarrow \pi^*$ transition

In this work we are concerned with the $n \rightarrow \pi^*$ vertical electronic transition of the acrolein (or propenal) molecule, $\text{CH}_2=\text{CH}-\text{CHO}$, between the two lower-lying singlet states (ground-state and first singlet excited-state). As represented pictorially in Fig.9.1 this excitation process is associated with the promotion of a lone pair electron of the oxygen atom (occupying the orbital n) to the empty π^* antibonding orbital of the CO group. In this work we have considered two different geometries for the molecule. A first one corresponds to a model geometry which has been used in previous works[18] and which allows to perform easily systematic studies for series of polyenals of increasing sizes, the acrolein molecule being the smallest one. The geometry considered corresponds to a *s-cis* configuration and the internuclear distances are given by : $d_{C-O} = 1.22$, $d_{C-C} = 1.45$, $d_{C=C} = 1.35$, $d_{C-H} = 1.1$, and $\widehat{XCY} = 120^\circ$ (for all combinations X,Y=C,H,O). To make contact with ex-

periments we also consider the experimental geometry as obtained by Blom and collaborators[19] using microwave spectroscopy. The geometry, given in Ref.[19] and represented in Fig.9.1, is that of the *s-trans* isomer which, at room temperature, represents the great majority of the molecules.[20] Note that in both cases the geometry corresponds to a planar structure and the molecule has a Cs symmetry group. Molecular orbitals are classified according to the two irreducible representations \mathbf{A}' (symmetric with respect to the plane) and \mathbf{A}'' (antisymmetric). The closed-shell ground-state is of symmetry \mathbf{A}' while the open-shell excited-state belongs to \mathbf{A}'' .

9.5 Computational details

Methods. The computational method at the heart of this study is the Fixed-Node Diffusion Monte Carlo (FN-DMC). In the Introduction, the key ideas of the method necessary to the understanding of the present work have already been presented. For details of FN-DMC the reader is referred, *e.g.*, to Ref.[21] or Ref.[1]. For comparisons, we also present results obtained with the Configuration Interaction (CI) method using single and double excitations on a mono-(SCF) or multi-(CASSCF) configurational references. CI being a standard and well-documented approach, we refer the reader interested by the details to the relevant literature.

Trial wavefunctions. The fundamental quantity governing the efficiency of our Monte Carlo simulations is the trial wavefunction. It enters the simulation via the drift vector, Eq.(9.1), and the local energy, Eq.(9.2). The key point is that “good” trial wavefunctions is associated with small fixed-node biases (errors due to approximate nodes) and small statistical errors on the energy estimator. In the limit of an exact trial wavefunction, both errors entirely vanish. This property is known as the zero-variance zero-bias property.[6] Here, we have chosen to employ a form consisting of the product of a Jastrow part and a Slater-type part consisting of a linear combination

of a small number of determinants

$$\psi_T(\vec{R}) = \exp \left[\sum_{\alpha} \sum_{\langle i,j \rangle} U(r_{i\alpha}, r_{j\alpha}, r_{ij}) \right] \sum_k c_k D_k^{\uparrow}(\vec{R}) D_k^{\downarrow}(\vec{R}) \quad (9.3)$$

where the sum over α denotes a sum over the nuclei, $\sum_{\langle i,j \rangle}$ a sum over the pair of electrons, c_k the coefficients of the multiconfigurational expansion, and D_k^{σ} ($\sigma = \uparrow$ or \downarrow) are determinants made of one-particle space-orbitals. The exponential Jastrow term which depends explicitly on interelectronic distances is introduced to reproduce the small-electron-electron-distance behavior of the exact wavefunction (CUSP condition) and, also, to incorporate explicit couplings between electron-nucleus and electron-electron coordinates. Here, we have chosen to use the form introduced in Ref.[6]. Regarding the determinantal part, we have used either a monoconfigurational form obtained via a Self-Consistent-Field (SCF) calculation or a multiconfigurational form obtained via a Complete Active Space Self-Consistent-Field (CASSCF) calculation. The CASSCF form is introduced to incorporate the main part of the multiconfigurational character of the exact wavefunction issued from the interference between the various low-energy components (“non-dynamic” correlation effects). This multiconfigurational character is known to play an important role for excited states. The choice of the active space is crucial to get a chemically relevant CASSCF wavefunction. Here, two choices have been considered. The first choice is the minimal one. It consists in considering only the two molecular orbitals associated with the oxygen lone pair n and the π^* orbital of the CO bond. The resulting active space built with two electrons within two orbitals is denoted as CAS(2,2). Because of the Cs symmetry it is of dimension 2 both for the ground- and excited- states. The second choice is based on the fact that the CO π bond is expected to interact with the π system of the C=C double bond. Letting aside the much more inert σ -bond system we propose to consider the active space built by distributing six electrons within the five molecular orbitals $n, \pi_{CO}, \pi_{CO}^*, \pi_{CC}, \pi_{CC}^*$ (denoted as the CAS(6,5) active space). Using symmetry considerations this active space consists of 52 and 48 determinants for the ground- and excited-state,

respectively.

To build the Slater-type multiconfigurational wave functions two types of approach have been considered. First, we have made standard CASSCF calculations for both states. In such calculations the molecular orbitals (both occupied and active) as well as the coefficients of the determinantal expansion are fully optimized (in the deterministic *ab initio* sense) in an *independent* way for the two states. Such calculations are referred here to as State Specific (SS-CASSCF). In the second approach, the CASSCF wavefunctions are optimized in a *correlated* way using a common set of molecular orbitals, the linear coefficients of the expansion being still taken as independent. In practice, this is realized by making stationary with respect to the various parameters a weighted average of the energies of the states under consideration (the standard choice consisting in taking equal weights is adopted here). This approach is known as a State-Averaged (SA) method.[22]

Regarding optimization of the trial wavefunction very little has been done here : Only a few parameters of the Jastrow factor have been optimized in order to reduce the statistical fluctuations. It has been done separately for both states using standard techniques (see, Ref.[23]). Let us emphasize that by optimizing only the Jastrow part as done here the nodal patterns of the various *ab initio* trial wavefunctions introduced in this work are kept unchanged.

Finally, let us give some practical details. The present FN-DMC calculations are all-electron calculations done with a very small time-step, $\tau = 0.00025$, to insure a proper treatment of the nodal hypersurfaces and to remove the time-step error within statistical uncertainties. For a system consisting of light atoms such as the acrolein some care has to be taken for properly reproducing the electron-nucleus cusp both for the core and valence electrons. Regarding the core region, we have replaced the $1s$ atomic orbitals (oxygen and carbon atoms) expanded over the gaussian basis set by the $1s$ Slater-type orbitals given in the Clementi and Roetti's Tables [24]. On the other hand, the valence molecular orbitals are also modified at short nuclear

distances to impose the nuclear cusp ; we do that using a short- r representation of the radial part of orbitals under the form $c(1 + \alpha r^2 + \beta r^3) \exp(-\gamma r)$, in the same spirit as Ref.[25]. For each trial wave function and simulation, the calculations are extensive and have been carried over a large number of processors, the total statistics representing about 10^9 Monte Carlo steps.

Basis sets. Three Gaussian basis sets of increasing size have been considered. They are extracted from the Roos ANO basis sets presented in Ref.[26]. Basis set #1 consists of (14s,9p) \rightarrow [3s,2p] for the oxygen and carbon atoms, and (8s) \rightarrow [2s] for H (44 atomic basis functions). Basis set #2 is given by (14s,9p,4d) \rightarrow [3s,2p,1d] for O and C, and (8s,4p) \rightarrow [2s,1p] for H (76 basis functions). Basis set #3 is given by (14s,9p,4d,3f) \rightarrow [4s,3p,2d,1f] for O and C, and (8s,4p,3d) \rightarrow [3s,2p,1d] for H (176 basis functions).

9.6 Results

9.6.1 Configuration Interaction

SCF, CASSCF and CISD (Configuration Interaction Single and Double) calculations for the model geometry are presented in Table 9.1. We also present EOM-CCSD (Equation-Of-Motion Coupled-Cluster Single and Double) calculations to get an alternative estimate of the transition energy. Note that the largest calculations presented here are the CAS(6,5)-CISD. For these calculations the transition energies are similar and vary between 3.42 eV and 3.59 eV. The EOM-CCSD calculations vary between 3.51 eV and 3.63 eV. Consequently, an estimate for the gap of about 3.5 eV at this geometry appears reasonable. In what follows, we will refer to this value as the reference energy.

Let us now comment in more detail the results obtained. First, let us begin with the calculations at the SCF and CASSCF levels (no treatment of the dynamic correlation effects). At the SCF level (one determinant for the closed-shell ground-state of symmetry \mathbf{A}' and a two-determinant CAS(2,2) structure for the open-shell excited-state to recover the \mathbf{A}'' symmetry) the

gap is found too small (about 2.80 eV), in error of about 20%. Note that the result obtained with the small basis (basis # 1) is rather different from those obtained with the two other basis sets. However, results with basis # 2 and # 3 are very similar. This observation turns out to be true for all data presented below, thus indicating that finite basis set effects are marginal here for basis sets of a quality comparable to basis # 2.

Using the “minimal” active space consisting of the oxygen lone pair and π_{CO}^* orbital [SA-CAS(2,2) calculation] the gap is seriously increased up to the value of about 3.3 eV. When the full “chemically relevant” active space is employed [CAS(6,5) calculations] the reference value for the gap is approximately obtained. Note the difference of results for the State Specific (SS) and State Averaged (SA) calculations. In the first case the gap found is found to be slightly too large by about 0.1 eV. In the second case the gap obtained is close to the correct value. Finally, let us comment the impact of dynamical correlation contributions by looking at the configuration interaction results. Starting from a single SCF reference (SR-CI) the CI result (about 3.3 eV) improves the SCF result (about 2.8 eV) but is not able to recover the reference value. This illustrates clearly the necessity of introducing the non-dynamical correlation contributions. At the CAS(2,2)-CISD state-specific level the value is slightly improved (gap of about 3.4 eV) but is still in error of about 0.1 eV. When using a CAS(2,2) reference function and a state-averaged approach the reference value is recovered. Using a CAS(6,5) reference function does not appear to change the results, both SS and SA approaches leading indeed to the reference value. However, because of the CAS(6,5)-CASSCF results discussed above, it should be suspected that this result is a consequence of some cancellation of errors between missing dynamic and non-dynamic correlation effects.

9.6.2 Fixed-Node Diffusion Monte Carlo

Our all-electron FN-DMC calculations for the ground- and first-excited state energies of the acrolein (model geometry) are presented in Table 9.2.

Trial wavefunctions having various nodal patterns have been used. The trial wavefunctions have been built from SCF, SS-CASSCF, and SA-CASSCF Slater part with molecular orbitals optimized (in the deterministic *ab initio* sense) using the three basis sets of increasing sizes. Monte Carlo simulations have been performed to reach the 0.05-0.10 eV accuracy on the transition energies.

A first important remark is the great quality of the FN-DMC total energies. For example, the best QMC value obtained here for E_0 is -191.841 ± 0.003 [FN-DMC with CAS(2,2)-SA nodes and basis #2 in Table 9.2]. To compare with, we have performed CCSD(T) calculations with large basis sets, the results obtained being as follows : $E_0 = -191.41067104$ (using cc-pCVDZ), $E_0 = -191.59862499$ (cc-pCVTZ), and $E_0 = -191.65139255$ (using cc-pCVQZ). This result exemplifies the fact that QMC is a highly accurate method for computing total energies, a point already emphasized in the introduction. A second important remark is that the fixed-node error magnitude as a function of the nodal pattern used (the nodes are changed when using different levels of theory and basis sets) is large at the scale of our problem. For E_0 the higher and lower values are $-191.719(3)$ and $-191.841(3)$, respectively. This makes a variation of $0.122(4)$ a.u, that is 3.3 eV. For E_1 the variation is 3.4 eV. As seen, the range of the fixed-node energy variation is of the order of the gap itself. This result illustrates clearly the need of some “constructive correlation” between the fixed-node error on both energy components if some meaningful result for the transition energy is desired. However, as already pointed out, this fundamental problem is not specific to QMC but is present in any electronic structure method. For example, in the case of the best CAS(2,2)-CISD calculations presented in Table 9.1, the variation of the total energies as a function of the basis set used is about 11.6 eV, more than three times the FN-DMC result and the gap in energy itself! Another important remark regarding the FN-DMC results of Table 9.2 is that the fixed-node total energies does not necessarily improve (*i.e.*, decrease) as the size of the basis set is increased. This result is a nice illustra-

tion of the fundamental point already stressed in the introduction regarding the fact that decreasing the total variational energy does not necessarily improve the nodes of the trial wavefunction (see, also, comment [17]).

Using the SCF nodes, the QMC gap obtained varies between 3.02 and 3.37 eV. Such values obtained with mono-determinantal wavefunctions are reminiscent of the Single Reference-CI calculations results presented above. Using CASSCF nodes the value of the gap is increased. With the SA-CAS(2,2) nodal pattern values close to the reference gap are obtained. However, the fluctuations of the results as a function of the basis set are large, namely 0.36 eV. At the SS-CAS(6,5) level the fluctuations are still larger (0.55 eV) and the average value of the gap seems to be too large (3.92 eV). This is also reminiscent of the SS-CAS(6,5) CASSCF calculations which gives a slightly too large value. Finally, with the SA-CAS(6,5) nodes the results obtained are much more coherent : Remarkably, the fluctuations of the results as a function of the basis set are almost suppressed and the average value is excellent.

A few general comments are now in order. As seen from our results the quality of the nodes of the trial wavefunction plays indeed a central role. Let us insist that imposing approximate nodes is the sole source of systematic error within the FN-DMC framework (the statistical error is controllable by making sufficient long simulations). In the case of the acrolein molecule, using “monodeterminantal” or “SCF” nodes leads to too small transition energies, although a large part of the correlation energy is recovered via FN-DMC. This result is also true in extensive *ab initio* CISD calculations based on a *mono-configurational* reference wavefunction. This important result suggests that there probably exists some relation between the nature of the correlation energy recovered by FN-DMC and the nature of the nodes employed. To be more precise, using SCF nodes the DMC process would recover essentially the dynamic correlation effects and very little of the non-dynamic correlation component. To recover this last contribution, we really need nodes which have some “multiconfigurational” character. Physically,

such statement makes sense but, mathematically, the situation is much less clear. Note that this idea is supported by some recent calculations done by Caffarel and Ramirez[27] on the chromium dimer at the equilibrium distance. It is known that the electronic structure of this molecule is very hard to recover and that Cr_2 is considered as a real nightmare for standard methods of quantum chemistry. In particular, the ground-state wavefunction has a very strong multiconfigurational character. It has been shown that using SCF nodes a very small fraction of the binding energy is recovered by FN-DMC. This result is similar to what has been obtained by Scuseria[28] using single-reference CCSD(T) calculations using very large basis sets. Here, also, by using “mono-configurational” nodes in FN-DMC it seems that only the dynamic correlation contribution is recovered.

Another remark is that “State Averaged” CASSCF nodes appears effective to get the correct estimate transition energy. Such a result can be understood by the fact that using the same set of molecular orbitals for the two states allows large compensations of the fixed-node error components corresponding to configuration space regions where the nodal structure is common to both states and not primarily involved in the excitation process. For the acrolein molecule we have found that by using a minimal active space the correct value is recovered but the sensitivity of the value to the basis set is important. When the active space contains all molecular orbitals which are critical in the excitation process, this sensitivity appears to be much less important. Due to the local character of these excitations, we expect that such a result could extend to a large class of organic compounds and, more generally, to any system where a chemically active space can be properly defined. Note that, to the best of our knowledge, the first use of SA-CASSCF wavefunctions for FN-DMC has been done by Schautz and collaborators[14],[13]. In these works, SA-CASSCF wavefunctions have been introduced essentially to avoid the so-called root flipping problem occurring in MCSCF calculations (optimization of a state in presence of lower states of the same symmetry). Here, the motivation is different : we employ a state-averaged approach to

get two CASSCF wavefunctions constructed from a common set of optimized molecular orbitals. In this way, we can expect a maximal similarity between the nodal hypersurfaces of both states and, thus, a minimization of the fixed-node error on the energy difference.

