JOINT INSTITUTE FOR NUCLEAR RESEARCH
Dzhelepov Laboratory of Nuclear Problems

FINAL REPORT ON THE INTEREST PROGRAMME

Monte Carlo vs Two Component DFT

Supervisor:
Dr. Evgeni Popov

Student:
Zlatomir Zhelev,
Sofia University “St. Kliment Ohridski”

Participation period:
October 20 – November 30, Wave 13

Dubna, 2025

Quantum Monte Carlo methods in positron annihiliation spectrscopy

Zlatomir Zhelev

December 1, 2025

Abstract

Here we give a brief review of what has been achieved with QMC methods in PAS. One of the limitations is that only valence electrons are included. We suggest a new procedure in order to calculate the contribution to the momentum density of annihilating pairs from the semi-core electrons, so that this is not estimated by standard Two-component Density Functional Theory (TCDFT). This procedure has the potential of being applied not only in the case of PAS, but also in other areas of condensed matter physics, quantum chemistry and etc.

1 Introduction

1.1 Introduction to PAS and TCDFT

Positron annihilation spectroscopy (PAS) is an experimental technique that is applicable in a large number of situations that range from the study of vacancies in materials to high-T_c superconductors [1]. This methods is based on injecting positrons in a material that is probed. The positrons annihilate with the electrons, which produces the well-known 511-KeV photons that can be detected. The detection of these photons is what yields the information from the experiment. One of the most used methods up to date is the Two-component density functional theory (TCDFT). The work of Boroński and Niemann [2] lays the foundation of modern approaches.

Following [2] we are trying to minimize:

$$E\left[n_-, n_+\right] = F\left[n_-\right] + F\left[n_+\right] + \int V_e\left(\mathbf{r}\right)\left(n_- - n_+\right) d^3\mathbf{r} - \int \frac{n_-\left(\mathbf{r}\right) n_+\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|} d^3\mathbf{r} d^3\mathbf{r}' + E_c^{e-p}\left[n_-, n_+\right] \tag{1}$$

This is the energy functional, which has to be minimized under the conditions that the number of electrons and positrons remains the same. This can done with the method of Lagrangian multipliers. The first two terms (Hohenberg-Kohn functional) $F\left[n\right]$ are a short-hand for:

$$F\left[n\right] = T\left[n\right] + \frac{1}{2}\int \frac{n\left(\mathbf{r}\right) n\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|} d^3\mathbf{r} d^3\mathbf{r}' + E_{xc}\left[n\right] \tag{2}$$

Here the first term T is the kinetic energy, the second is the Coulomb term integral and the third E_{xc} being the exchange-correlation energy.

In (1) $V_e\left(\mathbf{r}\right)$ is an external potential and the integral has significance of electrostatic energy of interaction of the system with an outside source. The fourth term is the Coulomb interaction of the electrons and positrons. The fifth term E_c^{e-p} is the electron-positron correlation energy functional.

Now following [3,4] we can write for the total annihilation rate λ and the momentum probability density distribution ρ for annihilating pairs as:

$$\lambda = \pi r_0^2 c \sum_i \int \langle i | \hat{n}_-\left(\mathbf{r}\right) \hat{n}_+\left(\mathbf{r}\right) | i \rangle \, d^3\mathbf{r} \tag{3}$$

$$\rho\left(\mathbf{p}\right) = \pi r_0^2 c \left| \langle f | \hat{O}_a\left(\mathbf{p}\right) | i \rangle \right| \tag{4}$$

Here $\hat{n}_- = \hat{\psi}_- \hat{\psi}_-$ and $\hat{n}_+ = \hat{\psi}_+ \hat{\psi}_+$ are the electron and positron density operators respectively, $\hat{\psi}_-$ and $\hat{\psi}_+$ being the electron and positron field operators. The operator $\hat{O}_a$ in (4) is an effective annihilation operator:

$$\hat{O}_a\left(\mathbf{p}\right) = \int e^{-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}} \hat{\psi}_-\left(\mathbf{r}\right) \hat{\psi}_+\left(\mathbf{r}\right) d^3\mathbf{r} \tag{5}$$

But instead of using operators in calculations we can try writing the annihilation rate in terms of density of electrons and positrons calculated from DFT [5] as:

$$\lambda = \pi r_0^2 c \int n_-\left(\mathbf{r}\right) n_+\left(\mathbf{r}\right) \gamma\left(n_-\left(\mathbf{r}\right)\right) d^3\mathbf{r} \tag{6}$$

Here γ is a function known as the enhancement factor. Its origin lies in the fact that the electrons and positron are not independent. One can classically have an intuition that as the positron is positively charged then the electrons will be attracted to it. This correlation increases the annihilation rate and thus is reflected by the $\gamma\left(n_{-}\left(\mathbf{r}\right)\right)$ factor. The accuracy of the prediction of the annihilation rate depends on the quality of the enhancement factor that is used.

