Application of quantum Monte Carlo methods to excitonic and electronic systems

Robert Matthew Lee
Churchill College, Cambridge

Dissertation submitted for the degree of Doctor of Philosophy
at the University of Cambridge
March 2011

Dedicated to the memory of

Charlotte and Leonard.

Preface

This dissertation describes work carried out between October 2007 and March 2011 in the Theory of Condensed Matter group at the Cavendish Laboratory, Cambridge under the supervision of Prof. R. J. Needs. The following sections of this thesis are included in work that has been published:

- **Section 1.5** and **Chapter 2**:
  R. M. Lee, G. J. Conduit, N. Nemec, P. López Ríos, and N. D. Drummond, *Strategies for improving the efficiency of quantum Monte Carlo calculations*, Phys. Rev. E **83**, 066706 (2011);

- **Chapter 3**:
  R. M. Lee, N. D. Drummond, and R. J. Needs, *Exciton-exciton interaction and biexciton formation in bilayer systems*, Phys. Rev. B **79**, 125308 (2009);

- **Chapter 4**:
  R. M. Lee and N. D. Drummond, *Ground state properties of the one-dimensional electron liquid*, Phys. Rev. B **83**, 245114 (2011).

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration with others, except where specifically indicated in the text. This dissertation has not been submitted in whole or in part for any other degree or diploma at this or any other university. This dissertation does not exceed the word limit of 60,000 words.

Robert Matthew Lee

Cambridge, March 2011

Summary

Application of quantum Monte Carlo methods to excitonic and electronic systems

Robert Matthew Lee

The work in this thesis is concerned with the application and development of quantum Monte Carlo (QMC) methods. We begin by proposing a technique to maximise the efficiency of the extrapolation of DMC results to zero time step, finding that a relative time step ratio of 1:4 is optimal.

We discuss the post-processing of QMC data and the calculation of accurate error bars by reblocking, setting out criteria for the choice of block length. We then quantify the effects of uncertainty in the correlation length on estimated error bars, finding that the frequency of outliers is significantly increased for short runs.

We then report QMC calculations of biexciton binding energies in bilayer systems. We have also calculated exciton-exciton interaction potentials, and radial distribution functions for electrons and holes in bound biexcitons. We find a larger region of biexciton stability than other recent work [C. Schindler and R. Zimmermann, Phys. Rev. B **78**, 045313 (2008)]. We also find that individual excitons retain their identity in bound biexcitons for large layer separations.

Finally, we give details of a QMC study of the one-dimensional homogeneous electron gas (1D HEG). We present calculations of the energy, pair correlation function, static structure factor (SSF), and momentum density (MD) for the 1D HEG. We observe peaks in the SSF at even-integer-multiples of the Fermi wave vector, which grow as the coupling is increased. Our MD results show an increase in the effective Fermi wave vector as the interaction strength is raised in the paramagnetic harmonic wire; this appears to be a result of the vanishing difference between the wave functions of the paramagnetic and ferromagnetic systems. We have extracted the Luttinger liquid exponent from our MDs by fitting to data around the Fermi wave vector, finding good agreement between the exponents of the ferromagnetic infinitely-thin and harmonic wires.

Acknowledgements

I am privileged to have worked under the supervision of Richard Needs during my time at Cambridge, and would like to thank him for all of the guidance and support that he has given me. I am also indebted to Neil Drummond, whose generosity with his time and wisdom made this work possible. I am also grateful to Gareth Conduit for his help with the analytical work on the time step extrapolation scheme of Sec. 1.5, and to Norbert Nemec for his help with determining the optimal choice of block length in the reblocking algorithm of Sec. 2.1.

I also wish to thank Mike Towler for, among other things, inviting me to lecture and lead expeditions at the picturesque and stimulating *Quantum Monte Carlo and the CASINO program V* Summer School at The Towler Institute in Tuscany. Others who have been in the group at Cambridge and who have both educated and entertained me are Pablo López Ríos, John Trail, Andrew Morris, Alston Misquitta, Priyanka Seth, and Gareth Griffiths.

