

# Study of small atoms using the Stochastic Variational Method

David Christoffer Hansen  
20051747



Department of Physics and Astronomy  
Faculty of Science  
Århus University  
Denmark  
13<sup>th</sup> of August, 2008

# Study of small atoms using the Stochastic Variational Method

David Christoffer Hansen  
20051747

Department of Physics and Astronomy, Århus University

Submitted for the degree of Bachelor of Physics  
13<sup>th</sup> of August 2008

## Abstract

In this paper, the Ritz theorem is expanded into the Stochastic variational method, and a detailed study of the theoretical background is made. A lightweight implementation of the stochastic variational method is made, using spherical gaussians only, and the algorithms involved are discussed. The implementation uses the Cholesky-Jacobi eigensolver for solving the generalized eigenvalue problem, and Brent's method for testing the trial functions needed for the stochastic variational method. The groundstate energy is calculated for the first four atoms in the periodic table, all of which should be well described by the spherical gaussian. The results are then compared with various other theoretical and experimental values, providing a guideline for the accuracy of the method and the viability of the stochastic variational method calculations without the use of supercomputers.

# CONTENTS

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	SVM . . . . .	2
<b>2</b>	<b>Theory</b>	<b>4</b>
2.1	The basic problem . . . . .	5
2.2	Matrix elements . . . . .	6
2.3	Optimizing the parameters . . . . .	9
2.4	The not so few body problem . . . . .	9
2.5	A lower bound . . . . .	10
2.6	The implementation . . . . .	11
<b>3</b>	<b>Practicalities</b>	<b>13</b>
<b>4</b>	<b>Results</b>	<b>14</b>
4.1	proof of concept: Positronium . . . . .	14
4.2	Hydrogen . . . . .	14
4.3	Helium . . . . .	15
4.4	Lithium . . . . .	16
4.5	Beryllium . . . . .	17
<b>5</b>	<b>Improvements of the SVM</b>	<b>20</b>
	SVM . . . . .	20
<b>6</b>	<b>Conclusion</b>	<b>22</b>
	<b>Bibliography</b>	<b>23</b>

# CHAPTER 1

## INTRODUCTION

Numeric solutions to the Schrödinger equation has been sought almost since the invention of quantum mechanics, and have been used with increasingly greater accuracy since the 1950's and the dawn of electronic computing. Many methods have been used, and even more have been proposed, leading to solution of extreme accuracy, above and beyond what can be obtained in experiments for the few-body system. Today the most successful methods are the Quantum Monte Carlo methods and the various variational based calculations. Using these, modern physics have achieved theoretical bounds for the ground state energy with a precision far beyond the best experiments. All these methods have a common goal: To achieve the greatest accuracy possible, at the lowest computational cost possible with the greatest ease possible. Many fail in this, and many schemes that appear sound in the world of exact arithmetic perform poorly in a limited precision context, needing more than double precision floating points to counter numeric instability, which will encumber any computation, by a factor 10 or more. The most accurate models are often custom made for a particular few body system, or are corrections to more general methods. In recent years the variational methods based on an expansion of Gaussian functions have proven to give reliable and accurate results, among these the stochastic variational method, which has been used to calculate systems with 2-7 particles.

### 1.1 SVM

The stochastic variational method was originally invented by V. I. Kukulin and V M Krasnopolsky in 1977<sup>[17]</sup>, but was given relatively little attention at the time, mainly due to lack of computational power needed for the method to be effective. It was not until K. Varga and Y. Suzuki used the method in 1994<sup>[19]</sup> and later demonstrated the full potential of the method in an article from 1995<sup>[14]</sup> that common use began. The book published in 1998<sup>[1]</sup> gave examples of its use in practically every type of few-body problem currently being studied, even subnuclear problems, using relativistic corrections to the kinetic energy operator. With the exponential increase in the available computational power in recent years, it has become possible to perform accurate calculations without the use of a super-computer, making the method practical for use in everyday calculations of modest size. In this paper, a lightweight implementation of the stochastic variational method is made using a high level functional language, using only spherical Gaussian, and tested on various 2-5 body atomic problems. The limits of this simple method is explored, using Beryllium as the limit of what can reasonably be achieved without adding angular momentum com-

mutations and relativistic corrections.

The program is written in Haskell using Glasgow extensions and can be found at <http://www.phys.au.dk/~dch05/Bachelor/> .

All values given are in atomic units unless specifically stated otherwise.

# CHAPTER 2

## THEORY

The Ritz theorem, commonly known as the variational theorem, is one of the most useful theorems of quantum mechanics. It states that, for any normalized function  $\Psi$ , the following holds

$$\langle \Psi | \hat{H} | \Psi \rangle \geq E_1 \quad (2.1)$$

where  $\hat{H}$  is the Hamiltonian of a given quantum mechanical system, and  $E_1$  the ground state eigenenergy. A proof of this can be found in [2]. This can further be generalized to the Mini-Max theorem, which states that for any Hermitian operator  $\hat{H}$  with discrete eigenvalues  $E_1 \leq E_2 \leq \dots$ , the eigenvalues ( $\epsilon_1 \leq \epsilon_2 \leq \dots \leq \epsilon_k$ ) of H in the restricted subspace  $\mathcal{V}_k$  of a linear combinations of K independent functions, the following holds:

$$E_i \leq \epsilon_i \quad 1 \leq i \leq k \quad (2.2)$$

a proof can be found in [8]. Equality only holds when the exact eigenfunctions of  $\hat{H}$  are contained in  $\mathcal{V}_k$ .

The practical offspring of this is obvious. We can easily obtain an upper bound for any eigenvalue we like, simply by making a guess. If a an upper bound close to the actual eigenvalue is desired, this guess most obviously be qualified, and relatively close to the true eigenfunction of the system. For small systems this is a simple task, and is a common exercise in introductory courses in quantum mechanics, usually leading to the correct answer within one or two percent using a simple modified hydrogen wave-function for each electron. To compete with the accuracy achieved in modern experiments, more elaborate schemes must be used, with a larger set of functions better capable of describing the eigenspace of the system. In this case, qualified guessing, involving functions made specifically to describe some physical feature is no longer practical. Instead, a set of functions known to describe the wave equations well is needed. The obvious choice of a Fourier or Taylor series cannot easily be used, as the inner products do not, in general, yield finite number. This could (at least partially) be solved doing all integration numerically, however this is a relatively expensive process (compared to a purely analytical approach), and prone to errors or at least statistical uncertainty. Further more, since we cannot (usually) perform our calculations on an infinite number of functions, we have to choose a subset which is a problem for any set of orthogonal functions, as we do not know which ones will contribute the most to the accurate solution resulting in slow convergence.

Thus, we have a set of requirements for our basis set:

- It must be overcomplete. This makes the choice of what particular subset we use less important, greatly reducing the problem.
- The matrix elements involved in the calculations must be analytically calculable. This provides us with machine-precision accuracy, and allows calculations to be done fast and efficiently.
- A subset of the set of functions must be able to describe the eigenfunctions well.

