(d) Compare your results with those from the exact diagonalization with and without the $4p4h$ excitation. Compare also your results to perturbation theory at different orders, in particular to second order. Discuss your results.

8.13 Derive the amplitude equations of Eq. (8.31) starting with

$$0 = \langle \Phi^{a_1 \dots a_n}_{i_1 \dots i_n} | \overline{H} | \Phi_0 \rangle.$$

8.14 Replace the Minnesota interaction model with realistic models for nuclear forces based on effective field theory. In particular replace the Minnesota interaction with the low-order (LO) contribution which includes a contact term and a one-pion exchange term only. The expressions are discussed in Sect. 8.2.4 and Eq. (8.4). Reference [62] contains a detailed compilation of all terms up to order NNLO, with tabulated values for all constants. When adding realistic interaction models we recommend that you use the many-body perturbation theory codes to second order in the interaction, see the code link at https://github.com/ManyBodyPhysics/LectureNotesPhysics/tree/master/Programs/Chapter8-programs/cpp/MBPT2/src/.

8.10 Solutions to Selected Exercises

8.1 To solve this problem, we start by introducing the shorthand label for single-particle states below the Fermi level F as $i, j, \dots \leq F$. For single-particle states above the Fermi level we reserve the labels $a, b, \dots > F$, while the labels $p, q, \dots$ represent any possible single particle state. Using the ansatz for the ground state $|\Phi_0$ as new reference vacuum state, the anticommutation relations are

$$\{a_p^\dagger, a_q\} = \delta_{pq}, p, q \leq F,$$

and

$$\{a_p, a_q^\dagger\} = \delta_{pq}, \ p, q > F.$$

It is easy to see then that

$$a_i|\Phi_0\rangle = |\Phi_i\rangle \neq 0, \qquad a_a^\dagger|\Phi_0\rangle = |\Phi^a\rangle \neq 0,$$

and

$$a_i^\dagger|\Phi_0\rangle = 0 \qquad a_a|\Phi_0\rangle = 0.$$

We can then rewrite the one-body Hamiltonian as

$$\begin{aligned}\hat{H}_0 &= \sum_{pq} \langle p|\, \hat{h}_0\, |q\rangle\, a_p^\dagger a_q \\ &= \sum_{pq} \langle p|\, \hat{h}_0\, |q\rangle\, \{a_p^\dagger a_q\} + \delta_{pq\in i} \sum_{pq} \langle p|\, \hat{h}_0\, |q\rangle \\ &= \sum_{pq} \langle p|\, \hat{h}_0\, |q\rangle\, \{a_p^\dagger a_q\} + \sum_{i} \langle i|\, \hat{h}_0\, |i\rangle\,,\end{aligned}$$

where the curly brackets represent normal-ordering with respect to the new vacuum state. Withe respect to the new vacuum reference state, the

8.2 Using our anti-commutation rules, Wick's theorem discussed in the appendix and the definition of the creation and annihilation operators from the previous problem, we can rewrite the set of creation and annihilation operators of

$$\hat{H}_I = \frac{1}{4} \sum_{pqrs} \langle pq|\, \hat{v}\, |rs\rangle\, a_p^\dagger a_q^\dagger a_s a_r$$

as

$$\begin{aligned}a_p^\dagger a_q^\dagger a_s a_r &= \{a_p^\dagger a_q^\dagger a_s a_r\} \\ &\quad + \{a_p^\dagger \overbrace{a_q^\dagger a_s} a_r\} + \{a_p^\dagger \overbrace{a_q^\dagger a_s a_r}\} + \{\overbrace{a_p^\dagger a_q^\dagger} a_s a_r\} \\ &\quad + \{\overbrace{a_p^\dagger a_q^\dagger a_s} a_r\} + \{\overbrace{a_p^\dagger \overbrace{a_q^\dagger a_s} a_r}\} + \{\overbrace{a_p^\dagger \overbrace{a_q^\dagger a_s a_r}}\} \\ &= \{a_p^\dagger a_q^\dagger a_s a_r\} + \delta_{qs\in i} \{a_p^\dagger a_r\} - \delta_{qr\in i} \{a_p^\dagger a_s\} - \delta_{ps\in i} \{a_q^\dagger a_r\} \\ &\quad + \delta_{pr\in i} \{a_q^\dagger a_s\} + \delta_{pr\in i}\delta_{qs\in i} - \delta_{ps\in i}\delta_{qr\in i}.\end{aligned}$$

Inserting the redefinition of the creation and annihilation operators with respect to the new vacuum state, we have

$$\begin{aligned}\hat{H}_I &= \frac{1}{4} \sum_{pqrs} \langle pq|\, \hat{v}\, |rs\rangle\, a_p^\dagger a_q^\dagger a_s a_r \\ &= \frac{1}{4} \sum_{pqrs} \langle pq|\, \hat{v}\, |rs\rangle\, \{a_p^\dagger a_q^\dagger a_s a_r\} + \frac{1}{4} \sum_{pqrs} \Big(\delta_{qs\in i} \langle pq|\, \hat{v}\, |rs\rangle\, \{a_p^\dagger a_r\} \\ &\quad - \delta_{qr\in i} \langle pq|\, \hat{v}\, |rs\rangle\, \{a_p^\dagger a_s\} - \delta_{ps\in i} \langle pq|\, \hat{v}\, |rs\rangle\, \{a_q^\dagger a_r\} \\ &\quad + \delta_{pr\in i} \langle pq|\, \hat{v}\, |rs\rangle\, \{a_q^\dagger a_s\} + \delta_{pr\in i}\delta_{qs\in i} - \delta_{ps\in i}\delta_{qr\in i} \Big)\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{4}\sum_{pqrs} \langle pq|\,\hat{v}\,|rs\rangle \left\{a_p^\dagger a_q^\dagger a_s a_r\right\} \\
&\quad + \frac{1}{4}\sum_{pqi} \Big(\langle pi|\,\hat{v}\,|qi\rangle - \langle pi|\,\hat{v}\,|iq\rangle - \langle ip|\,\hat{v}\,|qi\rangle + \langle ip|\,\hat{v}\,|iq\rangle \Big) \left\{a_p^\dagger a_q\right\} \\
&\quad + \frac{1}{4}\sum_{ij} \Big(\langle ij|\,\hat{v}\,|ij\rangle - \langle ij|\,\hat{v}\,|ji\rangle \Big) \\
&= \frac{1}{4}\sum_{pqrs} \langle pq|\,\hat{v}\,|rs\rangle \left\{a_p^\dagger a_q^\dagger a_s a_r\right\} + \sum_{pqi} \langle pi|\,\hat{v}\,|qi\rangle \left\{a_p^\dagger a_q\right\} + \frac{1}{2}\sum_{ij} \langle ij|\,\hat{v}\,|ij\rangle \,.
\end{aligned}
$$

Summing up, we obtain a two-body part defined as

$$
\hat{V}_N = \frac{1}{4}\sum_{pqrs} \langle pq|\,\hat{v}\,|rs\rangle \left\{a_p^\dagger a_q^\dagger a_s a_r\right\},
$$

a one-body part given by

$$
\hat{F}_N = \sum_{pqi} \langle pi|\,\hat{v}\,|qi\rangle \left\{a_p^\dagger a_q\right\},
$$

and finally the so-called reference energy

$$
E_{\text{ref}} = \frac{1}{2}\sum_{ij} \langle ij|\,\hat{v}\,|ij\rangle \,.
$$

which is the energy expectation value for the reference state. Thus, our normal-ordered Hamiltonian with at most a two-body nucleon-nucleon interaction is defined as

$$
\hat{H}_N = \frac{1}{4}\sum_{pqrs} \langle pq|\,\hat{v}\,|rs\rangle \left\{a_p^\dagger a_q^\dagger a_s a_r\right\} + \sum_{pq} f_q^p \left\{a_p^\dagger a_q\right\} = \hat{V}_N + \hat{F}_N,
$$

with

$$
\hat{F}_N = \sum_{pq} f_q^p \left\{a_p^\dagger a_q\right\},
$$

and

$$
\hat{V}_N = \frac{1}{4}\sum_{pqrs} \langle pq|\,\hat{v}\,|rs\rangle \left\{a_p^\dagger a_q^\dagger a_s a_r\right\},
$$

where

$$f_q^p = \langle p|\,\hat{h}_0\,|q\rangle + \sum_i \langle pi|\,\hat{v}\,|qi\rangle$$

8.4 The following python code sets up the quantum numbers for both infinite nuclear matter and neutron matter employing a cutoff in the value of n. The full code can be found at https://github.com/ManyBodyPhysics/LectureNotesPhysics/tree/master/Programs/Chapter8-programs/python/spstatescc.py.

```
from numpy import *

nmax =2
nshell = 3*nmax*nmax
count = 1
tzmin = 1

print "Symmetric nuclear matter:"
print "a, nx, ny, nz, sz, tz, nx^2 + ny^2 + nz^2"
for n in range(nshell):
   for nx in range(-nmax,nmax+1):
       for ny in range(-nmax,nmax+1):
         for nz in range(-nmax, nmax+1):
            for sz in range(-1,1+1):
               tz = 1
               for tz in range(-tzmin,tzmin+1):
                  e = nx*nx + ny*ny + nz*nz
                  if e == n:
                     if sz != 0:
                        if tz != 0:
                           print count, " ",nx," ",ny, " ",nz,"
                               ",sz," ",tz," ",e
                           count += 1


nmax =1
nshell = 3*nmax*nmax
count = 1
tzmin = 1
print "-----------------------------------"
print "Neutron matter:"
print "a, nx, ny, nz, sz, nx^2 + ny^2 + nz^2"
for n in range(nshell):
   for nx in range(-nmax,nmax+1):
       for ny in range(-nmax,nmax+1):
         for nz in range(-nmax, nmax+1):
            for sz in range(-1,1+1):
               e = nx*nx + ny*ny + nz*nz
               if e == n:
                  if sz != 0:
                     print count, " ",nx," ",ny, " ",sz," ",tz,"
                          ",e
                     count += 1
```

Acknowledgements We are much indebted to Carlo Barbieri, Scott Bogner, Alex Brown, David Dean, Heiko Hergert, Dean Lee, Titus Morris, Thomas Papenbrock, Nathan Parzuchowski, Piotr Piecuch and Fei Yuan for many discussions on many-body theory. Computational resources were provided by Michigan State University and the Norwegian Notur project (Supercomputing grant NN2977K). This work was supported by NSF Grant No. PHY-1404159 (Michigan State University), by the Oce of Nuclear Physics, U.S. Department of Energy, under grants DE-FG02-96ER40963, DE-SC0008499 (NUCLEI SciDAC collaboration) and the Field Work Proposal ERKBP57 at Oak Ridge National Laboratory (ORNL).

This research used also resources of the Oak Ridge Leadership Computing Facility located at ORNL, which is supported by the Office of Science of the Department of Energy under Contract No. DE-AC05-00OR22725.

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The U.S. Government retains a nonexclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. Government or any U.S. Government agency. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan.

Appendix, Wick's Theorem

Wick's theorem is based on two fundamental concepts, namely *normal ordering* and *contraction*. The normal-ordered form of $\hat{A}\hat{B}..\hat{X}\hat{Y}$, where the individual terms are either a creation or annihilation operator, is defined as

$$\left\{\hat{A}\hat{B}..\hat{X}\hat{Y}\right\} \equiv (-1)^p \,[\text{creation operators}] \cdot [\text{annihilation operators}]\,. \tag{8.59}$$

The p subscript denotes the number of permutations that is needed to transform the original string into the normal-ordered form. A contraction between to arbitrary operators $\hat{X}$ and $\hat{Y}$ is defined as

$$\hat{X}\hat{Y} \equiv \langle 0|\hat{X}\hat{Y}|0\rangle. \tag{8.60}$$

It is also possible to contract operators inside a normal ordered products. We define the original relative position between two operators in a normal ordered product as p, the so-called permutation number. This is the number of permutations needed to bring one of the two operators next to the other one. A contraction between two operators with $p \neq 0$ inside a normal ordered is defined as

$$\left\{\overbrace{\hat{A}\hat{B}..\hat{X}\hat{Y}}\right\} = (-1)^p \left\{\hat{A}\hat{B}..\hat{X}\hat{Y}\right\}. \tag{8.61}$$

In the general case with m contractions, the procedure is similar, and the prefactor changes to

$$(-1)^{p_1+p_2+..+p_m}. \tag{8.62}$$

Wick's theorem states that every string of creation and annihilation operators can be written as a sum of normal ordered products with all possible ways of contractions,

$$\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z} = \left\{\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z}\right\} \tag{8.63}$$

$$+ \sum_{[1]} \left\{\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z}\right\} \tag{8.64}$$

$$+ \sum_{[2]} \left\{\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z}\right\} \tag{8.65}$$

$$+ \ldots \tag{8.66}$$

$$+ \sum_{[\frac{N}{2}]} \left\{\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z}\right\}. \tag{8.67}$$

The $\sum_{[m]}$ means the sum over all terms with m contractions, while $\left[\frac{N}{2}\right]$ means the largest integer that not do not exceeds $\frac{N}{2}$ where N is the number of creation and annihilation operators. When N is even,

$$\left[\frac{N}{2}\right] = \frac{N}{2}, \tag{8.68}$$

and the last sum in Eq. (8.63) is over fully contracted terms. When N is odd,

$$\left[\frac{N}{2}\right] \neq \frac{N}{2}, \tag{8.69}$$

and none of the terms in Eq. (8.63) are fully contracted.

An important extension of Wick's theorem allow us to define contractions between normal-ordered strings of operators. This is the so-called generalized Wick's theorem,

$$\left\{\hat{A}\hat{B}\hat{C}\hat{D}..\right\}\left\{\hat{R}\hat{X}\hat{Y}\hat{Z}..\right\} = \left\{\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z}\right\} \tag{8.70}$$

$$+ \sum_{[1]} \left\{\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z}\right\} \tag{8.71}$$

$$+ \sum_{[2]} \left\{ \overbrace{\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z}}^{} \right\} \tag{8.72}$$

$$+ \ldots \tag{8.73}$$

Turning back to the many-body problem, the vacuum expectation value of products of creation and annihilation operators can be written, according to Wick's theorem in Eq. (8.63), as a sum over normal ordered products with all possible numbers and combinations of contractions,

$$\langle 0|\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z}|0\rangle = \langle 0| \left\{ \hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z} \right\} |0\rangle \tag{8.74}$$

$$+ \sum_{[1]} \langle 0| \left\{ \overbrace{\hat{A}\hat{B}}^{}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z} \right\} |0\rangle \tag{8.75}$$

$$+ \sum_{[2]} \langle 0| \left\{ \overbrace{\hat{A}\hat{B}\overbrace{\hat{C}\hat{D}}^{}}^{}..\hat{R}\hat{X}\hat{Y}\hat{Z} \right\} |0\rangle \tag{8.76}$$

$$+ \ldots \tag{8.77}$$

$$+ \sum_{[\frac{N}{2}]} \langle 0| \left\{ \overbrace{\hat{A}\hat{B}\overbrace{\hat{C}\hat{D}}^{}}^{}.. \overbrace{\hat{R}\hat{X}\hat{Y}\hat{Z}}^{} \right\} |0\rangle. \tag{8.78}$$

All vacuum expectation values of normal ordered products without fully contracted terms are zero. Hence, the only contributions to the expectation value are those terms that *is* fully contracted,

$$\langle 0|\hat{A}\hat{B}\hat{C}\hat{D}..\hat{R}\hat{X}\hat{Y}\hat{Z}|0\rangle = \sum_{[all]} \langle 0| \left\{ \overbrace{\hat{A}\hat{B}\overbrace{\hat{C}\hat{D}}^{}}^{}.. \overbrace{\hat{R}\hat{X}\hat{Y}\hat{Z}}^{} \right\} |0\rangle \tag{8.79}$$

$$= \sum_{[all]} \overbrace{\hat{A}\hat{B}\overbrace{\hat{C}\hat{D}}^{}}^{}.. \overbrace{\hat{R}\hat{X}\hat{Y}\hat{Z}}^{}. \tag{8.80}$$

To obtain fully contracted terms, Eq. (8.68) must hold. When the number of creation and annihilation operators is odd, the vacuum expectation value can be set to zero at once. When the number is even, the expectation value is simply the sum of terms with all possible combinations of fully contracted terms. Observing that the only contractions that give nonzero contributions are

$$\overbrace{a_\alpha a_\beta^\dagger}^{} = \delta_{\alpha\beta}, \tag{8.81}$$

the terms that contribute are reduced even more.

Wick's theorem provides us with an algebraic method for easy determination of the terms that contribute to the matrix element.

References

1. S.L. Shapiro, S.A. Teukolsky, *Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects* (Wiley, New York, 1983)
2. J.M. Lattimer, M. Prakash, Astrophys. J. **550**, 426 (2001)
3. J.M. Lattimer, M. Prakash, Phys. Rep. **442**, 109 (2007)
4. A.W. Steiner, J.M. Lattimer, E.F. Brown, Astrophys J. **722**, 33 (2010)
5. A.W. Steiner, S. Gandolfi, Phys. Rev. Lett. **108**, 081102 (2012)
6. J.M. Lattimer, Ann. Rev. Nucl. Part. Sci. **62**, 485 (2012)
7. F. Weber, *Pulsars as Astrophysical Laboratories for Nuclear and Particle Physics* (Institute of Physics Publishing, London, 1999)
8. H. Heiselberg, M. Hjorth-Jensen, Phys. Rep. **328**, 237 (2000)
9. S. Weinberg, Phys. Lett. B **251**, 288 (1990)
10. S. Weinberg, Nucl. Phys. B **363**, 3 (1991)
11. C. Ordonez, U. van Kolck, Phys. Lett. B **291**, 459 (1992)
12. C. Ordonez, L. Ray, U. van Kolck, Phys. Rev. Lett. **72**, 1982 (1994)
13. U. van Kolck, Phys. Rev. C **49**, 2932 (1994)
14. U. van Kolck, Prog. Part. Nucl. Phys. **43**, 337 (1999)
15. R. Machleidt, D. Entem, Phys. Rep. **503**, 1 (2011)
16. E. Epelbaum, H.W. Hammer, U.G. Meißner, Rev. Mod. Phys. **81**, 1773 (2009)
17. G. Hagen, T. Papenbrock, M. Hjorth-Jensen, D.J. Dean, Rep. Prog. Phys. **77**, 096302 (2014)
18. A. Ekström, G.R. Jansen, K.A. Wendt, G. Hagen, T. Papenbrock, B.D. Carlsson, C. Forssén, M. Hjorth-Jensen, P. Navrátil, W. Nazarewicz, Phys. Rev. C **91**, 051301 (2015)
19. G. Hagen, M. Hjorth-Jensen, G.R. Jansen, T. Papenbrock, Phys. Scr. **91**, 063006 (2016)
20. L. Coraggio, J.W. Holt, N. Itaco, R. Machleidt, L.E. Marcucci, F. Sammarruca, Phys. Rev. C **89**, 044321 (2014)
21. F. Sammarruca, L. Coraggio, J.W. Holt, N. Itaco, R. Machleidt, L.E. Marcucci, Phys. Rev. C **91**, 054311 (2015)
22. R. Roth, S. Binder, K. Vobig, A. Calci, J. Langhammer, P. Navrátil, Phys. Rev. Lett. **109**, 052501 (2012)
23. S. Binder, P. Piecuch, A. Calci, J. Langhammer, P. Navrátil, R. Roth, Phys. Rev. C **88**, 054319 (2013)
24. H. Hergert, S. Bogner, T. Morris, A. Schwenk, K. Tsukiyama, Phys. Rep. **621**, 165 (2016)
25. P. Navratil, S. Quaglioni, G. Hupin, C. Romero-Redondo, A. Calci, Phys. Scr. **91**, 053002 (2016)
26. A. Carbone, A. Polls, A. Rios, Phys. Rev. C **88**, 044302 (2013)
27. T. Inoue, S. Aoki, T. Doi, T. Hatsuda, Y. Ikeda, N. Ishii, K. Murano, H. Nemura, K. Sasaki, Phys. Rev. Lett. **111**, 112503 (2013)
28. B.D. Day, Rev. Mod. Phys. **39**, 719 (1967)
29. K.A. Brueckner, C.A. Levinson, H.M. Mahmoud, Phys. Rev. **95**, 217 (1954)
30. K.A. Brueckner, Phys. Rev. **100**, 36 (1955)
31. M. Baldo, G.F. Burgio, Rep. Prog. Phys. **75**, 026301 (2012)
32. M. Baldo, A. Polls, A. Rios, H.J. Schulze, I. Vidaña, Phys. Rev. C **86**, 064001 (2012)
33. L. Coraggio, J.W. Holt, N. Itaco, R. Machleidt, F. Sammarruca, Phys. Rev. C **87**, 014322 (2013)
34. R.J. Bartlett, M. Musiał, Rev. Mod. Phys. **79**, 291 (2007)
35. I. Shavitt, R.J. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, Cambridge, 2009)

36. G. Baardsen, A. Ekström, G. Hagen, M. Hjorth-Jensen, Phys. Rev. C **88**, 054312 (2013)
37. G. Hagen, T. Papenbrock, A. Ekström, K.A. Wendt, G. Baardsen, S. Gandolfi, M. Hjorth-Jensen, C.J. Horowitz, Phys. Rev. C **89**, 014319 (2014)
38. G. Hagen, G.R. Jansen, T. Papenbrock, Phys. Rev. Lett. **117**, 172501 (2016)
39. J. Carlson, J. Morales, V.R. Pandharipande, D.G. Ravenhall, Phys. Rev. C **68**, 025802 (2003)
40. S. Gandolfi, A.Y. Illarionov, K.E. Schmidt, F. Pederiva, S. Fantoni, Phys. Rev. C **79**, 054005 (2009)
41. S. Gandolfi, A.Y. Illarionov, F. Pederiva, K.E. Schmidt, S. Fantoni, Phys. Rev. C **80**, 045802 (2009)
42. A. Gezerlis, J. Carlson, Phys. Rev. C **81**, 025803 (2010)
43. S. Gandolfi, J. Carlson, S. Reddy, Phys. Rev. C **85**, 032801 (2012)
44. A. Lovato, O. Benhar, S. Fantoni, K.E. Schmidt, Phys. Rev. C **85**, 024003 (2012)
45. A. Gezerlis, I. Tews, E. Epelbaum, S. Gandolfi, K. Hebeler, A. Nogga, A. Schwenk, Phys. Rev. Lett. **111**, 032501 (2013)
46. J. Carlson, S. Gandolfi, F. Pederiva, S.C. Pieper, R. Schiavilla, K.E. Schmidt, R.B. Wiringa, Rev. Mod. Phys. **87**, 1067 (2015)
47. V. Somà, A. Cipollone, C. Barbieri, P. Navrátil, T. Duguet, Phys. Rev. C **89**, 061301 (2014)
48. W. Dickhoff, C. Barbieri, Prog. Part. Nucl. Phys. **52**, 377 (2004)
49. T.D. Morris, N.M. Parzuchowski, S.K. Bogner, Phys. Rev. C **92**, 034331 (2015)
50. M. Prakash, I. Bombaci, M. Prakash, P.J. Ellis, J.M. Lattimer, R. Knorren, Phys. Rep. **280**, 1 (1997)
51. C.J. Horowitz, E.F. Brown, Y. Kim, W.G. Lynch, R. Michaels, A. Ono, J. Piekarewicz, M.B. Tsang, H.H. Wolter, J. Phys. G **41**, 093001 (2014)
52. J.P. Blaizot, G. Ripka, *Quantum Theory of Finite Systems* (MIT press, Cambridge, 1986)
53. W. Dickhoff, D. Van Neck, *Many-Body Theory Exposed!: Propagator Description of Quantum Mechanics in Many-Body Systems*, 2nd edn. (World Scientific, New Jersey, 2008)
54. R.D. Mattuck, *A guide to Feynman Diagrams in the Many-Body Problem*, 2nd edn. (Dover, New York, 1992)
55. S. Binder, A. Ekström, G. Hagen, T. Papenbrock, K.A. Wendt, Phys. Rev. C **93**, 044332 (2016)
56. K.S. McElvain, W.C. Haxton, Nuclear Physics without High-Momentum Potentials: Direct Construction of the Effective Interaction from Scattering Observables (2016). ArXiv:1607.06863
57. D. Thompson, M. Lemere, Y. Tang, Nucl. Phys. A **286**, 53 (1977)
58. S. Weinberg, Phys. Lett. B **363**, 288 (1990)
59. C. Ordóñez, L. Ray, U. van Kolck, Phys. Rev. C **53**, 2086 (1996)
60. U. van Kolck, Phys. Rev. C **49**, 2932 (1994)
61. A. Ekström, G. Baardsen, C. Forssén, G. Hagen, M. Hjorth-Jensen, G.R. Jansen, R. Machleidt, W. Nazarewicz, T. Papenbrock, J. Sarich, S.M. Wild, Phys. Rev. Lett. **110**, 192502 (2013)
62. B.D. Carlsson, A. Ekström, C. Forssén, D.F. Strömberg, G.R. Jansen, O. Lilja, M. Lindby, B.A. Mattsson, K.A. Wendt, Phys. Rev. X **6**, 011019 (2016)
63. E. Epelbaum, A. Nogga, W. Glöckle, H. Kamada, U.G. Meißner, H. Witała, Phys. Rev. C **66**, 064001 (2002)
64. P. Navrátil, S. Quaglioni, I. Stetcu, B.R. Barrett, J. Phys. G **36**(8), 083101 (2009)
65. G. Golub, C. Van Loan, *Matrix Computations* (John Hopkins University Press, Baltimore, 1996)
66. C. Lanczos, J. Res. Natl. Bur. Stand. **45**, 255 (1950)
67. E.R. Davidson, Comput. Phys. Commun. **53**, 49 (1989)
68. E.R. Davidson, Comput. Phys. **5**, 519 (1993)
69. F. Coester, Nucl. Phys. **7**, 421 (1958)
70. F. Coester, H. Kümmel, Nucl. Phys. **17**, 477 (1960)
71. H. Kümmel, K.H. Lührmann, J.G. Zabolitzky, Phys. Rep. **36**, 1 (1978)
72. B.H. Brandow, Rev. Mod. Phys. **39**, 771 (1967)
73. R.J. Bartlett, Ann. Rev. Phys. Chem. **32**, 359 (1981)

74. G.R. Jansen, M.D. Schuster, A. Signoracci, G. Hagen, P. Navrátil, Phys. Rev. C **94**, 011301 (2016)
75. A. Baran, A. Bulgac, M.M. Forbes, G. Hagen, W. Nazarewicz, N. Schunck, M.V. Stoitsov, Phys. Rev. C **78**, 014318 (2008)
76. L.S. Blackford, J. Demmel, I. Du, G. Henry, M. Heroux, L. Kaufman, A. Lumsdaine, A. Petitet, R.C. Whaley, ACM Trans. Math. Softw. **28**, 135 (2002)
77. W. Gropp, E. Lusk, A. Skjellum, *Using MPI* (MIT Press, Cambridge, 1999)
78. E. Gabriel, G.E. Fagg, G. Bosilca, T. Angskun, J.J. Dongarra, J.M. Squyres, V. Sahay, P. Kambadur, B. Barrett, A. Lumsdaine, R.H. Castain, D.J. Daniel, R.L. Graham, T.S. Woodall, in *Proceedings, 11th European PVM/MPI Users' Group Meeting*, Budapest (2004), p. 97
79. R.C. Martin, *Clean Code: A Handbook of Agile Software Craftsmanship* (Prentice Hall, Upper Saddle River, 2015)
80. C. Gros, Zeit. Phys. B **86**, 359 (1992)
81. C. Gros, Phys. Rev. B **53**, 6865 (1996)
82. J.J. Shepherd, A. Grüneis, G.H. Booth, G. Kresse, A. Alavi, Phys. Rev. B **86**, 035111 (2012)
83. J.J. Shepherd, J. Chem. Phys. **145**, 031104 (2016)
84. C. Lin, F.H. Zong, D.M. Ceperley, Phys. Rev. E **64**, 016702 (2001)
85. B. Chapman, G. Jost, R. van der Pas, *Using OpenMP. Portable Shared Memory Parallel Programming* (MIT Press, Cambridge, 2008)

Chapter 9
Variational and Diffusion Monte Carlo Approaches to the Nuclear Few- and Many-Body Problem

Francesco Pederiva, Alessandro Roggero, and Kevin E. Schmidt

9.1 Monte Carlo Methods in Quantum Many-Body Physics

9.1.1 Expectations in Quantum Mechanics

In the previous chapters the authors pointed out in several different ways that the non-relativistic quantum many-body problem is equivalent to the solution of a very complicated differential equation, the many-body Schrödinger equation.

As it was illustrated, in the few-body case ($A < 6$) it possible to find compute exact solutions. At the very least, one can expand the eigenfunctions on a basis set including $\mathscr{M}$ elements, diagonalize the Hamiltonian matrix, and try to reach convergence as a function of $\mathscr{M}$. Unfortunately, this procedure becomes more and more expensive when the number of bodies A increases. There are many ingenuous ways to improve the speed of convergence and the quality of the results. The price to pay often is the introduction of more or less controlled approximations.