The probability distribution of annihilating pairs is calculated using DFT quantities [5] as:

$$\rho\left(\mathbf{p}\right) = \pi r_0^2 c \sum_i \gamma_i \left| \int e^{-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}} \psi_+\left(\mathbf{r}\right) \psi_i\left(\mathbf{r}\right) d^3\mathbf{r} \right|^2 \tag{7}$$

Here ψ_i are the electrons orbitals and ψ_+ is the positron orbital. Here γ_j is a state dependent enhancement factor. It is calculated as:

$$\gamma_i = \frac{\lambda_i}{\lambda_i^{IPM}} \tag{8}$$

Here λ_i is the annihilation rate coming only from the electron in the i-th state, which is:

$$\lambda_i = \pi r_0^2 c \int n_+\left(\mathbf{r}\right) n_i\left(\mathbf{r}\right) \gamma\left(n_-\left(\mathbf{r}\right)\right) d^3\mathbf{r}, \tag{9}$$

where n_i is the density only due to one electron in the i-th orbital. As for λ_i^{IPM} it is a special case of calculating the annihilation rate in the independent particle approximation. That is when we take the enhancement factor $\gamma\left(n_-\left(\mathbf{r}\right)\right) = 1$.

There have been softwares developed to solve the modified Kohn-Sham equations that are obtained from (1), as well as calculate the annihilation rates and the momentum probability density of annihilating pairs. One of the most used ones is ABINIT.

1.2 QMC in PAS

But there is no universal functional that works very well for all calculations. Thus there should be a search for a parameter-free or at least have as few parameters as possible. One of the recent candidates for this task are Quantum Monte Carlo (QMC) methods - such as Variational Monte Carlo (VMC) and Diffusion Monte Carlo (DMC).

For the first time QMC has been used for calculating the positron lifetimes of positron in solids by K. Simula, et. al. [6]. Using their notation, they consider the following wavefunction:

$$\Psi_{SJ}\left(\mathbf{R}\right) = e^{J\left(\mathbf{R}\right)} \left[\phi^l\left(\mathbf{r}_{i\uparrow}\right)\right] \left[\phi^m\left(\mathbf{r}_{j\downarrow}\right)\right] \phi\left(\mathbf{r}_+\right) \tag{10}$$

Where $J\left(\mathbf{R}\right)$ is the Jastrow factor, $[\phi]$ is a shorthand for Slater determinant of electron orbitals, $\uparrow$ signifies spin-up, $\downarrow$ signifies spin-down, $\phi\left(\mathbf{r}_+\right)$ is the positron orbital. The authors of the paper [6] have also considered backflow displacement, but have found that it is negligible and thus will not be discussed further.

The authors use twist averaging. It is a generalization of periodic boundary conditions that allows a particle to have a different phase when it is translated by a distance corresponding to the size of the periodic boundary conditions L_i [10]. It can be written as:

$$\psi\left(\mathbf{r} + L_i \mathbf{e}_i\right) = e^{i\theta_i} \psi\left(\mathbf{r}\right) \tag{11}$$

This can easily be generalized to many particles. Thus from (11) we are see that components of the Bloch vectors are modified to be:

$$k_i = \frac{2\pi n_i}{L_i} + \frac{\theta_i}{L_i} \tag{12}$$

Essentially (12) means that we are changing the positions of the Bloch vectors in k-space. Thus we are exploring displaced points in the k-space. When we use twist averaging of a quantity, i.e.

$$\langle A \rangle = \frac{\int A\left(\theta\right) d^n\theta}{\int d^n\theta} = \frac{1}{\left(2\pi\right)^n} \int A\left(\theta\right) d^n\theta \tag{13}$$

where $A\left(\theta\right)$ is the average of the quantity for the given twist angles θ, yields better convergence due do the fact that for calculating properties for an infinite system we must calculate an integral over the whole first Brillouine zone, and these displacements allow us to do so better. Hence it corrects finite size cell effects, and we may not need using such a large cell as otherwise.

As for the positron orbital they use the one corresponding to $\mathbf{k} = 0$. It is the only relevant as experiments [1] show that when a positron enters a sample it thermalizes very quickly and thus there is no reason to consider any other state than the one corresponding to $\mathbf{k} = 0$. The authors [6] then optimize the Jastrow factors using VMC, as for the optimization of the wave function through DMC they used the fixed node approximation. They then a result from [11] that the annihilation rate can be written as:

$$\lambda = \pi r_0^2 c \sum_{i=1}^{N_e} \frac{\langle\Psi|\, \hat{O}_i^s \delta\left(\mathbf{r}_i - \mathbf{r}_+\right)|\Psi\rangle}{\langle\Psi|\Psi\rangle} = 100.9 g\left(0\right)\frac{N_e^{\uparrow}}{V} \tag{14}$$

where N_e is the number of electrons in the simulation cell of volume V, $g\left(r\right)$ is the translationally and rotationally averaged the pair correlation function (PCF) and r_0 is the clasically calculated electron radius. Here and hereafter $\hat{O}_i^s$ is the spin-projection operator to the spin of the electron-positron par.