The correlated Gaussian basis [7] fulfills all these requirements. The basis has not been proven complete in general, only for various quantum mechanical systems [12] [13] but has been used with considerable success in calculating various quantum mechanical systems, making it perfectly suitable.

The simplest of the correlated gaussians is the spherical gaussian described by

$$\Psi = c \cdot \exp \left( -\frac{1}{2} \sum_{j>i=1}^N \alpha_{ij} (\mathbf{r}_i - \mathbf{r}_j)^2 \right) \quad (2.3)$$

For obvious reasons a set of these functions is not capable of fully describing a system with angular momentum, such as larger atoms and molecules. It should, however, be perfectly capable of describing the ground states and first few excited states of simpler systems, such as small atoms, where the electrons are known to reside in s-orbitals.

## 2.1 THE BASIC PROBLEM

In order to project the first  $K$  eigenfunctions of  $\hat{H}$  onto our subspace  $\mathcal{V}_K$ , we must first rephrase the problem:

For a given set of linear independent functions in the subspace  $\mathcal{V}_K$  we know that

$$\hat{H}|\Psi_n \rangle = \epsilon_n |\Psi_n \rangle \quad (2.4)$$

where

$$\Psi_n = \sum_{j=1}^K c_{nj} \Psi_j \quad (2.5)$$

thus we get that

$$\sum_{j=1}^K \hat{H} |c_{nj} \Psi_j \rangle = \sum_{j=1}^K \epsilon_n c_{nj} |\Psi_j \rangle \quad (2.6)$$

which leads to

$$\sum_{j=1}^K \langle \Psi_i | \hat{H} | \Psi_j \rangle = \sum_{j=1}^K \epsilon_n c_{nj} \langle \Psi_i | \Psi_j \rangle \quad (2.7)$$

If we let the matrices  $\mathcal{H}$  and  $\mathcal{B}$  be defined by

$$\mathcal{H}_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle \quad \mathcal{B}_{ij} = \langle \Psi_i | \Psi_j \rangle \quad (2.8)$$

where  $\mathcal{H}_{ij}$  refers to the element of  $\mathcal{H}$  in the  $i$ 'th row in the  $j$ 'th column. we can write this equation as

$$\mathcal{H}\mathbf{c}_n = \epsilon_n \mathcal{B}\mathbf{c}_n \quad (2.9)$$

where  $\mathbf{c}_n$  is a vector. This is otherwise known as the generalized eigenvalue problem. Several methods solving this problem are known, with the Lancos and QZ algorithms being the most widely used. As accuracy is a very important part of this algorithm, another scheme will be used.

By using Cholesky decomposition, where  $\mathcal{B} = LL^T$ , where L is a lower triangular matrix, the problem is reduced

$$\mathcal{H}\mathbf{c}_n = \epsilon_n \mathcal{L}L^T\mathbf{c}_n \Leftrightarrow \quad (2.10)$$

$$L^{-1}\mathcal{H}\mathbf{c}_n = \epsilon_n \cdot L^T\mathbf{c}_n \Leftrightarrow \quad (2.11)$$

$$L^{-1}\mathcal{H}L^{-1}L^T\mathbf{c}_n = \epsilon_n \cdot L^T\mathbf{c}_n \quad (2.12)$$

. leaving us with the regular eigenvalue problem for a symmetrical matrix, with  $L^T\mathbf{c}_n$  as the eigenvector. The true eigenvector can be reached by a simple matrix multiplication with the inverse of L,  $L^{-1}$ .

Cholesky decomposition requires that the matrix is semi-definite, meaning that for any vector  $\mathbf{v}$  the following holds:

$$\mathbf{v}^T B \mathbf{v} \geq 0 \quad (2.13)$$

As B is the Gram-matrix of our set of linear independent functions, it is positive-definite, and thus the Cholesky-decomposition can always be performed.

At this point, any of the myriad of eigenvalue algorithms created could be used to solve our problem. The Jacobi-algorithm has proven itself as an extremely stable way of computing the eigenvalue problem [6], and is thus very well suited for use in quantum mechanics.

An alternative to using Cholesky decomposition is to simply create a set of orthogonal functions ( $\phi_1, \phi_2, \dots, \phi_K$ ) from the basis set using the Gram-Schmidt formula:

$$\phi_{K+1} = \frac{\Psi_K - \sum_{i=1}^K \phi_i \langle \phi_i | \Psi_K \rangle}{\sqrt{\langle \Psi_K | \Psi_K \rangle - \sum_{i=1}^K \langle \phi_i | \Psi_K \rangle^2}} \quad (2.14)$$

This would make B equal to the unity matrix, and thus reduce the problem to the normal eigenvalue problem. Like the Cholesky decomposition this could be updated in  $O(K^2)$  time. The cost is thus very similar. While the Gram-Schmidt method does not require a matrix multiplication, the evaluation of the H matrix elements takes equally longer time.

## 2.2 MATRIX ELEMENTS

In order to separate center of mass motion, we use the center of mass coordinates, which are related to the inter-particle coordinates by a linear transformation U:

$$\mathbf{x}_i = \sum_{j=1}^N (U_{ij} \mathbf{r}_j) \quad (i = 1, \dots, N) \quad (2.15)$$



where

$$U_J = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ \frac{m_1}{m_{12}} & \frac{m_2}{m_{12}} & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \cdots & \cdots \\ \frac{m_1}{m_{12 \cdots N-1}} & \frac{m_2}{m_{12 \cdots N-1}} & \cdots & \cdots & -1 \\ \frac{m_1}{m_{12 \cdots N}} & \frac{m_2}{m_{12 \cdots N}} & \cdots & \cdots & 1 \end{pmatrix} \quad (2.16)$$

In these coordinates, the momentum  $\mathbf{p}_i$  can be given in terms of the operator  $\pi_j = -i\hbar \frac{\partial}{\partial \mathbf{x}_j}$ :

$$\mathbf{p}_i = \sum_{j=1}^N U_{ji} \pi_j \quad (2.17)$$

The center of mass kinetic energy is then given by  $T_{cm} = \frac{\pi_N^2}{m_{12 \cdots N}}$  This gives us the kinetic energy operator as

$$\sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} - T_{cm} = \frac{1}{2} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \Lambda_{ij} \pi_i \cdot \pi_j \quad (2.18)$$

where  $\Lambda$  is a matrix given by

$$\Lambda_{ij} = \sum_{k=1}^N U_{ik} U_{jk} \frac{1}{m_k} \quad (i, j = 1, \dots, N-1) \quad (2.19)$$

In terms of  $\mathbf{x}$ , the interparticle-distance needed to calculate the potential energy is given by

$$\mathbf{r}_i - \mathbf{r}_j = \sum_{k=1}^{N-1} ((U^{-1})_{ik} - (U^{-1})_{jk})_{\mathbf{k}} \equiv \widetilde{w}^{(ij)} \mathbf{x} \quad (2.20)$$

