

Méthode Monte Carlo pour les électrons sur réseau

André-Marie Tremblay

Département de physique et

Centre de recherche en physique du solide

Université de Sherbrooke, Sherbrooke, Québec J1K 2R1

1. Introduction.

Il existe un grand nombre d'approches aux simulations Monte Carlo pour les systèmes quantiques. On peut, en gros, dégager trois grandes catégories: Monte Carlo variationnel, fonctions de Green Monte Carlo, et Monte Carlo pour les intégrales de chemin. Chacune de ces approches a plusieurs variantes. Nous nous contenterons ici d'une brève description, qui devrait permettre de situer dans un contexte plus global la méthode qui sera décrite plus en détail dans les pages qui suivent.

Schématiquement dans la première approche, le Monte Carlo variationnel, on cherche à obtenir les paramètres d'une fonction d'onde variationnelle en minimisant l'énergie. Le calcul de l'énergie, ainsi que celui des autres observables, nécessite l'évaluation d'intégrales multidimensionnelles, ce qui peut se faire efficacement par les méthodes Monte Carlo¹. Dans la deuxième approche, fonctions de Green Monte Carlo, on utilise l'analogie entre l'équation de Schrödinger et l'équation de diffusion pour calculer la fonction d'onde du fondamental par une marche aléatoire. La fonction de Green qui définit cette marche aléatoire est elle-même obtenue par la méthode Monte Carlo. En un certain sens, on se sert ici de la méthode Monte Carlo pour résoudre une équation intégrale par itération, l'aspect Monte Carlo intervenant évidemment dans le calcul des intégrales. Kalos¹ donne une introduction simple à cette méthode². Finalement, la méthode des intégrales de chemin transforme un problème de mécanique quantique en un problème qui ressemble à l'évaluation d'une fonction de partition, problème avec lequel nous sommes déjà familiers.

La méthode que nous allons décrire ici s'apparente à la méthode des intégrales de chemin. Dans toutes les approches aux problèmes quantiques contenant plus d'une particule, nous faisons face à deux cas très différents: les bosons et les fermions. On peut même ajouter un troisième cas aux précédents, celui des spins, dont les relations de commutation ne sont ni bosoniques ni

fermioniques. Ces problèmes de spin se présentent lorsque les degrés de liberté de translation ne sont pas importants, comme c'est le cas dans les matériaux isolants. Nous ne discuterons ici que du cas des fermions, qui s'avère être le plus difficile. En effet, l'antisymétrie de la fonction d'onde empêche en général de définir les quantités positives-définies nécessaires en simulation Monte Carlo, que ce soit la fonction de Green, ou le poids servant à l'échantillonnage selon l'importance. C'est le fameux "problème de signe" dont l'origine précise est en général mal connue et qui dépend des détails de la méthode de simulation. Comme nous le verrons dans ce qui suit, ce problème représente encore un obstacle majeur à l'utilisation de la méthode Monte Carlo pour les problèmes fermioniques.

Dans ce qui suit, nous motiverons d'abord les calculs par le problème physique dit des "électrons fortement corrélés". Ce problème a repris de l'importance depuis la découverte des supraconducteurs à haute température de transition. Nous décrirons ensuite en détail la méthode dite "du déterminant", développée il y a une dizaine d'années en Californie^{3,4,5,6}, pour étudier le modèle de Hubbard. Ce modèle est un exemple particulièrement simple d'électrons fortement corrélés qui est l'objet d'un renouveau d'attention dans le contexte des supraconducteurs à haute température de transition. Nous concluons par quelques exemples de résultats.

2. Les fermions sur réseau: motivation physique.

Un des grands problèmes de l'heure en physique de la matière condensée est celui des électrons fortement corrélés. Que veut-on dire par là? Dans beaucoup de matériaux, les propriétés électroniques peuvent être expliquées par des modèles d'électrons se propageant dans des bandes. Les interactions, qu'elles soient du type électron-électron ou électron-phonon, sont faibles et on peut en tenir compte par la théorie des perturbations. Dans le cas des électrons fortement corrélés, ce n'est plus le cas. Les supraconducteurs à haute température de transition semblent être une classe de matériaux qui nécessite des modèles où les électrons sont fortement corrélés. Sans entrer dans les détails des motivations physiques pour ce dernier énoncé, mentionnons brièvement qu'une des raisons qui suggère l'emploi de modèles d'électrons fortement corrélés est le fait qu'ils deviennent des isolants antiferromagnétiques dans un domaine de paramètres où la théorie des bandes prédit qu'ils devraient être des métaux. Il y a plusieurs modèles typiques pour les électrons fortement corrélés. On utilise généralement utilisé le plus simple d'entre eux, soit le modèle de Hubbard⁷. Bien qu'il soit le plus simple, ce modèle n'est toujours pas résolu dans l'ensemble du domaine de paramètres d'intérêt. Ce modèle est brièvement décrit dans la section suivante, où on indique ensuite pourquoi les diagonalisations exactes sont difficiles. Une bonne partie du reste de cet article est basé sur une présentation antérieure.⁸

2.1. Les électrons fortement corrélés dans des bandes étroites: Modèle de Hubbard.

Tant la structure que les propriétés physiques des supraconducteurs à haute température de transition suggèrent qu'il est essentiel de commencer par comprendre des modèles bi-dimensionnels. En effet, tous les supraconducteurs à haute température de transition ont en commun des plans de cuivre-oxygène. Se restreignant aux basses énergies, et donc à la bande la plus près du niveau de Fermi où l'interaction de Coulomb est écrantée, le modèle le plus simple a la forme suivante,

$$\hat{H} = -t \sum_{\langle i,j \rangle, \sigma} (c_{i,\sigma}^+ c_{j,\sigma} + h.c.) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}. \quad (1)$$

Les indices i et j dénotent les sites d'un réseau carré bi-dimensionnel et les parenthèses $\langle \rangle$ indiquent que la première somme est restreinte aux premiers voisins. L'opérateur $c_{j,\sigma}$ détruit un électron de spin σ sur le site j alors que $c_{i,\sigma}^+$ crée un électron de même spin sur le site i . En d'autres mots, le premier terme de l'Hamiltonien représente le fait qu'un électron sur un site peut sauter sur un autre site: c'est le terme d'énergie cinétique. On suppose que l'écrantage est le plus efficace possible et donc le dernier terme de l'Hamiltonien représente l'énergie de répulsion de Coulomb qui n'agit que lorsque deux électrons occupent le même site. Comme il n'y a qu'une orbitale effective par site, le principe d'exclusion de Pauli ne permet pas à un site d'être occupé par plusieurs électrons de même spin, mais deux électrons de spin opposé peuvent se retrouver sur le même site.

Un mot sur la notation: les opérateurs de seconde quantification, comme \hat{H} ci-dessus, seront toujours coiffés d'un accent circonflexe, à moins qu'il ne s'agisse des opérateurs de création et d'annihilation eux-mêmes qui seront toujours représentés par la lettre c .

2.2. La difficulté des diagonalisations exactes.

Considérons un réseau bi-dimensionnel de N sites. Chaque site peut être soit vide, soit occupé par un électron dont le spin peut être en haut ou en bas, soit occupé par deux électrons de spin opposé. Ceci veut dire que chaque site peut avoir quatre états possibles, et donc que pour N sites, l'espace des états à plusieurs électrons a une dimension 4^N . Pour $N = 16$, cela veut déjà dire 2^{32} états (rappelons que un mégaoctet de mémoire contient environ 2^{20} octets). L'Hamiltonien de Hubbard (1) définit une matrice qui doit être diagonalisée dans cet espace des états. En groupant les états avec un nombre fixe d'électrons, on rend cet Hamiltonien diagonal par bloc. De plus, seulement les états complètement antisymétriques sont permis physiquement (Principe de Pauli), ce qui réduit le nombre d'états à considérer, pour N électrons dont N^σ sont de spin σ , à

$$\prod_{\sigma = \pm 1} \frac{N!}{(N - N^\sigma)!(N^\sigma)!}.$$

Néanmoins, pour un nombre d'électrons comparable au nombre de sites (densité finie...) cet espace de vecteurs d'états est encore énorme. De plus, il y a des éléments de matrice hors diagonaux entre toutes les paires de vecteurs d'états qui ne diffèrent que par des déplacements d'électrons d'un site voisin à l'autre, le spin étant conservé dans les déplacements. Par exemple, si un vecteur d'état rangé peut être obtenu d'un vecteur d'état colonne en déplaçant un seul électron d'un seul site et en préservant l'ordre de remplissage des autres sites (important à cause de l'antisymétrie), alors il y a un élément de matrice égal à $-t$ entre ces deux vecteurs d'état dans la matrice Hamiltonienne. Il y a aussi une contribution diagonale à la matrice Hamiltonienne égale à U fois le nombre de sites occupés par deux électrons de spin opposés.

La structure de la matrice Hamiltonienne serait simple si l'Hamiltonien ne contenait qu'un seul terme: celui d'énergie cinétique (diagonalisable par transformée de Fourier) ou celui d'énergie potentielle (diagonalisable dans l'espace des positions). Lorsque les deux termes sont comparables, la matrice Hamiltonienne est éparsée, comme nous venons de le voir, mais elle n'a pas une structure simple et l'espace des états est exponentiellement grand. En pratique donc, les diagonalisations exactes doivent utiliser les symétries au maximum pour simplifier le problème et il est difficile d'aller au-delà de réseaux 4×4 à cause de la croissance exponentielle du problème avec la taille. Par contre, la méthode Monte Carlo ne nécessite pas la mémorisation de tout l'espace des états et, de plus, le nombre d'opérations arithmétiques nécessaires à la solution n'augmente que de façon polynomiale avec la taille du réseau et l'inverse de la température. L'article de Dagotto *et al.*⁹ compare les résultats obtenus à la fois par des méthodes Monte Carlo et par diagonalisations exactes pour les supraconducteurs à haute température de transition. Les avantages et inconvénients de chaque méthode deviennent clairs. Nous reviendrons plus loin aux avantages et désavantages de la méthode Monte Carlo.

3. Monte Carlo quantique par la méthode des déterminants (en bref).

Nous décrivons brièvement dans cette section l'algorithme connu sous le nom "méthode du déterminant" ("determinantal approach") ou sous l'acronyme BSS, suivant les initiales de ses découvreurs³, Blankenbecler, Scalapino et Sugar. Au début, cette approche a été surtout appliquée au modèle de Hubbard par Hirsch⁴ qui proposa aussi des améliorations. Il y a quelques années, White *et al.*⁵ ont trouvé la solution à un problème de longue date, celui de la stabilité des calculs à basse température. Plusieurs des résultats obtenus pour le modèle de Hubbard par cette approche ont été résumés dans la littérature.¹⁰

La section 3.1 décrit la décomposition de Trotter, communément utilisée pour réduire les problèmes de mécanique quantique à des problèmes d'intégrale fonctionnelle. Cette décomposition a été décrite dans d'autres cours de la présente école d'été. (Voir par exemple le cours de José Lopez). Il s'agit d'une technique commune à presque toutes les approches numériques. La transformation de Hubbard-Stratonovich de la section 3.2 permet ensuite d'éliminer les fermions du problème, tel que démontré en 3.3. Le problème est alors réduit à une trace sur des variables classiques qui ressemble au calcul d'une fonction de partition. Après avoir donné une interprétation physique en 3.4 et décrit comment en principe calculer des observables dans la section 3.5, on montre en

3.6 que l'échantillonnage selon l'importance n'est plus aussi efficace que dans le cas de variables classiques parce que le caractère fermionique des degrés de liberté introduit des signes négatifs qui ne permettent plus d'avoir des "poids de Boltzmann" positifs. C'est le fameux "problème de signe". Plusieurs détails de calcul sont relégués à des annexes. Ce n'est que dans la section 4 que l'on discutera plus en détail les problèmes pratiques qui surgissent lors d'une simulation, entre autres celui de la stabilité aux basses températures et celui de l'accélération de l'exécution par des techniques de matrices éparses. Des exemples de résultats obtenus par cette méthode sont donnés dans la section 5. Les problèmes d'ergodicité et les problèmes de signe sont discutés, avec des exemples, dans la dernière section.

3.1. Formule de Trotter.

Dans l'ensemble grand-canonique, on calcule la valeur moyenne d'une observable \mathbf{O} en utilisant,

$$\langle \hat{O} \rangle = \frac{\text{Tr}\{\exp[-\beta (\hat{H} - \mu \hat{N})] \hat{O}\}}{\text{Tr}\{\exp[-\beta (\hat{H} - \mu \hat{N})]\}} \quad (2)$$

où \hat{H} est l'Hamiltonien de Hubbard (1) et \hat{N} donne le nombre d'électrons dans l'état sur lequel cet opérateur agit. La trace est faite dans l'espace des états décrits dans la section précédente. L'opérateur $\hat{H} - \mu \hat{N}$ contient une partie diagonale \hat{V} contenant l'énergie d'interaction, qui ne commute pas avec la partie hors-diagonale \hat{K} représentant l'énergie cinétique. Tel que mentionné ci-haut, \hat{K} ou \hat{V} par eux-mêmes seraient faciles à diagonaliser. On voit donc que si seulement l'exponentielle d'un ou l'autre de ces opérateurs plutôt que l'exponentielle de leur somme apparaissait dans l'équation (2) le problème serait simplifié. Ceci peut se faire en utilisant la décomposition dite de Trotter¹¹:

$$e^{-\beta(\hat{H} - \mu \hat{N})} = \prod_{l=1}^{N_\tau} e^{-\Delta \tau (\hat{K}_l + \hat{V}_l)} \approx \prod_{l=1}^{N_\tau} e^{-\Delta \tau \hat{K}_l} e^{-\Delta \tau \hat{V}_l} + O((\Delta \tau)^2). \quad (3)$$

où $N_\tau \Delta \tau$ est égal à β . L'entier N_τ , noté P dans les notes de Lopez, est d'habitude appelé "le nombre de tranches de temps imaginaire". Cette appellation de temps imaginaire est courante en physique statistique quantique et vient du fait que $\exp(-\beta \hat{H})$ a la même forme que l'opérateur d'évolution $\exp(i\hat{H}t)$ à condition que le temps t soit considéré comme imaginaire. Dans le reste de ces notes, lorsque nous parlerons de temps, nous voudrions toujours dire temps imaginaire. Le fait que l'erreur soit d'ordre $(\Delta \tau)^2$, lorsque la relation (3) ci-dessus est utilisée dans une trace, peut être prouvé explicitement (Fye¹² et Suzuki¹³). Une preuve abrégée du résultat est donnée dans l'annexe A. Il existe beaucoup d'autres décompositions de Trotter possibles¹² mais en pratique, c'est celle ci-dessus qui est utilisée. Notons que la propriété cyclique de la trace rend inutile¹² les décompositions hermitiques du type $e^{-\Delta \tau \hat{K}_l / 2} e^{-\Delta \tau \hat{V}_l} e^{-\Delta \tau \hat{K}_l / 2}$ que l'on rencontre souvent dans d'autres contextes.

Pour chacune des N_τ tranches de temps, nous avons un réseau de N sites dans l'espace. Nous travaillons donc maintenant effectivement avec un réseau de $N_\tau N$ sites existant dans l'espace des positions et du temps imaginaire. Plus la température est basse, plus il sera nécessaire de prendre de tranches de temps imaginaire pour obtenir une précision donnée. À une température donnée, le nombre de tranches de temps augmente avec la précision désirée, ce qui rendra le calcul d'autant plus long. C'est en partie ce qui fait la difficulté des simulations dans le domaine quantique: pour obtenir un résultat exact, même pour un seul site, on doit avoir en principe un réseau de taille infinie dans la direction du temps imaginaire. En pratique, on procède par extrapolation ou on fait le calcul pour une valeur de $\Delta \tau$ suffisamment petite pour que l'erreur systématique d'ordre $(\Delta \tau)^2$ soit plus petite ou du même ordre que l'erreur statistique inhérente aux simulations Monte Carlo. Nous verrons plus bas des exemples de ce genre de calculs (Figure 10 et annexe A). Pour le genre de décomposition décrit par l'équation (3), la valeur de $\Delta \tau$ nécessaire pour obtenir une précision donnée est indépendante du nombre de sites dans l'espace, tel que démontré par Fye.¹²

3.2. Découplage Hubbard-Stratonovich des interactions. Degrés de liberté de jauge.

On peut rendre l'Hamiltonien quadratique en opérateurs de fermions. Malgré le prix qu'il faut payer, soit l'introduction d'un nombre de variables classiques égal au nombre de sites dans le réseau espace-temps, cela en vaut la peine. En effet lorsque l'Hamiltonien est quadratique en fermions il devient possible de faire la trace sur les variables fermioniques. Ne subsistent alors que

des variables classiques qu'on peut alors traiter par des méthodes analogues à celles discutées, par exemple, dans ce volume par Alain Mailhot.

Pour chaque tranche de temps ℓ , définissons

$$\hat{V}_{i,\ell} \equiv U \hat{n}_{i,\ell \uparrow} \hat{n}_{i,\ell \downarrow} - \mu (\hat{n}_{i,\ell \uparrow} + \hat{n}_{i,\ell \downarrow}).$$

Notez que pour les opérateurs fermioniques l'indice ℓ ne dénote pour le moment que la position dans le produit de Trotter. Comme les opérateurs de nombre commutent, nous pouvons aussi écrire sans approximation supplémentaire,

$$e^{-\Delta \tau \sum_i \hat{V}_{i,\ell}} = \prod_{i=1}^N e^{-\Delta \tau \hat{V}_{i,\ell}}.$$

Notant par un indice grec ν chacun des sites i, ℓ sur le réseau espace-temps, il y a un terme de la forme

$$e^{-\Delta \tau \hat{V}_\nu} = e^{-\Delta \tau \left[U \left(\hat{n}_{\nu \uparrow} - \frac{1}{2} \right) \left(\hat{n}_{\nu \downarrow} - \frac{1}{2} \right) - \left(\mu - \frac{U}{2} \right) (\hat{n}_{\nu \uparrow} + \hat{n}_{\nu \downarrow}) - \frac{U}{4} \right]}$$

pour chaque site du réseau espace-temps. Chacun de ces termes peut être rendu quadratique en opérateur de fermion en utilisant la relation

$$e^{-\Delta \tau \left[U \left(\hat{n}_{\nu \uparrow} - \frac{1}{2} \right) \left(\hat{n}_{\nu \downarrow} - \frac{1}{2} \right) \right]} = e^{-\Delta \tau \frac{U}{4}} \int_{-\Lambda}^{\Lambda} dx P(x) e^{x(\hat{n}_{\nu \uparrow} - \hat{n}_{\nu \downarrow})}$$

qu'il est facile de satisfaire avec un grand nombre de fonctions $P(x)$. En effet, considérant les quatre états possibles sur chaque site, soient $|\uparrow\rangle|\downarrow\rangle = |0\rangle|0\rangle, |0\rangle|1\rangle, |1\rangle|0\rangle, |1\rangle|1\rangle$, la fonction $P(x)$, choisie paire en x , n'a que deux conditions à satisfaire pour que l'équation précédente soit vraie:

$$1 = \int_{-\Lambda}^{\Lambda} dx P(x)$$

$$e^{\Delta \tau \frac{U}{2}} = \int_{-\Lambda}^{\Lambda} dx P(x) \cosh(x).$$

Physiquement, on peut considérer le champ de Hubbard et Stratonovich x comme étant un boson fictif qui est à l'origine de l'interaction de Hubbard U , par analogie avec le photon qui est à l'origine de l'interaction de Coulomb en électrodynamique quantique. Le choix $\Lambda = \infty$ et

$$P(x) = \frac{1}{\sqrt{2\pi \Delta \tau U}} e^{-\frac{x^2}{2 \Delta \tau U}}$$

correspond à ce qui est normalement connu sous le nom de transformation de Hubbard-Stratonovich. Étant donné les propriétés bien connues des gaussiennes, ce choix est évidemment commode pour les calculs analytiques. Pour les calculs numériques cependant, il est préférable d'utiliser une variable discrète plutôt que continue. Ceci a été fait par Hirsch¹⁴ qui, pour $U > 0$, a suggéré le choix,

$$P(x) = \frac{1}{2} [\delta(x - \lambda) + \delta(-x - \lambda)]$$

$$\cosh(\lambda) = e^{\Delta \tau \frac{U}{2}}$$

Explicitement, le choix de Hirsch correspond à utiliser,

$$e^{-\Delta \tau \left[U \left(\hat{n}_{v\uparrow} - \frac{1}{2} \right) \left(\hat{n}_{v\downarrow} - \frac{1}{2} \right) \right]} = e^{-\Delta \tau \frac{U}{4}} \frac{1}{2} \sum_{x_v = \pm 1} e^{\lambda x_v (\hat{n}_{v\uparrow} - \hat{n}_{v\downarrow})}. \quad (4)$$

Même en se restreignant à des transformations discrètes, il existe un grand nombre de choix possibles pour découpler les termes d'interaction comme nous venons de le faire. En effet, l'équation précédente montre bien que le champ x_v se couple à la composante z du spin des fermions au site d'espace-temps v . Comme l'axe de quantification local est complètement arbitraire, on s'attend à ce que l'identité suivante soit aussi satisfaite¹⁵:

$$e^{-\Delta \tau \left[U \left(\hat{n}_{v\uparrow} - \frac{1}{2} \right) \left(\hat{n}_{v\downarrow} - \frac{1}{2} \right) \right]} = e^{-\Delta \tau \frac{U}{4}} \frac{1}{2} \sum_{x_v = \pm 1} e^{-\lambda x_v (2\hat{s} \cdot \vec{l})} \quad (5)$$

où \vec{l} est un vecteur unité de direction arbitraire. Le choix de \vec{l} correspond donc à un choix de jauge.¹⁶ Cette dernière identité est prouvée dans l'annexe B. Des transformations correspondant à des rotations dans l'espace dit de Nambu sont aussi possibles.¹⁵

Comme toutes les transformations de Hubbard-Stratonovich décrites ci-dessus sont exactes, les résultats physiques ne dépendent pas de la transformation choisie lorsqu'un nombre infini d'itérations Monte Carlo est fait. Cependant, plusieurs quantités importantes dans l'utilisation pratique de l'algorithme dépendent du choix de transformation: par exemple, "le signe du déterminant" que nous rencontrerons dans la section suivante, et l'efficacité avec laquelle

l'algorithme Monte Carlo explore l'espace des phases (ces problèmes sont discutés plus en détail dans la section 6.2). Malheureusement, il n'y a pas de façon systématique de savoir à l'avance quel est le choix de la transformation de Hubbard-Stratonovich le plus approprié pour une situation donnée. Des résultats additionnels récents sur cette transformation apparaissent dans l'annexe C.