F. Pederiva (✉)
Physics Department, University of Trento, and INFN-TIFPA, via Sommarive 14, 38123 Trento, Italy
e-mail: francesco.pederiva@unitn.it

A. Roggero
Institute for Nuclear Theory, University of Washington, Seattle, WA 98195-1550, USA
e-mail: roggero@uw.edu

K.E. Schmidt
Department of Physics and Astronomy, Arizona State University, Tempe, AZ 85283-1506, USA
e-mail: kevin.schmidt@asu.edu

M. Hjorth-Jensen et al. (eds.), *An Advanced Course in Computational Nuclear Physics*, Lecture Notes in Physics 936, DOI 10.1007/978-3-319-53336-0_9

All these approaches have one common feature: they end up with some closed expression for the eigenfunctions. However, we should remember that the wavefunction per se is not an observable. In order to make predictions to be compared with experiments, we only need a way to compute *expectations* of operators $\hat{O}$ describing the observables we are interested in.

Given a many-body Hamiltonian $\hat{H}$, we might want, for instance, to look for the ground state eigenfunction and eigenvalue. This means that we want to solve the following equation:

$$\hat{H}|\Psi_0\rangle = E_0|\Psi_0\rangle. \tag{9.1}$$

At this point we to provide a representation of the Hilbert space in term of some basis set. This set will be denoted as $\{|X\rangle\}$. Its elements could be eigenstates of the position or of the momentum operators, or eigenstates of a simpler Hamiltonian of which we know the exact spectrum. In order to make the notation less cumbersome, we will assume that the quantum numbers X characterizing the basis states are in the continuum. In the case of a discrete spectrum, integrals in the following have to be replaced by sums over all their possible values, without any loss of generality. As an example, X could include the positions or the momenta of A nucleons, and their spin and isospin values.

All the physical information we need about the time-independent problem is then included in integrals of the form:

$$\langle O\rangle \equiv \langle\Psi_0|\hat{O}\Psi_0\rangle = \frac{\displaystyle\int dXdX'\langle\Psi_0|X\rangle\langle X|\hat{O}|X'\rangle\langle X'|\Psi_0\rangle}{\displaystyle\int dX|\langle X|\Psi_0\rangle|^2}. \tag{9.2}$$

These integrals are apparently as hard to solve as the Schrödinger equation itself, even if we had access to the explicit form of the wavefunction. Is there any real gain in reformulating the problem this way?

We can first notice that expectations can in general be written in a slightly different form, independent of the nature of the operator $\hat{O}$:

$$\langle O\rangle = \frac{\displaystyle\int dX|\langle X|\Psi_0\rangle|^2\frac{\langle X|\hat{O}\Psi_0\rangle}{\langle X|\Psi_0\rangle}}{\displaystyle\int dX|\langle X|\Psi_0\rangle|^2}. \tag{9.3}$$

For the moment we will just assume that the quotient appearing at numerator of the expectation is always well defined, and we will later discuss this aspect in more detail. The standard quantum mechanical interpretation of the wavefunction tells us

that the quantity:

$$P[X] = \frac{|\langle X|\Psi_0\rangle|^2}{\int dX|\langle X|\Psi_0\rangle|^2}, \tag{9.4}$$

is the probability density of finding the system in the state $|X\rangle$ labeled by the set of quantum numbers X. Thereby, the expectation integral has the general form:

$$\langle O\rangle = \int dXP[X]\frac{\langle X|\hat{O}\Psi_0\rangle}{\langle X|\Psi_0\rangle}, \tag{9.5}$$

i.e. the average of what we will call the *local* operator $O_{loc} \equiv \frac{\langle X|\hat{O}\Psi_0\rangle}{\langle X|\Psi_0\rangle}$ weighted with the probability of finding the system in a given state $|X\rangle$. Integrals like that in Eq. (9.5) have a direct physical interpretation. In a measurement process what we would observe is essentially the result of a *sampling process* of $P[X]$. The expectation of our operator is approximated by:

$$\langle O\rangle \simeq \frac{1}{M}\sum_{k=1}^{M} O(X_k), \tag{9.6}$$

where M is the number of measurements performed, and $O(X_k)$ is a shorthand notation to indicate the value assumed by the observable $\hat{O}$ in the state labeled by the quantum numbers X_k. The laws of statistics also give us a way of estimating a *statistical* error on $\langle O\rangle$, and we know that the error decreases by increasing the number of measurements.

There is here an important point to notice: in a physical measurement process we have *no direct knowledge of the wavefunction*, we just *sample* its squared modulus!

This argument suggests that if we had a numerical way of sampling the squared modulus of a wavefunction, we could in principle compute expectations and make comparisons with experiments without needing an explicit expression of the wavefunction itself. Quantum Monte Carlo methods aim exactly at solving the many-body Schrödinger equation by sampling its solutions, eventually without any need of an explicit analytical form.

The remainder of this chapter will be organized as follows. First we will discuss how to perform calculations based on an accurate, explicit ansatz for the wavefunction of an A-body system interacting via a purely central potential, exploiting the variational principle of quantum mechanics (Variational Monte Carlo methods). Then we will discuss how to sample the exact ground state of the system by projecting it out of an initial ansatz (Projection Monte Carlo methods). Finally, we will see how these methods need to be extended when we are interested in studying Hamiltonians that have an explicit dependence on the spin and isospin states of the particles, as it happens for the modern interactions employed in nuclear physics.

9.2 Variational Wavefunctions and VMC for Central Potentials

9.2.1 Coordinate Space Formulation

As previously discussed, we are in principle free to choose any representation of the Hilbert space of the system we like, in order to compute expectations. The most convenient choice, for a system of particles interacting via a purely central potential, with no explicit dependence on the spin or isospin state, is to use the eigenstates of the position operator. If $R = \mathbf{r}_1, \ldots \mathbf{r}_A$ are the coordinates of the A (identical)[1] particles of mass m constituting the system, we have that:

$$|X\rangle \equiv |R\rangle \tag{9.7}$$

with the normalization:

$$\langle R'|R\rangle = \delta(R - R')\,. \tag{9.8}$$

Notice that we are here considering a $3A$-dimensional Cartesian space, without decomposing it in the product of A3-dimensional spaces. In this representation the wavefunction is simply given by:

$$\langle R|\Psi_0\rangle \equiv \Psi_0(R) = \Psi_0(\mathbf{r}_1, \ldots \mathbf{r}_A). \tag{9.9}$$

The Hamiltonian instead reads:

$$\hat{H} = \sum_{i=1}^{A} \frac{p_i^2}{2m} + V(\mathbf{r}_1, \ldots \mathbf{r}_A), \tag{9.10}$$

or

$$\hat{H} = \int dR|R\rangle \left[-\frac{\hbar^2}{2m} \sum_{i=1}^{A} \nabla_i^2 + V(\mathbf{r}_1, \ldots \mathbf{r}_A) \right] \langle R|\,, \tag{9.11}$$

where V is the interparticle potential. Substituting this form into Eq. (9.1), operating from the left with $\langle R|$ gives the Schrödinger differential equation

$$\left[-\frac{\hbar^2}{2m} \sum_{i=1}^{A} \nabla_i^2 + V(\mathbf{r}_1, \ldots \mathbf{r}_A) \right] \Psi_0(R) = E_0 \Psi_0(R)\,. \tag{9.12}$$

[1] We will always refer to systems of identical particle throughout the text. The generalization to mixtures is normally straightforward, and it will not be discussed here.

We will often use the same symbol for the Hilbert space operator and its differential form and write this simply as $\hat{H}\Psi_0(R) = E_0\Psi_0(R)$; whether the operator or differential form is used can be discerned readily from context. In this representation the states of the Hilbert space are sampled by sampling the particle positions from the squared modulus of the wavefunction $|\Psi_0(R)|^2$.

9.2.2 Variational Principle and Variational Wavefunctions

As already seen in the previous chapters, one of the possible ways to approximate a solution of the many-body Schrödinger equation is to exploit the variational principle. Given a *trial state* $|\Psi_T\rangle$, the following inequality holds:

$$E_T = \frac{\langle\Psi_T|\hat{H}\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} \geq E_0, \tag{9.13}$$

where E_0 is the ground state eigenvalue of the Hamiltonian $\hat{H}$. The equality holds if and only if $|\Psi_T\rangle = |\Psi_0\rangle$. The variational principle holds for the ground state, but also for excited states, provided that $|\Psi_T\rangle$ is orthogonal to all the eigenstates having eigenvalue lower than that of the state one wants to approximate.

In coordinate space the formulation of the variational principle can be directly transformed in a form equivalent to that of Eq. (9.5):

$$E_T = \frac{\int dR|\Psi_T(R)|^2 \frac{\hat{H}\Psi_T(R)}{\Psi_T(R)}}{\int dR|\Psi_T(R)|^2} \geq E_0, \tag{9.14}$$

where $\frac{\hat{H}\Psi_T(R)}{\Psi_T(R)}$ is called the *local energy*. Contrary to what happens in functional minimization approaches (such as the Hartree-Fock method), the variational principle is used to determine the best trial wavefunction within a class defined by some proper ansatz. The wavefunction will depend on a set of *variational parameters* $\{\alpha\}$. The solution of the variational problem will therefore be given by the solution of the Euler problem:

$$\frac{\delta E_T(\{\alpha\})}{\delta\{\alpha\}} = 0. \tag{9.15}$$

This means that in order to find the variational solution to the Schrödinger problem we need to evaluate many times the integral of Eq. (9.14) using different values of the variational parameters, and find the minimum trial eigenvalue.

9.2.3 Monte Carlo Evaluation of Integrals

The integral in Eq. (9.14) is in general defined in a $3A$-dimensional space. Since particles interact, we expect that the solution cannot be expressed as a product of single particle functions, and therefore the integral cannot be factorized in a product of simpler integrals. In this sense, the problem is strictly analogous to that of a classical gas at finite temperature $\beta = 1/K_B T$. In that case, given a classical Hamiltonian $H(p,q) = \sum_{i=1}^{A} \frac{p_i^2}{2m} + V(q_1 \dots q_A)$, the average energy of the system is given by:

$$E = \frac{3A}{2} K_B T + \frac{1}{Z} \int dq_1 \cdots dq_A V(q_1 \cdots q_A) e^{-\beta V(q_1 \cdots q_A)}, \tag{9.16}$$

where

$$Z \equiv \int dq_1 \cdots dq_A e^{-\beta V(q_1 \cdots q_A)} \tag{9.17}$$

is the *configurational partition function* of the system. Also in this case the integral to be evaluated is of the same form as Eq. (9.14). We can distinguish in the integrand the product of a *probability density*:

$$P(q_1 \dots q_A) = \frac{e^{-\beta V(q_1 \cdots q_A)}}{Z}, \tag{9.18}$$

and a function to be integrated which is the potential energy V. For classical systems we have a quite intuitive way of proceeding, which is at the basis of statistical mechanics. If we are able to compute (or measure) the potential for some given set of particle coordinates, and we average over many different configurations (sets of particle positions), we will obtain the estimate of the potential energy we need.

This fact can be easily formalized by making use of the Central Limit Theorem. Given a probability density $P[X]$ defined in a suitable event space X, let us consider an arbitrary function $F(X)$. One can define a stochastic variable:

$$S_N(F) = \frac{1}{N} \sum_{i=1}^{N} F(X_i), \tag{9.19}$$

where the events X_i are assumed to be *statistically independent*, and are distributed according to $P[X]$. The stochastic variable $S_N(F)$ will in turn have its own probability density $P[S_N]$, which in general depends on the index N. The Central Limit Theorem

states that for large N the probability density $P[S_N]$ will be a Gaussian, namely:

$$\lim_{N\to\infty} P[S_N] = \frac{1}{\sqrt{2\pi\sigma_N^2(F)}} \exp\left\{-\frac{(S_N - \langle F\rangle)^2}{2\sigma_N^2(F)}\right\}, \tag{9.20}$$

where we define the expectation of F as:

$$\begin{aligned} \langle F\rangle &= \int P[X]F(X)dX, \\ \langle F^2\rangle &= \int P[X]F^2(X)dX, \end{aligned} \tag{9.21}$$

and

$$\sigma_N^2(F) = \frac{1}{N}\left[\langle F^2\rangle - \langle F\rangle^2\right] \tag{9.22}$$

is the variance of the Gaussian. The reported average is estimated as $S_N(F)$, while $\langle F^2\rangle - \langle F\rangle^2$ is estimated by $\frac{N}{N-1}\left[S_N(F^2) - S_N^2(F)\right]$. This well known result is at the basis of all measurement theory. Averages over a set of measurements of a system provide the correct expectation of the measured quantity with an error that can be in turn estimated, and that decreases with the square root of the number of measurements N.

This result is very important from the point of view of numerical evaluation of integrals. If we had a way to numerically sample an arbitrary probability density $P[X]$, we could easily estimate integrals like that in Eq. (9.14). The statistical error associated with the estimate would decrease as the square root of the sampled points *regardless of the dimensionality of the system*.

For a classical system, configurations might be generated by solving Newton's equations, possibly adding a thermostat in order to be consistent with the canonical averaging. However, this is not certainly possible for a quantum system. The solution is to use an artificial dynamics, provided that it generates (at least in some limit) configurations that are distributed according to the probability density we want to use. Once again, in order to simplify the following description we will work in the space of the coordinates of the A particles, but the argument can be generalized to arbitrary spaces.

A very detailed description of what follows in this section can be found in the book of Kalos and Whitlock [1] and references therein.

We start defining a *transition matrix* $T_k(R_{k+1} \leftarrow R_k)$ expressing the probability that in the k-th step of the dynamics the system moves from the configuration R to a configuration R'. If at the first step the system is in a configuration R_0, sampled from an arbitrary distribution $P_0[R_0]$, the probability density of finding the system

in a configuration R_1 at the next step will be given by:

$$P_1[R_1] = \int dR_0 P_0[R_0] T_0(R_1 \leftarrow R_0). \tag{9.23}$$

We the introduce an integral operator $\hat{T}_0$ such that:

$$P_1[R_1] = \hat{T}_0 P[R_0]. \tag{9.24}$$

With this notation, the probability density of the configuration at an arbitrary step k will become:

$$P_k[R_k] = \hat{T}_{k-1} P[R_{k-1}] = \hat{T}_{k-1} \cdots \hat{T}_1 \hat{T}_0 P_0[R_0]. \tag{9.25}$$

The sequence of stochastic variables R_k generated at each step of this procedure is called a *Markov Chain*. Let us assume that $\hat{T}_k$ does not depend on the index k. What we will generate is then a *stationary* Markov Chain, for which the probability density generated at each step will only depend on the transition matrix and the probability density of the first element. In fact:

$$P_k[R_k] = \hat{T} P[R_{k-1}] = \hat{T} \cdots \hat{T}\hat{T} P_0[R_0] = \hat{T}^k P_0[R_0]. \tag{9.26}$$

Under these assumptions one might wonder if the sequence is convergent (in functional sense), i.e. if a limiting probability density $P_\infty[R]$ exists. It is interesting to notice that if such function exists, it has to be an eigenvector of the integral operator $\hat{T}$. In fact, since we assume $\hat{T}$ to be independent of k we have:

$$\lim_{k\to\infty} \hat{T} P_k[R_k] = \lim_{k\to\infty} P_{k+1}[R_{k+1}]$$

$$\hat{T} P_\infty[R] = P_\infty[R].$$

It is also easy to realize that the eigenvalue is indeed 1. In fact, let us consider the general relation:

$$\hat{T} P_\infty[R] = \gamma P_\infty[R]. \tag{9.27}$$

The recursive application of $\hat{T}$ would give:

$$\hat{T}^k P_\infty[R] = \gamma^k P_\infty[R]. \tag{9.28}$$

If $\gamma \neq 1$ we would lose the normalization property of $P\infty[R]$.

These properties of stationary Markov chains can be exploited to sample a generic probability density $P[R]$. In fact, if we can determine the transition operator that has as eigenvector a *given* $P_\infty[R]$, a repeated application of such operator to an *arbitrary* initial distribution of points will eventually generate a chain in which

each element is distributed according to $P_\infty[R]$. There is a simple recipe to construct such transition operator. We will assume that we have at hand a transition operator $\hat{\bar{T}}$ that we can sample (it could be as simple as a uniform probability within a given volume). We will split the searched transition operator in the product of $\hat{\bar{T}}$ and an unknown factor $\hat{A}$ that we will call "acceptance probability", defined in such a way that:

$$\hat{\bar{T}}\hat{A} = \hat{T}. \tag{9.29}$$

In order for the system to preserve its equilibrium state once the probability distribution is reached, we expect that the dynamics described by the random walk will not change the density of sampled points anywhere in the events space. Transitions carrying away from a state R to anywhere must be balanced by transitions leading from anywhere to the same state R:

$$\int dR' P(R) T(R' \leftarrow R) = \int dR' P(R') T(R \leftarrow R') \,. \tag{9.30}$$

One way to enforce this condition is to impose the more stringent *detailed balance* condition, which requires the *integrands* in Eq. (9.30) be equal:

$$P(R) T(R' \leftarrow R) = P(R') T(R \leftarrow R') \,. \tag{9.31}$$

The detailed balance condition can be in turn recast into a requirement on the acceptance probability. In fact:

$$\frac{A(R' \leftarrow R)}{A(R \leftarrow R')} = \frac{P(R')}{P(R)} \frac{\bar{T}(R \leftarrow R')}{\bar{T}(R' \leftarrow R)}. \tag{9.32}$$

The quantities on the r.h.s. of Eq. (9.32) are all known. The configuration R' has to be sampled originating in R from the given transition probability $\bar{T}(R \leftarrow R')$. The probability density $P(R)$ is the one we actually want to asymptotically sample. If we interpret the A values to be probabilities to actually keep the transition, then maximizing the possible A values leads to the slightly modified version of Eq. (9.32):

$$A(R' \leftarrow R) = \min\left(\frac{P(R')}{P(R)} \frac{\bar{T}(R \leftarrow R')}{\bar{T}(R' \leftarrow R)}, 1 \right). \tag{9.33}$$

This expression is often called the *acceptance ratio*. In practice, it represents the probability according to which we have to *accept* the new configuration as the new member of the Markov chain, rather than keeping the original point as the next

point in the chain.[2] Further analysis shows that existence and uniqueness of the correct eigenvalue 1 solution and therefore convergence to the correct distribution will be guaranteed if (1) every allowed state can be reached from any other by a finite sequence of transitions and (2) there are no cycle of states. The latter is guaranteed if there are any transitions that leave the system in the same state, that is any rejections.

There is a case in which Eq. (9.32) further simplifies. If the transition matrix is taken to be symmetric in the arguments R and R', the ratio becomes unity, and one is left with:

$$A(R' \leftarrow R) = \min\left(\frac{P(R')}{P(R)}, 1\right). \tag{9.34}$$

At this point we have all the ingredients to describe an algorithm that performs a Monte Carlo evaluation of an integral such that of Eq. (9.14). In the following we will describe the simplest version, i.e. the so called "Metropolis-Hastings algorithm" [2, 3].

1. Start from an arbitrary configuration of the A particles. If the potential has a strongly repulsive core one has to pay attention to avoid overlapping pairs.
2. Sweep over the coordinates and generate new positions according to some transition probability. A simple choice is a uniform displacement within a cube of side Δ, i.e.:

$$\bar{T}(R' \leftarrow R) = \begin{cases} \frac{1}{\Delta} \text{ if } & |R_i'^{\alpha} - R_i^{\alpha}| < \frac{\Delta}{2} \\ \\ 0 \text{ otherwise} \end{cases} \tag{9.35}$$

with $\alpha = x, y, z$, and $i = 1 \ldots A$. This choice has the advantage of being symmetric. If we imagine to store our configuration in an array $R[0\ldots2][0\ldots A-1]$ the implementation of this step would read:

```
MC_Move()
for i ∈ {0, A − 1} do
    for j ∈ {0, 2} do
        R_new[i][j] ← R[i][j] + (rand() − 0.5) ∗ Δ
    end for
end for
```

[2]The standard jargon refers to this as a "rejection" event. However one has not to be confused: this is the result of a *reversed* move, and generates a new element in the chain coincident with the starting point.

We will assume that the function rand() generates a random number uniformly distributed in $[0, 1)$.

3. At this point we need to evaluate the acceptance ratio. This is easily done with our choice of the transition matrix, since we only need to evaluate the probability densities in R and R':

$$A(R' \leftarrow R) = \min\left(\frac{|\Psi_T(R')|^2}{|\Psi_T(R)|^2}, 1\right). \tag{9.36}$$

4. Next we need to decide whether we keep the proposed configuration as the next element in the chain or if we want to resort to the original one. If we define $\text{acc} = A(R' \leftarrow R)$, then:

```
Accept_reject()
ξ = rand()
if acc > ξ then
    R[i][j] ← R_new[i][j]
end if
```

5. According to the Central Limit Theorem, we now need to cumulate the values of the rest of the integrand. In the case of our variational calculation we need to sum up the local energies. Notice that this step has to be taken whatever the result of the procedure described at the previous point. If we want to estimate the statistical error, we also need to cumulate the *square* of the local energy.

```
Acuest()
eloc ← Ĥ Ψ_T(R) / Ψ_T(R)
ecum ← ecum + eloc
ecum2 ← ecum2 + eloc * eloc
```

6. Steps 2–5 need to be repeated N_{steps} times, where N_{steps} must be sufficiently large to provide a small enough statistical error. The final estimate of the energy is given by $\langle E\rangle \pm \Delta E$, where:

$$\begin{aligned}\langle E\rangle &= \frac{1}{N_{steps}} \cdot \text{ecum} \\ \Delta E &= \sqrt{\frac{1}{N_{steps}-1}\left(\frac{1}{N_{steps}} \cdot \text{ecum2} - \langle E\rangle^2\right)}\end{aligned} \tag{9.37}$$

Notice that this algorithm could in principle be used to evaluate arbitrary integrals. In fact, it is always possible to multiply and divide the integrand by a probability density $P(X)$ that can be used to sample the values of X:

$$I = \int F(X)dX = \int P[X]\frac{F(X)}{P[X]} \tag{9.38}$$

9.2.3.1 Autocorrelations

The main hypothesis underlying the Central Limit Theorem is that data used to construct the averages are sampled independently. While in a measurement process this is a quite reasonable assumption, in the case of the computation of an integral by means of any method based on the Markov chain theory (including the Metropolis-Hastings method) this requirement is not satisfied by construction. In fact, data are sampled based on a transition matrix, and the resulting random walk has a certain degree of memory of the past events. What are the consequences of such memory? Let us consider a sequence of points $X_1, X_2, \cdots, X_N$ sampled via the Metropolis algorithm from some probability density $P[X]$. If we assume these data not to be independent, we have to consider the joint probability for the specific realization of the chain in order to estimate the integral of a given function F:

$$I = \frac{1}{N}\sum_{i=1}^{N}\int dX_1, dX_2, \cdots dX_N P[X_1, X_2, \cdots X_N]F(X_i). \tag{9.39}$$

If the samples are independent then $P[X_1, X_2, \cdots X_N] = P[X_1]P[X_2]\cdots P[X_N]$, and we are in the case previously discussed. However, since we can arbitrarily exchange the indexes of the integration variables, we can easily see that the value of I is unchanged despite the presence of correlations. By construction, in a Markov process two consecutive samples will always be correlated to each other. This seems to be inconsistent to the use we want to make of these samples, i.e. to apply the Central Limit Theorem to integration. However, we can hope that after a certain number of steps memory is lost, and data will become effectively independent. Is it possible to estimate this typical *autocorrelation length*? Based on the previous argument one can define a measure of the autocorrelation by looking at the variance of the expectation of F with respect to P:

$$(\Delta I)^2 = \left\langle \frac{1}{N^2}\sum_{i=1}^{N}F(X_i)\sum_{i=1}^{N}F(X_j)\right\rangle - \langle F\rangle^2. \tag{9.40}$$

The corresponding standard deviation is the estimate of the statistical error on the integral of F. The first term can be recast in the following way:

$$\begin{aligned}\left\langle \frac{1}{N^2}\sum_{i=1}^{N}F(X_i)\sum_{i=1}^{N}F(X_j)\right\rangle \\ = \frac{1}{N^2}\sum_{i,j=1}^{N}\int P[X_1,X_2,\cdots X_N]F(X_i)F(X_j)dX_1\cdots dX_N \\ = \frac{1}{N^2}\sum_{i,j=1}^{N}\langle F(X_i)F(X_j)\rangle. \end{aligned} \tag{9.41}$$

Since the Markov chain is stationary, this quantity is expected to depend only on the difference of the indexes $\tau = i - j$. We will then define an *autocorrelation coefficient*:

$$c(F)_\tau = \frac{\langle F(X_i)F(X_{i+\tau})\rangle - \langle F\rangle^2}{\langle F^2\rangle - \langle F\rangle^2}. \tag{9.42}$$

The coefficient is normalized to the variance $\sigma^2(F)$, in such a way that $C(F)_0 = 1$. Correlation coefficients are related to the average of the product of the F in the following way:

$$\langle F(X_i)F(X_{i+\tau})\rangle = c(F)_\tau\sigma^2(F) + \langle F\rangle^2. \tag{9.43}$$

We can use the previous expression to estimate the error on I:

$$\begin{aligned}(\Delta I)^2 = \frac{1}{N^2}\sum_{i,j=1}^{N}\langle F(X_i)F(X_j)\rangle - \langle F\rangle^2 \\ = \frac{1}{N}\sigma^2(F)\sum_{\tau=1}^{N}c(F)_\tau + \langle F\rangle^2 - \langle F\rangle^2 = \frac{\sigma^2(F)}{N}\sum_{\tau=1}^{N}c(F)_\tau. \end{aligned} \tag{9.44}$$

As it can be seen the error not only depends on the variance of F, but also on the sum over all the autocorrelation coefficients of F. This is the main consequence of having autocorrelated samples: the statistical error is underestimated by the variance of F, and needs to be corrected by a factor that depends on the autocorrelation length.

Usually the coefficients $c(F)_\tau$ have an exponential decay. If we approximate them as $c(F)_\tau \sim \exp(-\tau/\bar{\tau})$, the sum of the coefficients can be approximated as:

$$\sum_{\tau=1}^{N}c(F)_\tau \sim \int_0^\infty d\tau e^{-\frac{\tau}{\bar{\tau}}} = \bar{\tau}. \tag{9.45}$$

This means that it is sufficient to fit the exponential decay of the autocorrelation coefficients in order to find an estimate of the characteristic autocorrelation length that corrects the estimate of the error on the integral. In particular the correct expression for the error is:

$$\Delta I \simeq \sqrt{\frac{1}{N-1}\sigma^2(F)\bar{\tau}}, \tag{9.46}$$

which has a simple interpretation: We are not generating N independent samples of the variable X during our Markov process, but rather $N/\bar{\tau}$ of them, and this number must be used as the correct count of events for the error estimation.

It is important to be extremely careful about the estimation of autocorrelations. in many cases an underestimation of the statistical errors leads to a wrong interpretation of the results and to wrong physical conclusions.

Autocorrelations also play a crucial role in choosing the step width Δ in the Metropolis-Hastings algorithm. A common criterion is to choose it in such a way that the fraction of accepted moves is about 50%. However, the ideal value is clearly the one minimizing the autocorrelations among samples, and quite often this value corresponds to acceptances of the order 30 or 40%.

Once the value of $\bar{\tau}$ has been estimated, it is possible to organize the calculation in such a way that the statistical error computed by the code is more realistic by using a *reblocking* technique. In practice the values of the quantity to be averaged are summed up in blocks of N_b elements each:

$$F_l^b = \sum_{i=1}^{N_b} F(X_i). \tag{9.47}$$

Then, the F_l^b are used as the data on which performing the computation of the variance and of the standard deviation. If $N_b \gtrsim \bar{\tau}$, the standard deviation will be corrected by the effects of the autocorrelation of the original data. Typically calculations store block values so that the values can be "reblocked" for example by combining pairs of blocks. The estimated error should be unchanged if the blocks are uncorrelated. In addition, the ratio of the block variance to the variance of the original function can be used to estimate the number of independent samples, and therefore the autocorrelation time.

9.2.4 Construction of the Wavefunction and Computational Procedures

When performing a variational calculation, the first step consists of deciding which model wavefunction we intend to use.

First of all we have to take care of the symmetry of the particles. Nucleons are Fermions, and therefore it is necessary to build an antisymmetric wavefunction. If the Hamiltonian does not contain terms acting on the spin or isospin state of a nucleon or of a pair of nucleons, each particle will preserve its own initial state. In this case it is easy to write an antisymmetric wavefunction simply using a product of Slater determinants, one for each species.

