ON A PROBABILISTIC INTERPRETATION OF SHAPE DERIVATIVES OF DIRICHLET GROUNDSTATES WITH APPLICATION TO FERMION NODES

MATHIAS ROUSSET

Abstract. This paper considers Schrödinger operators, and presents a probabilistic interpretation of the variation (or shape derivative) of the Dirichlet groundstate energy when the associated domain is perturbed. This interpretation relies on the distribution on the boundary of a stopped random process with Feynman-Kac weights. Practical computations require in addition the explicit approximation of the normal derivative of the groundstate on the boundary. We then propose to use this formulation in the case of the so-called fixed node approximation of Fermion groundstates, defined by the bottom eigenelements of the Schrödinger operator of a Fermionic system with Dirichlet conditions on the nodes (the set of zeros) of an initially guessed skew-symmetric function. We show that shape derivatives of the fixed node energy vanishes if and only if either (i) the distribution on the nodes of the stopped random process is symmetric; or (ii) the nodes are exactly the zeros of a skew-symmetric eigenfunction of the operator. We propose an approximation of the shape derivative of the fixed node energy that can be computed with a Monte-Carlo algorithm, which can be referred to as Nodal Monte-Carlo (NMC). The latter approximation of the shape derivative also vanishes if and only if either (i) or (ii) holds.

Mathematics Subject Classification. 60H30, 65C35, 65C05, 35P99.

Received July 8, 2009. Revised February 16, 2010. Published online August 26, 2010.

1. INTRODUCTION AND RESULTS

Throughout this paper, we consider a Schrödinger operator in $\mathbb{R}^d$ of the form:

$$H = -\frac{\Delta}{2} + V,$$  \hspace{2cm} (1.1)

with a smooth potential $V$ going to infinity at infinity, and acting on real valued functions generically denoted with the letter $\psi$ (‘wave functions’). Such functions $\psi$ will be defined up to a real valued multiplicative constant (e.g. in eigenvalue and/or variational problems). We also consider a general family

$$\theta \mapsto \Omega_\theta$$ \hspace{2cm} (1.2)

Keywords and phrases. Schrödinger operator, groundstate, shape derivatives, Feynman-Kac formula, quantum Monte-Carlo methods, Fermion nodes, fixed node approximation.

1 INRIA Lille, Nord Europe & Université Lille 1, Villeneuve d’Ascq, France. mathias.rousset@inria.fr

Article published by EDP Sciences © EDP Sciences, SMAI 2010
of open smooth domains in $\mathbb{R}^d$ depending sufficiently smoothly of a parameter $\theta \in \mathbb{R}^p$. The boundary will be denoted $\partial \Omega_{\theta}$. Gradients in the space $\mathbb{R}^d$ will be denoted $\nabla$, and gradients with respect to $\theta$, $\nabla \theta$. The Dirichlet groundstate and its Dirichlet groundstate energy $(\psi_\theta^*, E_\theta^*)$, are then defined as the unique bottom eigenelements of $H$, solution to the variational problem:

$$E_\theta^* := \inf \left( \frac{\int_{\Omega_{\theta}} \psi H(\psi)}{\int_{\Omega_{\theta}} \psi^2}, \psi|_{\partial \Omega_{\theta}} = 0 \right) = \frac{\int_{\Omega_{\theta}} \psi_\theta^* H(\psi_\theta^*)}{\int_{\Omega_{\theta}} (\psi_\theta^*)^2}. \quad (1.3)$$

Calculus of variations detailed in Section 2 then yields the shape derivative of the groundstate energy through the formula:

$$\nabla \theta E_\theta^* = -\frac{1}{2} \int_{\partial \Omega_{\theta}} |\nabla \psi_\theta^*|^2 r_\theta \, d\sigma, \quad (1.4)$$

where in the above $\sigma$ is the usual surface measure induced by the canonical Euclidean structure $\mathbb{R}^d$, and $r_\theta$ is the shape derivative, i.e. the field

$$r_\theta : \partial \Omega_{\theta} \to \mathbb{R}^p$$

such that formally the boundary variation writes down:

$$\partial \Omega_{\theta + \delta \theta} = \{ x + n(x)r_\theta(x) \cdot d\theta \mid x \in \partial \Omega_{\theta} \},$$

where $n(x)$ is the exterior normal vector at $x \in \partial \Omega_{\theta}$. If $\{\psi_\theta\}_{\theta \in \mathbb{R}^p}$ is a smooth family of smooth functions such that $\psi_\theta(x) = 0$ for $x \in \partial \Omega_{\theta}$, then the shape derivative $r_\theta$ can also be defined through

$$r_\theta(x) \nabla \psi_\theta(x) \cdot n(x) = -\nabla \theta \psi_\theta(x) \quad \forall x \in \partial \Omega_{\theta}. \quad (1.5)$$

Formula (1.5) can be proved as follows: from the Dirichlet conditions, one has for any small $h$:

$$0 = \psi_{\theta + h}^* \left( x + h \cdot r_\theta(x)n(x) + O(|h|^2) \right) - \psi_\theta^* = h \cdot r_\theta(x) \nabla \psi_\theta^*(x) \cdot n(x) + h \cdot \nabla \theta \psi_\theta^*(x) + O(h).$$

Differentiability in formula (1.4) is a classical result of abstract analytic perturbation of linear operators (see [24]), but can be proved directly with the variational formulation as in [19].

In Section 3, we introduce a standard Wiener process (Brownian motion)

$$t \mapsto W_t,$$

with some given initial distribution in $\Omega_{\theta}$. The first exit time of the domain $\Omega_{\theta}$ is denoted by

$$\tau := \inf \{ t \geq 0 \mid W_t \in \partial \Omega_{\theta} \}. \quad (1.6)$$

The long time probability distribution of the latter process with Feynman-Kac weights, and conditioned to remain in the domain $\Omega_{\theta}$ is denoted $d\eta_{\theta}^*$

$$\int_{\Omega_{\theta}} \varphi \, d\eta_{\theta}^* := \lim_{T \to +\infty} \frac{\mathbb{E} \left( \varphi(W_T) \mathbb{1}_{T \leq \tau} e^{-\int_0^T V(W_s) \, ds} \right)}{\mathbb{E} \left( \mathbb{1}_{T \leq \tau} e^{-\int_0^T V(W_s) \, ds} \right)}. \quad (1.7)$$
It has a probability density function given by the signed groundstate $\psi^*_\theta$:

$$\int_{\Omega_\theta} \varphi \, d\eta^*_\theta = \frac{\int_{\Omega_\theta} \varphi \, \psi^*_\theta \, dx}{\int_{\Omega_\theta} \psi^*_\theta \, dx}, \quad (1.8)$$

and the exponential rate of the evolution of the weighted extinction probability yields the groundstate energy:

$$\lim_{T \to +\infty} -\frac{1}{T} \ln \mathbb{E} \left( \mathbb{I}_{T \leq \tau} e^{-\int_0^\tau \psi^*_\theta \, (W_s) \, ds} \right) = E^{*}_{\theta}. \quad (1.9)$$

The probabilistic interpretations (1.7)–(1.9) have some variants (see (3.4), (3.5)) where a drift is added to the Wiener process and the range of the potential $V$ in the Feynman-Kac weight $e^{-\int_0^\tau \psi^*_\theta \, (W_s) \, ds}$ is reduced. This leads to Monte-Carlo methods with some importance sampling variance reduction which can efficiently compute the couple $(\psi^*_\theta, E^{*}_{\theta})$. This method has been widely used and studied in many fields. Special care is required to treat the weight $e^{-\int_0^\tau \psi^*_\theta \, (W_s) \, ds}$ when averages are computed. For instance when several random processes are simulated in parallel, some re-sampling of the Dirichlet groundstate (1.3) yields a similar formula but at finite time (as opposed to large time, which is the case here).

Then, we consider the weighted distribution of the Wiener process at the hitting time $\tau$, when the process is initially distributed according to $\eta^*_\theta$ defined in (1.7). The latter distribution is denoted $d\mu^*_{\theta, \lambda}\lambda$ and reads

$$\int_{\partial \Omega_\theta} \varphi \, d\mu^*_{\theta, \lambda} := \mathbb{E} \left( \varphi(W_\tau) \, \mathbb{I}_{\tau < +\infty} e^{-\int_0^\tau (V(W_s) - \lambda) \, ds} \mid \text{Law}(W_0) = \eta^*_\theta \right)$$

$$= \lim_{T \to +\infty} \frac{\mathbb{E} \left( \varphi(W_\tau) \, \mathbb{I}_{\tau < +\infty} e^{-\int_0^\tau (V(W_s) - \lambda) \, ds} \right)}{\mathbb{E} \left( \mathbb{I}_{\tau < +\infty} e^{-\int_0^\tau (V(W_s) - \lambda) \, ds} \right)} \quad (1.10)$$

which verifies

$$\int_{\partial \Omega_\theta} \varphi \, d\mu^*_{\theta, \lambda} = -\frac{1}{2(E^{*}_{\theta} - \lambda)} \int_{\Omega_\theta} \psi^*_\theta \int_{\partial \Omega_\theta} \varphi \nabla \psi^*_\theta \cdot n \, d\sigma. \quad (1.11)$$

Formula (1.11) holds for any $\lambda < E^{*}_{\theta}$ and is the grounding formula of this paper. Note that it can be related to the Dirichlet energy variation $\nabla \theta E^{*}_{\theta}$ in (1.4) by remarking that $|\nabla \psi^*_\theta \cdot n| = |\nabla \psi^*_\theta|$. Up to our knowledge, the formula (1.11) has never been pointed out in the literature, although the sensitivity analysis carried out in [12] yields a similar formula but at finite time (as opposed to large time, which is the case here).