3.3. Trace sur les fermions.

Pour être plus spécifique, nous utilisons dans ce qui suit la transformation de Hubbard-Stratonovich (4) avec quantification selon l'axe z . Les résultats seront facilement généralisables. La fonction de partition à évaluer prend maintenant la forme,

$$\begin{aligned}
Z &= \sum_{\{x_{i,\ell}\}} \prod_{\sigma = \pm 1} Tr \left[\prod_{\ell=1}^{N_\tau} e^{-\Delta \tau \hat{k}_\ell^\sigma} e^{\sum_{i=1}^N \hat{n}_{i,\ell,\sigma} \left[\lambda \sigma x_{i,\ell} + \Delta \tau \left(\mu - \frac{U}{2} \right) \right]} \right] \\
&\equiv \sum_{\{x_{i,\ell}\}} \prod_{\sigma = \pm 1} Tr \left[\prod_{\ell=1}^{N_\tau} \hat{D}_\ell^\sigma(\{x_{i,\ell}\}) \right] \\
&\equiv \sum_{\{x_{i,\ell}\}} \prod_{\sigma = \pm 1} Tr \left[\hat{U}^\sigma(N_\tau, 0; \{x_{i,\ell}\}) \right]
\end{aligned} \tag{6}$$

où la somme est sur toutes les configurations $\{x_{i,\ell}\}$ possibles des champs de Hubbard-Stratonovich. La trace Tr est effectuée sur les variables fermioniques et le facteur $(1/2)^{NN_\tau}$ a été abandonné puisqu'il se simplifie dans les calculs de moyennes thermodynamiques. Cette trace peut être faite exactement, laissant ainsi une fonction à évaluer qui ne dépendra que des variables classiques commutantes $\{x_{i,\ell}\}$. Pour chaque configuration, la dernière égalité ci-dessus définit l'opérateur d'évolution du temps imaginaire 0 jusqu'au temps imaginaire $\beta = N_\tau \Delta \tau$. En général, nous aurons

$$\hat{U}(\ell_1, \ell_2) \equiv \hat{D}_{\ell_1} \dots \hat{D}_{\ell_2+1} \tag{6a}$$

pour effectuer l'évolution du temps imaginaire ℓ_2 au temps imaginaire ℓ_1 . L'égalité du milieu dans l'équation (6) définit l'opérateur d'évolution du temps imaginaire $\ell - 1$ au temps imaginaire ℓ pour la ℓ ième tranche de temps imaginaire d'épaisseur $\Delta \tau$:

$$\hat{D}_\ell^\sigma(\{x_{i,\ell}\}) \equiv e^{-\Delta \tau c_{i,\sigma}^+ K_{i,j} c_{j,\sigma}} e^{c_{i,\sigma}^+ (V_{i,j})_\ell^\sigma c_{j,\sigma}} . \tag{6b}$$

Dans cette dernière équation on a défini

$K_{i,j} = -t$ lorsque les sites i et j sont premiers voisins et $K_{i,j} = 0$ autrement,

$$(V_{i,j})_{\ell}^{\sigma} = \delta_{i,j} \left[\sigma \lambda x_{i,\ell} + \Delta \tau \left(\mu - \frac{U}{2} \right) \right] \quad (7)$$

On peut donc calculer la trace si on connaît le résultat pour

$$\text{Tr} \left[e^{-\sum_{i,j} c_i^{\dagger} A_{i,j} c_j - \sum_{m,n} c_m^{\dagger} B_{m,n} c_n} \right].$$

Ce calcul est souvent fait avec des variables de Grassmann. Il est cependant possible d'évaluer cette trace avec des méthodes purement Hamiltoniennes, ce que nous ferons dans ce qui suit, en suivant l'approche des Réfs. 4,6

Notons tout d'abord le résultat plus simple,

$$\text{Tr} \left[e^{-\sum_{i,j} c_i^{\dagger} A_{i,j} c_j} \right] = \text{Tr} \left[e^{-\sum_{\alpha} c_{\alpha}^{\dagger} a_{\alpha} c_{\alpha}} \right] = \text{Tr} \left[\prod_{\alpha=1}^N e^{-c_{\alpha}^{\dagger} a_{\alpha} c_{\alpha}} \right].$$

La première égalité est obtenue en passant dans la base où la matrice hermitienne \mathbf{A} est diagonale. En notation matricielle,

$$\mathbf{c}^{\dagger} \mathbf{A} \mathbf{c} = (\mathbf{c}^{\dagger} \mathbf{T}^{\dagger}) \mathbf{T} \mathbf{A} \mathbf{T} (\mathbf{T} \mathbf{c}) = \tilde{\mathbf{c}}^{\dagger} \tilde{\mathbf{A}} \tilde{\mathbf{c}}$$

où $\tilde{\mathbf{A}}$ est diagonale avec éléments a_{α} et où les $\tilde{\mathbf{c}}^{\dagger}$ et $\tilde{\mathbf{c}}$ obéissent aux relations de commutation canoniques parce que la transformation \mathbf{T} est unitaire. La dernière trace peut être évaluée facilement dans la base diagonale parce que chaque terme du produit n'agit que sur un seul état propre à une particule qui ne peut être que vide ou occupé. La trace prend donc la forme

$$\prod_{\alpha=1}^N [1 + e^{-a_{\alpha}}] = \det[\mathbf{I} + e^{-\mathbf{A}}]$$

où on reconnaît dans le membre de gauche un résultat obtenu dans les cours de mécanique statistique élémentaire. La dernière égalité exprime le résultat sous une forme qui est indépendante de la base choisie. Notons que lorsqu'on développe le produit dans le membre de gauche,

$$1 + e^{-a_1} + e^{-a_2} + \dots + e^{-a_N} + e^{-a_1} e^{-a_2} + e^{-a_1} e^{-a_3} + \dots$$

on obtient 2^N termes correspondant au nombre d'états attendu dans l'ensemble grand canonique pour N sites qui peuvent être soit vides soit occupés par un fermion de spin donné (il y a un produit de cette forme pour les spins up et un autre pour les spins down, donnant les 4^N états attendus). Dans la somme, on reconnaît la contribution venant d'un état avec aucun fermion, celle venant des N états avec un seul fermion, celle des $N(N - 1) / 2$ états avec deux fermions, *etc.*

Pour évaluer la trace qui nous intéresse dans le cas général, il suffit de considérer l'opérateur

$$e^{-\sum_{i,j} c_i^\dagger A_{i,j} c_j - \sum_{m,n} c_m^\dagger B_{m,n} c_n}$$

et de noter que dans ce cas les états propres gauche de la matrice généralement non-hermitique ($e^{-A} e^{-B}$) jouent le même rôle que ceux de la matrice e^{-A} dans le cas précédent. Une preuve complète apparaît dans l'annexe D. On a donc finalement que

$$Z = \sum_{\{x_{i,\ell}\}} \prod_{\sigma = \pm 1} \det[\mathbf{I} + \mathbf{B}_{N_\tau}^\sigma \mathbf{B}_{N_\tau-1}^\sigma \dots \mathbf{B}_1^\sigma] \quad (8)$$

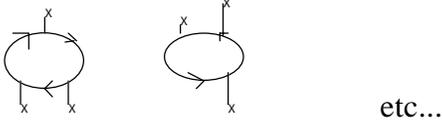
où les matrices \mathbf{B}_ℓ^σ de taille $N \times N$ sont définies par (7) et par

$$\mathbf{B}_\ell^\sigma \equiv e^{-\Delta \tau \mathbf{K}} e^{\mathbf{V}_\ell^\sigma} \quad (9)$$

3.4. Interprétation physique.

Cette section contient du matériel plus avancé mais elle n'est pas vraiment nécessaire pour la lecture du reste de ces notes. Seule la définition (10) sera utilisée dans le texte principal. Les preuves ne sont qu'esquissées.

On remarque que le problème équivalent que nous avons maintenant à résoudre est celui d'électrons sans interaction se déplaçant dans un champ magnétique pointant dans la direction z et dont la valeur à chaque site i, ℓ du réseau espace-temps est déterminée par le $x_{i,\ell}$ correspondant. On moyenne ensuite sur toutes les valeurs de $x_{i,\ell}$ possibles. Pour ceux qui sont familiers avec le langage diagrammatique, on peut noter que pour une configuration donnée des $x_{i,\ell}$ on résout exactement le problème suivant:



qui, après moyennage (trace) sur les champs aléatoires, redonne la série originale du Hamiltonien de Hubbard,



Nous notons que l'expression (8) pour la fonction de partition se réécrit aussi sous la forme,

$$Z = \sum_{\{x_{i,\ell}\}} \prod_{\sigma = \pm 1} \det \left[\left(G^\sigma(\{x_{i,\ell}\}) \right)^{-1} \right] = \sum_{\{x_{i,\ell}\}} \prod_{\sigma = \pm 1} e^{\text{Tr}[\ln(S^\sigma(\{x_{i,\ell}\}))]}$$

où la matrice G^σ de taille $(N_\tau N) \times (N_\tau N)$ et la matrice S^σ sont définies par

$$G^\sigma \equiv \begin{bmatrix} \mathbf{I} & 0 & 0 & \dots & 0 & B_{N_\tau}^\sigma \\ -B_1^\sigma & \mathbf{I} & 0 & \dots & 0 & 0 \\ 0 & -B_2^\sigma & \mathbf{I} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \mathbf{I} & 0 \\ 0 & 0 & 0 & \dots & -B_{N_\tau-1}^\sigma & \mathbf{I} \end{bmatrix}^{-1} \equiv (S^\sigma)^{-1} \quad (10)$$

G^σ joue le rôle de propagateur dans l'espace-temps pour le problème d'électrons sans interaction (*i.e.* dans une configuration $\{x_{i,\ell}\}$ donnée). En effet, définissant

$$c_{i,\sigma}(\ell_1) \equiv e^{\ell_1 \Delta \tau (\hat{H} - \mu N)} c_{i,\sigma} e^{-\ell_1 \Delta \tau (\hat{H} - \mu N)} \quad (11a)$$

et

$$c_{i,\sigma}^+(\ell_1) \equiv e^{\ell_1 \Delta \tau (\hat{H} - \mu N)} c_{i,\sigma}^+ e^{-\ell_1 \Delta \tau (\hat{H} - \mu N)} \quad (11b)$$

(qui n'est pas l'adjoint de l'opérateur précédent) on démontre en annexe F que pour le problème à un électron et $\ell_1 > \ell_2$ on a,

$$\langle c_{i,\sigma}(\ell_1)c_{j,\sigma}^+(\ell_2) \rangle_{\{x\}}^0 = \left[\frac{1}{\mathbf{I} + \mathbf{B}_{\ell_1} \dots \mathbf{B}_1 \mathbf{B}_{N_\tau} \dots \mathbf{B}_{\ell_1+1}} \mathbf{B}_{\ell_1} \dots \mathbf{B}_{\ell_2+1} \right]_{i,j}$$

ce qui est un élément de la matrice définie ci-dessus par l'équation (10). Ceci permet de faire l'identification suivante,

$$G^\sigma(i, \ell_1; j, \ell_2) \equiv \langle c_{i,\sigma}(\ell_1)c_{j,\sigma}^+(\ell_2) \rangle_{\{x\}}^0$$

pour $\ell_1 > \ell_2$. La notation $\langle \rangle_{\{x\}}^0$ rappelle que la moyenne thermodynamique se fait pour des électrons indépendants se déplaçant dans le champ $\{x_{i,\ell}\}$. Comme nous le verrons dans les sections suivantes, toutes les observables pourront être exprimées (théorème de Wick) en termes de ces fonctions de Green, ou propagateurs.

3.5. Calcul des observables.

Utilisant l'expression (6) pour la fonction de partition après la transformation de Hubbard Stratonovich, la valeur moyenne d'une observable peut s'écrire sous la forme,

$$\langle \hat{O} \rangle = \frac{\sum_{\{x_{i,\ell}\}} Tr \left[\hat{O} \prod_{\sigma = \pm 1} \hat{U}^\sigma(N_\tau, 0; \{x_{i,\ell}\}) \right]}{\sum_{\{x_{i,\ell}\}} Tr \left[\prod_{\sigma = \pm 1} \hat{U}^\sigma(N_\tau, 0; \{x_{i,\ell}\}) \right]}$$

qu'on peut trivialement réarranger comme suit,

$$\langle \hat{O} \rangle = \frac{\sum_{\{x_{i,\ell}\}} \frac{Tr \left[\hat{O} \prod_{\sigma = \pm 1} \hat{U}^\sigma(N_\tau, 0; \{x_{i,\ell}\}) \right]}{Tr \left[\prod_{\sigma = \pm 1} \hat{U}^\sigma(N_\tau, 0; \{x_{i,\ell}\}) \right]}}{\sum_{\{x_{i,\ell}\}} Tr \left[\prod_{\sigma = \pm 1} \hat{U}^\sigma(N_\tau, 0; \{x_{i,\ell}\}) \right]} Tr \left[\prod_{\sigma = \pm 1} \hat{U}^\sigma(N_\tau, 0; \{x_{i,\ell}\}) \right]$$

Utilisant l'expression (8) trouvée précédemment pour la fonction de partition, l'expression précédente devient

$$\langle \hat{O} \rangle = \frac{\sum_{\{x_i, \ell\}} \langle \hat{O} \rangle_{\{x\}}^0 \prod_{\sigma = \pm 1} \det[\mathbf{I} + \mathbf{B}_{N_\tau}^\sigma \mathbf{B}_{N_\tau - 1}^\sigma \dots \mathbf{B}_1^\sigma]}{\sum_{\{x_i, \ell\}} \prod_{\sigma = \pm 1} \det[\mathbf{I} + \mathbf{B}_{N_\tau}^\sigma \mathbf{B}_{N_\tau - 1}^\sigma \dots \mathbf{B}_1^\sigma]} \quad (12)$$

où la définition de la moyenne $\langle \rangle_{\{x\}}^0$ apparaît clairement lorsqu'on compare les deux équations précédentes. Cette notation rappelle qu'il s'agit d'évaluer la valeur moyenne de l'opérateur \hat{O} pour des électrons indépendants se déplaçant dans le champ de Hubbard-Stratonovich $\{x_{i,\ell}\}$. Ce type de valeurs moyennes est facile à évaluer en utilisant le théorème de Wick, que nous redémontrons explicitement plus bas.

Notons d'abord que la quantité

$$\prod_{\sigma = \pm 1} \det[\mathbf{I} + \mathbf{B}_{N_\tau}^\sigma \mathbf{B}_{N_\tau - 1}^\sigma \dots \mathbf{B}_1^\sigma]$$

joue maintenant le même rôle que le poids de Boltzmann dans les simulations Monte Carlo pour les problèmes classiques. Cette quantité n'est pas toujours positive, tel que discuté dans la sections suivante.

Pour ce qui est des observables, considérons d'abord l'exemple suivant (où on laisse tomber les variables de spin pour simplifier la notation puisque, dans la jauge choisie ici, les valeurs moyennes pour chaque type de spin factorisent) :

$$\langle c_i c_j^+ \rangle_{\{x\}}^0 = \frac{\text{Tr}[c_i c_j^+ \hat{D}_{N_\tau} \hat{D}_{N_\tau - 1} \dots \hat{D}_1]}{\text{Tr}[\hat{D}_{N_\tau} \hat{D}_{N_\tau - 1} \dots \hat{D}_1]}.$$

Pour évaluer cette quantité dans le cas général, il suffit de travailler dans la base des états propres gauche du produit de matrices $\mathbf{B}_{N_\tau} \mathbf{B}_{N_\tau - 1} \dots \mathbf{B}_1$, suivant le type de preuves faites en annexe D. Afin de simplifier la présentation, nous considérons le cas où un seul opérateur d'évolution à une particule $e^{-c^+ A c}$ intervient à la place des \hat{D}_i . Dans ce cas, passant à la base des états propres de \mathbf{A} , on écrit

$$\langle c_i c_j^+ \rangle_{\{x\}}^0 = \frac{\text{Tr} \left[c_i c_j^+ \prod_{\alpha=1}^N e^{-c_\alpha^+ a_\alpha c_\alpha} \right]}{\text{Tr} \left[\prod_{\alpha=1}^N e^{-c_\alpha^+ a_\alpha c_\alpha} \right]}.$$

En développant les opérateurs sur cette même base d'états propres,

$$c_i = \sum_{\alpha} \langle i | \alpha \rangle c_{\alpha}$$

$$c_i^+ = \sum_{\alpha} \langle \alpha | i \rangle c_{\alpha}^+$$

on trouve que

$$\langle c_i c_j^+ \rangle_{\{x\}}^0 = \sum_{\beta} \langle i | \beta \rangle \langle \beta | j \rangle \frac{\text{Tr} \left[c_{\beta} c_{\beta}^+ \prod_{\alpha=1}^N e^{-c_{\alpha}^+ a_{\alpha} c_{\alpha}} \right]}{\text{Tr} \left[\prod_{\alpha=1}^N e^{-c_{\alpha}^+ a_{\alpha} c_{\alpha}} \right]}$$

où on a utilisé le fait que $\text{Tr} \left[c_{\beta} c_{\gamma}^+ \prod_{\alpha=1}^N e^{-c_{\alpha}^+ a_{\alpha} c_{\alpha}} \right]$ s'annule lorsque $\beta \neq \gamma$. Évaluant la trace dans cette même base d'états propres, tous les termes du numérateur et du dénominateur se simplifient, sauf celui où β apparaît. L'opérateur $c_{\beta} c_{\beta}^+$ implique que seul l'état où β est inoccupé contribue à la trace au numérateur. Il reste donc,

$$\langle c_i c_j^+ \rangle_{\{x\}}^0 = \sum_{\beta} \langle i | \beta \rangle \langle \beta | j \rangle \frac{1}{1 + e^{-a_{\beta}}} = \langle i | \frac{1}{1 + e^{-\Lambda}} | j \rangle \equiv \left[\frac{1}{1 + e^{-\Lambda}} \right]_{i,j}$$

et dans le cas général qui nous intéresse ici,

$$\langle c_i c_j^+ \rangle_{\{x\}}^0 = \left[\frac{\mathbf{I}}{\mathbf{I} + \mathbf{B}_{N_r} \mathbf{B}_{N_r-1} \cdots \mathbf{B}_1} \right]_{i,j} \quad (13)$$

comme nous le prouverons en plus de détails dans l'annexe E.