To build the determinants one needs some single particle orbitals. There are several possible choices. For nuclei linear combinations of Gaussians or the eigenstates of the harmonic oscillator are definitely an option. Another choice might be that of using orbitals coming from a Hartree-Fock calculation. In this case the orbitals contain some information about the fact that nucleons interact, but there usually is no consistency between the Hamiltonian used to compute the orbitals and the Hamiltonian we are interested in.

The basic starting point is then a wavefunction of the form:

$$\varphi(R) = \det[\phi_j(\mathbf{r}_{p_i^\uparrow})]\det[\phi_j(\mathbf{r}_{p_i^\downarrow})]\det[\phi_j(\mathbf{r}_{n_i^\uparrow})]\det[\phi_j(\mathbf{r}_{n_i^\downarrow})], \tag{9.48}$$

If we just limited ourselves to this kind of wavefunction we would miss most of the interesting physics that happens when particles are close together. A seen in the previous chapters, a very important role is played by the *short range correlations*, that should introduce the many-body effects due to repulsion/attraction of particles at short distance. Contrarily to what one does in other methods, such as coupled clusters, in Quantum Monte Carlo calculations it is easier to work with wavefunctions containing *explicit* two-, three- or many-body correlations.

Here we will use the so-called *Jastrow* factor, i.e. a product of two-body functions that helps to reproduce the correlations from the pair-wise potential. The simplest version of a trial wavefunction therefore reads:

$$\Psi_T(R) = \varphi(R)\prod_{i<j}^{A} f(r_{ij}), \tag{9.49}$$

where $R = (\mathbf{r_1}, \cdots, \mathbf{r_A})$, and f is the so called *Jastrow function* (JF). How do we determine the JF? We have some information that we can exploit. In particular we might seek for analytic forms of f that satisfy what is commonly called the *cusp condition*,(see e.g. [4]) i.e. we must have:

$$\frac{\hat{H}f(r_{ij})}{f(r_{ij})} < \infty \tag{9.50}$$

everywhere in space. It is easy to realize that satisfying the cusp condition helps to prevent the local energy from fluctuating too much even in presence of a divergence of the potential, thereby reducing the variance and the statistical error. Usually in nuclear physics problems it is customary to take a further step. Recognizing that at small separations, the many-body Schrödinger equation is dominated by the short-

range pair potential, the two-body problem is solved to determine the f. In particular one can solve the following Schrödinger equation in relative coordinates:

$$-\frac{\hbar^2}{2m}\nabla^2 + qV(r)f(r) = \epsilon f(r), \tag{9.51}$$

and impose the boundary condition that the function becomes a constant at a distance h from the origin, where other parts of the Hamiltonian become important. The quantities q and h are two variational parameters. One could in principle consider a third variational parameter in the Jastrow factor by using a modified Jastrow function $\tilde{f}$ such that:

$$\tilde{f}(r) = e^{-b\log f(r_{ij})} \tag{9.52}$$

The function f is usually determined by numerically solving Eq. (9.51) with the Numerov or Runge-Kutta methods. One has to be careful that the resulting table has to be interpolated to compute the function at an arbitrary distance. Therefore it is important to choose an appropriate number of points (usually of the order of a few thousands). Single particle orbitals can also be either tabulated or computed analytically. Tabulation guarantees in general a faster computation at the price of a loss in numerical accuracy.

In the code it is necessary to compute derivatives of the wavefunction in order to estimate the local energy. This can be done either numerically or analytically. A very good test for checking that there are no major mistakes either in the Monte Carlo evaluation of integrals or in the computation of the local energy is to use the so-called *Jackson-Feenberg* identity for the kinetic energy.

The expectation of the kinetic energy is an integral of the form:

$$\langle T\rangle = \frac{-\frac{\hbar^2}{2m}\int_\Omega dR\Psi^*(R)\nabla^2\Psi(R)}{\int_\Omega dR|\Psi(R)|^2}, \tag{9.53}$$

where Ω is the integration volume. Integrating the numerator by parts one gets:

$$\langle T\rangle = \frac{\frac{\hbar^2}{2m}\int_\Omega dR\nabla\Psi^*(R)\cdot\nabla\Psi(R)}{\int_\Omega dR|\Psi(R)|^2} - \frac{\frac{\hbar^2}{2m}\int_{S(\Omega)} dS\Psi^*(R)\nabla\Psi(R)}{\int_\Omega dR|\Psi(R)|^2}. \tag{9.54}$$

The surface term is zero if the wavefunction is well behaved. We are therefore left with the integral:

$$\langle T \rangle = \frac{\frac{\hbar^2}{2m}\int_\Omega dR \nabla\Psi^*(R)\nabla\Psi(R)}{\int_\Omega dR|\Psi(R)|^2} = \frac{\frac{\hbar^2}{2m}\int_\Omega dR|\Psi(R)|^2 \frac{\nabla\Psi(R)}{\Psi(R)}\cdot\frac{\nabla\Psi^*(R)}{\Psi^*(R)}}{\int_\Omega dR|\Psi(R)|^2}. \tag{9.55}$$

We can sum Eqs. (9.53) and (9.55), and divide by 2 in order to obtain a new kinetic energy estimator:

$$\langle T \rangle_{JF} = \frac{\frac{\hbar^2}{4m}\int_\Omega dR|\Psi(R)|^2 \left[\frac{\nabla\Psi^*(R)}{\Psi^*(R)}\cdot\frac{\nabla\Psi(R)}{\Psi(R)} - \frac{\nabla^2\Psi(R)}{\Psi(R)}\right]}{\int_\Omega dR|\Psi(R)|^2}. \tag{9.56}$$

This is the *Jackson-Feenberg* kinetic energy estimator. It is easy to see that configuration by configuration the value of the integrand of T and T_{JF} are different. However, they have to be the same on average (i.e. always within the current statistical error). The equivalence of the two estimators checks the integration procedure, the correctness of the implementation of the boundary conditions, and the computation of derivatives. If any of these quantities are wrong, the two estimates of the kinetic energy will differ. This is an extremely useful consistency check, and should always be used in a variational calculation.

At this point it is necessary to perform several calculations varying the parameters in the wavefunction, and looking for a minimum of the energy. In the next subsection we will describe algorithms that allow for performing this search in an automatic way. However, when the number of parameters is small, it is also possible in principle to perform a scan on a grid.

In Fig. 9.1, as an example, we report the behavior of the variational energy computed in a ^{4}He nucleus, modeled with a two body Minnesota potential, and a wavefunction containing only a central Jastrow product. The spatial part of the orbital is an s-wave Gaussian with half width equal to 1.1 fm. The energies have been computed for a fixed value of the healing distance $h = 3.1$ fm as a function of the quencher parameter q, keeping fixed the amplitude parameter $b = 1$. Each run consists of an average over 6.4×10^5 samples, preceded by 6.4×10^4 equilibration steps.

As it can be seen, there is a clear minimum of the energy. The minimum can be determined with sufficient accuracy by fitting the resulting curve. A fit with a quadratic function predicts a minimum at $q \sim 1.4$. The corresponding eigenvalue is $E_T = -15.31(4)$ MeV.[3] The procedure should be repeated for different values of all other variational parameters until an absolute minimum is found.

[3]The number in parenthesis indicates the statistical error on the last figure.

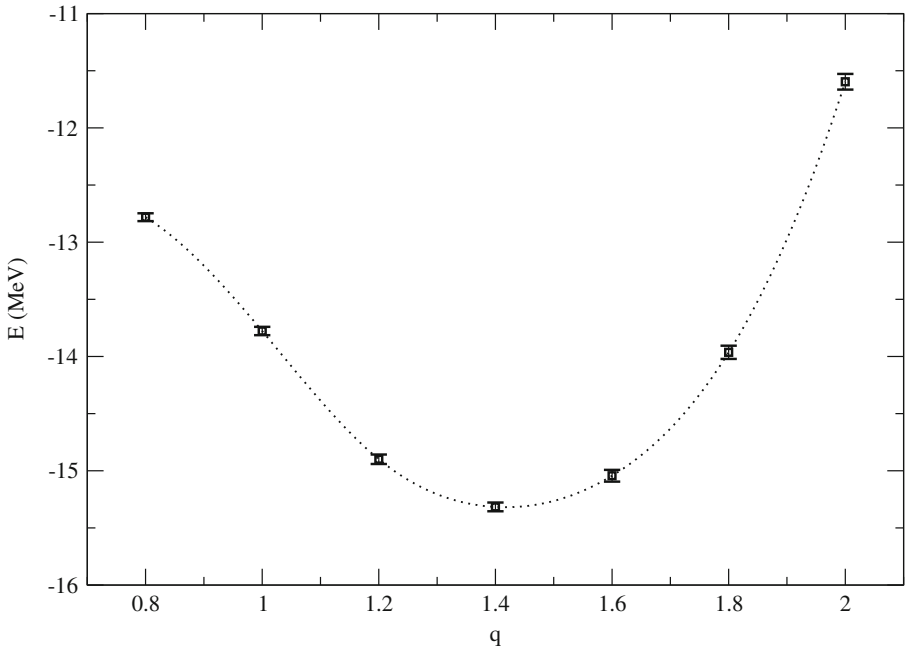

Fig. 9.1 An example of variational minimization of the energy. The estimate of the binding energy of a ^{4}He nucleus described by the Minnesota potential is here plotted as a function of the quencher parameter q, for a fixed value of the healing distance h (see text). The *dotted line* serves as a guide for the eye

The variational wavefunctions can be made arbitrarily richer in structure in order to improve the results, including what our physical intuition suggests as important terms to describe correlations. We will later discuss how to construct trial wavefunctions for realistic nuclear Hamiltonians. A full variational calculation for the ^{4}He nucleus with the Minnesota potential, including Jastrow factors with a spin/isospin dependence would give a binding energy $E_T = -25.52(4)$ MeV. As an example, the optimized Jastrow function for the central channel of the Minnesota potential is shown in Fig. 9.2.

9.2.5 Wave Function Optimization

9.2.5.1 Reweighting Methods

The brute-force optimization of the trial wave function becomes quite cumbersome with more than a few parameters. In general the problem is equivalent to searching an absolute minimum in a multi-dimensional space, and does not admit a simple solution. If one is interested in a quick search for local minima, it is possible to compute the gradient of the energy in the parameter space, and use for instance some variant of the steepest descent method. Computation of gradients is based on the so-called "reweighting method". If we have a trial function depending on a set of parameters $\{\alpha\}$, and another depending on a set $\{\alpha + \delta\alpha\}$, it s not necessary to

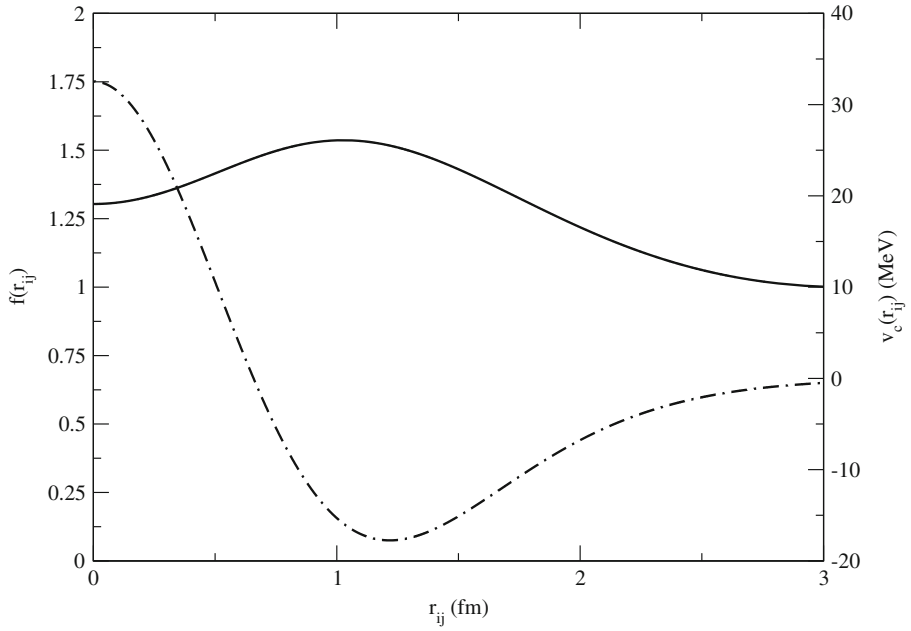

Fig. 9.2 The central channel of the Minnesota potential (*dashed-dotted line*), and the corresponding numerical Jastrow function (*solid line*) evaluated for $h = 3.1$, $q = 1.4$, and $b = 1$

perform two independent calculations to compute the difference (which would also be affected by rather large statistical errors). In fact, the following identity holds):

$$\frac{\int dR|\Psi_T(R,\{\alpha+\delta\alpha\})|^2 O(R)}{\int dR|\Psi_T(R,\{\alpha+\delta\alpha\})|^2} = \frac{\int dR|\Psi_T(R,\{\alpha\})|^2 \frac{|\Psi_T(R,\{\alpha+\delta\alpha\}|^2}{|\Psi_T(R,\{\alpha\})|^2} O(R)}{\int dR|\Psi_T(R,\{\alpha\})|^2 \frac{|\Psi_T(R,\{\alpha+\delta\alpha\})|^2}{|\Psi_T(R,\{\alpha+\delta\alpha\})|^2}} \tag{9.57}$$

It is therefore possible to use the configurations sampled from a trial wavefunction with a given parametrization $\{\alpha\}$ to compute expectations over a wavefunction with a different parametrization $\{\alpha+\delta\alpha\}$ by simply reweighting the values of the operator with the ration between the square moduli of the two wavefunctions:

$$\begin{aligned}\langle O_{\{\alpha+\delta\alpha\}}\rangle &\equiv \frac{\langle\Psi_T(R_k,\{\alpha+\delta\alpha\})|O(R_k)|\Psi_T(R_k,\{\alpha+\delta\alpha\})\rangle}{\langle\Psi_T(R_k,\{\alpha+\delta\alpha\})|\Psi_T(R_k,\{\alpha+\delta\alpha\})\rangle} \\ &= \frac{\sum_k \frac{|\Psi_T(R_k,\{\alpha+\delta\alpha\}|^2}{|\Psi_T(R_k,\{\alpha\}|^2} O(R_k)}{\sum_k \frac{|\Psi_T(R_k,\{\alpha+\delta\alpha\}|^2}{|\Psi_T(R_k,\{\alpha\}|^2}}\end{aligned} \tag{9.58}$$

where the R_k are sampled from $|\Psi_T(R,\{\alpha\})|^2$. Besides the obvious advantage of avoiding multiple calculations to compute the derivatives, the use of this reweighting technique allows direct computation of expectations of the gradients in the parameter space with very high accuracy. The access to gradients opens the way to the use of automated minimization algorithms such as the already mentioned

steepest descent method, the Levemberg-Marquardt algorithm [5] or the Linear Method [6] briefly sketched below.

9.2.5.2 Power Method

There is another class of algorithms that have been recently introduced, and based on the power method. We will here discuss in particular the algorithm due to Sandro Sorella [7]. This algorithm was originally discussed in terms of the Lanczos method, but for a single multiplication by his propagator it becomes equivalent to the simpler power method that we discuss here.

For Λ larger than the largest eigenvalue of the eigenvectors contained in $|\psi_n\rangle$, operating with $\Lambda - H$ will multiply the ground state by a larger number than any other state. Therefore iterating the equation

$$|\psi_{n+1}\rangle = (\Lambda - H)|\psi_n\rangle \tag{9.59}$$

will converge to the ground state. One way to implement this is to use a set of test functions (which, in principle, should be complete), $|\phi_m\rangle$. This gives the set of equations

$$\langle\phi_m|\psi_{n+1}\rangle = \langle\phi_m|(\Lambda - H)|\psi_n\rangle\,. \tag{9.60}$$

In his original paper Sorella assumes $|\psi_n\rangle = |\Psi_T\rangle$, and next approximates $|\psi_{n+1}\rangle$ as a linear combination of the original state and the derivatives with respect to the parameters

$$|\psi_{n+1}\rangle \simeq \Delta\alpha_0|\Psi_T\rangle + \sum_{n=1} \Delta\alpha_n \partial_{\alpha_n}|\psi_T\rangle \equiv \sum_{n=0} O^n|\psi_T\rangle\Delta\alpha_n \tag{9.61}$$

and he uses the same functions for $|\phi_m\rangle$, so that

$$|\phi_m\rangle = O^m|\Psi_T\rangle\,. \tag{9.62}$$

When evaluated in the position representation, the O^m for $m > 0$ correspond to multiplying by the derivative of the logarithm of the trial function. Substituting these expressions, and dividing by $\langle\Psi_T|\Psi_T\rangle$, Eq. (9.60) becomes

$$\frac{\langle\Psi_T|O^m(\Lambda - H)|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle} = \sum_{n=0} \frac{\langle\Psi_T|O^m O^n|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle}\Delta\alpha_n\,. \tag{9.63}$$

The expectation values can be calculated and the linear equations solved to get $\Delta\alpha_n$.

Alternatively, the $m = 0$ and $n = 0$ terms can be separated. Writing the trial function expectation of an operator O as $\langle O \rangle$, Eq. (9.63) becomes

$$\langle \Lambda - H \rangle = \Delta\alpha_0 + \sum_{n=1} \langle O^n \rangle \Delta\alpha_n \;\; m = 0 \tag{9.64}$$

$$\langle O^m (\Lambda - H) \rangle = \langle O^m \rangle \Delta\alpha_0 + \sum_{n=1} \langle O^m O^n \rangle \Delta\alpha_n \; m > 0 \,. \tag{9.65}$$

Substituting Eq. (9.64) into Eq. (9.65) gives

$$\langle O^m (\Lambda - H) \rangle - \langle \Lambda - H \rangle \langle O^m \rangle = \sum_{n=1} \left[\langle O^m O^n \rangle - \langle O^m \rangle \langle O^n \rangle \right] \Delta\alpha_n \,. \tag{9.66}$$

Solving gives $\Delta\alpha_{n>0}$ and Eq. (9.64) then gives the value for $\Delta\alpha_0$.

In either case, the result gives an approximation to the next trial function as a linear combination of the original function and its parameter derivatives. The new parameters are chosen to give this same linear combination as the first two terms in the Taylor series. Since dividing the approximate expression for $|\psi_{n+1}\rangle$ by $\Delta\alpha_0$ gives an expression that is the first two terms in the Taylor series, the new parameters are

$$\alpha_{n>0}^{(\text{new})} = \alpha_n^{(\text{old})} + \frac{\Delta\alpha_{n>0}}{\Delta\alpha_0} \tag{9.67}$$

More recently Toulouse and Umrigar [6] proposed a much more efficient method where the Hamiltonian is diagonalized in the reduced space spanned by the $|\phi_m\rangle$. The parameter variation is then given by the solution of the generalized eigenvalue equation

$$\sum_{n=0} \frac{\langle \Psi_T | O^m H O^n | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \Delta\alpha_n = \Delta E \sum_{n=0} \frac{\langle \Psi_T | O^m O^n | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} \Delta\alpha_n \,. \tag{9.68}$$

with the lowest eigenvalue $\Delta E^m in$:

$$\alpha_{n>0}^{(\text{new})} = \alpha_n^{(\text{old})} + \frac{\Delta\alpha_{n>0}^{min}}{\Delta\alpha_0^{min}} . \tag{9.69}$$

The gradient of the local energy is required for the expectation values appearing in Eq. (9.68), and can be efficiently estimated using the reweighting technique presented in the previous section.

When the parameters are far away from the minimum this approach can be less stable than the previous one giving rise to large parameter variations that invalidate the linear approximation Eq. (9.61). A quick strategy is then to use the solution of Eq. (9.66) early on in the optimization process and then switch to Eq. (9.68) when the resulting norm of the variation is below some threshold.

9.3 Projection Monte Carlo Methods in Coordinate Space

9.3.1 General Formulation

Variational calculations provide only an upper bound for the ground-state eigenvalue of a given Hamiltonian. However, it is possible to use Monte Carlo algorithms to actually solve the Schrödinger equation for an arbitrary number of interacting particles. This class of algorithms is based on the idea of imaginary time propagation.

Let us consider a Hamiltonian $\hat{H}$. The imaginary time evolution of an arbitrary state is defined starting by the standard time-dependent Schrödinger equation:

$$-i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = \hat{H}|\Psi(t)\rangle. \tag{9.70}$$

It is possible to Wick rotate, and introduce an *imaginary time* $\tau = \frac{it}{\hbar}$. The time-dependent Schrödinger equation is transformed into an imaginary-time-dependent equation:

$$-\frac{\partial}{\partial \tau}|\Psi(\tau)\rangle = \hat{H}|\Psi(\tau)\rangle, \tag{9.71}$$

where τ is defined as an *inverse energy* that parametrizes the propagation of the quantum state. The formal solution can be written using the imaginary time propagator

$$|\Psi(\tau)\rangle = e^{-\tau\hat{H}}|\Psi(0)\rangle \tag{9.72}$$

It is possible to expand the initial state $|\Psi(0)\rangle$ in eigenstates $|\phi_n\rangle$ of the Hamiltonian itself, such that $\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$. The imaginary time propagation of $|\Psi(0)\rangle = \sum_n c_n|\phi_n\rangle$ becomes:

$$|\Psi(\tau)\rangle = e^{-\tau\hat{H}}\sum_n c_n|\phi_n\rangle = \sum_n c_n e^{-\tau E_n}|\phi_n\rangle \tag{9.73}$$

Let us now consider the limit of the propagation for $\tau \to \infty$. The coefficients of the expansion $c_n e^{-\tau E_n}$ will either decrease (if $E_n > 0$) or increase (if $E_n < 0$) with the imaginary time, but in the limit the coefficient corresponding to the ground state of $\hat{H}$, i.e. $c_0 e^{-\tau E_0}$ will be dominant. This means that the imaginary time propagator has the interesting property of filtering out of an arbitrary state in the Hilbert space the ground state of a given Hamiltonian, provided that the state is not orthogonal to the ground state to begin with. We want to stress a very important point. The ground state we are referring to is the *mathematical* ground state of the Hamiltonian $\hat{H}$. The *physical* ground state needs to take into account the symmetry of the particles, either bosons or fermions. It is very easy to convince oneself that such mathematical

ground state is always a nodeless function (i.e. it is zero nowhere but possibly on the boundaries of the domain of existence of the wavefunction expressed in some representation). This is because the propagator is a positive definite function, at least for a Hamiltonian of the standard form $\hat{H} = \hat{T} + \hat{V}$, where $\hat{T}$ is the kinetic energy of a system of free particles and $\hat{V}$ is a local potential. In this case the eigenvector corresponding to the largest eigenvalue of the propagator is positive definite within the domain that defines the system. The largest eigenvalue of the propagator corresponds to the lowest eigenvalue of $\hat{H}$.

Notice that the imaginary time propagator is hermitian not unitary, and the normalization of the projected ground state is not guaranteed in general. By means of a small change in the propagator definition it is possible to guarantee the normalization of the projected ground state. In fact, let us define the propagator as:

$$|\Psi(\tau)\rangle = e^{-\tau(H-E_0)}|\Psi(0)\rangle. \tag{9.74}$$

It is easy to realize that in this case the amplitude of the component of the initial state along the ground state is preserved (while all other amplitudes decrease exponentially), and therefore the projected state is normalizable.

We will later discuss in detail the implications of these properties as concerns the application of imaginary-time propagation to many-fermion systems.

9.3.2 Imaginary Time Propagator in Coordinate Representation

We will focus on a practical implementation of imaginary time propagation, and we will limit ourselves to a system of bosons (or Boltzmannions) which do admit a ground-state wavefunction that is positive definite. We will also consider Hamiltonians of the form mentioned in the previous subsection, in which the interaction is local. In this case the propagator is easily represented in coordinates. Formally we would have:

$$\langle R|\Psi(\tau)\rangle = \int dR' \langle R|e^{-\tau(H-E_0)}|R'\rangle\langle R'|\Psi(0)\rangle, \tag{9.75}$$

where we have inserted a complete set of position eigenstates. The propagator

$$\langle R|e^{-\tau(\hat{H}-E_0)}|R'\rangle \tag{9.76}$$

seems to be still quite difficult to evaluate. However, let us break up the imaginary time interval τ in two equal intervals $\tau/2$. We can write

$$\langle R|e^{-\tau(\hat{H}-E_0)}|R'\rangle = \langle R|e^{-\frac{\tau}{2}(\hat{H}-E_0)}e^{-\frac{\tau}{2}(\hat{H}-E_0)}|R'\rangle, \tag{9.77}$$

since $\hat{H}$ obviously commutes with itself. Inserting a complete set we obtain:

$$\langle R|e^{-\tau(\hat{H}-E_0)}|R'\rangle = \int dR''\langle R|e^{-\frac{\tau}{2}(\hat{H}-E_0)}|R''\rangle\langle R''|e^{-\frac{\tau}{2}(\hat{H}-E_0)}|R'\rangle, \tag{9.78}$$

This process can be iterated for an arbitrary large number of times M:

$$\langle R|e^{-\tau(\hat{H}-E_0)}|R'\rangle = \int\cdots\int dR''\cdots dR^M\langle R|e^{-\frac{\tau}{2}(\hat{H}-E_0)}|R''\rangle\cdots\langle R^M|e^{-\frac{\tau}{2}(\hat{H}-E_0)}|R'\rangle. \tag{9.79}$$

Each of the factors in the integrand corresponds to a propagation for a *short* imaginary time $\Delta\tau = \tau/M$. In this case we can split the propagator using the Trotter-Suzuki formula:

$$e^{-\frac{\Delta\tau}{2}(\hat{H}-E_0)} \sim e^{-\frac{\Delta\tau}{2}(\hat{V}-E_0)}e^{-\Delta\tau\hat{T}}e^{-\frac{\Delta\tau}{2}(\hat{V}-E_0)} + o(\Delta\tau^3) \tag{9.80}$$

The representation in coordinates of each factor is known. The factors containing the potential, under the hypotheses made, are diagonal in the coordinates themselves, and simply become:

$$e^{-\frac{\Delta\tau}{2}(\hat{V}-E_0)}|R\rangle = |R\rangle e^{-\frac{\Delta\tau}{2}(V(R)-E_0)}, \tag{9.81}$$

while the kinetic term is the propagator of a set of A free particles obeying the equation:

$$-\frac{\partial}{\partial\tau}\Psi(R,t) = -\frac{\hbar^2}{2m}\nabla^2\Psi(R,t) \tag{9.82}$$

This is a classical *free diffusion* equation. If we interpret $\Psi(R,t)$ as a the density of the A particles, its evolution in time will be given by the well known diffusion law:

$$\Psi(R,t) = \frac{1}{(2\pi\frac{\hbar^2}{m}\Delta\tau)^{\frac{3A}{2}}}\int dR' e^{-\frac{(R-R')^2}{2\frac{\hbar^2}{m}\Delta\tau}}\Psi(R',0). \tag{9.83}$$

The short-time approximation for the propagator, correct at order $\Delta\tau$, will then become:

$$\langle R|e^{-\frac{\Delta\tau}{2}(\hat{H}-E_0)}|R'\rangle \sim \frac{1}{(2\pi\frac{\hbar^2}{m}\Delta\tau)^{\frac{3A}{2}}}e^{-\frac{\Delta\tau}{2}(V(R)-E_0)}e^{-\frac{(R-R')^2}{2\frac{\hbar^2}{m}\Delta\tau}}e^{-\frac{\Delta\tau}{2}(V(R)-E_0)}. \tag{9.84}$$

At this point it is possible to proceed in different ways. By substituting Eq. (9.84) in Eq. (9.79), one obtains an integral in which the integrand is a function of M replicas of the coordinates of the particles in the system. The ground-state expectation value of an operator that is a function of the coordinates can then be computed using on the left and on the right the imaginary time propagation started from an arbitrary state $\Psi(R,0)$. The resulting expression is:

$$\langle\phi_0|O(R)|\phi_0\rangle = \lim_{\tau\to\infty}\langle\Psi(R,\tau)|O(R)|\Psi(R,\tau)\rangle \sim \left(\frac{1}{(2\pi\frac{\hbar^2}{m}\Delta\tau)^{\frac{3A}{2}}}\right)^M \times$$

$$\times\int\int\cdots\int dR\,dR'\cdots dR^M\Psi(R,0)e^{-\frac{\Delta\tau}{2}(V(R)-E_0)}e^{-\frac{(R-R')^2}{2\frac{\hbar^2}{m}\Delta\tau}}e^{-\Delta\tau(V(R')-E_0)}\cdots O(R^{M/2})$$

$$\cdots e^{-\Delta\tau(V(R^{M-1})-E_0)}e^{-\frac{(R^{M-1}-R^M)^2}{2\frac{\hbar^2}{m}\Delta\tau}}e^{-\frac{\Delta\tau}{2}(V(R^M)-E_0)}\Psi(R^M,0) \tag{9.85}$$

This expression is reminiscent of a path-integral formulation of the problem. The integral can in principle be computed by means of a Metropolis-like algorithm, and gives the ground-state expectation of an arbitrary observable, provided that the number of slices M used is large enough to guarantee a correct filtering of the ground state. This method is known as Path Integral Ground State Monte Carlo (PIGS-MC) [8].