This leads us to a new representation of our spherical gaussian basis function:

$$\Psi = c \cdot \exp\left(-\frac{1}{2} \widetilde{\mathbf{x}} A \mathbf{x}\right) \quad (2.21)$$

where  $A$  is a positive definite matrix, and the relationship between  $A$  and the  $\alpha$  is given by

$$A_{kl} = \sum_{j>i=1}^N \alpha_{ij} w_k^{(ij)} \widetilde{w}_l^{(ij)} \quad (2.22)$$

In this notation the overlap of the spherical gaussians is given by

$$\langle \Psi_{A'} | \Psi_A \rangle = \left( \frac{(2\pi)^{N-1}}{\det(A + A')} \right)^{\frac{3}{2}} \quad (2.23)$$

The kinetic energy is given by

$$\langle \Psi_{A'} | \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} - T_{cm} | \Psi_A \rangle = \frac{3}{2} \hbar^2 \text{Tr} (A(A + A')^{-1} A' \Lambda) \langle \Psi_{A'} | \Psi_A \rangle \quad (2.24)$$

The matrix element of any two-body potential is given by

$$\langle \Psi_{A'} | V(\mathbf{r}_i - \mathbf{r}_j) | \Psi_A \rangle = \langle \Psi_{A'} | \Psi_A \rangle \left( \frac{c_{ij}}{2\pi} \right)^{\frac{3}{2}} \int V(\mathbf{r}) e^{-\frac{1}{2}c_{ij}r^2} d\mathbf{r} \quad (2.25)$$

where

$$\frac{1}{c_{ij}} = \widetilde{w^{(ij)}} (A + A')^{-1} w^{(ij)} \quad (2.26)$$

This gives us that the matrix element of the coulomb potential is given by

$$\langle \Psi_{A'} | \frac{1}{4\pi\epsilon_0} \frac{q_i \cdot q_j}{|\mathbf{r}_i - \mathbf{r}_j|} | \Psi_A \rangle = \frac{q_i q_j}{4\pi\epsilon_0} \sqrt{\frac{2c_{ij}}{\pi}} \langle \Psi_{A'} | \Psi_A \rangle \quad (2.27)$$

, the matrix elements of the distance to the center of mass operator is given by

$$\langle \Psi_{A'} | \mathbf{r}_i - \mathbf{x}_N | \Psi_A \rangle = 2\sqrt{\frac{2}{c_i\pi}} \langle \Psi_{A'} | \Psi_A \rangle \quad (2.28)$$

where  $c_i$  is defined like  $c_{ij}$  with the difference that

$$\widetilde{w^{(i)\mathbf{x}}} = \sum_{k=1}^{N-1} (U^{-1})_{ik} \mathbf{x}_k \quad (2.29)$$

The same operator, squared gives us

$$\langle \Psi_{A'} | (\mathbf{r}_i - \mathbf{x}_N)^2 | \Psi_A \rangle = \frac{3}{c_i} \langle \Psi_{A'} | \Psi_A \rangle \quad (2.30)$$

An extremely useful property of the spherical gaussians is that for any transformation  $\mathcal{T}\mathbf{x} = T\mathbf{x}$

$$\mathcal{T}\Psi_A = \exp\left(-\frac{1}{2}\tilde{T}AT\right) \quad (2.31)$$

This allows us to easily ensure proper symmetry of the wave equation. The symmetry of a given wavefunction is given by the operator  $\hat{\mathcal{P}}$  so the actual basis function becomes

$$\Psi_k = c_k \cdot \hat{\mathcal{P}} \cdot e^{-\frac{1}{2}\tilde{\mathbf{x}}A_k\mathbf{x}} \quad (2.32)$$

The symmetry of fermions is most easily assured by letting

$$\mathcal{P} = \frac{1}{\sqrt{N!}} \sum_p \hat{P} \cdot (-1)^p \quad (2.33)$$

where  $\hat{P}$  is the permutation operator, such that  $\mathbf{r}_i \rightarrow \mathbf{r}_{p_i}$ , interchanging particle indices of identical particles, over every possible permutation  $p$ . In single particle coordinates, this operator is particular simple, becoming

$$\hat{P}\mathbf{r} = T_P\mathbf{r} \quad (2.34)$$

where

$$(T_P)_{ij} = \delta_{jP_i} \quad (i, j = 1, \dots, N) \quad (2.35)$$

In center of mass coordinates this changes to

$$(T_P)_{ij} = \sum_{k=1}^N U_{ik}(U^{-1})_{Pkj} \quad (i, j = 1, \dots, N) \quad (2.36)$$

which results in

$$\hat{P}\Psi_k = c_k \cdot e^{-\frac{1}{2}\tilde{\mathbf{x}}\tilde{T}_P A_k T_P \mathbf{x}} \quad (2.37)$$

As the number of identical particles rises, the number of particle interchanges needed to assure the symmetry rises exponentially, making it a practical impossibility to calculate large systems, without approximations to the behavior of the particles.

## 2.3 OPTIMIZING THE PARAMETERS

The problem of optimizing the individual parameters of our basis functions is anything but trivial. For a simple system of 5 particles in a basis of size  $K = 200$ , 2000 parameters need to be optimized. While this is by no means impossible (if impractical, as each evaluation of the expectation value of the energy is possibly a  $O(K^3)$  operation unless great care is taken), we have the chance of encountering a local minima, rather than a global, and are thus extremely likely to find an excited state rather than the basis state of the system. Various stochastic methods exist that solve this problem, however such schemes are computationally expensive, and a full and accurate optimization is not needed: The basis is overcomplete, meaning that no single function is irreplaceable. The the simplest stochastic optimization scheme imaginable can be used: trial and error.

Various schemes of quasi-random numbers have been used in quantum mechanics. None of these are very suited for the current purpose however, as all of them tend to include points that add little or nothing to the approximation of the wave function, and thus lead to an extremely large basis size. Instead, pseudo-random numbers in a physical interval are chosen, as this has been shown to give the fastest convergence<sup>[9]</sup>. The Portable Combined Random Generator of L'Ecuyer<sup>[10]</sup> is used, which has a period of roughly  $2.30584 \cdot 10^{18}$ .

The physical interval from which the basis functions are chosen, is vital for the convergence. It is noted that the expectation value relevant to distance is given by  $\langle r \rangle = \sqrt{4/(\pi\alpha)}$ , which provides a good starting point for a guess, as bounds for the expected distance between particles in the system can often be estimated.

The scheme can be further improved by simply choosing a number of trial functions, from which the best is chosen to go into the basis. As will be show later, comparing the trial functions is considerably faster than adding a function to the basis, making it possible to choose from a rather large set of trial functions.

The main disadvantage over a fully random progression scheme is that it is harder to make reasonable assumption as to the required basis size for a given convergence, and that any hope of interpolating the expectation energy must be abandoned.