Finally, in Table 9.3 the configuration interaction and FN-DMC results at the experimental equilibrium geometry[19] are presented. The largest CI calculation [CAS(6,5)-CISD basis #2 (SA)] leads to a value of 3.84 eV. The FN-DMC result is 3.86 ± 0.07 eV, in full agreement with the *ab initio* transition energy. To the best of our knowledge, the experimental data of the literature fall into the range (3.69, 3.75) eV.[29],[30],[31] According to these data, the *ab initio* value is found to be slightly greater than expected. The FN-DMC value is also slightly greater. However, considering the statistical character of the result, there is a probability of about 6% that our transition energy falls into the previous experimental interval.

9.7 Summary and Perspectives

In this work we have presented FN-DMC calculations of the energy gap of a prototypical electronic transition in a simple organic molecule having a carbonyl group. To elucidate the impact of the fixed-node approximation on the computation of electronic energy differences and, more specifically, on the possible existence of *compensation of errors*, we have performed FN-DMC simulations under various conditions. These conditions include basis sets of various sizes, CASSCF wavefunctions with different active spaces, and molecular orbitals optimized for each state separately or molecular orbitals common to both states. Changing these conditions leads to changing the nodal hypersurfaces and, therefore, the magnitude of the fixed-node error. We emphasize that, in contrast with some other groups[13], [14],[15],[16] we propose here not to change the nodal structure of the various *ab initio* trial wavefunctions used [in other words, no optimization of the determinantal part in Eq.(9.3)]. Besides avoiding the optimization of many (hundred's) of parameters (always a not-so-easy task, particularly for large systems), our

major motivation is to avoid as much as possible the loss of the coherence between the two nodal patterns. In agreement with the common wisdom of the domain, it has been found that the impact of each of the previous factors is very important. More precisely, by changing the various conditions the fixed-node error on both total energies can vary by an amount which is of the order of magnitude of the energy difference. This result -quite common in standard DFT or *ab initio*-wavefunction based methods- confirms the fact that any realistic FN-DMC calculation must be based on a large and controlled compensation of errors. In other words, the error made on the location of the nodes for a given state must be tightly correlated to the error made for the other state. To achieve this, we have proposed to use the same set of optimized molecular orbitals for describing the two states involved in the transition (the so-called “State Averaged” approach of *ab initio* approach). By doing this the total energies obtained are less accurate but we introduce some “coherence” between the nodal shapes of the two states. Our FN-DMC results show, at least for the system studied, that this compensation works very well. Indeed, it has been found that in a state-average calculation using a chemically meaningful active space, the energy gap is nearly independent on the basis set, despite the fact that fixed-node errors on separate total energies vary substantially. Such a result is particularly interesting since it suggests that meaningful results can be obtained by using a simple strategy based on standard and compact *ab initio* trial wavefunctions without resorting to large-scale optimizations which will become increasingly difficult for large systems. However, to validate our approach it is clear that the next step will consist in investigating larger molecular systems.

Acknowledgments We would like to thank Dr. Sophie Hoyau for helpful references and IDRIS (CNRS, Orsay), CCRT (CEA/DAM, Ile-de-France), and CALMIP (Université de Toulouse) for computational support.

Bibliographie

- [1] W.M.C. Foulkes, L. Mitáš, R.J. Needs, and G. Rajogopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- [2] K. E. Schmidt and J. W. Moskowitz, *J. Chem. Phys.* **93**, 4172 (1990).
- [3] P.J. Reynolds, D.M. Ceperley, B.J. Alder, and W.A. Lester Jr., *J. Chem. Phys.* **77**, 5593 (1982).
- [4] J. C. Grossman, *J. Chem. Phys.* **117**, 1434 (2002).
- [5] A.J. Williamson, R.Q. Hood, and J.C. Grossman, *Phys. Rev. Lett.* **87**, 246406 (2001).
- [6] R. Assaraf and M. Caffarel, *J. Chem. Phys.* **119**, 10536 (2003).
- [7] B. Jeziorski, R. Moszynski, and K. Szalewicz, *Chem. Rev.* **94**, 1887 (1994).
- [8] J.C. Grossman, M. Rohlfing, L. Mitas, S.G. Louie, and M.L. Cohen, *Phys. Rev. Lett.* **86**, 472 (2001).
- [9] A. Aspuru-Guzik, O. El Akramine, J.C. Grossman, and W.A. Lester, Jr. *J. Chem. Phys.* **120**, 3049 (2004).
- [10] N.D. Drummond, A.J. Williamson, R.J. Needs, and G. Galli, *Phys. Rev. Lett.* **95**, 096801 (2005).
- [11] M.L. Tiago, P.R.C. Kent, R.Q. Hood, and F.A. Reboredo, *J. Chem. Phys.* **129**, 084311 (2008).
- [12] M. Caffarel, R. Hernández-Lamonedá, A. Scemama, and A. Ramirez-Solis, *Phys. Rev. Lett.* **99**, 153001 (2007).
- [13] F. Schautz and C. Filippi, *J. Chem. Phys.* **120**, 10931 (2004).

- [14] F. Schautz, F. Buda, and C. Filippi, *J. Chem. Phys.* **121**, 5836 (2004).
- [15] F. Cordova, L. Joubert Doriol, A. Ipatov, and M.E. Casida, *J. Chem. Phys.* **127**, 164111 (2008).
- [16] E. Tapavicza, I. Tavernelli, U. Rothlisberger, C. Filippi, and M.E. Casida, *J. Chem. Phys.* **129**, 124108 (2008).
- [17] Decreasing the variational energy does not lead systematically to a decrease of the fixed-node error. For example, it is quite easy to build a “bad” trial wavefunction (associated with a high VMC energy) with exact nodes which will lead to an exact FN-DMC energy, and a “very good” (much lower VMC energy) trial wavefunction with slightly wrong nodes which will lead to a biased higher FN-DMC energy.
- [18] J. Pitarch-Ruiz, S. Evangelisti, and D. Maynau, *Chem. Phys. Lett.* **372**, 22 (2003).
- [19] C. E. Blom, G. Grassi, and A. Bauder, *J. Am. Chem. Soc.* **106**, 7427 (1984).
- [20] C. E. Blom, R. P. Møller, and Hs. H. Gønthard, *Chem. Phys. Lett.* **73**, 483 (1980).
- [21] B.L. Hammond, W.A. Lester Jr., and P.J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, World Scientific Lecture and Course Notes in Chemistry Vol.1 (1994).
- [22] K. Docken and J. Hinze, *J. Chem. Phys.* **57**, 4928 (1972); H.-J. Werner and P. J. Knowles, *J. Chem. Phys.* **82**, 5053 (1985).
- [23] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, *Phys. Rev. Lett.* **60**, 1709 (1988).
- [24] E. Clementi and C. Roetti, *Atomic Data and Nuclear Data Tables* **14** 177 (1974).
- [25] S. Manten and A. Lüchow, *J. Chem. Phys.* **115**, 5362 (2001).
- [26] P. O. Widmark, P. A. Malmqvist, and B. Roos, *Theor. Chim. Acta* **77**, 291 (1990).

- [27] M. Caffarel and A. Ramirez-Solis, "Fixed-Node Quantum Monte Carlo for Chemistry" in *The proceedings of the 14th International Conference : "Recent Progress in Many-Body Theories"*, World Scientific ed. by R.F. Bishop, C.E. Campbell, J.W. Clark, and S. Fantoni, (2008).
- [28] G. E. Scuseria, *J. Chem. Phys.* **94**, 442 (1991).
- [29] A. Luthy, *Z. Phys. Chem.* **107**, 284 (1923).
- [30] F. E. Blacet, W. G. Young, and J. G. Roof, *J. Am. Chem. Soc.* **59**, 608 (1937).
- [31] K. Inuzuka, *Bull. Chem. Soc. Jpn.* **33**, 678 (1960).

FIGURE CAPTION

- Fig. 1 Pictorial representation of the excitation process corresponding to the promotion of a lone pair electron of the oxygen atom (orbital n) to the empty π^* antibonding orbital of the CO group. The geometry considered is the experimental geometry proposed by Blom and collaborators[19].

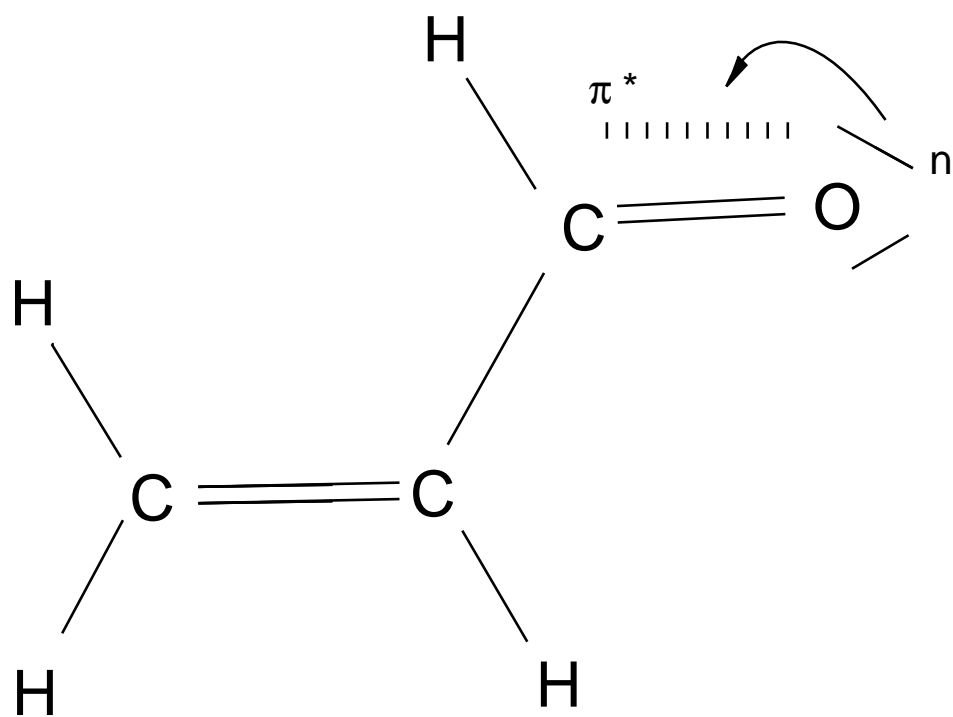


FIG. 9.1 -

TAB. 9.1 – All-electron *ab initio* calculations of the ground- and first-excited state energies of the acrolein molecule using different levels of theory. The geometry of the molecule is the model geometry described in the text.

	E_0 (a.u.)	E_1 (a.u.)	$\Delta = E_1 - E_0$ (eV)
CASSCF			
SCF for E_0 ; CAS(2,2) for E_1 basis #1	-190.70336212	-190.60950539	2.55 eV
SCF for E_0 ; CAS(2,2) for E_1 basis #2	-190.81156558	-190.70744919	2.83 eV
SCF for E_0 ; CAS(2,2) for E_1 basis #3	-190.83558573	-190.73282405	2.80 eV
CAS(2,2) State-Averaged (SA) basis #1	-190.69055231	-190.57594043	3.12 eV
CAS(2,2) State-Averaged (SA) basis #2	-190.80427004	-190.68039963	3.37 eV
CAS(2,2) State-Averaged (SA) basis #3	-190.82377795	-190.70148301	3.33 eV
CAS(6,5) ^a State-Averaged (SA) basis #1	-190.76639154	-190.64370613	3.34 eV
CAS(6,5) ^a State-Averaged (SA) basis #2	-190.87061536	-190.74188978	3.50 eV
CAS(6,5) ^a State-Averaged (SA) basis #3	-190.88959034	-190.76208712	3.47 eV
CAS(6,5) ^a State Specific (SS) basis #1	-190.77480828	-190.64806778	3.45 eV
CAS(6,5) ^a State Specific (SS) basis #2	-190.88157716	-190.74801729	3.63 eV
CAS(6,5) ^a State Specific (SS) basis #3	-190.90082219	-190.76883560	3.59 eV
CASSCF-CISD			
SR-CISD basis #1 (SS)	-191.05652956	-190.93663809	3.26 eV
SR-CISD basis #2 (SS)	-191.33909115	-191.21625317	3.34 eV
SR-CISD basis #3 (SS)	-191.48446942	-191.36425878	3.27 eV
CAS(2,2)-CISD basis #1 (SA)	-191.05316181	-190.92750265	3.42 eV
CAS(2,2)-CISD basis #2 (SA)	-191.33510751	-191.20593613	3.51 eV
CAS(2,2)-CISD basis #3 (SA)	-191.47976806	-191.35299762	3.45 eV
CAS(2,2)-CISD basis #1 (SS)	-191.05793972	-190.93693525	3.29 eV
CAS(2,2)-CISD basis #2 (SS)	-191.34134483	-191.21667357	3.39 eV
CAS(2,2)-CISD basis #3 (SS)	-191.48715772	-191.36472344	3.33 eV
CAS(6,5)-CISD basis #1 (SA)	-191.08041005	-190.95368831	3.42 eV
CAS(6,5)-CISD basis #2 (SA)	-191.36518607	-191.23539368	3.53 eV
CAS(6,5)-CISD basis #1 (SS)	-191.08279209	-190.95456833	3.49 eV
CAS(6,5)-CISD basis #2 (SS)	-191.36884345	-191.23684021	3.59 eV
EOM-CCSD :			
231			
basis #1			3.51 eV
basis #2			3.63 eV
basis #3			3.61 eV

^a 52 determinants for the ground-state and 48 determinants for the excited-state.

TABLE 9.2 – All-electron FN-DMC calculations of the ground- and first-excited state energies of the acrolein molecule using different nodal structures. The geometry of the molecule is the model geometry described in the text.

	E_0 (a.u.)	E_1 (a.u.)	$\Delta = E_1 - E_0$ (eV)
Quantum Monte Carlo :			
FN-DMC[nodes : SCF]			
Basis #1	-191.814 ± 0.0045	-191.690 ± 0.0023	3.37 ± 0.14 eV
Basis #2	-191.819 ± 0.0020	-191.708 ± 0.0042	3.02 ± 0.13 eV
Basis #3	-191.729 ± 0.0037	-191.608 ± 0.0021	3.29 ± 0.12 eV
FN-DMC[nodes :CAS(2,2)SA]			
Basis #1	-191.803 ± 0.0015	-191.665 ± 0.0013	3.76 ± 0.05 eV
Basis #2	-191.841 ± 0.0031	-191.716 ± 0.0042	3.40 ± 0.14 eV
Basis #3	-191.763 ± 0.0021	-191.634 ± 0.0022	3.51 ± 0.08 eV
FN-DMC[nodes :CAS(6,5) ^a SS]			
Basis #1	-191.795 ± 0.0033	-191.663 ± 0.0034	3.59 ± 0.13 eV
Basis #2	-191.755 ± 0.0022	-191.603 ± 0.0031	4.14 ± 0.10 eV
Basis #3	-191.744 ± 0.0048	-191.596 ± 0.0021	4.03 ± 0.14 eV
FN-DMC[nodes :CAS(6,5) ^a SA]			
Basis #1	-191.791 ± 0.0026	-191.663 ± 0.0025	3.48 ± 0.10 eV
Basis #2	-191.719 ± 0.0027	-191.592 ± 0.0026	3.46 ± 0.10 eV
Basis #3	-191.739 ± 0.0034	-191.611 ± 0.0023	3.48 ± 0.11 eV

^a 52 determinants for the ground-state and 48 determinants for the excited-state.

^b More precisely, SCF nodes for E_0 and CAS(2,2) nodes for E_1 .

TAB. 9.3 – CI and FN-DMC calculations of the transition energy at experimental equilibrium geometry. Comparison with experiment.