Approximations of $\nabla \theta E^{*}_{\theta}$ with Monte-Carlo methods can then be carried out using an approximating sequence of the Dirichlet groundstate (1.3) $\psi^n_{\theta} \xrightarrow{N \to \infty} \psi^*_\theta$, and random samples of size $N$ approximating the probabilistic formulations (1.7)–(1.10), denoted $\eta^N_{\theta} \xrightarrow{N \to \infty} \eta^*_\theta$ and $\mu^N_{\theta, \lambda} \xrightarrow{N \to \infty} \mu^*_{\theta, \lambda}$. The following identity can then be used:

$$\nabla \theta E^{*}_{\theta} = \frac{(E^{*}_{\theta} - \lambda)}{\int_{\Omega_\theta} \psi^*_\theta \, d\eta^*_\theta} \int_{\partial \Omega_\theta} r_\theta \nabla \psi^*_\theta \cdot n \, d\mu^*_{\theta, \lambda} \sim \frac{(E^{*}_{\theta} - \lambda)}{\int_{\Omega_\theta} \psi^*_\theta \, d\eta^N_{\theta}} \int_{\partial \Omega_\theta} r_\theta \nabla \psi^n_{\theta} \cdot n \, d\mu^N_{\theta, \lambda}. \quad (1.12)$$

The main limitation of computing $\nabla \theta E^{*}_{\theta}$ with the Monte-Carlo technique suggested above is the necessity of an analytical approximation $\psi^*_\theta$ of the groundstate $\psi^*_\theta$. Especially, the pointwise convergence of the normal derivative $\nabla \psi^n_{\theta} \cdot n$ may be hard to achieve. However, for practical situations in high dimension, we do not know any alternative point of view. A clear motivating example of a high dimensional problem is the case of Fermionic systems where the so-called fixed node approximation is used.
In Section 4, we introduce Fermionic groundstates \((\psi^*_{F}, E^*_{F})\) associated to a finite symmetry group \(S \subset O(\mathbb{R}^d)\) of the Hamiltonian \(H\) in (1.1); where \(O(\mathbb{R}^d)\) denotes the group of isometries. \(S\) is simply the permutation group of identical particles for physical systems. Fermionic groundstates are the solutions to the variational problem:

\[
E^*_{F} := \inf \left( \int_{\mathbb{R}^d} \psi H (\psi) \frac{d^d x}{\psi^2}, \quad \forall S \in S, \quad \psi \circ S = \det(S) \psi \right) = \int_{\Omega_\theta} \psi^*_F H (\psi^*_F) \frac{d^d x}{\int_{\Omega_\theta} (\psi^*_F)^2}. \tag{1.13}
\]

Any function \(\psi\) verifying the symmetry property

\[
\forall S \in S, \quad \psi \circ S = \det(S) \psi
\]

will be called **skew-symmetric**, whereas any function \(\psi\) verifying

\[
\forall S \in S, \quad \psi \circ S = \psi
\]

will be called **symmetric**. Note that existence of \((\psi^*_{F}, E^*_{F})\) follows in our context from the fact that \(H\) has a discrete spectrum, and is a classical result of spectral theory for more general potential \(V\) (see references in [6]); but uniqueness does not hold in general. In practice, \((\psi^*_{F}, E^*_{F})\) is computed using a parameterization of skew-symmetric functions, and a numerical optimization procedure. This is the main problem of computational Quantum Chemistry, and forms a huge scientific field. We refer to [7] for a mathematical introduction with a consequent bibliography. See also the following two typical papers [27,29] involving wave function optimization using a Monte-Carlo method. Monte Carlo methods in computational Quantum Chemistry are referred to as Quantum Monte Carlo (QMC) methods. For physical systems, the parameterization is given as a finite sum of Slater determinants multiplied by a strictly positive symmetric factor (called the Jastrow factor). The result of the latter optimization relies crucially on the quality of the parameterization, and will be called the **trial wave function**, which is a skew-symmetric function denoted \(\psi^{I}_{\theta}\). We will then consider a family of skew-symmetric functions \({\psi^{I}_{\theta}}\) \(\theta \in \mathbb{R}^p\), with an explicit analytical expression, which includes \(\psi^{I}_{\theta_0}\). The latter is used to define the **nodal domains**

\[
\Omega_\theta = \mathcal{N}^+_\theta \cup \mathcal{N}^-_\theta,
\]

where:

\[
\begin{align*}
\mathcal{N}^+_\theta &= \{ x \in \mathbb{R}^d \mid \psi^I_\theta(x) > 0 \} \\
\mathcal{N}^-_\theta &= \{ x \in \mathbb{R}^d \mid \psi^I_\theta(x) < 0 \} \\
\partial \mathcal{N}_\theta &= \{ x \in \mathbb{R}^d \mid \psi^I_\theta(x) = 0 \}.
\end{align*}
\tag{1.14}
\]

The **fixed node approximation** consists then in computing with a Monte-Carlo method the Dirichlet groundstate

\[
(\psi^{'F}_{\theta}, E^{'F}_{\theta}) = (\psi^{I}_{\theta}, E^*_{\theta})
\]

of the variational problem with Dirichlet conditions in \(\mathcal{N}^+_\theta \cup \mathcal{N}^-_\theta\). We refer the reader to [10,11] for historical papers on the Monte-Carlo computation of fixed node groundstates. The latter variational problem reads
explicitly:

\[ E_{\theta}^{\text{FN}} := \inf \left( \int_{\mathbb{R}^d} \psi H (\psi), \psi|_{\partial N_{\theta}} = 0 \right) = \frac{\int_{\mathbb{R}^d} \psi_0^{\text{FN}} H (\psi_0^{\text{FN}})}{\int_{\mathbb{R}^d} (\psi_0^{\text{FN}})^2} \]

\[ = \inf \left( \int_{N_{\theta}^+} \psi H (\psi), \psi|_{\partial N_{\theta}} = 0 \right) = \frac{\int_{N_{\theta}^+} \psi_0^{\text{FN}} H (\psi_0^{\text{FN}})}{\int_{N_{\theta}^+} (\psi_0^{\text{FN}})^2} \]  

(1.15)

of the trial wave function \( \psi_0 \) can be minimized towards the Fermionic groundstate energy \( E_{\overline{\theta}}^{\text{F}} \) using some optimization procedure, and the fixed node energy \( E_{\theta}^{\text{FN}} \leq E_{\overline{\theta}}^{\text{F}} \) associated with \( \psi_0 \) can be computed using a Monte-Carlo method associated with (1.9). However, a parameter \( \theta \) optimizing the energy \( E_{\theta}^{\text{F}} \) of the trial wave function does not in general, for a given parameterization, optimize the fixed node groundstate energy \( E_{\theta}^{\text{FN}} \).

An open problem is now to develop an algorithm that can directly minimize \( E_{\theta}^{\text{FN}} \). To precise this idea, let us consider the formulation of the exact Fermionic groundstate \( (\psi_0^F, E_{\theta}^F) \) as a variational problem involving the fixed node groundstate \( (\psi_0^{\text{FN}}, E_{\theta}^{\text{FN}}) \) and the nodal surface \( \partial N_{\theta} \), that is to say:

\[ E_{\overline{\theta}}^{\text{F}} := \inf (E_{\theta}^{\text{FN}}, \psi_0^{\text{FN}} \text{ solution of (1.15), } \psi_0^F \text{ skew-symmetric}) = \frac{\int_{\mathbb{R}^d} \psi_0^F H (\psi_0^F)}{\int_{\mathbb{R}^d} (\psi_0^F)^2} \]  

(1.16)

An approach to solve the variational problem (1.17) consists in the computation of the shape derivative of the fixed node groundstate

\[ \nabla_{\theta} E_{\theta}^{\text{FN}} \]  

(1.18)

using a Monte-Carlo estimation based on (1.12). In this context, the key formulas (1.11)–(1.12) can be rewritten as follows. \( t \mapsto W_{\tau}^{+} \) denotes a Brownian motion in \( N_{\theta}^+ \), and \( \tau^+ \) the hitting time of \( \partial N_{\theta} \). Then the probability measure \( d\eta_0^{\text{FN}} \) on \( N_{\theta}^+ \) is defined by:

\[ \int_{N_{\theta}^+} \varphi d\eta_0^{\text{FN}} := \lim_{T \to +\infty} \frac{\mathbb{E} \left( \varphi (W_{T}^{+}) \mathbb{I} \beta_{T} \leq T + e^{-\int_{0}^{T} V (W_{s}^{+}) ds} \right)}{\mathbb{E} \left( \mathbb{I} \beta_{T} \leq T + e^{-\int_{0}^{T} V (W_{s}^{+}) ds} \right)}, \]  

(1.19)

and the measure \( d\mu_{\theta, \lambda}^{\text{FN}} \) on \( \partial N_{\theta} \) for \( \lambda < E_{\theta}^{\text{FN}} \) is defined by:

\[ \int_{\partial N_{\theta}} \varphi d\mu_{\theta, \lambda}^{\text{FN}} := \lim_{T \to +\infty} \frac{\mathbb{E} \left( \varphi (W_{T}^{+}) \mathbb{I} \beta_{T} \leq T + e^{-\int_{0}^{T} (V (W_{s}^{+}) - \lambda) ds} \right)}{\mathbb{E} \left( \mathbb{I} \beta_{T} \leq T + e^{-\int_{0}^{T} (V (W_{s}^{+}) - \lambda) ds} \right)}. \]  