They then sample the PCF with QMC by binning the electron-positron distances. Due to the fact that very rarely do they have the distance between the electrons and protons approach zero, there is a very poor statistics for $r \approx 0$. Thus what they do is for an exponent of the form:

$$g\left(r\right) = \exp\left(p\left(r\right)\right) = \exp\left(a_0 + a_1 r + ... + a_N r^N\right) \tag{15}$$

The coefficient value of the coefficient a_1 has to be chosen so that the Kimball cusp condition for the PCF is satisfied [12], but in a different context. Here we want to obtain the cusp condition for the pair correlation function between an electron and a positron.

Following similar arguments as in the original paper [12] and due the fact we are looking at a single electron and a single positron from translational and rotational invariance of the Hamiltonian the wave function should depend only on the relative distance r between them $\Psi = \Psi\left(\mathbf{r}\right)$. Letting $\mu = \frac{m_e}{2}$ be the reduced mass we have:

$$\left(-\frac{\hbar^2}{2\mu}\frac{\partial^2}{\partial r^2} - \frac{\hbar^2}{\mu r}\frac{\partial}{\partial r} - \frac{e^2}{r}\right)\Psi = \epsilon\Psi \tag{16}$$

Where Ψ is the wave function containing the positron. Then if we suppose that the wave function is an analytic function and expanding:

$$\Psi\left(r\right) = a + br + O\left(r^2\right) \tag{17}$$

Using (17) in (16) and grouping the terms that behave as $O\left(\frac{1}{r}\right)$, we have:

$$\frac{\hbar^2 b}{\mu r} + \frac{e^2 a}{r} = 0 \implies \frac{b}{a} = \frac{1}{2a_0} \tag{18}$$

Now for the pair correlation function we have $g \propto \Psi^*\Psi$ from which we have:

$$g \propto a^2 + 2abr + O\left(r^2\right) \tag{19}$$

Which through differentiation at the 0, (18) and switching to atomic units:

$$\left.\frac{\partial g}{\partial r}\right|_{r=0} = -g\left(0\right) \tag{20}$$

In the original paper there is $-\frac{1}{a_0}$ due to not being in atomic units and because Kimball is considering the PCF cusp condition for electron-electron, whereas here we want electron-positron.

From (15) and (19) we can easily calculate that:

$$a_1 = -1 \tag{21}$$

This is why the authors [6] have set $a_1 = -1$. They have not said what order N polynomial they have fit. This depends on the amount of data one has, as for maximizing the accuracy we must use a high enough order N, but not so high that we have Runge's phenomenon [13].

They [6] have shown that on average they obtain better results than TCDFT for the computed lifetimes of positrons in different materials.

One year later in [7] the authors have attempted to compute the positon momentum density of annihilating pairs using QMC. They repeat the same procedure for calculating orbitals, twist averaging and everything else.

As for calculating the momemntum density of annihilating pairs they use the followin formula due to [14] (although it is in a slightly different form):

$$\rho\left(\mathbf{p}\right) = \pi r_0^2 c \int \Psi^*\left(\mathbf{r}_1, \mathbf{r}_1, ..., \mathbf{r}_N\right)\left(e^{\frac{i}{\hbar}\mathbf{p}\cdot\left(\mathbf{r}_+ - \mathbf{r}_1\right)}\hat{O}_i^s\right) \Psi\left(\mathbf{r}_+, \mathbf{r}_+, ..., \mathbf{r}_N\right) d^{3N}\mathbf{r} \tag{22}$$

Which can immediately written in a form that can be used for the Metropolis algorithm:

$$\rho\left(\mathbf{p}\right) = \pi r_0^2 c \left\langle \frac{\Psi^*\left(e^{\frac{i}{\hbar}\mathbf{p}\cdot\left(\mathbf{r}_+ - \mathbf{r}_1\right)}\hat{O}_i^s\right)\Psi}{\left|\Psi\left(\mathbf{r}_+, \mathbf{r}_1, ..., \mathbf{r}_N\right)\right|^2} \right\rangle_{|\Psi^2|} \tag{23}$$

The procedure is the multiply and divide by the quantity in the denominator and thus we obtain

(23) from (22). More about such procedures and QMC can be seen in [15].

After calculating (19) and projecting it to 1-D Doppler, they plot the data to be compared with both experiments and TCDFT methods. They show that for the most the part they achieve a higher accuracy for the most part in the distribution than standard DFT techniques.

1.3 Problems and directions for improvement

To cite the authors from [6]: "With pseudopotentials, the annihilation and screening interactions due to core electrons are not considered. We calculate reference DFT results and estimate the annihilation rate Γ_c due to core electrons using a number of enhancement functionals".

Also citing them from [7]: "We compute the core electron Doppler projections and their relative intensity as compared to total spectra with density functional theory (DFT) based methods.".

Thus in both cases they are not using only the QMC methods, but are also using DFT to estimate the contribution due to inner electrons. Also in [7] they mention that they are not considering elements who have d-electron subshells, because the annihilation rate there is non-negligible.

All of this combined with the high-quality results, they have obtained we think that the next step is model contribution from semi-core electrons.