	E_0 (a.u.)	E_1 (a.u.)	$\Delta = E_1 - E_0$ (eV)
FN-DMC[nodes :CAS(6,5)SA] Basis #3	-191.8504 ± 0.0020	-191.7086 ± 0.0023	3.86 ± 0.07 eV
CAS(2,2)-CISD basis #1 (SA)	-191.05867538	-190.92319083	3.69 eV
CAS(2,2)-CISD basis #2 (SA)	-191.34228392	-191.20106117	3.84 eV
CAS(2,2)-CISD basis #3 (SA)	-191.48744101	-191.34853105	3.78 eV
CAS(6,5)-CISD basis #1 (SA)	-191.08525253	-190.94917907	3.70 eV
CAS(6,5)-CISD basis #2 (SA)	-191.37158070	-191.23030715	3.84 eV
Experimental results			3.69-3.75eV ^a

^a Refs.[29],[30],[31]

Chapitre 10

Conclusion générale

Dans ce travail de thèse, nous avons proposé deux contributions méthodologiques pour les approches de type Monte Carlo quantique (QMC) appliquées à l'étude de la structure électronique des molécules. Ces contributions visent à permettre une application des méthodes QMC beaucoup plus "standardisée" qu'elles ne le sont actuellement, c'est à dire qui soient applicables à n'importe quel système moléculaire sans effort particulier. La philosophie générale sous-tendant notre travail est d'utiliser des fonctions d'essai simples mais "physiques" et dont l'optimisation reste très limitée.

Dans la première partie de notre travail nous avons proposé et implémenté notre nouvelle fonction d'onde de type multi-Jastrow dans le code QMC=Chem. Nous avons validé notre approche en testant nos idées sur les atomes d'oxygène, de fluor, de soufre et de cuivre, ainsi que sur la molécule HF. Nous avons pu constater pour les atomes une amélioration de l'énergie variationnelle associée à la nouvelle fonction d'onde, avec des résultats particulièrement spectaculaires pour l'atome le plus lourd, le cuivre, pour lequel l'utilisation de Jastrow différents dans chacune des couches paraît particulièrement pertinent. Néanmoins, avec cette fonction multi-Jastrow nous n'avons pas observé d'amélioration (abaissement) de l'énergie fixed-node DMC. Il faut cependant rester prudent concernant ce dernier point. En effet, nous avons constaté qu'en annulant le coefficient du cusp electron-noyau pour

l'atome de fluor, nous augmentions l'énergie fixed-node DMC. La structure nodale reste donc encore bien mystérieuse mais notons que notre nouvelle fonction d'onde offre un moyen de l'étudier plus en détail. Dans le cas de la molécule HF nous avons réussi à modifier l'énergie fixed-node et établi grâce à nos Jastrows orbitaux une courbe de potentiel de très bonne qualité.

Dans la seconde partie nous avons testé notre stratégie pour calculer les différences d'énergie par l'usage d'une fonction d'onde State-Averaged-CASSCF et nous avons validé cette approche pour la molécule d'acroléine. L'étape suivante sera de valider nos idées sur des systèmes de tailles plus importantes et de coupler les deux parties de ce travail de thèse (fonctions multi-Jastrow et différences d'énergie obtenues à partir d'orbitales moléculaires SA-CASSCF). Ainsi, nous envisageons prochainement l'étude de la transition $n \rightarrow \pi^*$ dans la molécule de tridecénal $C_{13}H_{14}O$ (100 électrons). Pour étudier cette différence d'énergie nous utiliserons des fonctions d'onde SA-CASSCF et nous rajouterons un Jastrow orbitalaire sur le groupe fonctionnel carbonyle.

Concernant notre fonction d'onde multi-Jastrow, nous nous sommes aperçus que l'étape d'optimisation n'était pas aussi simple que cela. Nous pensons qu'il est encore possible d'effectuer un travail dans ce sens pour la rendre plus aisée. De plus, nous suggérons qu'un travail supplémentaire soit effectué sur la forme du Jastrow afin d'analyser plus en détail l'importance et la pertinence des coefficients. Nous rappelons que le point important pour construire une telle fonction d'onde locale est de conserver l'antisymétrie de la fonction d'onde grâce une symétrie sur les composantes du Jastrow. D'autres formes de Jastrow pourraient être proposées sur ce modèle. En utilisant notre fonction d'onde, il est désormais possible d'amorcer notre projet de bibliothèque de Jastrow qui nous permettra de construire de manière automatisée des fonctions d'onde de qualité constante pour des systèmes de plus en plus complexes. (construction "morçeau par morçeau" des fonctions d'onde). Cette étape ne pourra être réalisée qu'après avoir validé l'utilisation des orbitales strictement localisées pour permettre un scaling linéaire

de notre algorithme.

A partir de ce moment là, nous aurons à notre disposition tous les outils nécessaires à l'étude des gros systèmes, qu'ils soient d'intérêt biologique ou autres.

ANNEXES

Annexe A

Dérivation mathématique de la formule de Feynman-Kac

La formule de Feynman-Kac établit un lien entre équation aux dérivées partielles et processus stochastiques. Pour simplifier notre exposé nous considérerons le cas unidimensionnel. L'évolution d'une probabilité de transition qui décrit un processus Markovien est donnée par l'équation backward de Kolmogorov :

$$\frac{\partial p}{\partial t} + \mu(x, t) \frac{\partial}{\partial x} p(x, t) + \frac{1}{2} \sigma^2(x, t) \frac{\partial^2 p(x, t)}{\partial x^2} = 0 \quad (\text{A.1})$$

$\mu(x, t)$ est le vecteur de dérive ou drift.

$\sigma^2(x, t)$ est le coefficient de diffusion.

Nous imposons comme conditions aux limites : $p(x, T) = p_0(x)$. La solution d'une telle équation nous est donnée par la formule de Feynman-Kac :

$$p(x, t) = E[p_0(X_T) | X_t = x] \quad (\text{A.2})$$

où X obéit à l'équation différentielle stochastique suivante :

$$dX = \mu(X, t)dt + \sigma(X, t)dB \quad (\text{A.3})$$

avec dB un mouvement brownien et la condition initiale $X(0) = x$

La preuve s'effectue grâce à la formule de Itô (voir la page 44 de [23]) appliquée

à $p(x, t)$. Celle-ci nous donne l'expression :

$$dp = \left(\frac{\partial p(x, t)}{\partial t} + \mu \frac{\partial p(x, t)}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 p(x, t)}{\partial x^2} \right) dt + \sigma(x, t) \frac{\partial p(x, t)}{\partial x} dB$$

Le terme entre parenthèses est nul par hypothèse. En intégrant l'expression ci-dessus nous avons :

$$\int_t^T dp = p(X_T, T) - p(x, t) = \int_t^T \sigma(x, t) \frac{\partial p(x, t)}{\partial x} dB$$

En prenant l'espérance (moyenne) nous avons :

$$p(x, t) = E[p(X_T, T)] - E\left[\int_t^T \sigma(x, t) \frac{\partial p(x, t)}{\partial x} dB\right]$$

L'espérance d'une intégrale de Itô est nulle. Nous avons donc la formule de Feynman-Kac :

$$p(x, t) = E[p(X_T, T)] = E[p_0(x)] = E(p_0(X_T) | X_t = x)$$

Ainsi, résoudre une équation aux dérivées partielles revient à calculer une intégrale stochastique. Cette intégrale peut être approchée en engendrant des trajectoires aléatoires par méthode de Monte Carlo. C'est le principe fondamental de la méthode "Diffusion Monte Carlo" qui assimile l'équation de Schrödinger à une équation de diffusion.

Annexe B

Existence de la décomposition SVD

B.1 Décomposition SVD

Soit une matrice A réelle de taille $n \times m$, A peut se décomposer en ;

$$A = QDT^t$$

- Q est une matrice aux colonnes orthogonales de taille $m \times n$ (i.e, $Q^tQ = Id$)
- T est une matrice orthogonale de taille $n \times n$ (i.e, $TT^t = T^tT = Id$)
- D est une matrice diagonale de tailles $n \times n$ dont les éléments sont positifs ou nuls.

Démonstration : La matrice A^*A est une matrice carrée de taille $m \times m$. Elle est hermitienne. Par le théorème spectral, elle est décomposable en valeurs propres positives ou nulles dans une base orthonormée, i.e. il existe une matrice orthogonale T de taille $m \times m$ telle que :

$$T^*A^*AT = \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \quad (\text{B.1})$$

où Σ est une matrice diagonale de taille $p \times p$ aux coefficients strictement positifs. En posant

$$T = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$$

avec T_1 de taille $p \times m$ et T_2 de taille $m - p \times m$, nous pouvons récrire :

$$\begin{pmatrix} T_1 \\ T_2 \end{pmatrix} A^* A \begin{pmatrix} T_1^* \\ T_2^* \end{pmatrix} = \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}$$

soit :

$$\begin{pmatrix} T_1^* A^* A T_1 & T_1^* A^* A T_2 \\ T_2^* A^* A T_1 & T_2^* A^* A T_2 \end{pmatrix} = \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix}$$

ce qui permet de définir :

$$T_1^* A^* A T_1 = \Sigma A T_2 = 0$$

Posons alors :

$$Q_1 = A T_1 \Sigma^{\frac{1}{2}}$$

Nous avons :

$$Q_1 \Sigma^{\frac{1}{2}} T_1^* = A T_1 \Sigma^{\frac{1}{2}} \Sigma^{\frac{1}{2}} T_1^* = A$$

Pour conclure, nous construisons :

$$Q = (Q_1 Q_2)$$

De telle sorte que Q soit orthogonale

$$D = \begin{pmatrix} \Sigma^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix}$$

de taille $n \times n$ Nous avons alors le résultat voulu :

$$A = Q D T^*$$

En effet

$$A = (Q_1 Q_2) \begin{pmatrix} \Sigma^{\frac{1}{2}} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} T_1^* \\ T_2^* \end{pmatrix} = (Q_1 Q_2) \begin{pmatrix} \Sigma^{\frac{1}{2}} T_1^* \\ 0 \end{pmatrix} = (Q_1 \Sigma^{\frac{1}{2}} T_1^*) = A$$

Cette décomposition s'effectue grâce à l'algorithme de Golub Reinsch. En effet, la preuve que nous avons exhibé précédemment pour démontrer l'existence d'une telle décomposition n'est pas numériquement stable. L'algorithme de Golub et Reinsch décompose la matrice A en valeurs singulières en deux étapes :

- une décomposition de la matrice A en une matrice bidiagonale B par la méthode de Householder
- la décomposition de la matrice B en valeurs singulières par la méthode QR

B.2 De la matrice A à la matrice B bidiagonale : La méthode de Householder

La méthode de Householder repose sur l'utilisation de matrices de Householder P , c'est à dire de la forme :

$$P = Id - 2w.w^t$$

avec w vecteur réel normé ($|w|^2 = 1$), ou pour un vecteur u quelconque :

$$P = Id - \frac{u.u^t}{H}$$

avec

$$H = \frac{1}{2}|u|^2$$

Une matrice de Householder est orthogonale :

$$\begin{aligned} P^2 &= (Id - 2w.w^t).(Id - 2w.w^t) \\ &= Id - 4w.w^t + 4w.(w^t.w).w^t \\ &= Id \end{aligned}$$

Donc $P = P^{-1}$ et puisque $P = P^t$, $P^t = P^{-1}$

En choisissant la matrice de Householder appropriée, nous pouvons en nous donnant une matrice A , annuler tout les éléments d'une colonne en dessous d'un élément situé sur ou sous la diagonale.

Nous allons ici pour l'exemple annuler les éléments de la première colonne de A en dessous du coefficient a_{11} .

Posons pour cela, x le vecteur composé de la première colonne de A et

$$u = x + \operatorname{sgn}(a_{11})|x|e_1$$

où

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

En appliquant P au vecteur x à gauche

$$\begin{aligned} P.x &= x - \frac{u}{H} \cdot (x + \operatorname{sgn}(a_{11})|x|e_1)^t \cdot x \\ &= x - \frac{2u \cdot (|x|^2 + \operatorname{sgn}(a_{11})|x|a_{11})}{2|x|^2 + \operatorname{sgn}(a_{11})2|x|a_{11}} \\ &= x - u \\ &= \operatorname{sgn}(a_{11})|x|e_1 \end{aligned}$$

$\operatorname{sgn}(a_{11})$ est introduit pour stabiliser l'algorithme.

En appliquant P au vecteur x , nous annulons toutes les composantes en dessous du premier élément de x . Nous pouvons faire de même avec les autres colonnes. De plus en posant

$$P = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & P_1 \end{array} \right)$$

où P_1 est une matrice de Householder, nous pouvons annuler tous éléments en dessous d'un éléments strictement en dessous de la diagonale.

De la même manière, en appliquant une matrice de Householder appropriée à droite de la matrice A , nous pouvons annuler tous les éléments d'une ligne situés à droite d'un éléments sur la diagonale (ou strictement à droite).

En appliquant des matrices de Householder appropriées à droite et à gauche de la matrice A , nous pouvons donc obtenir une matrice bidiagonale B .

$$Q^{(n)} \dots Q^{(1)} A S^{(1)} \dots S^{(n)} = B \quad (\text{B.2})$$

En posant

$$Q_A = (Q^{(n)} \dots Q^{(1)})^t \quad (\text{B.3})$$

et

$$S_A = (S^{(n)} \dots S^{(1)}) \quad (\text{B.4})$$

nous avons :

$$A = Q_A B S_A^t \quad (\text{B.5})$$

Nous allons montrer comment passer de la matrice A à la matrice bidiagonale B . Une matrice bidiagonale est une matrice de la forme :

$$B = \begin{pmatrix} b_{11} & b_{12} & 0 & 0 \dots & 0 \\ 0 & b_{22} & b_{23} & 0 \dots & 0 \\ \vdots & 0 & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & & & & \end{pmatrix}$$

B.3 Décomposition de B en SVD par la méthode QR

Nous allons maintenant décomposer la matrice B en valeur singulière grâce à la méthode QR. Les valeurs singulière de B seront celle de A .

En effet en décomposant B en valeur singulière :

$$B = Q_B D S_B^t$$

et en remplaçant dans B.5, nous obtenons :

$$A = Q_A Q_B D S_A S_A^t S_B^t A = Q D S^t \quad (\text{B.6})$$

avec $Q = Q_A Q_B$ et $S = S_B S_A$ la méthode QR consiste à décomposer une matrice réelle M en :

$$M = QR$$

avec Q matrice orthogonale et R matrice triangulaire supérieure. Une telle décomposition s'effectue en appliquant à droite de la matrice M des réductions de Householder.

Pour décomposer en valeur singulière nous utilisons la diagonalisation par la méthode QR. Il s'agit d'une méthode itérative.

- nous appliquons à droite une transformation de Householder afin d'obtenir une matrice triangulaire inférieure.
- Puis nous appliquons à gauche la même transformation. Nous obtenons une matrice quelconque mais dont les valeurs sont les plus concentrées sur la diagonale

Nous appliquons plusieurs fois ce procédé pour converger afin d'obtenir la matrice diagonale D de notre SVD. Les transformations appliquées ont été faites par des matrices orthogonales qui permettent de reconstituer Q_B et S_B

Annexe C

Usage des comatrices pour le calcul des déterminants

Dans tout ce qui suit, les matrices sont carrées de taille n . Soit la matrice A . La comatrice de A permet de calculer très facilement des déterminants d'une matrice obtenu en changeant une ligne ou une colonne de A . En effet, soit ΔA la comatrice de A . Supposons que nous ayons changé la i^{me} colonne de A pour obtenir la matrice B . Nous avons alors :

$$|B| = \sum_{j=1, n} B_{jk}(-1)^{j+k} \Delta_{jk} A \quad (\text{C.1})$$

Le déterminant de B s'obtient alors par simple combinaison linéaire. Lorsque nous avons plusieurs déterminants de ce type à calculer, nous effectuons une fois pour toute le calcul de la comatrice et nous appliquons alors la formule ci-dessus.

La comatrice de A est donnée par la formule.

$$A^{-1} = \frac{\Delta A^t}{|A|} \quad (\text{C.2})$$

L'inverse et le déterminant de A se trouvent par un pivot de Gauss s'effectuant en un nombre d'opérations d'ordre n^3 .

Nous avons pour notre part abandonner cette technique pour calculer des

déterminants car nous ne cherchons pas que des déterminants dont on aurait changé une seule ligne ou une seule colonne mais plusieurs à la fois. Dans ce cas là le calcul des déterminants par la comatrice ne fonctionne plus.

Annexe D

Formules du Jastrow commun et de ses dérivées

Nous présentons ici les formules concernant le Jastrow commun et ses dérivées dont nous avons eu besoin pour l'implémentation des algorithmes QMC et d'optimisation.

