On a probabilistic interpretation of shape derivatives of Dirichlet groundstates with application to Fermion nodes.

Mathias Rousset

On a probabilistic interpretation of shape derivatives of Dirichlet groundstates with application to Fermion nodes.

Mathias Rousset[*]

January 27, 2013

Contents

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

[*]INRIA Lille - Nord Europe & Université Lille 1, Villeneuve d'Ascq, France. E-mail: `mathias.rousset@inria.fr`

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.

1 Introduction and results

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

$$H = -\frac{\Delta}{2} + V, \tag{1}$$

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

$$\theta \mapsto \Omega_\theta \tag{2}$$

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 ∇, and gradients with respect to θ, ∇_θ. 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:

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

$$= \frac{\displaystyle\int_{\Omega_\theta} \psi_\theta^* H(\psi_\theta^*)}{\displaystyle\int_{\Omega_\theta} (\psi_\theta^*)^2}. \tag{3}$$

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

$$\nabla_\theta E_\theta^* = -\frac{1}{2} \int_{\partial\Omega_\theta} |\nabla\psi_\theta^*|^2 \, r_\theta \, d\sigma, \tag{4}$$

where in the above σ is the usual surface measure induced by the canonical Euclidean structure $\mathbb{R}^d$, and r_θ 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:

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

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_θ can also be defined through

$$r_\theta(x)\nabla\psi_\theta(x) \cdot n(x) = -\nabla_\theta\psi_\theta(x) \qquad \forall x \in \partial\Omega_\theta. \tag{5}$$

Formula (5) can be proved as follows: from the Dirichlet conditions, one has for any small h:

$$\begin{aligned}
0 &= \psi^*_{\theta+h}\left(x + h \cdot r_\theta(x)n(x) + \mathrm{O}(|h|^2)\right) - \psi^*_\theta(x) \\
&= h \cdot r_\theta(x)\nabla\psi^*_\theta(x) \cdot n(x) + h \cdot \nabla_\theta\psi^*_\theta(x) + \mathrm{O}(h).
\end{aligned}$$

Differentiability in formula (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 Ω_θ. The first exit time of the domain Ω_θ is denoted by

$$\tau := \inf\left(t \geq 0 | W_t \in \partial\Omega_\theta\right). \tag{6}$$

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

$$\int_{\Omega_\theta} \varphi \, d\eta^*_\theta \ := \lim_{T\to+\infty} \frac{\mathbb{E}\left(\varphi(W_T)\mathbf{1}_{T\leq\tau}\mathrm{e}^{-\int_0^T V(W_s)ds}\right)}{\mathbb{E}\left(\mathbf{1}_{T\leq\tau}\mathrm{e}^{-\int_0^T V(W_s)ds}\right)}. \tag{7}$$

It has a probability density function given by the signed groundstate ψ^*_θ:

$$\int_{\Omega_\theta} \varphi \, d\eta^*_\theta \ = \frac{\displaystyle\int_{\Omega_\theta} \varphi \, \psi^*_\theta \, dx}{\displaystyle\int_{\Omega_\theta} \psi^*_\theta \, dx}, \tag{8}$$

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

$$\lim_{T \to +\infty} -\frac{1}{T} \ln \mathbb{E}\left(\mathbb{1}_{T \leq \tau} \mathrm{e}^{-\int_0^T V(W_s)ds}\right) = E_\theta^*. \tag{9}$$

The probabilistic interpretations (7)-(8)-(9) have some variants (see (34)-(35)) where a drift is added to the Wiener process and the range of the potential V in the Feynman-Kac weight $\mathrm{e}^{-\int_0^T V(W_s)ds}$ 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 $\mathrm{e}^{-\int_0^T V(W_s)ds}$ when averages are computed. For instance when several random processes are simulated in parallel, some *re-sampling* of the set of processes has to be carried out at regular time intervals, according to the weights associated with each process. We refer the reader to [21, 1, 2] for applications in Quantum Chemistry (Diffusion Monte-Carlo or Pure Diffusion Monte-Carlo methods), to [16, 17] for applications in Bayesian statistics (Sequential Monte-Carlo methods), and to [14, 13, 15, 26] for the associated mathematical analysis.

Then, we consider the weighted distribution of the Wiener process at the hitting time τ, when the process is initially distributed according to η_θ^* defined in (7). The latter distribution is denoted $d\mu_{\theta,\lambda}^*$ and reads

$$\begin{aligned}
\int_{\partial\Omega_\theta} \varphi \, d\mu_{\theta,\lambda}^* := \ & \mathbb{E}\left(\varphi(W_\tau)\mathbb{1}_{\tau < +\infty} \mathrm{e}^{-\int_0^\tau (V(W_s)-\lambda)\,ds} \,|\, \mathrm{Law}(W_0) = \eta_\theta^*\right) \\
= \ & \lim_{T \to +\infty} \frac{\mathbb{E}\left(\varphi(W_\tau)\mathbb{1}_{T \leq \tau < +\infty} \mathrm{e}^{-\int_0^\tau (V(W_s)-\lambda)\,ds}\right)}{\mathbb{E}\left(\mathbb{1}_{T \leq \tau} \mathrm{e}^{-\int_0^T (V(W_s)-\lambda)\,ds}\right)}
\end{aligned} \tag{10}$$

which verifies

$$\int_{\partial\Omega_\theta} \varphi \, d\mu_{\theta,\lambda}^* = -\frac{1}{2(E_\theta^* - \lambda)\int_{\Omega_\theta} \psi_\theta^*} \int_{\partial\Omega_\theta} \varphi \nabla\psi_\theta^* \cdot n \, d\sigma. \tag{11}$$

Formula (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 (4) by remarking that $|\nabla\psi_\theta^* \cdot n| = |\nabla\psi_\theta^*|$. Up to our knowledge, the formula (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).

4

Approximations of $\nabla_\theta E_\theta^*$ with Monte-Carlo methods can then be carried out using an approximating sequence of the Dirichlet groundstate (3) $\psi_\theta^n \xrightarrow{n\to\infty} \psi_\theta^*$, and random samples of size N approximating the probabilistic formulations (7)-(10), denoted $\eta_\theta^N \xrightarrow{N\to\infty} \eta_\theta^*$ and $\mu_{\theta,\lambda}^N \xrightarrow{N\to\infty} \mu_{\theta,\lambda}^*$. The following identity can then be used:

$$
\begin{aligned}
\nabla_\theta E_\theta^* &= \frac{(E_\theta^* - \lambda)}{\int_{\Omega_\theta} \psi_\theta^* \, d\eta_\theta^*} \int_{\partial\Omega_\theta} r_\theta \nabla \psi_\theta^* . n \, d\mu_{\theta,\lambda}^* \\
&\sim \frac{(E_\theta^* - \lambda)}{\int_{\Omega_\theta} \psi_\theta^n \, d\eta_\theta^N} \int_{\partial\Omega_\theta} r_\theta \nabla \psi_\theta^n . n \, d\mu_{\theta,\lambda}^N
\end{aligned}
\tag{12}
$$

The main limitation of computing $\nabla_\theta E_\theta^*$ with the Monte-Carlo technique suggested above is the necessity of an analytical approximation ψ_θ^n of the groundstate ψ_θ^*. Especially, the pointwise convergence of the normal derivative $\nabla \psi_\theta^n . n$ 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 (ψ_F^*, E_F^*) associated to a finite symmetry group $\mathcal{S} \subset O(\mathbb{R}^d)$ of the Hamiltonian H in (1); where $O(\mathbb{R}^d)$ denotes the group of isometries. $\mathcal{S}$ is simply the permutation goup of identical particles for physical systems. Fermionic groundstates are the solutions to the variational problem:

$$
\begin{aligned}
E_F^* &:= \inf\left(\frac{\displaystyle\int_{\mathbb{R}^d} \psi H(\psi)}{\displaystyle\int_{\mathbb{R}^d} \psi^2}, \quad \forall S \in \mathcal{S}, \ \psi \circ S = \det(S)\,\psi \right) \\
&= \frac{\displaystyle\int_{\Omega_\theta} \psi_F^* H(\psi_F^*)}{\displaystyle\int_{\Omega_\theta} (\psi_F^*)^2}.
\end{aligned}
\tag{13}
$$

Any function ψ verifying the symmetry property

$$
\forall S \in \mathcal{S}, \ \psi \circ S = \det(S)\psi
$$

will be called *skew-symmetric*, whereas any function ψ verifying

$$
\forall S \in \mathcal{S}, \ \psi \circ S = \psi
$$

will be called *symmetric*. Note that existence of $(\psi_{\mathrm{F}}^*, E_{\mathrm{F}}^*)$ 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, $(\psi_{\mathrm{F}}^*, E_{\mathrm{F}}^*)$ is computed using a parametrization 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 [29, 28] 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 parametrization 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 parametrization, and will be called the *trial wave function*, which is a skew-symmetric function denoted $\psi_{\theta_0}^{\mathrm{I}}$. We will then consider a family of skew-symmetric functions $\left\{\psi_\theta^{\mathrm{I}}\right\}_{\theta \in \mathbb{R}^p}$, with an explicit analytical expression, which includes $\psi_{\theta_0}^{\mathrm{I}}$. The latter is used to define the *nodal domains*