I would also like to thank Michael Rutter for answering my questions about computing and language, and Tracey Ingham for working incredibly hard to keep the department running.

I am grateful to the other students of TCM for making my time in Cambridge so enjoyable; John Biggins, Jamie Blundell, Jonathan Edge, David O'Regan, and Alex Silver have all provided helpful discussion in addition to making these last three years lively and entertaining.

Finally, I thank my mother Stephanie, who could not have given me more.

Contents

Preface **i**

Abstract **iii**

Acknowledgements **v**

1 Quantum Monte Carlo **1**

1.1 Monte Carlo integration 1

1.2 Quantum Monte Carlo (QMC) 5

1.3 Variational Monte Carlo (VMC) 5

1.4 Diffusion Monte Carlo (DMC) 7

1.5 Maximising the efficiency of DMC time step extrapolation 14

2 Random errors in QMC **21**

2.1 Reblocking 22

2.2 VMC calculations 27

2.3 Gaussian model 29

2.4 Results 31

3 QMC with biexcitons **39**

3.1 Introduction 39

3.2 Biexciton binding energies 43

3.2.1 Model system 43

3.2.2 Results 47

3.3 Exciton-exciton interaction 53

3.4 RDFs in biexcitons 56

3.5 Conclusions 60

4 QMC with the one-dimensional electron liquid **63**

4.1 Introduction 63

4.2 Models 65

4.2.1 Hamiltonian 65

4.2.2 Infinitely-thin wire 66

4.2.3 Harmonic wire 66

4.3 Details of calculations 68

4.4 Results 72

4.4.1 Energy 72

4.4.2 Pair-correlation function 76

4.4.3 Static structure factor 79

4.4.4 Finite-size effects 85

4.4.5 Momentum density 92

4.4.6 Tomonaga-Luttinger liquid parameters 95

4.5 Conclusions 99

5 Conclusions **101**

A Appendix **105**

A.1 Fit to the exciton-exciton potential 105

A.2 Derivation of the quasi-1D interaction 107

A.3 Pair correlation function fitting parameters 110

A.4 Outliers from distributions with heavy tails 112

Chapter 1

Quantum Monte Carlo

This section is intended to serve as a brief introduction to the variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) methods. The reader is directed to the existing literature for a more comprehensive overview [1–6]. We begin with a general description of stochastic integration before describing the details of the quantum Monte Carlo (QMC) methods used in this thesis, finally going on to propose a new scheme for efficiently extrapolating DMC results to zero time step. We use Hartree atomic units ($\hbar = |e| = m_e = 4\pi\epsilon_0 = 1$) throughout except where explicitly stated. All of the QMC calculations presented here were performed using the CASINO code [3].

1.1 Monte Carlo integration

A wide range of problems relevant to physics may be formulated as integrals. In a large number of cases, particularly in systems with many degrees of freedom, the solution is inaccessible to analytical methods. When the integrand is known, however, one can numerically estimate such integrals. The choices of where to evaluate the integrand and how to use the resulting values to produce an estimate are where the various existing methods of numerical integration differ.

The simplest numerical methods are those which sample the integrand on a uniform grid. The integral is estimated by interpolating between the sampling points using some function for which the integral is known. Depending on the form of the integrand, such methods have the potential to be exact, *e.g.,* Simpson's rule is exact for third and lower degree polynomials [7]. For a general problem, however, the accuracy depends on how well the interpolating functions resemble the integrand. The difference δ between the numerical estimate and the exact integral in 1D is $\mathcal{O}(h^p)$, where h is the grid-spacing and p depends on the order of the method. It follows that for an integral over d-dimensions, reducing δ by a factor of x requires a factor of $x^{d/p}$ more grid points. This scaling makes solution of problems involving many degrees of freedom impractical using grid-based methods [4]. More sophisticated methods such as Gaussian quadrature essentially suffer from the same problem when the integrand is not of a convenient form. Furthermore, uniform sampling in phase space has the potential to miss features occurring on length scales less than $2h$.