D.1 Notations

Nous utiliserons les lettres minuscules i et j pour les indices portant sur les électrons. Nous utiliserons des lettres majuscules I pour labeller les noyaux. r_{ij} est la distance entre deux électrons i et j . r_{iI} est la distance entre l'électron i et le noyau I . Nous décrirons les coordonnées cartésiennes x, y et z de l'électron i par le triplet (x_i^1, x_i^2, x_i^3) . Nous utiliserons la lettre l pour décrire l'indice supérieur : une coordonnée d'un électron sera de la forme : x_i^l . \sum_I décrit la somme portant sur tous les noyaux. \sum_i décrit la somme portant sur tous les électrons. $\sum_{i<j}$ décrit la somme portant sur toutes les paires d'électrons i et j *distinctes*.

Ecrivant le Jastrow commun sous la forme e^J nous présentons maintenant les formules pour :

- le terme J ,

- la dérivée spatiale de J suivant une coordonnée : $\partial_{x_i^l} J$
- la dérivée double spatiale de J suivant une coordonnée : $\partial_{x_i^l}^2 J$
- la dérivée de J suivant un paramètre c_p : $\partial_{c_p} J$
- la dérivée de $\partial_{x_i^l} J$ suivant un paramètre c_p : $\partial_{x_i^l, c_p} J$
- la dérivée de $\partial_{x_i^l}^2 J$ suivant un paramètre c_p : $\partial_{x_i^l, c_p}^2 J$
- la dérivée double de J suivant deux paramètres c_1 et c_2 : $\partial_{c_1, c_2} J$

Ces formules sont présentées successivement pour la forme “simple” du Jastrow puis pour la forme “générale”.

D.2 Forme simple

Soit le terme de Jastrow e^J

D.2.1 Expression du terme J

$$J = \sum_{i < j} \frac{a_{\sigma_{ij}} r_{ij}}{1 + b_{\sigma_{ij}}^2 r_{ij}} - \sum_I \sum_i p_I r_{iI} \quad (\text{D.1})$$

D.2.2 Dérivée spatiale de J suivant une coordonnée

$$\partial_{x_i^l} J = \sum_j \frac{a_{\sigma_{ij}} (x_i^l - x_j^l)}{r_{ij} (1 + b_{\sigma_{12}}^2 r_{ij})^2} - \frac{p_I (x_i^l - x_I^l)}{r_{iI}} \quad (\text{D.2})$$

D.2.3 Dérivée double spatiale de J suivant une coordonnée

$$\partial_{x_i^l}^2 J = \sum_j 2 \frac{a_{\sigma_{ij}}}{r_{ij} (1 + b_{\sigma_{12}}^2 r_{ij})^3} - \frac{2p_I}{r_{iI}} \quad (\text{D.3})$$

D.2.4 Dérivée de J suivant un paramètre c_p

$$\partial_{a_{\sigma_{12}}} J = \sum_{i < j} \frac{r_{ij}}{1 + b_{\sigma_{ij}}^2 r_{ij}} \quad (\text{D.4})$$

$$\partial_{b_{\sigma_{12}}} J = -2a_{\sigma_{12}} b_{\sigma_{12}} \sum_{i < j} \left(\frac{r_{ij}}{1 + b_{\sigma_{ij}}^2 r_{ij}} \right)^2 \quad (\text{D.5})$$

$$\partial_{p_I} J = - \sum_i r_{iI} \quad (\text{D.6})$$

D.2.5 Dérivée de $\partial_{x_i^l} J$ suivant un paramètre c_p

$$\partial_{x_i^l, a_{\sigma_{12}}} J = \sum_j \frac{x_i^l - x_j^l}{r_{ij}(1 + b_{\sigma_{12}}^2 r_{ij})^2} \quad (\text{D.7})$$

$$\partial_{x_i^l, b_{\sigma_{12}}} J = -4a_{\sigma_{12}} b_{\sigma_{12}} \sum_j \frac{x_i^l - x_j^l}{(1 + b_{\sigma_{12}}^2 r_{ij})^3} \quad (\text{D.8})$$

$$\partial_{x_i^l, p_I} J = - \frac{(x_i^l - x_I^l)}{r_{iI}} \quad (\text{D.9})$$

D.2.6 Dérivée de $\partial_{x_i^l}^2 J$ suivant un paramètre c_p

$$\partial_{x_i^l, a_{\sigma_{12}}}^2 J = \sum_j 2 \frac{1}{r_{ij}(1 + b_{\sigma_{12}}^2 r_{ij})^3} \quad (\text{D.10})$$

$$\partial_{x_i^l, b_{\sigma_{12}}}^2 J = \sum_j -12a_{\sigma_{12}} b_{\sigma_{12}} \frac{1}{r_{ij}(1 + b_{\sigma_{12}}^2 r_{ij})^3} \quad (\text{D.11})$$

$$\partial_{x_i^l, p_I}^2 J = \frac{-2}{r_{iI}} \quad (\text{D.12})$$

D.2.7 Dérivée double de J suivant deux paramètres c_1 et c_2

$$\partial_{a_{\sigma_{12}} a_{\sigma_3 \sigma_4}} J = 0 \quad (\text{D.13})$$

$$\forall c_i, \partial_{p_I c_i} J = 0 \quad (\text{D.14})$$

$$si(\sigma_{12}) \neq (\sigma_3, \sigma_4) \partial_{a_{\sigma_{12}}, b_{\sigma_3 \sigma_4}} J = 0 \quad (\text{D.15})$$

$$\partial_{a_{\sigma_{12}}, b_{\sigma_{12}}} J = -2b_{\sigma_{12}} \sum_{i < j} \left(\frac{r_{ij}}{1 + b_{\sigma_{ij}}^2 r_{ij}} \right)^2 \quad (\text{D.16})$$

$$\partial_{b_{\sigma_{12}}, b_{\sigma_{12}}} J = 2 \sum_{i < j} \frac{r_{ij}^2 (3b_{\sigma_{12}}^2 r_{ij} - 1)}{(1 + b_{\sigma_{ij}}^2 r_{ij})^3} \quad (\text{D.17})$$

D.3 Forme ‘‘g n rale’’ pour J

Nous pr sentons ici l’expression de J et de ses d riv es pour un Jastrow ayant la forme ‘‘g n rale’’.

D.3.1 Expression de J

Le terme de Jastrow est de la forme e^J . J a pour expression :

$$J = \sum_I \sum_{i < j} [s_{ij}(X_{ij}) - p_I(X_{iI}) - p_I(X_{jI}) + g_1^I X_{iI}^2 X_{jI}^2 + g_2^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2] \quad (\text{D.18})$$

avec

$$X_{ij} = \frac{r_{ij}}{1 + b_{\sigma_{12}}^2 r_{ij}} \quad (\text{D.19})$$

$$X_{iI} = \frac{r_{iI}}{1 + b_I r_{iI}} \quad (\text{D.20})$$

et

$$s(X) = \frac{1}{N_I} (e_{1\sigma_{12}} X + e_{2\sigma_{12}} X^2 + e_{3\sigma_{12}} X^3 + e_{4\sigma_{12}} X^4) \quad (\text{D.21})$$

$$p_I(X) = \frac{1}{N_e - 1} (f_{1I} X + f_{2I} X^2 + f_{3I} X^3 + f_{4I} X^4) \quad (\text{D.22})$$

o  N_I est le nombre de noyaux, N_e le nombre d’ lectrons. Les coefficients e_k se distinguent suivant que les  lectrons sont de spins oppos s ou non. De m me, pour les coefficients $b_{\sigma_{12}}$.

D.3.2 D riv e spatiale de J suivant une coordonn e

$$\partial_{x_i^l} J = \sum_I \sum_{i < j} [Y_{ij}^l s^{\nabla}(X_{ij}) - Y_{ij}^l p_I^{\nabla}(X_{iI}) + 2Y_{iI}^l g_1^I X_{iI}^2 X_{jI}^2 + 2Y_{ij}^l g_2^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + Y_{iI}^l g_2^I (X_{iI}^2 X_{ij}^2)] \quad (\text{D.23})$$

avec

$$Y_{ij}^l = \frac{\partial X_{ij}}{\partial x_i^l} \frac{1}{X_{ij}} \quad (\text{D.24})$$

et

$$Y_{iI}^l = \frac{\partial X_{iI}}{\partial x_i^l} \frac{1}{X_{iI}} \quad (\text{D.25})$$

et

$$s^\nabla(X) = \frac{1}{N_I} (e_{1\sigma_{12}} X + 2e_{2\sigma_{12}} X^2 + 3e_{3\sigma_{12}} X^3 + 4e_{4\sigma_{12}} X^4) \quad (\text{D.26})$$

$$p_I^\nabla(X) = \frac{1}{N_e - 1} (f_{1I} X + 2f_{2I} X^2 + 3f_{3I} X^3 + 4f_{4I} X^4) \quad (\text{D.27})$$

D.3.3 Dérivée double spatiale de J suivant une coordonnée

(D.28)

$$\begin{aligned} \partial_{x_i^l}^2 J &= \sum_I \sum_{i < j} [Z_{ij}^l s^\Delta(X_{ij}) + (Y_{ij}^l)^2 s_{ij}^\Delta(X_{ij}) - Z_{ij}^l p_I^\Delta(X_{iI}) - (Y_{ij}^l)^2 p_I^\Delta(X_{iI}) \\ &+ 4(Y_{iI}^l)^2 g_1^I X_{iI}^2 X_{jI}^2 + 2Z_{iI}^l g_1^I X_{iI}^2 X_{jI}^2 + 4(Y_{ij}^l)^2 g_2^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 2Z_{ij}^l g_2^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + Z_{iI}^l g_2^I (X_{iI}^2 X_{ij}^2) + (Y_{iI}^l)^2 g_2^I (X_{iI}^2 X_{ij}^2)] \end{aligned} \quad (\text{D.29})$$

avec

$$Z_{ij}^l = \frac{\partial Y_{ij}^l}{\partial x_i^l} \quad (\text{D.30})$$

et

$$Z_{iI}^l = \frac{\partial Y_{iI}^l}{\partial x_i^l} \quad (\text{D.31})$$

et

$$s^\Delta(X) = \frac{1}{N_I} (e_{1\sigma_{12}} X + 4e_{2\sigma_{12}} X^2 + 9e_{3\sigma_{12}} X^3 + 16e_{4\sigma_{12}} X^4) \quad (\text{D.32})$$

$$p_I^\Delta(X) = \frac{1}{N_e - 1} (f_{1I} X + 4f_{2I} X^2 + 9f_{3I} X^3 + 16f_{4I} X^4) \quad (\text{D.33})$$

D.4 Dérivées suivant les paramètres

D.4.1 Dérivée simple de J suivant un paramètre c_p

Pour $k = 1, 4$:

$$J_{e_k} = \frac{1}{N_I} \sum_I \sum_{i < j} [X_{ij}^k] \quad (\text{D.34})$$

Pour $k = 1, 4$:

$$J_{f_{kI}} = \frac{-1}{N_e - 1} \sum_{i < j} [X_{iI}^k] \quad (\text{D.35})$$

$$J_{g_1^I} = \sum_{i < j} [X_{iI}^2 X_{jI}^2] \quad (\text{D.36})$$

$$J_{g_2^I} = \sum_{i < j} [(X_{iI}^2 + X_{jI}^2) X_{ij}^2] \quad (\text{D.37})$$

D.4.2 Dérivée de $\partial_{x_i^l} J$ suivant un paramètre c_p

Posons

$$T_{ij, c_p} = \frac{\partial X_{ij}}{\partial c_p} \frac{1}{X_{ij}} \quad (\text{D.38})$$

et

$$T_{iI, c_p} = \frac{\partial X_{iI}}{\partial c_p} \frac{1}{X_{iI}} \quad (\text{D.39})$$

$$\partial_{x_i^l, e_{1, \sigma_{12}}} J = \sum_I \sum_{j \neq i_0} Y_{iI}^l \frac{1}{N_I} X_{ij} \quad (\text{D.40})$$

$$\partial_{x_i^l, e_{2, \sigma_{12}}} J = \sum_I \sum_{j \neq i_0} Y_{iI}^l \frac{2}{N_I} X_{ij}^2 \quad (\text{D.41})$$

$$\partial_{x_i^l, e_{3, \sigma_{12}}} J = \sum_I \sum_{j \neq i_0} Y_{iI}^l \frac{3}{N_I} X_{ij}^3 \quad (\text{D.42})$$

$$\partial_{x_i^l, e_{4, \sigma_{12}}} J = \sum_I \sum_{j \neq i_0} Y_{iI}^l \frac{4}{N_I} X_{ij}^4 \quad (\text{D.43})$$

$$\partial_{x_i^l, f_{1I}} J = -Y_{iI}^l (X_{iI}) \quad (\text{D.44})$$

$$\partial_{x_i^l, f_{2I}} J = -Y_{iI}^l 2(X_{iI}^2) \quad (\text{D.45})$$

$$\partial_{x_i^l, f_{3I}} J = -Y_{iI}^l 3(X_{iI}^3) \quad (\text{D.46})$$

$$\partial_{x_i^l, f_{4I}} J = \sum_{j \neq i_0} -4Y_{iI}^l (X_{iI}^4) \quad (\text{D.47})$$

$$\partial_{x_i^l, g_{1, \sigma_{12}}^I} J = 2Y_{iI}^l X_{iI}^2 X_{jI}^2 \quad (\text{D.48})$$

$$\partial_{x_i^l, g_{2, \sigma_{12}}^I} J = 2Y_{iI}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 2Y_{ij}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \quad (\text{D.49})$$

$$\begin{aligned}
\partial_{x_i^l, b_{\sigma_{12}}} J &= \sum_I [Y_{ij}^l \frac{1}{N_I} (e_{1, \sigma_{ij}} X_{ij} + 2e_{2, \sigma_{ij}} X_{ij}^2 + 3e_{3, \sigma_{ij}} X_{ij}^3 + 4e_{4, \sigma_{ij}} X_{ij}^4) \\
&+ Y_{ij}^l T_{ij, b_n} \frac{1}{N_I} (e_{1, \sigma_{ij}} X_{ij} + 4e_{2, \sigma_{ij}} X_{ij}^2 + 9e_{3, \sigma_{ij}} X_{ij}^3 + 16e_{4, \sigma_{ij}} X_{ij}^4) \\
&+ 2Y_{ij}^l g_{2, \sigma_i, \sigma_j}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 2T_{ij, b_n} g_{2, \sigma_i, \sigma_j}^I (2Y_{ij}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\
&+ (2Y_{iI}^l X_{iI}^2) X_{ij}^2)] \quad (D.50)
\end{aligned}$$

$$\begin{aligned}
\partial_{x_i^l, b_I} J &= [-Y_{iI}^l p_I \frac{1}{N_e - 1} (f_{1I} X_{iI} + 2f_{2I} X_{iI}^2 + 3f_{3I} X_{iI}^3 + 4f_{4I} X_{iI}^4) \\
&- T_{iI, b_{nI}} Y_{iI}^l p_I \frac{1}{N_e - 1} (f_{1I} X_{iI} + 4f_{2I} X_{iI}^2 + 9f_{3I} X_{iI}^3 + 16f_{4I} X_{iI}^4) + 2(Y_{iI}^l) g_{1, \sigma_i, \sigma_j}^I X_{iI}^2 X_{jI}^2 \\
&+ 2(T_{iI, b_{nI}} + T_{jI, b_{nI}}) g_{1, \sigma_i, \sigma_j}^I 2Y_{iI}^l X_{iI}^2 X_{jI}^2 + g_{2, \sigma_i, \sigma_j}^I (2Y_{iI}^l X_{iI}^2) X_{ij}^2 \\
&+ g_{2, \sigma_i, \sigma_j}^I (4Y_{iI}^l T_{iI, b_{nI}} X_{iI}^2) X_{ij}^2 + g_{2, \sigma_i, \sigma_j}^I (2T_{iI, b_{nI}} X_{iI}^2 + 2T_{jI, b_{nI}} X_{jI}^2) X_{ij}^2 2Y_{ij}^l] \quad (D.51)
\end{aligned}$$