## 2.4 THE NOT SO FEW BODY PROBLEM

As with any other algorithm for calculating a full quantum mechanical system, the symmetry requirements for an N-body problem result in a  $O(N!)$  time dependency, on top of

the  $O(N^2)$  time from calculating the inter-particle potentials. Thus, any algorithm hoping to simulate systems of 10 or more particles, must be easily distributable. This can easily be done with the SVM method, as practically every part of the algorithm is separable from the others. The individual matrix elements and trial functions could easily be distributed, as could the diagonalization of the matrix. This also makes the SVM well suited for calculations on GPU's, that are inherently parallel in nature. All major GPU manufacturers have SDKs for deploying software on their hardware platforms, and the new generations of GPU's have double precision numerics, allowing for fast, high precision. As the GPU's are inherently designed for linear algebra, they are an order of magnitude faster than normal CPU's, allowing normal computers to perform calculations on a relatively large number of particles.

An additional step can be taken in order to improve performance considerably when calculating the matrix elements. If in a given trial function, only a single element  $\alpha_{ij}$  is changed to  $\alpha_{ij} + \lambda$ , it is possible to compute the change in the inverse and determinant using the Sherman-Morrison formula:

The changes in the matrix A is given by  $A \rightarrow A + \lambda w^{(ij)} \tilde{w}^{(ij)}$ , where  $w^{(ij)}$  is given by formula 2.20. The inverse and determinant of the matrix B given by  $B = A + A'$  is then changed via the formulas

$$(B + \lambda w^{(ij)} \tilde{w}^{(ij)})^{-1} = B^{-1} - \frac{\lambda}{1 + \lambda w^{(ij)} \tilde{w}^{(ij)} B^{-1} w^{(ij)}} B^{-1} w^{(ij)} \tilde{w}^{(ij)} B^{-1} \quad (2.38)$$

and

$$\det(B + \lambda w^{(ij)} \tilde{w}^{(ij)}) = (1 + \lambda w^{(ij)} \tilde{w}^{(ij)} B^{-1} w^{(ij)}) \det(B) \quad (2.39)$$

This changes the complexity of calculating the determinant and B (both used extensively in calculating the matrix elements) from  $O((N - 1)^3)$  to  $O((N - 1)^2)$  for an N particle system, allowing more trial functions to be tested. This is, however, a somewhat limited way of changing the basis functions, and is mainly desirable in larger (4 or more particles) systems, where the evaluation of the matrix elements is the primary computational cost.

## 2.5 A LOWER BOUND

While the Ritz theorem gives us an upper bound for the ground state energy, having a lower bound would be equally desirable, as it would provide the ability to estimate the error of the calculated ground state energy. The variance of the energy expectation value is given by

$$\sigma^2 = \frac{\langle \Psi | H^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} - E^2 \quad (2.40)$$

where  $E$  is the expectation value of the upper bound. It is then known that there is at least one exact eigenvalue in the interval  $[E - \sigma, E + \sigma]$ , according to [8]. The problem with the lower bound is that there is no guarantee that the value inside the bounds is not an excited eigenstate, and thus not the ground state we're looking for. This is especially problematic with the higher excited states where the eigenvalues are often very close to each other. The systems in this paper are already very well theoretically known, and these values will provide the lower bound. When expanding a basis that is already reasonably

close to the ground state, it may be advantageous to minimize the variance rather than the energy, although this requires the computation of more matrix elements.

## 2.6 THE IMPLEMENTATION

Requiring a full diagonalization every time the energy of the basis, expended with a new trial function is to be calculated would make this method far too cumbersome for computation of large basis sizes on a regular computer within a reasonable time frame. Instead, it is noted that given the matrix  $L^{-1}HL^{-1T}$  of size  $K \times K$  (which shall henceforth be denoted  $H'_K$ ), where  $L$  is the Cholesky-decomposition of  $B$  given by formula 2.8, the diagonal matrix  $D_K$  containing the eigenvalues of  $H'_K$  in the diagonal, and the matrix  $Q_K$  where  $H'_K = Q_K D_K Q_K^T$ , calculating the eigenvalues of  $H'_{K+1}$ , created by expanding the basis with the function  $\Psi_K$  can be done as follows: Let  $Q_{K+1}$  be the matrix  $Q_K$  expanded by having zeros in the  $K+1$ 'th row and column, kept in the last of the diagonal, where we put a one. The matrix  $Q_{K+1}^T H'_{K+1} Q_{K+1}$  will then be almost diagonal:

$$\begin{pmatrix} \epsilon_1 & 0 & \cdots & 0 & h_1 \\ 0 & \epsilon_2 & \cdots & 0 & h_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \epsilon_K & h_K \\ h_1 & h_2 & \cdots & h_K & h_{K+1} \end{pmatrix} \quad (2.41)$$

with the eigenvalues given as the solutions  $E$  to the equation

$$E - h_{k+1} = \sum_{i=1}^K \frac{h_i^2}{E - \epsilon_i} \quad (2.42)$$

Notice that, since we already know  $Q_K^T H'_K Q_K$ , it is possible to calculate  $Q_{K+1}^T H'_{K+1} Q_{K+1}$  in  $O((K+1)^2)$  time, thus allowing us to calculate the energy for each trial function in  $O((K+1)^2)$  time, a dramatic improvement over a full diagonalization in each step.

From the definition of the Cholesky decomposition, it is obvious that when expanding the basis, it is possible to also update the cholesky decomposition  $L$  of  $B$  in time  $O((K+1)^2)$ , and its inverse  $L^{-1}$  as well. It is slightly less obvious that if we also save the previous iteration of  $L^{-1}H$ , it is also possible to update  $L^{-1}HL^{-1T}$  within this time frame, thus allowing us to test a trial function in  $O((K+1)^2)$  time, as opposed to the  $O((K+1)^3)$  time it takes to do a full diagonalization.

Brent's method[11] is used as the rootfinding algorithm of choice for solving 2.42.

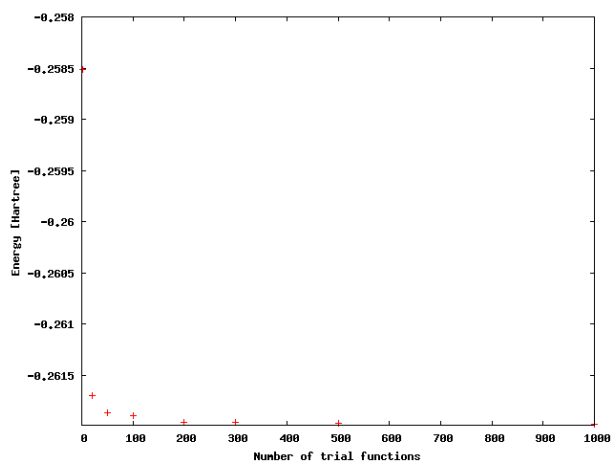
When updating the basis, we use the Jacobi Method. This is not the fastest of the eigenvalue algorithms, but holds several advantages to other, faster, methods. First of all it is easily modified to yield the matrix of eigenvectors  $Q$ . Second, it is highly stable and extremely accurate. Thirdly, fast parallel algorithms exist, allowing one to distribute almost all parts of the stochastic variational method. It should however be noted that while the Jacobi algorithm is generally a good choice, the used implementation converges when the root-mean-square of the off-diagonal elements is below  $10^{-6}$ , a bound that could be improved, allowing the Jacobi method to reach accuracy close to machine precision. While the check for orthogonality ensures that  $B$  is generally well-conditioned, the addition of a refinement procedure would allow the method to deal with less well-conditioned matrices,

and thus reduce our demand of non-orthogonality. In order to assure that  $A$  is positive definite, it is diagonalized by the Jacobi method, and the eigenvalues are tested to be positive.