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

where:

$$\begin{cases} \mathcal{N}_\theta^+ = \left\{x \in \mathbb{R}^d \,|\, \psi_\theta^{\mathrm{I}}(x) > 0\right\} \\ \mathcal{N}_\theta^- = \left\{x \in \mathbb{R}^d \,|\, \psi_\theta^{\mathrm{I}}(x) < 0\right\} \\ \partial\mathcal{N}_\theta = \left\{x \in \mathbb{R}^d \,|\, \psi_\theta^{\mathrm{I}}(x) = 0\right\}. \end{cases} \tag{14}$$

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

$$(\psi_\theta^{FN}, E_\theta^{FN}) = (\psi_\theta^*, E_\theta^*)$$

of the variational problem with Dirichlet conditions in $\mathcal{N}_\theta^+ \cup \mathcal{N}_\theta^-$. We refer the reader to [9, 11] for historical papers on the Monte-Carlo computation

of fixed node groundstates. The latter variational problem reads explicitely:

$$
\begin{aligned}
E_\theta^{\mathrm{FN}} \quad &:= \inf \left(\frac{\displaystyle\int_{\mathbb{R}^d} \psi H\left(\psi\right)}{\displaystyle\int_{\mathbb{R}^d} \psi^2}, \quad \psi|_{\partial\mathcal{N}_\theta} = 0 \right) \\
&= \frac{\displaystyle\int_{\mathbb{R}^d} \psi_\theta^{\mathrm{FN}} H\left(\psi_\theta^{\mathrm{FN}}\right)}{\displaystyle\int_{\mathbb{R}^d} (\psi_\theta^{\mathrm{FN}})^2} \\
&= \inf \left(\frac{\displaystyle\int_{\mathcal{N}_\theta^+} \psi H\left(\psi\right)}{\displaystyle\int_{\mathcal{N}_\theta^+} \psi^2}, \quad \psi|_{\partial\mathcal{N}_\theta} = 0 \right) \\
&= \frac{\displaystyle\int_{\mathcal{N}_\theta^+} \psi_\theta^{\mathrm{FN}} H\left(\psi_\theta^{\mathrm{FN}}\right)}{\displaystyle\int_{\mathcal{N}_\theta^+} (\psi_\theta^{\mathrm{FN}})^2},
\end{aligned}
\tag{15}
$$

the same definition holding in $\mathcal{N}_\theta^-$ by symmetry. From now on, $\mathcal{N}_\theta^+$ and $\mathcal{N}_\theta^-$ will be called respectively the positive and negative *nodal domains*, $\partial\mathcal{N}_\theta$ the *nodal surface* or *nodes*, and $(E_\theta^{FN}, \psi_\theta^{FN})$ the *fixed node groundstate* elements. Now the key problem is the following: the energy (called Variational Monte Carlo, or in short VMC energy in the QMC literature)

$$
E_\theta^{\mathrm{I}} = \frac{\displaystyle\int_{\mathcal{N}_\theta^+} \psi_\theta^{\mathrm{I}} H\left(\psi_\theta^{\mathrm{I}}\right)}{\displaystyle\int_{\mathcal{N}_\theta^+} \left(\psi_\theta^{\mathrm{I}}\right)^2}
\tag{16}
$$

of the trial wave function ψ_θ^{I} can be minimized towards the Fermionic groundstate energy E_F^* using some optimization procedure, and the fixed node energy $E_\theta^{\mathrm{FN}} \leq E_\theta^{\mathrm{I}}$ associated with ψ_θ^{I} can be computed using a Monte-Carlo method associated with (9). However, a parameter θ optimizing the energy E_θ^{I} of the trial wave function does not in general, for a given parametrization, optimize the fixed node groundstate energy E_θ^{FN}. An open problem is now to develop an algorithm that can directly minimize E_θ^{FN}. To precise this idea, let us consider the formulation of the exact Fermionic groundstate (ψ_F^*, E_F^*) as a variational problem involving the fixed node groundstate $(\psi_\theta^{\mathrm{FN}}, E_\theta^{\mathrm{FN}})$

7

and the nodal surface $\partial \mathcal{N}_\theta$, that is to say:

$$
\begin{aligned}
E_{\mathrm{F}}^* \quad &:= \inf \left(E_\theta^{\mathrm{FN}}, \quad \psi_\theta^{\mathrm{FN}} \text{ solution of (15)}, \quad \psi_{\mathrm{I}}^\theta \text{ skew-symmetric} \right) \\
&= \frac{\displaystyle\int_{\mathbb{R}^d} \psi_{\mathrm{F}}^* H\left(\psi_{\mathrm{F}}^*\right)}{\displaystyle\int_{\mathbb{R}^d} \left(\psi_{\mathrm{F}}^*\right)^2}.
\end{aligned} \tag{17}
$$

An approach to solve the variational problem (17) consists in the computation of the shape derivative of the fixed node groundstate

$$
\nabla_\theta E_\theta^{\mathrm{FN}} \tag{18}
$$

using a Monte-Carlo estimation based on (12). In this context, the key formulas (11)-(12) can be rewritten as follows. $t \mapsto W_t^+$ denotes a Brownian motion in $\mathcal{N}_\theta^+$, and τ^+ the hitting time of $\partial \mathcal{N}_\theta$. Then the probability measure $d\eta_\theta^{\mathrm{FN}}$ on $\mathcal{N}_\theta^+$ is defined by:

$$
\int_{\mathcal{N}_\theta^+} \varphi \, d\eta_\theta^{\mathrm{FN}} := \lim_{T \to +\infty} \frac{\mathbb{E}\left(\varphi(W_T^+) \mathbf{1}_{T \leq \tau^+} \mathrm{e}^{-\int_0^T V(W_s^+)ds} \right)}{\mathbb{E}\left(\mathbf{1}_{T \leq \tau^+} \mathrm{e}^{-\int_0^T V(W_s^+)ds} \right)}, \tag{19}
$$

and the measure $d\mu_{\theta,\lambda}^{\mathrm{FN}}$ on $\partial \mathcal{N}_\theta$ for $\lambda < E_\theta^{\mathrm{FN}}$ is defined by:

$$
\int_{\partial \mathcal{N}_\theta} \varphi \, d\mu_{\theta,\lambda}^{\mathrm{FN}} := \lim_{T \to +\infty} \frac{\mathbb{E}\left(\varphi(W_{\tau^+}^+) \mathbf{1}_{T \leq \tau^+ < +\infty} \mathrm{e}^{-\int_0^{\tau^+} \left(V(W_s^+) - \lambda \right) ds} \right)}{\mathbb{E}\left(\mathbf{1}_{T \leq \tau^+} \mathrm{e}^{-\int_0^T \left(V(W_s^+) - \lambda \right) ds} \right)}. \tag{20}
$$

We will show that (12) becomes:

$$
\begin{aligned}
\nabla_\theta E_\theta^{\mathrm{FN}} &= \frac{(E_\theta^{\mathrm{FN}} - \lambda)}{\int_{\mathcal{N}_\theta^+} \psi_\theta^{\mathrm{FN}} \, d\eta_\theta^{\mathrm{FN}}} \int_{\partial \mathcal{N}_\theta} r_\theta^+ \nabla^+ \psi_\theta^{\mathrm{FN}} \cdot n_+ \, d\mu_{\theta,\lambda}^{\mathrm{FN}} \\
&= \frac{2(E_\theta^{\mathrm{FN}} - \lambda)}{\int_{\mathcal{N}_\theta^+} \psi_\theta^{\mathrm{FN}} \, d\eta_\theta^{\mathrm{FN}}} \int_{\partial \mathcal{N}_\theta} r_\theta^+ \nabla^{\mathrm{sy}} \psi_\theta^{\mathrm{FN}} \cdot n_+ \, d\mu_{\theta,\lambda}^{\mathrm{FN}}.
\end{aligned} \tag{21}
$$

In the above, r_θ^+ is the shape derivative of $\mathcal{N}_\theta^+$, n_+ is the associated exterior normal vector, and $\nabla^{\mathrm{sy}} \psi_\theta^{\mathrm{FN}} \cdot n_+$ is the symmetrization of the normal groundstate gradient, that is to say:

$$
\nabla^{\mathrm{sy}} \psi \cdot n_+ = \frac{1}{2} \left(\nabla^+ \psi \cdot n_+ - \nabla^- \psi \cdot n_- \right), \tag{22}
$$

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

1. The measure $\mu_{\theta,\lambda}^{\mathrm{FN}}$ on $\partial\mathcal{N}_\theta$ is symmetric (*i.e.* invariant by the action of $\mathcal{S}$).