D.4.3 Dérivée de $\partial_{x_i^l}^2 J$ suivant un paramètre c_p

$$\partial_{x_i^l, e_{1, \sigma_{12}}}^2 J = \sum_I \sum_{j \neq i_0} [Z_{iI}^l \frac{1}{N_I} X_{ij} + (Y_{iI}^l)^2 \frac{1}{N_I} X_{ij}] \quad (D.52)$$

$$\partial_{x_i^l, e_{2, \sigma_{12}}}^2 J = \sum_I \sum_{j \neq i_0} [Z_{iI}^l \frac{2}{N_I} X_{ij}^2 + (Y_{iI}^l)^2 \frac{4}{N_I} X_{ij}^2] \quad (D.53)$$

$$\partial_{x_i^l, e_{3, \sigma_{12}}}^2 J = \sum_I \sum_{j \neq i_0} [Z_{iI}^l \frac{3}{N_I} X_{ij}^3 + (Y_{iI}^l)^2 \frac{9}{N_I} X_{ij}^3] \quad (D.54)$$

$$\partial_{x_i^l, e_{4, \sigma_{12}}}^2 J = \sum_I \sum_{j \neq i_0} [Z_{iI}^l \frac{4}{N_I} X_{ij}^4 + (Y_{iI}^l)^2 \frac{16}{N_I} X_{ij}^4] \quad (D.55)$$

$$\partial_{x_i^l, f_{1I}}^2 J = -Z_{iI}^l (X_{iI}) - (Y_{iI}^l)^2 (X_{iI}) \quad (D.56)$$

$$\partial_{x_i^l, f_{2I}}^2 J = -Z_{iI}^l 2(X_{iI}^2) - (Y_{iI}^l 4(X_{iI}^2))^2 \quad (D.57)$$

$$\partial_{x_i^l, f_{3I}}^2 J = -Z_{iI}^l 3(X_{iI}^3) - (Y_{iI}^l 9(X_{iI}^3))^2 \quad (D.58)$$

$$\partial_{x_i^l, f_{4I}}^2 J = -4Z_{iI}^l (X_{iI} - 16(Y_{iI}^l (X_{iI}^4))^2) \quad (D.59)$$

$$\partial_{x_i^l, g_{1, \sigma_{12}}}^2 J = 2Z_{iI}^l X_{iI}^2 X_{jI}^2 + 4(Y_{iI}^l X_{iI}^2)^2 X_{jI}^2 \quad (D.60)$$

$$\begin{aligned}
\partial_{x_i^l, g_{2, \sigma_{12}}}^2 J &= 2Y_{iI}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 4(Y_{iI}^l)^2 (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \quad (D.61) \\
&+ 2Y_{ij}^l X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 2Y_{iI}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 2Y_{ij}^l X_{iI}^2 + X_{jI}^2) X_{ij}^2
\end{aligned}$$

$$\begin{aligned}
\partial_{x_i^l, b_{\sigma_{12}}}^2 J &= \sum_I [Y_{ij}^l \frac{1}{N_I} (e_{1, \sigma_{ij}} X_{ij} + 2e_{2, \sigma_{ij}} X_{ij}^2 + 3e_{3, \sigma_{ij}} X_{ij}^3 + 4e_{4, \sigma_{ij}} X_{ij}^4) \\
&+ Y_{ij}^l T_{ij, b_n} \frac{1}{N_I} (e_{1, \sigma_{ij}} X_{ij} + 4e_{2, \sigma_{ij}} X_{ij}^2 + 9e_{3, \sigma_{ij}} X_{ij}^3 + 16e_{4, \sigma_{ij}} X_{ij}^4) \\
&+ 2Y_{ij}^l g_{2, \sigma_i, \sigma_j}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\
&+ 2T_{ij, b_n} g_{2, \sigma_i, \sigma_j}^I (2Y_{ij}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + (2Y_{iI}^l X_{iI}^2) X_{ij}^2)] \quad (D.62)
\end{aligned}$$

D.4.4 Dérivée double de J suivant deux paramètres c_1 et c_2

Pour deux paramètres c_1 et c_2 parmi $e_1, e_2, e_3, e_4, f_{1I}, f_{2I}, f_{3I}, f_{4I}, g_1^I, g_2^I$ éventuellement identiques on a :

$$J_{c_1, c_2} = 0 \quad (D.63)$$

Posons

$$\partial^2 \tilde{X}_{ij, l} = \frac{\partial \frac{\partial X_{ij}}{\partial c_l} \frac{1}{X_{ij}}}{\partial c_l} \quad (D.64)$$

et

$$\partial^2 \tilde{X}_{iI, l} = \frac{\partial \frac{\partial X_{iI}}{\partial c_l} \frac{1}{X_{iI}}}{\partial c_l} \quad (D.65)$$

$$\partial_{b_{\sigma_{12}}, e_{1, \sigma_{12}}} J = \sum_I \sum_{i < j} T_{ij, b_n} \frac{1}{N_I} X_{ij} \quad (D.66)$$

$$\partial_{b_{\sigma_{12}}, e_{2, \sigma_{12}}} J = \sum_I \sum_{i < j} T_{ij, b_n} \frac{2}{N_I} X_{ij}^2 \quad (D.67)$$

$$\partial_{b_{\sigma_{12}}, e_{3, \sigma_{12}}} J = \sum_I \sum_{i < j} T_{ij, b_n} \frac{3}{N_I} X_{ij}^3 \quad (D.68)$$

$$\partial_{b_{\sigma_{12}}, e_{4, \sigma_{12}}} J = \sum_I \sum_{i < j} T_{ij, b_n} \frac{4}{N_I} X_{ij}^4 \quad (D.69)$$

$$\partial_{b_I, f_{1I}} J = \sum_{i < j} -\frac{1}{N_e - 1} (T_{iI, c_p} X_{iI} + T_{jI, c_p} X_{jI}) \quad (D.70)$$

$$\partial_{b_I, f_{2I}} J = \sum_{i < j} -\frac{2}{N_e - 1} (T_{iI, c_p} X_{iI}^2 + T_{jI, c_p} X_{jI}^2) \quad (D.71)$$

$$\partial_{b_I, f_{3I}} J = \sum_{i < j} -\frac{3}{N_e - 1} (T_{iI, c_p} X_{iI}^3 + T_{jI, c_p} X_{jI}^3) \quad (D.72)$$

$$\partial_{b_I, f_{4I}} J = \sum_{i < j} -\frac{4}{N_e - 1} (T_{iI, c_p} X_{iI}^4 + T_{jI, c_p} X_{jI}^4) \quad (D.73)$$

$$\partial_{b_I, g_{1, \sigma_{12}}^I} J = \sum_{i < j} T_{jI, c_p} X_{iI}^2 X_{jI}^2 \quad (\text{D.74})$$

$$\partial_{b_I, g_{2, \sigma_{12}}^I} J = \sum_{i < j} (2T_{iI, c_p} X_{iI}^2 + 2T_{jI, c_p} X_{jI}^2) X_{ij}^2 \quad (\text{D.75})$$

$$\partial_{b_{\sigma_{12}}, g_{2, \sigma_{12}}^I} J = \sum_I \sum_{i < j} 2T_{ij, c_p} (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \quad (\text{D.76})$$

$$\begin{aligned} \partial_{b_{\sigma_{12}}, b_{\sigma_{12}}} J &= \sum_I \sum_{i < j} [V_{ij, c_p} \frac{1}{N_I} (e_{1, \sigma_{ij}} X_{ij} + 2e_{2, \sigma_{ij}} X_{ij}^2 + 3e_{3, \sigma_{ij}} X_{ij}^3 + 4e_{4, \sigma_{ij}} X_{ij}^4) \\ &+ (T_{ij, b_n})^2 \frac{1}{N_I} (e_{1, \sigma_{ij}} X_{ij} + 4e_{2, \sigma_{ij}} X_{ij}^2 + 9e_{3, \sigma_{ij}} X_{ij}^3 + 16e_{4, \sigma_{ij}} X_{ij}^4) \\ &+ 2V_{ij, c_p} g_{2, \sigma_i, \sigma_j}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 4(T_{ij, b_n})^2 g_{2, \sigma_i, \sigma_j}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2] \end{aligned}$$

$$\begin{aligned} \partial_{b_I, b_I} J &= \sum_{i < j} [-V_{iI, c_p} p_I \frac{1}{N_e - 1} (f_{1I} X_{iI} + 2f_{2I} X_{iI}^2 + 3f_{3I} X_{iI}^3 + 4e_{4I} X_{iI}^4) \\ &- V_{jI, c_p} p_I \frac{1}{N_e - 1} (f_{1I} X_{jI} + 2f_{2I} X_{jI}^2 + 3f_{3I} X_{jI}^3 + 4f_{4I} X_{jI}^4) \\ &- (T_{iI, b_{nI}})^2 p_I \frac{1}{N_e - 1} (f_{1I} X_{iI} + 4f_{2I} X_{iI}^2 + 9f_{3I} X_{iI}^3 + 16e_{4I} X_{iI}^4) \\ &- (T_{jI, b_{nI}})^2 p_I \frac{1}{N_e - 1} (f_{1I} X_{jI} + 4f_{2I} X_{jI}^2 + 9f_{3I} X_{jI}^3 + 16f_{4I} X_{jI}^4) \\ &+ 2(V_{iI, c_p} + V_{jI, c_p}) g_{1, \sigma_i, \sigma_j}^I X_{iI}^2 X_{jI}^2 + (2(T_{iI, b_{nI}} + T_{jI, b_{nI}}))^2 g_{1, \sigma_i, \sigma_j}^I X_{iI}^2 X_{jI}^2 \\ &+ g_{2, \sigma_i, \sigma_j}^I (2V_{iI, c_p} X_{iI}^2 + 2V_{jI, c_p} X_{jI}^2) X_{ij}^2] \\ &+ g_{2, \sigma_i, \sigma_j}^I ((2T_{iI, b_{nI}})^2 X_{iI}^2 + (2T_{jI, b_{nI}})^2 X_{jI}^2) X_{ij}^2 \end{aligned} \quad (\text{D.78})$$

$$\partial_{b_I, b_{\sigma_{12}}} J = \sum_{i < j} [g_{2, \sigma_i, \sigma_j}^I 2T_{ij, b_n} (2T_{iI, b_{nI}} X_{iI}^2 + 2T_{jI, b_{nI}} X_{jI}^2) X_{ij}^2] \quad (\text{D.79})$$

Annexe E

Formules du Jastrow orbitalaire et de ses dérivées

Nous présentons ici les formules concernant le Jastrow orbitalaire et ses dérivées dont nous avons eu besoin pour l'implémentation des méthodes QMC et d'optimisation.

E.1 Notations

Nous reprenons les notations de la partie précédente, à savoir : nous utiliserons les lettres minuscules i et j pour les indices portant les électrons, nous utiliserons des lettres majuscules I pour labeller les noyaux, r_{ij} est la distance entre deux électrons i et j , r_{iI} est la distance entre l'électron i et le noyau I . Nous décrirons les coordonnées cartésiennes x, y et z de l'électron i par le triplet (x_i^1, x_i^2, x_i^3) . Nous utiliserons la lettre l pour décrire l'indice supérieur : une coordonnée d'un électron sera donc de la forme : x_i^l . \sum_I décrit la somme portant sur tous les noyaux. \sum_i décrit la somme portant sur tous les électrons. $\sum_{i<j}$ décrit la somme portant sur toutes les paires d'électrons i et j *distinctes*. $\sum_{j\neq i}$ décrit la somme portant sur tous les électrons j distincts de i . L'indice n désigne le numéro du Jastrow.

Nous considérons le terme de Jastrow orbitalaire e^J . Nous présentons

dans cette partie les formules pour :

- le terme J ,
- la dérivée spatiale de J suivant une coordonnée : $\partial_{x_i^l} J$
- la dérivée double spatiale de J suivant une coordonnée : $\partial_{x_i^l}^2 J$
- la dérivée de J suivant un paramètre c_p : $\partial_{c_p} J$
- la dérivée de $\partial_{x_i^l} J$ suivant un paramètre c_p : $\partial_{x_i^l, c_p} J$
- la dérivée de $\partial_{x_i^l}^2 J$ suivant un paramètre c_p : $\partial_{x_i^l, c_p}^2 J$
- la dérivée double de J suivant deux paramètres c_1 et c_2 : $\partial_{c_1, c_2} J$

Ces formules sont successivement présentées pour la forme “simple” du Jastrow puis pour la forme “générale”.

E.2 Forme simple

Soit le terme de Jastrow e^J

E.2.1 Expression du terme J

$$J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} \frac{a_n r_{ij}}{1 + b_n^2 r_{ij}} - \sum_I p_{nI} r_{iI} \quad (\text{E.1})$$

E.2.2 Dérivée spatiale de J suivant une coordonnée

Il y a désormais deux formules différentes

- la dérivée s’effectue suivant l’électron \vec{r}_i , premier argument dans la fonction J

$$\partial_{x_i^l} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} \frac{a_n (x_i^l - x_j^l)}{r_{ij} (1 + b_n^2 r_{ij})^2} - \sum_I p_{nI} \frac{(x_i^l - x_I^l)}{r_{iI}} \quad (\text{E.2})$$

- la dérivée s’effectue suivant un des électrons \vec{r}_j

$$\partial_{x_j^l} J_n(\vec{r}_i | \dots) = \frac{1}{2} \frac{a_n (x_j^l - x_i^l)}{r_{ij} (1 + b_n^2 r_{ij})^2} \quad (\text{E.3})$$

E.2.3 Dérivée double spatiale de J suivant une coordonnée

Comme précédemment il y a deux formules différentes :

- la dérivée s'effectue suivant l'électron \vec{r}_i , premier argument dans la fonction J

$$\partial_{x_i}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} \left[\frac{2a_n}{r_{ij}(1 + b_n^2 r_{ij})^3} \right] - \frac{2p_{nI}}{r_{iI}} \quad (\text{E.4})$$

- la dérivée s'effectue suivant un des électrons \vec{r}_j

$$\partial_{x_j}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \frac{2a_n}{r_{ij}(1 + b_n^2 r_{ij})^3} \quad (\text{E.5})$$

E.2.4 Dérivée de J suivant un paramètre c_p

$$\partial_{a_n} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} \frac{r_{ij}}{1 + b_n^2 r_{ij}} \quad (\text{E.6})$$

$$\partial_{b_n} J_n(\vec{r}_i | \dots) = -\frac{1}{2} 2a_n b_n \sum_{j \neq i} \left(\frac{r_{ij}}{1 + b_n^2 r_{ij}} \right)^2 \quad (\text{E.7})$$

$$\partial_{p_{nI}} J_n(\vec{r}_i | \dots) = -r_{iI} \quad (\text{E.8})$$

E.2.5 Dérivée de $\partial_{x_i} J$ suivant un paramètre c_p

Il y a deux formules différentes

- la dérivée s'effectue suivant l'électron \vec{r}_i , premier argument dans la fonction J

$$\partial_{x_i^l, a_n} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} \frac{x_i - x_j}{r_{ij}(1 + b_n^2 r_{ij})^2} \quad (\text{E.9})$$

$$\partial_{x_i^l, b_n} J_n(\vec{r}_i | \dots) = -4a_n b_n \frac{1}{2} \sum_{j \neq i} \frac{x_i - x_j}{(1 + b_n^2 r_{ij})^3} \quad (\text{E.10})$$

$$\partial_{x_i^l, p_{nI}} J_n(\vec{r}_i | \dots) = -\frac{(x_i - x_j)}{r_{ij}} \quad (\text{E.11})$$

- la dérivée s'effectue suivant un des électrons \vec{r}_j

$$\partial_{x_j^l, a_n} J_n(\vec{r}_i | \dots) = \frac{1}{2} \frac{x_j - x_i}{r_{ij}(1 + b_n^2 r_{ij})^2} \quad (\text{E.12})$$

$$\partial_{x_j^l, b_n} J_n(\vec{r}_i | \dots) = -4a_n b_n \frac{1}{2} \frac{x_j - x_i}{(1 + b_n^2 r_{ij})^3} \quad (\text{E.13})$$

$$\partial_{x_j^l, p_I} J_n(\vec{r}_i | \dots) = 0 \quad (\text{E.14})$$

E.2.6 Dérivée simple du Laplacien du jastrow suivant les paramètres

Il y a deux formules différentes

- la dérivée s'effectue suivant l'électron \vec{r}_i , premier argument dans la fonction J

$$\partial_{x_j^l, a_n}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} 2 \frac{1}{r_{ij}(1 + b_n^2 r_{ij})^3} \quad (\text{E.15})$$

$$\partial_{x_j^l, b_n}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} -12a_n b_n \frac{1}{r_{ij}(1 + b_n^2 r_{ij})^3} \quad (\text{E.16})$$

$$\partial_{x_j^l, p_{nI}}^2 J_n(\vec{r}_i | \dots) = \frac{-2}{r_{iI}} \quad (\text{E.17})$$

- la dérivée s'effectue suivant un des électrons \vec{r}_j

$$\partial_{x_j^l, a_n}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \frac{2}{r_{ij}(1 + b_n^2 r_{ij})^3} \quad (\text{E.18})$$

$$\partial_{x_j^l, b_n}^2 J_n(\vec{r}_i | \dots) = -\frac{1}{2} 12a_n b_n \frac{1}{r_{ij}(1 + b_n^2 r_{ij})^3} \quad (\text{E.19})$$

$$\partial_{x_j^l, p_{nI}}^2 J_n(\vec{r}_i | \dots) = \frac{-2}{r_{iI}} \quad (\text{E.20})$$

E.2.7 Dérivée de J suivant les paramètres c_1 et c_2

$$\partial_{a_n, a_n} J_n(\vec{r}_i | \dots) = 0 \quad (\text{E.21})$$

$$\forall c_i \partial_{p_I, c_i} J_n(\vec{r}_i | \dots) = 0 \quad (\text{E.22})$$

$$\partial_{a_n, b_n} J_n(\vec{r}_i | \dots) = -2b_n \frac{1}{2} \sum_{j \neq i} \left(\frac{r_{ij}}{1 + b_n^2 r_{ij}} \right)^2 \quad (\text{E.23})$$

$$\partial_{b_n, b_n} J_n(\vec{r}_i | \dots) = 2 \frac{1}{2} \sum_{j \neq i} \frac{r_{ij}^2 (3b_n^2 r_{ij} - 1)}{(1 + b_n r_{ij})^3} \quad (\text{E.24})$$

E.3 Forme “générale”

Nous présentons ici l’expression de J et de ses dérivées pour un Jastrow à la forme “générale”.