# CHAPTER 3

## PRACTICALITIES

The number of trial functions used in the calculation is a fine balance. With enough trial functions, convergence will be faster, and a smaller basis is needed. Too many trial functions leads to redundant calculations, and a slower implementation. This becomes even more important when the evaluation of matrix elements become the dominant performance drain. With more particles the “sweet spot” between basis size and number of trial functions is changed, and for many-particle systems it will be advantageous to stop using trial functions at all. In the following tests, we are using the Ps- ion as an example, with a basis size of 100.



We notice that while there is an immediate gain in accuracy by using trial functions, using more than 100 seems redundant, compared to the cost of simply increasing the basis size.

In order to ensure convergence, a trial size of 200 is used in all calculations.

# CHAPTER 4

## RESULTS

In order to prove (or at least plausify) the correctness of the implementation, several examples have been made. It should be noted, that given the correlated nature of the gaussians and the large variance on the  $\langle r \rangle$  operator, it is not immediatly possible to visualize the wave functions of the various systems. What can, however, be visualized are the uncorrelated interpartile wavefunctions  $\Psi_{ij} = \sum_{k=1}^N c_k \cdot e^{-\frac{1}{2}\alpha_{ij}}$

### 4.1 PROOF OF CONCEPT: POSITRONIUM

The various electron-positron systems have been the source of much attention over the years, and many theoretical calculations have been made. This makes positronium a good benchmark for comparing the SVM with other numeric methods, as have indeed been done by K. Varga and Y Suzuki[14]. Here, it is used mainly to demonstrate the correctness of the implementation.

System	Energy	$\langle r \rangle$	$\sqrt{\langle r^2 \rangle}$
<i>Ps</i>	-0.25	1.732	1.5
Orig	-0.25		1.732
<i>Ps</i> <sup>+</sup>	-0.262005(0)	4.378973	5.106183
Orig	-2620004		4.592
<i>Ps</i> 2	-0.515769	3.181038	3.578653
Orig	-0.515989		3.608

The reference values are the SVM calculations made in [14] There is excellent agreement between the results and the reference values, although in the case of *Ps*<sup>+</sup> the  $\sqrt{\langle r^2 \rangle}$  value appears somewhat larger than the result obtained in the original.

### 4.2 HYDROGEN

Hydrogen is well known both from experiments and actual measurements. It is one of the few quantum mechanical systems for which a pure analytical solution has been found, which makes it somewhat uninteresting in our case, kept as a benchmark for our calculations. Measurements place this energy to be -0.499732 Hartree. Even using a small basis of



merely 30 basis states we get a value  $-0.4997278$ , which is in excellent agreement with the experimental value, and almost within the limits of the hyperfine splitting.

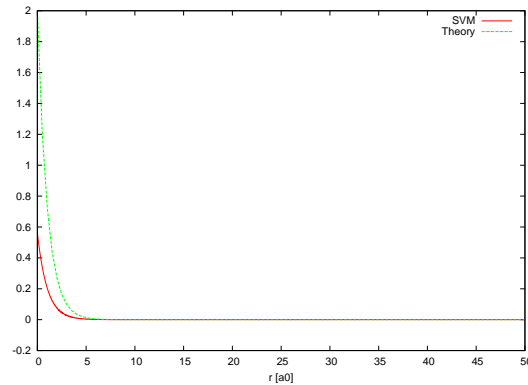


FIGURE 4.1: Calculated wave function of hydrogen vs. experimental value

### 4.3 HELIUM

Accurate theoretical models for the hydrogen atom have been available for the past 50 years, has not been considered a relevant problem for quite some time.

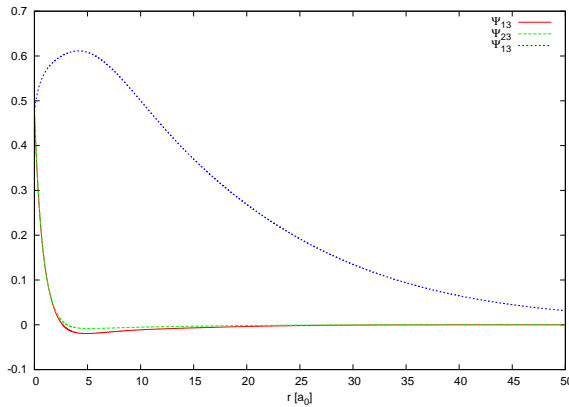


FIGURE 4.2: Non-correlated wave functions of He

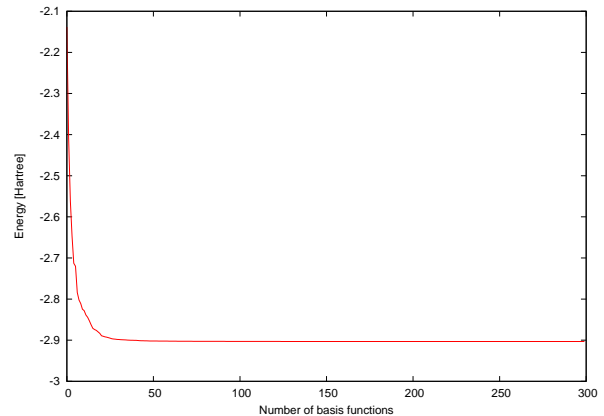


FIGURE 4.3: Energy convergence of He

Energy	$\langle r_1 \rangle$	$\langle r_2 \rangle$	$\sqrt{\langle r_1^2 \rangle}$	$\sqrt{\langle r_2^2 \rangle}$
$-2.903195(0)$	0.9295	0.9295	1.09253	1.09253

The non-relativistic groundstate of helium is  $-2.9037244$  Hartree<sup>[18]</sup>, and thus the value achieved is accurate to within 2‰. The  $\langle r \rangle$  values are smaller than those of hydrogen, as expected due to the larger charge of helium, and the fact that both electrons are in the same shell. It is curious to note that the non-correlated wavefunctions between the two electrons ( $\Psi_{13}$  and  $\Psi_{23}$ ) and the core show in 4.2 are very similar to the  $2s$  wave function of hydrogen, and the wavefunction between the electrons ( $\Psi_{12}$ ) is very similar to  $1p$  wave function.