We will see that Monte Carlo (MC) integration does not suffer from poor scaling of the error with dimensionality and that stochastic sampling can avoid the problems associated with uniform sampling such as those due to assumptions about the form of the integrand. For stochastic methods, the 'error' refers to the width of the distribution of estimates rather than the absolute deviation from the exact integral. Statistical errors are in general preferred to systematic errors because they may usually be estimated reliably [3] — we discuss in Chapter 2 how uncertainty in the estimated statistical error influences results.

We may demonstrate the more important properties of MC integration with a general example. Consider a definite integral of the form

$$I = \int_V \mathrm{d}\mathbf{R}\, g(\mathbf{R}) \;, \tag{1.1}$$

where $\mathbf{R}$ is a multi-dimensional vector, (this will later be a vector with dN elements comprising the d-dimensional coordinates of N particles), V defines the region of configuration space in which we are interested, and $g(\mathbf{R})$ is some nontrivial function. Provided that $g(\mathbf{R})$ may be evaluated at any point inside V (or, for methods such as the Gill-Miller scheme, that its value is already known at some subset of points), Eq. (1.1) can be estimated numerically [8].

Monte Carlo integration of Eq. (1.1) proceeds by decomposing $g(\mathbf{R})$ into two functions,

$$I = \int_V \mathrm{d}\mathbf{R}\, f(\mathbf{R}) P(\mathbf{R}) \;, \tag{1.2}$$

where $P(\mathbf{R})$ is a probability density function (PDF), *i.e.,*

$$\begin{aligned} P(\mathbf{R}) &= |P(\mathbf{R})| , & \mathbf{R} \in V , \\ \int_V \mathrm{d}\mathbf{R}\, P(\mathbf{R}) &= 1 , & \end{aligned} \tag{1.3}$$

and $f(\mathbf{R})$ may be any function that we are able to evaluate throughout V. In writing Eq. (1.2), we have performed an *importance sampling* transformation, so-called because $P(\mathbf{R})$ determines the density of sampling points over which $f(\mathbf{R})$ is averaged. The estimate of I is given by

$$I \approx \frac{1}{M} \sum_{i=1}^{M} f(\mathbf{R}_i) , \tag{1.4}$$

where $\{\mathbf{R}_i\}$ are M configurations distributed inside V according to $P(\mathbf{R})$, and the estimate becomes increasingly accurate as M grows. The configurations are usually generated by the Metropolis algorithm [9], which is described later in this section. In what immediately follows, we assume that the configurations $\{\mathbf{R}_i\}$ are independent — this is often not the case, an issue discussed in Chapter 2. Assuming that the central limit theorem holds for the samples $\{f(\mathbf{R}_i)\}$, a suitable measure of the error in the estimate of Eq. (1.4) is $\sigma_f/\sqrt{M}$, where σ_f is given by

$$\sigma_f^2 = \frac{1}{M-1} \sum_{i=1}^{M} \left(f(\mathbf{R}_i) - \langle f(\mathbf{R}) \rangle\right)^2 . \tag{1.5}$$

The choice of importance sampling transformation in Eq. (1.2) thus determines σ_f, the prefactor with which the error scales [4]. An intelligent choice of importance sampling transformation may greatly improve the efficiency of a calculation. For example, if it is possible to choose $P(\mathbf{R})$ such that $f(\mathbf{R})$ is constant within the limits of the integral [*i.e.,* the choice $P(\mathbf{R}) = |g(\mathbf{R})| / \int |g(\mathbf{R})|\ \mathrm{d}\mathbf{R}$ under the condition $g(\mathbf{R}) \geq 0$ for $\mathbf{R} \in V$], we could find I exactly by this method — in principle using a single evaluation of $f(\mathbf{R})$. However, in addition to the restriction $g(\mathbf{R}) = |g(\mathbf{R})|$, this requires knowledge of the normalisation of $P(\mathbf{R})$, which in this case is essentially the integral we are attempting to compute.

The advantages of MC integration over other methods are most significant in high-dimensional ($d \gtrsim 4$) spaces; we have already noted that the error scales as $M^{-1/2}$, independently of the dimensionality. Furthermore, when the integrand is sampled at points drawn from a continuous PDF, there is no regular grid-spacing through which finer features are systematically lost, so there has been no assumption about the smoothness of the integrand.