2 Methods

2.1 Slater determinant approximations for Metropolis algorithm

We would want to have approximate the Slater determinant so that it can be computationally not so heavy when considering semi-core electrons. Calculating the Slater determinant through an LU decomposition scales as $O\left(N^3\right)$. This means that adding semi-core electrons can increase drastically the computational time.

We are going to be looking at semi-core electrons that have negligible overlap with valence and semi-core electrons outside of the atom to which they belong. Let us consider a system of l-atoms, each of them has m_l semi-core electrons (m_l might be 0) and n_l valence electrons. Semi-core electrons is a somewhat of a relative concept, but in this context these are the non-valence that we want to model. By the subscript sc we will specify semi-core electrons, by the subscript v we will specify valence electrons, by the superscript (k) we will specify the k-th atom. Thus $\psi_{v,a}^{(k)}$ specifies the orbital of the a-th valence electron on the k-th atom.

From now on for the sake of brevity we will write drop the l subscript and assume that if we are looking at the l-th atom then $n \equiv n_l$ is the number of valence electrons. The same holds for all other quantities as well. Then we can write a Slater determinant (up to a normalization constant) as:

$$\Psi\left(\{\mathbf{r}\}\right) = \begin{vmatrix} \psi_{sc,1}^{(1)}\left(\mathbf{r}_{sc,1}^{(1)}\right) & ... & \psi_{sc,1}^{(1)}\left(\mathbf{r}_{v,n}^{(l)}\right) \\ ... & & ... \\ \psi_{sc,m}^{(1)}\left(\mathbf{r}_{sc,1}^{(1)}\right) & ... & \psi_{sc,m}^{(1)}\left(\mathbf{r}_{v,n}^{(l)}\right) \\ ... & & ... \\ \psi_{v,1}^{(1)}\left(\mathbf{r}_{sc,1}^{(1)}\right) & ... & \psi_{v,1}^{(1)}\left(\mathbf{r}_{v,n}^{(l)}\right) \\ ... & & ... \\ \psi_{v,n}^{(l)}\left(\mathbf{r}_{sc,1}^{(1)}\right) & ... & \psi_{v,n}^{(l)}\left(\mathbf{r}_{v,n}^{(l)}\right) \end{vmatrix} \tag{24}$$

This condition that there is a negligible overlap between these electrons can be written as:

$$\begin{aligned} \int \psi_{sc,a}^{(i)*}\left(\mathbf{r}\right)\psi_{v,b}^{(j)}\left(\mathbf{r}\right)d^3\mathbf{r} \approx 0, \qquad i \neq j \\ \int \psi_{sc,a}^{(i)*}\left(\mathbf{r}\right)\psi_{sc,b}^{(j)}\left(\mathbf{r}\right)d^3\mathbf{r} \approx 0, \qquad i \neq j \end{aligned} \tag{25}$$

Hence we can physically assert the amplitude of finding a semi-core electron is negligible at the positions where the amplitude of finding an external valence electron is non-negligible and vice-versa. Or from (11) we can mathematically write:

$$\begin{aligned} \psi_{sc,a}^{(i)}\left(\mathbf{r}_{v,b}^{(j)}\right) \approx 0, \qquad i \neq j \\ \psi_{v,b}^{(i)}\left(\mathbf{r}_{sc,a}^{(j)}\right) \approx 0, \qquad i \neq j \end{aligned} \tag{26}$$

This allows us to make the matrix of the Slater determinant sparser, as we set most of the matrix elements on certain rows/columns as being equal to zero.

$$\begin{vmatrix} \psi_{sc,1}^{(1)}\left(\mathbf{r}_{sc,1}^{(1)}\right) & ... & \psi_{v,n}^{(1)}\left(\mathbf{r}_{v,n}^{(1)}\right) & ... & 0 \\ ... & & ... & & ... \\ \psi_{sc,m}^{(1)}\left(\mathbf{r}_{sc,1}^{(1)}\right) & ... & \psi_{sc,m}^{(1)}\left(\mathbf{r}_{v,n}^{(1)}\right) & ... & 0 \\ \psi_{v,1}^{(1)}\left(\mathbf{r}_{sc,1}^{(1)}\right) & ... & \psi_{v,1}^{(1)}\left(\mathbf{r}_{v,n}^{(1)}\right) & ... & \psi_{v,1}^{(1)}\left(\mathbf{r}_{v,n}^{(l)}\right) \\ ... & & ... & & \\ 0 & ... & 0 & ... & \psi_{v,n}^{(l)}\left(\mathbf{r}_{v,n}^{(l)}\right) \end{vmatrix} \tag{27}$$

This greatly reduces the amount of components of the Slater determinant. For example are modeling a material containing Ga, which can be written as $[\mathrm{Ar}]\,3d^{10}4s^{2}4p^{1}$, depending on what we call valence electrons it has either 1 or 3. Chemically it has three valence electrons. So if we are to consider the effects of the $3d^{10}$ subshell then the number of electrons that we have to treat for each atom increases from 3 to 13.

Due to the increase in complexity the direct calculation of the Slater determinant might not be a viable option. But using the approximations from (26) and (27), we can reduce the matrix to being calculable through the use of sparse matrix determinants.