2. When defined on the whole space $\mathbb{R}^d$, the gradient of the fixed node goundstate $\nabla\psi_\theta^{\mathrm{FN}}$ is continuous on $\partial\mathcal{N}_\theta$.

3. The fixed node goundstate $\psi_\theta^{\mathrm{FN}}$ is a skew-symmetric eigenfunction of H on $\mathbb{R}^d$.

4. The fixed node energy variation vanishes $\nabla_\theta E_\theta^{\mathrm{FN}} = 0$ for any parametrization $\theta \mapsto \partial\mathcal{N}_\theta$ 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 $\nabla_\theta E_\theta^{\mathrm{FN}}$ using (21) requires a Monte-Carlo estimator of the measures $(\mu_{\theta,\lambda}^{\mathrm{FN}}, \eta_\theta^{\mathrm{FN}})$ on the one hand, and an analytical approximation of the fixed node groundstate elements $(\psi_\theta^{\mathrm{FN}}, \nabla^{\mathrm{sy}}\psi_\theta^{\mathrm{FN}} \cdot n_+)$ on the other hand (see also (12)). A key remark is that the elements $(\psi_\theta^{\mathrm{FN}}, \nabla^{\mathrm{sy}}\psi_\theta^{\mathrm{FN}} \cdot n_+)$ and any approximation with a skew-symmetric function smooth on $\mathbb{R}^d$, for instance with $(\psi_\theta^{\mathrm{I}}, \nabla\psi_\theta^{\mathrm{I}} \cdot n_+)$, share the same symmetry properties. This suggests the following approximation of the energy variation $\nabla_\theta E_\theta^{\mathrm{FN}}$ in (21):

$$\nabla_\theta E_\theta^{\mathrm{FN}} \sim \widehat{\nabla_\theta E_\theta^{\mathrm{FN}}} = \frac{2(E_\theta^{\mathrm{FN}} - \lambda)}{\int_{\mathcal{N}_\theta^+} \psi_\theta^{\mathrm{I}} d\eta_\theta^{\mathrm{FN}}} \int_{\partial\mathcal{N}_\theta} r_\theta^+ \nabla\psi_\theta^{\mathrm{I}} \cdot n_+ d\mu_{\theta,\lambda}^{\mathrm{FN}} \qquad (23)$$

$$= -\frac{2(E_\theta^{\mathrm{FN}} - \lambda)}{\int_{\mathcal{N}_\theta^+} \psi_\theta^{\mathrm{I}} d\eta_\theta^{\mathrm{FN}}} \int_{\partial\mathcal{N}_\theta} \nabla_\theta\psi_\theta^{\mathrm{I}} d\mu_{\theta,\lambda}^{\mathrm{FN}} \qquad (24)$$

$$= \frac{\int_{\mathcal{N}_\theta^+} \left(H - E_\theta^{\mathrm{FN}}\right) (\nabla_\theta\psi_\theta^{\mathrm{I}}) d\eta_\theta^{\mathrm{FN}}}{\int_{\mathcal{N}_\theta^+} \psi_\theta^{\mathrm{I}} d\eta_\theta^{\mathrm{FN}}}. \qquad (25)$$

In the same way as for the exact expression $\nabla_\theta E_\theta^{\mathrm{FN}}$, the latter vanishes $(\widehat{\nabla_\theta E_\theta^{\mathrm{FN}}} = 0)$ if 1, 2, or 3 holds. In return, if $\widehat{\nabla_\theta E_\theta^{\mathrm{FN}}} = 0$ for any parametrization of the nodes then 1, 2 or 3 holds. Such algorithms may be referred to as

9

Nodal Monte-Carlo. They will require variance reduction techniques exploiting the symmetry structure, in principle such that the variance of $\widehat{\nabla_\theta E_\theta^{\mathrm{FN}}}$ scales appropriately to 0 in the limit where $\mu_{\theta,\lambda}^{\mathrm{FN}}$ 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 (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 Remark 4.2). This explains the necessity of the fixed node approximation. The issue of optimizing the nodes of the trial wave function ψ_θ^{I} in the fixed node approximation was pointed out in [10], 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_θ^{I} in (16), as in [29, 28]; or (ii) the minimization of the variance of the local energy

$$E_L := V - (\psi_\theta^{\mathrm{I}})^{-1} \frac{\Delta}{2}(\psi_\theta^{\mathrm{I}}) \tag{26}$$

as in [29, 22]. As a consequence, an efficient method optimizing directly the fixed node energy E_θ^{FN} in (15), that is to say the nodes of the trial wave function ψ_θ^{I}, remains an unsolved problem and motivates the content of this paper. However, methods to approximately compute the gradient $\nabla_\theta E_\theta^{\mathrm{FN}}$ 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 $\nabla_R E_{\mathrm{F},R}^*$ of the groundstate energy with respect to some parameter R parametrizing 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: $\psi_\theta^{\mathrm{I}} \equiv \psi_{R,\theta}^{\mathrm{I}}$. Well established methods are available for the exact and variance reduced computation of the variational energy gradient:

$$\nabla_R E_{\theta,R}^{\mathrm{I}} \tag{27}$$

where $E_{\theta,R}^{\mathrm{I}}$ is defined by (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 [3]; see also some applications in [27]. The case of the fixed node energy gradient:

$$\nabla_R E_{\theta,R}^{\mathrm{FN}} \tag{28}$$

where $E_{\theta,R}^{\mathrm{FN}}$ is defined by (15) is more intricate, and requires some approximation since the derivatives $\nabla_R \psi_{\theta,R}^{\mathrm{FN}}$ of the fixed node groundstate remain unknown. Note that computing the gradients $\nabla_\theta E_\theta^{\mathrm{I}}$ or $\nabla_\theta E_{\theta,R}^{\mathrm{FN}}$ can be seen as a particular case of computing respectively (27) or (28), since the nodes of the trial wave function $\psi_{R,\theta}^{\mathrm{I}}$ depend on R *a priori* . An approximate formula has already been proposed to compute (28), see for instance equation (54) in [3], (10) in [8], (10) in [5], and (14) in [4]. The terms due to the variation of the nodes (which amounts to evaluate $\nabla_\theta E_{\theta,R}^{\mathrm{FN}}$) are called 'nodal Pulay terms', and were particularly pointed out in [5, 4]. In the above references, the formula used for the calculation is exactly the formula (25). However, the interpretation in terms of a stochastic process stopped on the nodes in (24)-(23), 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): [21, 1].

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 (4) is proven formally, and references are given for rigorous proofs.

Consider the Schrödinger operator (1) defined on $\mathbb{R}^d$ with a smooth potential V bounded from below. H defines a self-adjoint operator on the Hilbert space $\mathbb{L}^2(\mathbb{R}^d)$. 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 parametrization of smooth domains (2) is considered for $\theta \in \mathbb{R}^p$ such that $\partial\Omega_\theta$ has a smooth boundary for any θ. One assumes then that there exists a set of diffeomorphisms smoothly indexed by θ and such that:

$$\begin{cases} (\theta, x) \mapsto R_\theta(x) \text{ is smooth in } \mathbb{R}^p \times \mathbb{R}^d \\[2mm] R_0 = \text{Id} \\[2mm] R_\theta = \text{Id outside some compact set} \\[2mm] \forall \theta \in \mathbb{R}^p, \quad \Omega_\theta = R_\theta(\Omega_0). \end{cases} \tag{29}$$

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:

$$\forall x \in \partial\Omega_\theta, \quad r_\theta(x) = \nabla_\theta R_\theta(x) \cdot n(x), \tag{30}$$

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

$$\partial\Omega_{\theta+h} \sim \{x + h \cdot r_\theta(x) n(x) \,|\, x \in \partial\Omega_\theta\}.$$

Now classical results of spectral theory ensures that the Hamiltonian (1) considered in $\mathbb{L}^2(\Omega_\theta)$ with Dirichlet boundary condition, is self-adjoint with domain $\mathcal{D}(H) \subset H_0^1(\Omega_\theta)$, where $H_0^1(\Omega_\theta)$ 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 ground-state ψ_θ^*, solution of the variational problem (3), or equivalently solution of the eigenvalue problem:

$$\begin{cases} H(\psi_\theta^*) = E_\theta^* \psi_\theta^* \\[2mm] \psi_\theta^*|_{\partial\Omega_\theta} = 0 \\[2mm] \psi_\theta^* > 0, \end{cases}$$