We now describe the role of the Metropolis algorithm [9] in the construction of the estimator of Eq. (1.4). The algorithm allows one to distribute points in d-dimensional space according to an arbitrary PDF without any knowledge of that function's normalisation.

Suppose we are following the above procedure to estimate I; this requires us to generate i configurations $\{\mathbf{R}_i\}$ distributed according to $P(\mathbf{R})$. The Metropolis algorithm directs us to take the following steps, starting at a random position $\mathbf{R}'$:

1. Propose a move, $\mathbf{R} \leftarrow \mathbf{R}'$, with probability density $T(\mathbf{R} \leftarrow \mathbf{R}')$.

2. Accept the move with probability
$$P_{\text{accept}}(\mathbf{R} \leftarrow \mathbf{R}') = \min\left\{1\ ,\ \frac{T(\mathbf{R}' \leftarrow \mathbf{R})P(\mathbf{R})}{T(\mathbf{R} \leftarrow \mathbf{R}')P(\mathbf{R}')}\right\}\ . \tag{1.6}$$

3. The new position is $\mathbf{R}$ or $\mathbf{R}'$ upon acceptance or rejection, respectively. This new position is appended to the set of configurations $\{\mathbf{R}_i\}$. Return to step 1 to propose the next move and repeat until sufficiently many samples have been collected.

The transition PDF, $T(\mathbf{R} \leftarrow \mathbf{R}')$, is another choice that affects the efficiency of the procedure — it is obvious that if we only ever propose moves very close to the starting position, we will explore phase space very slowly and inefficiently. Equally, proposing moves where $|\mathbf{R} - \mathbf{R}'|$ is very large can make $P_{\text{accept}}(\mathbf{R} \leftarrow \mathbf{R}')$ small, resulting in many consecutive steps in the walk being at the same position due to move rejection. Clearly some sort of compromise is required. A common choice of $T(\mathbf{R} \leftarrow \mathbf{R}')$ is a product of Gaussians (one for each element of $\mathbf{R}$) centred on $\mathbf{R}'$. The width of the transition PDF can then be varied to find the maximum efficiency.

In practical applications, it is necessary to discard all of the data gathered before the system has reached equilibrium — this is the point at which the density has started to correctly describe the underlying distribution. The length of this equilibration period is system-dependent and may be estimated as the number of steps such that the root mean square (RMS) distance diffused by a particle is of the order of the largest physical length scale in the problem. The RMS distance diffused is $\sqrt{2DdnA\tau}$, where $D = 1/(2m)$ is the diffusion constant (m is the particle mass and $m = 1$ for electrons), d is the dimensionality, n is the number of steps, A is the Metropolis acceptance ratio, and τ is the time step. Note that τ is the *variance* of $T(\mathbf{R} \leftarrow \mathbf{R}')$, the (Gaussian) transition PDF [2].

Similarly, for an accurate estimate of the error bar one must correct for serial correlation, as discussed in Chapter 2. It is simple to show that the above procedure yields the correct distribution in the asymptotic limit [1].

1.2 Quantum Monte Carlo (QMC)

The application of MC integration to the calculation of expectation values in quantum mechanics is referred to as quantum Monte Carlo. In Sections 1.3 and 1.4, we outline two widely-used QMC algorithms. In Sec. 1.5, we describe a new scheme for minimising the statistical error in the DMC energy extrapolated to zero time step.