Pour une observable typique, nous aurons à évaluer des quantités comme $\langle c_1^+ c_2 c_3^+ c_4 \rangle_{\{x\}}^0$. En répétant le même genre de preuve que ci-dessus, on voit que dans le développement en fonctions propres les indices de chaque opérateur de destruction doivent être égaux à ceux d'un ou l'autre des opérateurs de création. Il y a deux façons d'accomplir cela. Pour chacun de ces cas, il n'y a qu'une somme sur deux indices qui reste et dans les traces il y a alors deux termes qui ne se

simplifient pas au numérateur et au dénominateur, sauf dans le cas où quatre indices sont égaux. Dans ce dernier cas, il n'y a qu'un terme qui ne se simplifie pas au numérateur et au dénominateur, mais on peut vérifier que le résultat pour quatre indices égaux a la même forme que lorsqu'on met les indices égaux dans la somme des contributions calculées pour le cas où les indices étaient différents. Retournant à la base originale, on trouve alors

$$\langle c_1^+ c_2 c_3^+ c_4 \rangle_{\{x\}}^0 = \langle c_1^+ c_4 \rangle_{\{x\}}^0 \langle c_2 c_3^+ \rangle_{\{x\}}^0 + \langle c_1^+ c_2 \rangle_{\{x\}}^0 \langle c_3^+ c_4 \rangle_{\{x\}}^0$$

ce qui n'est rien d'autre que le théorème de Wick pour des électrons indépendants. En général, toutes les paires d'opérateurs de création et d'annihilation doivent apparaître à droite et le signe de chaque contribution dépend du nombre de permutations d'opérateurs fermioniques.

3.6. Poids pour l'échantillonnage selon l'importance: le "signe du déterminant".

Tel que mentionné à la section précédente, la quantité $\prod_{\sigma = \pm 1} \det[\mathbf{I} + \mathbf{B}_{N_\tau}^\sigma \mathbf{B}_{N_\tau - 1}^\sigma \dots \mathbf{B}_1^\sigma]$

joue le rôle d'un poids de Boltzmann dans le calcul de la trace sur les variables de Hubbard-Stratonovich. Pour appliquer l'algorithme de Métropolis, du bain de chaleur ou tout autre algorithme, il faut que le poids soit positif. On utilise donc la valeur absolue de la quantité précédente,

$$P(\{x_{i,\ell}\}) = \left| \prod_{\sigma = \pm 1} \det[\mathbf{I} + \mathbf{B}_{N_\tau}^\sigma \mathbf{B}_{N_\tau - 1}^\sigma \dots \mathbf{B}_1^\sigma] \right| = \left| \prod_{\sigma = \pm 1} \det[(G^\sigma)^{-1}] \right|$$

comme poids de Boltzmann et on tient compte du signe de $\prod_{\sigma = \pm 1} \det[(G^\sigma)^{-1}]$ en l'incluant dans

l'observable. Bien que cette approche ne soit pas une approximation, il devient difficile d'obtenir des bonnes statistiques lorsque $\prod_{\sigma = \pm 1} \det[(G^\sigma)^{-1}]$ devient trop souvent négatif.¹⁷ Dans l'algorithme

du bain de chaleur, une nouvelle configuration est générée en faisant par exemple le changement $x_{i,\ell} \rightarrow -x_{i,\ell}$ sur un site i, ℓ du réseau espace-temps. Ce changement est accepté avec une probabilité

$$P(x_{i,\ell} \rightarrow -x_{i,\ell}) = \frac{|R_\uparrow R_\downarrow|}{1 + |R_\uparrow R_\downarrow|}$$

où

$$R_\sigma = \frac{\det[(G^{\sigma'})^{-1}]}{\det[(G^\sigma)^{-1}]}$$

la quantité $G^{\sigma'}$ référant à la nouvelle configuration et la quantité G^σ à la précédente.

On définit

$$s(\{x_{i,\ell}\}) = 1 \times \operatorname{sgn} \left(\prod_{\sigma = \pm 1} \det[(G^\sigma)^{-1}] \right)$$

comme le "signe" (sgn) associé à une configuration donnée des champs de Hubbard-Stratonovich. En gardant seulement les configurations Monte Carlo $\{x_{MC}\}$ générées par l'algorithme ci-dessus, (échantillonnage selon l'importance) on trouve, à partir de l'expression (12) pour la valeur moyenne, que

$$\langle \hat{O} \rangle = \frac{\sum_{\{x\}} \langle \hat{O} \rangle_{\{x\}}^0 s(\{x_{i,\ell}\})}{\sum_{\{x_{MC}\}} s(\{x_{i,\ell}\})}. \quad (14)$$

Évidemment, lorsque la "moyenne du signe" est petite, le dénominateur cause des problèmes!

4. Les calculs en pratique.

Pour rendre les calculs possibles en pratique, il y a plusieurs trucs à utiliser. Ceux décrits dans la section 4.1 sont dûs à BSS³ alors que ceux de la section 4.2 sont plus récents.⁵

4.1. Un cycle de mise-à-jour.

Les sections précédentes montrent que la matrice G^σ , de taille $(N_\tau N) \times (N_\tau N)$, est nécessaire à la fois pour calculer les quantité observables et aussi pour obtenir le poids pour l'échantillonnage selon l'importance. Avant de décrire comment une nouvelle configuration des champs de Hubbard-Stratonovich $\{x_{i,\ell}\}$ est obtenue à partir d'une autre en balayant un après l'autre les sites i, ℓ du réseau espace-temps, (un cycle de mise-à-jour), il est utile de prouver certaines

identités qui permettent d'obtenir facilement comment la matrice G^σ change lorsqu'un champ de Hubbard-Stratonovich est modifié sur un site du réseau. Nous suivons l'approche de Fye et Hirsch¹⁸ qui est elle-même inspirée des formules de Morrison-Woodbury-Shermann¹⁹ utilisées en analyse numérique.

Se souvenant des équations (9) et (10), définissons d'abord la matrice

$$g^\sigma = \left[(G^\sigma)^{-1} e^{-V^\sigma} \right]^{-1} = \begin{bmatrix} e^{-V_1^\sigma} & 0 & 0 & \dots & 0 & e^{-\Delta \tau \mathbf{K}} \\ -e^{-\Delta \tau \mathbf{K}} & e^{-V_2^\sigma} & 0 & \dots & 0 & 0 \\ 0 & -e^{-\Delta \tau \mathbf{K}} & e^{-V_3^\sigma} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & e^{-V_{N-1}^\sigma} & 0 \\ 0 & 0 & 0 & \dots & -e^{-\Delta \tau \mathbf{K}} & e^{-V_N^\sigma} \end{bmatrix}^{-1} = e^{V^\sigma} G^\sigma.$$

L'avantage de cette définition est que les champs de Hubbard-Stratonovich n'apparaissent que sur la diagonale de la matrice. Comme dans les sections précédentes, les matrices de taille $N \times N$ sont écrites en caractère gras. Utilisant l'identité matricielle générale

$$\frac{1}{X + Y} = \frac{1}{X} - \frac{1}{X} Y \frac{1}{X + Y} \quad (15)$$

on peut donc dire que si on connaît la matrice g^σ pour une configuration des $\{x_{i,\ell}\}$ donnée, alors pour une configuration différente $\{x'_{i,\ell}\}$ la matrice g'^σ correspondante sera donnée par l'expression suivante, où on laisse tomber l'indice de spin σ pour alléger la notation:

$$g' = g - g \left[e^{-V'} - e^{-V} \right] g'.$$

Retournant à la matrice originale G , et utilisant le fait que e^V est diagonale, on obtient,

$$G' = (e^{-V'+V}) G - (e^{-V'+V}) G [1 - e^{-V+V'}] G'. \quad (16)$$

Bien qu'une modification des $\{x_{i,\ell}\}$ sur un site modifie tous les éléments de matrice de la fonction de Green du réseau espace-temps, on remarque, à partir de l'équation précédente, que les G' ayant les deux indices de temps identiques ne dépendent que des G ayant aussi leurs deux indices de temps identiques. Ceci est une conséquence du fait que lorsqu'une seule variable de Hubbard-Stratonovich a été modifiée, alors la matrice $[1 - e^{-V+V'}]$ est nulle partout, sauf pour un élément sur la diagonale faisant intervenir le même indice de temps que celui de la variable $x_{i,\ell}$ modifiée.

Résolvant la dernière équation pour G' , un peu d'algèbre matricielle élémentaire permet d'écrire le résultat sous la forme

$$G' = \frac{1}{1 + (G - 1)[1 - e^{V' - V}]} G \quad (17)$$

que nous utiliserons ci-dessous. Une autre forme de cette expression, qui sera celle utilisée en pratique, est obtenue en manipulant les expressions pour faire apparaître le plus possible la combinaison $[1 - e^{-V + V'}]$ à gauche et à droite des matrices. Ceci est utile parce que cette matrice n'a qu'un seul élément non nul. Partant de la dernière expression (17), on multiplie le membre de droite par la gauche à l'aide de la matrice identité écrite sous la forme

$$1 + (G - 1)[1 - e^{V' - V}] - (G - 1)[1 - e^{V' - V}]$$

ce qui nous donne le résultat désiré,

$$G' = G - (G - 1) \left\{ [1 - e^{V' - V}] \frac{1}{1 + (G - 1)[1 - e^{V' - V}]} \right\} G. \quad (18)$$

Nous avons aussi besoin de connaître

$$R = \frac{\det[(G')^{-1}]}{\det[(G)^{-1}]} = \frac{\det[G]}{\det[G']}.$$

Utilisant le résultat (17) pour G' en fonction de G ainsi que $\det(AB) = \det(A) \det(B)$, on trouve que

$$R = \det[1 + (G - 1)[1 - e^{V' - V}]]. \quad (19)$$

Cette dernière expression peut se simplifier de la façon suivante: supposons qu'une seule variable de Hubbard-Stratonovich ait été modifiée. Alors, tel qu'expliqué di-dessus, la matrice $[1 - e^{-V + V'}]$ est nulle partout, sauf pour un élément sur la diagonale faisant intervenir le même indice de temps et d'espace que ceux de la variable $x_{i,\ell}$ modifiée. Ceci veut dire que la matrice $(G - 1)[1 - e^{-V + V'}]$ est zéro partout sauf dans une seule colonne, ce qui implique que le déterminant ne peut être que linéaire en $[1 - e^{-V + V'}]_{i,\ell;i,\ell}$ qui est le seul élément non nul de la

matrice $[1 - e^{-V + V'}]$. Pour simplifier la notation, nous introduisons les définitions suivantes: $W \equiv [1 - e^{-V + V'}]$ pour la matrice, et $w \equiv [1 - e^{-V + V'}]_{i,\ell;i,\ell}$ pour le seul élément non nul de la matrice W . Comme le déterminant est linéaire en w , on peut écrire,

$$\begin{aligned}
R &= 1 + \left. \frac{\partial (\det[1 + (G - 1)W])}{\partial w} \right|_{w=0} w \\
&= 1 + \left. \frac{\partial (e^{Tr[\ln(1 + (G - 1)W)])}}{\partial w} \right|_{w=0} w \\
&= 1 + e^{Tr[\ln(1 + (G - 1)W)]} \left. \frac{\partial (Tr[\ln(1 + (G - 1)W)])}{\partial w} \right|_{w=0} w \\
&= 1 + Tr \left[\left. \frac{1}{1 + (G - 1)W} \right|_{w=0} (G - 1) \left. \frac{\partial W}{\partial w} \right|_{w=0} \right] w
\end{aligned}$$

et donc, finalement, comme $\partial W / \partial w$ n'a qu'un élément non nul et que c'est la trace qui doit être calculée, il ne reste que

$$R = 1 + [G - 1]_{i,\ell;i,\ell} [1 - e^{-V + V'}]_{i,\ell;i,\ell}. \quad (20)$$

Nous pouvons maintenant procéder à la description d'une mise-à-jour.

Initialisation:

D'abord, la matrice \mathbf{K} est initialisée, puis les variables de Hubbard-Stratonovich sont choisies de façon aléatoire sur chaque site. Ceci permet d'initialiser les matrices \mathbf{V}_ℓ^σ puis de calculer la fonction de Green au temps $\ell = 0$, c'est-à-dire, comme à l'équation (13),

$$\mathbf{G}_{i,j}^\sigma(0,0) \equiv G^\sigma(i,0; j,0^-) = \left[\frac{\mathbf{I}}{\mathbf{I} + \mathbf{B}_{N_\tau}^\sigma \mathbf{B}_{N_\tau-1}^\sigma \cdots \mathbf{B}_1^\sigma} \right]_{i,j} \quad (21)$$

Génération des configurations de $\{x_{i,\ell}\}$ pour l'échantillonnage selon l'importance:

A. Balayant les sites l'un après l'autre dans une tranche de temps ℓ donnée, un changement $x_{i,\ell} \rightarrow -x_{i,\ell}$ est effectué. Ce changement est accepté avec une probabilité

$$P(x_{i,\ell} \rightarrow -x_{i,\ell}) = \frac{|R_{\uparrow} R_{\downarrow}|}{1 + |R_{\uparrow} R_{\downarrow}|}$$

où en utilisant les résultats (20) ci-haut,

$$R_{\sigma} = \frac{\det[(G^{\sigma})^{-1}]}{\det[(G^{\sigma})^{-1}]} = 1 + (1 - \mathbf{G}_{i,i}^{\sigma}(\ell, \ell)) \left[e^{(\mathbf{v}_{i,i}^{\sigma})_{\ell}^{\sigma} - (\mathbf{v}_{i,i}^{\sigma})_{\ell}^{\sigma}} - 1 \right].$$

Ce calcul est très rapide puisqu'il n'implique pas de multiplication de matrice.

B. Si ce changement n'est pas accepté, on passe au site suivant dans la même tranche de temps. Si ce changement est accepté, $(\mathbf{V}_{i,i}^{\sigma})_{\ell}^{\sigma}$ est mise-à-jour et la fonction de Green doit être recalculée avant de passer au site suivant. Comme nous n'avons besoin que des éléments diagonaux en indice de tranche de temps pour calculer le rapport de déterminants ci-haut, on a besoin que de $\mathbf{G}_{i,i}^{\sigma}(\ell, \ell)$. Ainsi, la formule pour la mise-à-jour de la fonction de Green (17) n'implique plus que la multiplication de matrices de taille $N \times N$, comme nous l'avons déjà noté. Pour être plus explicite, l'expression à utiliser est obtenue de l'expression (18) ci-dessus réécrite pour le cas particulier qui nous intéresse, soit

$$\mathbf{G}_{m,n}^{\prime\sigma} = \mathbf{G}_{m,n}^{\sigma} + (\mathbf{G}_{m,i}^{\sigma} - \mathbf{1}_{m,i}) \left\{ \left[e^{(\mathbf{v}_{i,i}^{\sigma})_{\ell}^{\sigma} - (\mathbf{v}_{i,i}^{\sigma})_{\ell}^{\sigma}} - 1 \right] \frac{1}{1 + (\mathbf{G}_{i,i}^{\sigma} - \mathbf{1}_{i,i}) \left[1 - e^{(\mathbf{v}_{i,i}^{\sigma})_{\ell}^{\sigma} - (\mathbf{v}_{i,i}^{\sigma})_{\ell}^{\sigma}} \right]} \right\} \mathbf{G}_{i,n}^{\sigma} \quad (22)$$

où il est sous-entendu que toutes les fonctions de Green font référence à la propagation à l'intérieur de la même tranche de temps ℓ . Il n'y a pas de sommation sur l'indice répété i car il s'agit du site où la variable Hubbard-Stratonovich vient d'être modifiée. Les multiplications des matrices apparaissant dans l'accolade $\{ \}$ de l'expression originale (18) sont réduites à un seul terme parce que la matrice $[1 - e^{-v + v'}]$ n'a qu'un élément non-nul. Pour se convaincre qu'un seul élément est nécessaire dans le calcul de la matrice inverse $[1 + (G - 1)[1 - e^{v' - v}]]^{-1}$, il suffit d'utiliser le développement en série généré par l'expression (15) avec $X = 1$ ce qui montre que, dans tous les termes, $(G - 1)$ est entouré à gauche et à droite par la matrice $[1 - e^{-v + v'}]$ qui sélectionne un seul élément diagonal de $(G - 1)$. Il résulte une série de Taylor scalaire qui peut être resommée pour obtenir la forme ci-dessus. La quantité entre accolade joue le rôle de la matrice t exacte pour la diffusion sur la seule "impureté" représentée par le champ de Hubbard-Stratonovich modifié au site i, ℓ . Malgré toutes ces simplifications, la mise-à-jour de la fonction de Green implique le calcul de N^2 termes.

C. Les étapes **A** et **B** sont répétées pour tous les points du réseau dans une tranche de temps donnée.

D. Pour passer à la tranche de temps suivante, on a besoin de la fonction de Green à la tranche de temps $\ell + 1$. La fonction de Green au temps zéro est donnée par (21). La généralisation de ce résultat à un temps quelconque est suggérée par la relation

$$Z = \sum_{\{x_{i,\ell}\}} Tr \left[\prod_{\sigma = \pm 1} \hat{U}^\sigma(\ell, 0; \{x_{i,\ell}\}) \hat{U}^\sigma(N_\tau, \ell; \{x_{i,\ell}\}) \right]$$

qui est une conséquence de la propriété cyclique de la trace. En effet, ayant vu ce résultat, où ℓ joue maintenant le rôle que 0 jouait auparavant, le lecteur ne sera pas surpris d'apprendre que

$$\mathbf{G}_{i,j}^\sigma(\ell, \ell) \equiv G^\sigma(i, \ell; j, \ell^-) = \left[\frac{\mathbf{I}}{\mathbf{I} + \mathbf{B}_\ell^\sigma \dots \mathbf{B}_1^\sigma \mathbf{B}_{N_\tau}^\sigma \dots \mathbf{B}_{\ell+1}^\sigma} \right]_{i,j} \quad (23a)$$

un résultat qui est prouvé de façon plus rigoureuse dans l'annexe E. Le calcul explicite de $\mathbf{G}_{i,j}^\sigma(\ell, \ell)$ utilisant cette formule est nécessaire à chaque fois qu'un certain nombre de tranches de temps $\bar{\ell}$ a été mis à jour à l'aide de la formule

$$\mathbf{G}^\sigma(\ell + 1, \ell + 1) = \mathbf{B}_{\ell+1}^\sigma \mathbf{G}^\sigma(\ell, \ell) \left(\mathbf{B}_{\ell+1}^\sigma \right)^{-1} \quad (23b)$$

qui est une façon beaucoup plus rapide, mais aussi beaucoup moins stable numériquement, d'obtenir la fonction de Green pour la tranche de temps suivante à partir des matrices \mathbf{B}_ℓ^σ et $\mathbf{G}^\sigma(\ell, \ell)$ qui viennent d'être calculées. On réfère parfois à la formule (23b) en disant que la fonction de Green est "enroulée" (du terme anglais *wrap-up*). Le nombre de tranches de temps $\bar{\ell}$ après lequel il faut recalculer au long la fonction de Green à partir de sa définition est donné en gros par la valeur empirique $\bar{\ell} \Delta \tau \approx 1/t$, où t est l'intégrale de saut du Hamiltonien de Hubbard.

Calcul des observables:

Comme dans toutes les simulations Monte Carlo, on doit laisser le système thermaliser en effectuant plusieurs des cycles de mise-à-jour décrits ci-dessus avant même de commencer à calculer les observables. Le nombre de cycles de thermalisation ("réchauffements") nécessaires dépend de la température, de la valeur des interactions et aussi de la précision avec laquelle on désire connaître une observable donnée. Il faut en principe vérifier qu'il n'y a plus de changements

systématiques des observables pour être certain qu'il y a eu thermalisation. Par changement systématique, on veut dire un changement en fonction de nombre de "mesures" qui est significativement plus grand que l'erreur statistique. En pratique, on note qu'en redémarrant la simulation à partir d'une configuration aléatoire différente des champs de Hubbard-Stratonovich, le résultat obtenu après le même nombre de mesures ne sera à l'intérieur des barres d'erreur, entre les différentes simulations, que si on a laissé le système suffisamment thermaliser avant de commencer à mesurer.