In the case of weakly overlapping semi-core and and valence electrons we can assume that:

$$\begin{aligned} \psi_{sc,a}^{(i)}\left(\mathbf{r}_{v,b}^{(i)}\right) &\approx 0 \\ \psi_{v,a}^{(i)}\left(\mathbf{r}_{sc,b}^{(i)}\right) &\approx 0 \end{aligned} \tag{28}$$

Which through the use of the famous formula from linear algebra that:

$$\det\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \det\left(A\right)\det\left(B\right) \tag{29}$$

By repeated use we can factorize the Slater determinant into a product of semi-core electron determinants and valence electrons.

But we will not suppose that the following approximation holds for the rest of the method.

Note: We use the following approximation only for the case of approximating fast the value of the wavefunction for a certain configuration in the Monte Carlo sampling. If we were to consider calculating ab initio the wavefunction of the system in this manner we might not obtain results thar are sufficiently accurate, as despite being small there is an exchange-correlation energy due to outer valence electrons.

2.2 Jastrow factors

Standard DFT schemes, although they yield very good results, they use pseudopotentials in their calculations. Pseudopotentials might be calculated in different ways (for example Dirac-Fock equations) depending on the what accuracy we want to achieve. But what all of them have in common is the lack of singularity in the potential. This means that there will be no cusp in the wavefunctions. This we can handle through the use of Jastrow factors to improve the wavefunction. These factor will be parametrized and we can perform variational Monte Carlo to improve them.

When we are considering using Jastrow factors we then suppose that we can write the wavefunction as (up to a normalization constant):

$$\Psi = \exp\left(J\left(\{\mathbf{r}_i\}, \{\mathbf{R}_a\}\right)\right)\left[\phi_i\right] \tag{30}$$

The coordinates $\{\mathbf{R}_a\}$ here specify the coordinates of the nuclei.
We must satisfy the cusp condition so that there are no divergences in the average energy from the Hamiltonian. The first to consider the cusp in the wave functions is Kato [8]. From [9] we can write the following cusp conditions that should be used:

$$\left(\frac{\partial J}{\partial r_{ij}}\right)_{r_{ij}=0} = \frac{1}{2} \tag{31}$$

$$\left(\frac{\partial J}{\partial r_{ij}}\right)_{r_{ij}=0} = \frac{1}{4} \tag{32}$$

$$\left(\frac{\partial J}{\partial r_{ij}}\right)_{r_{ij}=0} = -Z \tag{33}$$

Here r_{ij} is the distance between the two particles considered (it can include electron-nucleus distances) and Z is the charge of the nucleus. It should be noted that r_{ij} is treated as an independent variable, i.e. we suppose that $J = J\left(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_{ij}\right)$. Equation (31) is for the antiparallel spin electron-electron cusp condition, (32) for the parallel spin electron-electron cusp condition and (33) is for the electron-nucleus cusp condition.

2.3 Monte Carlo optimization

Here we will be following mainly the material that is covered brilliantly in [15].

The idea behind Variational Monte Carlo (VMC) is that we parametrize the wavefunction and then we optimize the parameters so that we approach the ground state of the system.

The average of any operator in the continuum can be written as:

$$\left\langle \hat{A} \right\rangle = \int \psi\left(\mathbf{r}\right) \hat{A} \psi\left(\mathbf{r}\right) d^n\mathbf{r} \tag{34}$$

which we can easily rewrite as:

$$\langle A \rangle = \int \frac{\left(\psi\left(\mathbf{r}\right) \hat{A} \psi\left(\mathbf{r}\right)\right)}{\left|\psi\left(\mathbf{r}\right)\right|^2} \left|\psi\left(\mathbf{r}\right)\right|^2 d^n\mathbf{r} \tag{35}$$

Assuming that $\left|\psi\left(\mathbf{r}\right)\right|^2$ is normalized, then it is a probability distribution and we can rewrite this as:

$$\langle A \rangle = \left\langle \frac{\left(\psi\left(\mathbf{r}\right) \hat{A} \psi\left(\mathbf{r}\right)\right)}{\left|\psi\left(\mathbf{r}\right)\right|^2} \right\rangle_{\left|\psi\left(\mathbf{r}\right)\right|^2} \equiv \left\langle A_l\left(\mathbf{r}\right) \right\rangle_{\left|\psi\left(\mathbf{r}\right)\right|^2} \tag{36}$$

Now we can sample configurations in order to calculate the average value:

$$\langle A \rangle \approx \frac{1}{N} \sum_{i=1}^{N} A_l\left(\mathbf{r}_i\right) \tag{37}$$

We write the approximate sign as this equality is reached only in the limit $N \to \infty$.

Metropolis, et. al. [16] have proposed an algorithm for efficient Monte Carlo sampling. If we are at $\mathbf{r}$ we propose a new position $\mathbf{r}'$, then we have an acceptance probability for this move of the form:

$$p_{acc} = \min\left(1, \frac{\left|\psi^2\left(\mathbf{r}'\right)\right|}{\left|\psi^2\left(\mathbf{r}'\right)\right|}\right) \tag{38}$$

This follows from the detailed balance condition in the system. Physically it represents that an equilibrium state is reversible. Thus the Metropolis algorithm is applicable for systems in the equilibrium state only. This simple algorithm lays the foundation for many of the Monte Carlo algorithms.