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

Lemma 2.1. *Consider domains* $\theta \mapsto \Omega_\theta$ *verifying* (29). *Then the Dirichlet energy* E_θ^* *solution of the variational problem* (3) *is differentiable with respect to the parameter* θ, *and the variation formula* (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:

$$(\nabla_\theta E_\theta^*)_{\theta=\theta_0} = \left(\nabla_\theta \frac{\displaystyle\int_{\Omega_\theta} \psi_\theta^* H\left(\psi_\theta^*\right)}{\displaystyle\int_{\Omega_\theta} \left(\psi_\theta^*\right)^2} \right)_{\theta=\theta_0}$$

$$= \left(\nabla_\theta \frac{\displaystyle\int_{\Omega_\theta} \psi_{\theta_0}^* H\left(\psi_{\theta_0}^*\right)}{\displaystyle\int_{\Omega_\theta} \left(\psi_{\theta_0}^*\right)^2} \right)_{\theta=\theta_0} + \left(\nabla_\theta \frac{\displaystyle\int_{\Omega_{\theta_0}} \psi_\theta^* H\left(\psi_\theta^*\right)}{\displaystyle\int_{\Omega_{\theta_0}} \left(\psi_\theta^*\right)^2} \right)_{\theta=\theta_0}.$$

Since $(\psi_\theta^*, E_\theta^*)$ is an eigenelement, it yields:

$$\nabla_\theta \frac{\displaystyle\int_{\Omega_\theta} \psi_{\theta_0}^* H\left(\psi_{\theta_0}^*\right)}{\displaystyle\int_{\Omega_\theta} \left(\psi_{\theta_0}^*\right)^2} = \nabla_\theta E_{\theta_0}^* = 0.$$

Then formal differentiation yields

$$\nabla_\theta \frac{\displaystyle\int_{\Omega_{\theta_0}} \psi_\theta^* H\left(\psi_\theta^*\right)}{\displaystyle\int_{\Omega_{\theta_0}} \left(\psi_\theta^*\right)^2} = \frac{\displaystyle\int_{\Omega_{\theta_0}} \nabla_\theta \psi_\theta^* (H - E_\theta^*) \left(\psi_\theta^*\right)}{\displaystyle\int_{\Omega_{\theta_0}} \left(\psi_\theta^*\right)^2} + \frac{\displaystyle\int_{\Omega_{\theta_0}} \psi_\theta^* (H - E_\theta^*) \left(\nabla_\theta \psi_\theta^*\right)}{\displaystyle\int_{\Omega_{\theta_0}} \left(\psi_\theta^*\right)^2},$$

and since $(\psi_\theta^*, E_\theta^*)$ is an eigenelement, the first term of the right hand side vanishes so that finally:

$$(\nabla_\theta E_\theta^*)_{\theta=\theta_0} = \frac{\displaystyle\int_{\Omega_{\theta_0}} \psi_{\theta_0}^* (-\frac{1}{2}\Delta + V - E_{\theta_0}^*) \left(\nabla_{\theta_0} \psi_{\theta_0}^*\right)}{\displaystyle\int_{\Omega_{\theta_0}} \left(\psi_{\theta_0}^*\right)^2}.$$

Now applying Green's integration by parts two times, and remarking that

13

$(H - E_{\theta_0}^*)(\psi_{\theta_0}^*) = 0$, we get

$$(\nabla_\theta E_\theta^*)_{\theta=\theta_0} = \frac{1}{2} \frac{\displaystyle\int_{\partial\Omega_{\theta_0}} \nabla\psi_{\theta_0}^* \cdot n \, \nabla_{\theta_0}\psi_{\theta_0}^* \, d\sigma}{\displaystyle\int_{\Omega_{\theta_0}} (\psi_{\theta_0}^*)^2}.$$

Then (5) applied to $\psi_{\theta_0}^*$ yields the result. The rigorous proof can be made using a change a variable with the diffeomophism R_θ, and then exploiting the variational formulation (see e.g. Theorem 2 in [12]). $\qquad\square$

3   Probabilistic interpretations

In this Section, the probabilistic formulations (8)-(9)-(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 τ from Ω_θ defined in (6). The classical probabilistic interpretation of the eigenelements $(\psi_\theta^*, E_\theta^*)$ is recalled in the following lemma:

Lemma 3.1. *Assume W_0 is distributed according to $\frac{\psi_{\mathrm{init}}(x)dx}{\int \psi_{\mathrm{init}}(x)dx}$ where $\psi_{\mathrm{init}} \in L^2(\Omega_\theta)$, and ψ_{init} is non vanishing in each connected component of Ω_θ. 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 (7)-(8)-(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 $\varphi \in C_c^\infty(\Omega_\theta)$ a smooth solution $(t,x) \mapsto u_t(x) \in C^\infty(\mathbb{R}^+ \times \Omega_\theta)$ of the parabolic problem:

$$\begin{cases} \partial_t u_t = -H(u_t) \\[2mm] u_t|_{\partial\Omega_\theta} = 0 \\[2mm] u_0 = \varphi \end{cases}$$

Then using Itô calculus, it yields for any $\phi \in C^\infty(\mathbb{R}^+ \times \Omega_\theta)$:

$$d\left(\phi_{T-t}(W_t)\mathrm{e}^{-\int_0^t V(W_s)ds}\right) =$$
$$\left(\frac{\Delta}{2} - V - \partial_t\right)(\phi_{T-t})(W_t)\mathrm{e}^{-\int_0^t V(W_s)\,ds}\,dt + \mathrm{e}^{-\int_0^t V(W_s)\,ds}\nabla\phi_{T-t}(W_t)\cdot dW_t,$$
$$(31)$$

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

$$\frac{\displaystyle\int_{\Omega_\theta} u_T(x)\psi_{\text{init}}(x)dx}{\displaystyle\int_{\Omega_\theta} \psi_{\text{init}}(x)dx} = \mathbb{E}\left(u_{T-T\wedge\tau}(W_{T\wedge\tau})e^{-\int_0^{T\wedge\tau} V(W_s)ds}\right)$$

$$= \mathbb{E}\left(\varphi(W_T)\mathbb{1}_{\tau \geq T}e^{-\int_0^T V(W_s)ds}\right).$$

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

$$\mathbb{E}\left(\varphi(W_T)\mathbb{1}_{\tau \geq T}e^{-\int_0^T V(W_s)ds}\right) = \sum_{n \geq 0} e^{-E_\theta^{*,n}T}\frac{\displaystyle\int_{\Omega_\theta} \psi_\theta^{*,n}\psi_{\text{init}}}{\displaystyle\int_{\Omega_\theta} \psi_{\text{init}}} \int_{\Omega_\theta} \psi_\theta^{*,n}\varphi. \quad (32)$$

Now using the fact that the groundstate has a spectral gap ($E_\theta^{*,1} > E_\theta^{*,0} = E_\theta^*$), the dominant term in (32) when $T \to +\infty$ enables to verify that the groundstate has a sign $\psi_\theta^{*,0} > 0$. Finally, taking the limit $\varphi \to \mathbf{1}$ by dominated convergence in the formula above, and then the leading term in the limit $T \to +\infty$ yields (9). (8) follows. $\qquad\square$

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

$$\begin{cases} dX_t = (\psi^{\mathrm{I}})^{-1}\nabla\psi^{\mathrm{I}}(X_t)dt + dW_t \\ \psi^{\mathrm{I}}|_{\partial\Omega_\theta} = 0, \end{cases} \quad (33)$$

where $\psi^{\mathrm{I}} > 0$ is a smooth function strictly positive in Ω_θ and vanishing on $\partial\Omega_\theta$. In [6], sufficient conditions on the behavior of ψ^{I} near $\partial\Omega_\theta$ and at infinity are given for (33) to be well-posed, and to verify the following variant of (8)-(9):

$$\lim_{T \to +\infty} \frac{\mathbb{E}\left(\varphi(X_T)e^{-\int_0^T E_L(X_s)ds}\right)}{\mathbb{E}\left(e^{-\int_0^T E_L(X_s)ds}\right)} = \frac{\displaystyle\int_{\Omega_\theta} \varphi\,\psi_\theta^*\psi^{\mathrm{I}}}{\displaystyle\int_{\Omega_\theta} \psi_\theta^*\psi^{\mathrm{I}}}, \quad (34)$$

$$\lim_{T \to +\infty} -\frac{1}{T}\ln\mathbb{E}\left(e^{-\int_0^T E_L(X_s)ds}\right) = E_\theta^*, \quad (35)$$

15

where the so-called local energy is defined by (26). The proof of the latter probabilistic interpretation (34)-(35) is similar to the proof of Lemma 3.1, and is based on the mapping of the Hamiltonian H to a weighted $\mathbb{L}^2$ space through:

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

Details and assumptions can be found in [6].