Après avoir laissé le système thermaliser, on peut commencer à évaluer les observables $\langle \hat{O} \rangle_{\{x\}}^\sigma$ qui seront ensuite moyennées sur les différentes configurations à l'aide de la formule (14).

Dans le cas où les observables sont évaluées à temps égal, les fonctions de Green sont utilisées à la fin de l'étape *C*, c'est-à-dire après la mise-à-jour de tous les sites d'une tranche de temps donnée, pour calculer les observables d'intérêt (incluant le fameux signe du déterminant). Les résultats sont accumulés pour obtenir éventuellement la moyenne (14) et l'écart-quadratique moyen. On réfère au calcul d'une observable après la mise-à-jour d'une tranche de temps en disant qu'on fait une "mesure".

L'évaluation de l'incertitude à l'aide de l'écart type pose quelques problèmes. En effet, pour pouvoir estimer l'erreur sur la moyenne, il faut que les données soient statistiquement indépendantes. Or ici chaque donnée, représentée par une mesure, est assez corrélée avec la précédente puisque les $x_{i,\ell}$ sont obtenus des $x_{i,\ell-1}$ à l'aide de seulement quelques changements locaux. Il faut donc calculer l'écart quadratique moyen de quantités qui sont plus statistiquement indépendantes que ne le sont les mesures individuelles entre elles. Pour obtenir l'indépendance statistique, on regroupe normalement un certain nombre de mesures (un "bloc" de mesures) selon la formule (14), ce qui nous donne un estimé du résultat, puis on calcule les fluctuations de ces différents estimés en supposant qu'ils sont statistiquement indépendants. L'erreur sur la moyenne est obtenue en divisant l'écart quadratique moyen par \sqrt{M} où M est le nombre d'estimés (blocs de mesures) utilisé pour évaluer les fluctuations. Si le nombre de mesures dans un bloc est trop petit, les différents blocs ne seront pas statistiquement indépendants et la division par \sqrt{M} donnera un estimé trop petit de l'erreur statistique. Par contre, si le nombre de mesures dans un bloc est comparable au nombre total de mesures, M sera très petit et notre estimé de l'erreur sera entaché d'une erreur statistique importante. Une des stratégies acceptables consiste à prendre M de l'ordre de 100 disons. Certains chercheurs dans ce domaine prennent $M = 10$ (A. Moreo par exemple). En général, il faut varier M ou calculer une fonction d'autocorrélation pour s'assurer que les mesures sont statistiquement indépendantes.

Remarques diverses:

L'erreur systématique due à la discrétisation $\Delta \tau$ peut être éliminée par extrapolation jusqu'à $\Delta \tau = 0$. Il suffit de faire un graphique de l'observable qu'on désire en fonction de $(\Delta \tau)^2$, et d'extrapoler la valeur à $(\Delta \tau)^2 = 0$ en utilisant une ligne droite. En pratique cela marche très bien comme nous le verrons plus loin: les corrections d'ordre plus élevé que $(\Delta \tau)^2$ sont généralement négligeables pour les valeurs de $\Delta \tau$ utilisées. Lorsque l'erreur statistique recherchée sur une observable est de l'ordre de 5%, $\Delta \tau = \frac{1}{8t}$ donne généralement une erreur statistique de l'ordre de l'erreur systématique. Il n'est alors pas nécessaire d'utiliser le processus d'extrapolation que nous venons de décrire. Après quelques vérifications, on cherche donc généralement à choisir les paramètres de la simulation pour être dans la situation où les extrapolations ne sont pas nécessaires.

Finalement, notons que le temps d'une simulation est proportionnel au nombre de balayages du réseau espace-temps, fois $N_\tau N^3$. Ce facteur vient du nombre total de sites NN_τ dans l'espace-temps, fois le nombre d'opérations nécessaire pour calculer la fonction de Green après chaque mise-à-jour d'un site, nombre qui est proportionnel à N^2 . Le calcul de la fonction de Green à une nouvelle tranche de temps, que ce soit par (23a) ou par (23b), est proportionnel à N^3 ce qui ne change donc pas la dépendance polynomiale en $N_\tau N^3$ mentionnée plus haut. Tout ceci ne tient pas compte des surplus de calcul provoqués par la stabilisation numérique décrite ci-dessous. Nous verrons que, malgré les calculs additionnels nécessaires pour maintenir la stabilité numérique, la dépendance en $N_\tau N^3$ est généralement dominante. Notons cependant que si, plutôt que d'utiliser (22), nous avions dû recalculer toute la fonction de Green espace-temps après chaque mise-à-jour, le temps de calcul aurait été proportionnel à $(N_\tau N)^3$, ce qui nous aurait fait perdre un facteur de $(N_\tau)^2$, un nombre assez grand!

4.2. Améliorations numériques nécessaires pour un programme qui fonctionne.

4.2.1. Stabilisation à basse température.

Des instabilités numériques apparaissent pour les simulations à basse température. L'origine de ces instabilités⁵ peut être comprise en considérant le problème sans interaction. En effet, dans ce cas, les valeurs propres de l'Hamiltonien vont de $-4t$ à $+4t$, ce qui veut dire que lorsque $\beta = 10/t$ par exemple, $e^{-\beta \hat{H}}$ a des valeurs propres qui vont de e^{-40} à e^{40} . Avec la précision habituelle de quinze chiffres sur les ordinateurs, il est donc très difficile d'évaluer $\mathbf{B}_{N_\tau}^\sigma \mathbf{B}_{N_\tau-1}^\sigma \dots \mathbf{B}_1^\sigma$ avec suffisamment de précision pour ne pas perdre d'information lorsque vient le temps de calculer

$$\left[\frac{\mathbf{I}}{\mathbf{I} + \mathbf{B}_{N_\tau}^\sigma \mathbf{B}_{N_\tau-1}^\sigma \dots \mathbf{B}_1^\sigma} \right]_{i,j}$$

où ni les valeurs propres trop petites, e^{-40} , ou trop grandes, e^{40} , ne contiennent d'information importante (la fonction de Green est dans ces cas soit nulle, soit égale à l'unité). Il n'y a que les valeurs propres autour de 1 (le niveau de Fermi) qui sont physiquement importantes.

Il existe plusieurs méthodes numériques pour régler ce genre de problème.¹⁹ Celle utilisée par White *et al.*⁵ procède comme suit. Supposons qu'on puisse calculer un produit de N_m matrices \mathbf{B} de façon stable. (Ce nombre N_m correspond généralement à la valeur de $\bar{\ell}$ définie sous l'équation (23b) puisque l'instabilité dans cette formule vient justement de la multiplication successive de matrices \mathbf{B} .) Les colonnes de la matrice résultante sont orthogonalisées par une procédure de Gram-Schmidt modifiée (qui est plus rapide que la procédure standard). En d'autres mots, on écrit

$$\mathbf{B}_{N_m} \dots \mathbf{B}_1 = \mathbf{QV}$$

où les colonnes de la matrice \mathbf{Q} sont orthogonales alors que la matrice \mathbf{V} est une matrice triangulaire supérieure où seulement le nombre 1 apparaît sur la diagonale. Chaque colonne de la matrice \mathbf{Q} est ensuite normalisée à l'unité pour pouvoir écrire cette matrice sous la forme \mathbf{UD} où \mathbf{U} est une matrice unitaire et \mathbf{D} une matrice diagonale contenant la norme de chaque colonne de \mathbf{Q} . L'utilité de cette approche vient du fait que les différents ordres de grandeur apparaissent maintenant tous dans \mathbf{D} . La matrice \mathbf{V} et la matrice \mathbf{U} peuvent être multipliées de façon stable.

Pour continuer le produit des N_m matrices \mathbf{B} suivantes par la gauche, on procède comme suit:

$$\{[\mathbf{B}_{2N_m} \dots \mathbf{B}_{N_m+1}]\mathbf{UD}\}\mathbf{V}$$

c'est-à-dire qu'on multiplie les matrices \mathbf{B} ensemble, puis on multiplie par \mathbf{U} . La multiplication par \mathbf{D} amplifie ensuite le fait que chaque colonne peut avoir une norme d'ordre de grandeur très différent. Pour isoler de nouveau ces ordres de grandeur, on orthogonalise encore une fois pour obtenir,

$$\{\mathbf{U}' \mathbf{D}' \mathbf{V}'\}\mathbf{V}.$$

Les matrices \mathbf{V} et \mathbf{V}' sont toutes deux des matrices triangulaires ayant l'unité sur la diagonale. Elles peuvent être multipliées de façon stable pour donner finalement

$$\{[\mathbf{B}_{2N_m} \dots \mathbf{B}_{N_m+1}]\mathbf{UD}\}\mathbf{V} = \mathbf{U}' \mathbf{D}' \mathbf{V}''$$

où $\mathbf{V}'' = \mathbf{V}' \mathbf{V}$. La non-associativité de la multiplication matricielle sur un ordinateur est ici flagrante: Les opérations doivent être absolument faites dans l'ordre prescrit.

Continuant ainsi jusqu'à ce que toutes les matrices \mathbf{B} aient été multipliées, on ajoute l'identité et on orthogonalise encore

$$\mathbf{I} + \mathbf{B}_{N_\tau} \dots \mathbf{B}_1 = \mathbf{U}_f \mathbf{D}_f \mathbf{V}_f.$$

avant d'inverser. En effet, le membre de droite est facile à inverser de façon stable car la seule matrice qui contient des ordres de grandeur très différents, la matrice \mathbf{D}_f , est inversible élément par élément de façon triviale:

$$\frac{\mathbf{I}}{\mathbf{I} + \mathbf{B}_{N_\tau} \dots \mathbf{B}_1} = \mathbf{V}_f^{-1} \mathbf{D}_f^{-1} \mathbf{U}_f^{-1}.$$

Le nombre d'opérations nécessaires pour la multiplication de matrices, tout comme pour l'orthogonalisation de Gram-Schmidt, est de l'ordre de N^3 lorsque la matrice est de taille $N \times N$. Il faut refaire ces orthogonalisations de l'ordre de N_τ/N_m fois lors d'un calcul de fonction de Green comme ci-dessus. De plus, ce genre de calcul doit être fait N_τ/N_m fois, soit à chaque fois que la fonction de Green doit être calculée au long avec la formule (23a) plutôt que "enroulée" avec la formule plus rapide (23b). La nécessité de stabilisation rajoute donc un temps de calcul proportionnel à $N^3(N_\tau/N_m)^2$. En pratique, les coefficients numériques sont tels que, dans la plupart des cas étudiés, le temps de l'algorithme continue à être proportionnel à $N_\tau N^3$. La vectorisation sur divers types d'ordinateurs (par exemple RISC/6000 avec l'option de compilation $-O$, ou Cray) peut modifier beaucoup ces estimés, rendant les grands systèmes plus accessibles qu'on aurait pu le croire a priori.

On peut aussi gagner un peu de vitesse⁵ en mémorisant les orthogonalisations successives \mathbf{UDV} et en mettant à jour le réseau en balayant d'abord les tranches de temps de 0 à N_τ , puis de N_τ à 0, ce qui permet de réutiliser les orthogonalisations \mathbf{UDV} antérieures. Cette façon de faire produit un algorithme dont l'exécution est en principe proportionnelle à $N^3(N_\tau/N_m)$.

4.2.2. Matrices éparses.

En utilisant les outils fournis par les compilateurs, on trouve que la sous-routine où cet algorithme passe le plus de temps est celle qui multiplie des matrices \mathbf{B}_ℓ^σ ensemble. Rappellant que $\mathbf{B}_\ell^\sigma \equiv e^{-\Delta \tau \mathbf{K}} e^{\mathbf{V}_\ell^\sigma}$, on voit tout de suite que le problème vient de la matrice $e^{-\Delta \tau \mathbf{K}}$ puisque l'autre est diagonale. En effet, la multiplication d'une matrice pleine demande de l'ordre de N^3

opérations. Ce nombre peut être rendu proportionnel à N^2 en utilisant encore une fois la décomposition de Trotter pour rendre la matrice d'énergie cinétique $e^{-\Delta \tau \mathbf{K}}$ éparse. En effet, considérons par exemple un réseau 4×4 . On regroupe les termes de saut en quatre matrices qui sont diagonales par bloc. Ceci se fait de la façon suivante:

$$\begin{array}{cccc} 1 & \text{---} & 2 & 3 & \text{---} & 4 \\ 5 & \text{---} & 6 & 7 & \text{---} & 8 \\ 9 & \text{---} & 10 & 11 & \text{---} & 12 \\ 13 & \text{---} & 14 & 15 & \text{---} & 16 \end{array}$$

où nous avons illustré les termes de saut qui appartiennent à la matrice \mathbf{K}_1^x dans la décomposition

$$\mathbf{K} = \mathbf{K}_1^x + \mathbf{K}_2^x + \mathbf{K}_1^y + \mathbf{K}_2^y.$$

Avec des conditions aux limites périodiques, les termes de saut appartenant à la matrice \mathbf{K}_2^x seraient ceux entre les sites 2 et 3, 1 et 4, 6 et 7 etc. Les liens dans la direction y verticale sont aussi arrangés de la même façon.

La matrice \mathbf{K}_1^x , par exemple, aura la forme

$$\mathbf{K}_1^x = \begin{bmatrix} 0 & -t & 0 & 0 & 0 \\ -t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -t & 0 \\ 0 & 0 & -t & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots \end{bmatrix}$$

qui est diagonale par bloc. Ceci veut dire que la matrice

$$e^{-\Delta \tau \mathbf{K}_1^x} = \begin{bmatrix} \cosh(t \Delta \tau) & -\sinh(t \Delta \tau) & 0 & 0 & 0 \\ -\sinh(t \Delta \tau) & \cosh(t \Delta \tau) & 0 & 0 & 0 \\ 0 & 0 & \cosh(t \Delta \tau) & -\sinh(t \Delta \tau) & 0 \\ 0 & 0 & -\sinh(t \Delta \tau) & \cosh(t \Delta \tau) & 0 \\ 0 & 0 & 0 & 0 & \dots \end{bmatrix}$$

est aussi diagonale par bloc. Son inverse est obtenu de façon triviale en changeant le signe de $\sinh(t \Delta \tau)$.

Le truc pour épargner du temps consiste maintenant à faire l'approximation suivante,

$$\begin{aligned}
e^{-\Delta \tau \mathbf{K}} &= e^{-\Delta \tau (\mathbf{K}_1^x + \mathbf{K}_2^x + \mathbf{K}_1^y + \mathbf{K}_2^y)} \\
&\approx e^{-\Delta \tau \mathbf{K}_1^x} e^{-\Delta \tau \mathbf{K}_2^x} e^{-\Delta \tau \mathbf{K}_1^y} e^{-\Delta \tau \mathbf{K}_2^y}
\end{aligned}$$

ce qui introduit une erreur de l'ordre de $(\Delta \tau)^2$, c'est-à-dire pas plus grande que celles déjà introduites auparavant. Le grand avantage que l'on retire est que $e^{-\Delta \tau \mathbf{K}_i^x} \times (\text{Matrice pleine})$ peut être calculé en un nombre d'opérations qui va comme N^2 au lieu de N^3 . Malgré les quatre multiplications qu'il faut faire de la sorte, il n'est pas nécessaire que N soit très grand pour que cette façon de faire devienne très avantageuse. Le gain en temps, linéaire en N , permet aussi de travailler à des $\Delta \tau$ plus petits, compensant largement l'erreur systématique additionnelle. En effet, diminuer $\Delta \tau$ d'un facteur deux double le temps de calcul, mais l'erreur systématique est diminuée d'un facteur quatre. Malgré toutes ces améliorations, les programmes que nous avons construits passent plus de la moitié de leur temps à multiplier des matrices \mathbf{B}_ℓ^σ ensemble.

5. Quelques résultats.

Les résultats présentés ici ont pour objet de démontrer qu'il est possible d'obtenir des résultats physiquement intéressants malgré la petitesse des systèmes étudiés. Les limites de la méthode seront discutées dans la section 6.

5.1. Ordre antiferromagnétique à demi-remplissage et facteur de structure magnétique.

Lorsqu'il y a un électron par site, il est bien connu que le modèle de Hubbard avec sauts aux premiers voisins prédit une transition de phase antiferromagnétique lorsqu'on traite ce problème dans l'approximation de la phase aléatoire généralisée (GRPA). Il y a cependant une complication additionnelle qui vient du fait qu'en deux dimensions les fluctuations thermiques interdisent les transitions de phase à symétrie brisée continue (Théorème de Mermin Wagner). La seule théorie analytique qui inclut cet effet jusqu'à présent est celle de la référence.²⁰

La figure 1 compare les résultats de simulations Monte Carlo pour le facteur de structure magnétique avec les résultats analytiques. Le facteur de structure magnétique S_{sp} et celui pour la charge S_{ch} sont définis par,

$$S_{ch}(\mathbf{q} \neq 0) = \frac{1}{N_s} \sum_{\mathbf{r}_i, \mathbf{r}_j} \exp(i\mathbf{q} \cdot (\mathbf{r}_i - \mathbf{r}_j)) \langle (n_{i,\uparrow} \mp n_{i,\downarrow})(n_{j,\uparrow} \mp n_{j,\downarrow}) \rangle \quad (24)$$

On choisit des unités où $t = 1$, et où le pas du réseau est aussi égal à l'unité. La figure 1 montre que le résultat Monte Carlo pour le facteur de structure magnétique au vecteur d'onde antiferromagnétique (π, π) sature à une valeur indépendante de la température lorsque celle-ci diminue. Cette valeur de saturation (les plateaux) dépend cependant fortement de la taille. Ceci est facile à comprendre à partir de la définition (24) ci-dessus: Lorsqu'il y a ordre à longue portée, la fonction de corrélation factorise et la double somme donne une contribution proportionnelle à N_s^2 que la division par N_s ne suffit pas à rendre indépendante de N_s . En d'autres mots, l'existence d'ordre à longue portée à température nulle se manifeste par une dépendance en taille du facteur de structure magnétique au vecteur d'onde où l'ordre s'établit. On peut aussi comprendre que le facteur de structure magnétique pour une taille donnée augmente, lorsque la température diminue, jusqu'à ce que la longueur de corrélation atteigne la taille du système. À ce moment, le résultat pour le système infini et celui pour le système fini commencent à différer, c'est-à-dire que le trait plein ne suit plus les résultats Monte Carlo. On peut cependant conclure que les résultats de la figure 1 démontrent clairement qu'il peut être possible de détecter l'existence d'une transition de phase à l'aide de simulations Monte Carlo. Il faut alors faire des études de taille finie, comme pour les systèmes classiques. En divisant la valeur au plateau par N_s , il est possible d'extrapoler la valeur du paramètre d'ordre pour un système infini à température nulle, comme il est fait dans la référence⁵. Cependant, lorsque l'ordre s'établit à des vecteurs d'ondes qui ne correspondent pas à un des vecteurs d'onde permis par la taille du réseau et par les conditions aux limites périodiques, les simulations sur des systèmes finis deviennent incapables de détecter cet ordre à longue portée. Ceci est suggéré aussi par les résultats de la figure 2 discutés dans le paragraphe suivant.