## 4.4 LITHIUM

The unrelativistic lithium groundstate is known theoreticly within 1n hartree, giving the value  $-7.478060326(10)$ , using a hybrid of variational calculations in Hylleraas coordinates and the superposition of correlated configurations method [16]. The used implementation of the SVM cannot hope to reach such accurate results as it is simply beynd the precision achievable with the used eigenvalue and rootfinding algorithms.

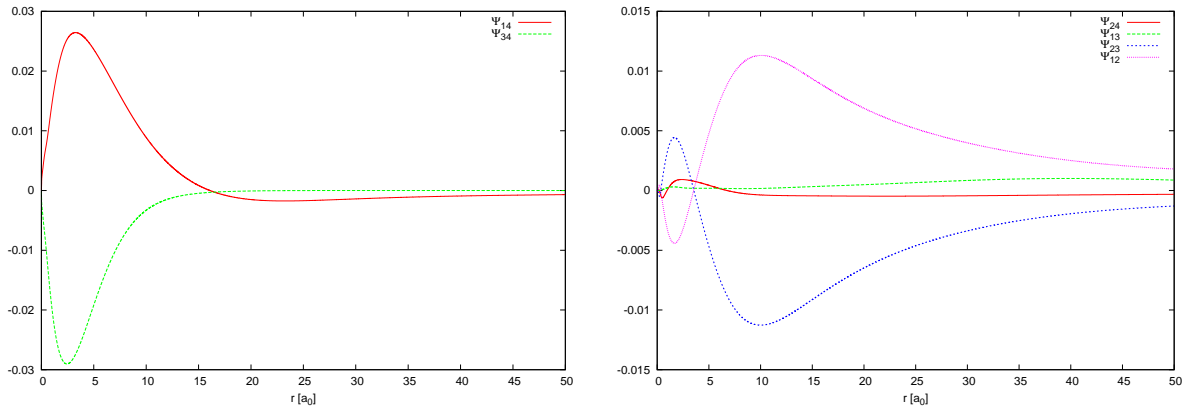


FIGURE 4.4: Non-correlated wave functions of Lithium

Energy	$\langle r_1 \rangle$	$\langle r_2 \rangle$	$\langle r_3 \rangle$	$\sqrt{\langle r_1^2 \rangle}$	$\sqrt{\langle r_2^2 \rangle}$	$\sqrt{\langle r_3^2 \rangle}$
$-7.4768004(8)$	2.2066	0.5753	2.2066	2.9902	0.6714	2.9902

The energy is in excellent agreement with the result optained in [16], and correct within  $60\mu Hartree$ . The expectation value for the distance to the center of mass does however seem to reveal a curious fact. The ground state of lithium is known to be  $1s^2 2s$ . The distances seen here seems to suggest a  $1s 2s^2$  state, though such a state is not known to exists, nor is there room in the energy to allow for an excited state. The values themselves are however perfectly valid. The normal size of the Lithium atom is around  $2.87a_0$ , and thus within the bounds of the expectation value. A calculation of  $Li^+$  does indeed show two core electrons in the same shell, with values of  $\langle r \rangle$  and  $\sqrt{\langle r^2 \rangle}$  corresponding to the values of  $\langle r_2 \rangle$  and  $\sqrt{\langle r_2^2 \rangle}$  of Li.

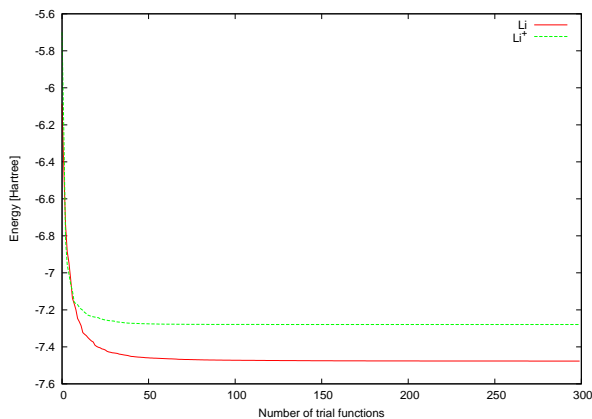


FIGURE 4.5: Convergence of Li and  $Li^+$

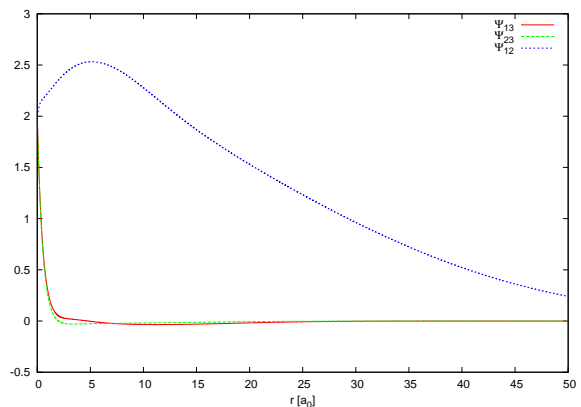


FIGURE 4.6: Wave functions of  $Li^+$

Energy	$\langle r1 \rangle$	$\langle r2 \rangle$	$\sqrt{\langle r1^2 \rangle}$	$\sqrt{\langle r2^2 \rangle}$
-7.2791703(2)	0.5728	0.5728	0.6681	0.6681

The difference between these two calculations yields a net ionisation energy of 0.19763616 a.u., which is in excellent agreement with the experimntal result of 0.19814 a.u..It is also a little on the low side, which is to be expected as fewer basis states are needed to acieve high accuracy when there are fewer particles involved. As could be expected, the uncorrelated wave functions for  $\text{Li}^+$  seen on figure 4.6, are virtually the same as those for helium, though somewhat more contracted due to the higher charge. They are, however very different from the wave functions of the neutral Lithium atom. This is presumably an effect of the symmetrization, as the two identical particles (referred to as particle 1 and 3) have extremely similar uncorrelated wave functions.

## 4.5 BERYLIUM

The beryllium groundstate has previously been calculated to an accuracy of  $(10 \pm 50)\mu$  hartree [15], giving the value -14.667353(2) This was achieved using a very extensive multiconfiguration Hartree-Fock calculation, with a Dirac-Fock-Breit relativistic correction. While no direct experimental result exists, the first two excitation levels are known with great precision, and adding these to a theoretical calculation of the helium-like  $\text{Be}^{2+}$  for which extremely accurate theoretical values exists, it is possible to gain a a semi-experimental value of  $-14.6693324$ . This is far beyond the scope of what this article hopes to achieve, and also, in terms of computation far beyond what will be needed. First of all, the relativistic corrections to the Schrödinger equation are not included though this is only around 3m hartree [15]. The lack of angular momentum means that the expected convergence is slower than what we could reasonably expect if we were using a more advanced form of gaussian, it also means that Beryllium is the largest atom for which a reasonable groundstate can be with spherical gaussians. This, combined with a limited computation time and basis size, means that a more fair comparison of the SVM capabilities is the Hartree-Fock energy of  $-14.5730230$ .