(1.20)
We will show that (1.12) becomes:

\[
\nabla_{\theta} E_{\theta}^{\text{FN}} = \frac{(E_{\theta}^{\text{FN}} - \lambda)}{I_{N_\theta}^+ \psi_{\theta, \lambda}^{\text{FN}} \partial N_\theta} \int_{\partial N_\theta} r_\theta^+ \nabla^+ \psi_{\theta}^{\text{FN}} \cdot n_+ d\mu_{\theta, \lambda}^{\text{FN}} \\
= \frac{2(E_{\theta}^{\text{FN}} - \lambda)}{I_{N_\theta}^+ \psi_{\theta, \lambda}^{\text{FN}} \partial N_\theta} \int_{\partial N_\theta} r_\theta^+ \nabla_{\text{sym}} \psi_{\theta}^{\text{FN}} \cdot n_+ d\mu_{\theta, \lambda}^{\text{FN}}.
\]

(1.21)

In the above, \( r_\theta^+ \) is the shape derivative of \( N_\theta^+ \), \( n_+ \) is the associated exterior normal vector, and \( \nabla_{\text{sym}} \psi_{\theta}^{\text{FN}} \cdot n_+ \) is the symmetrization of the normal groundstate gradient, that is to say:

\[
\nabla_{\text{sym}} \psi_{\theta}^{\text{FN}} \cdot n_+ = \frac{1}{2} \left( \nabla^+ \psi_{\theta} \cdot n_+ - \nabla^- \psi_{\theta} \cdot n_- \right),
\]

(1.22)

where \( \nabla^+ \psi \cdot n_+ \) (resp. \( \nabla^- \psi \cdot n_- \)) is the exterior normal derivative in \( N_\theta^+ \) (resp. \( N_\theta^- \)). By construction, \( r_\theta^+ \) is a skew-symmetric field on \( \partial N_\theta \), and \( \nabla_{\text{sym}} \psi_{\theta} \cdot n_+ \) is symmetric. Our main result concerns then the link between (i) a symmetry breaking of the measure \( \mu_{\theta, \lambda}^{\text{FN}} \), (ii) the fact that \( \psi_{\theta}^{\text{FN}} \) is an eigenfunction, and (iii) local extrema of \( \theta \to E_{\theta}^{\text{FN}} \). The link between (i), (ii) and (iii) can be stated through the following equivalent assertions:

1. The measure \( \mu_{\theta, \lambda}^{\text{FN}} \) on \( \partial N_\theta \) is symmetric (i.e. invariant by the action of \( S \)).
2. When defined on the whole space \( \mathbb{R}^d \), the gradient of the fixed node groundstate \( \nabla \psi_{\theta}^{\text{FN}} \) is continuous on \( \partial N_\theta \).
3. The fixed node groundstate \( \psi_{\theta}^{\text{FN}} \) is a skew-symmetric eigenfunction of \( H \) on \( \mathbb{R}^d \).
4. The fixed node energy variation vanishes \( \nabla_{\theta} E_{\theta}^{\text{FN}} = 0 \) for any parameterization \( \theta \to \partial N_\theta \) of the nodal surface.

This yields a probabilistic characterization of the nodes (set of zeros) of skew-symmetric eigenstates of \( H \) through a symmetry argument. This is an original result. The practical computation of \( \nabla_{\theta} E_{\theta}^{\text{FN}} \) using (1.21) requires a Monte-Carlo estimator of the measures \( (\mu_{\theta, \lambda}^{\text{FN}}, \eta_\theta^{\text{FN}}) \) on the one hand, and an analytical approximation of the fixed node groundstate elements \( (\psi_{\theta}^{\text{FN}}, \nabla_{\text{sym}} \psi_{\theta}^{\text{FN}} \cdot n_+) \) on the other hand (see also (1.12)). A key remark is that the elements \( (\psi_{\theta}^{\text{FN}}, \nabla_{\text{sym}} \psi_{\theta}^{\text{FN}} \cdot n_+) \) share the same symmetry properties. This suggests the following approximation of the energy variation \( \nabla_{\theta} E_{\theta}^{\text{FN}} \) in (1.21):

\[
\nabla_{\theta} E_{\theta}^{\text{FN}} \sim \overline{\nabla_{\theta} E_{\theta}^{\text{FN}}} = \frac{2(E_{\theta}^{\text{FN}} - \lambda)}{I_{N_\theta}^+ \psi_{\theta, \lambda}^{\text{FN}} \partial N_\theta} \int_{\partial N_\theta} r_\theta^+ \psi_{\theta, \lambda}^{\text{FN}} \cdot n_+ d\mu_{\theta, \lambda}^{\text{FN}}
\]

(1.23)

\[
= \frac{2(E_{\theta}^{\text{FN}} - \lambda)}{I_{N_\theta}^+ \psi_{\theta, \lambda}^{\text{FN}} \partial N_\theta} \int_{\partial N_\theta} \nabla_{\theta} \psi_{\theta, \lambda}^{\text{FN}} \cdot n_+ d\mu_{\theta, \lambda}^{\text{FN}}
\]

(1.24)

\[
= \int_{N_\theta^+} \overline{H - E_{\theta}^{\text{FN}}} \left( \nabla_{\theta} \psi_{\theta, \lambda}^{\text{FN}} \right) d\eta_\theta^{\text{FN}}
\]

(1.25)

In the same way as for the exact expression \( \nabla_{\theta} E_{\theta}^{\text{FN}} \), the latter vanishes (\( \overline{\nabla_{\theta} E_{\theta}^{\text{FN}}} = 0 \)) if 1, 2, or 3 holds. In return, if \( \nabla_{\theta} E_{\theta}^{\text{FN}} = 0 \) for any parameterization of the nodes then 1, 2 or 3 holds. Such algorithms may be referred to as Nodal Monte-Carlo. They will require variance reduction techniques exploiting the symmetry structure, in principle such that the variance of \( \nabla_{\theta} E_{\theta}^{\text{FN}} \) scales appropriately to 0 in the limit where \( \mu_{\theta, \lambda}^{\text{FN}} \) becomes symmetric (in other words, we seek for an ‘asymptotically scaling variance reduction’, also called ‘zero-variance estimation’ in the QMC literature). Ideas are given for future work on this matter.
Let us now position the content of this paper in the context of the QMC literature. First, we recall that efficiently computing in high dimension $d \gg 1$ the Fermionic groundstate (1.13) using Monte-Carlo methods is a fundamental problem with many applications; for instance it amounts to solve the eigenvalue problem for excited eigenstates, where classical power methods fail. A general solution is known to be intractable, and is usually referred to as the *sign problem* (see Rem. 4.2). This explains the necessity of the fixed node approximation. The issue of optimizing the nodes of the trial wave function $\psi_\theta^I$ in the fixed node approximation was pointed out in [9], where a long discussion on the structure of Fermion nodes and appropriate (from this perspective) trial wave functions is provided. Yet state-of-the art numerical methods optimizing the trial wave function is based on either, (i) the minimization of the VMC energy $E_\theta^I$ in (1.16), as in [27,29]; or (ii) the minimization of the variance of the local energy

$$E_L := V - (\psi_\theta^I)^{-1}\frac{\Delta}{2}(\psi_\theta^I)$$

(1.26)

as in [22,29]. As a consequence, an efficient method optimizing directly the fixed node energy $E_{\theta,R}^I$ in (1.15), that is to say the nodes of the trial wave function $\psi_{\theta,R}^I$, remains an unsolved problem and motivates the content of this paper. However, methods to approximately compute the gradient $\nabla_\theta E_{\theta,R}^I$ were already suggested in the QMC literature in the more general context of the calculation of physical properties (or “forces”). The main goal is to compute the derivative $\nabla_R E_{\theta,R}^I$ of the groundstate energy with respect to some parameter $R$ parameterizing the Hamiltonian, typically the potential energy $V \equiv V_R$. A classical example is the following: $R$ is the vector of the nuclei-nuclei distances in molecules. The trial wave function now also depends on $R$: $\psi_\theta^I \equiv \psi_{R,\theta}^I$. Well established methods are available for the exact and variance reduced computation of the variational energy gradient:

$$\nabla_R E_{\theta,R}^I$$

(1.27)

where $E_{\theta,R}^I$ is defined by (1.16). For instance, methods using coupling (correlated sampling) to reduce variance were tackled in [18], and a general construction of variance reduced estimators (‘zero variance-zero bias’ estimators) was proposed in [2]; see also some applications in [28]. The case of the fixed node energy gradient:

$$\nabla_R E_{\theta,R}^{\text{FN}}$$

(1.28)

where $E_{\theta,R}^{\text{FN}}$ is defined by (1.15) is more intricate, and requires some approximation since the derivatives $\nabla_R E_{\theta,R}^{\text{FN}}$ of the fixed node groundstate remain unknown. Note that computing the gradients $\nabla_\theta E_\theta^I$ or $\nabla_\theta E_{\theta,R}^{\text{FN}}$ can be seen as a particular case of computing respectively (1.27) or (1.28), since the nodes of the trial wave function $\psi_{R,\theta}^I$ depend on $R \ a \ priori$. An approximate formula has already been proposed to compute (1.28), see for instance equation (54) in [2], (10) in [8], (10) in [4], and (14) in [5]. The terms due to the variation of the nodes (which amounts to evaluate $\nabla_\theta E_{\theta,R}^{\text{FN}}$) are called ‘nodal Pulay terms’, and were particularly pointed out in [4,5]. In the above references, the formula used for the calculation is exactly the formula (1.25). However, the interpretation in terms of a stochastic process stopped on the nodes in (1.23)–(1.24), the analysis of the symmetry of the associated distribution on the nodes, and the suggestion of an associated Nodal Monte Carlo method are new results.