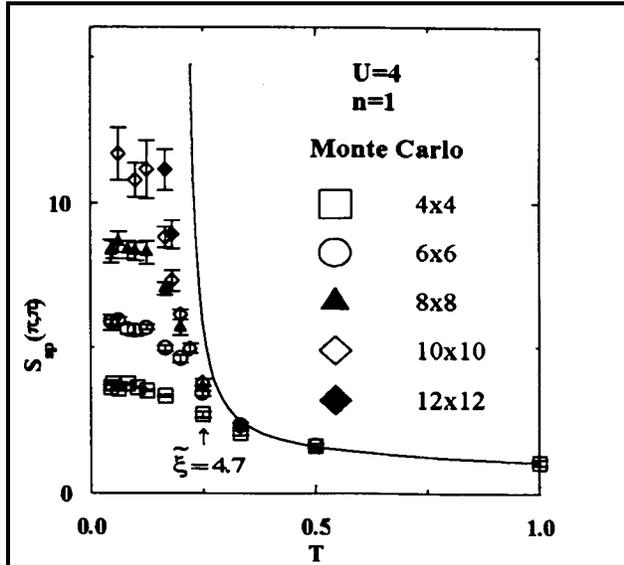


Figure 1: Facteur de structure magnétique au vecteur d'onde (π, π) en fonction de la température pour différentes tailles de système. La ligne pleine est le résultat d'une approche analytique,²⁰ et les symboles viennent de simulations Monte Carlo de la référence.⁵

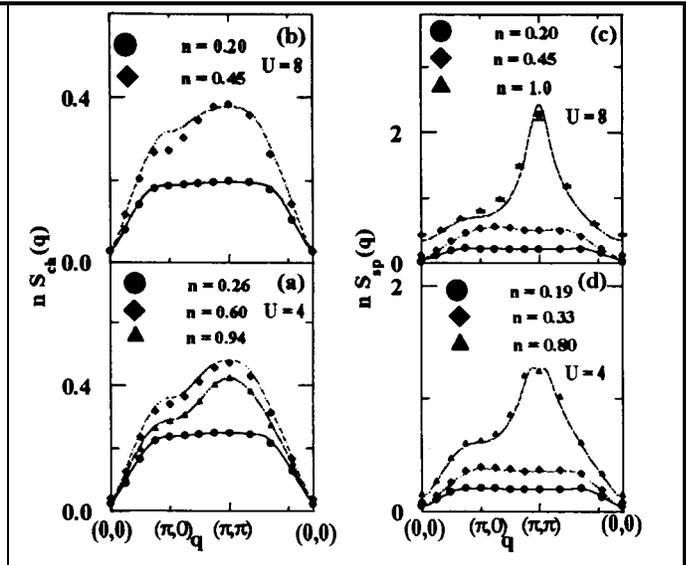


Figure 2: Facteur de structure de charge (à gauche) et de spin (à droite) dans un régime où les résultats ne dépendent plus de la taille. Les simulations et les calculs analytiques (ligne continue) viennent de la référence²⁰.

La figure 2 montre des résultats dans une région de paramètres où les effets de taille ne sont pas très importants. Les résultats analytiques et les résultats Monte Carlo sont alors essentiellement identiques. La seule différence se situe au niveau du facteur de structure magnétique qui, pour $U = 4$ et remplissage 0.8, commence à montrer des pics à une position incommensurable dans la théorie pour le système infini alors que les simulations Monte Carlo ne peuvent détecter cette tendance. Ceci s'explique par la petitesse des systèmes combinée aux conditions aux limites périodiques qui ne permettent d'étudier par simulations que quelques valeurs des vecteurs d'onde. Si un ordre à longue portée se développait à ce vecteur d'onde incommensurable, il deviendrait pratiquement impossible de le détecter par Monte Carlo.

5.2. Corrélations supraconductrices.

Le paramètre d'ordre supraconducteur s'écrit sous la forme,

$$\Delta_{\alpha}^{+} \equiv \sum_i \Delta_{\alpha}^{+}(\mathbf{r}_i) \equiv \frac{1}{2\sqrt{N}} \sum_{i,v} g^{\alpha}(v) c_{i\uparrow}^{+} c_{i+v\downarrow}^{+} \quad (25)$$

où la somme sur v est une somme sur les sites voisins de i . La valeur du poids $g^{\alpha}(v)$ détermine la symétrie spatiale (s , s étendue, d , p ...) et la symétrie de spin (triplet ou singulet) de la paire de Cooper. Pour les paramètres d'ordre singulets, la fonction de poids est paire dans l'espace, $g^{\alpha}(v) = g^{\alpha}(-v)$, et les symétries spatiales habituellement considérées sont:

$$\begin{aligned}
s : & \quad g^s(\mathbf{v}) = \delta_{\mathbf{v},0} , \\
s \text{ étendue, (dénotée } S): & \quad g^S(\mathbf{v}) = \delta_{\mathbf{v},\hat{x}} + \delta_{\mathbf{v},-\hat{x}} + \delta_{\mathbf{v},\hat{y}} + \delta_{\mathbf{v},-\hat{y}} \\
d : & \quad g^d(\mathbf{v}) = \delta_{\mathbf{v},\hat{x}} + \delta_{\mathbf{v},-\hat{x}} - \delta_{\mathbf{v},\hat{y}} - \delta_{\mathbf{v},-\hat{y}} . \quad (26)
\end{aligned}$$

Pour les paramètres d'ordre triplets, la fonction de poids est impaire dans l'espace, $g^\alpha(\mathbf{v}) = -g^\alpha(-\mathbf{v})$, et la symétrie habituellement considérée est p_x (composante $S_z = 0$), ou plus spécifiquement $g^{p_x}(\mathbf{v}) = \delta_{\mathbf{v},\hat{x}} - \delta_{\mathbf{v},-\hat{x}}$. Il est aussi commode de travailler dans l'espace de Fourier

$$\Delta_\alpha^+ = \frac{1}{2\sqrt{N}} \sum_{\mathbf{k}} F^\alpha(\mathbf{k}) c_{\mathbf{k}\uparrow}^+ c_{-\mathbf{k}\downarrow}^+ \quad (27)$$

où

$$F^\alpha(\mathbf{k}) = \sum_{\mathbf{v}} g^\alpha(\mathbf{v}) e^{i\mathbf{k}\cdot\mathbf{x}_{\mathbf{v}}} . \quad (28)$$

Pour S étendue, nous avons $F^\alpha(\mathbf{k}) = F^S(\mathbf{k}) = 2(\cos(k_x) + \cos(k_y))$, pour les ondes d nous avons $F^d(\mathbf{k}) = 2(\cos(k_x) - \cos(k_y))$, pour les ondes p_x $F^{p_x}(\mathbf{k}) = 2i \sin(k_x)$ et finalement pour les ondes s , $F^s(\mathbf{k}) = 1$.

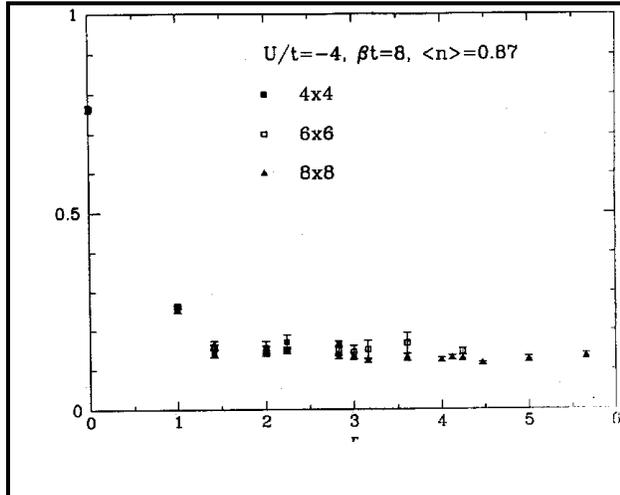


Figure 3: Fonction de corrélation $\langle \Delta_s^+(\mathbf{r}) \Delta_s(0) \rangle$ pour le paramètre d'ordre supraconducteur de type s , dans le cas $U < 0$ où il y a de l'ordre à longue portée. Cette fonction de corrélation sature à une valeur finie à grande distance. Figure tirée de la référence¹⁰.

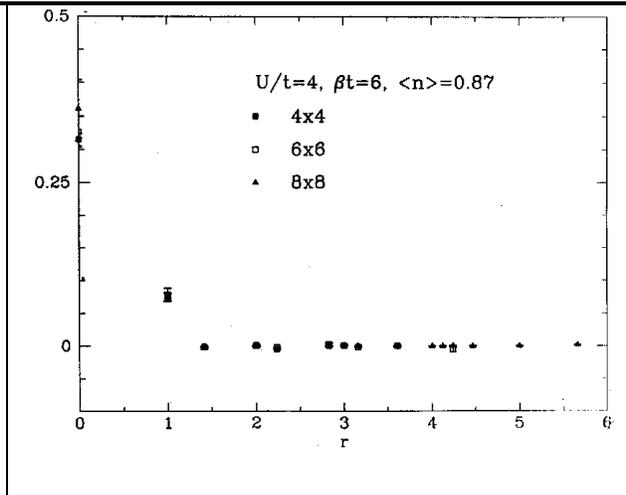


Figure 4: Fonction de corrélation $\langle \Delta_d^+(\mathbf{r}) \Delta_d(0) \rangle$ pour le paramètre d'ordre supraconducteur de type d , dans le cas $U > 0$ où on ne sait pas s'il y a de l'ordre à longue portée. Cette fonction de corrélation sature à une valeur nulle à grande distance. Figure tirée de la référence²¹.

Est-ce que l'approche Monte Carlo peut détecter la supraconductivité? Le cas où U est négatif est un cas où on sait qu'il y a de la supraconductivité. On peut donc commencer par s'assurer que les simulations Monte Carlo peuvent détecter sa présence. La figure 3 démontre¹⁰ que les simulations peuvent effectivement détecter de l'ordre supraconducteur à longue portée. En effet, pour des systèmes 4×4 , 6×6 et 8×8 on trouve que la fonction de corrélation $\langle \Delta_s^+(\mathbf{r}) \Delta_s(0) \rangle$ tend vers une valeur constante non nulle à mesure que \mathbf{r} augmente plutôt que de décroître vers zéro. Évidemment le système est trop petit pour faire la différence entre un véritable ordre à longue portée et un ordre algébrique du type Kosterlitz-Thouless.

Le cas $U > 0$ devrait correspondre aux supraconducteurs à haute température de transition. La question est de savoir si vraiment ce modèle peut prédire de la supraconductivité. Si tel est le cas, diverses raisons théoriques, que nous ne pouvons expliquer ici, suggèrent que cette supraconductivité devrait être de type d . Or, la figure 4 montre²¹ bien que $\langle \Delta_d^+(\mathbf{r}) \Delta_d(0) \rangle$ décroît vers zéro avec la distance, ce qui n'est pas très encourageant pour la supraconductivité!

En fait, on peut même montrer²² qu'à remplissage faible, la fonction de corrélation $\langle \Delta_d^+ \Delta_d \rangle$ se comporte comme si toutes les corrélations venaient de la matrice T , un effet trivial. Ceci est illustré par les figures 5a et 5b. Il n'y a que près du demi-remplissage $\langle n \rangle = 1$ que de légères déviations par rapport au calcul de matrice T se produisent.

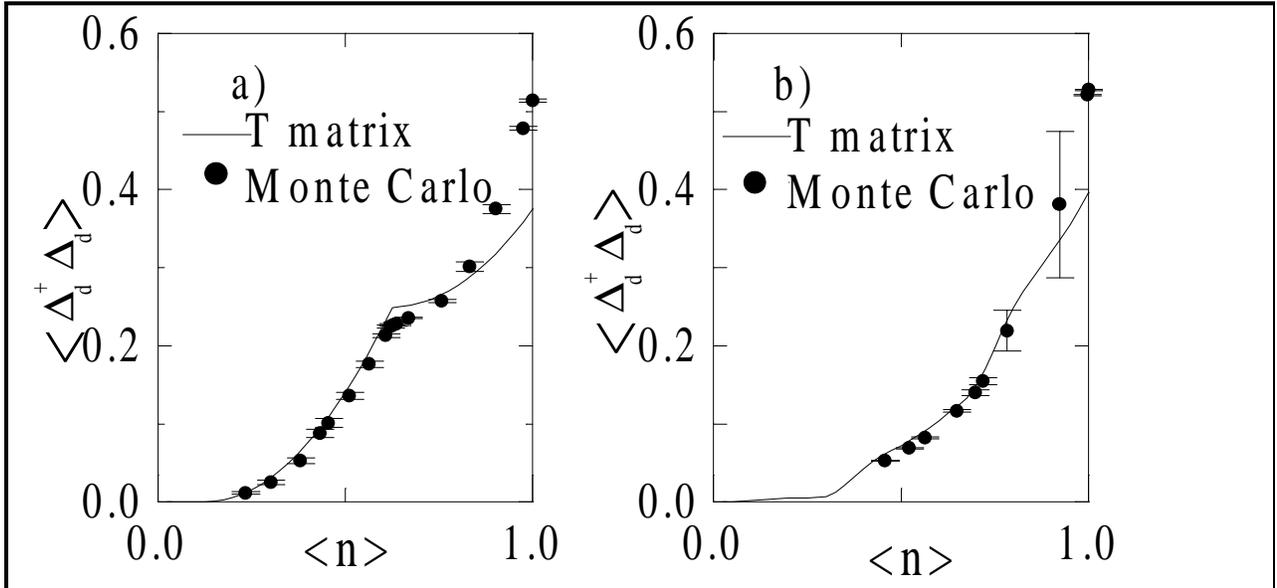


Figure 5: (a) Fonction de corrélation $\langle \Delta_d^+ \Delta_d \rangle$ pour le paramètre d'ordre supraconducteur de type d , dans le cas $U = 4$, $\beta = 6$, pour un système de taille 4×4 et $\Delta \tau = 1/10$. (Figure tirée de la référence²²). Les points représentent les résultats des simulations Monte Carlo alors que le trait plein est un calcul de matrice T . (b) Même chose qu'en (a) mais pour un système de taille 8×8 et $\Delta \tau = 1/8$.

Afin de donner une idée du temps nécessaire pour ces simulations, mentionnons que pour les réseaux 8×8 chaque point nécessite 8 heures de calcul sur une machine tournant à environ 9 MFlops (IBM RS6000 modèle 320). Pour un nombre fixe de mesures les barres d'erreur sont plus grandes là où le signe moyen du déterminant est plus petit.

5.3. Détection de la supraconductivité sans connaître la symétrie de la paire de Cooper.

Il est aussi possible de détecter la présence de supraconductivité sans connaître la symétrie du paramètre d'ordre. Ceci est très utile puisque, *à priori*, on ne peut être certain du type d'ordre qui s'installera. En général, pour trouver s'il y a de la supraconductivité il suffit de chercher la présence de l'effet Meissner, phénomène qui caractérise le mieux la supraconductivité. Ceci se fait à l'aide de la fonction de corrélation courant-courant.²³ En effet, on sait que la réponse linéaire du courant j_x à un potentiel vecteur A_x statique est donnée par

$$\langle j_x(\mathbf{q}, i\omega_n = 0) \rangle_A = -e^2 (\langle -k_x \rangle - \Lambda_{xx}(\mathbf{q}, i\omega_n = 0)) A_x(\mathbf{q}, i\omega_n = 0) \quad (29)$$

où le terme diamagnétique est donné, pour le cas des sauts aux premiers voisins, par

$$k_x(i) \equiv -t \sum_{\sigma} (c_{i+x,\sigma}^+ c_{i,\sigma} + c_{i,\sigma}^+ c_{i+x,\sigma})$$

et la fonction de corrélation courant-courant Λ_{xx} par les expressions,

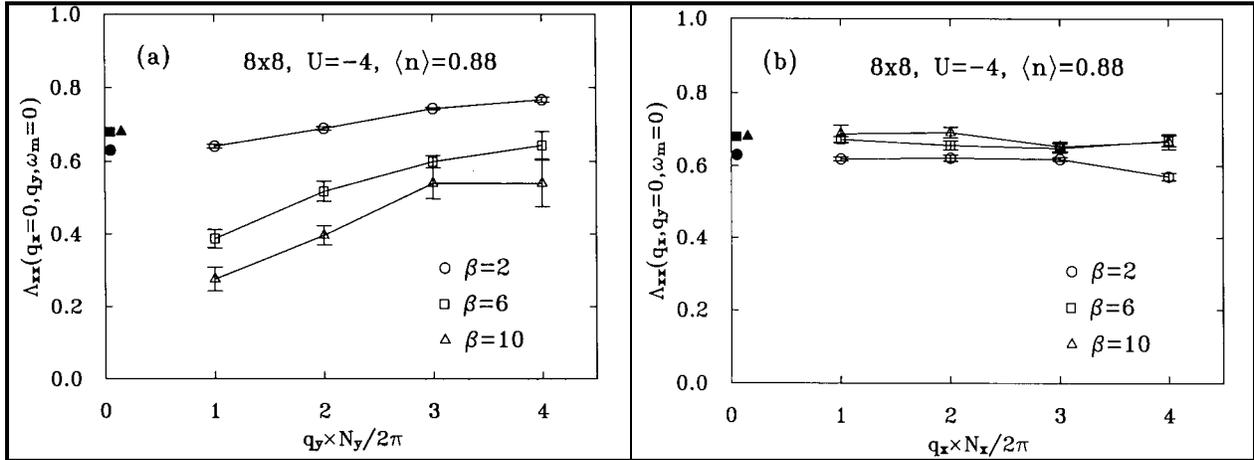


Figure 6: Figure et résultats tirés de la référence²³. Résultats Monte Carlo pour le modèle de Hubbard attractif (a) Réponse transversale $\Lambda_{xx}(q_x=0, q_y, i\omega_n=0)$ vs q_y ; (b) Réponse longitudinale $\Lambda_{xx}(q_x, q_y=0, i\omega_n=0)$ vs q_x pour $U = -4$, $\langle n \rangle \cong 0.875$ et $\beta = 2, 6$ et 10 . La contribution diamagnétique $\langle -k_x \rangle$ est indiquée par les symboles pleins.

$$\Lambda_{xx} \equiv \frac{1}{N} \int_0^\beta d\tau \langle j_x^p(\mathbf{q}, \tau) j_x^p(-\mathbf{q}, 0) \rangle$$

$$j_x^p(\mathbf{q}) \equiv it \sum_{i,\sigma} e^{-i\mathbf{q}\cdot\mathbf{r}_i} (c_{i+x,\sigma}^+ c_{i,\sigma} - c_{i,\sigma}^+ c_{i+x,\sigma}).$$

L'effet Meissner se manifeste par le fait que la réponse statique à un potentiel vecteur transversal est donnée par l'équation de London, $j = \frac{n_s e^2}{m} A$, alors que la réponse à un potentiel vecteur longitudinal s'annule par invariance de jauge. En d'autres mots, l'existence d'ordre à longue portée hors-diagonal rend la limite $\mathbf{q} \rightarrow 0$ non uniforme: Lorsque le courant est dans la direction x , la limite $q_x = 0, q_y \rightarrow 0$ correspond à la réponse transversale et on doit obtenir l'équation de London si le système est supraconducteur. Par contre, la limite $q_y = 0, q_x \rightarrow 0$ correspond à la réponse longitudinale qui doit toujours s'annuler à cause de l'invariance de jauge. Dans ce dernier cas, $\langle -k_x \rangle = \Lambda_{xx}(\mathbf{q}, \omega)$ dans l'équation pour la réponse (29) alors que pour la réponse transversale on obtient $\langle -k_x \rangle \neq \Lambda_{xx}(\mathbf{q}, \omega)$ si et seulement si il y a supraconductivité.

Les figures 6(a) et 6(b) vérifient ces résultats pour le cas attractif, ($U < 0$). Les efforts pour détecter la supraconductivité dans le cas $U > 0$ à l'aide de cette méthode ont échoué.²³

6. Limites de cette approche Monte Carlo quantique.

De loin le principal problème limitant l'utilité de cette méthode Monte Carlo quantique est lié au fait que le poids de Boltzmann pour l'échantillonnage selon l'importance n'est pas optimal

lorsque le signe du déterminant fermionique n'est pas positif, tel que discuté à la section 3.6. L'échantillonnage devient en fait très inefficace lorsque le signe du déterminant change de positif à négatif. L'annexe C explique comment ce problème de signe peut être relié à la transformation de Hubbard-Stratonovich. Nous démontrons ceci d'une autre façon dans la section suivante. Comme le signe moyen du déterminant est une quantité intermédiaire du calcul qui n'est pas une observable physique, il peut dépendre de la transformation de Hubbard-Stratonovich, même dans la limite où la somme sur les configurations est faite au complet, sans approximation. Nous verrons aussi que la transformation de Hubbard-Stratonovich influence l'ergodicité de l'algorithme. En d'autres mots, bien que les observables ne puissent pas dépendre de la transformation Hubbard-Stratonovich dans la limite où la somme est faite sur toutes les configurations, en pratique, l'efficacité de l'échantillonnage peut en dépendre. Finalement, dans la section 6.3 nous montrerons qu'il peut être très important de choisir les bons estimateurs statistiques pour les quantités physiques à mesurer.