Next the probabilistic interpretation of the shape derivative given by formula (11) is proven.

Proposition 3.2. *Assume W_0 is distributed according to $\frac{\psi_{\mathrm{init}}(x)dx}{\int \psi_{\mathrm{init}}(x)dx}$ where $\psi_{\mathrm{init}} \in L^2(\Omega_\theta)$, and ψ_{init} is non vanishing in each connected component of Ω_θ. Assume the boundary $\partial\Omega_\theta$ is smooth and uniformly Lipschitz, and consider the measure defined for any $\lambda < E_\theta^*$ by (10). Then (11) holds.*

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

$$
\begin{cases}
\partial_t h_t(\varphi) = -(H-\lambda)(h_t(\varphi)) \\
h_t(\varphi)|_{\partial\Omega_\theta} = \varphi|_{\partial\Omega_\theta} \\
h_0(\varphi) = \varphi
\end{cases}
\tag{36}
$$

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

$$
\begin{cases}
(H-\lambda)(h_\infty(\varphi)) = 0 \\
h_\infty(\varphi)|_{\partial\Omega_\theta} = \varphi.
\end{cases}
\tag{37}
$$

This is a classical consequence of spectral theory, but we recall briefly the basic arguments. Existence of a smooth solution of (37) follows from the fact that $(H-\lambda)^{-1}$ can be extended to a bounded operator of $\mathbb{L}^2(\Omega_\theta)$ so that:

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

which indeed is solution of (37). Note that in the above $(H-\lambda)^{-1}$ implicitly refer to the operator with homogenous Dirichlet boundary condition, so that $(H-\lambda)^{-1} \circ (H-\lambda) \neq \mathrm{Id}$ when operating on test function with inhomogenous boundary conditions. Then the homogenous solution of (36) has been solved using spectral decomposition in (32), proving that such homogenous solution

is unique and vanishes exponentially fast when $\lambda < E_\theta^*$. The latter analysis of the homogenous case proves uniqueness in (36)-(37), as well as exponential long time convergence of the time dependent equation (36).

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

$$
\frac{\int_{\Omega_\theta} h_\infty(\varphi)\psi_{\text{init}}}{\int_{\Omega_\theta} \psi_{\text{init}}} := \mathbb{E}\left(\varphi(W_\tau)\mathbb{1}_{\tau<+\infty}e^{-\int_0^\tau (V(W_s)-\lambda)\,ds}\right) \tag{38}
$$

where $h_\infty(\varphi)$ is the solution to the elliptic partial differential equation with inhomogenous Dirichlet conditions (37). Indeed, using (31) with test function $h_{T-t}(\varphi)$, and stopped at time $T \wedge \tau = \inf(T, \tau)$ yields:

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

Now the event $\mathbb{1}_{\tau=+\infty}$ has null probability if $t \mapsto W_t$ is recurrent, and if the latter is transient then using $\lim_\infty V = +\infty$:

$$
\lim_{T\to+\infty} \mathbb{E}\left(\mathbb{1}_{\tau=+\infty}e^{-\int_0^{T\wedge\tau}(V(W_s)-\lambda)\,ds}\right) = 0;
$$

so that taking the limit $T \to +\infty$ leads to (38).

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 ψ_θ^* is the groundstate:

$$
\psi_\theta^* = \frac{1}{E_\theta^* - \lambda}(H - \lambda)(\psi_\theta^*),
$$

so that integration by parts yields:

$$
\begin{aligned}
\int_{\Omega_\theta} h_\infty(\varphi)\psi_\theta^* &= -\frac{1}{E_\theta^* - \lambda}\int_{\partial\Omega_\theta} \frac{1}{2}\varphi\nabla\psi_\theta^* \cdot n\,d\sigma \\
&\quad + \frac{1}{E_\theta^* - \lambda}\int_{\Omega_\theta} \frac{1}{2}\nabla\psi_\theta^* \cdot \nabla h_\infty(\varphi) + (V - \lambda)h_\infty(\varphi)\psi_\theta^* \\
&= -\frac{1}{E_\theta^* - \lambda}\int_{\partial\Omega_\theta} \frac{1}{2}\varphi\nabla\psi_\theta^* \cdot n\,d\sigma + 0.
\end{aligned}
$$

17

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

$$\mathbb{E}\left(\varphi(W_\tau)\mathbb{1}_{T\leq\tau<+\infty}\mathrm{e}^{-\int_0^\tau(V(W_s)-\lambda)\,ds}\right) =$$
$$\mathbb{E}\left(\mathbb{E}\left(\varphi(\widetilde{W}_{\widetilde\tau})\mathbb{1}_{\widetilde\tau<+\infty}\mathrm{e}^{-\int_0^{\widetilde\tau}(V(\widetilde{W}_s)-\lambda)\,ds}\big|\widetilde{W}_0=W_T\right)\mathbb{1}_{T\leq\tau}\mathrm{e}^{-\int_0^T(V(W_s)-\lambda)\,ds}\right),$$

where $((\widetilde{W}_t)_{t\geq0},\widetilde\tau)$ is an independent copy of $(W_t)_{t\geq0},\tau)$. Together with *Step* 2 and *Step* 3 with the probabilistic interpretation of the groundstate ψ_θ^* in (10)-(8), it completes the proof. $\qquad\square$

4   Fermion groundstates

Consider the Schrödinger operator (1). Fermionic groundstates (ψ_F^*, E_F^*) associated to (1) are defined with respect to a finite symmetry group $\mathcal{S}\subset O(\mathbb{R}^d)$ 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

$$\forall S\in\mathcal{S}\quad V\circ S=V.$$

Fermion systems appears in the special case when $\mathcal{S}$ 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:

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

"Fermion" groundstates are then the solutions to the variational problem (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:

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

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

Since in this paper, we restrict to smooth operators with compact resolvent, the existence of (ψ_F^*, E_F^*) follows directly from the fact that the spectrum is discrete, and smoothness from the fact ψ_F^* satisfies an eigenvalue problem of a smooth elliptic operator (see [20]).

Remark 4.2 (The sign problem). Computing directly (ψ_F^*, E_F^*) using Monte-Carlo methods is an untractable 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 (32) with:

- a skew-symmetric test function ψ^{I},

- a non-symmetric positive initial condition $\psi_{\mathrm{init}} > 0$,

yields:

$$\mathbb{E}\left(\psi^{\mathrm{I}}(W_T)\mathrm{e}^{-\int_0^T V(W_s)ds}\right) = \sum_{\psi^{*,n} \in \mathcal{H}} \mathrm{e}^{-E^{*,n}T} \frac{\int_{\mathbb{R}^d} \psi^{*,n} \psi_{\mathrm{init}}}{\int_{\mathbb{R}^d} \psi_{\mathrm{init}}} \int_{\mathbb{R}^d} \psi^{*,n} \psi^{\mathrm{I}},$$

and in principle if φ is symmetric:

$$\lim_{T \to +\infty} \frac{\mathbb{E}\left(\varphi(W_T)\psi^{\mathrm{I}}(W_T)\mathrm{e}^{-\int_0^T V(W_s)ds}\right)}{\mathbb{E}\left(\psi^{\mathrm{I}}(W_T)\mathrm{e}^{-\int_0^T V(W_s)ds}\right)} = \frac{\int_{\mathbb{R}^d} \psi_{\mathrm{F}}^* \psi^{\mathrm{I}} \varphi}{\int_{\mathbb{R}^d} \psi_{\mathrm{F}}^* \psi^{\mathrm{I}}}.$$

Unfortunately, this computation relies crucially on the rate of vanishing of the normalisation which is due to the non-symmetry of the initial condition $\psi_{\mathrm{init}} > 0$. 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, (ψ_F^*, E_F^*) can be approximated using a hybrid methodology in two steps. First, a *trial wave function* is obtained using an analytical parametrization of $\mathcal{H}$, usually of the form:

$$\psi_{\alpha,\theta}^I := J_\alpha \psi_\theta^{\mathrm{skew}}. \tag{39}$$

In the above, $J_\alpha > 0$ is a strictly positive symmetric part called the *Jastrow factor* and parametrized by $\alpha \in \mathbb{R}^n$; and $\psi_\theta^{\mathrm{skew}}$ is the skew-symmetric part parametrized by $\theta \in \mathbb{R}^p$. The most convenient numerical optimization method is then used to solve (13) in the space of parameters formed by (α, θ), and the solution of the optimization procedure is denoted with the parameters (α_0, θ_0). This yields the set of functions $\left\{\psi_\theta^{\mathrm{I}}\right\}_{\theta \in \mathbb{R}^p}$ defining the nodal domains through

$$\psi_\theta^{\mathrm{I}} := J_{\alpha_0} \psi_\theta^{\mathrm{skew}}.$$

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

$$\det\left(\phi_j(x_i)\right)_{i,j=1\ldots N},$$

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

The link with Section 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 $(\psi_\theta^{\mathrm{FN}}, E_\theta^{\mathrm{FN}}) = (\psi_\theta^*, E_\theta^*)$ of the variational problem with Dirichlet conditions (15). Such a computation is made using the probabilistic interpretations (8)-(9), or more usually in practice using the variant (34)-(35). This method is known under the DMC acronym (Diffusion Monte-Carlo) in Computational Chemistry, and has been widely studied, see for instance [21, 2].