E.3.1 Expression de J

Le terme de Jastrow est de la forme e^J . J a pour expression : Soit le terme de jastrow $e^{J_n(\vec{r}_i | \dots)}$ avec :

$$J_n(\vec{r}_i, \dots) = \frac{1}{2} \sum_{j \neq i} [s_n(X_{ij})] - \sum_I [p_{n,I}(X_{iI})] + \frac{1}{2} \sum_I \sum_{j \neq i} [g_{1_n}^I X_{iI}^2 X_{jI}^2 + g_{2_n}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2] \quad (\text{E.25})$$

avec

$$X_{ij} = \frac{r_{ij}}{1 + b_n r_{ij}} \quad (\text{E.26})$$

$$X_{iI} = \frac{r_{iI}}{1 + b_{nI} r_{iI}} \quad (\text{E.27})$$

et

$$s_n(X) = (e_{1_n} X + e_{2_n} X^2 + e_{3_n} X^3 + e_{4_n} X^4) \quad (\text{E.28})$$

$$p_{n,I}(X) = (f_{1I_n} X + f_{2I_n} X^2 + f_{3I_n} X^3 + f_{4I_n} X^4) \quad (\text{E.29})$$

E.3.2 Dérivée spatiale de J suivant une coordonnée

- la dérivée s'effectue suivant l'électron \vec{r}_i , premier argument dans la fonction J

$$\begin{aligned} \partial_{x_i^l} J_n(\vec{r}_i | \dots) &= \frac{1}{2} \sum_{j \neq i} [Y_{ij}^l s_n^\nabla(X_{ij})] - \sum_I [Y_{ij}^l p_{n,I}^\nabla(X_{iI})] \\ &+ \frac{1}{2} \sum_I \sum_{j \neq i} [2Y_{iI}^l g_{1n}^I X_{iI}^2 X_{jI}^2 + 2Y_{ij}^l g_{2n}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + Y_{iI}^l g_{2n}^I (X_{iI}^2 X_{ij}^2)] \end{aligned} \quad (\text{E.30})$$

- la dérivée s'effectue suivant un des électrons \vec{r}_j

avec

$$Y_{ij}^l = \frac{\partial X_{ij}}{\partial x_i} \frac{1}{X_{ij}} \quad (\text{E.31})$$

et

$$Y_{iI}^l = \frac{\partial X_{iI}}{\partial x_i} \frac{1}{X_{iI}} \quad (\text{E.32})$$

et

$$s_n^\nabla(X) = e_{1n} X + 2e_{2n} X^2 + 3e_{3n} X^3 + 4e_{4n} X^4 \quad (\text{E.33})$$

$$p_{n,I}^\nabla(X) = f_{1In} X + 2f_{2In} X^2 + 3f_{3In} X^3 + 4f_{4In} X^4 \quad (\text{E.34})$$

E.3.3 Dérivée double spatiale de J suivant une coordonnée

$$\begin{aligned} \partial_{x_i^l}^2 J_n(\vec{r}_i | \dots) &= \frac{1}{2} \sum_{j \neq i} [Z_{ij}^l s_n^\nabla(X_{ij}) + (Y_{ij}^l)^2 s_n^\Delta(X_{ij})] - \sum_I [Z_{ij}^l p_{n,I}^\nabla(X_{iI}) - (Y_{ij}^l)^2 p_I^\Delta(X_{iI})] \\ &+ \frac{1}{2} \sum_I \sum_{j \neq i} [4(Y_{iI}^l)^2 g_{1n}^I X_{iI}^2 X_{jI}^2 + 2Z_{iI}^l g_{1n}^I X_{iI}^2 X_{jI}^2 + 4(Y_{ij}^l)^2 g_{2n}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2] \\ &+ 2Z_{ij}^l g_{2n}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + Z_{iI}^l g_{2n}^I (X_{iI}^2 X_{ij}^2) + (Y_{iI}^l)^2 g_{2n}^I (X_{iI}^2 X_{ij}^2) \end{aligned} \quad (\text{E.35})$$

avec

$$Z_{ij}^l = \frac{\partial Y_{ij}^l}{\partial x_i} \quad (\text{E.36})$$

et

$$Z_{iI}^l = \frac{\partial Y_{iI}^l}{\partial x_i} \quad (\text{E.37})$$

et

$$s_n^\Delta(X) = (e_{1_n}X + 4e_{2_n}X^2 + 9e_{3_n}X^3 + 16e_{4_n}X^4) \quad (\text{E.38})$$

$$p_{n,I}^\Delta(X) = f_{1_{nI}}X + 4f_{2_{nI}}X^2 + 9f_{3_{nI}}X^3 + 16f_{4_{nI}}X^4 \quad (\text{E.39})$$

E.3.4 Dérivée de J suivant un paramètre c_p

Pour $k=1,4$

$$\partial_{e_{k_n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} [X_{ij}]^k \quad (\text{E.40})$$

Pour $k=1,4$

$$\partial_{f_{k_n}} J_n(\vec{r}_i | \dots) = - \sum_I [X_{iI}]^k \quad (\text{E.41})$$

$$\partial_{g_{1_n}^I} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_I \sum_{j \neq i} [X_{iI}^2 X_{jI}^2] \quad (\text{E.42})$$

$$\partial_{g_{2_n}^I} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_I \sum_{j \neq i} [(X_{iI}^2 + X_{jI}^2) X_{ij}^2] \quad (\text{E.43})$$

Posons

$$T_{ij,c_p} = \frac{\partial X_{ij}}{\partial c_p} \frac{1}{X_{ij}}$$

$$T_{iI,c_p} = \frac{\partial X_{iI}}{\partial c_p} \frac{1}{X_{iI}}$$

$$\begin{aligned} \partial_{b_n} J_n(\vec{r}_i | \dots) &= \frac{1}{2} \sum_{j \neq i} [T_{ij,b_n} s_n^\nabla(X_{ij})] \\ &+ \frac{1}{2} \sum_I \sum_{j \neq i} [2T_{ij,b_n} g_{2_n}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2] \end{aligned} \quad (\text{E.44})$$

$$\begin{aligned} \partial_{b_{nI}} J_n(\vec{r}_i | \dots) &= - \sum_I [T_{ij} p_{n,I}^\nabla(X_{iI})] \\ &+ \frac{1}{2} \sum_I \sum_{j \neq i} [2(T_{iI,b_{nI}} + T_{jI,b_{nI}}) g_{1_n}^I X_{iI}^2 X_{jI}^2 + 2T_{iI,b_{nI}} g_{2_n}^I (X_{iI}^2 X_{ij}^2 + T_{jI,b_{nI}} g_{2_n}^I (X_{jI}^2 X_{ij}^2)] \end{aligned} \quad (\text{E.45})$$

E.3.5 Dérivée de $\partial_{x_i^l} J$ suivant un paramètre c_p

Posons

$$Y_{ij}^l = \frac{\partial X_{ij}}{\partial x_i^l} \frac{1}{X_{ij}} \quad (\text{E.46})$$

et

$$Y_{iI}^l = \frac{\partial X_{iI}}{\partial x_i^l} \frac{1}{X_{iI}} \quad (\text{E.47})$$

Posons

$$\nabla \tilde{X}_{ij,l} = \frac{\partial X_{ij}}{\partial x_i} \frac{1}{X_{ij}} \quad (\text{E.48})$$

et

$$\nabla \tilde{X}_{iI,l} = \frac{\partial X_{iI}}{\partial x_i} \frac{1}{X_{iI}} \quad (\text{E.49})$$

- la dérivée s'effectue suivant l'électron \vec{r}_i , premier argument dans la fonction J_n

$$\partial_{x_i^l, e_{1n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} Y_{ij}^l X_{ij} \quad (\text{E.50})$$

$$\partial_{x_i^l, e_{2n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} Y_{ij}^l 2X_{ij}^2 \quad (\text{E.51})$$

$$\partial_{x_i^l, e_{3n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} Y_{ij}^l 3X_{ij}^3 \quad (\text{E.52})$$

$$\partial_{x_i^l, e_{4n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} Y_{ij}^l 4X_{ij}^4 \quad (\text{E.53})$$

$$\partial_{x_i^l, f_{1In}} J_n(\vec{r}_i | \dots) = -Y_{iI}^l (X_{iI}) \quad (\text{E.54})$$

$$\partial_{x_i^l, f_{2In}} J_n(\vec{r}_i | \dots) = -Y_{iI}^l 2(X_{iI}^2) \quad (\text{E.55})$$

$$\partial_{x_i^l, f_{3In}} J_n(\vec{r}_i | \dots) = -Y_{iI}^l 3(X_{iI}^3) \quad (\text{E.56})$$

$$\partial_{x_i^l, f_{4In}} J_n(\vec{r}_i | \dots) = -4Y_{iI}^l (X_{iI}^4) \quad (\text{E.57})$$

$$\partial_{x_i^l, g_{1n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} 2Y_{iI}^l X_{iI}^2 X_{jI}^2 \quad (\text{E.58})$$

$$\partial_{x_i^l, g_{2n}^I} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} 2Y_{iI}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 2Y_{ij}^l X_{iI}^2 + X_{jI}^2) X_{ij}^2 \quad (\text{E.59})$$

$$\begin{aligned} \partial_{x_i^l, b_n} J_n(\vec{r}_i | \dots) &= \sum_{j \neq i} [\nabla \tilde{X}_{ij,l} (e_{1n} X_{ij} + 2e_{2n} X_{ij}^2 + 3e_{3n} X_{ij}^3 + 4e_{4n} X_{ij}^4) \\ &+ Y_{ij}^l T_{ij,b_n} (e_{1n} X_{ij} + 4e_{2n} X_{ij}^2 + 9e_{3n} X_{ij}^3 + 16e_{4n} X_{ij}^4) \\ &+ 2\nabla \tilde{X}_{ij,l} g_{2n}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 2T_{ij,b_n} g_{2n}^I (2Y_{ij}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + (2Y_{iI}^l X_{iI}^2) X_{ij}^2)] \end{aligned} \quad (\text{E.60})$$

$$\begin{aligned} \partial_{x_i^l, b_{nI}} J_n(\vec{r}_i | \dots) &= [-\nabla \tilde{X}_{iI,l} p_I (f_{1I_n} X_{iI} + 2f_{2I_n} X_{iI}^2 + 3f_{3I_n} X_{iI}^3 + 4f_{4I_n} X_{iI}^4) \\ &- T_{iI,b_{nI}} \nabla \tilde{X}_{iI,l} p_I (f_{1I_n} X_{iI} + 4f_{2I_n} X_{iI}^2 + 9f_{3I_n} X_{iI}^3 + 16f_{4I_n} X_{iI}^4) \\ &+ 2(\nabla \tilde{X}_{iI,l}) g_{1n}^I X_{iI}^2 X_{jI}^2 \\ &+ 2(T_{iI,b_{nI}} + T_{jI,b_{nI}}) g_{1n}^I 2Y_{iI}^l X_{iI}^2 X_{jI}^2 \\ &+ \frac{1}{2} \sum_{j \neq i} [g_{2n}^I (2\nabla \tilde{X}_{iI,l} X_{iI}^2) X_{ij}^2 \\ &+ g_{2n}^I (4\nabla \tilde{X}_{iI,l} T_{iI,b_{nI}} X_{iI}^2) X_{ij}^2 + g_{2n}^I (2T_{iI,b_{nI}} X_{iI}^2 \\ &+ 2T_{jI,b_{nI}} X_{jI}^2) X_{ij}^2 2\nabla \tilde{X}_{ij,l}] \end{aligned} \quad (\text{E.61})$$

- la dérivée s'effectue suivant un des électrons \vec{r}_j

$$\partial_{x_j^l, e_{1n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} Y_{iI}^l X_{ij} \quad (\text{E.62})$$

$$\partial_{x_j^l, e_{2n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} Y_{iI}^l 2X_{ij}^2 \quad (\text{E.63})$$

$$\partial_{x_j^l, e_{3n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} Y_{iI}^l 3X_{ij}^3 \quad (\text{E.64})$$

$$\partial_{x_j^l, e_{4n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} Y_{iI}^l 4X_{ij}^4 \quad (\text{E.65})$$

$$\partial_{x_j^l, g_{1n}^I} J_n(\vec{r}_i | \dots) = \frac{1}{2} 2Y_{iI}^l X_{iI}^2 X_{jI}^2 \quad (\text{E.66})$$

$$\partial_{x_j^l, g_{2n}^I} J_n(\vec{r}_i | \dots) = \frac{1}{2} 2Y_{iI}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 2Y_{ij}^l X_{iI}^2 + X_{jI}^2) X_{ij}^2 \quad (\text{E.67})$$

$$\begin{aligned} \partial_{x_j^l, b_n} J_n(\vec{r}_i | \dots) &= \frac{1}{2} [\nabla \tilde{X}_{ij,l} (e_{1n} X_{ij} + 2e_{2n} X_{ij}^2 + 3e_{3n} X_{ij}^3 + 4e_{4n} X_{ij}^4) \\ &+ Y_{ij}^l T_{ij,b_n} (e_{1n} X_{ij} + 4e_{2n} X_{ij}^2 + 9e_{3n} X_{ij}^3 + 16e_{4n} X_{ij}^4) \\ &+ \frac{1}{2} [2\nabla \tilde{X}_{ij,l} g_{2n}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 2T_{ij,b_n} g_{2n}^I (2Y_{ij}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + (2Y_{iI}^l X_{iI}^2) X_{ij}^2)] \end{aligned} \quad (\text{E.68})$$