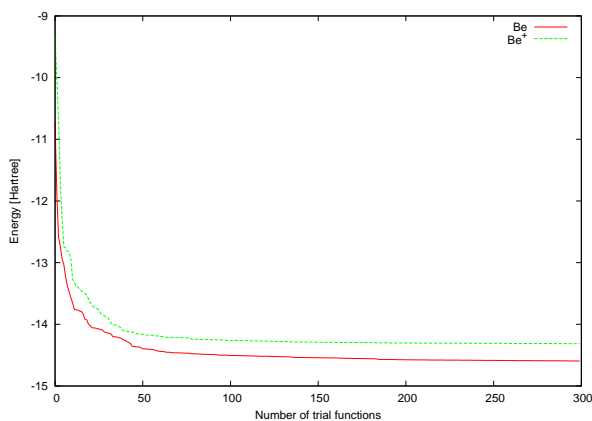


FIGURE 4.7: Convergence of Be and  $\text{Be}^+$

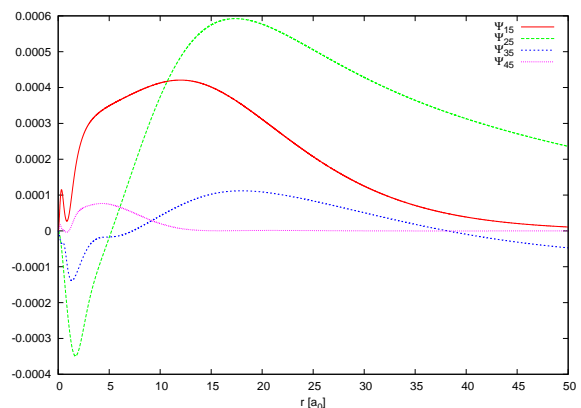


FIGURE 4.8: Non-correlated electron-core wave functions of Be

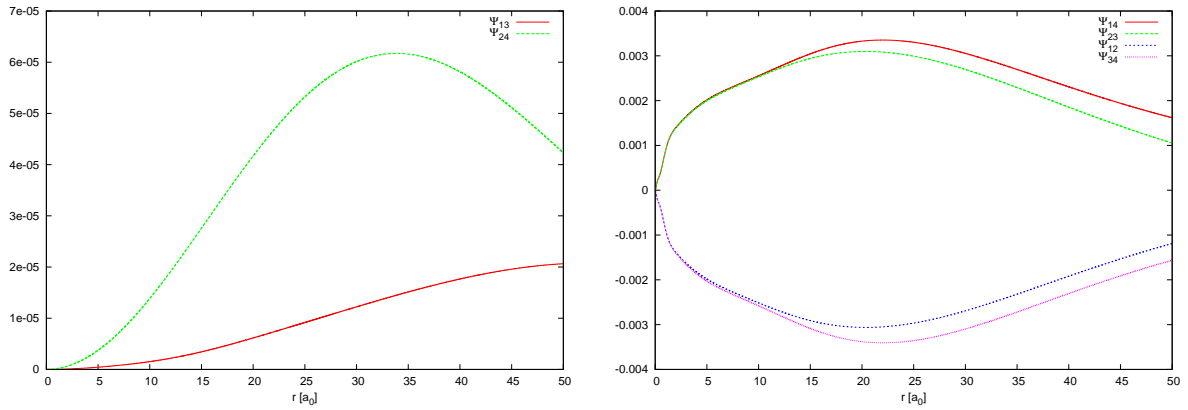


FIGURE 4.9: Non-correlated electron-electron wave functions of Be

System	Energy	$\langle r_1 \rangle$	$\langle r_2 \rangle$	$\langle r_3 \rangle$	$\langle r_4 \rangle$
Be	-14.58502(8)	1.6794	1.6199	1.6794	1.6199
Be+	-14.31360(0)	1.3604	0.4169	1.3604	
		$\sqrt{\langle r_1^2 \rangle}$	$\sqrt{\langle r_2^2 \rangle}$	$\sqrt{\langle r_3^2 \rangle}$	$\sqrt{\langle r_4^2 \rangle}$
Be		2.3410	2.2469	2.3410	2.2469
Be+		1.8016	0.4851	1.8016	

The energy found is somewhat than what was predicted by the Hartree-Fock method, and judging by the convergence, a larger basis size may in fact have given a substantially better value. Several memory leaks in the program used had not been removed at the time of writing, and thus a larger basis was not possible due to the memory and time limitations. The expectation values of the distance to the center of mass are within a very reasonable spectrum, as we see two pairs with different expectation values with could correspond to the two shells in beryllium. The actual values are reasonable, the normal size of the Beryllium atom is around  $2.1165a_0$  which corresponds very nicely with what we have found.

While the non-correlated wave functions between the electrons are, pairwise, very similar, no such symmetry can be found in the functions between the electron and the core. Given a larger basis size, it is likely that a greater similarity between the wavefunctions of the electrons residing in the same shells would have been observed.

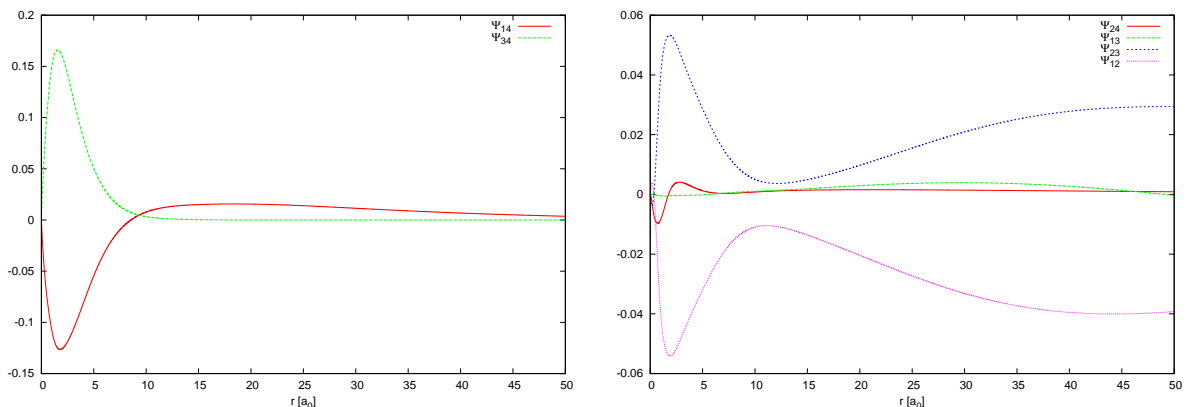


FIGURE 4.10: Non-correlated wave functions of  $\text{Be}^+$

Like Lithium, the  $\langle r \rangle$  values do not seem to follow the Shell model. For both Be and  $\text{Be}^+$  it is clearly seen that Be is smaller than Lithium which is to be expected, as the core charge of Be is higher than that of Li, and thus likely to bind the electrons closer. The experimental ionisation energy of Be is 0.34261 [5], which is somewhat higher than the 0.28470 estimate from the SVM. As with lithium, this is mainly due to the fact that a higher number of basis states were needed for an accurate estimate of the groundstate of Be.