As for VMC we can either try minimizing the energy or minimizing the variance of the energy. From theorem that can be found in most Quantum Mechanics texts, for example [17], through variational methods we always have the following relation:

$$E_{var} \geq E_0 \tag{39}$$

Where E_0 is the ground state energy. Hence lowering the energy of the system is equivalent to approaching the ground state wavefunction. The equality is reached only for the ground state wavefunction.

But we can write the variance of the energy as:

$$\sigma_E = \langle\psi| \left(\hat{H} - E_0\right)^2 |\psi\rangle\,, \tag{40}$$

assuming that the wavefunction is normalized, otherwise we must divide by it's modulus. We can easily see that the ground state of the system has zero variance. Thus instead of minimizing the energy we may minimize the variance.

There are many methods for both energy minimization and variance minimization and for a comprehensive overview one should definitely look at [15]. Here we will look at only the one that is to be of most use in crystals, namely Sorella's stochastic reconfiguration technique.

This technique is an equivalent to the gradient-descent method that has wide applications. First one has 'generalized forces', the negative gradient of the energy. If we are using s variational parameters then we may write the variational parameter vector $\boldsymbol{\alpha} = \left(\alpha_1, ..., \alpha_s\right)$, then we write the generalized force vector as:

$$\mathbf{f} = -\nabla_{\boldsymbol{\alpha}} E \Leftrightarrow f_i = -\frac{\partial E}{\partial \alpha_i} \tag{41}$$

Then using our physical intuition in order to reach the equilibrium state the system must follow the direction of the forces, which allows us to vary the coefficients as:

$$\alpha_i' = \alpha_i + \delta\alpha_i, \qquad \delta\alpha_i = \Delta f_i \tag{42}$$

Here Δ is a small proportionality constant determining the size of the step. Then by Taylor expanding the energy, we have:

$$E_{\boldsymbol{\alpha}'} = E_{\boldsymbol{\alpha}} + \sum_i \frac{\partial E}{\partial \alpha_i} \delta\alpha_i + O\left(\Delta^2\right) \tag{43}$$

When we choose Δ small enough we can neglect second and thereon terms in the expansion and using (41) and (42) we obtain:

$$E_{\boldsymbol{\alpha}'} - E_{\boldsymbol{\alpha}} = -\Delta \sum_i f_i^2 \leq 0 \tag{44}$$

Thus by consecutive iterations we can minimize the energy.

As for DMC good discussion can be found in [15, 20] as well as how the implementation is to be done. We will restrict ourselves only to a very samll discussion.

The main idea is that is based on a mathematical analogy between the Schrödinger equation and the diffusion equation. The Schrödinger equation can be thought of as a diffusion in imaginary time. We then sample a lot of walkers and evolve them according to the quantum dynamics of the system. The only problem is that this method scales exponentially with the number of electrons considered in the system, but the computation time is reduced by the so called Fixed-Node Approximation.

The reason why we want to have a very accurate density of the system is because we want our nodal surface to be as accurate as possible as it determines the accuracy of the DMC method.

2.4 Approximation of splitting the optimization procedure

Based on everything we have said in 2.1. we now make the following approximation. We will separate the semi-core electrons in an atom and the corresponding valence electrons from the rest of the system. Then we optimize the wave function so as to minimize energy \ variance of energy for the corresponding wavefunctions for electrons on a given atom. We can either use the potential obtained from DFT calculations, or we can rewrite the Hamiltonian as follows:

$$H = T + \left(V_{DFT}\left(\mathbf{r}\right) - V_{Ps}\left(\mathbf{r}\right)\right) + V_{nuc-el}\left(\mathbf{r}\right) \tag{45}$$

The following terms are - T - kinetic energy of all the electrons, V_{DFT} is the potential as computed from DFT, V_{Ps} is the pseudopotential of the atomic core along with the inner electrons, and V_{nuc-el} is a new approximation for the nucleus potential along with core electrons.

The idea behind the new Hamiltonian is that when going closer to the nucleus the pseudopotential may yield results that are not satisfactory as it gives the potential produced at large enough distances. The new approximate term V_{nuc-el} will depend from a given situation to another, so a general recipe can not be given. One can do the Monte Carlo calculations using a pseudopotential which is more accurate closer to the nucleus than the pseudopotential used in the DFT calculations.

After we have optimized the semi-core wavefunctions then we can move on to optimizing the Jastrow factors of the whole system containing - nuclei, valence electrons, semi-core electrons and the positron.

Here we can neglect Jastrow factors between valence electrons and semi-core electrons on different atoms, as well as between semi-core and semi-core electrons on different atoms. We won't optimize the Jastrow factors between valence and semi-core electrons on the same atom, as this we have done so, as said in the previous section. This is the idea behind the splitting of optimization procedure.