E.3.6 Dérivée de $\partial_{x_i}^2 J$ suivant un paramètre c_p

- la dérivée s'effectue suivant l'électron \vec{r}_i , premier argument dans la fonction J

$$\partial_{x_i^l, e_{1n}}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} [\nabla^2 \tilde{X}_{iI} X_{ij} + (Y_{iI}^l)^2 X_{ij}] \quad (\text{E.69})$$

$$\partial_{x_i^l, e_{2n}}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} [\nabla^2 \tilde{X}_{iI} 2X_{ij}^2 + (Y_{iI}^l)^2 4X_{ij}^2] \quad (\text{E.70})$$

$$\partial_{x_i^l, e_{3n}}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} [\nabla^2 \tilde{X}_{iI} 3X_{ij}^3 + (Y_{iI}^l)^2 9X_{ij}^3] \quad (\text{E.71})$$

$$\partial_{x_i^l, e_{4n}}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} [\nabla^2 \tilde{X}_{iI} 4X_{ij}^4 + (Y_{iI}^l)^2 16X_{ij}^4] \quad (\text{E.72})$$

$$\partial_{x_i^l, f_{1In}}^2 J_n(\vec{r}_i | \dots) = -\nabla^2 \tilde{X}_{iI} (X_{iI}) - (Y_{iI}^l)^2 (X_{iI}) \quad (\text{E.73})$$

$$\partial_{x_i^l, f_{2In}}^2 J_n(\vec{r}_i | \dots) = -\nabla^2 \tilde{X}_{iI} 2(X_{iI}^2) - (Y_{iI}^l)^2 4(X_{iI}^2)^2 \quad (\text{E.74})$$

$$\partial_{x_i^l, f_{3In}}^2 J_n(\vec{r}_i | \dots) = -\nabla^2 \tilde{X}_{iI} 3(X_{iI}^3) - (Y_{iI}^l)^2 9(X_{iI}^3)^2 \quad (\text{E.75})$$

$$\partial_{x_i^l, f_{4In}}^2 J_n(\vec{r}_i | \dots) = -4\nabla^2 \tilde{X}_{iI} (X_{iI} - 16(Y_{iI}^l)^2 (X_{iI}^4)^2) \quad (\text{E.76})$$

$$\partial_{x_i^l, g_{1n}^I}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} 2\nabla^2 \tilde{X}_{iI} X_{iI}^2 X_{jI}^2 + 4(Y_{iI}^l X_{iI}^2)^2 X_{jI}^2 \quad (\text{E.77})$$

$$\begin{aligned} \partial_{x_i^l, g_{2n}^I}^2 J_n(\vec{r}_i | \dots) &= \frac{1}{2} \sum_{j \neq i} [2Y_{iI}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 4(Y_{iI}^l)^2 (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 2Y_{ij}^l X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 2Y_{iI}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 2Y_{ij}^l X_{iI}^2 + X_{jI}^2) X_{ij}^2] \end{aligned} \quad (\text{E.78})$$

$$\begin{aligned} \partial_{x_i^l, b_n}^2 J_n(\vec{r}_i | \dots) &= \frac{1}{2} \sum_{j \neq i} [\nabla \tilde{X}_{ij, l} (e_{1, \sigma_k, \sigma_l} X_{ij} + 2e_{2, \sigma_k, \sigma_l} X_{ij}^2 + 3e_{3, \sigma_k, \sigma_l} X_{ij}^3 + 4e_{4, \sigma_k, \sigma_l} X_{ij}^4) \\ &+ Y_{ij}^l T_{ij, b_n} (e_{1, \sigma_k, \sigma_l} X_{ij} + 4e_{2, \sigma_k, \sigma_l} X_{ij}^2 + 9e_{3, \sigma_k, \sigma_l} X_{ij}^3 + 16e_{4, \sigma_k, \sigma_l} X_{ij}^4) \\ &+ 2\nabla \tilde{X}_{ij, l} g_{2, \sigma_i, \sigma_j}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 2T_{ij, b_n} g_{2, \sigma_i, \sigma_j}^I (2Y_{ij}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + (2Y_{iI}^l X_{iI}^2) X_{ij}^2)] \end{aligned} \quad (\text{E.79})$$

- la dérivée s'effectue suivant un des électrons \vec{r}_j

$$\partial_{x_i^l, e_{1n}}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} [\nabla^2 \tilde{X}_{iI} X_{ij} + (Y_{iI}^l)^2 X_{ij}] \quad (\text{E.80})$$

$$\partial_{x_j^l, e_{2n}}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} [\nabla^2 \tilde{X}_{iI} 2X_{ij}^2 + (Y_{iI}^l)^2 4X_{ij}^2] \quad (\text{E.81})$$

$$\partial_{x_j^l, e_{3n}}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} [\nabla^2 \tilde{X}_{iI} 3X_{ij}^3 + (Y_{iI}^l)^2 9X_{ij}^3] \quad (\text{E.82})$$

$$\partial_{x_j^l, e_{4n}}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} [\nabla^2 \tilde{X}_{iI} 4X_{ij}^4 + (Y_{iI}^l)^2 16X_{ij}^4] \quad (\text{E.83})$$

$$\partial_{x_j^l, g_{1n}^I}^2 J_n(\vec{r}_i | \dots) = \frac{1}{2} 2\nabla^2 \tilde{X}_{iI} X_{iI}^2 X_{jI}^2 + 4(Y_{iI}^l X_{iI}^2)^2 X_{jI}^2 \quad (\text{E.84})$$

$$\begin{aligned} \partial_{x_j^l, g_{2n}^I}^2 J_n(\vec{r}_i | \dots) &= \frac{1}{2} [2Y_{iI}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 4(Y_{iI}^l)^2 (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 2Y_{ij}^l X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 2Y_{iI}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + 2Y_{ij}^l X_{iI}^2 + X_{jI}^2) X_{ij}^2] \end{aligned} \quad (\text{E.85})$$

$$\begin{aligned} \partial_{x_j^l, b_n}^2 J_n(\vec{r}_i | \dots) &= \frac{1}{2} [\nabla \tilde{X}_{ij,l} (e_{1,\sigma_k,\sigma_l} X_{ij} + 2e_{2,\sigma_k,\sigma_l} X_{ij}^2 + 3e_{3,\sigma_k,\sigma_l} X_{ij}^3 + 4e_{4,\sigma_k,\sigma_l} X_{ij}^4) \\ &+ Y_{ij}^l T_{ij,b_n} (e_{1,\sigma_k,\sigma_l} X_{ij} + 4e_{2,\sigma_k,\sigma_l} X_{ij}^2 + 9e_{3,\sigma_k,\sigma_l} X_{ij}^3 + 16e_{4,\sigma_k,\sigma_l} X_{ij}^4) \\ &+ 2\nabla \tilde{X}_{ij,l} g_{2,\sigma_i,\sigma_j}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 2T_{ij,b_n} g_{2,\sigma_i,\sigma_j}^I (2Y_{ij}^l (X_{iI}^2 + X_{jI}^2) X_{ij}^2 + (2Y_{iI}^l X_{iI}^2) X_{ij}^2)] \end{aligned} \quad (\text{E.86})$$

E.3.7 Dérivée double de J suivant deux paramètres c_1 et c_2

Nous choisissons deux paramètres c_1 et c_2 parmi $e_{1n}, e_{2n}, e_{3n}, e_{4n}, f_{1n}, f_{2n}, f_{3n}, f_{4n}, g_{1n}^I, g_{2n}^I$ éventuellement identiques.

Lorsque c_1 est un paramètre appartenant à un Jastrow n et c_2 à un Jastrow distinct m on a :

$$\partial_{c_1, c_2} J_n(\vec{r}_i | \dots) = 0 \quad (\text{E.87})$$

Posons

$$V_{ij, c_p} = \frac{\partial T_{ij, c_p}}{\partial c_p} \quad (\text{E.88})$$

et

$$V_{iI, c_p} = \frac{\partial T_{iI, c_p}}{\partial c_p} \quad (\text{E.89})$$

$$\partial_{b_n, e_{1n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} T_{ij, b_n} X_{ij} \quad (\text{E.90})$$

$$\partial_{b_n, e_{2n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} T_{ij, b_n} 2X_{ij}^2 \quad (\text{E.91})$$

$$\partial_{b_n, e_{3n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} T_{ij, b_n} 3X_{ij}^3 \quad (\text{E.92})$$

$$\partial_{b_n, e_{4n}} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} T_{ij, b_n} 4X_{ij}^4 \quad (\text{E.93})$$

$$\partial_{b_{nI}, f_{1I}} J_n(\vec{r}_i | \dots) = -1(\partial \tilde{X}_{iI, l} X_{iI} + \partial \tilde{X}_{jI, l} X_{jI}) \quad (\text{E.94})$$

$$\partial_{b_{nI}, f_{2I}} J_n(\vec{r}_i | \dots) = -2(\partial \tilde{X}_{iI, l} X_{iI}^2 + \partial \tilde{X}_{jI, l} X_{jI}^2) \quad (\text{E.95})$$

$$\partial_{b_{nI}, f_{3I}} J_n(\vec{r}_i | \dots) = -3(\partial \tilde{X}_{iI, l} X_{iI}^3 + \partial \tilde{X}_{jI, l} X_{jI}^3) \quad (\text{E.96})$$

$$\partial_{b_{nI}, f_{4I}} J_n(\vec{r}_i | \dots) = -4(\partial \tilde{X}_{iI, l} X_{iI}^4 + \partial \tilde{X}_{jI, l} X_{jI}^4) \quad (\text{E.97})$$

$$\partial_{b_{nI}, g_{1n}^I} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} \partial \tilde{X}_{jI, l} X_{iI}^2 X_{jI}^2 \quad (\text{E.98})$$

$$\partial_{b_{nI}, g_{2n}^I} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} (2\partial \tilde{X}_{iI, l} X_{iI}^2 + 2\partial \tilde{X}_{jI, l} X_{jI}^2) X_{ij}^2 \quad (\text{E.99})$$

$$\partial_{b_n, g_{2n}^I} J_n(\vec{r}_i | \dots) = \frac{1}{2} \sum_{j \neq i} 2\partial \tilde{X}_{ij, l} (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \quad (\text{E.100})$$

$$\begin{aligned} \partial_{b_n, b_n} J_n(\vec{r}_i | \dots) &= \frac{1}{2} \sum_I \sum_{j \neq i} [V_{ij, b_{nI}} \frac{1}{N_I} (e_{1n} X_{ij} + 2e_{2n} X_{ij}^2 + 3e_{3n} X_{ij}^3 + 4e_{4n} X_{ij}^4) \\ &+ (T_{ij, b_n})^2 \frac{1}{N_I} (e_{1n} X_{ij} + 4e_{2n} X_{ij}^2 + 9e_{3n} X_{ij}^3 + 16e_{4n} X_{ij}^4) \\ &+ 2V_{ij, b_{nI}} g_{2n}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2 \\ &+ 4(T_{ij, b_n})^2 g_{2n}^I (X_{iI}^2 + X_{jI}^2) X_{ij}^2] \end{aligned} \quad (\text{E.101})$$

$$\begin{aligned}
\partial_{b_{nI}, b_{nI}} J_n(\vec{r}_i | \dots) &= \sum_I [-V_{iI, b_{nI}} p_I (f_{1I} X_{iI} + 2f_{2I} X_{iI}^2 + 3f_{3I} X_{iI}^3 + 4e_{4I} X_{iI}^4) \\
&- V_{jI, b_{nI}} p_I (f_{1I} X_{jI} + 2f_{2I} X_{jI}^2 + 3f_{3I} X_{jI}^3 + 4f_{4I} X_{jI}^4) \\
&- (T_{iI, b_{nI}})^2 p_I (f_{1I} X_{iI} + 4f_{2I} X_{iI}^2 + 9f_{3I} X_{iI}^3 + 16e_{4I} X_{iI}^4) \\
&- (T_{jI, b_{nI}})^2 p_I (f_{1I} X_{jI} + 4f_{2I} X_{jI}^2 + 9f_{3I} X_{jI}^3 + 16f_{4I} X_{jI}^4) \\
&+ 2(V_{iI, b_{nI}} + V_{jI, b_{nI}}) g_{1n}^I X_{iI}^2 X_{jI}^2 \\
&+ (2(T_{iI, b_{nI}} + T_{jI, b_{nI}}))^2 g_{1n}^I X_{iI}^2 X_{jI}^2 \\
&+ g_{2n}^I (2V_{iI, b_{nI}} X_{iI}^2 + 2V_{jI, b_{nI}} X_{jI}^2) X_{ij}^2] \\
&+ g_{2n}^I ((2T_{iI, b_{nI}})^2 X_{iI}^2 + (2T_{jI, b_{nI}})^2 X_{jI}^2) X_{ij}^2] \quad (\text{E.102})
\end{aligned}$$

$$\begin{aligned}
\partial_{b_{nI}, b_{nI}} J_n(\vec{r}_i | \dots) &= \frac{1}{2} \sum_I \sum_{j \neq i} [g_{2n}^I 2T_{ij, b_n} (2T_{iI, b_{nI}} X_{iI}^2 + 2T_{jI, b_{nI}} X_{jI}^2) X_{ij}^2] \\
&\quad (\text{E.103})
\end{aligned}$$

Bibliography

- [1] J. Casulleras and J. Boronat, Phys. Rev. Lett. **84**, 3121 (2000).
- [2] J. von Stecher, C. H. Greene, and D. Blume, Phys. Rev. A **76**, 053613 (2007).
- [3] N. Trivedi, D. M. Ceperley, Phys. Rev. B **40**, 2737 (1989).
- [4] W. R. Magro and D. M. Ceperley, Phys. Rev. Lett. **73**, 826 (1994).
- [5] J. Boronat, J. Casulleras, and J. Navarro, Phys. Rev. B **50**, 3427 (1994).
- [6] J. Boronat and J. Casulleras, Phys. Rev. B **49**, 8920 (1994).
- [7] S. Moroni, S. Fantoni, and G. Senatore, Europhys. Lett. **30** 93-98 (1995).
- [8] S. Giorgini, J. Boronat, and J. Casulleras, Phys. Rev. Lett. **77**, 2754 (1996).
- [9] S. Moroni, G. Senatore, and S. Fantoni, Phys. Rev. B **55**, 1040 (1997).
- [10] J. Boronat and J. Casulleras, Europhys. Lett. **38** 291-296 (1997).
- [11] S. Moroni and M. Boninsegni, Europhys. Lett. **40** 287 (1997).
- [12] E. Strepparola, R. Nifosì and M. P. Tosi, J. Phys. : Cond. Mat. **10** 11645 (1998).
- [13] J. Boronat and J. Casulleras, Phys. Rev. B **59**, 8844 (1999).
- [14] F. Mazzanti, A. Polls, J. Boronat, and J. Casulleras, Phys. Rev. Lett. **92**, 085301 (2004).
- [15] F. Mazzanti, A. Polls, J. Boronat, and J. Casulleras, Phys. Rev. Lett. **92**, 085301 (2004).
- [16] L. Vranješ, J. Boronat, J. Casulleras, and C. Cazorla, Phys. Rev. Lett. **95**, 145302 (2005).
- [17] J. M. Marín, J. Boronat, and J. Casulleras, Phys. Rev. B **71**, 144518 (2005).
- [18] S. Ujevic and S. A. Vitiello, J. Phys. : Cond. Mat. **19** 116212 (2007).
- [19] V. Apaja and M. Saarela, Europhys. Lett. **84** 40003 (2008).

- [20] A. J. Williamson, R. Q. Hood, and J. C. Grossman, Phys. Rev. Lett. **87**, 246406 (2001).
- [21] A. Sarsa, S. Fantoni, K. E. Schmidt, and F. Pederiva, Phys. Rev. C **68**, 024308 (2003).
- [22] A. Fabrocini, S. Fantoni, A. Yu. Illarionov, and K. E. Schmidt, Phys. Rev. Lett. **95**, 192501 (2005).
- [23] B. Øksendal, “*Stochastic differential Equations : An Introduction with Applications*”, Springer (2007).
- [24] N. Carr, “*The big switch, rewiring the world from Edison to Google*”, Editor W. W. Norton & Co Ltd (2008).
- [25] I. Foster. Science, vol. 308, May 6, (2005).
- [26] D. Chappel sponsored by Microsoft, “*A short introduction to cloud platform an enterprised oriented view*”, D. Chappel and associate, www.davidchappel.com (2008).
- [27] I. Foster, “*There’s Grid in them thar Clouds*”, www.ianfostertypad.com, (2008).
- [28] J. Markoff, “*Software via the Internet : Microsoft in Cloud Computing*”, New York Times, September 3, (2007).
- [29] S. Baker, “*Google and the wisdom of clouds*”, BusinessWeek, December 16, (2007).
- [30] R. King, “*How cloud computing is changing business*”, BusinessWeek, August 4, (2008).
- [31] S. Lohr, “*Cloud computing gains steam with new I.B.M Gear*”, New York Times, April 23, (2008).
- [32] B. Johson, “*Cloud computing is a trap, warns GNU foudner Richard Stallman*”, The Guardian, September 29, (2008).
- [33] B. Worthen, “*Larry Ellison s Brilliant Anti-Cloud Computing Rant*”, The Wall Street Journal, [http ://blogs.wsj.com/biztech/](http://blogs.wsj.com/biztech/), September 25, (2005).