1.3 Variational Monte Carlo (VMC)

The ground state energy E_0 of a quantum system may be written as

$$E_0 = \frac{\int d\mathbf{R}\, \psi_0^*(\mathbf{R})\hat{H}\psi_0(\mathbf{R})}{\int d\mathbf{R}\, |\psi_0(\mathbf{R})|^2} \ , \tag{1.7}$$

where ψ_0 is the ground state wave function and $\hat{H}$ is the Hamiltonian operator. The vector $\mathbf{R}$ contains the positions of all of the particles in the system. Since the exact functional form of ψ_0 is in general very difficult to find, we appeal to the variational principle to allow us to make an estimate. Using a trial wave function $\psi(\mathbf{R})$ of our choosing, a variational estimate E is given by

$$E = \frac{\int d\mathbf{R}\, \psi^*(\mathbf{R})\hat{H}\psi(\mathbf{R})}{\int d\mathbf{R}\, |\psi\,(\mathbf{R})|^2} \geq E_0 \ , \tag{1.8}$$

where equality holds if and only if $\psi = \psi_0$. In order to evaluate Eq. (1.8), we formulate it as a MC integral,

$$E = \frac{\int d\mathbf{R}\, |\psi(\mathbf{R})|^2 E_L(\mathbf{R})}{\int d\mathbf{R}\, |\psi(\mathbf{R})|^2} \ , \tag{1.9}$$

where $E_L = (\hat{H}\psi)/\psi$ is the local energy. We obtain a VMC estimate by using the Metropolis algorithm to generate a set of configurations distributed according to $|\psi(\mathbf{R})|^2$ and then averaging the values of E_L over those configurations. Decomposing the integrand in this way is convenient because the conventional interpretation of $|\psi(\mathbf{R})|^2$ is that it is a PDF. Assuming that the transition PDF is symmetric, the Metropolis acceptance probability is

$$P_{\text{accept}}(\mathbf{R} \leftarrow \mathbf{R}') = \frac{|\psi(\mathbf{R})|^2}{|\psi(\mathbf{R}')|^2} , \tag{1.10}$$

depending only on the magnitude of the trial wave function at the initial and proposed positions. The density of configurations is thus large where $|\psi|^2$ is large.

The choice of ψ is the only real difficulty; constructing an accurate trial wave function for most real problems is far from trivial. However, if a trial wave function satisfies the relevant boundary conditions and includes some free parameters $\{\alpha_i\}$, one may systematically improve ψ by minimising one of several cost functions with respect to the $\{\alpha_i\}$. For example, one can minimise the variance of the local energy

$$\sigma^2 = \frac{\int \mathrm{d}\mathbf{R} \, |\psi(\alpha_i)|^2 (E_L(\alpha_i) - \langle E_L(\alpha_i) \rangle)^2}{\int \mathrm{d}\mathbf{R} \, |\psi(\alpha_i)|^2} , \tag{1.11}$$

where the functional dependence of ψ and E_L on $\mathbf{R}$ has been omitted here for clarity. In practice, wave function properties such as the energy variance of Eq. (1.11) are often calculated using correlated sampling, where the configurations generated using some set of parameters $\{\alpha_i\}$ are used to evaluate a quantity for different parameters $\{\alpha_i'\}$ by applying appropriate weights to the configurations.

There are several other choices of objective function available, such as the total energy or the mean absolute deviation (MAD) of the local energy from the median. Variance minimisation has historically been favoured since it is relatively straightforward to produce numerically-robust implementations and the variance has a minimum at every eigenstate [10–13]. Energy minimisation has grown in popularity in recent years due to the development and implementation of stable algorithms, and the general belief that minimum-energy wave functions have more desirable properties than those with minimum variance [4,14]. In particular, the statistical efficiency of the more intensive DMC method depends on the deviation of the VMC energy from the exact ground state energy [15].

In practice, the best choice of optimisation procedure is system-dependent, and the optimisation stage appears to be the most problematic part of the majority of QMC calculations. We discuss briefly in Chapter 3 a potential problem with variance minimisation in finite electron-hole systems; if the wave function describes some regions of configuration space better than others, one might expect variance minimisation to favour the high-quality regions even if they should only contribute a small amount to the ground state. However, if ψ is capable of describing the exact ground state, minimisation of the energy, variance or MAD should produce ψ_0.