However, there is a simpler way to implement the imaginary time propagation. Let us expand the initial state from which we want to project the ground state in eigenstates of the position:

$$|\Psi\rangle \simeq \sum_i \Psi(R_i)|R_i\rangle \tag{9.86}$$

We will call each of these points in coordinate space a *walker*, and we will refer to the whole ensemble of points as to the *population* of walkers. If we apply the short-time propagator to each walker, it is easy to understand its effect. We will call the application of the short-time propagator to the walker population an *imaginary time step* (or simply a *time step*). Each time step originates a new *generation* of walkers.

The Gaussian factor in the propagator tells us the probability that a walker positioned in R' is displaced to a new position R. Since the probability density is a Gaussian of variance $\sigma^2 = \frac{\hbar^2}{m}\Delta\tau$, the RMS displacement will be proportional to $\sqrt{\Delta\tau}$ times a constant, which plays the role of a *diffusion constant* D, equal to $\frac{\hbar^2}{m}$. For each coordinate of each particle we need to extract a random number η distributed as:

$$P[\eta] = \frac{1}{\sqrt{2\pi D\Delta\tau}}e^{-\frac{\eta^2}{2D\Delta\tau}} \tag{9.87}$$

and add it to the original coordinate.

```
DMC_Move()
for i ∈ {0, A − 1} do
    for j ∈ {0, 2} do
        R_new[i][j] ← R[i][j] + DΔτ[rgaus()]
    end for
end for
```

The function rgaus() generating normally-distributed random numbers is now universally available as a library routine, but it can easily be implemented starting from a uniform distribution by using the Box-Muller formula. The part of the propagator depending on the potential has a slightly different interpretation. In the classical analogy we could say that the factor $W = e^{-\frac{\Delta\tau}{2}(V(R^M)-E_0)}$ represents the probability of a process to occur by which new points might be created in the time interval $\Delta\tau$ (if $W > 1$) or destroyed (if $W < 1$), or in other words, a process related to the presence of a source or a sink of walkers. W is interpreted as the average number of walkers that this process would generate over time at the position R. As we will later see, this creation/absorption (or *branching*) process is related to the fact that the normalization of the propagated state is not preserved. Since we cannot work with a non-integer number of walkers, we can use the following strategy

1. use the quantity W as a weight for the contribution to the estimates from the walker at a given position. Since in the short-time propagator we have two such factors, one from the initial position and on from the final position of the walker, we can use the product of the two as the total weight:

$$W = \exp\left\{-\Delta\tau\left[\frac{V(R) + V(R')}{2} - E_0\right]\right\} \tag{9.88}$$

   Estimates will be integrals of the form $\langle O\rangle = \int \phi_0(R)O(R)dR$, and they can be computed as:

$$\langle O\rangle = \frac{\sum_l^{N_{wk}} W_{kl}O(R_{kl})}{\sum_l^{N_{wk}} W_{kl}}, \tag{9.89}$$

   where N_{wk} is the number of walkers in a given generation. We will discuss later the specific form of the function O for interesting cases.
2. In order to generate a number of points that is correct on average, we can sample N_{mult}, the number of points to be generated for the next generation, in the following way:

$$N_{mult} = \text{int}(W + \xi), \tag{9.90}$$

where int() is the function truncating the argument to an integer, and ξ is a random number in $[0, 1)$. N_{mult} could be ≥ 1, in which case the next generation will contain *Nmult* copies of the walker, or 0, in which case the walker is suppressed.

The projection of the ground state will be achieved when propagating for a sufficiently long imaginary time. This means that we need to evolve the population of walkers for a large number of time steps, and eventually we will sample a density of points with a distribution *proportional* to the ground-state wavefunction. In the initial stage of the run, the energy and other estimators will have a value that is still strongly biased by the initial state. This means that the initial part of the propagation should be excluded from the averages. There is no automatic recipe to choose how much of the walk should be discarded. Usually it is convenient to monitor some observable (typically energy) and try to see where its value stops having a systematic trend as a function of the imaginary time.

How is the constant E_0 fixed? In principle it should be equal to the ground-state energy. This would mean that we need to know the solution of the problem. . . before solving it! In practice it is not strictly necessary to use the exact value of E_0, but it is sufficient to use a realistic variational estimate. The value of E_0 can also be used to reduce the fluctuations in the population of walkers due to the branching process, at the cost of introducing additional bias. For example, it is possible to modify the weight of a given configuration in the following way:

$$\tilde{W} = \frac{N_t}{N_g} \exp\left\{-\Delta\tau \left[\frac{V(R) + V(R')}{2} - E_0\right]\right\} \tag{9.91}$$

where N_t is a "target" number of walkers in the population and N_g is the number of walkers in the current generation. This modified weight reacts to the variations of the population, increasing or decreasing the weight depending on whether N_g is smaller or larger than N_t, respectively. This modification obviously introduces a bias in the results, since it modifies the propagator. However, this bias will be linearly decreasing with the time-step $\Delta\tau$. The weight can also be rewritten as:

$$\tilde{W} = \exp\left\{-\Delta\tau \left[\frac{V(R) + V(R')}{2} - \tilde{E}\right]\right\}, \tag{9.92}$$

where

$$\tilde{E} = E_0 + \frac{1}{\Delta\tau} \log\left(\frac{N_t}{N_g}\right) \tag{9.93}$$

Therefore, at each generation the constant can be modified to keep the size of the population under control.

The weight can also be used to estimate the energy. In fact if we take the logarithm of both members of Eq. (9.92) we obtain:

$$\log \tilde{W} = -\Delta\tau \left[\frac{V(R) + V(R')}{2} - \tilde{E} \right] \tag{9.94}$$

from which we obtain:

$$E_0 = \frac{1}{\Delta\tau} \log \left(\frac{N_g \tilde{W}}{N_t} \right) + \frac{V(R) + V(R')}{2}. \tag{9.95}$$

This is the so-called *growth energy* estimator, and it can be used in principle to evaluate the ground-state eigenvalue.

A simpler way of evaluating the energy is to use a test function $\Psi_T(R)$. In this case the idea is to evaluate the following matrix element:

$$\langle E \rangle = \frac{\langle \phi_0 | \hat{H} | \Psi_T \rangle}{\langle \phi_0 | \Psi_T \rangle} = \frac{\int dR \phi_0(R) \hat{H} \Psi_T(R)}{\int dR \phi_0(R) \Psi_T(R)} \tag{9.96}$$

Both numerator and denominator integrals are suitable for Monte Carlo evaluation. The probability density that we sample is $\phi_0(R)$, and the functions to be cumulated following the recipe in Eq. (9.89) are $\hat{H}\Psi_T(R)$ and $\Psi_T(R)$. The latter is necessary whenever $\Psi_T(R)$ is not normalized. We will then have

$$\langle E \rangle = \frac{\langle \hat{H} \Psi_T \rangle}{\langle \Psi_T \rangle}. \tag{9.97}$$

However, due to the hermiticity of the hamiltonian, one has:

$$\langle E \rangle = \frac{\langle \phi_0 | \hat{H} | \Psi_T \rangle}{\langle \phi_0 | \Psi_T \rangle} = \frac{\langle \Psi_T | \hat{H} | \phi_0 \rangle}{\langle \phi_0 | \Psi_T \rangle} = E_0, \tag{9.98}$$

independent of the choice of $\Psi_T(R)$. This is the most practical way to evaluate the energy eigenvalue and its standard deviation. Other observables can be evaluated in a similar way. However the results will always depend on the choice of the test function. We will discuss this aspect later.

A last important remark remains to be made. In devising the algorithm we are making some approximations. First of all the imaginary time propagator is not exact, but is correct only at order $\Delta\tau^2$. This means that for any finite imaginary time step value, the answer will be biased of an amount proportional to $\Delta\tau^2$. The same holds for the population size whenever one wants to apply population control as described above. For any finite target population N_t there will be a bias on the answer of order $1/N_t$. These biases can be corrected by performing several simulations with different values of $\Delta\tau$ and N_t, and then extrapolating to $\Delta\tau \to 0$ and $1/N_t \to 0$. As we will

show in the last part of this chapter, methods exist to completely eliminate the time step bias. However, it is possible to reduce the bias with some minor modifications in the propagator and by introducing an acceptance/rejection mechanism (cite CYRUS TIME STEP).

9.3.3 Application to the Harmonic Oscillator

A very simple illustration of the sense of the algorithm can be made by implementing to the one-dimensional harmonic oscillator. We consider the Hamiltonian:

$$\hat{H} = -\frac{1}{2}\frac{\partial^2}{\partial x^2} + \frac{1}{2}x^2 \tag{9.99}$$

The ground-state eigenvalue is $E_0 = \frac{1}{2}$ and the ground-state eigenfunction is the Gaussian $\Psi_0(x) = \frac{1}{\pi^{1/4}} e^{-\frac{x^2}{2}}$. As we have illustrated in the previous section, the propagation can start from any distribution of points with a density not orthogonal to the ground state. A very simple choice in this case is a constant. In Fig. 9.3 we can see how the histogram of the walkers evolves as a function of the imaginary time applying the algorithm described in the previous section, including population control. The initial uniform distribution of walkers in the interval $[-6, 6]$ is transformed into the correct Gaussian density. The mechanism that leads to this

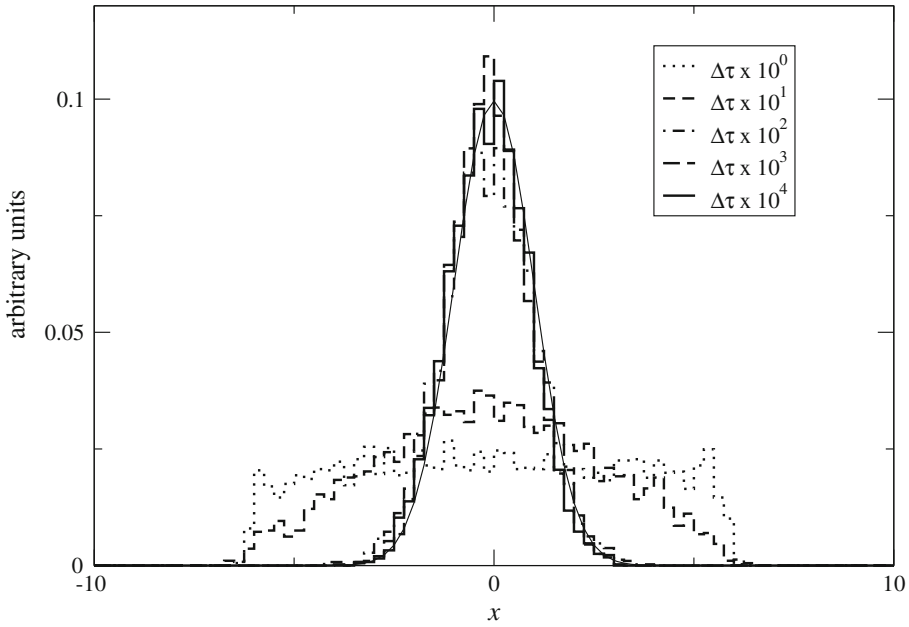

Fig. 9.3 Histogram of the walker population after N DMC imaginary time steps for the harmonic oscillator Hamiltonian described in text. Here we used $\Delta\tau = 10^{-3}$, with a target population of 4000 walkers

result is easy to understand. Any walker finding itself after diffusion in a region where the potential is larger than the eigenvalue will tend to be suppressed, while walkers near the origin will tend to multiply themselves. This will result in a histogram peaked at the origin ad decaying fast to zero when moving away from it.

In order to estimate the energy we need a test function. An approximation to the ground state might be given by the function:

$$\Psi_T(x) = \frac{1}{1+x^2}. \tag{9.100}$$

We can therefore estimate the energy by means of the following quotient [see Eq. (9.97)]:

$$\langle E \rangle = \frac{\sum_i w(x_i) \frac{1-3x_i^2}{(1+x_i^2)^3} + \frac{1}{2}\frac{x_i^2}{1+x_i^2}}{\sum_i w(x_i) \frac{1}{1+x_i^2}}, \tag{9.101}$$

where the sums runs first over all the generations (i.e. the imaginary time steps performed) and then over all the walkers belonging to a given generation.

In Fig. 9.4 we show the logarithm of the energy estimator averaged over each single generation as a function of the imaginary time. As we would expect from the general behavior of the coefficients of the excited states as function of the imaginary

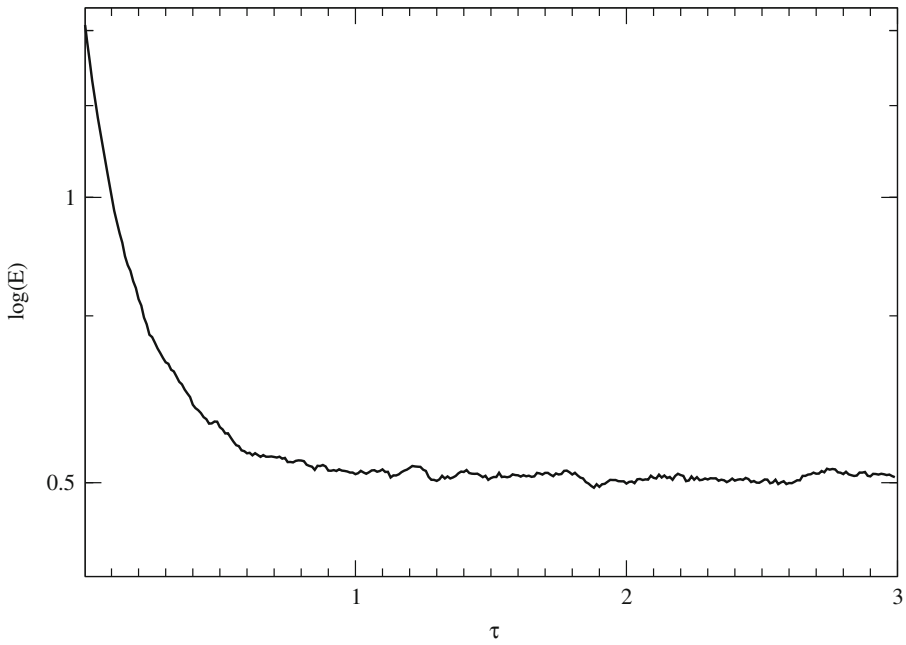

Fig. 9.4 Logarithm of the estimated energy averaged over a single generation as a function of the imaginary time in a run with a population target of 4000 walkers, and with an imaginary time step $\Delta\tau = 10^{-4}$

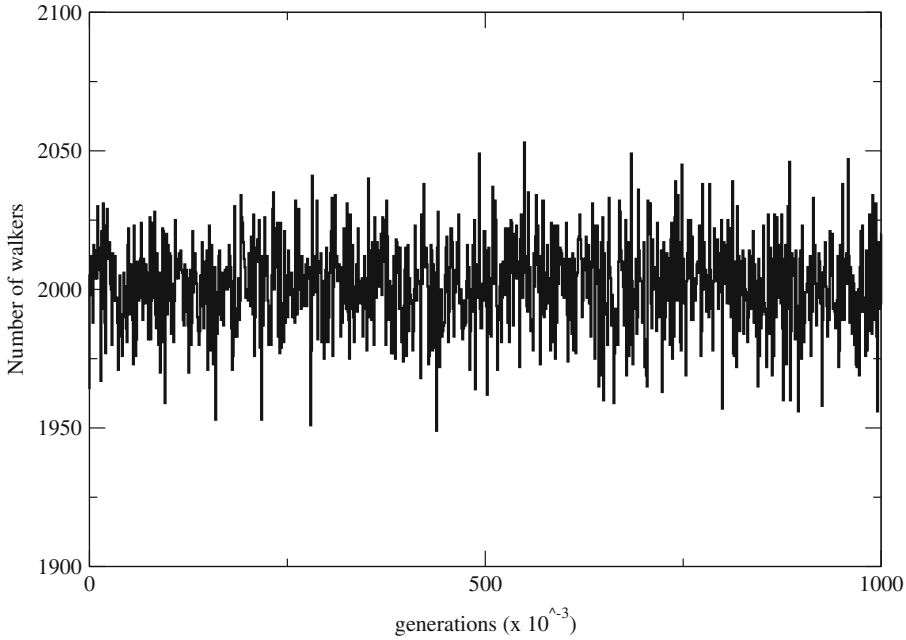

Fig. 9.5 Typical fluctuations of the walker population in a DMC run for the one dimensional harmonic oscillator. The target population in this case is 2000. The imaginary time step is set to $\Delta\tau = 0.075$

time, we see a clear exponential decay of the energy towards the exact eigenvalue. The figure clearly shows how the transient is not made up of a single exponential. The initial state needs includes a large number of excited states, that all need to be projected out before reaching the ground state. In Fig. 9.5 the typical behavior of the fluctuation in the walker number is reported. In the specific case the time step was set to $\Delta\tau = 0.3$. Nevertheless, the walker number never departs from the target by more than 3%. This is the effect of the population control procedure described in the previous subsection. Unfortunately population control alone is not sufficient to guarantee a stable calculation. In presence of particle-particle interactions that diverge at the origin fluctuations in the number of walkers become extremely wide. This is the reason why it is necessary to introduce the so-called importance sampling, that we will discuss in a later section.

Finally, in Fig. 9.6 we show one of the points discussed in the previous section, that is the bias of the result due to the finite imaginary time step. The difference between the energy estimate and the exact eigenvalue is plotted as a function of $\Delta\tau$ for a target population of 2000 walkers and a total of 10^5 generations for each value of $\Delta\tau$. The observed bias is quite small, but well outside of the statistical error. The dependence on $\Delta\tau$ is quadratic, as expected from the analysis of the propagator. Interpolating the data with a function of the form $E = E_0 + \alpha(\Delta\tau)^2$ we predict E_0 to be $(-3 \pm 1) \times 10^{-5}$. As it can be seen there is still a small residual bias due to the finiteness of the population. Further extrapolation would be needed to recover the exact answer.

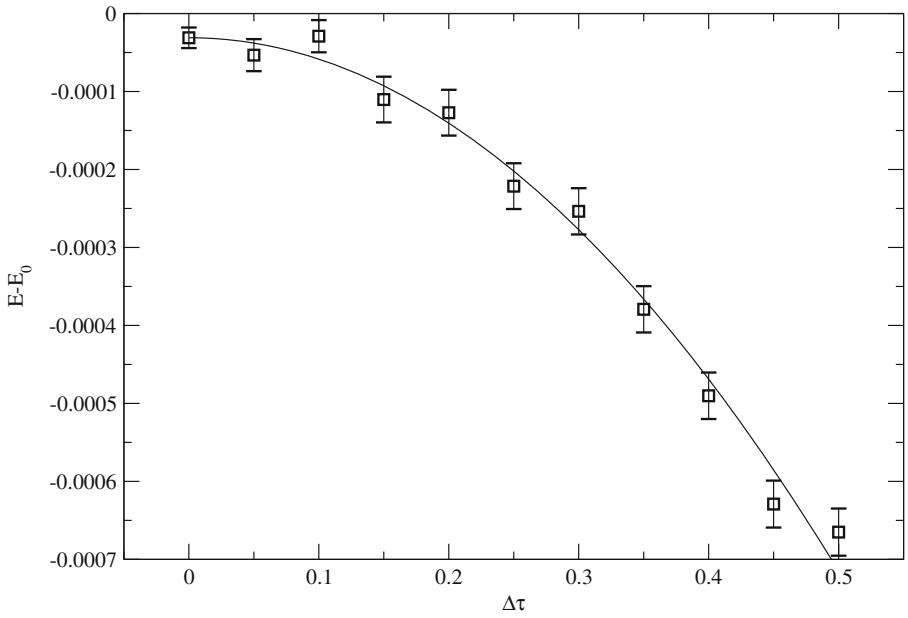

Fig. 9.6 Illustration of the imaginary time-step extrapolation procedure. The energy is computed for different values of $\Delta\tau$, and the results are fitted with a linear function. The intercept will give the correct prediction for the eigenvalue. Notice that the results should still be extrapolated for an infinite population. Here we use a target number of walkers equal to 2000, and the runs consist of 10^5 generations each. Errorbars refer to one standard deviation. The plotted value for $\Delta\tau = 0$ and the corresponding errorbar are obtained from the linear fit of the data

9.3.4 Importance Sampling

The simple diffusion algorithm we have illustrated above suffers of a substantial deficiency when particles interact with a potential having a repulsive or attractive core. Since the free particle diffusion propagator does not have any information about the potential, particles have no restrictions to come close to each other. This means that the weights will suffer of large fluctuations whenever a pair of particles find themselves at short distance. The consequent fluctuations in the population make the computation unmanageable.

The use of an importance function to guide the diffusion process [9] was the key to make Diffusion Monte Carlo algorithms usable. The idea is to give up on the request of sampling the ground-state wavefunction, and rather try to sample a distribution that, asymptotically in imaginary time, is the product of the ground-state wavefunction and of a known function that is the best possible approximation to the ground state obtained, for instance, by means of a variational calculation. We will call this function Ψ_G. Starting from Eq. (9.75) we can multiply both sides by Ψ_G and obtain:

$$\Psi_G(R)\Psi(R,\Delta\tau) = \int dR' G'(R',R,\Delta\tau)\Psi_G(R)\Psi(R',0), \tag{9.102}$$

where we have defined:

$$G'(R,R',\Delta\tau) = \frac{1}{(2\pi\frac{\hbar^2}{m}\Delta\tau)^{\frac{3A}{2}}} e^{-\Delta\tau(V(R')-E_0)} e^{-\frac{(R'-R)^2}{2\frac{\hbar^2}{m}\Delta\tau}}. \tag{9.103}$$

Since all the expressions we have written are correct at order $\Delta\tau$, for our purposes we can assume the equivalence of G' and G. We can multiply and divide the integrand in Eq. (9.102) by $\Psi_G(R')$ to obtain:

$$\Psi_G(R)\Psi(R,\Delta\tau) = \int dR' G'(R',R,\Delta\tau)\frac{\Psi_G(R)}{\Psi_G(R')}\Psi_G(R')\Psi(R',0),. \tag{9.104}$$

In Eq. (9.104) we can identify a new walker density to be propagated, namely:

$$f(R,\tau) = \Psi_G(R)\Psi(R,\tau), \tag{9.105}$$

and the corresponding propagator:

$$\tilde{G}(R,R',\Delta\tau) = G'(R,R',\tau)\frac{\Psi_G(R)}{\Psi_G(R')}. \tag{9.106}$$

The quotient of the wavefunctions can be included in the weight, and provides a correction that prevents the walkers to excessively multiply or die near the divergent points of the potential This point is better illustrated considering the short time limit it is possible to expand the ratio of the guiding functions. At first order in $\Delta\tau$ the result is:

$$\tilde{G}(R,R',\Delta\tau) \simeq G_0(R,R',\tau)\left[1 + \frac{\nabla\Psi_G(R')}{\Psi_G(R')}(R-R') + \cdots\right] \tag{9.107}$$

At the same order we can regard the terms in bracket as the expansion of an exponential and write:

$$\tilde{G}(R,R',\Delta\tau) \simeq G_0(R,R',\tau) e^{\frac{\nabla\Psi_G(R')}{\Psi_G(R')}(R-R')} \tag{9.108}$$

This can be combined with the Gaussian factor in G_0, and by completing the square (which introduces a term at order $\Delta\tau 2$), the propagator is modified as follows:

$$\tilde{G}(R,R',\Delta\tau) \simeq \frac{1}{(2\pi\frac{\hbar^2}{m}\Delta\tau)^{\frac{3A}{2}}} e^{-\frac{\Delta\tau}{2}(V(R')-E_0)} e^{-\frac{(R-R'-\frac{\hbar^2}{m}\Delta\tau\frac{\nabla\Psi_G(R')}{\Psi_G(R')})^2}{2\frac{\hbar^2}{m}\Delta\tau}} e^{-\frac{\Delta\tau}{2}(V(R)-E_0)}. \tag{9.109}$$

The same expansion can be performed to compute the *change in normalization* of the propagated density after a time step. The change in normalization is given by:

$$\mathcal{N} = \int dR\tilde{G}(R, R', \tau), \tag{9.110}$$

i.e. the total weight of the final points R that can be reached starting from R'. Once more we can expand the ratio of the guiding functions in the propagator, but this time up to second order:

$$\tilde{G}(R, R', \Delta\tau) \simeq G_0(R, R', \tau)\left[1 + \frac{\nabla\Psi_G(R')}{\Psi_G(R')}(R - R') + \frac{1}{2}\frac{\partial_{i\alpha}\partial_{j\beta}\Psi_G(R')(R - R')_{i\alpha}(R - R')_{j\beta}}{\Psi_G(R')} + \cdots\right] \tag{9.111}$$

Inserting the previous equation in Eq. (9.110) we can see that after integrating over R the terms containing odd powers of $(R - R')$ disappear by parity. We are therefore left with:

$$\mathcal{N} = e^{-\Delta\tau[V(R')-E_0]}\left[1 + \frac{1}{2}\frac{\nabla^2\Psi_G(R')}{\Psi_G(R')}\frac{\hbar^2}{m}\Delta\tau + \cdots\right] \tag{9.112}$$

We can now use the same trick used above to write the expression in square parenthesis as an exponential. The result is:

$$\mathcal{N} = \exp\left[-\Delta\tau\left(V(R') - \frac{\hbar^2}{2m}\frac{\nabla^2\Psi_G(R')}{\Psi_G(R')} - E_0 z\right)\right] \tag{9.113}$$

In the previous expression it is possible to immediately recognize the local energy. In fact, when using importance sampling, the normalization assumes the expression:

$$\mathcal{N} = \exp\left[-\Delta\tau\left(\frac{\hat{H}\Psi_G(R')}{\Psi_G(R')} - E_0\right)\right] \tag{9.114}$$

This is the new form of the weight factor that one needs to compute in order to determine the multiplicity of the walker at a given position. It is immediately clear that the fact that in the exponential we have the difference between the local energy, instead of the potential energy, and the reference eigenvalue E_0 essentially resolves the issue related to the fluctuations of the population related to a divergent behavior of the interaction. In fact, if we knew the exact solution the exponent would be identically zero, and the population would be absolutely stable. However, by means of an accurate variational calculation it is possible to obtain a very good approximation of the ground-state wavefunction, thereby reducing the fluctuations in the population to a minimum.

The algorithm including importance sampling is modified in the following way.

1. For each walker, and for each coordinate perform a "drift" move along the gradient of the guiding function. This displacement is deterministic.

```
DMC_Drift()
for i ∈ {0, A − 1} do
    for j ∈ {0, 2} do
        R_drift[i][j] ← R[i][j] + ∇Ψ_G(R)/Ψ_G(R) |_[i][j] DΔτ
    end for
end for
```

2. Cycle again over coordinates and diffuse the position from R_{drift} as in the non-importance sampled case.
3. Compute the new multiplicity of the walker and the weight to assign to estimators using

$$W = \exp\left[-\Delta\tau\left(\frac{\hat{H}\Psi_G(R')}{\Psi_G(R')} - E_0\right)\right] \tag{9.115}$$

In this way the walkers will asymptotically sample the distribution:

$$f(R) = \Psi_G(R)\phi_0(R). \tag{9.116}$$

This means that it is possible to evaluate integrals of the form:

$$\langle O\rangle = \frac{\int dR f(R)O(R)}{\int dR f(R)}. \tag{9.117}$$

As in the previous case the evaluation of the exact energy eigenvalue can be easily obtained by using the local energy. In fact, the matrix element of the Hamiltonian between the guiding function[4] and the ground-state wavefunction is:

$$\langle E\rangle = \frac{\langle\phi_0|\hat{H}|\Psi_G\rangle}{\langle\phi_0|\Psi_G\rangle} = \frac{\int dR f(R)\frac{\hat{H}\Psi_G(R)}{\Psi_G(R)}}{\int dR f(R)} \tag{9.118}$$

Once more, because of the hermiticity of the Hamiltonian we will have that $\langle E\rangle = E_0$. All other estimators will be matrix elements of the operator between Ψ_G and ϕ_0.

[4]It is always possible to project the energy from a function Ψ_T other than Ψ_G, by introducing a further weighing factor $\frac{\Psi_T}{\Psi_G}$. However this is very rarely used in standard applications.

9.3.5 The Fermion Sign Problem

As we have mentioned, imaginary time propagation projects out of an arbitrary initial state the absolute (mathematical) ground state of a given Hamiltonian $\hat{H}$, which is always a nodeless function. One might correctly object that if the initial state is chosen in such a way not to have any overlap with this ground state, the projection will correctly give back some excited state of $\hat{H}$. More rigorously, if our initial state has components only within a certain subspace of the total Hilbert space, which could be selected, for instance, by the wavefunction symmetry, then imaginary time propagation will end up projecting out the eigenstate with lowest eigenvalue within that given subspace.