We can thus focus on optimizing positron-nucleus, electrons-nucleus, positron-electrons and maybe 3-body Jastrow factors.

2.5 Group theoretical aspects of Jastrow factors

2.5.1 Group theoretical aspects in crystals

Although we have (27) which will greatly speed up our calculation of the wavefunction, which is necessary for the Metropolis algorithm. The sheer amount of electrons requires a lot of Jastrow factors. However we can reduce this number from group theoretical arguments.

The Jastrow factors for semi-core and valence electrons around nucleus that are a linear combination of the elementary translation of the space group should be the same. As shifting by such translations can at most result in a change in phase of the wavefunction [18], but the Jastrow factors are real and thus they don't change.

In this way the space group of the material determines how many Jastrow factors are needed to be optimized for the electrons around a single atom.

Physically this is the case as the semi-core electrons experience the same potential from the valence electrons and all the other atoms and electrons in a crystal by this type of translations.

This implies that we may need to use QMC, for the electrons belonging to the same atom, only for a few atoms. This greatly reduces the amount of computational time, as this does not scale with how many cells we are to use.

Similar arguments can be used for the Jastrow factor between different valence electrons, valence electrons-nuclei, valence electrons-positron and positron-nuclei.

Note: Everything said up to now is valid only for the case that potential can truly be assumed to be periodic. It may not be periodic for example in the case when there are are transitions between different energetic levels for the valence electrons or there is a dislocation that occurs in the material. We will not be considering such cases.

2.5.2 Group theoretical aspects for molecules

Molecules are characterized by point groups, instead of space groups as in crystals. In this case we have to look only at the operations transforming one atom of the molecule into an equivalent, which by similar arguments should yield the same Jastrow factors between semi-core and valence electrons belonging to the same atom.

2.6 Self-consistent Monte Carlo optimization

After what has been said in 2.4. and 2.5. we now consider the possibility of a self-consistent Monte Carlo scheme. By this we perform the following steps:

1) Optimize the Jastrow factors between semi-core and valence electrons on the same atom.

2) Optimize the Jastrow factors between the positron, electrons and nuclei.

3) Repeat until the desired accuracy is achieved.

A possible problem with this procedure is that there might be a slow convergence, but this has to be tested with real calculations.

3 Steps to practical realization

1. One should first calculate the orbitals of the system through DFT. Twist averaging is suggested as it improves the accuracy of desired quantities and we want the nodal surface that is going to be used in DMC to be as accurate as possible, as this is the one of the sources of errors in the fixed-node approximation [15].

2. With the obtained orbitals, one should first change from the plane-wave basis to the blip basis [19], as in order to obtain the value at a point of the wavefunction we have to sum over all the plane-waves. This procedure becomes increasingly expensive with increase in accuracy of the wavefunction.

We can thus consider thus consider interpolating the wavefunction by splines. One of the most useful ones are the B-splines [13]. In our case the blip functions are cubic B-splines. In three dimensions around any points $\mathbf{r}$ there are only 64-blip coefficients, no matter how many plane-waves are used [19]. We can write the wavefunction in plane-wave representation and in blip basis representation at a given point $\mathbf{r}$, following the notation from [19], as:

$$\psi\left(\mathbf{r}\right)=\sum_{\mathbf{k}}c_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{r}} \tag{46}$$

$$\psi\left(\mathbf{r}\right)=\sum_{s}a_{s}\Theta_{s}\left(\mathbf{r}\right) \tag{47}$$

We can obtain coefficients from one representation into the other through:

$$a_s=\sum_{\mathbf{k}}c_{\mathbf{k}}\gamma_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{48}$$

where the coefficients $\gamma_{\mathbf{k}}$ are given by:

$$\gamma_{\mathbf{k}}=\frac{3}{k^4}\left(3-4\cos k-\cos 2k\right) \tag{49}$$

Here (48) represents nothing else but a Discrete Fourier Transform and we can calculate them efficiently and fast [19].

3. Next we have to incorporate the algorithm for calculating Slater determinants as in (27). Here we have to implement a sparse determinant calculation method.

4. We have to calculate the positron wave function. We want a wavefunction that is accurate

enough in the semi-core electrons region. This can be done in two ways.

The first is to calculate the electrons and positron orbitals, using a pseudopotential that treats the semi-core electrons as "valence" electrons. This can be done done with a suitably chosen pseudopotential, as if we choose one that treats the semi-core electrons as "frozen" in the core then the potential near that region is different and the wavefunction in that region will not be very accurate.

The second approach is to calculate the orbitals for the electrons (both valence and semi-valence) and then calculate the orbital of the positron when placed in such a system, supposing that the positron does not appreciably distrurb the electron density.

The second approach will be computationally more accurate and yield great errors, thus it is the preferable method.

5. Good Jastrow factors have to be chosen, as the accuracy of VMC depends on them [15]. They have to satisfy the described cusp conditions in **2.3**.

6. Now one has to perform the Quantum Monte Carlo methods. First we optimize with VMC the parameters in the Jastrow factors. One can use either energy minimization or variance minimization as described in **2.4**. We split the optimization procedure as described in **2.4**. From this we obtain the optimized wavefunction for the system Ψ_{VMC}.