As explained in the introduction, the key problem is that the nodal surface $\partial \mathcal{N}_\theta = (\psi_\theta^{\mathrm{I}})^{-1}(0)$ may be different from $(\psi_F^*)^{-1}(0)$. The open question is thus now to carrry out numerical methods associated to the variational problem (17). In this context, a direct minimization of E_θ^{FN} in (17) requires the computation of:

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

using formula (4) and the probabilistic interpretation (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 (14) is a smooth manifold, a sufficient condition being:

$$\forall x \in \partial \mathcal{N}_\theta, \quad \nabla \psi_\theta^{\mathrm{I}}(x) \neq 0. \tag{40}$$

Moereover, we assume that the assumptions of Section 2 on the mapping $\theta \mapsto \mathcal{N}_\theta$ apply, namely that there is a diffeomorphic mapping $\theta \mapsto R_\theta$ associated to the domain $\mathcal{N}_\theta$ such that (29) holds.

A fundamental remark concerns the symmetry of the normal derivative of the fixed node groundstate $\nabla^+ \psi_\theta^{\mathrm{FN}}.n_+$, (where ∇^+ and n_+ refer to the domain $\mathcal{N}_\theta^+$, and will be defined below), or equivalently the measure $\mu_{\theta,\lambda}^*$ in (11) defined on $\partial \mathcal{N}_\theta$ for stochastic processes evolving in $\mathcal{N}_\theta^+$. The latter indeed presents a *symmetry breaking*, in the sense that they are *only* invariant by the action of the symmetry sub-group

$$\mathcal{S}^+ := \{ S \in \mathcal{S} \,|\, \det(S) = 1 \}$$

on $\partial \mathcal{N}_\theta$. Before going further, we will precise notations in appropriate definitions and lemmas.

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

$$S(\partial \mathcal{N}_\theta) = \partial \mathcal{N}_\theta.$$

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

$$\psi_\theta^{\mathrm{I}}(S(x)) = -\psi_\theta^{\mathrm{I}}(x)$$

so that $\psi_\theta^{\mathrm{I}}(x) = 0$ is equivalent to $\psi_\theta^{\mathrm{I}}(S(x)) = 0$. $\square$

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

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

Two traces of $\nabla \psi$ on $\partial \mathcal{N}_\theta$ can then be defined depending if the reference domain is $\mathcal{N}_\theta^+$ or $\mathcal{N}_\theta^-$. The definition of associated symmetric and skew-symmetric traces then follows.

Definition 5.2. *Let ψ a skew-symmetric function in $\mathbb{R}^d$ with smooth restrictions and Dirichlet boundary conditions on $\overline{\mathcal{N}}_\theta^+$ and $\overline{\mathcal{N}}_\theta^-$. n_+ denotes the exterior normal of $\partial \mathcal{N}_\theta^+$, n_- denotes the exterior normal of $\partial \mathcal{N}_\theta^-$, so that:*

$$n_+ = -n_-. \tag{41}$$

21

$\nabla^+\psi \cdot n_+$ denotes the exterior normal derivative of ψ in $\mathcal{N}_\theta^+$, and $\nabla^-\psi \cdot n_-$ the exterior normal derivative of ψ in $\mathcal{N}_\theta^-$. Then the skew-symmetrization of the normal derivative is defined by

$$\nabla^{\text{sk}}\psi \cdot n = \frac{1}{2}\left(\nabla^+\psi \cdot n_+ + \nabla^-\psi \cdot n_-\right), \tag{42}$$

and the plain symmetrization with respect to n_+ is defined by (22).

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

Lemma 5.3. *Let ψ a skew-symmetric function in $\mathbb{R}^d$ with smooth restrictions on $\overline{\mathcal{N}}_\theta^+$ and $\overline{\mathcal{N}}_\theta^-$, and Dirichlet boundary conditions on $\partial\mathcal{N}_\theta$. Then $\nabla^{\text{sk}}\psi \cdot n$ is skew-symmetric, and $\nabla^{\text{sy}}\psi \cdot n_+$ is symmetric, under the action of $\mathcal{S}$ on $\partial\mathcal{N}_\theta$. If $\varphi \in C_c^\infty(\mathbb{R}^d)$, one has the integration by parts formula:*

$$\frac{1}{2}\int_{\mathbb{R}^d} \nabla\varphi \cdot \nabla\psi = \int_{\partial\mathcal{N}_\theta} \varphi\nabla^{\text{sk}}\psi \cdot n\, d\sigma - \int_{\mathbb{R}^d} \varphi\frac{\Delta}{2}(\psi). \tag{43}$$

Moreover, $\nabla^{\text{sk}}\psi \cdot n = 0$ on $\partial\mathcal{N}_\theta$ if and only if $\nabla^+\psi \cdot n_+$ or $\nabla^-\psi \cdot n_-$ are symmetric. In the opposite case, $\nabla^+\psi \cdot n_+$ and $\nabla^-\psi \cdot n_-$ are invariant only under the action of the special sub-group $\mathcal{S}^+$ and a "symmetry breaking" occurs.

Proof. By skew-symmetry of ψ, one has for any $S \in \mathcal{S}$ in $\mathcal{N}_\theta^+ \cup \mathcal{N}_\theta^-$:

$$S\,\nabla\psi \circ S = \det(S)\,\nabla\psi,$$

and then on $\partial\mathcal{N}_\theta$,

$$S\,n_+ \circ S = \det(S)\,n_+$$

$$S\,n_- \circ S = \det(S)\,n_-,$$

so that since $S^T S = \text{Id}$, it yields on $\partial\mathcal{N}_\theta$ for any $S \in \mathcal{S}$:

$$\begin{cases} \nabla^+\psi \cdot n_+ \circ S = \nabla^+\psi \cdot n_+ & \text{if } \det(S) = 1 \\[2mm] \nabla^-\psi \cdot n_- \circ S = \nabla^-\psi \cdot n_- & \text{if } \det(S) = 1 \\[2mm] \nabla^+\psi \cdot n_+ \circ S = -\nabla^-\psi \cdot n_- & \text{if } \det(S) = -1. \end{cases} \tag{44}$$

This yields the symmetry properties of $\nabla\psi^{\text{sk}} \cdot n$ and $\nabla\psi^{\text{sy}} \cdot n_+$ with respect to $\mathcal{S}$, and of $\nabla^+\psi \cdot n_+$ and $\nabla^-\psi \cdot n_-$ with respect to $\mathcal{S}^+$.

The integration by parts formula is obtained by applying seperately on $\mathcal{N}_\theta^+$ and $\mathcal{N}_\theta^-$ the classical Green's identity.

Finally from the symmetry properties (44), $\nabla^-\psi \cdot n_- = -\nabla^+\psi \cdot n_+$ on $\partial\mathcal{N}_\theta$ if and only if $\nabla^+\psi \cdot n_+ \circ S = \nabla^+\psi \cdot n_+$ or $\nabla^-\psi \cdot n_- \circ S = \nabla^-\psi \cdot n_-$ for any $S \in \mathcal{S}$ with $\det(S) = -1$. $\qquad\square$

One can now apply these remarks to the fixed node groundstate $\psi_\theta^{\mathrm{FN}}$ solution of (15).