For fermionic systems, a popular choice for ψ is the Slater-Jastrow form [3],

$$\psi(\mathbf{R}) = \exp[J(\mathbf{R})]\mathrm{det}_{\uparrow}[\phi_n(\mathbf{r}_i^{\uparrow})]\mathrm{det}_{\downarrow}[\phi_n(\mathbf{r}_j^{\downarrow})] \ , \tag{1.12}$$

where the Jastrow function $J(\mathbf{R})$ is a function of interparticle distances [16] and $\mathrm{det}_{\uparrow}[\phi_n(\mathbf{r}_i^{\uparrow})]$ is a determinant of single particle orbitals $\{\phi_n\}$ for the electrons with spin $\uparrow$. The orbitals ϕ_n can be taken, for example, from density functional theory or Hartree-Fock calculations, and can contain free parameters for optimisation. The Jastrow factor is symmetric with respect to particle exchange. The sum over spin coordinates in the expectation value of Eq. (1.8) has already been performed, so that Eq. (1.12) is spin-independent. The use of a product of up- and down-spin determinants instead of a spin-dependent wave function does not alter the expectation values of spin-independent operators [1].

The elegance and strength of the method stem from the freedom to choose whatever trial wave function one wishes. One can of course use a sum of determinants rather than a single one, and recent years have seen accuracy improve through the introduction of backflow transformations, pairing orbitals, and new forms of Jastrow factor [6, 16]. However, it is difficult to achieve equivalent accuracy for different systems, making VMC energy differences poor. As a result, VMC is usually seen as the starting point for more expensive and accurate diffusion Monte Carlo calculations.

1.4 Diffusion Monte Carlo (DMC)

The DMC distribution is represented by the density of points in configuration space, so is more flexible than the VMC wave function.

The DMC method is an approach for solving the imaginary time Schrödinger equation

$$(\hat{H} - E_T)\Phi(\mathbf{R}, \tau) = -\frac{\partial\Phi(\mathbf{R}, \tau)}{\partial\tau} \ , \tag{1.13}$$

where $\tau = it$ is the imaginary time and E_T is a fixed reference energy, which we may choose freely. The wave function Φ depends on τ in addition to the particle positions.

The Hamiltonian is

$$\hat{H} = \hat{T} + \hat{U} = -\frac{\nabla^2_{\mathbf{R}}}{2} + U(\mathbf{R}) \ , \tag{1.14}$$

where the potential U depends in principle on the positions of all of the particles in the system (and is often Coulombic), and $-2\hat{T} = \nabla^2_{\mathbf{R}}$ is the Laplacian with respect to all of the particle coordinates.

The method projects out the ground state component of the initial wave function by propagating a distribution of walkers in imaginary time. This may be understood by expressing the solution to Eq. (1.13) as a sum over the eigenstates $\{\phi_i\}$ of $\hat{H}$ (which is always possible since $\hat{H}$ is Hermitian),

$$\Phi(\mathbf{R}, \tau) = \sum_{i=0}^{\infty} c_i \phi_i(\mathbf{R}) e^{(E_T - E_i)\tau} \ , \tag{1.15}$$

where $\{c_i\}$ is the set of expansion coefficients. Each component of Φ decays exponentially with imaginary time, unlike the corresponding expansion in real time for which the phase of each component rotates. This is convenient if we are interested in the ground state properties of systems; since E_0 is the lowest eigenvalue of $\hat{H}$ the ϕ_0 component dominates exponentially as $\tau \to \infty$.

With an appropriate method for evolving configurations in imaginary time, one could therefore find the ground state wave function (and energy, *etc.*) from any starting distribution that has nonzero overlap with the ground state. The exact Green's function of Eq. (1.13) is

$$G_{\text{exact}}(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau) = \left\langle \mathbf{R} \middle| e^{-\delta\tau(\hat{H} - E_T)} \middle| \mathbf{R}' \right\rangle \ , \tag{1.16}$$

where $\delta\tau$ is an interval in imaginary time. The explicit form of Eq. (1.16) is in general unknown. However, if either the kinetic term $\hat{T}$ or the potential term $\hat{U}$ were omitted from Eq. (1.14), the Green's function would be known exactly. We may use this fact to construct an approximation to Eq. (1.16) that is exact in the limit $\delta\tau \to 0$. We do this by rewriting Eq. (1.16) as