This seems to be particularly useful when thinking of applying DMC-like algorithm to the study of many-fermion systems, as the nuclear systems we are interested in. The antisymmetry property of the fermionic ground state suggests that it should be sufficient to start from an arbitrary antisymmetric state $|\Psi_A\rangle$ (provided it is not orthogonal to the fermion ground state) to obtain the sought solution. In fact, one might speculate that antisymmetry itself would guarantee that there is no overlap with the symmetric ground state since the beginning:

$$\begin{aligned}\lim_{\tau\to\infty} e^{-\tau(\hat{H}-E_0^A)}|\Psi_A(0)\rangle &= \sum_n e^{-\tau(E_n-E_0^A)}\langle\phi_n|\Psi_A\rangle|\phi_n\rangle = \\ &= \langle\phi_0^A|\Psi_A\rangle|\phi_0^A\rangle \\ &\quad + \lim_{\tau\to\infty}\langle\phi_0|\Psi_A\rangle|\phi_0\rangle e^{-\tau(E_0-E_0^A)} \qquad (9.119)\end{aligned}$$

However, this abstract formulation forgets that eventually we need to *sample a probability density* in order to operate with a Monte Carlo integration, and any excited state will have a wavefunction changing sign somewhere, thereby breaking this requirement. If we had an exact knowledge of the *nodal surface* of the ground state (i.e. of the set of points such that $\phi_0^A(0) = 0$), we could use an antisymmetric function $\Psi_G^A(R)$ having the same nodal surface, and obtain by importance function the required positive definite density to sample:

$$\langle E\rangle = \frac{\langle\phi_0^A|\hat{H}|\Psi_G^A\rangle}{\langle\phi_0^A|\Psi_G^A\rangle} = \frac{\int dR\phi_0^A(R)\Psi_G^A(R)\dfrac{\hat{H}\Psi_G^A(R)}{\Psi_G^A(R)}}{\int dR\phi_0^A(R)\Psi_G^A(R)}. \qquad (9.120)$$

If $\Psi_G^a(R)$ does not have the same nodal surface as $\phi_0^A(R)$, we are once again in trouble.

We might have then the idea of separately sampling the positive and the negative part of the wave function. It is always possible to split an antisymmetric function as:

$$\psi^A(R) = \Psi^+(R) - \Psi^-(R), \qquad (9.121)$$

where both Ψ^+ and Ψ^- are positive definite functions. It is easy to see that each one, by linearity, is a solution of the Schrödinger equation with the same eigenvalue as the fermionic ground state. We can call $|R^+\rangle$ the walkers sampling the positive part and $|R^-\rangle$ the walkers sampling the negative part of Ψ^A. The energy expectation could be computed as:

$$E_0^A = \frac{\int dR^+ f^+(R^+) \frac{\hat{H}\Psi_G^A(R^+)}{\Psi_G^A(R^+)} - \int dR^+ f^-(R^-) \frac{\hat{H}\Psi_G^A(R^-)}{\Psi_G^A(R^-)}}{\int dR^+ f^+(R^+)\Psi_G^A(R^+) - \int dR^+ f^-(R^-)\Psi_G^A(R^-)}, \tag{9.122}$$

where $f^{\pm}$, as above, has the meaning of the importance sampled density of walkers. However, once more we have to notice that since both f^+ and f^- will obey the same imaginary time Schrödinger equation, the two densities will both converge to the ground-state density for $\hat{H}$. This means that both the numerator and the denominator of Eq. (9.122) will tend to 0 in the limit $\tau \to \infty$, and the ratio becomes undetermined. The major effect that one can observe during the calculation is that the variance of the energy will become exponentially large, and the integral will be dominated by statistical noise. This is the so called *fermion sign problem*. For some authors there is a prove that the computation of estimates such as Eq. (9.122) is an NP complex problem [10], and a solution will always require computer time that is exponentially increasing with the dimension of the system. However there are hints that by using methods that break this *plus/minus symmetry*, based on correlated dynamics and cancellation methods it is possible to reduce the cost to a polynomial dependence [11, 12].

9.3.5.1 Fixed-Node Approximation

A possible way of circumventing the sign problem in the case in which the antisymmetric ground-state wavefunction has to be real is to use some artificial boundary conditions [13].

We can define a nodal pocket $\Omega(R)$ as the set of points that can be reached from R without crossing the nodal surface at any point. For a standard Hamiltonian we can expect that for any pair of points R', R not on the nodal surface of the wavefunction, there exist a permutation P of the coordinates such that $PR' \in \Omega(R)$. This in turn means that all the space (but for the nodal surface, which has zero measure) can be covered by summing over all the permutations of the points lying in a single nodal pocket $\Omega(R)$. This the so-called *tiling theorem*. The tiling theorem implies that the fermion ground-state eigenvalue of the Schrödinger equation solved inside any $\Omega(R)$ is the same as the eigenvalue of the problem solved on the whole space.

The prove of the tiling theorem is quite simple. If the tiling property does not hold for the antisymmetric ground state $\phi_0^A(R)$, then $\sum_P \Omega(PR)$ will not completely cover the space, leaving out some regions. This means that somewhere

there are two regions $\Omega(R_a)$ and $\Omega(R_b)$ that share part of the nodal surface and are not equivalent. It is then possible to construct a function with a lower eigenvalue in the region $\Omega(R_a) \bigcup \Omega(R_b)$ by simply removing the common node and solving for the ground state of $\hat{H}$ within that region. Let us call ϕ^0_{ab} this function. Constructing an antisymmetric function $\Psi_A(R) = \sum_P (-1)^P \phi^0_{ab}$ we will have an antisymmetric function with an eigenvalue lower than that of $\phi^A_0(R)$, thereby violating the assumption that ϕ^0_A is the antisymmetric ground state of $\hat{H}$.

By the same kind of construction it is also possible to prove that the solution of the Schrödinger equation within a given nodal pocket $\Omega(R)$ is always an upper bound of the true antisymmetric eigenvalue, and that the exact result is recovered if and only if the nodal surface of the wavefunction generated by replicating the pocket coincides with that of the exact eigenfunction.

The previous considerations suggest that solving for the ground state of a given Hamiltonian within a nodal pocket $\Omega(R)$ will provide an upper bound of the energy of a many-fermion system, which can in principle can by improved by improving the nodal structure of the test function used to determine the boundary conditions. This is called the *fixed-node approximation*. In order to have zero density at the nodal surface we have to assume that at the border of the nodal pocket an infinite absorbing potential exists, such that walkers never cross that surface. From the point of view of the algorithm this introduces a very tiny modification in the code. We have to remember that we can solve for the ground state in *any* pocket. This means that we do not need to care either of the initial position of the walkers or of the associated sign of the wavefunction. We said that the fixed-node approximation corresponds to modify the Hamiltonian as follows

$$\hat{\tilde{H}} = \hat{H} + V_\Omega(R), \tag{9.123}$$

where

$$V_\Omega(R) = \begin{cases} \infty & \text{if} \quad R \in S(\Omega) \\ 0 & \text{otherwise} \end{cases} \tag{9.124}$$

This means that every time the walker crosses the border of the nodal pocket $S(\Omega)$ its weight becomes zero, and the walker is simply canceled from the population. Fixed node calculations are presently very widely employed especially in quantum chemistry and solid state physics applications (for a review of applications to many electron systems see [14]). When the wavefunction needs to be complex it is no longer possible to define a nodal surface, and a different kind of approach has to be used. This will be discussed in the next section concerning the applications to the nuclear physics case.

9.4 Quantum Monte Carlo for Nuclear Hamiltonians in Coordinate Space

9.4.1 General Auxiliary Field Formalism

We begin by looking at the auxiliary field formalism without importance sampling. All such diffusion Monte Carlo methods can be formulated as

$$|\Psi(t+\Delta t)\rangle = \int dX P(X) T(X) |\Psi(t)\rangle \tag{9.125}$$

where X is a set of variables which will become our auxiliary fields, $P(X)$ is a probability density,

$$\begin{aligned} P(X) &\geq 0 \\ \int dX P(X) &= 1\,, \end{aligned} \tag{9.126}$$

and $T(X)$ is an operator that operates in the Hilbert space of $|\Psi(t)\rangle$. We are free to choose the variables X, the probability density $P(X)$, and the operator $T(X)$ subject only to the constraint that the integral gives the desired propagator

$$e^{-(H-E_T)\Delta t} = \int dX P(X) T(X)\,, \tag{9.127}$$

at least in the limit that $\Delta t \to 0$.

In diffusion Monte Carlo methods, we represent the state $|\Psi(t)\rangle$ as a linear combination of basis states which obviously must span the Hilbert space. These can be a complete set. An example is the position eigenstates used for diffusion Monte Carlo for central potentials. They can also form an overcomplete set such as or the position and spin/isospin bases used in the nuclear GFMC method and the position and overcomplete outer product of single particle spinor basis used in AFDMC, or the overcomplete single particle bases used in auxiliary field methods such as those developed by Zhang and coworkers. For either case, we can denote these basis states as possible "walkers." We will denote one of these walker states as $|RS\rangle$ since we will be applying the method to systems where the basis is given by the positions of the particles, R, and a spinor for each spin-isospin of the particles, S.

The state, $|\Psi(t)\rangle$, at time t is represented in diffusion Monte Carlo methods as a linear combination of walker states

$$|\Psi(t)\rangle = \sum_{i=1}^{N_W} w_i |R_i S_i\rangle \tag{9.128}$$

where w_i is a coefficient, often called the weight, and N_W is the number of walkers.

The key ingredient to implementing a diffusion Monte Carlo method is to choose the walker basis and the operator $T(X)$ such that when $T(X)$ operates on a walker basis state, it gives one and only one new walker basis state. That is we want

$$T(X)|RS\rangle = W(X,R,S)|R'S'\rangle \tag{9.129}$$

where $|R'S'\rangle$ is normalized in the same way as $|RS\rangle$, and $W(X,R,S)$ is the change in the normalization from the propagation.

Once we have arranged for Eq. (9.129) to be true, we can implement the diffusion Monte Carlo by starting with $|\Psi(0)\rangle$ written, as in Eq. (9.128), as any, not unreasonable, linear combination of walkers. For each walker, we sample X values from $P(X)$, and use Eq. (9.129) to propagate to a new walker $|R'_iS'_i\rangle$, with a new weight w'_i given by the proportionality constant of Eq. (9.129) multiplied by the original weight w_i. We branch on the magnitude of the weight, so usually, after branching, $w'_i = 1$, where we are ignoring the fermion sign or phase problem for now and assuming that all of the weights are greater than or equal to zero. We will deal with the fermion case below.

9.4.2 Operator Expectations and Importance Sampling

9.4.2.1 Mixed Averages

Diffusion Monte Carlo methods efficiently calculate ground-state mixed averages

$$\bar{O}_{\text{mixed}} = \frac{\langle\Psi_T|O|\Psi(t)\rangle}{\langle\Psi_T|\Psi(t)\rangle} \tag{9.130}$$

where $|\Psi_T\rangle$ is trial state. If O is the Hamiltonian, operating on $|\Psi(t)\rangle$ shows that the result is the ground-state energy for large t. For other operators, for which the ground state is not an eigenstate, either approximate extrapolation methods or forward walking or its equivalent must be used to extract the correct ground-state expectation value.

Given a set of walkers as in Eq. (9.128), the mixed estimate can be calculated by

$$\bar{O}_{\text{mixed}} \simeq \frac{\sum_{i=1}^{N_w} w_i\langle\Psi_T|O|R_iS_i\rangle}{\sum_{i=1}^{N_w} w_i\langle\Psi_T|R_iS_i\rangle} \tag{9.131}$$

where the right hand side differs from the correct result because of statistical errors from the sampling which decreases as $N_W^{-1/2}$, and possible population size bias which decreases as N_W^{-1}. Statistical errors can be minimized by reducing the variance through importance sampling. Population bias also can be controlled with importance sampling, and, since it decays faster with population size, can be readily detected and removed by either taking larger numbers of walkers or extrapolation.

Efficient Monte Carlo methods need to have low variance so that the statistical error bars can be made small. For our walker propagation, this means that we should sample new walkers not only corresponding to the weight they will receive from our algorithm, but with this weight multiplied by their expected survival probability. The imaginary time Schrödinger equation is self adjoint, so the optimum importance function is the desired function. Typically, a trial function that can be efficiently evaluated is determined variationally and used as an approximation to the optimum trial function. Usually this trial wave function is used as the importance function. Sometimes a different importance function is used, so we will write this more general case.

9.4.2.2 Importance Sampling

To add importance sampling, we arrange to sample our walkers from a new state which we call $|\Psi_I \Psi(t)\rangle$ such that

$$\langle RS|\Psi_I\Psi(t)\rangle = \langle \Psi_I|RS\rangle\langle RS|\Psi(t)\rangle \tag{9.132}$$

so that

$$|\Psi_I\Psi(t)\rangle = \sum_{i=1}^{N_w} w_i |R_i S_i\rangle \tag{9.133}$$

An alternative way of looking at this is that the sampling probability for the walkers at $R_i S_i$ has been modified so that

$$|\Psi(t)\rangle = \sum_{i=1}^{N_w} w_i \langle \Psi_I|R_i S_i\rangle^{-1} |R_i S_i\rangle \,. \tag{9.134}$$

Calculating a mixed average now becomes

$$\bar{O}_{\text{mixed}} = \frac{\sum_{i=1}^{N_w} w_i \frac{\langle \Psi_T|R_i S_i\rangle}{\langle \Psi_I|R_i S_i\rangle} \frac{\langle \Psi_T|O|R_i S_i\rangle}{\langle \Psi_T|R_i S_i\rangle}}{\sum_{i=1}^{N_w} w_i \frac{\langle \Psi_T|R_i S_i\rangle}{\langle \Psi_I|R_i S_i\rangle}} \,. \tag{9.135}$$

For the usual case where $|\Psi_I\rangle = |\Psi_T\rangle$, and $w_i = 1$, we have

$$\bar{O}_{\text{mixed}} = \frac{1}{N_w} \sum_{i=1}^{N_w} \frac{\langle \Psi_T|O|R_i S_i\rangle}{\langle \Psi_T|R_i S_i\rangle} \,. \tag{9.136}$$

We substitute Eqs. (9.132) and (9.133) into Eq. (9.125)

$$|\Psi_I \Psi(t+\Delta t)\rangle = \sum_{i=1}^{N_w} w_i \int dX P(X) \frac{\langle \Psi_I | R_i' S_i' \rangle}{\langle \Psi_I | R_i S_i \rangle} T(X) |R_i S_i\rangle$$
$$= \sum_{i=1}^{N_w} w_i \int dX P(X) \frac{\langle \Psi_I | T(X) | R_i S_i \rangle}{\langle \Psi_I | R_i S_i \rangle} \frac{T(X)}{W(X, R_i, S_i)} |R_i S_i\rangle \quad (9.137)$$

where $|R_i' S_i'\rangle$ is defined as in Eq. (9.129). Notice that the operator $T(X)/W(X, R_i, S_i)$ operating on $|R_i S_i\rangle$ gives a normalized walker. The additional weight of this walker is given by $P(X)\frac{\langle \Psi_I | T(X) | R_i S_i \rangle}{\langle \Psi_I | R_i S_i \rangle}$. We want to minimize fluctuations in this weight factor, and to do this we normalize it and sample from the normalized distribution. The normalization will be the weight.

We write

$$\mathcal{N} = \int dX P(X) \frac{\langle \Psi_I | T(X) | R_i S_i \rangle}{\langle \Psi_I | R_i S_i \rangle}$$
$$= \frac{\langle \Psi_I | e^{-(H - E_T)\Delta t} | R_i S_i \rangle}{\langle \Psi_I | R_i S_i \rangle}$$
$$= e^{-(E_L(R_i, S_i) - E_T)\Delta t} + O(\Delta t^2) \quad (9.138)$$

where the local energy $E_L(R_i, S_i)$ is defined by

$$E_L(R_i, S_i) = \frac{\langle \Psi_I | H | R_i S_i \rangle}{\langle \Psi_I | R_i S_i \rangle} \quad (9.139)$$

and we now sample X variables from the normalized distribution

$$\tilde{P}(X) = \mathcal{N}^{-1} P(X) \frac{\langle \Psi_I | T(X) | R_i S_i \rangle}{\langle \Psi_I | R_i S_i \rangle} . \quad (9.140)$$

The importance sampled diffusion Monte Carlo in the auxiliary field formalism becomes

$$|\Psi_I \Psi(t+\Delta t)\rangle = \sum_{i=1}^{N_w} w_i \int dX \tilde{P}(X) e^{-(E_L(R_i, S_i) - E_T)\Delta t} \frac{T(X)}{W(X, R_i, S_i)} |R_i S_i\rangle . \quad (9.141)$$

We propagate a walker by sampling an X value from $\tilde{P}(X)$, we include the local energy expression in the weight, and construct the new normalized walker position and spin state as $W^{-1}(X, R_i, S_i) T(X) |R_i S_i\rangle$. In each of the equations above, the ratio of the wave function terms gives the walker weight. In Eq. (9.141) these terms have been combined to give a weight that depends on the local energy expectation value.

All of the expectation values and weights contain ratios of trial wave functions so that any normalization factor multiplying the $|RS\rangle$ cancels and any convenient normalization can be used. We can therefore drop the W factors and normalize our walker kets at the end of a step. Typically just the walker positions are stored and the walker spinors are normalized to have magnitude 1.

9.4.2.3 Importance Sampling with a Hubbard-Stratonovich Transformation

We often have Hamiltonians where the Hubbard-Stratonovich transformation

$$e^{\frac{O^2}{2}} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-\frac{x^2}{2}} e^{xO} \tag{9.142}$$

can be used to write a propagator in the form of Eq. (9.129). Examples are writing the kinetic energy as an integral over translations, or writing terms like $\boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j = (\boldsymbol{\sigma}_i + \boldsymbol{\sigma}_j)^2 - 6$ as an integral over spin rotations.

Since we primarily use the Hubbard-Stratonovich transformation to define our auxiliary fields, it is useful to work out how importance sampling can be included within the short-time approximation for this particular case. We begin with a Hamiltonian that is quadratic in a set of N_O operators (which for our nuclear problems will be momentum and spin-isospin operators) O_n,

$$H = \frac{1}{2} \sum_{n=1}^{N_O} \lambda_n O_n^2 \tag{9.143}$$

so that the imaginary time propagator is

$$\begin{aligned} e^{-H\Delta t} = & \int dx \frac{1}{(2\pi)^{N_O/2}} e^{-\frac{1}{2}\sum_{n=1}^{N_O} x_n^2} e^{-i\sum_{n=1}^{N_O} x_n \sqrt{\lambda_n \Delta t} O_n} \\ & + O(\Delta t^2) \end{aligned} \tag{9.144}$$

where the Δt^2 terms comes from the possible non-commutativity of the O_n.

As before, we choose our walker basis and the operators O_n such that operating on a walker, $|RS\rangle$, with a term sampled from the integrand, gives a result proportional to another walker

$$e^{-i\sum_{n=1}^{N_O} x_n \sqrt{\lambda \Delta t} O_n} |RS\rangle = W(\{x_n\}, R, S) |R'S'\rangle \tag{9.145}$$

where $\{x_n\}$ represents the set of sampled x_n values.

We now sample $\tilde{P}(X)$ which is

$$\begin{aligned}\tilde{P}(X) &= \mathcal{N}^{-1} e^{-\frac{1}{2}\sum_{n=1}^{N_O} x_n^2} \frac{\langle \Psi_T | e^{-i\sum_{n=1}^{N_O} x_n \sqrt{\lambda_n \Delta t} O_n} | RS \rangle}{\langle \Psi_T | RS \rangle} \\ &= \mathcal{N}^{-1} e^{-\frac{1}{2}\sum_{n=1}^{N_O} x_n^2} \left(1 - i \sum_{n=1}^{N_O} x_n \sqrt{\lambda_n \Delta t} \frac{\langle \Psi_T | O_n | RS \rangle}{\langle \Psi_T | RS \rangle} \right. \\ &\quad \left. - \frac{1}{2} \sum_{n=1,m=1}^{N_O} x_n x_m \sqrt{\lambda_m \lambda_n} \Delta t \frac{\langle \Psi_T | O_n O_m | RS \rangle}{\langle \Psi_T | RS \rangle} + \ldots \right)\end{aligned} \tag{9.146}$$

Notice that if we were to expand $T(X)/W(X, R_i, S_i)$ it would have the form $1 + O(x_n \Delta t^{1/2}) + O(x_n x_m \Delta t) + \ldots$. Therefore if we drop terms of order Δt^2, the $O(\Delta t)$ term of $P(X)$ contributes only when it multiplies the 1 term from $T(X)/W(X, R_i, S_i)$. We can therefore integrate it over X without changing the result to this order in Δt. This term cancels the normalization, so that

$$\begin{aligned}\tilde{P}(X) &= e^{-\frac{1}{2}\sum_{n=1}^{N_O} x_n^2} \left(1 - i \sum_{n=1}^{N_O} x_n \sqrt{\lambda_n \Delta t} \frac{\langle \Psi_T | O_n | RS \rangle}{\langle \Psi_T | RS \rangle} + O(\Delta t^{3/2}) \right) \\ &= e^{-\frac{1}{2}\sum_{n=1}^{N_O} x_n^2} e^{-i\sum_{n=1}^{N_O} x_n \sqrt{\lambda_n \Delta t} \frac{\langle \Psi_T | O_n | RS \rangle}{\langle \Psi_T | RS \rangle} + \sum_{n=1}^{N_O} \lambda_n [\frac{\langle \Psi_T | O_n | RS \rangle}{\langle \Psi_T | RS \rangle}]^2} + O(\Delta t^{3/2}) \\ &= \exp\left\{ -\frac{1}{2} \sum_{n=1}^{N_O} \left[x_n + i \sqrt{\lambda_n \Delta t} \frac{\langle \Psi_T | O_n | RS \rangle}{\langle \Psi_T | RS \rangle} \right]^2 \right\}\end{aligned} \tag{9.147}$$

where in the last line, we have written the linear term in x in the exponent, and included a canceling term so that only the linear term survives integration to order Δt.

We sample our expression by sampling x_n from the shifted gaussian (Again, we assume here that $i\sqrt{\lambda_n \Delta t}\langle O_n \rangle$ is real. We will discuss what to do for the complex case below.)

$$x_n = \chi_n - i\sqrt{\lambda_n \Delta t}\langle O_n \rangle \tag{9.148}$$

where χ_n is sampled from a gaussian with unit variance. The new unnormalized ket is

$$|R'S'\rangle = e^{-i\sum_{n=1}^{N_O} x_n \sqrt{\lambda \Delta t} O_n} |RS\rangle \langle \Psi_T | RS \rangle \tag{9.149}$$

and its weight is given by the local energy expression

$$W(R', S') = e^{-[\langle H \rangle - E_T]\Delta t} \tag{9.150}$$

9.4.3 Application to Standard Diffusion Monte Carlo

9.4.3.1 Diffusion Monte Carlo Without Importance Sampling

It is helpful to apply the formalism above to derive the well known central potential diffusion Monte Carlo algorithm [9]. The Hamiltonian is

$$H = \sum_{j=1}^{A} \sum_{\alpha=1}^{3} \frac{p_{j\alpha}^2}{2m} + V(R) \tag{9.151}$$

where $p_{j\alpha}$ and R operate on Hilbert space, and $p_{j\alpha}$ is the α component of the momentum operator for the jth particle. Making the short-time approximation, the propagator can be written as

$$e^{-(H-E_T)\Delta t} = e^{\sum_{j=1}^{A} \sum_{\alpha=1}^{3} \frac{p_{j\alpha}^2}{2m} \Delta t} e^{-[V(R)-E_T]\Delta t} + O(\Delta t^2)\,. \tag{9.152}$$

Since the Hamiltonian does not operate on the spin, we can drop the spin variable from the our walker expressions and take just a position basis $|R\rangle$. Operating with the potential term

$$e^{-[V(R)-E_T]\Delta t}|R_j\rangle = e^{-[V(R_j)-E_T]\Delta t}|R_j\rangle \tag{9.153}$$

clearly satisfies Eq. (9.129). The kinetic energy part of the propagator does not satisfy Eq. (9.129). However, by using the Hubbard-Stratonovich transformation, we can write the kinetic energy in terms of the translation operators $e^{-\frac{i}{\hbar} p_{j\beta} a}$. We introduce the auxiliary field or Hubbard-Stratonovich variables, $x_{j\alpha}$, and write

$$\begin{aligned} & e^{-\sum_{j=1}^{A} \sum_{\alpha=1}^{3} \frac{p_{j\alpha}^2}{2m} \Delta t} \\ & \quad = \prod_{j\alpha} \frac{1}{(2\pi)^{3/2}} \int dx_{j\alpha} e^{-\frac{x_{j\alpha}^2}{2}} e^{-\frac{i}{\hbar} \mathbf{p}_{j\alpha} x_{j\alpha} \sqrt{\frac{\hbar^2 \Delta t}{m}}} \end{aligned} \tag{9.154}$$

With this definition, X is the set $\{x_{j\alpha}\}$, for the A particles,

$$P(X) = \prod_{j\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_{j\alpha}^2}{2}}\,, \tag{9.155}$$

and

$$T(X)|R\rangle = e^{-[V(R)-E_T]\Delta t}|R+\Delta R\rangle \tag{9.156}$$

where $R' = R + \Delta R$ is given by translating each particle's position in R

$$r'_{j\alpha} = r_{j\alpha} + x_{j\alpha}\frac{\hbar^2\Delta t}{m}\,. \tag{9.157}$$

This is identical to the standard diffusion Monte Carlo algorithm without importance sampling. We move each particle with a gaussian distribution of variance $\frac{\hbar^2\Delta t}{m}$, and include a weight of $e^{-[V(R)-E_T]\Delta t}$. We would then include branching on the weight to complete the algorithm.

While the Hubbard-Stratonovich transformation is the most common, there are many other possibilities. For example, the propagator for the relativistic kinetic energy $\sqrt{p^2c^2+m^2c^4}-mc^2$ can be sampled by using

$$e^{-\left[\sqrt{p^2c^2+m^2c^4}-mc^2\right]\Delta t} = \int d^3x f(x) e^{-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{x}} \tag{9.158}$$

with

$$\begin{aligned} f(x) &= \int \frac{d^3p}{(2\pi)^3} e^{\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{x}} e^{-\left[\sqrt{p^2c^2+m^2c^4}-mc^2\right]\Delta t} \\ &= e^{mc^2\Delta t} K_2\left(\frac{mc}{\hbar}\sqrt{x^2+c^2\Delta t^2}\right) \end{aligned} \tag{9.159}$$

where K_2 is the modified Bessel function of order 2 [15].

9.4.3.2 Importance Sampled Diffusion Monte Carlo in the Auxiliary Field Formulism

We break up the Hamiltonian as a kinetic and potential part. The potential part gives the usual $e^{-V(R)\Delta t}$ weight, and we need to work only with the importance sampled kinetic energy part. The kinetic energy operator is already written as a sum of squares,

$$KE = \sum_{j\alpha}\frac{p_{j\alpha}^2}{2m} \tag{9.160}$$

where j is the particle label and α is the x, y, or z coordinate. We can identify $\lambda_{j\alpha} = m^{-1}$, and $O_{j\alpha} = p_{j\alpha}$. Substituting this into our previous formalism, we have

$$\begin{aligned} i\sqrt{\lambda_{j\alpha}\Delta t}\langle O_{j\alpha}\rangle &= i\sqrt{\frac{\Delta t}{m}}\frac{\langle\Psi_T|p_{j\alpha}|RS\rangle}{\langle\Psi_T|RS\rangle} \\ &= -\sqrt{\frac{\hbar^2\Delta t}{m}}\frac{\partial_{j\alpha}\langle\Psi_T|RS\rangle}{\langle\Psi_T|RS\rangle}\,. \end{aligned} \tag{9.161}$$

The sampled value of $x_{j\alpha}$ will be

$$x_{j\alpha} = \chi_{j\alpha} + \sqrt{\frac{\hbar^2\Delta t}{m}}\frac{\partial_{j\alpha}\langle\Psi_T|RS\rangle}{\langle\Psi_T|RS\rangle} \tag{9.162}$$

where the $\chi_{j\alpha}$ are sampled from a gaussian with unit variance. The new walker will be

$$|R'S'\rangle = e^{-\frac{i}{\hbar}\sum_{j\alpha}x_{j\alpha}\sqrt{\frac{\hbar^2\Delta t}{m}}p_{j\alpha}}|RS\rangle\,. \tag{9.163}$$

Since $e^{-\frac{i}{\hbar}p_{j\alpha}a}$ is the translation operator that translates the ket's $j\alpha$ position coordinate by a. We have

$$\begin{aligned} S' &= S \\ R'_{j\alpha} &= R_{j\alpha} + x_{j\alpha}\sqrt{\frac{\hbar^2\Delta t}{m}} \\ &= R_{j\alpha} + \chi_{j\alpha}\sqrt{\frac{\hbar^2\Delta t}{m}} + \frac{\hbar^2\Delta t}{m}\frac{\partial_{j\alpha}\langle\Psi_T|RS\rangle}{\langle\Psi_T|RS\rangle} \end{aligned} \tag{9.164}$$

which is the standard diffusion Monte Carlo propagation. The weight factor is the local energy.