Lemma 5.4. *Let $\psi_\theta^{\mathrm{FN}}$ be the solution of (15) with a smooth boundary $\partial\mathcal{N}_\theta$. Then the following assertions are equivalent:*

1. *$\psi_\theta^{\mathrm{FN}}$ is an eigenfunction of H in $\mathbb{L}^2(\mathbb{R}^d)$.*

2. *$\nabla^{\mathrm{sk}}\psi_\theta^{\mathrm{FN}} \cdot n$, as defined by (42), vanishes on $\partial\mathcal{N}_\theta$.*

3. *$\nabla^+\psi_\theta^{\mathrm{FN}} \cdot n_+$ or $\nabla^-\psi_\theta^{\mathrm{FN}} \cdot n_-$ are symmetric under the action of $\mathcal{S}$ on $\partial\mathcal{N}_\theta$.*

Proof. The third point and the second point are equivalent by Lemma 5.3. Now, if $\varphi \in C_c^\infty(\mathbb{R}^d)$, integration by parts (43) yields:

$$\int_{\mathbb{R}^d} (H - E_\theta^{\mathrm{FN}})(\varphi)\psi_\theta^{\mathrm{FN}} = \int_{\partial\mathcal{N}_\theta} \varphi\nabla^{\mathrm{sk}}\psi_\theta^{\mathrm{FN}} \cdot n d\sigma,$$

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

In the present context, the formula of the shape derivative of the Dirichlet groundstate can be written with symmetrized normal derivatives. We denote the shape derivative r_θ^+ (resp. r_θ^-) of the nodes as defined in (30), with respect to the normal n_+ (resp. n_-), so that

$$r_\theta^+ = -r_\theta^-.$$

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

$$\nabla_\theta E_\theta^{\mathrm{FN}} = -\frac{1}{\int_{\mathcal{N}_\theta^+} (\psi_\theta^*)^2} \int_{\partial\Omega_\theta} \left(\nabla^{\mathrm{sk}}\psi_\theta^{\mathrm{FN}} \cdot n\right)\left(\nabla^{\mathrm{sy}}\psi_\theta^{\mathrm{FN}} \cdot n_+\right) r_\theta^+ \, d\sigma. \qquad (45)$$

Moreover, $\nabla_\theta E_\theta^{\mathrm{FN}} = 0$ for any parametrization $\theta \mapsto \partial\Omega_\theta$ of the nodes if and only if the fixed node groundstate $\psi_\theta^{\mathrm{FN}}$ is an eigenstate of H in $\mathbb{R}^d$.

Proof. The shape derivative formula (4) can be decomposed as the sum of the part due to $\mathcal{N}_\theta^+$ and to $\mathcal{N}_\theta^-$:

$$\nabla_\theta E_\theta^{\mathrm{FN}} = -\frac{1}{2\int_{\mathcal{N}_\theta^+ \cup \mathcal{N}_\theta^-} (\psi_\theta^*)^2} \left(\int_{\partial\mathcal{N}_\theta} \left(\left|\nabla^+\psi_\theta^{\mathrm{FN}}\right|^2 r_\theta^+ + \left|\nabla^-\psi_\theta^{\mathrm{FN}}\right|^2 r_\theta^- \right) d\sigma \right).$$

Since $\psi_\theta^* = 0$ on $\partial\mathcal{N}_\theta$, then $\left|\nabla^+\psi_\theta^{\mathrm{FN}}\right| = \left|\nabla^+\psi_\theta^{\mathrm{FN}} \cdot n_+\right|$ and symmetrically $\left|\nabla^-\psi_\theta^{\mathrm{FN}}\right| = \left|\nabla^-\psi_\theta^{\mathrm{FN}} \cdot n_-\right|$. We get:

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

where in the last line we have used $r_\theta^+ = -r_\theta^-$. Then (45) follows. As a consequence, $\nabla_\theta E_\theta^{\mathrm{FN}} = 0$ if and only if $\nabla^+\psi_\theta^{\mathrm{FN}} \cdot n_+ = \nabla^-\psi_\theta^{\mathrm{FN}} \cdot n_-$, which is equivalent by Lemma 5.4 to the fact that $\psi_\theta^{\mathrm{FN}}$ is an eigenfunction. $\square$

We can now state the main result of this paper, which consists in a characterization of the nodes of eigenstates through the symmetry of a random stopped process, and suggests a method to evaluate the shape derivative of the fixed node energy $\nabla_\theta E_\theta^{\mathrm{FN}}$.

Theorem 5.6. *Consider a Brownian motion $t \mapsto W_t^+$ in $\mathcal{N}_\theta^+$, and τ^+ the hitting time of $\partial\mathcal{N}_\theta$. Consider the measure $\mu_{\theta,\lambda}^{\mathrm{FN}}$ on $\partial\mathcal{N}_\theta$ defined for any $\lambda < E_\theta^{\mathrm{FN}}$ by (20). Then the fixed node groundstate $\psi_\theta^{\mathrm{FN}}$ is an eigenfunction of H in $\mathbb{R}^d$, if and only if $\mu_{\theta,\lambda}^{\mathrm{FN}}$ is invariant under the symmetry group $\mathcal{S}$. Moreover, the shape derivative of the fixed node energy E_θ^{FN} is given by (21).*

Proof. The proof consists in a direct application of Lemma 3.1 and Proposition 3.2. Indeed, (11) yields the identity of measures on $\partial\mathcal{N}_\theta$:

$$d\mu_{\theta,\lambda}^{\mathrm{FN}} = -\frac{\nabla^+\psi_\theta^{\mathrm{FN}} \cdot n_+ \, d\sigma}{2(E_\theta^{\mathrm{FN}} - \lambda)\int_{\mathcal{N}_\theta^+} \left(\psi_\theta^{\mathrm{FN}}\right)^2};$$

and Lemma 5.4 enables to characterize eigenfunctions of H from the symmetry of $\eta_\theta^{\mathrm{FN},\lambda}$. Next, since r_θ^+ on $\partial\mathcal{N}_\theta$ is skew-symmetric, and the integral on $\partial\mathcal{N}_\theta$ of integrable skew-symmetric functions vanishes, it yields:

$$\int_{\partial\mathcal{N}_\theta} r_\theta^+ \nabla^{\mathrm{sy}}\psi_\theta^{\mathrm{FN}} \cdot n_+ \nabla^+\psi_\theta^{\mathrm{FN}} \cdot n_+ \, d\sigma = \int_{\partial\mathcal{N}_\theta} r_\theta^+ \nabla^{\mathrm{sy}}\psi_\theta^{\mathrm{FN}} \cdot n_+ \nabla^{\mathrm{sk}}\psi_\theta^{\mathrm{FN}} \cdot n \, d\sigma.$$

Then (45) with (34) yield the result (21). $\square$

24

6   Comments on Monte-Carlo methods

Theorem 5.6 suggests a Monte-Carlo general strategy to approximate the fixed node energy variation (21) by using the probabilistic interpretation (20)-(19). As already commented in the introduction, computing (20) with Monte-Carlo methods is a well known topic, and computing (19) can be carried out using independent stopped processes. In fact, the most straightforward limitation of the method consists in approximating on the nodes $\partial \mathcal{N}_\theta$ the symmetrized gradient

$$\nabla^{\mathrm{sy}} \psi_\theta^{\mathrm{FN}} \sim \nabla \psi_\theta^n$$

by a sequence $\psi_\theta^n \xrightarrow{n \to +\infty} \psi_\theta^{\mathrm{FN}}$ 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 $\nabla \psi_\theta^n$ converges in a pointwise sense towards $\nabla^{\mathrm{sy}} \psi_\theta^{\mathrm{FN}}$, which is necessary when approximating the measure $\eta_\theta^{\mathrm{FN}, \lambda}$ using a Monte-Carlo sample. However, in the context of optimization, it is fundamental to remark that the stationary nodal surfaces solution of $\nabla_\theta E_\theta^{\mathrm{FN}} = 0$ for any shape derivative are not modified when approximating $\nabla^{\mathrm{sy}} \psi_\theta^{\mathrm{FN}} \cdot n_+$ by any symmetric field $\nabla \psi_\theta^n \cdot n^+$. This is the meaning of the following proposition.

Proposition 6.1. *Consider any smooth skew-symmetric function ψ_θ^{I} in $\mathbb{R}^d$ with zeros given by $\partial \mathcal{N}_\theta$. Then the approximation $\widehat{\nabla_\theta E_\theta^{\mathrm{FN}}}$ given by (23) yields (24)-(25). Moreover the fixed point equation*

$$\widehat{\nabla_\theta E_\theta^{\mathrm{FN}}} = 0$$

holds for any shape variation r_θ^+ of the nodes $\partial \mathcal{N}_\theta$ if and only if $\psi_\theta^{\mathrm{FN}}$ is an eigenfunction of the Hamiltonian H in $\mathbb{R}^d$.