$$\begin{aligned} G_{\text{exact}}(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau) &= \left\langle \mathbf{R} \middle| e^{-\delta\tau(\hat{T} + \hat{U} - E_T)} \middle| \mathbf{R}' \right\rangle \\ &= \left\langle \mathbf{R} \middle| e^{-\delta\tau(\hat{U} - E_T)/2} e^{-\delta\tau\hat{T}} e^{-\delta\tau(\hat{U} - E_T)/2} + \mathcal{O}(\delta\tau^3) \middle| \mathbf{R}' \right\rangle \\ &\approx e^{-\delta\tau(U(\mathbf{R}) - E_T)/2} \left\langle \mathbf{R} \middle| e^{-\delta\tau\hat{T}} \middle| \mathbf{R}' \right\rangle e^{-\delta\tau(U(\mathbf{R}') - E_T)/2} \ , \end{aligned} \tag{1.17}$$

which does have a known explicit form. We will refer to the last line of Eq. (1.17) as $G_{\text{approx}}(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau)$, which is given in full by

$$G_{\text{approx}}(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau) = G_{\text{U}}(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau) G_{\text{T}}(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau) \ , \tag{1.18}$$

where we have defined

$$G_{\text{U}}(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau) = \exp\left(-\frac{\delta\tau}{2}\left[U(\mathbf{R}) + U(\mathbf{R}') - 2E_T\right]\right) \tag{1.19}$$

and

$$G_{\text{T}}(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau) = \frac{1}{(2\pi\delta\tau)^{3N/2}} \exp\left(-\frac{|\mathbf{R} - \mathbf{R}'|^2}{2\delta\tau}\right) \ , \tag{1.20}$$

where N is the number of particles. We have assumed in writing Eq. (1.20) that the system is 3-dimensional. One could then perform a DMC calculation by considering $G_{\text{approx}}(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau)$ as the probability of a configuration going from $\mathbf{R}'$ to $\mathbf{R}$ in time $\delta\tau$, and this would be accurate for very small $\delta\tau$. In practice, the distribution of walkers could be generated initially from VMC. Propagation would then be carried out by repeatedly displacing the walkers by a Gaussian random number with variance $\delta\tau$, as described by Eq. (1.20), and assigning each walker a statistical weight, given by Eq. (1.19).

The simple algorithm described above suffers from two problems. The first is that the potential term $U(\mathbf{R})$ can vary dramatically throughout configuration space, and indeed it contains Coulomb divergences for many systems. This can result in one walker dominating over the rest, leading to stability and sampling problems [4, 7, 17].

The second problem is that the lowest energy state of the Hamiltonian (1.14) is generally bosonic, and indeed we have not mentioned fermionic antisymmetry anywhere in the explanation above. In fact, in writing down Eq. (1.20), we assumed that Φ is a PDF satisfying the conditions of Eq. (1.3). One clearly cannot satisfy $\Phi = |\Phi|$ for all $\mathbf{R}$ for fermionic systems in general.

Implementations of QMC that do not take any measures to prevent the distribution from decaying to the bosonic ground state will therefore suffer from what is known as the fermion sign problem when attempting to simulate fermionic systems. In such cases, the fermionic component of the initial distribution will vanish exponentially relative to the bosonic component, by exactly the same argument that we used to explain why the excited states decay with imaginary time. It is, in practice, impossible to avoid including a nonzero bosonic component in a representation of a wave function with a finite distribution of configurations.

Despite the severity of the problem, methods are in use today which attempt (and formally succeed) to exactly solve variants of Eq. (1.13). The release-node method involves assigning walkers negative statistical weights when they cross an odd number of trial wave function nodes, so that expectation values may be thought of in terms of differences between two populations of opposite sign [18]. The distribution as a whole (considering only the absolute values of the weights) tends to the bosonic ground state, so it is necessary to achieve equilibrium and to accumulate expectation values before the signal-to-noise ratio becomes too poor for the fermionic distribution to be extracted. Similarly, there exist approaches where walkers are again given signs, but then pairs of walkers with opposite signs are annihilated when some measure (*e.g.,* the Euclidian norm of the distance between walkers in configuration space) reaches a certain threshold. A major source of difficulty for algorithms involving cancellation of walkers is setting a suitable criterion for annihilation; new methods working in a discrete antisymmetrised space have shown great promise, although the scaling of computational cost with system size is unclear [19–21]. However, exact methods for solving Eq. (1.13) for fermions in general scale exponentially with system size [1, 18, 22, 23].