The following classical textbooks are recommended:

- about spectral theory of operators: [25];
- about elliptic theory of Partial Differential Equations: [20];
- about random processes and Feynman-Kac representations: [23];
- about Monte-Carlo methods in Quantum Chemistry (QMC): [1,21].
2. Shape derivatives of Dirichlet groundstates

In this section, some notations and results are recalled concerning Dirichlet groundstates of Schrödinger operators with a smooth potential. Then, formula (1.4) is proven formally, and references are given for rigorous proofs.

Consider the Schrödinger operator (1.1) defined on \( \mathbb{R}^d \) with a smooth potential \( V \) bounded from below. \( H \) defines a self-adjoint operator on the Hilbert space \( L^2(\mathbb{R}^d) \). For simplicity, \( V \) is assumed to go to infinity at infinity such that \( H \) has a compact resolvent, and thus a purely discrete spectrum. Generalization to operators involving continuous spectrum, although of fundamental importance in Quantum Chemistry, is left as technical extensions of the present work. Then a parameterization of smooth domains (1.2) is considered for \( \theta \in \mathbb{R}^p \) such that \( \partial \Omega_\theta \) has a smooth boundary for any \( \theta \). One assumes then that there exists a set of diffeomorphisms smoothly indexed by \( \theta \) and such that:

\[
\begin{align*}
(\theta, x) &\mapsto R_\theta(x) \text{ is smooth in } \mathbb{R}^p \times \mathbb{R}^d \\
R_0 &= \text{Id} \\
R_\theta &= \text{Id} \text{ outside some compact set} \\
\forall \theta \in \mathbb{R}^p, \quad \Omega_\theta &= R_\theta(\Omega_0).
\end{align*}
\]

(2.1)

In this setting, the shape derivative of \( \theta \mapsto \Omega_\theta \) can be defined as the smooth field:

\[
r_\theta : \partial \Omega_\theta \to \mathbb{R}^p,
\]

verifying:

\[
\forall x \in \partial \Omega_\theta, \quad r_\theta(x) = \nabla_\theta R_\theta(x) \cdot n(x),
\]

and thus locally for small \( h \in \mathbb{R}^p \):

\[
\partial \Omega_{\theta+h} \sim \{ x + h \cdot r_\theta(x)n(x) \mid x \in \partial \Omega_\theta \}.
\]

Now classical results of spectral theory ensures that the Hamiltonian (1.1) considered in \( L^2(\Omega_\theta) \) with Dirichlet boundary condition, is self-adjoint with domain \( \mathcal{D}(H) \subset H^1_0(\Omega_\theta) \), where \( H^1_0(\Omega_\theta) \) is the usual Sobolev space of function with Dirichlet conditions and square integrable first order derivatives. Moreover, \( H \) has a unique (up to a multiplicative constant) signed groundstate \( \psi_0^* \), solution of the variational problem (1.3), or equivalently solution of the eigenvalue problem:

\[
\begin{align*}
H(\psi_0^*) &= E_0^* \psi_0^* \\
\psi_0^*|_{\partial \Omega_\theta} &= 0 \\
\psi_0^* > 0,
\end{align*}
\]

where \( \cdot|_{\partial \Omega_\theta} \) denotes the usual trace operator on the boundary. The regularity of \( V \) then ensures that \( \psi_0^* \) is smooth on \( \Omega_\theta \) and that \( \nabla \psi_0^* \cdot n \) is smooth on \( \partial \Omega_\theta \). The following derivative formula can now be stated.

Lemma 2.1. Consider domains \( \theta \mapsto \Omega_\theta \) verifying (2.1). Then the Dirichlet energy \( E_0^* \) solution of the variational problem (1.3) is differentiable with respect to the parameter \( \theta \), and the variation formula (1.4) holds.

Proof. The formal computation is detailed, and references are given for the rigorous result, which is not much more involved but less instructive from our point of view. Fix \( \theta_0 \in \mathbb{R}^p \), and consider the formal chain rule:

\[
(\nabla_\theta E_0^*)_{\theta=\theta_0} = \left( \nabla_\theta \int_{\Omega_\theta} \frac{\psi_0^* H(\psi_0^*)}{(\psi_0^*)^2} \right)_{\theta=\theta_0} = \left( \nabla_\theta \int_{\Omega_\theta} \psi_0^* H(\psi_0^*) \right)_{\theta=\theta_0} \cdot \frac{1}{\left( \int_{\Omega_\theta} (\psi_0^*)^2 \right)_{\theta=\theta_0}}.
\]
ON A PROBABILISTIC INTERPRETATION OF SHAPE DERIVATIVES

Since \((\psi_0^*, E_0^*)\) is an eigenelement, it yields:

\[
\nabla_\theta \int_{\Omega_0} \psi_0^* H (\psi_0^*) \int_{\Omega_0} (\psi_0^*)^2 = \nabla_\theta E_0^* = 0.
\]

Then formal differentiation yields

\[
\nabla_\theta \int_{\Omega_0} \psi_0^* H (\psi_0^*) \int_{\Omega_0} (\psi_0^*)^2 = \int_{\Omega_0} \nabla_\theta \psi_0^* (H - E_0^*) (\psi_0^*) + \int_{\Omega_0} \psi_0^* (H - E_0^*) (\nabla_\theta \psi_0^*)
\]

and since \((\psi_0^*, E_0^*)\) is an eigenelement, the first term of the right hand side vanishes so that finally:

\[
(\nabla_\theta E_0^*)_{\theta = \theta_0} = \int_{\Omega_0} \psi_0^* \left(-\frac{1}{2} \Delta + V - E_0^*\right) (\nabla_\theta \psi_0^*) \int_{\Omega_0} (\psi_0^*)^2.
\]

Now applying Green’s integration by parts two times, and remarking that \((H - E_0^*) (\psi_0^*) = 0\), we get

\[
(\nabla_\theta E_0^*)_{\theta = \theta_0} = \frac{1}{2} \int_{\partial \Omega_0} \nabla \psi_0^* \cdot n \nabla_\theta \psi_0^*\ d\sigma
\]

Then (1.5) applied to \(\psi_0^*\) yields the result. The rigorous proof can be made using a change a variable with the diffeomorphism \(R_\theta\), and then exploiting the variational formulation (see e.g. Thm. 2 in [12]).

\[\square\]

3. Probabilistic interpretations

In this section, the probabilistic formulations (1.8)-(1.9)-(1.11) are proven and detailed. Associated Monte-Carlo methods are recalled with some references.

Consider notations and assumptions of Section 2. Let \(t \mapsto W_t\) be a standard Wiener process with exit time \(\tau\) from \(\Omega_0\) defined in (1.6). The classical probabilistic interpretation of the eigenelements \((\psi_0^*, E_0^*)\) is recalled in the following lemma:

**Lemma 3.1.** Assume \(W_0\) is distributed according to \(\frac{\psi_{\text{init}}(x)\ dx}{\int \psi_{\text{init}}(x)\ dx}\) where \(\psi_{\text{init}} \in L^2(\Omega_0)\), and \(\psi_{\text{init}}\) is non vanishing in each connected component of \(\Omega_0\). The groundstate elements can be expressed through the long time behavior of the process with Feynman-Kac weights and conditioned by large exit times. This yields the formulas (1.7)–(1.9).

**Proof.** The result follows from the classical representation of parabolic equations through the Feynman-Kac formula (see [23]). Let us recall the different steps of the argument. First, consider \(\varphi \in C^\infty_c(\Omega_0)\) a smooth solution \((t, x) \mapsto u_t(x) \in C^\infty(\mathbb{R}^+ \times \Omega_0)\) of the parabolic problem:

\[
\begin{align*}
\partial_t u_t &= -H(u_t) \\
u_t|_{\partial \Omega_0} &= 0 \\
u_0 &= \varphi.
\end{align*}
\]
Then using Itô calculus, it yields for any $\phi \in C^\infty(\mathbb{R}^+ \times \Omega_\theta)$:

$$d\left(\phi_{T-t}(W_t)e^{-\int_t^T V(W_s)\,ds}\right) = \left(\frac{\Delta}{2} - V - \partial_t\right)\left(\phi_{T-t}(W_t)e^{-\int_t^T V(W_s)\,ds}\right)dt + e^{-\int_t^T V(W_s)\,ds}\nabla\phi_{T-t}(W_t)\cdot dW_t,$$

(3.1)

so that applying the latter computation to $\phi = u$ between time 0 and the stopping time $T \wedge \tau = \inf(T, \tau)$ yields:

$$\mathbb{E}\left(\int_{\Omega_\theta} u_T(x)\psi_{\text{init}}(x)\,dx\right) = \mathbb{E}\left(\int_{\Omega_\theta} \psi_{\text{init}}(x)\,dx\right) = \mathbb{E}\left(\int_{\Omega_\theta} \psi_{\text{init}}(x)\,dx\right) = \mathbb{E}\left(\varphi(W_T)\mathbb{I}_{\tau \geq T} e^{-\int_0^T V(W_s)\,ds}\right).$$

Now denoting by $(\psi_\theta^{\ast,n}, E_\theta^{\ast,n})_{n \geq 0}$ the full spectrum of $H$ normalized in $L^2(\Omega_\theta)$, the above Feynman-Kac representation reads:

$$\mathbb{E}\left(\varphi(W_T)\mathbb{I}_{\tau \geq T} e^{-\int_0^T V(W_s)\,ds}\right) = \sum_{n \geq 0} e^{-E_\theta^{\ast,n}T} \int_{\Omega_\theta} \mathbb{E}\left(\psi_\theta^{\ast,n}\psi_{\text{init}}\right) \int_{\Omega_\theta} \psi_\theta^{\ast,n}\varphi. \tag{3.2}$$

Now using the fact that the groundstate has a spectral gap ($E_\theta^{1} > E_\theta^{0} = E_\theta^\ast$), the dominant term in (3.2) when $T \to +\infty$ enables to verify that the groundstate has a sign $\psi_\theta^{\ast,0} > 0$. Finally, taking the limit $\varphi \to \mathbb{I}$ by dominated convergence in the formula above, and then the leading term in the limit $T \to +\infty$ yields (1.9), (1.8) follows.