9.4.4 *Fixed-Phase Importance-Sampled Diffusion Monte Carlo*

The fixed-phase approximation [16] was developed to extend the fixed-node approximation to electrons in a magnetic field where the ground-state wave function is complex. The approximation enforces the trial function's phase as the phase for the calculated ground state. Diffusion Monte Carlo is used to sample the magnitude of the ground state.

If the walker phase has been chosen so that $\langle\Psi_T|R\rangle$ is real, the fixed-phase approximation requires that after propagation $\langle\Psi_T|R'\rangle$ would also be real since an

imaginary part would correspond to the calculated ground-state having a different phase than the trial function. Therefore in the implementation of the fixed-phase approximation we discard the imaginary part of the weight of a propagated walker. For an arbitrary initial phase, we discard the imaginary part of the ratio $\frac{\langle\Psi_T|R'\rangle}{\langle\Psi_T|R\rangle}$ which means that the we replace the importance sampled factor in Eq. (9.146) with its real part

$$\frac{\langle\Psi_T|e^{-i\sum_{n=1}^{N_O} x_n\sqrt{\lambda_n\Delta t}O_n}|RS\rangle}{\langle\Psi_T|RS\rangle} \rightarrow \mathrm{Re}\left[\frac{\langle\Psi_T|e^{-i\sum_{n=1}^{N_O} x_n\sqrt{\lambda_n\Delta t}O_n}|RS\rangle}{\langle\Psi_T|RS\rangle}\right]. \tag{9.165}$$

The fixed-phase algorithm for propagating a walker is then

1. Propagate to the new position (the spin does not change with a central potential)

$$\begin{aligned} S' &= S \\ R'_{j\alpha} &= R_{j\alpha} + \chi_{j\alpha}\sqrt{\frac{\hbar^2\Delta t}{m}} + \frac{\hbar^2\Delta t}{m}\mathrm{Re}\left[\frac{\partial_{j\alpha}\langle\Psi_T|RS\rangle}{\langle\Psi_T|RS\rangle}\right] \end{aligned} \tag{9.166}$$

2. Include a weight factor for the walker of

$$W = e^{-(\mathrm{Re}\langle H\rangle - E_T)\Delta t} \tag{9.167}$$

This is identical to the fixed-phase algorithm of Ortiz et al.

We will see that similar approximations can be used for our spin-isospin dependent problems.

9.4.5 Application to Quadratic Forms

Quadratic forms in operators that change from one walker to another can be diagonalized to produce the sum of squares needed for Eq. (9.143). That is for

$$H = \frac{1}{2}\sum_{ij} O_i A_{ij} O_j \tag{9.168}$$

with A_{nm} real and symmetric, we can calculate the normalized real eigenvectors and eigenvalues of the matrix A,

$$\sum_j A_{ij}\psi_j^{(n)} = \lambda_n|\psi_i\rangle$$
$$\sum_j \psi_j^{(n)}\psi_j^{(m)} = \delta_{nm}\,. \tag{9.169}$$

The matrix is then

$$A_{ij} = \sum_n \psi_i^{(n)}\lambda_n\psi_j^{(n)} \tag{9.170}$$

and substituting back we have

$$H = \frac{1}{2}\sum_n \lambda_n \mathcal{O}_n^2$$
$$\mathcal{O}_n = \sum_j \psi_j^{(n)} O_j\,, \tag{9.171}$$

which is now in the form of Eq. (9.143).

9.4.6 Auxiliary Field Breakups

There are many possible ways to break up the nuclear Hamiltonian using the auxiliary field formalism. As a concrete example let's look at the spinor propagator when we have a spin-exchange potential between A neutrons

$$V = \sum_{i<j} v^\sigma(r_{ij})\boldsymbol{\sigma}_i\cdot\boldsymbol{\sigma}_j \tag{9.172}$$

Taking the operators to be the x, y, and z components of the Pauli operators for each particle, we have a quadratic form in these $3A$ operators. Since walker gives the positions of the neutrons, we know the value of $v^\sigma(r_{ij})$ for all pairs. We can then write

$$V = \frac{1}{2}\sum_{ij}^{A} B_{ij}\sigma_{ix}\sigma_{jx} + \frac{1}{2}\sum_{ij}^{A} B_{ij}\sigma_{iy}\sigma_{jy} + \frac{1}{2}\sum_{ij}^{A} B_{ij}\sigma_{iz}\sigma_{jz} \tag{9.173}$$

where $B_{ii} = 0$, and $B_{ij} = v^{\sigma}(r_{ij})$ for $i \neq j$. Finding the eigenvectors $\psi_i^{(n)}$ and eigenvalues λ_n of the B matrix, we can write

$$V = \frac{1}{2}\sum_n \lambda_n (\mathscr{O}_{nx})^2 + \frac{1}{2}\sum_n \lambda_n (\mathscr{O}_{ny})^2 + \frac{1}{2}\sum_n \lambda_n (\mathscr{O}_{nz})^2$$

$$\mathscr{O}_{nx} = \sum_{i=1}^{A} \psi_i^{(n)} \sigma_{ix}$$

$$\mathscr{O}_{ny} = \sum_{i=1}^{A} \psi_i^{(n)} \sigma_{iy}$$

$$\mathscr{O}_{nz} = \sum_{i=1}^{A} \psi_i^{(n)} \sigma_{iz}\,. \tag{9.174}$$

Using the Hubbard-Stratonovich transformation would give us $3A$ auxiliary fields.

We can modify this transformation. For example, the diagonal elements of the B matrix are zero. Adding a nonzero diagonal term B_{jj}, would give us additional terms proportional to $\sigma_{jx}^2 = \sigma_{jy}^2 = \sigma_{jz}^2 = 1$, that is, these would be additional purely central terms. Subtracting a corresponding central contribution would then give an identical interaction, but different eigenvectors and therefore different spin rotation operators.

Another alternative would be to look at each term in the sum separately as a quadratic form of two operators. The resulting 2×2 matrices have two eigenvalues and eigenvectors so that

$$v^{\sigma}(r_{ij})\sigma_{ix}\sigma_{jx} = \frac{1}{4}v^{\sigma}(r_{ij})[\sigma_{ix} + \sigma_{jx}]^2 - \frac{1}{4}v^{\sigma}(r_{ij})[\sigma_{ix} - \sigma_{jx}]^2 \tag{9.175}$$

and each of the $3A(A-1)/2$ terms would require 2 auxiliary fields or $3A(A-1)$ total. We can reduce the number of auxiliary fields by including diagonal terms to our 2×2 matrix equal to the off diagonal terms. These make the eigenvector $(1,-1)/\sqrt{2}$ have a zero eigenvalue, which then does not contribute

$$v^{\sigma}(r_{ij})\sigma_{ix}\sigma_{jx} = \frac{1}{2}v^{\sigma}(r_{ij})[\sigma_{ix} + \sigma_{jx}]^2 - v^{\sigma}(r_{ij}) \tag{9.176}$$

where the second term on the right hand side is a central potential counter term that would be added to the physical central potential, and $3A(A-1)/2$ auxiliary fields would be required. This form could also be derived by expanding the square $[\sigma_{ix} + \sigma_{jx}]^2$.

Each of these breakups gives the same net propagator after integration of the auxiliary fields. If a good importance function is used, and the sampling can be carried out, we would expect the local energy for a complete step to have

low variance, and therefore the propagation to have low variance. The trade off then would be the complexity of constructing the operator combination versus the number of auxiliary fields needed in the propagation. In our work to date, we have used the full diagonalization to minimize the number of auxiliary field integrations. The cost of the diagonalization is order A^3 which is the same order as the cost for calculating a Slater determinant for the trial functions we need. However, it is easy to imagine having more complicated Hamiltonians where the cost of full diagonalization would be prohibitive (for example adding Δ degrees of freedom to the nuclei) and a simpler breakup using more auxiliary fields would be more efficient.

The best break up will be the one which optimizes the accuracy and variance of the results for a given amount of computational resources.

9.4.7 AFDMC with the v_6' Potential for Nuclear Matter

The Argonne v_6' potential includes central, spin and isospin exchange, and tensor interactions. Writing out the components, the Hamiltonian is

$$\begin{aligned} H = & \sum_{i\alpha} \frac{p_{i\alpha}^2}{2m} + \sum_{i<j} v^c(r_{ij}) + \sum_{i<j,\alpha\beta} \left\{ v^\sigma(r_{ij})\delta_{\alpha\beta} + v^t(r_{ij}) \left[3\hat{\alpha}\cdot\hat{r}_{ij}\hat{\beta}\cdot\hat{r}_{ij} - \delta_{\alpha\beta} \right] \right\} \sigma_{i\alpha}\sigma_{j\beta} \\ & + \sum_{i<j,\alpha\beta\gamma} \left\{ v^{\sigma\tau}(r_{ij})\delta_{\alpha\beta} + v^{t\tau}(r_{ij}) \left[3\hat{\alpha}\cdot\hat{r}_{ij}\hat{\beta}\cdot\hat{r}_{ij} - \delta_{\alpha\beta} \right] \right\} [\sigma_{i\alpha}\tau_{i\gamma}][\sigma_{j\beta}\tau_{j\gamma}] \\ & + \sum_{i<j,\gamma} v^\tau(r_{ij})\tau_{i\gamma}\tau_{j\gamma} \end{aligned} \tag{9.177}$$

where α and β refer to the x, y, and z components and $\hat{\alpha}\hat{\beta}$ are the corresponding unit vectors. We work in a position basis. The potential is quadratic in the 15A spin-isospin operators $\sigma_{i\alpha}$, $\tau_{i\gamma}$, $\sigma_{i\alpha}\tau_{i\gamma}$. Since each spin-isospin operator can rotate the corresponding spin-isospinor the natural basis is the overcomplete basis of the outer product of these spin-isospinors—one for each particle. A walker consists of an overall weight factor, and x, y, and z coordinates and four complex numbers for the components of $|p\uparrow\rangle$, $|p\downarrow\rangle$, $|n\uparrow\rangle$, $|n\downarrow\rangle$ for each of the A particles.

9.4.7.1 The v_6' Hamiltonian as a Sum of Operator Squares

We now follow Sect. 9.4.5 and define matrices

$$\begin{aligned} C^{\sigma}_{i\alpha,j\beta} &= v^{\sigma}(r_{ij})\delta_{\alpha\beta} + v^{t}(r_{ij})\left[3\hat{\alpha}\cdot\hat{r}_{ij}\hat{\beta}\cdot\hat{r}_{ij} - \delta_{\alpha\beta}\right] \\ C^{\sigma\tau}_{i\alpha,j\beta} &= v^{\sigma\tau}(r_{ij})\delta_{\alpha\beta} + v^{t\tau}(r_{ij})\left[3\hat{\alpha}\cdot\hat{r}_{ij}\hat{\beta}\cdot\hat{r}_{ij} - \delta_{\alpha\beta}\right] \\ C^{\tau}_{i,j} &= v^{\tau}(r_{ij}) \end{aligned} \tag{9.178}$$

which have zero matrix elements when $i = j$. Their eigenvalues and normalized eigenvectors are defined as

$$\begin{aligned} \sum_{j\beta} C^{\sigma}_{i\alpha,j\beta}\psi^{\sigma\ (n)}_{j\beta} &= \lambda^{\sigma}_{n}\psi^{\sigma\ (n)}_{i\alpha} \\ \sum_{j\beta} C^{\sigma\tau}_{i\alpha,j\beta}\psi^{\sigma\ (n)}_{j\beta} &= \lambda^{\sigma\tau}_{n}\psi^{\sigma\tau\ (n)}_{i\alpha} \\ \sum_{j} C^{\tau}_{i,j}\psi^{\tau\ (n)}_{j} &= \lambda^{\tau}_{n}\psi^{\tau\ (n)}_{i} \end{aligned} \tag{9.179}$$

with operator combinations

$$\begin{aligned} O^{\sigma}_{n} &= \sum_{i\alpha}\psi^{\sigma\ (n)}_{i\alpha}\sigma_{i\alpha} \\ O^{\sigma\tau}_{n\beta} &= \sum_{i\alpha}\psi^{\sigma\tau\ (n)}_{i\alpha}\sigma_{i\alpha}\tau_{i\beta} \\ O^{\tau}_{n\alpha} &= \sum_{i}\psi^{\tau\ (n)}_{i}\tau_{i\alpha} \end{aligned} \tag{9.180}$$

The Hamiltonian becomes

$$\begin{aligned} H = &\sum_{i=1}^{A}\sum_{\alpha=1}^{3}\frac{p^2_{i\alpha}}{2m} + \sum_{i<j} v^{c}(r_{ij}) + \frac{1}{2}\sum_{n=1}^{3A}\lambda^{\sigma}_{n}(O^{\sigma}_{n})^2 \\ &+\frac{1}{2}\sum_{n=1}^{A}\sum_{\alpha=1}^{3}\lambda^{\tau}_{n}(O^{\tau}_{n\alpha})^2 + \frac{1}{2}\sum_{n=1}^{3A}\sum_{\alpha=1}^{3}\lambda^{\sigma\tau}_{n}(O^{\sigma\tau}_{n\alpha})^2 \end{aligned} \tag{9.181}$$

This is, of course, identical to the original Hamiltonian given in Eq. (9.177), but now it is in a form that makes the propagator easy to sample using auxiliary fields.

9.4.7.2 Complex Auxiliary Fields

In realistic nuclear physics problems, the fermion sign problem necessarily becomes a phase problem since conservation of angular momentum requires that flipping a spin changes the orbital angular momentum, which induces an angular phase to the wave function. Various fixed-phase approximations can be used. The Hubbard-Stratonovich transformation integrates the auxiliary field over all real values with a gaussian weight. With importance sampling, the gaussian for x_n is shifted by $i\sqrt{\lambda_n \Delta t}\langle O_n\rangle$ as shown in the last line of Eq. (9.147). As Zhang and Krakauer [17] showed for electronic structure problems, it is equally valid to integrate the auxiliary field over any shifted contour, and by shifting the contour so that x_n becomes complex and takes on the values $x_n = z + i\sqrt{\lambda_n \Delta t}\langle O_n\rangle$, $-\infty < z < \infty$. Integrating over these values does not change the result. However, now this factor is real. We implement the fixed phase approximation by taking the real part of $\langle H\rangle$.

Note that this method cannot be used for the momentum operator. This is because the operator $e^{-\frac{i}{\hbar}p_{j\alpha}a}$ is not bounded if a has an imaginary part. We therefore implement the kinetic energy terms exactly as in the central potential fixed-phase approximation.

There are of course other possible approximations that can be used. The auxiliary fields can be kept real. We find that the approximation is more accurate with Zhang-Krakauer prescription for auxiliary fields for the spin operators.

9.4.7.3 The v_6' Algorithm

We can now give the complete algorithm used for the v_6' potential.

1. We begin with a set of walkers $|R_i S_i\rangle$ which we sample from our trial function magnitude squared, $|\langle RS|\Psi_T\rangle|^2$, with Metropolis Monte Carlo. The walkers consist of the $3A$ coordinates of the A particles, and A 4-component normalized spinors.
2. For each walker in turn we calculate the C^σ, C^τ and $C^{\sigma\tau}$ matrices, their eigenvalues, and their eigenvectors.
3. From the trial function and spinor values we evaluate $\langle\sigma_{j\alpha}\rangle$, $\langle\sigma_{j\alpha}\tau_{j\beta}\rangle$, $\langle\tau_{j\alpha}\rangle$, $\langle p_{j\alpha}\rangle$, and $\langle H\rangle$.
4. We sample the complex values for the spin-isospin auxiliary fields

$$x_n = \chi_n + i\sqrt{\lambda_n \Delta t}\langle O_n\rangle \tag{9.182}$$

and transform our walker spinors using

$$|RS'\rangle = e^{-i\sum_{n=1}^{N_O} x_n\sqrt{\lambda\Delta t}O_n}|RS\rangle \tag{9.183}$$

and normalize the spinors

5. We sample the new positions from

$$r'_{j\alpha} = r_{j\alpha} + \chi_{j\alpha}\sqrt{\frac{\hbar^2\Delta t}{m}} + \frac{\hbar^2\Delta t}{m}\mathrm{Re}\frac{\partial_{j\alpha}\langle\Psi_T|RS\rangle}{\langle\Psi_T|RS\rangle} \tag{9.184}$$

6. The weight of the new walker is given by $\mathrm{W} = e^{-[\mathrm{Re}\langle H\rangle - E_T]\Delta t}$
7. We branch on the walker weight, taking the number of new walkers to be the integer part of W plus a uniform random value on $(0, 1)$. If the weight W is negative, we discard the walker.

9.4.8 Isospin-Independent Spin-Orbit Interaction

Without isospin exchange, the spin orbit term for particles j and k is

$$\frac{1}{4\hbar}v_{LS}(r_{jk})[(\boldsymbol{r}_j - \boldsymbol{r}_k)\times(\boldsymbol{p}_j - \boldsymbol{p}_k)]\cdot(\boldsymbol{\sigma}_j + \boldsymbol{\sigma}_k) \tag{9.185}$$

We can write the kinetic energy plus spin-orbit interaction Hamiltonian as

$$\begin{aligned}
&\sum_{j\alpha}\frac{p_{j\alpha}^2}{2m} + \frac{1}{4\hbar}\sum_{j<k}v_{LS}(r_{jk})[(\boldsymbol{r}_j - \boldsymbol{r}_k)\times(\boldsymbol{p}_j - \boldsymbol{p}_k)]\cdot(\boldsymbol{\sigma}_j + \boldsymbol{\sigma}_k) \\
&\quad = \sum_{j\alpha}\frac{(p_{j\alpha} + \frac{m}{4\hbar}\sum_{k\neq j}v_{LS}(r_{jk})[(\boldsymbol{\sigma}_j + \boldsymbol{\sigma}_k)\times(\boldsymbol{r}_j - \boldsymbol{r}_k)]_\alpha)^2}{2m} + V_{\mathrm{Counter}} \\
&\quad V_{\mathrm{Counter}} = -\frac{1}{2m}\sum_{j\alpha}\left[\frac{m}{4\hbar}\sum_{k\neq j}v_{LS}(r_{jk})[(\boldsymbol{\sigma}_j + \boldsymbol{\sigma}_k)\times(\boldsymbol{r}_j - \boldsymbol{r}_k)]_\alpha\right]^2
\end{aligned} \tag{9.186}$$

where the counter terms subtract off the unwanted interaction from completing the square. The counter terms do not depend on $\boldsymbol{p}_j$, so they can be included with the rest of the local potential, and will contribute to the drift and the local energy for that part. However, we will see that the local energy part is canceled below (that is the final weight will be just the correct total local energy which does not include the counter terms).

Using the Hubbard-Stratonovich break up with importance sampling, we have $\lambda_{j\alpha} = m^{-1}$, and

$$i\sqrt{\lambda_{j\alpha}\Delta t}\langle O_{j\alpha}\rangle = -\sqrt{\frac{\hbar^2\Delta t}{m}}\frac{\partial_{j\alpha}\langle\Psi_T|RS\rangle}{\langle\Psi_T|RS\rangle} + i\sqrt{\frac{m\Delta t}{16\hbar^2}}\sum_{k\neq j}[(\langle\boldsymbol{\sigma}_j\rangle + \langle\boldsymbol{\sigma}_k\rangle)\times\boldsymbol{r}_{jk}]_\alpha v_{LS}(r_{jk})\,. \tag{9.187}$$

The sampled value of $x_{j\alpha}$ will be

$$x_{j\alpha} = \chi_{j\alpha} + \sqrt{\frac{\hbar^2 \Delta t}{m}} \frac{\partial_{j\alpha} \langle \Psi_T | RS \rangle}{\langle \Psi_T | RS \rangle} - i \sqrt{\frac{m \Delta t}{16 \hbar^2}} \sum_{k \neq j} [(\langle \boldsymbol{\sigma}_j \rangle + \langle \boldsymbol{\sigma}_k \rangle) \times \boldsymbol{r}_{jk}]_\alpha v_{LS}(r_{jk}) \,. \tag{9.188}$$

where our fixed-phase like approximation will modify this to keep the translation real, so that

$$x_{j\alpha} = \chi_{j\alpha} + \mathrm{Re} \left\{ \sqrt{\frac{\hbar^2 \Delta t}{m}} \frac{\partial_{j\alpha} \langle \Psi_T | RS \rangle}{\langle \Psi_T | RS \rangle} - i \sqrt{\frac{m \Delta t}{16 \hbar^2}} \sum_{k \neq j} [(\langle \boldsymbol{\sigma}_j \rangle + \langle \boldsymbol{\sigma}_k \rangle) \times \boldsymbol{r}_{jk}]_\alpha v_{LS}(r_{jk}) \right\} . \tag{9.189}$$

The walker propagator is

$$|R'S'\rangle = e^{-\frac{i}{\hbar} \sum_{j\alpha} x_{j\alpha} \sqrt{\frac{\hbar^2 \Delta t}{m}} p_{j\alpha}} e^{i \sum_{j\alpha} x_{j\alpha} \sqrt{\frac{m \Delta t}{16 \hbar^2}} \sum_{k \neq j} [(\boldsymbol{\sigma}_j + \boldsymbol{\sigma}_k) \times \boldsymbol{r}_{jk}]_\alpha v_{LS}(r_{jk})} |RS\rangle \tag{9.190}$$

The local energy term for the spin orbit will contain the kinetic energy, the spin orbit, and the negative of the counter terms. Therefore, the counter term contribution cancels in the weight, and the final weight is the local energy.

9.5 GFMC with Full Spin-Isospin Summation

As mentioned above, current high quality trial wave functions for the coordinate space nuclear Hamiltonians require the same computational complexity to calculate either one or all of the spin-isospin amplitudes at a specified position for the particles. Very roughly for A nucleons, each of which can be a proton or neutron with spin up or down, the number of spin-isospin amplitudes is 4^A. Symmetries can lower this factor but not change its overall exponential character (see Table 9.1).

Typically these calculations are done in either a good charge or good isospin basis. In a good charge basis, with A nucleons, with Z protons, the number of combinations of protons and neutrons is $\frac{A!}{Z!(A-Z)!}$, while the tensor force can flip any of the spins so there are 2^A spin states. The total number of allowed spin-isospin states is the product of these factors. Sometimes the initial calculations are done with a Hamiltonian that conserves isospin and the charge symmetry breaking components are added perturbatively. In this case the number of states can be further reduced. Since $T_z = \frac{2Z-A}{2}$, the number of isospin states T states for a given $T_z \leq T$ is given by the difference in the number of charge states with $T_z = T$ and $T_z = T + 1$, which is $\frac{A!}{(\frac{A}{2}-T)!(\frac{A}{2}+T)!} \frac{2T+1}{\frac{A}{2}+T+1}$.

Table 9.1 The number of spin-isospin amplitudes for the ground states of some representative nuclei

Nucleus	Spin	Charge states	Total	Isospin/T reversal
^{4}He	16	6	96	16
^{8}Be	256	70	17,920	1792
^{12}C	4096	924	3,784,704	270,336
^{16}O	65,536	12,870	8.4×10^8	4.7×10^7

Time-reversal invariant states have a further factor of 2 reduction, since in that case, the time reversal operator

$$\mathscr{T} = \left[\prod_{i=1}^{A} \sigma_{ix}\sigma_{iz}\right] K \tag{9.191}$$

relates the amplitudes of the states given by flipping all the spins. Here K is the complex conjugating operator.

To see how this works, we can look at a straightforward generalization of a Jastrow-Slater trial state,

$$|\Psi_T\rangle = \left[\mathscr{S}\prod_{i<j}\sum_{p} f_{ij}^{(p)} O_{ij}^{(p)}\right]|\Phi\rangle \tag{9.192}$$

where $|\Phi\rangle$ is a model state, typically one or a small linear combination of antisymmetric products of single particle orbitals. The p sum is over the same sort of operators as those in the potential (usually operators with gradients are either omitted or kept only at lowest order), with the Jastrow correlations $f_{ij}^{(p)}$ depending only on the spatial operator $|\boldsymbol{r}_i - \boldsymbol{r}_j|$, while the $O_{ij}^{(p)}$ contain spin-isospin operators and the unit vector operators $\frac{\boldsymbol{r}_i - \boldsymbol{r}_j}{|\boldsymbol{r}_i - \boldsymbol{r}_j|}$. The $\mathscr{S}$ is a symmetrizing operator applied to the Jastrow product, since the operators in general do not commute, so that the trial function is properly antisymmetric under interchange.

To form a trial wave function we take the inner product with $\langle RS|$ to obtain $\Psi_T(R,S) = \langle RS|\Psi_T\rangle$. The spatial operators operating to the left on their eigenstate $\langle RS|$ are replaced by their eigenvalues. This leaves just the spin-isospin matrix elements. The model state is evaluated for all possible spin-isospin states as enumerated above, $\langle RS'|\Phi\rangle$. In our spin-isospin basis, each of the operators $\langle S''|O_{ij}^{(p)}|S'\rangle$ is a sparse matrix which can either be tabulated or easily calculated as needed. For example, in the charge basis, acting on a single basis state, the interaction can change the spins of a pair to any of the four values. If the particles of the pair are a neutron and a proton, they can be interchanged. This shows that there are at most either 4 or 8 nonzero entries per row or column of the matrix representation. The construction of the Jastrow product is obtained by these repeated sparse-matrix multiplications.

The symmetrizing operator has the factorial of the number of pairs terms. It would be prohibitive to calculate explicitly. However, the commutator terms are small, so the sum over orders of the operators is done by Monte Carlo sampling.

Since much of the computational time is spent in evaluating the trial wave functions, wave functions that include more complicated correlations as well as alpha particle clustering are often included. The simplest wave function above is adequate for the alpha particle.

A GFMC calculation uses walkers given by positions for all the particles, and amplitudes for each of the possible spin-isospin states in the basis.

In the simplest GFMC implementation, the so-called primitive approximation can be used. Here the propagator is

$$\left[\prod_{i<j} e^{-\frac{1}{2}\tau \sum_p v_{ij}^{(p)}}\right] e^{-\tau \sum_i \frac{p_i^2}{2m}} \left[\prod_{i<j} e^{-\frac{1}{2}\tau \sum_p v_{ij}^{(p)}}\right] \tag{9.193}$$

where the opposite order of the pairs is taken in the two products to minimize the time-step errors. The exponentials of the pair operators can be written as a linear combination of pair operators, and these are then operated on the walker states giving new amplitudes. The kinetic energy term is implemented by sampling a gaussian to give new positions.

9.6 General Projection Algorithms in Fock Space and Non-local Interactions

In recent years, a number of projection algorithms working in a discrete Fock space (configuration space) rather than in coordinate space have been proposed [18–23]. While more similar to more standard many-body techniques like Coupled Cluster (CC) and Many Body Perturbation Theory already covered in previous chapters the adoption of statistical techniques in a configuration space has some advantage. First of all Monte Carlo techniques can be implemented with a much milder scaling with the system size enabling the possibility with a much larger number of basis states that build up the total Hilbert space. Contrary to e.g. CC theory we can ensure that the final QMC estimate for the ground-state energy would be an upper bound of the true eigenvalue, thus providing useful benchmark results. Also, working on a finite many-body space allows practical calculations with non-local interactions, like those developed within the Chiral Effective Field Theory approach to nuclear forces, in a far more controllable way than not with the continuous coordinate-space formulation exposed so far (as was done in [24]). Finally, another great advantage of performing the Monte Carlo on a discrete Hilbert space is the possibility to devise an efficient strategy to reduce the impact of the sign-problem by using cancellation techniques [18–20] in an analogous fashion to what was sketched at the end of Sect. 9.3.5. Unfortunately we won't have space here to cover these aspects.

9.6.1 Fock Space Formulation of Diffusion Monte Carlo

To set the stage let us take a finite set $\mathscr{S}$ of single-particle (sp) states of size $\mathscr{N}_s$ and consider a general second-quantized fermionic Hamiltonian including two and possibly many-body interactions

$$H = \sum_{\alpha\in\mathscr{S}} \epsilon_\alpha a^\dagger_\alpha a_\alpha + \sum_{\alpha\beta\gamma\delta\in\mathscr{S}} V_{\alpha\beta\gamma\delta} a^\dagger_\alpha a^\dagger_\beta a_\delta a_\gamma + \dots \,. \tag{9.194}$$

In this expression Greek letter indices indicates sp states (ie. α is a collective label for all sp quantum numbers), the operator $a^\dagger_\alpha$ (a_α) creates (destroys) a particle in the sp state α and the $V_{\alpha\beta\gamma\delta}$ are general (anti-symmetrized) two-body interaction matrix elements:

$$V_{\alpha\beta\gamma\delta} = \langle\alpha\beta|\hat{V}|\gamma\delta\rangle - \langle\alpha\beta|\hat{V}|\gamma\delta\rangle. \tag{9.195}$$

For an N-fermion system the resulting Fock space would be spanned by the full set of N-particle Slater determinants that can be generated using the sp orbitals $\alpha \in \mathscr{S}$. We will denote these Slater-determinants in the occupation number basis by $|\mathbf{n}\rangle$, where $\mathbf{n} \equiv \{n_\alpha\}$ and $n_\alpha = 0, 1$ are occupation number of the single-particle orbital α satisfying $\sum_\alpha n_\alpha = N$. For example in a system composed by two identical fermions and with $\mathscr{N}_s = 4$ available sp states we will write

$$|0110\rangle \equiv a^\dagger_3 a^\dagger_2 |0\rangle \tag{9.196}$$

where $|0\rangle$ is our vacuum state (that can be conveniently set to the Hartree-Fock ground state Φ_{HF}), while $a^\dagger_2$ and $a^\dagger_3$ creates a particle in sp state 2 and 3 respectively.