Proof. Since $\partial \mathcal{N}_\theta$ is defined implicitly by $\psi_\theta^{\mathrm{I}} = 0$, deriving $\psi_{\theta+h}^{\mathrm{I}}(x + r_{\theta+h}^+ n_+) = 0$ with respect to h yields on $\partial \mathcal{N}_\theta$:

$$\nabla_\theta \psi_\theta^{\mathrm{I}} + r_\theta^+ \nabla \psi_\theta^{\mathrm{I}} \cdot n_+ = \nabla_\theta \psi_\theta^{\mathrm{I}} - r_\theta^+ \left| \nabla \psi_\theta^{\mathrm{I}} \right| = 0,$$

where in the last line we have used the fact that $\nabla \psi_\theta^{\mathrm{I}} \cdot n_+ = - \left| \nabla \psi_\theta^{\mathrm{I}} \right|$ on $\partial \mathcal{N}_\theta$ since $\mathcal{N}_\theta^+ = \left\{ x \in \mathbb{R}^d \,|\, \psi_\theta^{\mathrm{I}}(x) > 0 \right\}$. This gives (24). Next (11) yields:

$$\int_{\partial \mathcal{N}_\theta^+} \nabla_\theta \psi_\theta^{\mathrm{I}} \, d\mu_{\theta,\lambda}^{\mathrm{FN}} = \quad - \frac{1}{2(E_\theta^{\mathrm{FN}} - \lambda) \displaystyle\int_{\mathcal{N}_\theta^+} \psi_\theta^{\mathrm{FN}}} \int_{\partial \mathcal{N}_\theta} \nabla_\theta \psi_\theta^{\mathrm{I}} \nabla^+ \psi_\theta^{\mathrm{FN}} \cdot n_+ \, d\sigma,$$

25

so that integration by parts in $\mathcal{N}_\theta^+$ gives (25).

Finally remark that $\nabla \psi_\theta^{\mathrm{I}} \cdot n_-$ is symmetric and that the shape derivative r_θ^+ is spanning all skew-symmetric field, so that the fixed point equation is equivalent to the fact that μ_θ^{FN} is symmetric. Then Lemma 5.4 enables to conclude. $\qquad \square$

Eventually, the following algorithm (without details) is suggested to estimate $\nabla_\theta E_\theta^{\mathrm{FN}}$. It might be referred to as Nodal Monte-Calo (NMC).

Algorithm 6.2 (NMC). *Consider a parametrization of wave functions of the form* (39). *Then the following steps are suggested:*

1. *Optimize the symmetric "Jastrow" factor J_α according to the variational problem with fixed node* (15), *and denote ψ_θ^{I} the obtained trial function.*

2. *Generate a sample according to the probability distribution $\eta_\theta^{\mathrm{FN}}$ in $\mathcal{N}_\theta^+$ defined in* (20). *Use for instance a long time trajectory of a drifted stochastic process of the type* (33) *with Feynman-Kac weights as in formula* (34).

3. *Use the latter sample as initial conditions; and sample independent Brownian motions. Stop them when they hit the nodes $\partial \mathcal{N}_\theta$. Then compute the measure $\mu_{\theta,\lambda}^{\mathrm{FN}}$ according to* (20).

4. *Estimate the variations of the fixed node energy $\nabla_\theta E_\theta^{\mathrm{FN}}$ with* (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 (20) to reduce the variance caused by the exponential weights. However, the classical stochastic differential equation (33) where ψ^{I} vanishes at the nodes cannot be used. Indeed, the repulsive drift at the nodes prevent the process from hitting the latter, and (20) no longer holds. Instead of a skew-symmetric guiding function ψ^{I}, we propose to use as a drift a symmetric function ψ_B, for instance an approximation of the *Bosonic* groundstate $\psi_{\mathrm{B}}^* > 0$ of H in $\mathbb{R}^d$ solution of the eigenvalue problem

$$H(\psi_{\mathrm{B}}^*) = E_{\mathrm{B}}^* \psi_{\mathrm{B}}^*.$$

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 μ_θ^{FN} in (20) and another random process stopped on the nodes $\partial \mathcal{N}_\theta$, with a distribution denoted $\widetilde{\mu}_\theta^{\mathrm{FN}}$. The goal is to design $\widetilde{\mu}_\theta^{\mathrm{FN}}$ so that the following two features hold: first computing the average (24) with $\widetilde{\mu}_\theta^{\mathrm{FN}}$ always yields 0; second the coupling is perfect in the special case when μ_θ^{FN} is symmetric (that is to say the Monte-Carlo method computing μ_θ^{FN} and $\widetilde{\mu}_\theta^{\mathrm{FN}}$ 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

References

[1] R. Assaraf and M. Caffarel. *A pedagogical introduction to Quantum Monte Carlo*, volume 74. Springer, 2000.

[2] R. Assaraf, M. Caffarel, and A. Khelif. Diffusion Monte-Carlo with a fixed number of walkers. *Phys. Rev. E*, 61(4):4566–4575, 2000.

[3] Roland Assaraf and Michel Caffarel. Zero-variance zero-bias principle for observables in quantum Monte Carlo: Application to forces. *The Journal of Chemical Physics*, 119(20):10536–10552, 2003.

27

[4] A. Badinski, P. D. Haynes, and R. J. Needs. Nodal Pulay terms for accurate diffusion quantum Monte Carlo forces. *Phys. Rev. B*, 77(8):085111, Feb 2008.

[5] A. Badinski and R. J. Needs. Total forces in the diffusion Monte Carlo method with nonlocal pseudopotentials. *Phys. Rev. B*, 78(3):035134, Jul 2008.

[6] E. Cancès, B. Jourdain, and T. Lelièvre. Quantum monte-carlo simulations of fermions. a mathematical analysis of the fixed-node approximation. *Math. Mod. and Meth. in App. Sci.*, 16:1403–1440, 2006.

[7] E. Cancès, C. Le Bris, and Y. Maday. *Méthodes mathématiques en chimie quantique: Une introduction.* Springer-Verlag, 2006.

[8] Mosé Casalegno, Massimo Mella, and Andrew M. Rappe. Computing accurate forces in quantum Monte Carlo using Pulay's corrections and energy minimization. *The Journal of Chemical Physics*, 118(16):7193–7201, 2003.

[9] D. Ceperley, G. V. Chester, and M. H. Kalos. Monte-Carlo simulation of a many-fermion study. *Phys. Rev. B*, 16:3081–3099, October 1977.

[10] D. M. Ceperley. Fermion nodes. *Journal of Statistical Physics*, 63:1237–1267, June 1991.

[11] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45(7):566–569, Aug 1980.

[12] C. Costantini, E. Gobet, and N. El Karoui. Boundary sensitivities for diffusion processes in time dependent domains. *Appl. Math. Optim.*, 54(2):159–187, 2006.

[13] P. Del Moral. *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications.* Springer Series Probability and its Applications. Springer, 2004.

[14] P. Del Moral and L. Miclo. Branching and Interacting Particle Systems approximations of Feynman-Kac formulae with applications to nonlinear filtering. *Lecture notes in Mathematics*, 1729:1–145, 2000.

[15] P. Del Moral and L. Miclo. Particle approximations of Lyapounov exponents connected to Schrödinger operators and Feynman-Kac semigroups. *ESAIM Proba. Stat.*, 7:171–208, 2003.

[16] A. Doucet, N. de Freitas, and N.J. Gordon. *Sequential Monte-Carlo Methods in Practice*. Series Statistics for Engineering and Information Science. Springer, 2001.

[17] A. Doucet, P. Del Moral, and A. Jasra. Sequential monte carlo samplers. *J. Roy. Stat.. Soc. B*, 68(3):411–436, 2006.

[18] C. Filippi and C. J. Umrigar. Correlated sampling in quantum Monte Carlo: A route to forces. *Phys. Rev. B*, 61(24):R16291–R16294, Jun 2000.

[19] J. Garcia Melian and J. Sabina De Lis. On the perurbation of eigenvalues for the p-laplacian. *C.R. Acad. Sci. Paris, Série 1*, 332:893–898, 2001.

[20] D. Gildbarg and N.S.Trudinger. *Elliptic Partial Differential Equation of Second Order*. Springer-Verlag, 1983.

[21] B.L. Hammond, W.A. Lester, and P.J. Reynolds. *Monte Carlo Methods in Ab Initio Quantum Chemistry*. World Scientific, 1994.

[22] Huang Hongxin and Shubin Liu. An improved algorithm of fixed-node quantum monte carlo method with self-optimization process. *Journal of Molecular Structure: THEOCHEM*, 726(1-3):93 – 97, 2005.

[23] I. Karatzas, , and S.E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[24] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, Berlin, second edition, 1976. Grundlehren der Mathematischen Wissenschaften, Band 132.

[25] M. Reed and B. Simon. *Methods of modern mathematical physics. IV. Analysis of operators*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1978.

[26] M. Rousset. On the control of an interacting particle estimation of Schrödinger ground states. *SIAM J. Math. Anal.*, 38(3):824–844 (electronic), 2006.

[27] J. Toulouse, R. Assaraf, and C. J. Umrigar. Zero-variance zero-bias quantum Monte Carlo estimators of the spherically and system-averaged pair density. *The Journal of Chemical Physics*, 126(24):244112, 2007.

[28] J. Toulouse and C. J. Umrigar. Optimization of quantum Monte Carlo wave functions by energy minimization. *The Journal of Chemical Physics*, 126(8):084102, 2007.

[29] C. J. Umrigar and C. Filippi. Energy and Variance Optimization of Many-Body Wave Functions. *Physical Review Letters*, 94(15):150201–+, April 2005.