La majorité des résultats de ce chapitre 6 sont tirés de la référence¹⁵. Nous illustrons notre propos en comparant les résultats obtenus à l'aide de deux choix de jauge différents pour la transformation de Hubbard-Stratonovich. Retournant à la définition (5) de la transformation Hubbard-Stratonovich discrète, nous paramétrisons un choix général de $\vec{\ell}$ à l'aide de

$$\vec{\ell}_{i,\tau} = \sin(\theta_{i,\tau})\vec{e}_x + \cos(\theta_{i,\tau})\vec{e}_z \quad ; \quad \theta_{i,\tau} = \mathbf{k} \cdot \mathbf{r}_i - \omega \tau.$$

Le choix $\theta_{i,\tau} = 0$ correspond à la jauge dite "standard", *i.e.* celle originellement proposée par Hirsch et couramment utilisée aujourd'hui. Le choix

$$\mathbf{k} = \left(\frac{\pi}{N_x}, \frac{\pi}{N_y} \right) \quad ; \quad \omega = 0$$

correspond à ce que nous appellerons la jauge spatialement symétrique.

6.1.Problèmes de signe. Effet de la jauge de la transformation Hubbard-Stratonovich.

Le signe moyen $\langle s \rangle$ du déterminant est défini comme la somme $1 + 1 + 1 - 1 \dots$ divisée par le nombre total de configurations étudiées, le signe devant chaque 1 correspondant à celui du déterminant pour une configuration donnée des champs de Hubbard-Stratonovich. Le signe moyen du déterminant n'est pas une quantité physiquement observable. Ce n'est qu'une quantité intermédiaire dans le calcul. Le lecteur ne sera donc pas surpris d'apprendre que la valeur de $\langle s \rangle$ puisse dépendre du choix de la transformation de Hubbard-Stratonovich. La figure 7 montre le signe moyen $\langle s \rangle$ du déterminant en fonction du remplissage pour la jauge spatialement symétrique

(ligne pleine) et pour la jauge standard (ligne pointillée). On constate que la pire situation se retrouve près du demi-remplissage. Bien que le signe soit en général plus grand pour la jauge standard, la figure montre que près du remplissage $\langle n \rangle = 0.75$ le signe peut être meilleur pour la jauge spatialement symétrique. Cependant, ce dernier résultat peut être trompeur, comme le montre la figure 8. En effet, sur cette figure on voit un cas où $\langle s \rangle$, bien que plus grand à haute température pour la jauge spatialement symétrique, devient éventuellement plus petit à basse température. Il est connu²⁴ que le signe moyen tombe exponentiellement avec l'inverse de la température β . La pente sur un graphique semi-log dépend de la jauge, comme nous le montre la figure 8. Dans la référence 15, plusieurs choix de jauge ont été étudiés. Il semble que, de plusieurs points de vue, la jauge spatialement symétrique et la jauge standard sont celles qui donnent les résultats les plus intéressants. C'est pourquoi dans cette section et celles qui suivent, nous n'avons comparé que ces deux jauges.

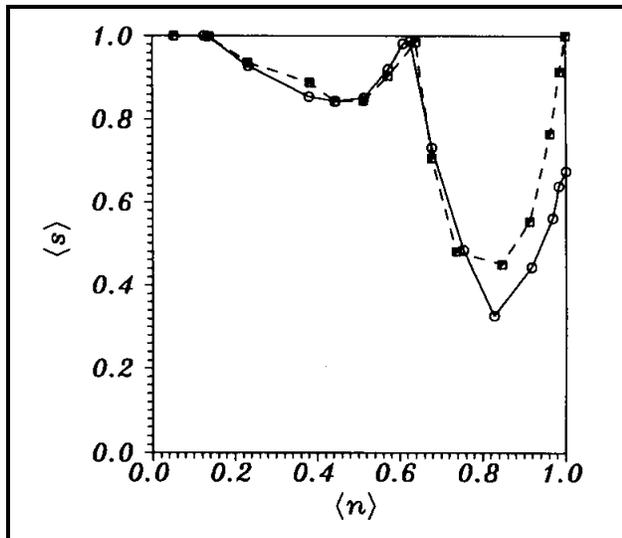


Figure 7: Signe moyen $\langle s \rangle$ du déterminant de fermions en fonction du remplissage $\langle n \rangle$ pour un réseau 4×4 à $U = 4$, $\beta = 4$ avec $\Delta \tau = 1/8$. 200 balayages complets du réseau espace-temps ont été faits avant de commencer à mesurer et 200 000 mesures ont été prises, une après chaque mise à jour d'une tranche de temps. Ces mesures sont groupées par blocs de 2 000 pour faire les statistiques. Un balayage du réseau espace-temps est fait à vide une fois sur deux. La ligne solide relie les résultats obtenus pour la jauge spatialement symétrique, alors que la ligne pointillée relie ceux obtenus pour la jauge standard. Figure tirée de la référence¹⁵.

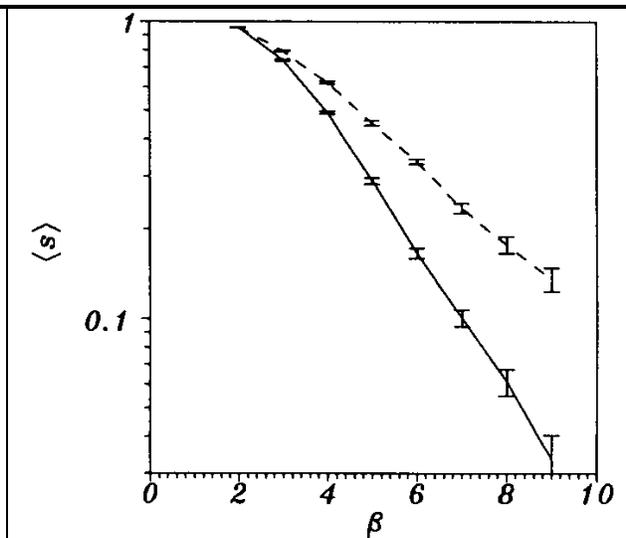


Figure 8: Graphique semi-logarithmique pour le signe moyen $\langle s \rangle$ du déterminant de fermions en fonction de l'inverse de la température β pour un système 2×2 à $U = 8$ avec 3 électrons et $\Delta \tau = 1/8$. Pour "réchauffer", 800 balayages complets du réseau espace-temps ont été faits. Ensuite, 800 000 mesures ont été prises et groupées par bloc de 8 000 pour faire les statistiques. Un balayage du réseau espace-temps est fait à vide une fois sur deux. La ligne solide relie les résultats obtenus pour la jauge spatialement symétrique, alors que la ligne pointillée relie ceux obtenus pour la jauge standard. Dans les deux cas on retrouve le comportement exponentiel asymptotique. Figure tirée de la référence¹⁵.

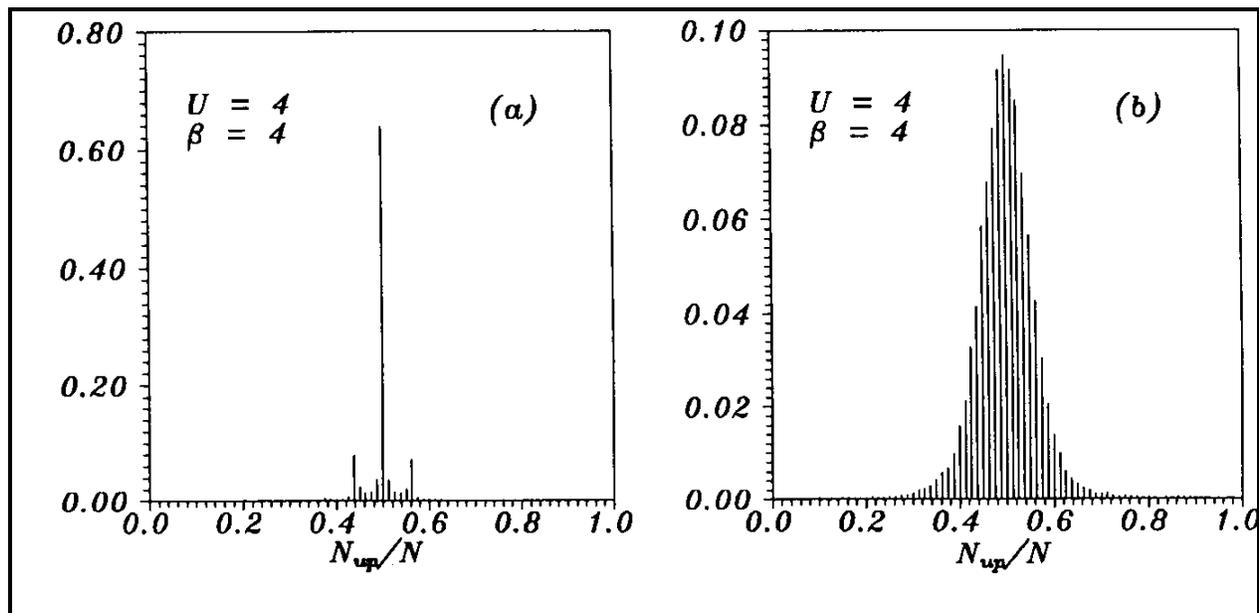


Figure 9: Histogrammes normalisés de N_{\uparrow} / N pour un système 4×4 à demi-rempli, $\beta = 4$, $U = 4$. Ces calculs ont été faits avec 400 balayages de thermalisation du réseau espace-temps au complet et 100 000 mesures groupées par blocs de 1 dans le cas (a), et par blocs de 5 dans le cas (b). Les résultats de la figure (a) sont pour la transformation de Hubbard-Stratonovich habituelle (jauge standard). Les résultats en (b) sont pour la jauge spatialement symétrique. La valeur de $\Delta \tau$ est la même dans les deux cas, soit $1 / 8$. Figure tirée de la référence¹⁵.

6.2. Efficacité de l'exploration de l'espace des phases et ergodicité.

L'algorithme décrit ici est bien sûr ergodique. En d'autres mots, toutes les configurations possibles des champs de Hubbard-Stratonovich peuvent être générées lorsque l'on permet à chaque variable discrète de prendre n'importe laquelle de ses deux valeurs possibles, comme nous le faisons ici. Cependant, ergodicité en théorie et ergodicité en pratique peuvent être deux choses très différentes. Comme chaque tranche de temps est mise à jour en changeant une variable à la fois à l'intérieur de cette tranche, la nouvelle configuration des champs de Hubbard-Stratonovich est très corrélée avec celle qui la précède. La figure 9(a) illustre le problème. À l'aide de l'ensemble des configurations générées lors d'une simulation Monte Carlo, on construit un histogramme de la fraction des spins qui sont dans la direction up, N_{\uparrow} / N . Comme le système est à demi-rempli et que les simulations sont faites dans l'ensemble grand-canonique, l'histogramme devrait plutôt ressembler à une cloche centrée autour de la valeur 0.5, comme l'histogramme de la figure 9(b). Les résultats de la figure 9(a) suggèrent plutôt que le système "colle" dans un sous-ensemble de l'espace des phases. Ce problème, que nous appellerons "collement", a été identifié et étudié en détail par Scalettar *et al.*²⁵ qui ont montré comment des changements globaux de la configuration des champs de Hubbard-Stratonovich pouvaient aider à améliorer l'efficacité avec laquelle l'espace des configurations est exploré. Un exemple de changement global qu'ils ont suggéré consiste à basculer la valeur de toutes les variables Hubbard-Stratonovich $x_{i,\ell}$ ayant un même indice spatial i , indépendamment de la coordonnée en temps

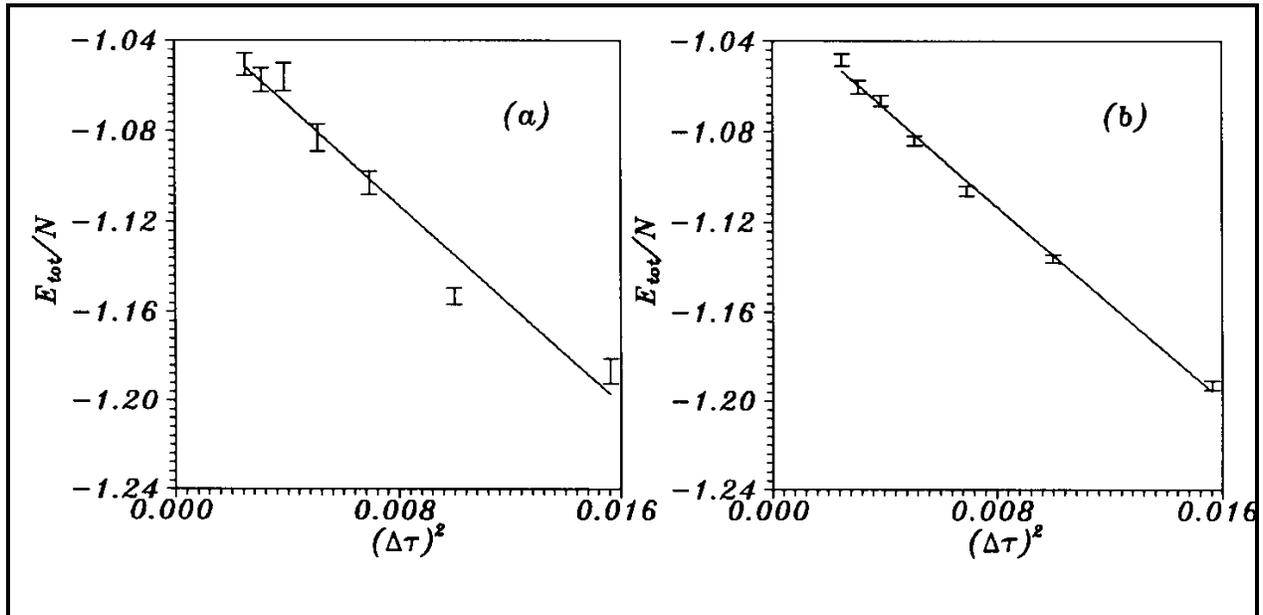


Figure 10: Énergie par site en fonction du carré de la grandeur de tranche de temps imaginaire $(\Delta \tau)^2$ pour un système 2×2 à $\beta = 4$, $U = 8$ et potentiel chimique fixe $\mu = 4$ correspondant au demi-remplissage. Les calculs sont faits pour 1 000 000 de mesures groupées par bloc de 10 000. De 800 à 2 000 balayages de thermalisation ont été faits, à mesure que $\Delta \tau$ diminuait. Les lignes solides représentent les meilleurs lissages des données par une droite. La figure (a) est pour la jauge standard, alors que la figure (b) est pour la jauge spatialement symétrique. Le signe moyen du déterminant $\langle s \rangle$ est à peu près égal à 0.6 dans ce dernier cas, alors qu'il est égal à l'unité dans le cas standard. Malgré le fait que le signe moyen soit plus petit pour la jauge spatialement symétrique (b), les résultats permettent une extrapolation plus fiable de la limite $\Delta \tau \rightarrow 0$ que les résultats (a) pour la jauge standard à cause du "collement" observé dans le cas standard. Figure tirée de la référence¹⁵.

imaginaire ℓ . On retrouve aussi ces stratégies de changement global dans le Monte Carlo classique.

Nous adoptons un point de vue un peu différent pour montrer que le collement, ou l'efficacité de l'exploration de l'espace des phases, dépend du choix des variables de Hubbard-Stratonovich.¹⁵ La figure 9(b) représente l'histogramme obtenu pour le même nombre de configurations qu'en (a), mais cette fois avec une transformation de Hubbard-Stratonovich dans la jauge spatialement symétrique. On voit que même si l'algorithme passe d'une configuration à l'autre simplement en basculant localement les variables Hubbard-Stratonovich, l'espace des phases est exploré de façon plus uniforme, du moins en ce qui concerne cette observable. À la section suivante 6.3, nous reviendrons au choix de l'estimateur pour une observable donnée.

Bien que la section précédente ait montré que du point de vue du signe moyen du déterminant il n'y a pas d'avantage majeur à choisir une jauge autre que la jauge standard, il semble ici que la jauge spatialement symétrique est vraiment préférable du point de vue de l'efficacité de l'exploration de l'espace des configurations. La figure 10 montre un exemple où, en effet, la

précision obtenue est meilleure dans le cas de la jauge spatialement symétrique Fig. 10(b) que dans le cas de la jauge standard Fig. 10(a), même si le signe moyen du déterminant est plus petit dans la jauge spatialement symétrique ($\langle s \rangle = 0.6$) que dans la jauge standard ($\langle s \rangle = 1$). Pour être plus spécifique, la figure 10 montre comment l'énergie du fondamental serait obtenue en éliminant l'effet de la discrétisation en temps imaginaire par une extrapolation $\Delta \tau \rightarrow 0$ (cf. annexe A). Malgré le fait que les mesures soient groupées en blocs de 10 000, les barres d'erreur dans la partie (a) sont très dispersées par rapport à la valeur moyenne attendue (ligne pleine), ce qui suggère qu'il y a du collement. En effet, lorsqu'il y a collement l'erreur statistique calculée sous-estime l'erreur réelle sur les points mesurés. En (b), l'erreur statistique semble être une meilleure mesure de l'erreur réelle sur chaque point. Il semble donc y avoir moins de collement dans ce cas.

Dans les deux cas (a) et (b), le nombre d'échantillons étudiés est le même. On serait donc porté à immédiatement déclarer vainqueur la jauge spatialement symétrique. Cependant, les calculs en (b) nécessitent environ 4 fois plus de temps qu'en (a). Si on calculait quatre fois plus longtemps en (a), les barres d'erreur diminueraient d'un facteur deux environ et les résultats obtenus dans les deux jauges deviendraient comparables. Pourquoi les calculs prennent-ils quatre fois plus de temps dans la jauge spatialement symétrique? Parce que le déterminant fermionique ne peut plus être factorisé en déterminant pour spins ups, multiplié par déterminant pour spins down. Le champ de Hubbard-Stratonovich ne conserve pas la composante z du spin dans la jauge spatialement symétrique à cause de la présence de la matrice de Pauli σ_x . Il faut donc travailler avec une matrice $2N_s \times 2N_s$ plutôt qu'avec deux matrices $N_s \times N_s$, ce qui veut dire que, pour N_s sites, le temps de calcul va comme $(2N_s)^3 N_\tau$ pour la jauge spatialement symétrique, et comme $2(N_s)^3 N_\tau$ pour la jauge standard.

6.3. Influence du choix de l'estimateur.

Même dans le cas de simulations Monte Carlo classiques, il est important de bien choisir la quantité mesurée (l'estimateur) de telle sorte qu'on puisse obtenir le plus rapidement possible des résultats statistiquement fiables. Par exemple, supposons que je veuille calculer la valeur moyenne de la quantité $\langle n_{i,\uparrow} n_{i,\downarrow} \rangle$. Je peux mesurer cette quantité sur le site i à chaque nouvelle configuration, mais une façon beaucoup plus efficace d'obtenir ce résultat est d'utiliser l'invariance sous translation et d'évaluer, pour chaque configuration générée par le Monte Carlo, la quantité

$$\sum_i \langle n_{i,\uparrow} n_{i,\downarrow} \rangle / N_s .$$

Les statistiques pour cette quantité seront bien meilleures que celles pour le calcul direct de $\langle n_{i,\uparrow} n_{i,\downarrow} \rangle$.

Il y a un aspect plus subtil à ce problème d'estimateur. En effet, la transformation de Hubbard-Stratonovich brise la symétrie de rotation. La trace sur toutes les configurations des variables de Hubbard-Stratonovich restaure cette symétrie mais, encore une fois, en pratique la symétrie de rotation n'est pas complètement restaurée lorsqu'il y a un nombre fini de configurations. En particulier, on peut vérifier l'inégalité suivante

$$S_z S_z \neq \frac{1}{2} (S_x S_x + S_y S_y)$$

pour une configuration donnée. Les moyennes sur les configurations de ces deux quantités ne sont égales qu'à l'intérieur de barres d'erreur. On a trouvé des cas où à haute température les barres d'erreur sur $\frac{1}{2} (S_x S_x + S_y S_y)$ sont plus petites que celles sur $S_z S_z$, alors que l'inverse est vrai à basse température! Dans le cas particulier qui nous intéressait²⁶, où on essayait de calculer la susceptibilité magnétique uniforme près du demi-remplissage, les résultats pouvaient être obtenus dix fois plus rapidement avec l'estimateur $S_z S_z$ qu'avec l'estimateur $\frac{1}{2} (S_x S_x + S_y S_y)$ parce que les barres d'erreur étaient environ trois fois plus petites avec $S_z S_z$ qu'avec $\frac{1}{2} (S_x S_x + S_y S_y)$. La référence 27 contient les citations à la littérature originale sur ce problème.