As expected the uncorrelated wave functions of  $\text{Be}^+$  are very similar to those of Lithium.

# CHAPTER 5

## IMPROVEMENTS OF THE SVM

A full study of the convergence rate of the Stochastic variational method has not yet been made, and would provide valuable information as to whether it would be suitable for computation of more-body-systems using either large scale distributed computing, or the computing power presumed available in the future.

### SVM

The used implementation of the SVM could easily be improved. Adding a refinement cycle would improve convergence dramatically.

Consider an already established basis of size  $K$ . While the trial functions have ensured that these functions independently are reasonable estimates of the optimal added parameter, we have no guarantee that previously added basis functions are still reasonably optimal. A refinement cycle could then be added, in which a number of trial functions were created and tested against the functions already in the basis. This could be done reasonably inexpensive, as matrix elements for the trial functions would only have to be created once, and calculations of the expectation energy when testing the trial functions could be made in  $O(N^2)$ , as shown earlier. This has proven to greatly contribute to the convergence as better accuracy is made without a direct increase in basis size.

There is another problem with the used implementation of the SVM. Both Brent's method and the Jacobi eigenvalue algorithm are iterative algorithms, containing within them an accuracy parameter determining when to converge. In the previous calculations this has been set to  $10^{-6}$ , as further decreasing this parameter gave an undesirable increase in computation time. With proper attention it would be possible to implement this to an accuracy comparable to the one achieved in modern experiments, however such an implementation would require more computational power, and would probably be better suited for use in a distributed environment.

It is obvious that the addition of angular momentum would allow the used implementation to cover a much broader range of problems, covering (in theory) every known atom and its various excitations. Additionally, by changing the exponential term to  $\exp(-\frac{1}{2}\tilde{\mathbf{x}}A\mathbf{x} - \mathbf{s} \cdot \mathbf{x})$  with the vector being an additional stochastic parameter  $\mathbf{s}$ , it would be possible to describe molecules. Both these additions would increase the computation time required, but greatly increase convergence in the respective systems.

The greatest advantage of the SVM is also its greatest disadvantage. In its general form, it does not benefit from much of the physical knowledge we may have of a system.

In the study of Beryllium, for instance, it is known that there are two shells, and that the electrons in each shell will be exactly the same in the non-relativistic approximation, which could greatly reduce computation time and increase convergence. The advantage here is of course that no such assumption ever enters into the calculations, which may lead to new and indeed surprising results.

The used implementation is not well suited for excited states as no optimization of the excited wave-functions every take place. After a rough basis optimization to the ground state, a number of functions could be added to optimize the desired excited state (provided the basis set is sufficiently large to allow for calculation of the wanted excited state). This makes the SVM capable of accurately determining spectra as commonly used in atomic and molecular spectroscopy.

With the addition of relativistic energy and the various relativistic corrections, it is possible to determine atomic spectra with great precision, using various realistic potentials, and thus even study subnuclear phenomena.

# CHAPTER 6

## CONCLUSION

In this paper, the theoretical framework behind the the stochastic variational method has been described in great detail along with it's relation to the generalized eigenvalue problem, and various improvements to the basic approach have been suggested. It has been futher shown that a lightwieght implementation, using only speherical gaussians is perfectly capable of describing various few body systems with a reasonable accuracy, yielding various expectation values with considerable ease. It was also shown, that although reasonable accuracy was obtained, the simple implementation used was not able to compete with the accuracy of the best theoretical values found using ther methods. Given the amount of computational power becoming available with the advent of general purpose computing on graphics processing units, making performance available on desktop computers that were previously only available on cluster computers, the stochastic variational method should be ideal for describing larger systems of 8 or more particles, although nano-Hartree precision will probably require greater than double precision floating points. A further sutdy could include the combination of the correlated gaussians with correlated hydrogen wave functions, possibly allowing for faster convergence what is currently experienced.



# BIBLIOGRAPHY

- [1] Y. Suzuki, K. Varga, Stochastic Variational Approach to Quantum-Mechanical Few-Body Problems, Springer 1998, ISBN 3-540-65152-7.
- [2] Introduction to Quantum Mechanics, Second Edition, International Edition by David . Griffiths, Pearson Education International 2005, ISBN 0-13-191175-9.
- [3] The wikipedia entry on Cholesky decomposition [http://en.wikipedia.org/wiki/Cholesky\\_decomposition](http://en.wikipedia.org/wiki/Cholesky_decomposition)
- [4] A New Implementation of a Parallel Jacobi Diagonalizer by I.J.Bush, STFC Daresbury Laboratory <http://www.cse.scitech.ac.uk/arc/bfg.shtml>
- [5] National Institute of Standards and Technology <http://physics.nist.gov/PhysRefData/PerTable/index.html>
- [6] P. I. Davies, N J. Higham, F Tisseur, Journal on Matrix Analysis and Applications, 2001, Vol. 23 , Issue 2 , 472 - 493
- [7] S. F. Boys, Proc. R. Soc. Lond. Ser. A 258, 402 (1960). K. Singer, Ibid. 258, 412 (1960).
- [8] Y. Suzuki, K. Varga, Stochastic Variational Approach to Quantum-Mechanical Few-Body Problems, Chapter 3, Springer 1998, ISBN 3-540-65152-7.
- [9] Y. Suzuki, K. Varga, Stochastic Variational Approach to Quantum-Mechanical Few-Body Problems, Chapter 4, Springer 1998, ISBN 3-540-65152-7.
- [10] P. L'Ecuyer, Communications of the ACM, 1988, Vol 31, issue 6, 724-774
- [11] R.P. Brent (1973). Algorithms for Minimization without Derivatives, Chapter 4. Prentice-Hall, Englewood Cliffs, NJ. ISBN 0-13-022335-2.
- [12] H. F. King, J. Chem. Phys. 46, 705 (1967)
- [13] International Journal of Quantum Chemistry, Vol. 61, 769-776 (1997)
- [14] K . Varga, Suzuki, Phys. Rev. C 52, 2885 - 2905 (1995)
- [15] A Maartensson-Pendril, S. A. Alexander, L. Adamowicz, N. Oliphant, J. Olsen, P. Oster, H. M. Quiney, S. Salomonson, D. Sundhold, Phys. Rev. A 43, 3355-3364 (1991)
- [16] D. K. McKenzie, G. W. F. drake, Phys. Rev. A 44, 6973-6976 (1991)

- [17] V. I. Kukulin, V. M. Krasnopolsky, J. Phys. G 3, 795 (1977)
- [18] V. I. Korobov, Phys. Rev. A 66, 024501 (2002)
- [19] K. Varga, Y. Suzuki, Nucl. Phys A571, 447 (1994)