- [34] Oracle Press Release, “*Oracle Unveils New Cloud Computing Products and Services*”, http://www.oracle.com/us/corporate/press/017521_EN
- [35] M. M. Hurley and P. A. Christiansen, *J. Chem. Phys.* **86**, 1069 (1987).
- [36] P. A. Christiansen, *J. Chem. Phys.* **88**, 4867 (1988).
- [37] P. A. Christiansen, *J. Chem. Phys.* **95**, 361 (1991).
- [38] L. Mitáš, E. L. Shirley, and D. M. Ceperley, *J. Chem. Phys.* **95**, 3467 (1991).
- [39] P. Belohorec, S. M. Rothstein, and J. Vrbik, *J. Chem. Phys.* **98**, 6401 (1993).
- [40] S. M. Rothstein, *Int. J. Quantum Chem.* **60**, 803 (1996).
- [41] S. M. Rothstein, in *Recent Advances in Quantum Monte Carlo Methods*, Part I edited by W. A. Lester (World Scientific, Singapore, 1997), p. 181.
- [42] L. Mitáš, in *Computer Simulation Studies in Condensed Matter Physics V*, edited by D. P. Landau, K. K. Mon, and H. B. Schuttler (Springer, Berlin, 1993), p. 94.
- [43] L. Mitáš, *Phys. Rev. A* **49**, 4411 (1994).
- [44] H. J. Flad and M. Dolg, *J. Chem. Phys.* **107**, 7951 (1997).
- [45] I. V. Ovcharenko, W. A. Lester, Jr., C. Xiao, and F. Hagelberg, *J. Chem. Phys.* **114**, 9028 (2001).
- [46] P. Sen and L. Mitáš, *Phys. Rev. B* **68**, 155404 (2003).
- [47] O. El Akramine, W. A. Lester, Jr., X. Krokidis, C. A. Taft, T. C. Guimaraes, A. C. Pavao, and R. Zhu, *Mol. Phys.* **101**, 277 (2003).
- [48] S. Sokolova and A. Luchów, *Chem. Phys. Lett.* **320**, 421 (2000).
- [49] L. Wagner and L. Mitáš, *Chem. Phys. Lett.* **370**, 412 (2003).
- [50] M. Caffarel, J-P. Daudey, J-L. Heully, and A. Ramírez-Solís, *J. Chem. Phys.* **123**, 094102 (2005).
- [51] A. Aspuru-Guzik, O. El Akramine, J. C. Grossman, and W. A. Lester, Jr., *J. Chem. Phys.* **120**, 3049 (2004).

- [52] J. Koseki, R. Maezono, M. Tazchikawa, M. D. Towler, and R. J. Needs, *J. Chem. Phys* **129**, 085103 (2008).
- [53] A. de la Lande, J. Maddaluno, A. Scemama, J. Pilmé, D. Salahub, O. Parisel, H. Gérard, M. Caffarel, J.-P. Piquemal, “Spin states can control dioxygen activation in metalloenzymes through spatial reorientation of oxygen lone pairs” (preprint).
- [54] J. C. Grossman, L. Mitáš, and K. Raghavachari, *Phys. Rev. Lett.* **75**, 3870 (1995).
- [55] J. C. Grossman and L. Mitáš, *Phys. Rev. Lett.* **74**, 1323 (1995).
- [56] L. Mitáš, J. C. Grossman, I. Stich, and J. Tobik, *Phys. Rev. Lett.* **84**, 1479 (2000).
- [57] T. Torelli and L. Mitáš, *Phys. Rev. Lett.* **85**, 1702 (2000).
- [58] G. Belomoin, E. Rogozhina, J. Therrien, P. V. Braun, L. Abuhassan, M. H. Nayfeh, L. Wagner, and L. Mitáš, *Phys. Rev. B* **65**, 193406 (2002).
- [59] A. Sarsa, J. Mur-Petit, A. Polls, and J. Navarro, *Phys. Rev. B* **68**, 224514 (2003).
- [60] E. Bodo, E. Coccia, D López-Durán and F. A. Gianturco, *Phys. Scr.* **76** No 3 C104-C110 (2007).
- [61] M. Neek-Ammal, G Tayebirad and R Asgari, *J. Phys. B : At. Mol. Opt. Phys.* **40** No 8 1509-1521 (2007).
- [62] E. Coccia, E. Bodo, and F. A. Gianturco, *Europhys. Lett.* **82** No 2 23001 (2008)
- [63] M. L. Tiago, P. R. C. Kent, R. Q. Hood, and F. A Reboredo *J. Chem Phys* **129**, 084311 (2008).
- [64] M. C. Gordillo, J. Boronoat, and J. Casulleras, *Phys. Rev. B* **61**, R878 (2000).
- [65] M. C. Gordillo, J. Boronoat, and J. Casulleras, *Phys. Rev. Lett.* **85**. 2348 (2000).

- [66] M. C. Gordillo, J. Boronoat, and J. Casulleras, Phys. Rev. B **65**, 014503 (2001).
- [67] M. C. Gordillo, J. Boronoat, and J. Casulleras, Low Temp. Phys. **126**, 199 (2002).
- [68] M. C. Gordillo, J. Boronoat, and J. Casulleras, Phys. Rev. B **68**, 0125421 (2003).
- [69] G. Cicero, J. C. Grossmann, and G. Galli, Phys. Rev. B **74**, 035425 (2006).
- [70] M. C. Gordillo, Phys. Rev. Lett. **96**, 216102 (2006).
- [71] M. C. Gordillo, J. Boronoat, and J. Casulleras, Phys. Rev. B **76**, 0193402 (2007).
- [72] M. C. Gordillo, Phys. Rev. Lett. **101**, 046102 (2008).
- [73] “*Electronic Materials : The Oligomer Approach*”, Edited by K. Mullen and G. Wegner , Wiley- VCH (2007).
- [74] W. A. Sokalski, “*Molecular Materials with Specific Interactions - Modeling and Design*” (Challenges and Advances in Computational Chemistry and Physics), Springer (2007).
- [75] N. C. Cohen, *Handbook on Molecular Modeling in Drug Design*, Academic Press (2006).
- [76] S. Manten and A. Luchów, J. Chem. Phys. **119**, 1307 (2003).
- [77] F. A. Reboredo and A. J. Williamson, Phys. Rev. B **71**, 121105 (2005).
- [78] J. Kussmann and C. Ochsenfeld, J. Chem. Phys. **128** 134104 (2008).
- [79] D. Alfé and M. J. Gillan, J. Phys. : Condens. Matter **16**, L305 (2004).
- [80] D. Maynau, S. Evangelisti, N. Guihéry, C. J. Calzado, J.-P. Malrieu, J. Chem. Phys **116** 10060 (2002).
- [81] M. Caffarel, “Stochastic methods in quantum mechanics” in *Numerical Determination of the Electronic Structure of Atoms, Diatomic and Polyatomic Molecules* (Kluwer Academic Publishers, 1989), pp.85-105.

- [82] B. L. Hammond, W. A. Lester Jr., and P. J. Reynolds in “*Monte Carlo Methods in Ab Initio Quantum Chemistry*”, World scientific Lecture and course notes in chemistry Vol.1 (1994).
- [83] W. M. C. Foulkes, L. Mitáš, R. J. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- [84] R. Brown *Philosophical Magazine N.S.* **6**, 161-166 (1829).
- [85] L. Bachelier, *Annales de l’Ecole Normale Supérieure* **17** (1900).
- [86] A. Einstein, *Ann. Phys.* **17**, 549-60 (1905).
- [87] P. Langevin, *C. R. Acad. Sci. (Paris)* **146**, 530-533 (1908).
- [88] A. D. Fokker, “*Over Brown’sche bewegingen in het stralingsveld, en waarschijnlijkheidsbeschouwingen in de stralingstheorie*”, Universiteit Leiden (1913).
- [89] M. Kac, *Transactions of the American Mathematical Society* **65** (1) 1-13 (1949).
- [90] J. F. Ramlay, *The American Mathematical Monthly* **76** (8) : 916-918 (1969).
- [91] N. Metropolis, S. Ulam, *Journal of the statistical association*, **44** 335-341 (1949).
- [92] N. Metropolis , “*The beginning of the Monte-Carlo method*”, Los Alamos Science (1987).
- [93] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21** 1087 (1953).
- [94] W.K. Hasting, *Biometrika*, 57(1) :97-109 (1970).
- [95] J. B. Anderson, *J. Chem. Phys.* **63** 4 (1975).
- [96] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester Jr., *J. Chem. Phys.* **77** 11 (1982).
- [97] M. Casula, C. Attacalite, S. Sorella, *J. Chem. Phys.* **121**, 7110 (2004).
- [98] F. Schautz, F. Buda, C.Filippi *J. Chem. Phys.* **121**, 5836 (2004).

- [99] A. J. Williamson, J. C. Grossman, and G. Galli, *J. Chem. Phys.* **117**, 6721 (2002).
- [100] A. Puzder, A. J. Williamson, J. C. Grossman, and G. Gali, *J. Chem. Phys.* **117**, 6721 (2002).
- [101] A. Puzder, A. J. Williamson, J. C. Grossman, and G. Gali, *Phys. Rev. Lett.* **88**, 097401 (2002).
- [102] G. Belomoin, E. Rogozhina, J. Therrien, P.V. Braun, L. Abuhassan, M.H. Nafeh, L. Wagner, and L. Mitáš, *Phys. Rev. B* **65**, 193406 (2002).
- [103] C. Filippi, S. B. Healy, P. Kratze, E. Pehlke, and M. Scheffler, *Phys. Rev. Lett.* **89**, 166102 (2002).
- [104] F. A. Matsen, *J. Adv. Quantum Chem.* **1**, 59-114 (1964).
- [105] D. Ceperley, G. V. Chester, and M. H. Kalos, *Phys Rev. B* **16**, 3081 (1977).
- [106] A. Scemama, T. Lelièvre, G. Stolz, E. Cancès, M. Caffarel, *J. Chem. Phys.* **125**, 114105 (2006).
- [107] D. M. Ceperley and B. J. Alder, *J. Chem. Phys.* **77**, 5593 (1982).
- [108] D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett* **45**, 566-569 (1980).
- [109] M. H. Kalos and F. Pederivia, *Phys. Rev. Lett.* **85** 3547,(2000).
- [110] R. Assaraf, M. Caffarel, A. Khelif, *J. Phys. A : Math. Theor.* **40**, p. 1181 (2007).
- [111] J. Toulouse and C. J Umrigar, *J. Chem. Phys.* **128**, 174101 (2008).
- [112] D. Bressanini, D. Ceperley and P. Reynold in “*Recent Advances in Quantum Monte Carlo Methods*” Part II edited by W. A. Lester *et al.* World Scientific, (2002).
- [113] S. Manten and A. Luchów, *J. Chem. Phys.* **115**, 5362 (2001).
- [114] W. Klopper, K. L. Bak, P. Jorgensen, J. Olsen, and T. Helgaker, *J. Phys. B* **32** R103 (1999).
- [115] H. F. Trotter, *Proc. Am. Math. Soc.* **10**, 545 (1959).

- [116] M. Suzuki, S. Miyashita, and A. Kuroda, Prog. Theor. Phys. **58**, 1377 (1977).
- [117] M. Suzuki, J. Math. Phys. **32**, 400 (1991).
- [118] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).
- [119] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).
- [120] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, R. G. Henning, Phys. Rev. Lett. **98** 110201 (2007).
- [121] J. Toulouse, C. J. Umrigar, J. Chem. Phys. **126**, 084102 (2007).
- [122] M. Caffarel, A. Scemama, T. Bouabça, QMC=Chem2008 is a package of QMC programs for electronic structure calculations, Toulouse, France 2008
- [123] R. Assaraf and M. Caffarel, Phys. Rev. Lett. **83**, No. 23, (1999).
- [124] R. Assaraf, M. Caffarel, J. Chem. Phys. **113**, 4028 (2000).
- [125] R. Assaraf, M. Caffarel, J. Chem. Phys. **119** 20, (2003).
- [126] C. J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005).
- [127] W.H. Press, S. A. Teukolsky, W.T. Vetterling, B.P. Flannery, “*Numerical Recipes in Fortran, The Art of Scientific Computing*”, Cambridge University Press, 59-63
- [128] G. H. Golub, Reinsch, Numer Math **14** 403-420 (1970).
- [129] A. S. Householder, Journal ACM, 5 (4) 339-342 (1958).
- [130] J. G. F. Francis, The Computer Journal, vol. 4, no. 3, 265-271 (1961).
- [131] J. G. F. The Computer Journal, vol. 4, no. 4, 332-345 (1962).
- [132] V. N. Kublanovskaya, USSR Computational Mathematics and Mathematical Physics, vol. 3, 637-657 (1961).
- [133] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, LAPACK Users’ Guide, Third (1999).

- [134] C. B. Lieber and G. W. Stewart *SIAM J. Numer. Anal.* **10** 241-256 (1973).
- [135] M. P. Nightingale and V. Melik-Alaverdian, *Phys. Rev. Lett.* **87** 043401 (2001).
- [136] T. Kato. *Comm. Pure Appl. Math.* **10** 151 (1957).
- [137] K. E. Schmidt and J. W. Moskowitz, *J. Chem. Phys.* **93**, 4173 (1990).
- [138] Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004
- [139] GAMESS-UK is a package of ab initio programs. See : "[http :www.cfs.dl.ac.uk/gamess-ukindex.shtml](http://www.cfs.dl.ac.uk/gamess-ukindex.shtml)", M.F. Guest, I. J. Bush, H.J.J. van Dam, P. Sherwood, J.M.H. Thomas, J.H. van Lenthe, R.W.A Havenith, J. Kendrick, "The GAMESS-UK electronic structure package : algorithms, developments and applications", *Molecular Physics* **103**, No. 6-8, 20 March-20 April 2005, 719-747
- [140] MOLPRO, version 2006.1, a package of ab initio programs, H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Sch \tilde{A} $\frac{1}{4}$ tz, P. Celani, T.

- Korona, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel and G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer and M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni and T. Thorsteinsson, see <http://www.molpro.net>.
- [141] G Karlström, R. Lindh, P.-Å. Malmqvist, B.O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady, L. Seijo, *Computational Material Science* **28**, 222 (2003)
- [142] T. Helgaker, P. Jørgensen, J. Olsen, “*Molecular electronic-structure theory*”, John Wiley & sons,LTD, chapitres 10,11,12
- [143] R. P. Feynman and M. Cohen, *Phys. Rev.* **12**, 1189-1204 (1956).
- [144] Y. Kwon, D. M. Ceperley and R. M. Martin, *Phys Rev. B* **58**, 6800-6806 (1998).
- [145] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, “*Numerical Recipes in Fortran, The Art of Scientific Computing*”, i Cambridge University Press, 65-67.
- [146] C. F. Bunge, J. A. Barrientos, A. Vivier-Bunge, and J. A. Cogordan, *Phys. Rev. A* **46**, 3691 (1992).
- [147] F. Schautz and C. Filippi, *J. Chem. Phys.* **120**, 10931 (2004).
- [148] F. Cordova, L. Joubert Doriol, A. Ipatov, and M. E. Casida, *J. Chem. Phys.* **127**, 164111 (2008).
- [149] E. Tapavicza, I. Tavernilli, U. Rothlisberger, C. Filippi, and M. E. Casida, *J. Chem. Phys.* **129**, 124108 (2008).