7. Conclusion.

Malgré tous les détails et trucs du métier décrits dans ces notes de cours, l'écriture d'un programme Monte Carlo utilisant cette méthode demeure assez complexe (quelques milliers de lignes) et l'utilisation de ce programme requiert une certaine expérience. En effet, comme pour les méthodes Monte Carlo classiques, il est nécessaire de se familiariser avec la thermalisation, les corrélations et en plus il faut s'habituer aux effets de l'erreur systématique due à la décomposition de Trotter ainsi qu'aux problèmes reliés au signe du déterminant. Ce dernier problème a tendance à surgir chaque fois que le régime de paramètres est intéressant pour la supraconductivité à haute température de transition! (Loi de Murphy généralisée?) Néanmoins, il n'y a pas vraiment de méthode numérique pour le modèle de Hubbard qui soit bien meilleure que celle expliquée ici. La méthode de renormalisation de la matrice densité, développée par White²⁸ et décrite ailleurs dans ces notes de cours, est définitivement supérieure pour les problèmes à une dimension. Cependant, en dimension plus élevée cette méthode de White n'a pas encore déclassé la méthode Monte Carlo. Mais qui connaît l'avenir?

Remerciements:

Je tiens à remercier plusieurs étudiants qui ont aidé à développer ce programme de simulations Monte Carlo au cours des ans. Par ordre chronologique, André Reid, Hugues Néliste, Christian Boily et Alain Veilleux. Je remercie particulièrement Liang Chen qui a développé avec moi les techniques de jauge pour améliorer l'ergodicité et qui a collaboré à l'obtention de la majorité des résultats de mon groupe. Je lui suis reconnaissant, ainsi qu'à Anne-Marie Daré et Yury Vilk, pour la permission d'utiliser certains de leurs résultats dans les sections 5 et 6. Je remercie aussi Anne-Marie Daré et Alain Veilleux pour leurs commentaires et pour les corrections qu'ils ont apportées au présent texte. Ce programme de recherche est financé en partie par le Conseil de Recherches en Sciences Naturelles et en Génie du Canada, et par le Fonds pour la Formation de Chercheurs et l'Aide à la Recherche du Québec. Je remercie aussi l'Institut Canadien de Recherches Avancées, section supraconductivité et la fondation Killam.

Annexe A

Erreur systématique due à la décomposition de Hubbard Stratonovich

La figure 10 montre un exemple d'analyse de l'erreur systématique. On y extrapole la valeur de l'énergie à $\Delta \tau = 0$ simplement en faisant un graphique de l'énergie en fonction de $(\Delta \tau)^2$. Empiriquement, la dépendance en $(\Delta \tau)^2$ est satisfaite pour des $\Delta \tau$ relativement grands ($\Delta \tau_{\max} = 1/8$ sur la figure. mais nous avons vu cette dépendance en $(\Delta \tau)^2$ à des valeurs encore plus grandes). Il est à remarquer que la pente en fonction de $(\Delta \tau)^2$ dépend en général de l'observable.

Ce problème de l'erreur systématique due à la décomposition de Hubbard Stratonovich a été discuté en détail par Fye¹² et Suzuki¹³. Nous donnons ici une preuve schématisée que l'erreur est bien d'ordre $(\Delta \tau)^2$ dans le calcul des observables. C'est ce qu'on veut dire lorsqu'on écrit

$$e^{-\beta(\hat{H} - \mu \hat{N})} = \prod_{l=1}^{N_\tau} e^{-\Delta \tau (\hat{K}_l + \hat{V}_l)} \approx \prod_{l=1}^{N_\tau} e^{-\Delta \tau \hat{K}_l} e^{-\Delta \tau \hat{V}_l} + O((\Delta \tau)^2). \quad (30)$$

Utilisant la formule de Weyl,

$$\exp(A + B) = \exp(A)\exp(B)\exp(-[A, B]/2)$$

on voit bien que,

$$e^{-\Delta \tau (\hat{K} + \hat{V})} = e^{-\Delta \tau \hat{K}} e^{-\Delta \tau \hat{V}} - \frac{1}{2} [\hat{K}, \hat{V}] (\Delta \tau)^2 + \dots$$

Comme le terme d'ordre $(\Delta \tau)^2$ intervient N_τ fois dans la première équation de cette annexe, l'erreur serait normalement d'ordre $N_\tau (\Delta \tau)^2 = \beta \Delta \tau$. Cependant, comme le commutateur $[\hat{K}, \hat{V}]$ est *réel* et anti-hermitien, nous avons que $\langle [\hat{K}, \hat{V}] \mathcal{O} \rangle = 0$. Ce dernier résultat doit être prouvé avec plus de soins,²⁹ mais finalement c'est lui qui permet d'écrire l'identité d'opérateurs (30) qui n'est vraiment valable que lorsque cette matrice densité (30) se retrouve dans une trace.

Pour prouver la disparition de la contribution anti-hermitienne nous suivons Fye¹². Utilisant d'abord la formule de Weyl deux fois et définissant $C_1 \equiv \frac{1}{2} [\hat{K}, \hat{V}]$ nous écrivons,

$$\prod_{l=1}^{N_\tau} e^{-\Delta \tau \hat{K}_l} e^{-\Delta \tau \hat{V}_l} = \prod_{l=1}^{N_\tau} \left(e^{-\Delta \tau (\hat{K}_l + \hat{V}_l + \Delta \tau C_1)} + O((\Delta \tau)^3) \right) = e^{-\beta(\hat{H} - \mu \hat{N} + \Delta \tau C_1)} + O((\Delta \tau)^2).$$

Utilisant maintenant la représentation d'interaction habituelle, nous pouvons écrire le membre de droite de cette dernière équation sous la forme

$$e^{-\beta(\hat{H} - \mu \hat{N} + \Delta \tau C_1)} = e^{-\beta(\hat{H} - \mu \hat{N})} (1 + \Delta \tau D_1(\beta)) + O((\Delta \tau)^2)$$

où l'opérateur anti-hermitien $D_1(\beta)$ est défini par

$$D_1(\beta) = \int_0^\beta d\tau C_1(\tau).$$

À l'aide de ces trois derniers résultats, on voit que la moyenne d'un opérateur \hat{O} quelconque peut s'écrire,

$$\langle \hat{O} \rangle_{approx} = \frac{\langle \hat{O} \rangle + (\Delta \tau) \langle D_1(\beta) \hat{O} \rangle}{1 + (\Delta \tau) \langle D_1(\beta) \rangle} + O((\Delta \tau)^2) \quad (31)$$

où $\langle \cdot \rangle_{approx}$ est calculée avec la matrice densité $\prod_{l=1}^{N_\tau} e^{-\Delta \tau \hat{K}_l} e^{-\Delta \tau \hat{V}_l}$. Pour montrer que les termes linéaires en $\Delta \tau$ s'annulent, nous avons d'abord besoin du lemme suivant:

$$\begin{aligned} \left[e^{-\beta(\hat{H} - \mu \hat{N})} D_1(\beta) \right]^+ &= \left[\int_0^\beta d\tau e^{-(\beta - \tau)(\hat{H} - \mu \hat{N})} C_1 e^{-\tau(\hat{H} - \mu \hat{N})} \right]^+ \\ &= \int_0^\beta d\tau e^{-\tau(\hat{H} - \mu \hat{N})} C_1^+ e^{-(\beta - \tau)(\hat{H} - \mu \hat{N})} \\ &= \int_0^\beta d\tau e^{-(\beta - \tau)(\hat{H} - \mu \hat{N})} C_1^+ e^{-\tau(\hat{H} - \mu \hat{N})} \end{aligned}$$

où la dernière égalité est prouvée en faisant le changement de variables $\tau \rightarrow \beta - \tau$. La quantité à gauche de cette équation a donc les mêmes propriétés d'hermiticité que C_1 et par conséquent, dans le cas qui nous concerne, le terme de gauche un opérateur anti-hermitien. Comme la trace d'un opérateur anti-hermitien avec une représentation en nombres réels s'annule, le terme en $\Delta \tau$ au dénominateur de la moyenne (31) disparaît. De même le terme linéaire en $\Delta \tau$ au numérateur s'annule puisque $Tr \left[e^{-\beta(\hat{H} - \mu \hat{N})} D_1(\beta) \hat{O} \right]$ est représentable par un nombre réel et

$$\begin{aligned} \left\{ \text{Tr} \left[e^{-\beta(\hat{H} - \mu \hat{N})} D_1(\beta) \hat{O} \right] \right\}^+ &= \text{Tr} \left[\hat{O}^+ \left(e^{-\beta(\hat{H} - \mu \hat{N})} D_1(\beta) \right)^+ \right] = - \text{Tr} \left[\hat{O} e^{-\beta(\hat{H} - \mu \hat{N})} D_1(\beta) \right] \\ &= - \text{Tr} \left[e^{-\beta(\hat{H} - \mu \hat{N})} D_1(\beta) \hat{O} \right] \end{aligned}$$

C.Q.F.D.

Il est utile de noter que si on cherche la moyenne d'un opérateur qui n'est pas hermitien, tel la fonction de Green, alors on peut toujours représenter cet opérateur comme étant la somme d'un opérateur hermitien, $(\hat{O} + \hat{O}^+) / 2$ et d'un opérateur anti-hermitien $(\hat{O} - \hat{O}^+) / 2$. Comme la moyenne de la partie anti-hermitienne évaluée avec la matrice densité exacte s'annulera, la meilleure façon d'estimer la valeur moyenne de l'opérateur non-hermitien sera d'utiliser,

$$\left\langle \frac{\hat{O} + \hat{O}^+}{2} \right\rangle_{approx} = \langle \hat{O} \rangle + O((\Delta \tau)^2).$$

Annexe B

Transformation de Hubbard-Stratonovich discrète avec choix de jauge

Pour prouver la relation

$$e^{-\Delta \tau \left[U \left(\hat{n}_{v\uparrow} - \frac{1}{2} \right) \left(\hat{n}_{v\downarrow} - \frac{1}{2} \right) \right]} = e^{-\Delta \tau \frac{U}{4}} \frac{1}{2} \sum_{x_v = \pm 1} e^{-\lambda x_v (2\hat{s} \cdot \vec{l})},$$

où \vec{l} est un vecteur unité de direction arbitraire, il suffit d'abord de vérifier, à l'aide des relations de commutation des fermions, que

$$(2\hat{S}_x)^2 = (2\hat{S}_y)^2 = (2\hat{S}_z)^2 = \hat{n}_\uparrow + \hat{n}_\downarrow - 2 \hat{n}_\uparrow \hat{n}_\downarrow$$

d'où on trouve que

$$(2\hat{S} \cdot \vec{l})^2 = \hat{n}_\uparrow + \hat{n}_\downarrow - 2 \hat{n}_\uparrow \hat{n}_\downarrow.$$

Le membre de gauche de l'identité à prouver s'écrit donc,

$$e^{-\Delta \tau \frac{U}{4}} \left[e^{-\Delta \tau \frac{U}{2} (2\hat{S} \cdot \vec{l})^2} \right]$$

alors que le membre de droite devient

$$e^{-\Delta \tau \frac{U}{4}} \cosh(2\hat{S} \cdot \vec{\ell} \lambda).$$

Comme $\hat{n}_\uparrow^2 = \hat{n}_\uparrow$, on prouve aussi que $\left[(2\hat{S} \cdot \vec{\ell})^2 \right]^2 = (2\hat{S} \cdot \vec{\ell})^2$ ce qui permet de réécrire pour le membre de gauche,

$$\left[e^{-\Delta \tau \frac{U}{2} (2\hat{S} \cdot \vec{\ell})^2} \right] = 1 + (2\hat{S} \cdot \vec{\ell})^2 \left[e^{-\frac{\Delta \tau U}{2}} - 1 \right]$$

et pour le membre de droite,

$$\cosh(2\hat{S} \cdot \vec{\ell} \Lambda) = 1 + (2\hat{S} \cdot \vec{\ell})^2 [\cosh(\Lambda) - 1]$$

ce qui implique que l'identité est satisfaite lorsque

$$\cosh(\Lambda) = e^{-\Delta \tau \frac{U}{2}}$$

comme dans le cas $\vec{\ell} = \vec{z}$ proposé initialement par Hirsch¹⁴.

Annexe C

Traitement symétrique des fermions dans la transformation de Hubbard-Stratonovich

Batrouni et de Forcrand³⁰ ont proposé une méthode qui traite les fermions des deux spins de façon symétrique. La transformation de Hubbard-Stratonovich utilisée est de la forme,

$$e^{-\Delta \tau U (\hat{n}_\uparrow \hat{n}_\downarrow - \frac{1}{2} \hat{n}_\uparrow - \frac{1}{2} \hat{n}_\downarrow)} = \int dy P(y) y^{(\hat{n}_\uparrow + \hat{n}_\downarrow)}$$

où les quatre états possibles $|\uparrow\rangle|\downarrow\rangle = |0\rangle|0\rangle, |0\rangle|1\rangle, |1\rangle|0\rangle, |1\rangle|1\rangle$ impliquent trois conditions pour la fonction $P(y)$:

$$\int dy P(y) = 1$$

$$\int dy P(y) y = e^{\Delta \tau U/2} .$$

$$\int dy P(y) y^2 = 1$$

Pour avoir un signe de déterminant positif, il suffit que $P(y)$ soit toujours positif. En effet, le déterminant sera formé de produits de $P(y)$ et d'un produit de déterminants de signes identiques pour les spins up et les spins down puisque ceux-ci interviennent de façon symétrique dans la combinaison $\hat{n}_\uparrow + \hat{n}_\downarrow$. Lorsque $P(y)$ est positif, on peut interpréter cette fonction comme si elle représentait une probabilité. Les équations précédentes nous disent que les deux premiers moments de cette probabilité sont,

$$\langle y \rangle = e^{\Delta \tau U/2}$$

$$\langle y^2 \rangle = 1 .$$

Or, les moments d'une probabilité obéissent à $\langle y \rangle^2 \leq \langle y^2 \rangle$ ce qui n'est possible que si $e^{\Delta \tau U/2} \leq 1$, ou $U < 0$. Il est bien connu qu'il n'y a pas de problème de signe lorsque⁴ $U < 0$ (puisque le déterminant pour les spins up a le même signe que celui pour les spins down) en conformité avec ce résultat. Batrouni et de Forcrand³⁰ ont suggéré que, même pour $U > 0$, il était intéressant de fonctionner avec ce type de transformation de Hubbard-Stratonovich où les spins sont traités de façon symétrique, mais où les $P(y)$ peuvent être négatifs. Avec leur approche, contrairement au cas habituel, le signe du déterminant peut facilement être calculé à partir de la configuration des y puisque le signe du déterminant fermionique ne vient que du fait que les $P(y)$ peuvent être négatifs. Tout ceci est bien expliqué dans la référence³⁰.

Annexe D

Évolution d'un état à N particules par un produit d'opérateurs d'évolution à une particule: théorèmes pour l'évaluation de la trace fermionique.

Nous utilisons la notation matricielle pour les opérateurs de création dans une base où N états à une particule sont possibles (dans le modèle de Hubbard à une bande, ceci correspond à N sites):

$$\mathbf{c}^+ \equiv (c_1^+, c_2^+, \dots, c_N^+)$$

La démonstration suit les références 6,4.

Lemme A: L'action de l'opérateur d'évolution à une particule $e^{-c^+ A c}$ sur un état à une particule $c_j^+ |0\rangle$ est donné par

$$e^{-c^+ A c} c_j^+ |0\rangle = (\mathbf{c}^+ e^{-A})_j e^{-c^+ A c} |0\rangle = (\mathbf{c}^+ e^{-A})_j |0\rangle.$$

Preuve:

Notons d'abord que

$$e^{-c^+ A c} = e^{-(c^+ T^+)(T A T^+)(T c)} \equiv e^{-\tilde{c}^+ \tilde{A} \tilde{c}}$$

où A est diagonalisée par la transformation unitaire T . \tilde{A} est la matrice diagonale correspondante et les opérateurs création et d'annihilation $\tilde{c}^+ \equiv c^+ T^+$, $\tilde{c} \equiv T c$. (soit $\tilde{c}^+ \equiv (\tilde{c}_{\alpha_1}^+, \tilde{c}_{\alpha_2}^+, \dots, \tilde{c}_{\alpha_N}^+)$...) obéissent aux relations de commutation canoniques. Le lemme peut être prouvé en notant que

$$e^{-\tilde{c}^+ \tilde{A} \tilde{c}} \tilde{c}_\alpha^+ = \tilde{c}_\alpha^+ e^{-a_\alpha} e^{-\tilde{c}^+ \tilde{A} \tilde{c}} = (\tilde{c}^+ e^{-\tilde{A}})_\alpha e^{-\tilde{c}^+ \tilde{A} \tilde{c}}$$

qu'il est facile de démontrer en utilisant le fait que \tilde{A} est diagonale et qu'il y a une particule de plus dans l'état α lorsque l'exponentielle agit après \tilde{c}_α^+ plutôt qu'avant. Le lemme est prouvé en multipliant ensuite la dernière équation par la droite avec $T_{\alpha j}$ et en sommant sur α .

Lemme B: Dans un état à N particules, l'opérateur $e^{-c^+ A c}$ fait évoluer chaque particule indépendamment, comme dans A ci-dessus, c'est-à-dire,

$$e^{-c^+ A c} c_j^+ c_i^+ |0\rangle = (\mathbf{c}^+ e^{-A})_j (\mathbf{c}^+ e^{-A})_i e^{-c^+ A c} |0\rangle = (\mathbf{c}^+ e^{-A})_j (\mathbf{c}^+ e^{-A})_i |0\rangle.$$

Preuve:

Lorsque $i = j$, les membres de droite et de gauche s'annulent tous les deux à cause du principe de Pauli. Lorsque i est différent de j , il suffit de commuter successivement l'opérateur d'évolution avec chacun des opérateurs de création, comme dans A. Physiquement, l'opérateur d'évolution $e^{-c^+ A c}$ étant un opérateur à une particule, l'état de chaque particule évolue indépendamment de toutes les autres particules.

Lemme C: L'opérateur $e^{-c^+ B c} e^{-c^+ A c}$ est un opérateur à une particule, faisant évoluer chaque état à une particule avec la matrice $e^{-B} e^{-A}$ (au sens défini par les deux lemmes précédents).

Preuve:

En appliquant le résultat **A** deux fois on trouve,

$$e^{-c^+ B c} e^{-c^+ A c} c^+ = e^{-c^+ B c} (c^+ e^{-A}) e^{-c^+ A c} = (c^+ e^{-B} e^{-A}) e^{-c^+ B c} e^{-c^+ A c}$$

où dans la deuxième égalité on a utilisé $e^{-c^+ B c} (e^{-A})_{i,j} = (e^{-A})_{i,j} e^{-c^+ B c}$ qui est vrai puisque $e^{-c^+ B c}$ n'est pas une matrice, mais un opérateur scalaire. Tout ceci prouve bien que la matrice $e^{-B} e^{-A}$, qui est en général non-hermitique, joue ici le même rôle que la matrice e^{-A} dans le cas où un seul opérateur d'évolution apparaît. Le résultat **B** montre que cela demeure vrai pour des états à plusieurs particules.

Lemme D:

$$Tr[e^{-c^+ B c} e^{-c^+ A c}] = \det[\mathbf{I} + e^{-B} e^{-A}]$$

Preuve:

Il suffit de passer dans la base des vecteurs propres gauche de la matrice $e^{-B} e^{-A}$, i.e.,

$$e^{-c^+ B c} e^{-c^+ A c} c_\alpha^+ |0\rangle = c_\alpha^+ (e^{-B} e^{-A})_\alpha |0\rangle.$$

Dans cette base, on peut écrire,

$$Tr[e^{-c^+ B c} e^{-c^+ A c}] = \prod_{\alpha=1}^N (1 + (e^{-B} e^{-A})_\alpha) = \det[\mathbf{I} + e^{-B} e^{-A}]$$

C.Q.F.D.