We can now use these states as a complete basis in our many-body Hilbert space and express a generic state in it as

$$|\Psi\rangle = \sum_{\mathbf{n}} \langle\mathbf{n}|\Psi\rangle|\mathbf{n}\rangle \equiv \sum_{\mathbf{n}} \Psi(\mathbf{n})|\mathbf{n}\rangle \tag{9.197}$$

where the sum is over all possible basis vectors that one can obtain from the $\mathscr{N}_S$ single-particle orbitals.

It is important to notice at this point that no assumption is made on the locality of the interaction, which translates into restrictions on the structure of the tensor $V_{\alpha\beta\gamma\delta}$. This shows already that possible non-local interactions can be cleanly incorporated in the formalism.

As was already introduced in Sect. 9.3.1, the core idea behind a Diffusion Monte Carlo algorithm is to extract ground-state informations on the system by evolving in imaginary-time an initial guess for the lowest eigenstate of the hamiltonian H:

$$\Psi_{\tau+\Delta\tau}(\mathbf{m}) = \sum_{\mathbf{n}} \langle\mathbf{m}|P|\mathbf{n}\rangle \Psi_\tau(\mathbf{n}). \tag{9.198}$$

with a suitable projection operator P [cf. Eq. (9.125) and discussion above]. In order to illustrate how the evolution in (9.198) can be implemented in a stochastic way, it will be useful first to express the matrix elements of P as follows

$$\langle \mathbf{m}|P|\mathbf{n}\rangle = p(\mathbf{m},\mathbf{n})g(\mathbf{n}) \tag{9.199}$$

with

$$g(\mathbf{n}) = \sum_{\mathbf{m}} \langle \mathbf{m}|P|\mathbf{n}\rangle \tag{9.200}$$

and

$$p(\mathbf{m},\mathbf{n}) = \frac{\langle \mathbf{m}|P|\mathbf{n}\rangle}{\sum_{\mathbf{m}} \langle \mathbf{m}|P|\mathbf{n}\rangle}. \tag{9.201}$$

At this point, provided the matrix elements $\langle \mathbf{m}|P|\mathbf{n}\rangle \geq 0$ we can interpret $p(\mathbf{m},\mathbf{n})$ for fixed $\mathbf{n}$ as (normalized) probability distribution for the states $\mathbf{m}$ and $g(\mathbf{n})$ as a weight factor. This is analogous to what was done in Sect. 9.3.2 for the conventional coordinate-space formulations where now p takes the place of the gaussian Eq. (9.87) while g replaces the weight Eq. (9.88).

Imagine now that at a given imaginary-time τ the wave-function Ψ_τ is non-negative in configuration space

$$\Psi_\tau(\mathbf{n}) \geq 0 \forall \mathbf{n}, \tag{9.202}$$

then we can represent it as an ensemble of configurations. Due to the non-negativity of the matrix elements of P, we also have that the evolution described in (9.198) preserves the signs

$$\Psi_{\tau+\Delta\tau}(\mathbf{m}) \geq 0 \ \forall \mathbf{m}. \tag{9.203}$$

This suggests the following procedure for the stochastic imaginary-time evolution:

1. walker starts at configuration $\mathbf{n}$ with weight $w(\mathbf{n})$
2. a new configuration $\mathbf{m}$ is chosen from the probability distribution $p(\mathbf{m},\mathbf{n})$
3. the walker's weight gets rescaled as $w(\mathbf{n}) \to w(\mathbf{m}) = w(\mathbf{n})g(\mathbf{n})$
4. repeat from 1.

In order to improve efficiency one can include a *branching* step where the new configuration in $\mathbf{m}$ is replicated according to its weight as explained in Sect. 9.3.2.

Expectation values of observables can then be estimated as usual [cf. Eq. (9.130)] with the mixed estimator

$$\langle O\rangle_{mixed} = \frac{\langle \Psi_T|O|\Psi(\tau)\rangle}{\langle \Psi_T|\Psi(\tau)\rangle} = \frac{\sum_l^{N_w} w(\mathbf{m}_l)\langle \Psi_T|O|\mathbf{m}_l\rangle}{\sum_l^{N_w} w(\mathbf{m}_l)\Psi_T(\mathbf{m}_l)} \tag{9.204}$$

where Ψ_T is a trial state and the sums run over the walker population of size N_w.

In practice we have to choose some form for the evolution operator that appears in (9.198), a common choice in discrete spaces is on operator very similar to the one already encountered in the discussion of Sect. 9.2.5.2:

$$\begin{aligned}\langle \mathbf{m}|P|\mathbf{n}\rangle &= \langle \mathbf{m}|1 - \Delta\tau\,(H - E_T)|\mathbf{n}\rangle \\ &= \delta_{\mathbf{m},\mathbf{n}} - \Delta\tau\langle \mathbf{m}|H - E_T|\mathbf{n}\rangle\end{aligned} \tag{9.205}$$

where E_T is an energy shift used in the simulation to preserve the norm of the solution (the constant E_0 introduced in Sect. 9.3.2). Convergence to the ground-state by repeated application of the projector P to the initial state $|\Psi_0\rangle$

$$|\Psi_{gs}\rangle = \lim_{M\to\infty} P^M|\Psi_0\rangle \tag{9.206}$$

is guaranteed provided that the eigenvalues of P lie between -1 and 1 in order to ensure the diagonal part remains positive definite. This requirement translates into a condition on the imaginary-time step $\Delta\tau$ which has to satisfy the bound

$$\Delta\tau < 2/(E_{max} - E_{min}) \tag{9.207}$$

where E_{max} and E_{min} are respectively the maximum and minimum eigenvalue of H in our finite basis. This upper bound becomes tighter and tighter as we increase the number of particle N and/or the number of sp-states $\mathcal{N}_s$. As a consequence the number M of iterations needed for convergence to the ground state increases dramatically. A way to deal with this problem is to employ a different algorithm proposed in [25] (see also [26, 27]) that allows us to sample directly from the exponential propagator

$$\langle \mathbf{m}|P|\mathbf{n}\rangle = \langle \mathbf{m}|e^{-\Delta\tau(H-E_T)}|\mathbf{n}\rangle \tag{9.208}$$

in analogy to Eq. (9.74), but now without any limitation on the choice of the imaginary time step $\Delta\tau$ that can be chosen arbitrarily large. We leave the discussion of its details in Sect. 9.6.4.

In our discussion so far we have assumed that the matrix elements on the projector that defines $p(\mathbf{m},\mathbf{n})$ in Eq. (9.201) are actually positive definite. Under general circumstances however this is not the case. This clearly prevents the interpretation of $p(\mathbf{m},\mathbf{n})$ as a probability distribution invalidating the naive approach

employed above. In order to circumvent the problem we can use the same idea behind the fixed node (phase) approximation introduced in Sect. 9.3.5.1.

Before continuing it is worth to mention that in principle one can still produce a stochastic evolution by absorbing the signs into the weight factor $g(\mathbf{n})$ while sampling off-diagonal moves using $|\langle \mathbf{m}|P|\mathbf{n}\rangle|$. However as briefly explained in Sect. 9.3.5.1 this is accompanied by an exponential decay of the signal to noise ratio as a function of the total projection time $\tau = M\Delta\tau$. Recently it was shown that by employing an annihilation step in the evolution this problem can be substantially alleviated [18, 20, 21]. At the end however these algorithms have still an exponential scaling with system-size, though with a reduced exponent.

9.6.2 *Importance Sampling and Fixed-Phase Approximation*

As we just mentioned, we can deal with the sign-problem in a way which is similar to standard coordinate-space QMC: we will use an initial ansatz Φ_T for the ground-state wave-function and use that to constrain the random walk in a region of the many-body Hilbert space where

$$\langle \mathbf{m}|P|\mathbf{n}\rangle \geq 0 \tag{9.209}$$

is satisfied. In order for this scheme to be practical one needs a systematic way for reducing the bias coming from this approximation, e.g. we want the bias to go to zero as the ansatz Φ_T goes towards the ground-state Ψ_{gs}. That's exactly what is done in coordinate-space fixed-node(fixed-phase) QMC simulations presented in the previous sections.

In this derivation we will follow the work in [26, 27] and generalize it to the case of complex-hermitian hamiltonians usually found in nuclear theory. Similarly to what was done in Sect. 9.4.4 the imaginary part of the solution is constrained to be the same of that of the trial wave-function

$$\Re[\Psi^*(\mathbf{n})\Phi_T(\mathbf{n})] = 0 \tag{9.210}$$

for every distribution $\Psi(\mathbf{n})$ sampled in the random walk. In this expression $\Re$ stands for the real part and * is complex-conjugation.

We start by defining for any configurations $\mathbf{n}$ and $\mathbf{m}$ for which $|\Phi_T(\mathbf{n})| \neq 0$ the following quantity:

$$\begin{aligned} \mathfrak{s}_{\mathbf{mn}} &= \text{sign}\ \Re\left[\Phi_T^*(\mathbf{m})H_{\mathbf{mn}}\Phi_T^*(\mathbf{n})^{-1}\right] \\ &= \text{sign}\ \frac{\Re\left[\Phi_T^*(\mathbf{m})H_{\mathbf{mn}}\Phi_T(\mathbf{n})\right]}{|\Phi_T(\mathbf{n})|^2} = \mathfrak{s}_{\mathbf{nm}}. \end{aligned} \tag{9.211}$$

Now define a one-parameter family of Hamiltonians $\mathscr{H}_\gamma$ defined over configurations $\mathbf{n}$ (again such that $|\Phi_T(\mathbf{n})| \neq 0$) with off-diagonal matrix elements given by

$$\langle \mathbf{m}|\mathscr{H}_\gamma|\mathbf{n}\rangle = \begin{cases} -\gamma\langle \mathbf{m}|H|\mathbf{n}\rangle & \mathfrak{s}(\mathbf{m},\mathbf{n}) > 0 \\ \langle \mathbf{m}|H|\mathbf{n}\rangle & \text{otherwise} \end{cases}, \tag{9.212}$$

while the diagonal terms are

$$\begin{aligned} \langle \mathbf{n}|\mathscr{H}_\gamma|\mathbf{n}\rangle &= \langle \mathbf{n}|H|\mathbf{n}\rangle + (1+\gamma) \sum_{\substack{\mathbf{m}\neq\mathbf{n} \\ \mathfrak{s}(\mathbf{m},\mathbf{n})>0}} \langle \mathbf{m}|H|\mathbf{n}\rangle \\ &= \langle \mathbf{n}|H|\mathbf{n}\rangle + \sum_{\mathbf{m}} h_{\mathbf{mn}}. \end{aligned} \tag{9.213}$$

In the limit where $\gamma \to -1$ we clearly recover the original Hamiltonian:

$$\mathscr{H}_{\gamma=-1} \equiv H. \tag{9.214}$$

We proceed to define a corresponding family of propagators $\mathscr{P}_\gamma$ for configurations $\mathbf{n}$ with $|\Phi_T(\mathbf{n})| \neq 0$ by

$$\langle \mathbf{m}|\mathscr{P}_\gamma|\mathbf{n}\rangle = \delta_{\mathbf{m},\mathbf{n}} - \Delta\tau \frac{\Re\left[\Phi_T^*(\mathbf{m})\langle \mathbf{m}|\mathscr{H}_\gamma - E_T|\mathbf{n}\rangle\Phi_T(\mathbf{n})\right]}{|\Phi_T(\mathbf{n})|^2}. \tag{9.215}$$

It is clear now that for any $\gamma \geq 0$ we have

$$\langle \mathbf{m}|\mathscr{P}_\gamma|\mathbf{n}\rangle \geq 0 \tag{9.216}$$

and so the propagator $\mathscr{P}$ is, by construction, free from the sign-problem. Performing the corresponding random-walk allows us to filter the state

$$\Phi_T(\mathbf{n})\phi_\gamma^0(\mathbf{n}), \tag{9.217}$$

where now $\phi_\gamma^0(\mathbf{n})$ is the ground-state of the hamiltonian $\mathscr{H}_\gamma$. The ground-state energy E_γ obtained following this procedure can be proved (the proof is left to the Appendix) to be a strict upper bound for the true ground-state energy E_0 of the true hamiltonian H. Moreover, this upper bound is tighter than the variational upper-bound provided by

$$E_T = \frac{\langle \Phi_T|H|\Phi_T\rangle}{\langle \Phi_T|\Phi_T\rangle} \geq E_0. \tag{9.218}$$

As you can show in Problem 9.5 any linear extrapolation of E_γ from any two values $\gamma \geq 0$ to $\gamma = -1$ (which would correspond to the original hamiltonian)

also provides an upper-bound on E_{gs} that is tighter than the individual E_γ's. A good compromise between the tightness of the upper-bound and the statistical noise in the extrapolation is to choose two values of γ: 0 and 1, thus giving the following energy estimator:

$$E_{extr} = 2E_{\gamma=0} - E_{\gamma=1} \tag{9.219}$$

To ensure the success of the proposed method a good choice for the importance function $|\Phi_T\rangle$ is critical.

9.6.3 Trial Wave-Functions from Coupled Cluster Ansatz

As have been pointed out before, a crucial role is played by the importance function Φ_T used to impose the constraint. This is especially true if we want to estimate expectation values of operators other than the energy (cf. discussion in Sect. 9.4.2.1).

Fundamental prerequisites for a viable importance function are

1. enough flexibility to be able to account for the relevant correlations in the system
2. availability of an efficient way to evaluate its overlap with states explored during the random walk

Within a Fock space formulation, an excellent choice for Φ_T that satisfy the first requirement is given by the wave function generated in a Coupled Cluster calculation. Starting from a reference state, which usually is the Hartree-Fock solution of the problem, CC theory allows to include dynamical correlations into a new state as

$$|\Psi_{CC}\rangle = e^{\hat{T}}|\Phi_{HF}\rangle. \tag{9.220}$$

In the above equation, correlations are introduced trough the excitation operator $\hat{T}$ which in CC theory is hierarchically divided as

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots \tag{9.221}$$

counting the number of creation/annihilation operators that compose them. The first two terms are:

$$\hat{T}_1 = \sum_{\alpha,\beta\in\mathscr{S}} t_\alpha^\beta a_\beta^\dagger a_\alpha \qquad \hat{T}_2 = \frac{1}{4}\sum_{\alpha,\beta,\gamma,\delta\in\mathscr{S}} t_{\alpha\beta}^{\gamma\delta} a_\gamma^\dagger a_\delta^\dagger a_\alpha a_\beta \quad \cdots \tag{9.222}$$

The final state $|\Psi_{CC}\rangle$ will then be uniquely identified by the coefficients t_α^β and $t_{\alpha\beta}^{\gamma\delta}$ corresponding to single and double particle-hole excitations respectively. The

exponentiated form of the CC wave-functions enables to effectively include some correlations up to the maximum N-particle N-hole in a relatively compact way.

But is the wave-function in Eq. (9.220) also quick to evaluate? In order to simplify the discussion we will focus here on the case of a homogeneous system that can be described dropping the one-particle-one-hole excitation operator $\hat{T}_1$ in the expansion (which do not contribute due to translational invariance).[5] In this situation the lowest order of CC theory is the Coupled Cluster Doubles (CCD) approximation.

To set the notation, we will express a generic Slater-Determinant state describing an M-particle-M-hole state as

$$|\mathbf{m}\rangle = a^\dagger_{p_1} \ldots, a^\dagger_{p_M} a_{h_1} \ldots, a_{h_M} |\Phi_{HF}\rangle \equiv |\Phi^{p_1,\ldots,p_M}_{h_1,\ldots,h_M}\rangle. \tag{9.223}$$

The required amplitude can then be expressed as a superposition of $M-2$ particle/hole states that can be generated from $\mathbf{m}$. Eventually (the proof is tedious but straightforward) one obtains:

$$\langle \mathbf{m}|\Psi_{CC}\rangle = \sum_{\gamma=2}^{M}\sum_{\mu<\nu}^{M}(-1)^{\gamma+\mu+\nu} t^{p_\mu p_\nu}_{h_1 h_\gamma} \Psi^{M-2}_{CC} \begin{pmatrix} p_1,p_2,\ldots,p_{\mu-1},p_{\mu+1},\ldots,p_{\nu-1},p_{\nu+1},\ldots,p_M \\ h_2\ldots,h_{\gamma-1},h_{\gamma+1},\ldots,h_M \end{pmatrix} \tag{9.224}$$

assuming $p_1 < p_2 < \cdots < p_M$ and $h_1 < h_2 < \cdots < h_M$. The normalization is fixed in such a way that $\langle \Phi_{HF}|\Psi_{CC}\rangle = 1$.

One way to implement Eq. (9.224) is for instance trough a recursive function that takes as input some K-particle-K-hole state and returns 1.0 for $K = 0$, the correct amplitude t^{ab}_{ij} for $K = 2$ and for $K > 2$ calls itself again removing two particle and two hole states. Clearly this approach becomes slow when states with large values of K are sampled often during the random walk. Just to give an idea, for calculations of pure neutron matter with soft Chiral EFT interactions we have $K \leq 6$ at densities $\rho \approx 0.08\,\mathrm{fm}^{-3}$ (cf. discussion in [28]) and the calculation can be made very efficient.

Within CC theory the coefficients $t^{\gamma\delta}_{\alpha\beta}$ appearing in the equations above are to be obtained as the self-consistent solutions of the following non-linear equation:

$$\langle \Phi^{\gamma\delta}_{\alpha\beta}|\hat{H}\left(1+\hat{T}_2+\frac{1}{2}\hat{T}_2^2\right)|\Phi_{HF}\rangle = \left(\frac{1}{4}\sum_{\alpha,\beta,\gamma,\delta\in\mathscr{S}}\langle\alpha\beta||\gamma\delta\rangle t^{\gamma\delta}_{\alpha\beta}\right) t^{\gamma\delta}_{\alpha\beta} \tag{9.225}$$

where $\langle\alpha\beta||\gamma\delta\rangle$ are the anti-symmetrized two-body matrix elements of the interaction defined in Eq. (9.195).

Solving Eq. (9.225) is in general a very expensive computational problem and within the fixed-node approach all that matters are the signs in Eq. (9.211). It could then be possible to find cheaper approximate ways to determine the doubles coeffi-

[5] Extension to singlets (p-h states) and triplets (3p-3h states) is simple.

cients $t_{\alpha\beta}^{\gamma\delta}$ while still preserving a good quality in the fixed-node approximation. A quite precise and very cheap approximation that have been used successfully is to obtain the coefficients within second order Moeller–Plesset perturbation theory:

$$t_{\alpha\beta}^{\gamma\delta} = \frac{\langle \alpha\beta||\gamma\delta\rangle}{\eta_\alpha + \eta_\beta - \eta_\gamma - \eta_\delta} \quad \text{with} \quad \eta_i = \epsilon_i + \sum_{k\in\mathscr{S}} \langle ik||ik\rangle \tag{9.226}$$

and ϵ_i are the single particle energies appearing in the one body part of the Hamiltonian Eq. (9.194). This is equivalent to truncating the self-consistent solution of (9.225) after the first iteration.

9.6.4 Propagator Sampling with No Time-Step Error

As we pointed out before, in simulations employing the linear propagator (9.205) raising the dimension of the basis set has a detrimental effect on the efficiency of the algorithm since in order to satisfy the bound Eq. (9.207) we are forced to employ an exceedingly small time step. Moreover, in practice values of τ much smaller than the maximum value are usually employed due to the difficulty in obtaining reliable estimates of E_{max} in realistic situations.

To further complicate the scenario, when lattice fixed-node(fixed-phase) methods are employed this maximum value is reduced even further because the diagonal matrix elements of P gets pushed towards the negative region by the addition of the sign-violating contributions $\sum_{\mathbf{m}} h_{\mathbf{mn}}$ in Eq. (9.213). If this method is used to control the sign-problem additional care has to be devoted in the choice of the time-step, greatly deteriorating the efficiency of the overall scheme.

In a discrete space however we can cope with the problem by using an algorithm firstly introduced by Trivedi and Ceperley [25], which shares similarities with the Domains Green's Function Monte Carlo by Kalos et al. [29]. The idea is to use directly (meaning sample from) the exponential propagator

$$P^{exp}(\tau, \mathbf{m}, \mathbf{n}) = \langle \mathbf{m}|e^{-\tau(H-E_T)}|\mathbf{n}\rangle, \tag{9.227}$$

that clearly has no problem with negative diagonal elements. These schemes usually come with the name of *continuous-time* evolution.

For simplicity let us forget the sign-problem for the time being and imagine we are working with the positive-definite importance-sampled greens function (9.215) with $\gamma = 0$ and the corresponding Hamiltonian $\widetilde{H}$ which then satisfies

$$\widetilde{H}_{\mathbf{m},\mathbf{n}} \leq 0 \qquad \forall\ \mathbf{m} \neq \mathbf{n}. \tag{9.228}$$

Furthermore, we will neglect the energy shift E_T since its addition is straightforward.

Recall that the propagator can be written as a product of a stochastic matrix $\widetilde{p}_{\mathbf{m},\mathbf{n}}$ and a weight factor $\widetilde{g}_{\mathbf{n}}$ (cf. Sect. 9.6.1):

$$\widetilde{P}_{\mathbf{m},\mathbf{n}}(\Delta\tau) = \delta_{\mathbf{m},\mathbf{n}} - \Delta\tau\widetilde{H}_{\mathbf{m},\mathbf{n}} = \widetilde{p}_{\mathbf{m},\mathbf{n}}\widetilde{g}_{\mathbf{n}} \tag{9.229}$$

where the two factors are given by:

$$\begin{aligned} \widetilde{p}_{\mathbf{m},\mathbf{n}} &= \frac{\widetilde{P}_{\mathbf{m},\mathbf{n}}(\Delta\tau)}{\widetilde{g}_{\mathbf{n}}}, \\ \widetilde{g}_{\mathbf{n}} &= \sum_{\mathbf{m}} \widetilde{P}_{\mathbf{m},\mathbf{n}}(\Delta\tau) = 1 - \Delta\tau E_L(\mathbf{n}) \end{aligned} \tag{9.230}$$

and in the last equation we have used the expression for the local energy

$$E_L(\mathbf{n}) = \frac{\langle \Phi_T|H|\mathbf{n}\rangle}{\langle \Phi_T|\mathbf{n}\rangle} = \sum_{\mathbf{m}} \frac{\Phi_T(\mathbf{m})\langle \mathbf{m}|H|\mathbf{n}\rangle}{\Phi_T|\mathbf{n}\rangle} \equiv \sum_{\mathbf{m}} \widetilde{H}_{\mathbf{m},\mathbf{n}}. \tag{9.231}$$

The continuous-time limit is recovered by applying M times $\widetilde{P}(\Delta\tau)$ and letting $\Delta\tau \to 0$ while preserving constant the product $\tau = M\Delta\tau$:

$$\lim_{M\to\infty} \widetilde{P}_{\mathbf{m},\mathbf{n}}(\tau)^M = \lim_{M\to\infty} \left(1 - \frac{\tau}{M}\widetilde{H}_{\mathbf{m},\mathbf{n}}\right)^M = \lim_{\Delta\tau\to 0} \left(1 - \Delta\tau\widetilde{H}_{\mathbf{m},\mathbf{n}}\right)^{\frac{\tau}{\Delta\tau}} = \langle \mathbf{m}|e^{-\tau\widetilde{H}}|\mathbf{n}\rangle. \tag{9.232}$$

Now note that if we let $\Delta\tau \to 0$ the probability to make a diagonal move in a single step among the M will accordingly go to ≈ 1, in fact:

$$P_{diag} = \frac{\widetilde{P}_{\mathbf{n},\mathbf{n}}(\Delta\tau)}{\widetilde{g}_{\mathbf{n}}} = \frac{1 - \Delta\tau\widetilde{H}_{\mathbf{n},\mathbf{n}}}{1 - \Delta\tau E_L(\mathbf{n})} \xrightarrow{\Delta\tau\to 0} 1 \tag{9.233}$$

since the local-energy E_L does not depend on the time step but just on the current configuration $\mathbf{n}$. Accordingly, the probability of making K consecutive diagonal moves will be:

$$\begin{aligned} P_{diag}^K &= \left(\frac{\widetilde{P}_{\mathbf{n},\mathbf{n}}(\Delta\tau)}{\widetilde{g}_{\mathbf{n}}}\right)^K = \left(\frac{1 - \Delta\tau\widetilde{H}_{\mathbf{n},\mathbf{n}}}{1 - \Delta\tau E_L(\mathbf{n})}\right)^K \\ &\xrightarrow{K\to\infty} \exp\left(\tau(E_L(\mathbf{n}) - \widetilde{H}_{\mathbf{n},\mathbf{n}})\right) = \exp\left(\tau\widetilde{H}_{\mathbf{n}}^{off}\right) = f_{\mathbf{n}}(\tau) \end{aligned} \tag{9.234}$$

where we have implicitly defined the off-diagonal sum

$$\widetilde{H}_{\mathbf{n}}^{off} = \sum_{\mathbf{m}\neq\mathbf{n}} \widetilde{H}_{\mathbf{m},\mathbf{n}} < 0 \tag{9.235}$$

and the inequality holds thanks to Eq. (9.228).

The elapsed time between consecutive off-diagonal moves is therefore distributed as an exponential distribution $f_{\mathbf{n}}(\tau)$ with average time given by

$$\int_0^\infty \tau f_{\mathbf{n}}(\tau) = -\frac{1}{\widetilde{H}_{\mathbf{n}}^{off}} = \left|\frac{1}{\widetilde{H}_{\mathbf{n}}^{off}}\right|. \tag{9.236}$$

We can then sample the time when the off-diagonal move happens by using a transformation technique: suppose we have a way to sample values ξ from a uniform distribution $g(\xi) = \text{const}$, due to conservation of probability the samples τ drawn from the wanted $f_{\mathbf{n}}(\tau)$ will satisfy:

$$|f(\tau)d\tau| = |g(\xi)d\xi| \quad \longrightarrow \left|\frac{d\xi(\tau)}{d\tau}\right| = f_{\mathbf{n}}(\tau) \tag{9.237}$$

where τ are the samples drawn from the wanted PDF $f_{\mathbf{n}}$. By solving now equation (9.237) for $\xi(\tau)$ and performing the inversion to $\tau = \tau(\xi)$ we obtain the following relation

$$\tau_\xi = \frac{log(\xi)}{\widetilde{H}_{\mathbf{n}}^{off}}. \tag{9.238}$$

that allows to sample exactly from $f_{\mathbf{n}}$ using only samples from a uniform distribution $\xi \in (0, 1)$.

Walkers undergoing such random walk accumulate weight during the K diagonal-moves as well as from performing the off-diagonal step. The weight coming from the diagonal moves is given by

$$w_{\mathbf{n}} = \widetilde{g}_{\mathbf{n}}^K = (1 - \Delta\tau E_L(\mathbf{n}))^K \xrightarrow{\Delta\tau\to 0} e^{-\tau E_L(\mathbf{n})}. \tag{9.239}$$

For the off-diagonal moves instead we have at least two options for sampling the new state $|\mathbf{m}\rangle$:

- heat-bath sampling:

$$P_1(\mathbf{m}, \mathbf{n}) = \widetilde{H}_{\mathbf{m},\mathbf{n}}/\widetilde{H}_{\mathbf{n}}^{off} \tag{9.240}$$

  1. new configuration $|\mathbf{m}\rangle$ is chosen using the normalized probability P_1
  2. the off-diagonal weight would be $w_{\mathbf{m},\mathbf{n}} = 1$

- uniform sampling:

$$P_2(\mathbf{m}, \mathbf{n}) = 1/N_{conn} \tag{9.241}$$

1. new configuration $|\mathbf{m}\rangle$ is chosen among the N_{conn} states connected to $|\mathbf{n}\rangle$
2. reweight the new walker using $w_{\mathbf{m},\mathbf{n}} = P_1(\mathbf{m},\mathbf{n})/P_2(\mathbf{m},\mathbf{n})$

The first option is clearly more expensive per iteration than the second since an explicit calculation of the off-diagonal sum $\widetilde{H}^{off}_{\mathbf{n}}$ is needed in order to normalize P_1. In the uniform sampling case however the weights $w_{\mathbf{m},\mathbf{n}}$ can have large fluctuations forcing the use of smaller time-steps to keep them under control. In our case since we already need to compute the off-diagonal sum in order to generate the fixed-phase hamiltonian Eqs. (9.212) and (9.213) the heat-bath sampling comes with no additional cost. It is worth noting that other choice can be made that are more efficient when fixed-node(phase) is not employed at all [30] or when the transformation that produces $\mathscr{H}_\gamma$ is carried out only approximately [31].