In practice however, a diffusion solution to a stochastic differential equation with a repulsive drift at the boundary $\partial \Omega$ is used:

$$\begin{cases}
\frac{dX_t}{dt} = (\psi^1)^{-1}\nabla\psi^1(X_t)dt + dW_t \\
\psi^1|_{\partial \Omega_\theta} = 0,
\end{cases} \tag{3.3}$$

where $\psi^1 > 0$ is a smooth function strictly positive in $\Omega_\theta$ and vanishing on $\partial \Omega_\theta$. In [6], sufficient conditions on the behavior of $\psi^1$ near $\partial \Omega_\theta$ and at infinity are given for (3.3) to be well-posed, and to verify the following variant of (1.8)–(1.9):

$$\lim_{T \to +\infty} \frac{\mathbb{E}\left(\varphi(X_T)e^{-\int_0^T E_L(X_s)\,ds}\right)}{\mathbb{E}\left(e^{-\int_0^T E_L(X_s)\,ds}\right)} = \int_{\Omega_\theta} \varphi \psi_0^\ast \psi^1, \tag{3.4}$$

$$\lim_{T \to +\infty} -\frac{1}{T} \ln \mathbb{E}\left(e^{-\int_0^T E_L(X_s)\,ds}\right) = E_0^\ast, \tag{3.5}$$

where the so-called local energy is defined by (1.26). The proof of the latter probabilistic interpretation (3.4)–(3.5) is similar to the proof of Lemma 3.1, and is based on the mapping of the Hamiltonian $H$ to a weighted $L^2$ space through:

$$H_t = (\psi^1)^{-1}H(\psi^1) = -\frac{\Delta}{2} - (\psi^1)^{-1}\nabla \psi^1 \cdot \nabla + E_L.$$

Details and assumptions can be found in [6].

Next the probabilistic interpretation of the shape derivative given by formula (1.11) is proven.

**Proposition 3.2.** Assume $W_0$ is distributed according to $\frac{\psi_{\text{init}}(x)\,dx}{\int \psi_{\text{init}}(x)\,dx}$ where $\psi_{\text{init}} \in L^2(\Omega_\theta)$, and $\psi_{\text{init}}$ is non vanishing in each connected component of $\Omega_\theta$. Assume the boundary $\partial \Omega_\theta$ is smooth and uniformly Lipschitz, and consider the measure defined for any $\lambda < E_0^\ast$ by (1.10). Then (1.11) holds.
ON A PROBABILISTIC INTERPRETATION OF SHAPE DERIVATIVES

Proof. Step 1. Let $\varphi \in C^\infty(\overline{\Omega}_\theta)$. We claim that the parabolic differential equation with inhomogeneous Dirichlet conditions

$$\begin{cases}
\partial_t h_t(\varphi) = -(H - \lambda) (h_t(\varphi)) \\
h_t(\varphi)|_{\partial \Omega_\theta} = \varphi|_{\partial \Omega_\theta} \\
h_0(\varphi) = \varphi
\end{cases} \quad (3.6)$$

has a unique smooth solution for $\lambda < E^*$ converging exponentially fast towards $h_\infty(\varphi)$ unique smooth solution of the elliptic inhomogeneous Dirichlet problem:

$$\begin{cases}
(H - \lambda) (h_\infty(\varphi)) = 0 \\
h_\infty(\varphi)|_{\partial \Omega_\theta} = \varphi.
\end{cases} \quad (3.7)$$

This is a classical consequence of spectral theory, but we recall briefly the basic arguments. Existence of a smooth solution of (3.7) follows from the fact that $(H - \lambda)^{-1}$ can be extended to a bounded operator of $L^2(\Omega_\theta)$ so that:

$$h_\infty(\varphi) = (H - \lambda)^{-1} (H - \lambda) (\varphi) - \varphi,$$

which indeed is solution of (3.7). Note that in the above $(H - \lambda)^{-1}$ implicitly refer to the operator with homogeneous Dirichlet boundary condition, so that $(H - \lambda)^{-1} \circ (H - \lambda) \neq \text{Id}$ when operating on test function with inhomogeneous boundary conditions. Then the homogeneous solution of (3.6) has been solved using spectral decomposition in (3.2), proving that such homogeneous solution is unique and vanishes exponentially fast when $\lambda < E^*_\theta$. The latter analysis of the homogeneous case proves uniqueness in (3.6)–(3.7), as well as exponential long time convergence of the time dependent equation (3.6).

Step 2. We claim that for any $\varphi \in C^\infty(\overline{\Omega}_\theta)$,

$$\int_{\Omega_\theta} h_\infty(\varphi) \psi_{\text{init}} \overline{\psi}_{\text{init}} := \mathbb{E} \left( \varphi(W_T) \mathbb{I}_{T < +\infty} e^{-\int_0^T (V(W_s) - \lambda)ds} \right) \quad (3.8)$$

where $h_\infty(\varphi)$ is the solution to the elliptic partial differential equation with inhomogeneous Dirichlet conditions (3.7). Indeed, using (3.1) with test function $h_{T-t}(\varphi)$, and stopped at time $T \wedge \tau = \inf(T, \tau)$ yields:

$$\int_{\Omega_\theta} h_T(\varphi) \psi_{\text{init}} \overline{\psi}_{\text{init}} := \mathbb{E} \left( h_{T \wedge \tau}(\varphi) (W_{T \wedge \tau}) e^{-\int_0^{T \wedge \tau} (V(W_s) - \lambda)ds} \right).$$

Now the event $\mathbb{I}_{\tau = +\infty}$ has null probability if $t \mapsto W_t$ is recurrent, and if the latter is transient then using $\lim V = +\infty$:

$$\lim_{T \to +\infty} \mathbb{E} \left( \mathbb{I}_{T = +\infty} e^{-\int_0^T (V(W_s) - \lambda)ds} \right) = 0;$$

so that taking the limit $T \to +\infty$ leads to (3.8).

Step 3. We claim that

$$\int_{\Omega_\theta} h_\infty(\varphi) \psi^*_\theta = -\frac{1}{2(E^*_\theta - \lambda)} \int_{\partial \Omega_\theta} \varphi \nabla \psi^*_\theta \cdot n d\sigma.$$

Since $\psi^*_\theta$ is the groundstate:

$$\psi^*_\theta = \frac{1}{E^*_\theta - \lambda} (H - \lambda) (\psi^*_\theta),$$
so that integration by parts yields:

$$
\int_{\Omega} h_\infty(\varphi) \psi^* \, d\sigma = -\frac{1}{E_0^* - \lambda} \int_{\partial\Omega_0} \frac{1}{2} \varphi \nabla \psi^* \cdot n \, d\sigma + \frac{1}{E_0^* - \lambda} \int_{\Omega_0} \frac{1}{2} \varphi \nabla h_\infty(\varphi) + (V - \lambda) h_\infty(\varphi) \psi^* \\
= -\frac{1}{E_0^* - \lambda} \int_{\Omega_0} \frac{1}{2} \varphi \nabla \psi^* \cdot n \, d\sigma + 0.
$$

**Step 4.** Now using the Markov property of Brownian motion yields:

$$
\mathbb{E}\left( \varphi(W_T) 1_{T < \tau < +\infty} e^{-\int_0^T (V(W_s) - \lambda) \, ds} \right) = \mathbb{E}\left( \varphi(W_T) 1_{\tau < +\infty} e^{-\int_0^\tau (V(W_s) - \lambda) \, ds} \mid W_0 = W_T \right) \\
\times 1_{T \leq \tau} e^{-\int_0^T (V(W_s) - \lambda) \, ds},
$$

where $$((\tilde{W}_t)_{t \geq 0}, \tilde{\tau})$$ is an independent copy of $$((W_t)_{t \geq 0}, \tau)$$. Together with Steps 2 and 3 with the probabilistic interpretation of the groundstate $$\psi^*$$ in (1.8)–(1.10), it completes the proof. □

4. **Fermion Groundstates**

Consider the Schrödinger operator (1.1). Fermionic groundstates $$(\psi^*_F, E^*_F)$$ associated to (1.1) are defined with respect to a finite symmetry group $$S \subset O(\mathbb{R}^d)$$ of the potential $$V$$, where $$O(\mathbb{R}^d)$$ denotes the group of isometries of $$\mathbb{R}^d$$. A symmetry group of $$V$$ is defined by the property

$$\forall S \in S \quad V \circ S = V.$$

Fermion systems appears in the special case when $$S$$ contains symmetries with odd parity:

$$\det(S) = -1,$$

so that $$H$$ can be defined (since the Laplacian operator commutes with isometries) to operate on the Hilbert space of skew-symmetric function:

$$H = \{ \psi \in L^2(\mathbb{R}^d) \mid \forall S \in S, \psi \circ S = \det(S) \psi \}.$$

“Fermion” groundstates are then the solutions to the variational problem (1.13).