Il est utile de réécrire le lemme **C** en terme de l'opérateur d'évolution à une particule. N'écrivant pas explicitement la dépendance en spin ou en $\{x_{i,\ell}\}$ pour alléger la notation et suivant les définitions (6), nous avons

$$\hat{U}(\ell, \ell) = 1 \quad ; \quad \hat{U}(\ell, \ell - 1) = \hat{D}_\ell$$

de telle sorte que

$$\hat{U}(\ell_1, \ell_2) = \hat{D}_{\ell_1} \dots \hat{D}_{\ell_2 + 1}.$$

Le lemme C nous montre alors que

$$\hat{U}(N_\tau, 0) (\mathbf{c}^+)_i = (\mathbf{c}^+ \mathbf{B}_{N_\tau} \mathbf{B}_{N_\tau - 1} \dots \mathbf{B}_1)_i \hat{U}(N_\tau, 0) \quad (32)$$

où les \mathbf{B}_ℓ sont définis par l'équation (9).

Annexe E

Fonction de Green à temps égal.

À la section 3.5, nous avons été confrontés au problème du calcul de quantités de la forme

$$\langle c_i c_j^+ \rangle_{\{x\}}^0 = \frac{\text{Tr}[c_i c_j^+ \hat{D}_{N_\tau} \hat{D}_{N_\tau - 1} \dots \hat{D}_1]}{\text{Tr}[\hat{D}_{N_\tau} \hat{D}_{N_\tau - 1} \dots \hat{D}_1]}.$$

Le résultat a été obtenu dans le cas où un seul opérateur à une particule apparaît et une généralisation plausible a été présentée à l'équation (13). Dans cette annexe, nous donnons une preuve plus rigoureuse du résultat pour le cas suivant,

$$\mathbf{G}_{i,j}^\sigma(\ell, \ell) \equiv G^\sigma(i, \ell; j, \ell^-) = \langle c_i(\ell) c_j^+(\ell) \rangle_{\{x\}}^0 \equiv \frac{\text{Tr}[c_i c_j^+ \hat{U}(\ell, 0) \hat{U}(N_\tau, \ell)]}{\text{Tr}[\hat{U}(\ell, 0) \hat{U}(N_\tau, \ell)]}$$

où les définitions (6) ont été utilisées. Passant dans la base des vecteurs propres gauches de $\mathbf{B}_\ell^\sigma \dots \mathbf{B}_1^\sigma \mathbf{B}_{N_\tau}^\sigma \dots \mathbf{B}_{\ell+1}^\sigma$ et utilisant le résultat (32) à la fin de l'annexe précédente on trouve, pour un état à deux particules par exemple,

$$\hat{U}(\ell, 0) \hat{U}(N_\tau, \ell) c_\alpha^+ c_\beta^+ |0\rangle = c_\alpha^+ (\mathbf{B}_\ell^\sigma \dots \mathbf{B}_1^\sigma \mathbf{B}_{N_\tau}^\sigma \dots \mathbf{B}_{\ell+1}^\sigma)_\alpha c_\beta^+ (\mathbf{B}_\ell^\sigma \dots \mathbf{B}_1^\sigma \mathbf{B}_{N_\tau}^\sigma \dots \mathbf{B}_{\ell+1}^\sigma)_\beta |0\rangle$$

puisque $\hat{U}(\ell, 0) \hat{U}(N_\tau, \ell) |0\rangle = |0\rangle$. Faisant le changement de base $c_i c_j^+ = \sum_{\beta \gamma} \langle i | \beta \rangle \langle \gamma | j \rangle c_\beta c_\gamma^+$, il n'y a que le terme $\beta = \gamma$ qui survit dans la trace

$$\begin{aligned}
Tr[c_i c_j^+ \hat{U}(\ell, 0) \hat{U}(N_\tau, \ell)] = \\
\sum_{\beta} \langle i | \beta \rangle \langle \beta | j \rangle \left[\langle 0 | c_{\beta} c_{\beta}^+ | 0 \rangle + \langle 0 | c_{\beta} c_{\beta} c_{\beta}^+ c_{\beta}^+ | 0 \rangle (\mathbf{B}_{\ell}^{\sigma} \dots \mathbf{B}_1^{\sigma} \mathbf{B}_{N_{\tau}}^{\sigma} \dots \mathbf{B}_{\ell+1}^{\sigma})_{\beta} \right] \\
\prod_{\alpha \neq \beta} \left[\langle 0 | 0 \rangle + \langle 0 | c_{\alpha} c_{\alpha}^+ | 0 \rangle (\mathbf{B}_{\ell}^{\sigma} \dots \mathbf{B}_1^{\sigma} \mathbf{B}_{N_{\tau}}^{\sigma} \dots \mathbf{B}_{\ell+1}^{\sigma})_{\alpha} \right]
\end{aligned}$$

Au dénominateur, tous les termes où $\alpha \neq \beta$ se simplifient avec le numérateur et il ne reste pour la fonction de Green que

$$\begin{aligned}
\frac{Tr[c_i c_j^+ \hat{U}(\ell, 0) \hat{U}(N_\tau, \ell)]}{Tr[\hat{U}(\ell, 0) \hat{U}(N_\tau, \ell)]} &= \sum_{\beta} \langle i | \beta \rangle \frac{1}{1 + (\mathbf{B}_{\ell}^{\sigma} \dots \mathbf{B}_1^{\sigma} \mathbf{B}_{N_{\tau}}^{\sigma} \dots \mathbf{B}_{\ell+1}^{\sigma})_{\beta}} \langle \beta | j \rangle \\
&= \left[\frac{\mathbf{I}}{\mathbf{I} + \mathbf{B}_{\ell}^{\sigma} \dots \mathbf{B}_1^{\sigma} \mathbf{B}_{N_{\tau}}^{\sigma} \dots \mathbf{B}_{\ell+1}^{\sigma}} \right]_{i, j} = \mathbf{G}_{i, j}^{\sigma}(\ell, \ell) \equiv G^{\sigma}(i, \ell; j, \ell^{-})
\end{aligned}$$

ce qui correspond bien aux équations (23) et (13).

Annexe F

Fonctions de Green à temps inégal.

Dans cette annexe, pour des raisons de commodité de notation, les indices de spin ne sont pas écrits explicitement. Nous utilisons la définition de Negele et Orland³¹ pour la fonction de Green:

$$G(i, \ell_1; j, \ell_2) = \left\langle T(c_i(\ell_1) c_j^+(\ell_2)) \right\rangle_{\{x\}}^0$$

où la dépendance en fréquence de Matsubara des opérateurs de création et d'annihilation est donnée par l'équation (11), l'indice supérieur 0 rappelle que les électrons sont indépendants, et l'indice inférieur indique que ces électrons peuvent se déplacer dans un champ qui dépend de la position et du temps imaginaire. Le produit chronologique T ordonne les opérateurs selon leur temps imaginaire, les plus grands temps étant à gauche. Pour être plus spécifique, la fonction de Green est définie comme suit,

$$G(i, \ell_1; j, \ell_2) = \theta(\ell_1 - \ell_2) \left\langle c_i(\ell_1) c_j^+(\ell_2) \right\rangle_{\{x\}}^0 - \theta(\ell_2 - \ell_1) \left\langle c_j^+(\ell_2) c_i(\ell_1) \right\rangle_{\{x\}}^0$$

avec $\theta(\tau) = 1$ pour $\tau > 0$ et $\theta(\tau) = 0$ pour $\tau < 0$.

Considérons le cas $\ell_1 > \ell_2$. Alors, utilisant les définitions (11) et faisant la décomposition de Trotter sur les opérateurs de la forme $e^{\ell_1 \Delta \tau (\hat{H} - \mu N)}$, la fonction de Green s'écrit,

$$\langle c_i(\ell_1) c_j^+(\ell_2) \rangle_{\{x\}}^0 = Z^{-1} \text{Tr} [\hat{U}(N_\tau, \ell_1) c_i \hat{U}(\ell_1, \ell_2) c_j^+ \hat{U}(\ell_2, 0)]$$

Utilisant la dernière équation, (32), de l'annexe D pour commuter l'opérateur d'évolution $\hat{U}(\ell_1, \ell_2)$ avec l'opérateur de création c_j^+ , cette dernière équation se réécrit,

$$\langle c_i(\ell_1) c_j^+(\ell_2) \rangle_{\{x\}}^0 = Z^{-1} \text{Tr} [\hat{U}(N_\tau, \ell_1) c_i c_k^+ \hat{U}(\ell_1, 0)] (\mathbf{B}_{\ell_1} \dots \mathbf{B}_{\ell_2+1})_{k,j}$$

avec une somme implicite sur l'indice k . Utilisant ensuite la propriété cyclique de la trace, on trouve

$$\langle c_i(\ell_1) c_j^+(\ell_2) \rangle_{\{x\}}^0 = \frac{\text{Tr} [c_i c_k^+ \hat{U}(\ell_1, 0) \hat{U}(N_\tau, \ell_1)]}{\text{Tr} [\hat{U}(\ell_1, 0) \hat{U}(N_\tau, \ell_1)]} (\mathbf{B}_{\ell_1} \dots \mathbf{B}_{\ell_2+1})_{k,j}.$$

Le rapport des traces n'est autre chose que la fonction de Green à temps égal que nous avons déjà rencontrée à la section 4.1, équation (23). Une preuve de sa forme explicite apparaît à l'annexe E. Utilisant ces résultats, ou l'équation (23), on voit que

$$\langle c_i(\ell_1) c_j^+(\ell_2) \rangle_{\{x\}}^0 = \left[\frac{\mathbf{I}}{\mathbf{I} + \mathbf{B}_{\ell_1} \dots \mathbf{B}_1 \mathbf{B}_{N_\tau} \dots \mathbf{B}_{\ell_1+1}} \mathbf{B}_{\ell_1} \dots \mathbf{B}_{\ell_2+1} \right]_{i,j}.$$

Ce même résultat peut en principe être obtenu³¹ en calculant l'élément correspondant de la matrice G définie par l'équation (10).

Avec quelques manipulations matricielles, et se souvenant que $\ell_1 > \ell_2$, on peut aussi réécrire cette équation sous la forme apparaissant dans Hirsch⁴, Eq.(3.25):

$$\langle c_i(\ell_1) c_j^+(\ell_2) \rangle_{\{x\}}^0 = \left[\frac{\mathbf{I}}{\mathbf{B}_{\ell_2+1}^{-1} \dots \mathbf{B}_{\ell_1}^{-1} [\mathbf{I} + \mathbf{B}_{\ell_1} \dots \mathbf{B}_1 \mathbf{B}_{N_\tau} \dots \mathbf{B}_{\ell_1+1}]} \right]_{i,j}$$

$$\langle c_i(\ell_1)c_j^+(\ell_2) \rangle_{\{x\}}^0 = \left[\mathbf{B}_{\ell_1} \cdots \mathbf{B}_{\ell_2+1} \frac{\mathbf{I}}{\mathbf{I} + \mathbf{B}_{\ell_2} \cdots \mathbf{B}_1 \mathbf{B}_{N_\tau} \cdots \mathbf{B}_{\ell_2+1}} \right]_{i,j}$$

Afin d'établir la connexion avec des résultats connus, considérons le cas où le champ extérieur ne dépend pas du temps imaginaire. Passant à la base où les \mathbf{B} sont diagonaux, on peut alors écrire,

$$\mathbf{B}_{i,j} = e^{-\Delta \tau (\varepsilon_j - \mu)} \delta_{i,j}$$

de telle sorte que l'équation encadrée ci-dessus pour la fonction de Green devient,

$$\langle c_i(\ell_1)c_j^+(\ell_2) \rangle_{\{x\}}^0 = \frac{1}{1 + e^{-\beta (\varepsilon_j - \mu)}} e^{-\Delta \tau (\ell_1 - \ell_2)(\varepsilon_j - \mu)} \delta_{i,j}$$

qui se réécrit sous la même forme que l'équation (2.75a) de la Ref.³¹:

$$\langle c_i(\ell_1)c_j^+(\ell_2) \rangle_{\{x\}}^0 = \left[1 - \frac{1}{1 + e^{\beta (\varepsilon_j - \mu)}} \right] e^{-[\Delta \tau (\ell_1 - \ell_2)](\varepsilon_j - \mu)} \delta_{i,j}.$$

Notons que lorsque $\ell_1 = \ell_2 + 0$ on retrouve le résultat bien connu:

$$\langle c_i(\ell_1)c_j^+(\ell_1) \rangle_{\{x\}}^0 = [1 - f(\varepsilon_j)] \delta_{i,j}$$

où f est la fonction de Fermi-Dirac. Ces définitions de la fonction de Green diffèrent par un signe de celle utilisée par plusieurs auteurs, dont Mahan par exemple. En fréquences de Matsubara, la fonction utilisée ici est de la forme (Eq.(2.131a) de la Réf.³¹).

$$G(i \omega_n) = \frac{-1}{i \omega_n - (\varepsilon - \mu)}.$$

Il faut aussi noter que la définition (10) correspond bien à l'équation (2.70) de la Réf.³¹

Références

1. Pour une introduction au Monte Carlo, incluant une discussion du Monte Carlo pour fonctions de Green, cf. M.H. Kalos et P.A. Whitlock, *Monte Carlo Methods*, (Wiley, New York, 1986).
2. D.M. Ceperley et M.H. Kalos, "Quantum Many-Body Problems" dans *Monte Carlo Methods in Statistical Physics*, K. Binder, Ed. (Springer Verlag, Berlin, 1979)
3. R. Blankenbecler, D.J. Scalapino, et R.L. Sugar, Phys. Rev. D24, 2278 (1981).
4. J.E. Hirsch, Phys. Rev. B31, 4403 (1985).
5. S.R. White, D.J. Scalapino, R.L. Sugar, E.Y. Loh, J.E. Gubernatis, R.T. Scalettar, Phys. Rev. B40, 506 (1989).
6. Pour une revue, voir E.Y. Loh, et J.E. Gubernatis, dans *Electron Phase Transitions*, p. 177-235, édité par W. Hauke, Y.V. Kopaev (Elsevier, Amsterdam, 1992)
7. J. Hubbard, Proc. Roy. Soc. A276, 238 (1963); *ibid.* A277, 237 (1964); *ibid.* A281, 410 (1964).
8. A.-M.S. Tremblay, C. Boily, Liang Chen, et A. Reid, dans *Supercomputing Symposium '90*, Ed. Dominique Pelletier, p.197-210 (École Polytechnique, Montréal, 1990).
9. E. Dagotto, A. Moreo, F. Ortolani, D. Poilblanc, et J. Riera, Phys. Rev. B 45, 10 741 (1992); E. Dagotto, Rev. Mod. Phys. (juillet 1994).
10. D.J. Scalapino "Does the Hubbard Model Have the Right Stuff?", pré-tirage, mars 93, USCB; D.J. Scalapino p. 314 dans *High Temperature Superconductivity, Proceedings of the Los Alamos Symposium, 1989*, Eds. K. Bedell, D. Coffey, D.E. Meltzer, D. Pines, et J.R. Schrieffer (Addison-Wesley, Redwood City, 1990).
11. H.F. Trotter, Proc. Am. Math. Soc. 10, 545 (1959); Les applications à la méthode Monte Carlo ont débutées avec M. Suzuki, Comm. Math. Phys. 51, 183 (1976) et M. Barma et B.S. Shastry, Phys. Rev. B 18, 3351 (1978).
12. R.M. Fye, Phys. Rev. B 33, 6271 (1986); R.M. Fye et R.T. Scalettar, Phys. Rev. B 36, 3833 (1987).
13. M. Suzuki, Phys. Lett. 113A, 299 (1985).
14. J.E. Hirsch, Phys. Rev. B 28, 4059 (1983).
15. Liang Chen, et A.-M.S. Tremblay, Int. J. Mod. Phys. B 6, 547 (1992).
16. Les différentes jauges peuvent être paramétrisées sans redondance par l'ensemble de toutes les directions dans \mathbf{R}^3 , *i.e.* par les éléments de RP^2 , soit la sphère unité S^2 dont les points antipodaux sont considérés identiques.

17. Il est à noter que dans le cas $U < 0$, on peut prouver⁴ que le déterminant pour les spins up a le même signe que pour les spins down, ce qui veut dire que le produit est toujours positif et qu'il n'y a pas de problème de signe. La symétrie particule-trou au demi-remplissage permet aussi de prouver que le produit des déterminants pour spins up et down est positif. Voir aussi l'annexe C.
18. R.M. Fye, and J.E. Hirsch, Phys. Rev. B **38**, 433 (1988).
19. W.H. Press, B.P. Flannery, S.A. Teukolsky, et W.T. Vetterling, *Numerical Recipes*, (Cambridge Univ. Press, Cambridge, 1989).
20. Y. Vilk, Liang Chen, A.-M.S. Tremblay, Phys. Rev. B (Mai 1994, sous presse)
21. A. Moreo, Phys. Rev. B **45**, 5059 (1992).
22. Anne-Marie Daré, Liang Chen, et A.-M.S. Tremblay, Phys. Rev. B **49**, 4106 (1994).
23. D.J. Scalapino, S.R. White et S.C. Zhang, Phys. Rev. Lett. **68**, 2830 (1992).
24. Sandro Sorella, Int. J. Mod. Phys. **B5**, 937 (1991).
25. R.T. Scalettar, R.M. Noack, and R.R.P. Singh, Phys. Rev. B **44**, 10 502 (1991).
27. Liang Chen, et A.-M.S. Tremblay, Phys. Rev. B **49**, 4338 (1994)
28. S.R. White et D.A. Huse, Phys. Rev. B **48**, 3844 (1993); S.R. White, Phys. Rev. Lett. **69**, 2863 (1992).
29. Dans le cas des décompositions du type hermitique $e^{-\Delta \tau \hat{K}_\ell / 2} e^{-\Delta \tau \hat{V}_\ell} e^{-\Delta \tau \hat{K}_\ell / 2}$, la première correction est donnée par un opérateur hermitique, de telle sorte que finalement la précision est aussi d'ordre $(\Delta \tau)^2$.
30. Ghassan George Batrouni, et Philippe de Forcrand, Phys. Rev. B **48**, 589 (1993).
31. J.W. Negele et H. Orland, *Quantum Many-Particle Systems*, (Addison-Wesley, Redwood City, 1988).

Table des Matières

1. Introduction.....	530
2. Les fermions sur réseau: motivation physique.....	531
2.1. Les électrons fortement corrélés dans des bandes étroites: Modèle de Hubbard.....	532
2.2. La difficulté des diagonalisations exactes.....	532
3. Monte Carlo quantique par la méthode des déterminants (en bref).....	533
3.1. Formule de Trotter.....	534
3.2. Découplage Hubbard-Stratonovich des interactions. Degrés de liberté de jauge.....	535
3.3. Trace sur les fermions.....	538
3.4. Interprétation physique.....	540
3.5. Calcul des observables.....	542
3.6. Poids pour l'échantillonnage selon l'importance: le "signe du déterminant".	545
4. Les calculs en pratique.....	546
4.1. Un cycle de mise-à-jour.....	546
4.2. Améliorations numériques nécessaires pour un programme qui fonctionne.	553
4.2.1. Stabilisation à basse température.....	553
4.2.2. Matrices éparées.....	555
5. Quelques résultats.....	557
5.1. Ordre antiferromagnétique à demi-remplissage et facteur de structure magnétique.....	557
5.2. Corrélations supraconductrices.....	559
5.3. Détection de la supraconductivité sans connaître la symétrie de la paire de Cooper.....	561
6. Limites de cette approche Monte Carlo quantique.....	562
6.1. Problèmes de signe. Effet de la jauge de la transformation Hubbard-Stratonovich.....	563
6.2. Efficacité de l'exploration de l'espace des phases et ergodicité.....	565
6.3. Influence du choix de l'estimateur.....	567
7. Conclusion.....	568
Annexe A.....	570
Erreur systématique due à la décomposition de Hubbard Stratonovich.....	570
Annexe B.....	572
Transformation de Hubbard-Stratonovich discrète avec choix de jauge.....	572
Annexe C.....	573

Traitement symétrique des fermions dans la transformation de Hubbard-Stratonovich	573
Annexe D.....	574
Évolution d'un état à N particules par un produit d'opérateurs d'évolution à une particule: théorèmes pour l'évaluation de la trace fermionique.	574
Annexe E.....	577
Fonction de Green à temps égal.....	577
Annexe F.....	578
Fonctions de Green à temps inégal.....	578
Références.....	581