Now we use the optimized wavefunction Ψ_{VMC} for the DMC. We are to use fixed-node approximation, which was our main reasons for using VMC in the first place as it improves the accuracy of this surface. We perform DMC with N-walkers.

7. Sample the integral considered in (23). In order to obtain the one dimensional distribution we have to sum over as in [7] in the following way:

$$\rho\left(p_z\right) = \int \rho\left(\mathbf{p}\right) dp_x dp_y \tag{50}$$

This is nothing more but the obtaining the one-dimensional probability distribution from the multidimensional functional.

4 Conclusion and outlook

Quantum Monte Carlo algorithms in PAS is an emerging method that is aimed at improving the TCDFT techniques, as it is parameter free. TCDFT will not be abolished as a method as we need an accurate initial wavefunction to obtain accurate Monte Carlo results. The articles originating this field [6,7] contain a lot of open problems that are thus to be solved.

Here we suggest a new algorithm as well as the steps to be performed in a calculation. It is aimed at modeling semi-core electrons annihilation through QMC methods. This method can be used in all other branches that have to model semi-core electrons in materials, the only restriction being the absence of a very strong correlation between semi-core electrons and valence electrons and small overlap with external electrons (25).

Of course concrete calculations have to be done to establish the accuracy of this algorithm, so that it can be compared with standard techniques such as DFT and the experimental data. This is the next step.

Some of the other open problems are for example modeling the annihilation with every electron in the material, including the deep core electrons. This task is very computationally heavy and algorithms reducing the time needed to execute all the calculations, would prove very beneficial.

QMC methods should also be developed to be used for calculation where traditional local functional DFT methods fail as for example the case of surface states [7].

5 Acknowledgements

I would like to thank Dr. Evgeni Popov and Dr. Aleksandar Donkov for the interesting and currently relevant topic they chose for this project, as well as for their support, tips that they provided and correction they have made.

6 References

[1] M. J. Puska, R. M. Nieminen, Theory of positrons in solids and on solid surfaces, Rev. Mod.

Phys. 66, 841 (1994)

[2] E. Boroński and R. M. Nieminen, Electron-positron density-functional theory, Phys. Rev. B 34, 3820 (1986)

[3] R. A. Ferrell, Theory of positron annihilation in solids, Rev. Mod. Phys. 28, 308 (1956)

[4] G. F. Gribakin, J. A. Young, and C. M. Surko, Positron-molecule interactions: Resonant attachment, annihilation, and bound states, Rev. Mod. Phys. 82, 2557 (2010)

[5] T. Torsti, et. al, Three real-space discretization techniques in electronic struc-ture calculations, arXiv:cond-mat/0601201 (2006)

[6] K. A. Simula, J. E. Muff, I. Makkonen, N. D. Drummond Quantum Monte Carlo study of positron lifetimes in solids, Phys. Rev. Lett. 129, 166403 (2022)

[7] K. A. Simula, J. Härkönen, I. Zhelezova, N. D. Drummond, F. Tuomisto, I. Makkonen, Quantum Monte Carlo study of Doppler broadening of positron annihilation radiation, Phys. Rev. B 108, 045201 (2023)

[8] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics, Commun. Pure Appl. Math. 10, 151 (1957)

[9] N. D. Drummond, M. D. Towler, R. J. Needs, Jastrow correlation factor for atoms, molecules, and solids, Phys. Rev. B 70, 235119 (2004)

[10] C. Lin, F. H. Zong, and D. M. Ceperley Twist-averaged boundary conditions in continuum quantum Monte Carlo algorithms, Phys. Rev. E 64, 016702

[11] G. F. Gribakin, J. A. Young, and C. M. Surko, Positron-molecule interactions: Resonant attachment, annihilation, and bound states, Rev. Mod. Phys. 82, 2557 (2010)

[12] J. C. Kimball, Short-range correlations and electron-gas response functions, Phys. Rev. A 7, 1648 (1973)

[13] Rainer Kress, Numerical Analysis, Springer (1998)

[14] G. Ryzhikh and J. Mitroy, Positron annihilation profiles for HPs and He (3Se) e+, Journal of Physics B: Atomic, Molecular and Optical Physics 32, 4051 (1999)

[15] Becca F, Sorella S. Quantum Monte Carlo Approaches for Correlated Systems. Cambridge University Press (2017)

[16] Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and Teller, E. 1957. Equations of state calculations by fast computing machines. J. Chem. Phys., 21, 1087–1092

[17] Weinberg S. Lectures on Quantum Mechanics. 2nd ed. Cambridge University Press, (2015)

[18] M. Tinkham, Group Theory and Quantum Mechanics, Dover (2003)

[19] D. Alfè and M. J. Gillan, Efficient localized basis set for quantum Monte Carlo calculations on condensed matter, Phys. Rev. B 70, 161101(R) (2004)

[20] Landau D, Binder K. A Guide to Monte Carlo Simulations in Statistical Physics. 5th ed. Cambridge University Press (2021)