**Example 4.1** (Fermions). In the case of $$N$$ physical quantum particles of Fermionic type in dimension 3, one has $$\mathbb{R}^d = \mathbb{R}^{3N}$$, and a potential of the form:

$$V(x) := \sum_{i=1}^N V_{\text{ext}}(x_i) + \sum_{1 \leq i < j \leq N} V_{\text{int}}(x_i - x_j),$$

where $$x = (x_1, \ldots, x_N)$$ are the 3 dimensional coordinates of the particles, $$V_{\text{ext}}$$ is an exterior smooth potential that goes to infinity at infinity, and $$V_{\text{int}}$$ a smooth interaction potential that vanishes at infinity. Then the discrete symmetry group is the permutation group $$S_N$$ of physical particles. See [6] for the mathematical analysis of Quantum Monte-Carlo (QMC) methods in this context.

Since in this paper, we restrict to smooth operators with compact resolvent, the existence of $$(\psi^*_F, E^*_F)$$ follows directly from the fact that the spectrum is discrete, and smoothness from the fact $$\psi^*_F$$ satisfies an eigenvalue problem of a smooth elliptic operator (see [20]).
Remark 4.2 (the sign problem). Computing directly \((\psi^*, E^*_F)\) using Monte-Carlo methods is an intractable problem known as the sign problem. The latter can be summarized as follows. Since \(H\) leaves invariant \(\mathcal{H}\), eigenfunctions of \(H\) in \(\mathbb{R}^d\) are either symmetric or skew-symmetric functions. Thus computing (3.2) with:

- a skew-symmetric test function \(\psi\),
- a non-symmetric positive initial condition \(\psi_{\text{init}} > 0\),

yields:

\[
\mathbb{E}\left(\psi^I(W_T) e^{-\int_0^T V(W_s) ds}\right) = \sum_{\psi^{*,n} \in \mathcal{H}} e^{-E^*,n_T} \int_{\mathbb{R}^d} \psi^{*,n}_{\text{init}} \int_{\mathbb{R}^d} \psi^{*,n}_I, \]

and in principle if \(\varphi\) is symmetric:

\[
\lim_{T \to +\infty} \frac{\mathbb{E}\left(\varphi(W_T) \psi^I(W_T) e^{-\int_0^T V(W_s) ds}\right)}{\mathbb{E}\left(\psi^I(W_T) e^{-\int_0^T V(W_s) ds}\right)} = \int_{\mathbb{R}^d} \psi^*_F \psi^{\dagger}\varphi \int_{\mathbb{R}^d} \psi^*_F \psi^{\dagger},
\]

Unfortunately, this computation relies crucially on the rate of vanishing of the normalization which is due to the non-symmetry of the initial condition \(\psi_{\text{init}} > 0\). Since stochastic processes used to compute such averages quickly forget their initial condition and have a symmetric dynamics in the full state space \(\mathbb{R}^d\), one is compelled to compute ratios of vanishing averages (of the type \(\frac{0}{0}\)), with Monte-Carlo estimators having a non-vanishing statistical variance. This leads to infinite variance when \(T\) is large. This forms the sign problem. This problem appears more generally when trying to solve higher eigenvalue problems with Monte-Carlo methods. Although, there is probably no general solution, solving the sign problem for particular situations in high dimension would be considered as a major breakthrough.

In practice, \((\psi^*_F, E^*_F)\) can be approximated using a hybrid methodology in two steps. First, a trial wave function is obtained using an analytical parameterization of \(\mathcal{H}\), usually of the form:

\[
\psi^I_{\alpha, \theta} := J_{\alpha} \psi^{\text{skew}}_{\theta}. \tag{4.1}
\]

In the above, \(J_{\alpha} > 0\) is a strictly positive symmetric part called the Jastrow factor and parameterized by \(\alpha \in \mathbb{R}^n\); and \(\psi^{\text{skew}}_{\theta}\) is the skew-symmetric part parameterized by \(\theta \in \mathbb{R}^p\). The most convenient numerical optimization method is then used to solve (1.13) in the space of parameters formed by \((\alpha, \theta)\), and the solution of the optimization procedure is denoted with the parameters \((\alpha_0, \theta_0)\). This yields the set of functions \(\{\psi^I_{\theta}\}_{\theta \in \mathbb{R}^p}\) defining the nodal domains through

\[
\psi^I_{\theta} := J_{\alpha_0} \psi^{\text{skew}}_{\theta}. \]

Example 4.3 (Fermions). In the case of \(N\) physical particles of Fermionic type, \(\psi^{\text{skew}}_{\theta}\) is built using a sum of Slater determinant, that is to say a sum of functions of the form:

\[
\det (\phi_j(x_i))_{i,j=1 \ldots N},
\]

where \((\phi_j)_{j=1 \ldots N}\) are \(N\) smooth functions of \(\mathbb{R}^3\).

The link with Sections 2 and 3 is made by posing:

\[
\Omega_{\theta} := \mathcal{N}_{\theta}^+ \cup \mathcal{N}_{\theta}^-.
\]

The Fixed Node Approximation consists in computing with a Monte-Carlo method the solution \((\psi^{FN}_{\theta}, E^{FN}_{\theta}) = (\psi^*_\theta, E^*_\theta)\) of the variational problem with Dirichlet conditions (1.15). Such a computation is made using
the probabilistic interpretations (1.8)–(1.9), or more usually in practice using the variant (3.4)–(3.5). This method is known under the DMC acronym (Diffusion Monte-Carlo) in Computational Chemistry, and has been widely studied, see for instance [3,21].

As explained in the introduction, the key problem is that the nodal surface $\partial N_\theta = (\psi_I^\dagger)^{-1}(0)$ may be different from $(\psi^\ast F)^{-1}(0)$. The open question is thus now to carry out numerical methods associated to the variational problem (1.17). In this context, a direct minimization of $E_{FN}^\theta$ in (1.17) requires the computation of:

$$\nabla_\theta E_{FN}^\theta$$

using formula (1.4) and the probabilistic interpretation (1.11).

5. Fixed node and symmetry breaking

In this section, we consider the context of Section 4, and we assume that the nodal surface defined in (1.14) is a smooth manifold, a sufficient condition being:

$$\forall x \in \partial N_\theta, \quad \nabla \psi_I^\dagger(x) \neq 0. \quad (5.1)$$

Moreover, we assume that the assumptions of Section 2 on the mapping $\theta \mapsto N_\theta$ apply, namely that there is a diffeomorphic mapping $\theta \mapsto R_\theta$ associated to the domain $N_\theta$ such that (2.1) holds.

A fundamental remark concerns the symmetry of the normal derivative of the fixed node groundstate $\nabla^+ \psi_{FN}^\dagger \cdot n_+$ (where $\nabla^+$ and $n_+$ refer to the domain $N_\theta^+$, and will be defined below), or equivalently the measure $\mu_{\theta, \lambda}^\ast$ in (1.11) defined on $\partial N_\theta$ for stochastic processes evolving in $N_\theta^+$. The latter indeed presents a symmetry breaking, in the sense that they are only invariant by the action of the symmetry sub-group

$$S^+ := \{ S \in S \mid \det(S) = 1 \}$$

on $\partial N_\theta$. Before going further, we will precise notations in appropriate definitions and lemmas.

**Lemma 5.1.** $S$ is a symmetry group of the nodal surface $\partial N_\theta$ in the sense that any space transformation $S \in S$ verify:

$$S(\partial N_\theta) = \partial N_\theta.$$ 

**Proof.** By skew-symmetry, it yields for any $x \in \mathbb{R}^d$

$$\psi_I^\dagger(S(x)) = -\psi_I^\dagger(x)$$

so that $\psi_I^\dagger(x) = 0$ is equivalent to $\psi_I^\dagger(S(x)) = 0$. □

So consider now $\psi$ a skew-symmetric function in $\mathbb{R}^d$ with Dirichlet boundary conditions on the nodes $\psi|_{\partial N_\theta} = 0$, and such that the restrictions $\psi|_{N_\theta^+} : N_\theta^+ \to \mathbb{R}$, on the one hand and $\psi|_{N_\theta^-} : N_\theta^- \to \mathbb{R}$, on the other hand are smooth on the associated closed domains. Remark that using skew-symmetry, for any $S \in S$ verifying $\det S = -1$, $\psi|_{N_\theta^-}$ is the image of $\psi|_{N_\theta^+}$ through the transformation:

$$\psi|_{N_\theta^-} = -\psi|_{N_\theta^+} \circ S.$$ 

Two traces of $\nabla \psi$ on $\partial N_\theta$ can then be defined depending if the reference domain is $N_\theta^+$ or $N_\theta^-$. The definition of associated symmetric and skew-symmetric traces then follows.
**Definition 5.2.** Let \( \psi \) a skew-symmetric function in \( \mathbb{R}^d \) with smooth restrictions and Dirichlet boundary conditions on \( \overline{N_\theta}^+ \) and \( \overline{N_\theta}^- \). \( n_+ \) denotes the exterior normal of \( \partial N_\theta^+ \), \( n_- \) denotes the exterior normal of \( \partial N_\theta^- \), so that:

\[
n_+ = -n_-.
\] (5.2)

\( \nabla^k \psi \cdot n_+ \) denotes the exterior normal derivative of \( \psi \) in \( \overline{N_\theta}^+ \), and \( \nabla^k \psi \cdot n_- \) the exterior normal derivative of \( \psi \) in \( \overline{N_\theta}^- \). Then the skew-symmetrization of the normal derivative is defined by

\[
\nabla^k \psi \cdot n = \frac{1}{2} (\nabla^k \psi \cdot n_+ + \nabla^k \psi \cdot n_-),
\] (5.3)

and the plain symmetrization with respect to \( n_+ \) is defined by (1.22).