Finally, in order for the measurements along the path to be unbiased we want to define equidistant "time-slices" along the random walk. In order to this we simply choose a target time-step τ_t at the beginning then for each move we first sample a value of τ_ξ from Eq. (9.238), if $\tau_\xi > \tau_t$ we set $\tau = \tau_t$ and use correspondingly a diagonal move if instead $\tau_\xi < \tau_t$ we have to sample an off-diagonal move. The process is preformed until the sum of all the sampled τ_ξ reaches the target time τ_t. The final algorithm for a single walker at $|\mathbf{n}\rangle$ is then as follows:

EXP_Move()
$\tau = \tau_t$
loop
  $E_L(\mathbf{n}) = \sum_{\mathbf{m}} \widetilde{H}_{\mathbf{m},\mathbf{n}}$ ▷ Eq. (9.231)
  $\widetilde{H}^{off}_{\mathbf{n}} = E_L(\mathbf{n}) - \widetilde{H}_{\mathbf{n},\mathbf{n}}$ ▷ Eq. (9.235)
  $\xi =$ rand()
  $\tau_\xi = log(\xi)/\widetilde{H}^{off}_{\mathbf{n}}$ ▷ Eq. (9.238)
  if $\tau_\xi \geq \tau$ **then**
    $w(\mathbf{n}) \rightarrow w(\mathbf{n}) \exp\left(-\tau\ E_L(\mathbf{n})\right)$
    exit
  end if
  $w(\mathbf{n}) \rightarrow w(\mathbf{n}) \exp\left(-\tau_\xi\ E_L(\mathbf{n})\right)$
  $\tau \rightarrow \tau - \tau_\xi$
  $\mathbf{m} \leftarrow$ HeatBath$[P_1, \mathbf{n}]$ ▷ Eq. (9.240)
  $\mathbf{n} \rightarrow \mathbf{m}$
end loop

where the function $HeatBath[P, \mathbf{n}]$ generates a new configuration according to the probability P (e.g. Eq. (9.240)) starting from the current state $\mathbf{n}$. In Problem 9.6 you will try to devise an implementation of this function.

As a final remark, it is evident that the most expensive part of the algorithm is the computation of the local energy $E_L(\mathbf{n}$ since it will require a sum over all states connected to $\mathbf{n}$ from the Hamiltonian and for each one $\mathbf{m}$ of these we have

to compute both the matrix element of the Hamiltonian and the overlap with the trial function $\Psi_T(\mathbf{m})$. The use of symmetries to reduce the size of the sum is thus of fundamental importance to reach medium-sized systems. We can show this for the simple case of a homogeneous system with only two-body interactions so that the connected states will be all the possible 2-particle-2-hole excitations that can be obtained from the initial state $|\mathbf{n}\rangle$. Neglecting the construction of the transformed matrix $\widetilde{H}$, we can then implement the calculation of the local energy as

```
EL_calc1()
E_L = 0
for i ∈ occ(n) do
    for j ∈ occ(n) do
        for a ∈ 𝒮 \ occ(n) do
            for b ∈ 𝒮 \ occ(n) do
                |m⟩ = a_a^† a_b^† a_i a_j |n⟩
                E_L = E_L + H̃_{m,n}
            end for
        end for
    end for
end for
E_L = E_L/4
```

where $occ(\mathbf{n})$ is the set of single-particle states that are occupied in the initial state $\mathbf{n}$. The above algorithm requires $O(N_{occ}^2 \mathcal{N}_s^2)$ evaluations of the Hamiltonian. Many of these are however equivalent to other ones or just zero. For instance all the terms with $i = j$ or $a = b$ give zero due to the Pauli principle. If we fix an ordering of the single particle orbitals in the many-body states and use anti-symmetrized matrix elements the configurations obtained interchanging e.g. $i \leftrightarrow j$ are equivalent. Finally if both momentum and spin are conserved, given the triple (i, j, a) there exist only one single particle state b allowed. An implementation like

```
EL_calc2()
E_L = 0
for i ∈ occ(n) do
    for j < i ∈ occ(n) do
        for a ∈ 𝒮 \ occ(n) do
            b ← FourthState[i, j, a]
            if b ∈ 𝒮 \ occ(n) and b < a then
                |m⟩ = a_a^† a_b^† a_i a_j |n⟩
                E_L = E_L + H̃_{m,n}
```

(continued)

```
            end if
        end for
    end for
end for
```

will take now only $O(N_{occ}^2 \mathcal{N}_s)$ evaluations of the Hamiltonian at most, and with a reduced prefactor with respect to the previous version. The function *FourthState* returns the only single particle state allowed by symmetry.

9.6.5 Results

The combination of imaginary time projection, use of importance function derived from CC calculations and no time-step error propagator make up the algorithm that goes under the name of Configuration Interaction Monte Carlo (CIMC). Actual calculations with CIMC require a substantial amount of CPU time. Here we report some results obtained by making use of a simplified Hamiltonian in which the nucleon-nucleon interaction is described by the Minnesota interaction. The system under investigation is homogeneous pure neutron matter (PNM). In QMC calculations PNM is typically modeled as a periodic system containing A neutrons. The cell size is adjusted in such a way that the average density of the system is ρ.

In Fig. 9.7 we show how the computed energy depends on the number of plane wave shells included in the model space. As it can be seen, it is necessary to pay attention to the convergence of the results, which can strongly depend on the specific details of the system. In this case, for instance, one can easily see how convergence is faster when the density is increased.

In Fig. 9.8 the energy computed by CIMC shown for the same neutron matter model as a function of the density (the so called "Equation of State" of neutron matter) is compared with the coupled theory results with doubles (CCD) only discussed in the previous chapter. In this calculation single-particle states up to $N_{\max} = 36$ have been used. The CIMCC and CCD results are converged to the fifth digit as function of $N_{\max}$. The agreement between the two methods is at the level of the third digit after the decimal point for neutron matter with the Minnesota interaction. This is a striking agreement between such different many-body methods, in particular for larger densities where correlations and contributions from states above and below the Fermi level play a larger role, as seen from the difference between the reference energy and the CIMC and CCD energies. Most likely, there will be larger differences between different many-body methods when proton correlations are brought in, as well as when more realistic interaction models will be used. Such results will be presented elsewhere. In the next two chapters we will add results using two additional many-body methods, the in-medium SRG approach described in Chap. 10 and the Green's function approach of Chap. 11.

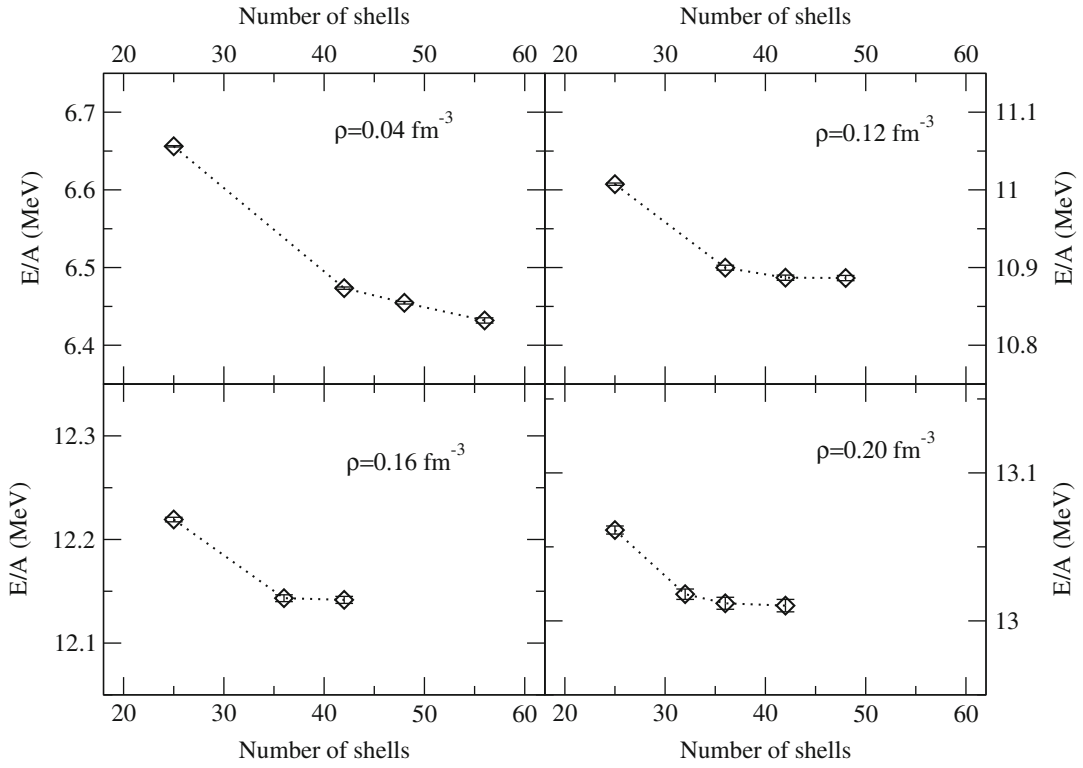

Fig. 9.7 Convergence of the CIMC energies as a function of the number of shells used for a periodic cell of 66 neutrons at different densities

9.7 Conclusions and Perspectives

Quantum Monte Carlo methods are still one of the most powerful tools to attack general many body problems, and in particular the many-nucleon problem. Despite the fact that the Fermion sign problem prevents us so far from having strictly exact results for the solution of the Schrödinger equation, the accuracy that can be reached is very high, and in any cases it constitutes the current benchmark.

Another important general feature of QMC calculations is that they provide a very flexible framework in which it is possible to explore from low temperature condensed helium, to trapped fermions, from atoms and molecules ad solid state devices to nuclei and eventually lattice QCD. It is not rare that technical improvements spread across different disciplines, and the development of the method itself is a common ground that is often the subject of interdisciplinary workshops and conferences.

In the field of nuclear physics it is possible that Fock-space based methods will eventually become the standard. Their main feature is the possibility of dealing with non-local interactions, which makes it possible to extend the use of QMC to the original formulations of ξ-EFT potentials, and a whole class of soft-core interactions that so far have never been used in this context. On the other hand, the availability of more and more accurate versions of the AFDMC codes will open the access of

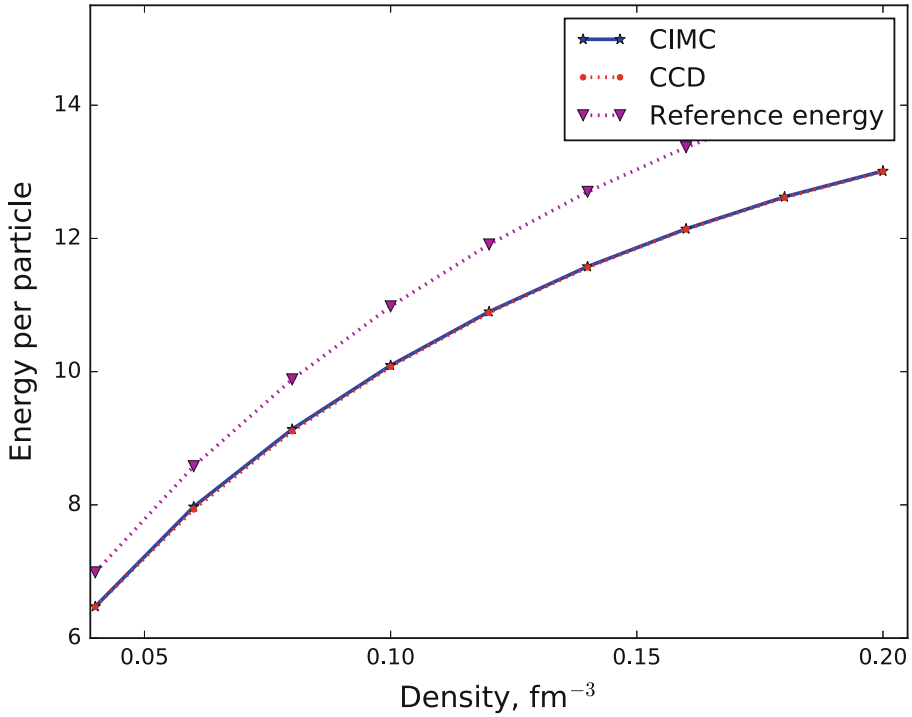

Fig. 9.8 Equation of state of neutron matter modeled as a periodic cell containing A = 66 neutrons using the CIMC method and coupled cluster theory with doubles correlations. Single-particle states up to $N_{\max} = 36$ have been included

accurate studies of the equation of state f neutron and nuclear matter, and of general baryonic matter of extreme importance for astrophysical applications, concerning in particular the physics of neutron stars. The possibility of extending accurate calculations to large A systems is also crucial for understanding the phenomenology of exotic beams.

In this chapter we did not deal with the problem of evaluating excited states and dynamical quantities within a QMC framework. Several methods are nowadays available, mostly based on the evaluation of the Laplace transform of a given response function by means of the calculation of imaginary time correlation functions. Many technical advances have been recently made in this field (see e.g. Refs. [32–35]), and the subject is still under very active investigation.

Finally, the hardest wall to climb remains the solution of the Fermion sign problem. Although there are claims that the problem is NP complete (which is true in general), thereby preventing any solution within standard classical computation, there are hints that many Hamiltonians of interest might admit a viable solution with polynomial scaling in A. This problem would definitely deserve more efforts than those that are presently devoted to its solution.

9.8 Problems

9.1 Evaluate the following integral by means of the Metropolis algorithm

$$I = \int_0^1 e^x - 1\, dx$$

sampling points from

1. P(x) = 1 for $x \in [0, 1]$
2. P(x) = x for $x \in [0, 1]$

- Compare the average and the statistical error for the cases 1) and 2). Which is the best estimate?
- Try to figure out a way to sample a probability density proportional to x^n, and reevaluate the integral I. How is the convergence and the statistical error behaving by increasing n? Try to give an explanation of the result.

9.2 Try to sketch the general proof that for a generic integral I defined as in problem 9.1 the best statistical error is obtained when sampling from a probability density proportional to F(x).

9.3 Consider the one-dimensional Hamiltonian:

$$\hat{H} = \frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{2}x^2$$

and consider the parametrized family of trial solutions $\psi(x, \alpha, \beta) = e^{-\alpha^2 x^2} - \beta$. Compute by means of the Metropolis algorithm the energy and the standard deviation of the energy as a function of α keeping $\beta = 0.01$. Is the minimum found at the same value than for $\beta = 0$? Why?

9.4 Prove that the propagator defined in the integrand of Eq. (9.83) is the Green's function of the differential equation (9.82).

9.5 Show that the given two fixed-phase energies E_a and E_b obtained using the hamiltonians $\mathscr{H}_\gamma$ defined in Eqs. (9.212) and (9.213) with $\gamma = a$ and $\gamma = b$ $(a, b \geq 0)$ the linear extrapolation to $\gamma = -1$ (remember that $\mathscr{H}_{-1} = H$) is still an upper bound. (Hint: show that E_γ is a convex function of the parameter γ).

9.6 Implement the function $HeatBath[P, \mathbf{n}]$ that appears in the algorithm EXP_Move() in Sect. 9.6.4.

9.7 In the case of nuclear Hamiltonians spin is not a conserved quantity. Referring to the calculation of the local energy in Sect. 9.6.4, how would you modify the subroutine EL_calc2 to take into account the non-conservation of spin?

Appendix

In this appendix we give the proof of the upper-bound property for the auxiliary hamiltonians $\mathscr{H}_\gamma$, defined in Sect. 9.6.2, for the general complex-hermitian case (see [27] for the original proof in the real symmetric case). We will concentrate in the simpler case $\gamma = 0$ in Eqs. (9.212) and (9.213), extension to the generic $\gamma \geq 0$ is then straightforward. In what follows we will use the shorthand $\mathscr{H}_{\gamma=0} \equiv \widetilde{\widetilde{H}}$. Let $\Psi(\mathbf{n})$ be any arbitrary wave function, our goal is to show that

$$\Re[\langle\Psi|\widetilde{H}|\Psi\rangle] \geq \Re\left[\langle\Psi|H|\Psi\rangle\right] . \tag{9.242}$$

Let us proceed by considering the following difference:

$$\begin{aligned}
\Re[\langle\Psi|\widetilde{H}|\Psi\rangle] - \Re\left[\langle\Psi|H|\Psi\rangle\right] &= \sum_{\mathbf{mn}} \Re\left[\Psi^*(\mathbf{m})(\widetilde{H}_{\mathbf{mn}} - H_{\mathbf{mn}})\Psi(\mathbf{n})\right] \\
&= \sum_{\mathbf{mn}} h_{\mathbf{mn}}|\Psi(\mathbf{n})|^2 + \sum_{\mathbf{m}\neq\mathbf{n}} \Re\left[\Psi^*(\mathbf{m})(\widetilde{H}_{\mathbf{mn}} - H_{\mathbf{mn}})\Psi(\mathbf{n})\right] \\
&= \sum_{\mathbf{n}} \sum_{\mathfrak{s}_{\mathbf{mn}}\neq -} |\Psi(\mathbf{n})|^2 \frac{\Re\left[\Phi_T^*(\mathbf{m})H_{\mathbf{mn}}\Phi_T(\mathbf{n})\right]}{|\Phi_T(\mathbf{n})|^2} - \Re\left[\Psi^*(\mathbf{m})H_{\mathbf{mn}}\Psi(\mathbf{n})\right]
\end{aligned} \tag{9.243}$$

where the second sum is over all $\mathbf{mn}$ pairs such that $\mathfrak{s}_{\mathbf{mn}}$ of (9.211) is positive-definite. The last term can now be rewritten as:

$$\begin{aligned}
\Re\left[\Psi^*(\mathbf{m})H_{\mathbf{mn}}\Psi(\mathbf{n})\right] &= \Re\left[\Psi^*(\mathbf{m})\Phi_T(\mathbf{m})\Phi_T(\mathbf{m})^{-1}H_{\mathbf{mn}}\Phi_T^*(\mathbf{n})^{-1}\Phi_T^*(\mathbf{n})\Psi(\mathbf{n})\right] \\
&= (\Psi^*(\mathbf{m})\Phi(\mathbf{m}))\Re\left[\Phi_T(\mathbf{m})^{-1}H_{\mathbf{mn}}\Phi_T^*(\mathbf{n})^{-1}\right](\Phi_T^*(\mathbf{n})\Psi(\mathbf{n})) \\
&= (\Psi^*(\mathbf{m})\Phi(\mathbf{m}))\Re\left[\frac{\Phi_T^*(\mathbf{m})}{|\Phi_T(\mathbf{m})|^2}H_{\mathbf{mn}}\frac{\Phi_T(\mathbf{n})}{|\Phi_T(\mathbf{n})|^2}\right](\Phi_T^*(\mathbf{n})\Psi(\mathbf{n}))
\end{aligned} \tag{9.244}$$

where in the second step we used the fact that by employing a real propagator we are imposing a fixed-phase constraint, ie $\Im(\Phi_T^*(\mathbf{n})\Psi(\mathbf{n})) = 0$ for every $\mathbf{n}$ explored in the random walk. The equation for the difference becomes:

$$\begin{aligned}
\Re[\langle\Psi|\widetilde{H}|\Psi\rangle] - \Re\left[\langle\Psi|H|\Psi\rangle\right] &= \sum_{\mathbf{mn}} \Re\left[\Psi^*(\mathbf{m})(\widetilde{H}_{\mathbf{mn}} - H_{\mathbf{mn}})\Psi(\mathbf{n})\right] \\
&= \sum_{\mathbf{n}} \sum_{\mathfrak{s}_{\mathbf{mn}}\neq -} \frac{\Re\left[\Phi_T^*(\mathbf{m})H_{\mathbf{mn}}\Phi_T(\mathbf{n})\right]}{|\Phi_T(\mathbf{n})|^2}\left(|\Psi(\mathbf{n})|^2 - \frac{(\Psi^*(\mathbf{m})\Phi_T(\mathbf{m}))(\Phi_T^*(\mathbf{n})\Psi(\mathbf{n}))}{|\Phi(\mathbf{m})|^2}\right) .
\end{aligned} \tag{9.245}$$

Using again the fixed-phase constraint (ie. $(\Phi_T^*(\mathbf{n})\Psi(\mathbf{n})) \equiv (\Phi_T(\mathbf{n})\Psi^*(\mathbf{n}))$) we can rewrite the numerator of the second term as:

$$(\Psi^*(\mathbf{m})\Phi_T(\mathbf{m}))(\Phi_T^*(\mathbf{n})\Psi(\mathbf{n})) = -\frac{1}{2}\big(|\Psi^*(\mathbf{m})\Phi_T(\mathbf{n}) - \Phi_T^*(\mathbf{m})\Psi(\mathbf{n})|^2 - |\Phi_T(\mathbf{n})|^2|\Psi(\mathbf{m})|^2 - |\Phi_T(\mathbf{m})|^2|\Psi(\mathbf{n})|^2\big) \tag{9.246}$$

and then we have:

$$\begin{aligned}
&\Re\left[\langle\Psi|\widetilde{H}|\Psi\rangle\right] - \Re\left[\langle\Psi|H|\Psi\rangle\right] = \sum_{\mathbf{mn}} \Re\left[\Psi^*(\mathbf{m})(\widetilde{H}_{\mathbf{mn}} - H_{\mathbf{mn}})\Psi(\mathbf{n})\right] \\
&= \sum_{\mathbf{n}} \sum_{s_{\mathbf{mn}}\neq -} \frac{\Re\left[\Phi_T^*(\mathbf{m})H_{\mathbf{mn}}\Phi_T(\mathbf{n})\right]}{|\Phi_T(\mathbf{n})|^2}\big(|\Psi(\mathbf{n})|^2 + \frac{|\Psi^*(\mathbf{m})\Phi_T(\mathbf{n}) - \Phi_T^*(\mathbf{m})\Psi(\mathbf{n})|^2}{2|\Phi_T(\mathbf{m})|^2} \\
&\quad - \frac{|\Phi_T(\mathbf{n})|^2|\Psi(\mathbf{m})|^2}{2|\Phi_T(\mathbf{m})|^2} - \frac{|\Psi(\mathbf{n})|^2}{2}\big) \\
&= \text{(positive terms)} \\
&\quad + \sum_{\mathbf{n}} \sum_{s_{\mathbf{mn}}\neq -} \frac{\Re\left[\Phi_T^*(\mathbf{m})H_{\mathbf{mn}}\Phi_T(\mathbf{n})\right]}{2|\Phi_T(\mathbf{n})|^2}\left(|\Psi(\mathbf{n})|^2 - \frac{|\Phi_T(\mathbf{n})|^2|\Psi(\mathbf{m})|^2}{|\Phi_T(\mathbf{m})|^2}\right)
\end{aligned} \tag{9.247}$$

Now we note that

$$\Re\left[\Phi_T^*(\mathbf{m})H_{\mathbf{mn}}\Phi_T(\mathbf{n})\right] = \Re\left[\Phi_T^*(\mathbf{n})H_{\mathbf{nm}}\Phi_T(\mathbf{m})\right]$$

for a complex-hermitian hamiltonian, we can then express the sums by allowing only unique **mn** combinations:

$$\begin{aligned}
&\Re\left[\langle\Psi|\widetilde{H}|\Psi\rangle\right] - \Re\left[\langle\Psi|H|\Psi\rangle\right] = \sum_{\mathbf{mn}} \Re\left[\Psi^*(\mathbf{m})(\widetilde{H}_{\mathbf{mn}} - H_{\mathbf{mn}})\Psi(\mathbf{n})\right] \\
&\quad = \text{(positive terms)} \\
&\quad + \sum_{\mathbf{n}} {\sum_{s_{\mathbf{mn}}\neq -}}' \Re\left[\Phi_T^*(\mathbf{m})H_{\mathbf{mn}}\Phi_T(\mathbf{n})\right]\left(\frac{|\Psi(\mathbf{n})|^2}{2|\Phi_T(\mathbf{n})|^2} + \frac{|\Psi(\mathbf{m})|^2}{2|\Phi_T(\mathbf{m})|^2}\right. \\
&\quad \left. - \frac{|\Phi_T(\mathbf{n})|^2|\Psi(\mathbf{m})|^2}{2|\Phi_T(\mathbf{m})|^2|\Phi_T(\mathbf{n})|^2} - \frac{|\Phi_T(\mathbf{m})|^2|\Psi(\mathbf{n})|^2}{2|\Phi_T(\mathbf{n})|^2|\Phi_T(\mathbf{m})|^2}\right) \\
&\quad = \text{(positive terms)}
\end{aligned} \tag{9.248}$$

which by definition is positive. The extension to the case with $\gamma > 0$ is straightforward since we are basically adding a positive constant to the difference.

References

1. M.H. Kalos, P.A. Whitlock, *Monte Carlo Methods* (Wiley-VCH, Weinheim, 2008)
2. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, J. Chem. Phys. **21**, 1087 (1953)
3. W.K. Hastings, Biometrika **57**, 97 (1970)
4. D. Ceperley, G.V. Chester, M.H. Kalos, Phys. Rev. B **16**, 3081 (1977)
5. M.P. Nightingale, C.J. Umrigar, in *Recent Advances in Quantum Monte Carlo Methods*, ed. by W.A. Lester Jr. (World Scientific, Singapore, 1996)
6. J. Toulouse, C.J. Umrigar, J. Chem. Phys. **126**, 084102 (2007)
7. S. Sorella, Phys. Rev. B **64**, 024512 (2001)
8. A. Sarsa, K.E. Schmidt, W.R. Magro, J. Chem. Phys. **113**, 1366 (2000)
9. J.B. Anderson, J. Chem. Phys. **65**, 4121 (1976)
10. M. Troyer, U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005)
11. M.H. Kalos, F. Pederiva, Phys. Rev. Lett. **85**, 3547 (2000)
12. R. Assaraf, M. Caffarel, A. Khelif, J. Phys. A: Math. Theor. **40**, 1181 (2007)
13. D.M. Ceperley, B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980)
14. W.M.C. Foulkes, L. Mitas, R.J. Needs, G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001)
15. J. Carlson, V.R. Pandharipande, R. Schiavilla, Phys. Rev. C **47**, 484 (1993)
16. G. Ortiz, D.M. Ceperley, R.M. Martin, Phys. Rev. Lett. **71**, 2777 (1993)
17. S. Zhang, H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003)
18. G.H. Booth, A.J.W. Thom, A. Alavi, J. Chem. Phys. **131**, 054106 (2009)
19. D. Cleland, G.H. Booth, A. Alavi, J. Chem. Phys. **132**, 041103 (2010)
20. F.R. Petruzielo, A.A. Holmes, H.J. Changlani, M.P. Nightingale, C.J. Umrigar, Phys. Rev. Lett. **109**, 230201 (2012)
21. G.H. Booth, A. Grüneis, G. Kresse, A. Alavi, Nature **493**, 365 (2013)
22. A. Mukherjee, Y. Alhassid, Phys. Rev. A **88**, 053622 (2013)
23. A. Roggero, A. Mukherjee, F. Pederiva, Phys. Rev. B **88**, 115138 (2013)
24. A. Roggero, A. Mukherjee, F. Pederiva, Phys. Rev. Lett. **112**, 221103 (2014)
25. N. Trivedi, D.M. Ceperley, Phys. Rev. B **41**, 4552 (1990)
26. S. Sorella, L. Capriotti, Phys. Rev. B **61**, 2599 (2000)
27. D.F.B. ten Haaf, H.J.M. van Bemmel, J.M.J. van Leeuwen, W. van Saarloos, D.M. Ceperley, Phys. Rev. B **51**, 13039 (1995)
28. E. Rrapaj, A. Roggero, J.W. Holt, Phys. Rev. C **93**, 065801 (2016)
29. M.H. Kalos, D. Levesque, L. Verlet, Phys. Rev. A **9**, 2178 (1974)
30. A. Holmes, H.J. Changlani, C.J. Umrigar, J. Chem. Theory Comput. **12**, 1561 (2016)
31. M. Kolodrubetz, B.K. Clark, Phys. Rev. B **86**, 075109 (2012)
32. E. Vitali, M. Rossi, L. Reatto, D.E. Galli, Phys. Rev. B **82**, 174510 (2010)
33. A. Roggero, F. Pederiva, G. Orlandini, Phys. Rev. B **88**, 094302 (2013)
34. A. Lovato, S. Gandolfi, R. Butler, J. Carlson, E. Lusk, S. C. Pieper, R. Schiavilla, Phys. Rev. Lett. **111**, 092501 (2013)
35. A. Lovato, S. Gandolfi, J. Carlson, S.C. Pieper, R. Schiavilla, Phys. Rev. C **91**,062501(R) (2015)