Note that \( \psi \in C^1(\mathbb{R}^d) \) if and only if \( \nabla^k \psi \cdot n = 0 \) on \( \partial N_\theta \), so that \( \nabla^k \psi \cdot n \) can be seen as the gradient discontinuity of \( \psi \) at \( \partial N_\theta \). One can now precise the idea of symmetry breaking on \( \partial N_\theta \).

**Lemma 5.3.** Let \( \psi \) a skew-symmetric function in \( \mathbb{R}^d \) with smooth restrictions on \( \overline{N_\theta}^+ \) and \( \overline{N_\theta}^- \), and Dirichlet boundary conditions on \( \partial N_\theta \). Then \( \nabla^k \psi \cdot n \) is skew-symmetric, and \( \nabla^k \psi \cdot n_+ \) is symmetric, under the action of \( S \) on \( \partial N_\theta \). If \( \varphi \in C^\infty_c(\mathbb{R}^d) \), one has the integration by parts formula:

\[
\frac{1}{2} \int_{\mathbb{R}^d} \nabla \varphi \cdot \nabla \psi = \int_{\partial N_\theta} \varphi \nabla^k \psi \cdot n d\sigma - \int_{\mathbb{R}^d} \varphi \frac{\Delta}{2} \psi. \quad (5.4)
\]

Moreover, \( \nabla^k \psi \cdot n = 0 \) on \( \partial N_\theta \) if and only if \( \nabla^k \psi \cdot n_+ \) or \( \nabla^k \psi \cdot n_- \) are symmetric. In the opposite case, \( \nabla^k \psi \cdot n_+ \) and \( \nabla^k \psi \cdot n_- \) are invariant only under the action of the special sub-group \( S^+ \) and a “symmetry breaking” occurs.

**Proof.** By skew-symmetry of \( \psi \), one has for any \( S \in S \) in \( \overline{N_\theta}^+ \cup \overline{N_\theta}^- \):

\[
S \nabla \psi \circ S = \det(S) \nabla \psi,
\]

and then on \( \partial N_\theta \),

\[
Sn_+ \circ S = \det(S) n_+ \quad \text{and} \quad Sn_- \circ S = \det(S) n_-,
\]

so that \( S^T S = \text{Id} \), it yields on \( \partial N_\theta \) for any \( S \in S \):

\[
\begin{cases}
\nabla^k \psi \cdot n_+ \circ S = \nabla^k \psi \cdot n_+ & \text{if } \det(S) = 1 \\
\nabla^k \psi \cdot n_- \circ S = -\nabla^k \psi \cdot n_- & \text{if } \det(S) = 1 \\
\nabla^k \psi \cdot n_+ \circ S = -\nabla^k \psi \cdot n_- & \text{if } \det(S) = -1.
\end{cases}
\] (5.5)

This yields the symmetry properties of \( \nabla^k \psi \cdot n \) and \( \nabla^{sk} \psi \cdot n_+ \) with respect to \( S \), and of \( \nabla^k \psi \cdot n_+ \) and \( \nabla^k \psi \cdot n_- \) with respect to \( S^+ \).

The integration by parts formula is obtained by applying separately on \( \overline{N_\theta}^+ \) and \( \overline{N_\theta}^- \) the classical Green’s identity.

Finally from the symmetry properties (5.5), \( \nabla^k \psi \cdot n_- = -\nabla^k \psi \cdot n_+ \) on \( \partial N_\theta \) if and only if \( \nabla^k \psi \cdot n_+ \circ S = \nabla^k \psi \cdot n_+ \) or \( \nabla^k \psi \cdot n_- \circ S = \nabla^k \psi \cdot n_- \) for any \( S \in S \) with \( \det(S) = -1 \). \( \square \)

One can now apply these remarks to the fixed node groundstate \( \psi^{\text{FN}}_\theta \) solution of (1.15).
Lemma 5.4. Let \( \psi_{\theta}^{FN} \) be the solution of (1.15) with a smooth boundary \( \partial N_{\theta} \). Then the following assertions are equivalent:

1. \( \psi_{\theta}^{FN} \) is an eigenfunction of \( H \) in \( L^{2}(\mathbb{R}^{d}) \).
2. \( \nabla^{sk}\psi_{\theta}^{FN} \cdot n \), as defined by (5.3), vanishes on \( \partial N_{\theta} \).
3. \( \nabla^{+}\psi_{\theta}^{FN} \cdot n_{+} \) or \( \nabla^{-}\psi_{\theta}^{FN} \cdot n_{-} \) are symmetric under the action of \( S \) on \( \partial N_{\theta} \).

Proof. The third point and the second point are equivalent by Lemma 5.3. Now, if \( \varphi \in C^{\infty}_{c}(\mathbb{R}^{d}) \), integration by parts (5.4) yields:

\[
\int_{\mathbb{R}^{d}}(H - E_{\theta}^{FN})(\varphi)\psi_{\theta}^{FN} = \int_{\partial N_{\theta}} \varphi \nabla^{sk}\psi_{\theta}^{FN} \cdot n \, ds,
\]

so that \( \psi_{\theta}^{FN} \in L^{2}(\mathbb{R}^{d}) \) is an eigenfunction if and only if \( \nabla^{sk}\psi_{\theta}^{FN} \cdot n = 0 \) on \( \partial N_{\theta} \).

Proposition 5.5. The shape derivatives of the fixed node groundstate \( \psi_{\theta}^{FN} \) with respect to the nodal parameter \( \theta \) reads

\[
\nabla_{\theta}E_{\theta}^{FN} = -\frac{1}{2} \int_{\partial N_{\theta}} \left( \frac{1}{(\psi_{\theta}^{FN})^{2}} \left( \int_{\partial N_{\theta}} \left( |\nabla^{+}\psi_{\theta}^{FN}|^{2} r_{\theta}^{+} + |\nabla^{-}\psi_{\theta}^{FN}|^{2} r_{\theta}^{-} \right) \, ds \right) \right).
\]

Moreover, \( \nabla_{\theta}E_{\theta}^{FN} = 0 \) for any parameterization \( \theta \mapsto \partial N_{\theta} \) of the nodes if and only if the fixed node groundstate \( \psi_{\theta}^{FN} \) is an eigenfunction of \( H \) in \( \mathbb{R}^{d} \).

Proof. The shape derivative formula (1.4) can be decomposed as the sum of the part due to \( N_{\theta}^{+} \) and to \( N_{\theta}^{-} \):

\[
\nabla_{\theta}E_{\theta}^{FN} = -\frac{1}{2} \int_{\partial N_{\theta}} \left( \frac{1}{(\psi_{\theta}^{FN})^{2}} \left( \int_{\partial N_{\theta}} \left( |\nabla^{+}\psi_{\theta}^{FN}|^{2} r_{\theta}^{+} + |\nabla^{-}\psi_{\theta}^{FN}|^{2} r_{\theta}^{-} \right) \, ds \right) \right).
\]

Since \( \psi_{\theta} = 0 \) on \( \partial N_{\theta} \), then \( |\nabla^{+}\psi_{\theta}^{FN}| = |\nabla^{+}\psi_{\theta}^{FN} \cdot n_{+}| \) and symmetrically \( |\nabla^{-}\psi_{\theta}^{FN}| = |\nabla^{-}\psi_{\theta}^{FN} \cdot n_{-}| \). We get:

\[
4r_{\theta}^{+} (\nabla^{sk}\psi_{\theta}^{FN} \cdot n_{+}) (\nabla^{sy}\psi_{\theta}^{FN} \cdot n_{+}) = r_{\theta}^{+} \left( |\nabla^{+}\psi_{\theta}^{FN}|^{2} - |\nabla^{-}\psi_{\theta}^{FN}|^{2} \right) = r_{\theta}^{+} |\nabla^{+}\psi_{\theta}^{FN}|^{2} + r_{\theta}^{-} |\nabla^{-}\psi_{\theta}^{FN}|^{2},
\]

where in the last line we have used \( r_{\theta}^{+} = -r_{\theta}^{-} \). Then (5.6) follows. As a consequence, \( \nabla_{\theta}E_{\theta}^{FN} = 0 \) if and only if \( \nabla^{+}\psi_{\theta}^{FN} \cdot n_{+} = \nabla^{-}\psi_{\theta}^{FN} \cdot n_{-} \), which is equivalent by Lemma 5.4 to the fact that \( \psi_{\theta}^{FN} \) is an eigenfunction.

Theorem 5.6. Consider a Brownian motion \( t \mapsto W_{t}^{+} \) in \( N_{\theta}^{+} \), and \( \tau^{+} \) the hitting time of \( \partial N_{\theta} \). Consider the measure \( \mu_{\theta,\lambda}^{FN} \) on \( \partial N_{\theta} \) defined for any \( \lambda < E_{\theta}^{FN} \) by (1.20). Then the fixed node groundstate \( \psi_{\theta}^{FN} \) is an eigenfunction of \( H \) in \( \mathbb{R}^{d} \), if and only if \( \mu_{\theta,\lambda}^{FN} \) is invariant under the symmetry group \( S \). Moreover, the shape derivative of the fixed node energy \( E_{\theta}^{FN} \) is given by (1.21).
Proof. The proof consists in a direct application of Lemma 3.1 and Proposition 3.2. Indeed, (1.11) yields the identity of measures on \( \partial N_\theta \):
\[
d\mu^{\text{FN}}_{\theta, \lambda} = -\frac{\nabla^+ \psi_{\theta}^{\text{FN}} \cdot n_+ d\sigma}{2(E_{\theta}^{\text{FN}} - \lambda) \int_{N_\theta^+} (\psi_{\theta}^{\text{FN}})^2};
\]
and Lemma 5.4 enables to characterize eigenfunctions of \( H \) from the symmetry of \( \eta_\theta^{\text{FN}} \). Next, since \( r_\theta^+ \) on \( \partial N_\theta \) is skew-symmetric, and the integral on \( \partial N_\theta \) of integrable skew-symmetric functions vanishes, it yields:
\[
\int_{\partial N_\theta} r_\theta^+ \nabla^\text{sy} \psi_{\theta}^{\text{FN}} \cdot n_+ \nabla^+ \psi_{\theta}^{\text{FN}} \cdot n_+ d\sigma = \int_{\partial N_\theta} r_\theta^+ \nabla^\text{sy} \psi_{\theta}^{\text{FN}} \cdot n_+ \nabla^\text{sk} \psi_{\theta}^{\text{FN}} \cdot n d\sigma.
\]
Then (5.6) with (3.4) yield the result (1.21).

6. COMMENTS ON MONTE-CARLO METHODS

Theorem 5.6 suggests a Monte-Carlo general strategy to approximate the fixed node energy variation (1.21) by using the probabilistic interpretation (1.20)–(1.19). As already commented in the introduction, computing (1.20) with Monte-Carlo methods is a well known topic, and computing (1.19) can be carried out using independent stopped processes. In fact, the most straightforward limitation of the method consists in approximating on the nodes \( \partial N_\theta \) the symmetrized gradient
\[
\nabla^\text{sy} \psi_{\theta}^{\text{FN}} \sim \nabla^{\text{sk}} \psi_{\theta}^{\text{FN}}
\]
by a sequence \( \psi^n_{\theta} \) of smooth approximating functions in \( \mathbb{R}^d \), usually converging in the sense of the energy norm. Indeed, there is no reason a priori that the gradient \( \nabla^\text{sk} \psi_{\theta}^{\text{FN}} \) converges in a pointwise sense towards \( \nabla^\text{sy} \psi_{\theta}^{\text{FN}} \), which is necessary when approximating the measure \( \eta_\theta^{\text{FN}} \) using a Monte-Carlo sample. However, in the context of optimization, it is fundamental to remark that the stationary nodal surfaces solution of \( \nabla E_{\theta}^{\text{FN}} = 0 \) for any shape derivative are not modified when approximating \( \nabla^\text{sy} \psi_{\theta}^{\text{FN}} \cdot n_+ \) by any symmetric field \( \nabla^{\text{sk}} \psi_{\theta}^{\text{FN}} \cdot n_+ \) in \( \mathbb{R}^d \).

This is the meaning of the following proposition.

Proposition 6.1. Consider any smooth skew-symmetric function \( \psi_{\theta}^{\text{FN}} \) in \( \mathbb{R}^d \) with zeros given by \( \partial N_\theta \). Then the approximation \( \nabla_{\theta} E_{\theta}^{\text{FN}} \) given by (1.23) yields (1.24)–(1.25). Moreover the fixed point equation
\[
\nabla_{\theta} E_{\theta}^{\text{FN}} = 0
\]
holds for any shape variation \( r_\theta^+ \) of the nodes \( \partial N_\theta \) if and only if \( \psi_{\theta}^{\text{FN}} \) is an eigenfunction of the Hamiltonian \( H \) in \( \mathbb{R}^d \).

Proof. Since \( \partial N_\theta \) is defined implicitly by \( \psi_{\theta}^{\text{FN}} = 0 \), deriving \( \psi_{\theta}^{\text{FN} + h}(x + r_\theta^+ h n_+) = 0 \) with respect to \( h \) yields on \( \partial N_\theta \):
\[
\nabla_{\theta} \psi_{\theta}^{\text{FN} + h} + r_\theta^+ \nabla \psi_{\theta}^{\text{FN}} \cdot n_+ = \nabla_{\theta} \psi_{\theta}^{\text{FN}} - r_\theta^+ |\nabla \psi_{\theta}^{\text{FN}}| = 0,
\]
where in the last line we have used the fact that \( \nabla \psi_{\theta}^{\text{FN}} \cdot n_+ = -|\nabla \psi_{\theta}^{\text{FN}}| \) on \( \partial N_\theta \) since \( N_\theta^+ = \{ x \in \mathbb{R}^d | \psi_{\theta}^{\text{FN}}(x) > 0 \} \). This gives (1.24). Next (1.11) yields:
\[
\int_{\partial N_{\theta}^+} \nabla_{\theta} \psi_{\theta}^{\text{FN}} d\mu^{\text{FN}}_{\theta, \lambda} = -\frac{1}{2(E_{\theta}^{\text{FN}} - \lambda) \int_{N_\theta^+} \psi_{\theta}^{\text{FN}} d\mu_{\theta}} \int_{\partial N_\theta} \nabla_{\theta} \psi_{\theta}^{\text{FN}} \nabla \psi_{\theta}^{\text{FN}} \cdot n_+ d\sigma,
\]
so that integration by parts in \( N_{\theta}^+ \) gives (1.25).

Finally remark that \( \nabla \psi_{\theta}^{\text{FN}} \cdot n_+ \) is symmetric and that the shape derivative \( r_\theta^+ \) is spanning all skew-symmetric field, so that the fixed point equation is equivalent to the fact that \( \mu_{\theta}^{\text{FN}} \) is symmetric. Then Lemma 5.4 enables to conclude.
Eventually, the following algorithm (without details) is suggested to estimate $\nabla_\theta E^{\text{FN}}_\theta$. It might be referred to as Nodal Monte-Carlo (NMC).

**Algorithm 6.2 (NMC).** Consider a parameterization of wave functions of the form (4.1). Then the following steps are suggested:

1. Optimize the symmetric “Jastrow” factor $J_\alpha$ according to the variational problem with fixed node (1.15), and denote $\psi^1_\alpha$ the obtained trial function.
2. Generate a sample according to the probability distribution $\eta_\theta^{\text{FN}}$ in $N^+_\theta$ defined in (1.20). Use for instance a long time trajectory of a drifted stochastic process of the type (3.3) with Feynman-Kac weights as in formula (3.4).
3. Use the latter sample as initial conditions; and sample independent Brownian motions. Stop them when they hit the nodes $\partial N_\theta$. Then compute the measure $\mu_\theta^{\text{FN}}$ according to (1.20).
4. Estimate the variations of the fixed node energy $\nabla_\theta E^{\text{FN}}_\theta$ with (1.24).

As a conclusion, we mention two possible strategies of variance reduction that will be necessary for efficient computations. Their development is left for future work.

- First, a drifted diffusion instead of plain Brownian motion may be used in (1.20) to reduce the variance caused by the exponential weights. However, the classical stochastic differential equation (3.3) where $\psi^1$ vanishes at the nodes cannot be used. Indeed, the repulsive drift at the nodes prevent the process from hitting the latter, and (1.20) no longer holds. Instead of a skew-symmetric guiding function $\psi^1$, we propose to use as a drift a symmetric function $\psi_B$, for instance an approximation of the Bosonic groundstate $\psi_B^*$ of $H$ in $\mathbb{R}^d$ solution of the eigenvalue problem

$$H(\psi_B^*) = E_B^* \psi_B^*.$$  

This may remove the variance of the exponential Feynman-Kac weights, while letting the walkers hit the nodal surface. Such variance reductions are referred to as importance sampling methods. Note that in known algorithms (e.g. DMC), such guided walkers will very quickly hit the nodal surface, causing high variance branching. However, in the proposed method, the algorithm is stopped when walkers have hit the nodal surface once, and no branching is performed, avoiding such kind of variance instability.

- Second, a key point would be to develop a coupling (in the probabilistic sense) between the random process used to compute $\mu_\theta^{\text{FN}}$ in (1.20) and another random process stopped on the nodes $\partial N_\theta$, with a distribution denoted $\tilde{\mu}_\theta^{\text{FN}}$. The goal is to design $\tilde{\mu}_\theta^{\text{FN}}$ so that the following two features hold: first computing the average (1.24) with $\tilde{\mu}_\theta^{\text{FN}}$ always yields 0; second the coupling is perfect in the special case when $\mu_\theta^{\text{FN}}$ is symmetric (that is to say the Monte-Carlo method computing $\mu_\theta^{\text{FN}}$ and $\tilde{\mu}_\theta^{\text{FN}}$ has to be the same with the same random numbers). In the physics terminology, this would yield a zero variance estimator, which we prefer to call asymptotic variance reduction, in the sense that thanks to the variance reduction, the variance of the estimator scales with the quantity to be computed when the latter becomes small.

**Acknowledgements.** The author would like to thank E. Cancès, B. Jourdain and T. Lelièvre for pointing out the problem and fruitful discussions, as well as the referees for very careful reading and useful suggestions.

**References**