We avoid the poor behaviour associated with the $U(\mathbf{R})$ term and (exponentially scaling solutions to) the sign-problem by making the only uncontrolled approximation in the method — this is the fixed node (FN) approximation, by which the DMC distribution is forced to share the nodes of ψ, the trial wave function [24]. Enforcing the FN approximation is essentially the same as placing an infinite repulsive potential barrier at the nodes of ψ; the DMC wave function Φ is forced to go to zero wherever $\psi = 0$ and the method satisfies a variational principle [17]. If the nodes of the trial wave function are exact then the FN DMC energy will be exact.

One can efficiently implement the FN approximation by working with the mixed distribution $f = \Phi\psi$, where ψ is the trial wave function, usually optimised at the VMC level, and Φ is the DMC wave function, which is the lowest-energy wave function that shares the nodes of ψ. One can obtain the appropriate importance-sampled equation by multiplying Eq. (1.13) by ψ,

$$-\psi\frac{\partial\Phi}{d\tau} = -\frac{1}{2}\psi\nabla^2\Phi + \frac{1}{2}\Phi\nabla^2\psi + \left(\frac{\hat{H}\psi}{\psi} - E_T\right)\psi\Phi \ , \tag{1.21}$$

and then rearranging terms to get

$$-\frac{1}{2}\nabla^2 f(\mathbf{R},\tau) + \nabla\cdot(\mathbf{V}(\mathbf{R})f(\mathbf{R},\tau)) + (E_L(\mathbf{R}) - E_T)f(\mathbf{R},\tau) = -\frac{\partial f(\mathbf{R},t)}{\partial\tau} \ , \tag{1.22}$$

where $\mathbf{V}(\mathbf{R}) = \psi^{-1}(\mathbf{R})\nabla\psi(\mathbf{R})$ is the drift velocity and $E_L(\mathbf{R}) = \psi^{-1}(\mathbf{R})\hat{H}\psi(\mathbf{R})$ is the local energy. Equation (1.22) is called the importance-sampled imaginary time Schrödinger equation [5].

The third term of Eq. (1.22) now depends on the local energy E_L rather than U alone. This is convenient because E_L is roughly constant in configuration space when ψ is accurate, removing the numerical awkwardness of a wildly fluctuating source-sink term.

The FN propagator is similar to the product of Eqs. (1.20) and (1.19), and illustrates why the importance sampling transformation of Eq. (1.22) is efficient. As before, consider approximating the exact Green's function by one that has an explicit form using the procedure of Eq. (1.17). The two factors that we obtain from performing the earlier analysis, this time with Eq. (1.22), are [2]

$$G_D(\mathbf{R}\leftarrow\mathbf{R}^{'},\delta\tau) = \frac{1}{(2\pi\delta\tau)^{3N/2}}\exp\left(-\frac{|\mathbf{R}-\mathbf{R}^{'}-\delta\tau\mathbf{V}(\mathbf{R}^{'})|^2}{2\delta\tau}\right) \ , \tag{1.23}$$

which accounts for the first two terms of Eq. (1.22), and is called the drift-diffusion Green's function, and

$$G_B(\mathbf{R}\leftarrow\mathbf{R}^{'},\delta\tau) = \exp\left(-\frac{\delta\tau}{2}\left[E_L(\mathbf{R}) + E_L(\mathbf{R}^{'}) - 2E_T\right]\right) \ , \tag{1.24}$$

which is called the branching factor. The error in the propagator introduced through the replacement of the exact Green's function with $G_D G_B$ is $\mathcal{O}(\delta\tau^2)$. This carries forward to give an error in the mixed distribution (and thus mixed estimators) of $\mathcal{O}(\delta\tau)$.

The initial population of walkers is generated at the VMC level, *i.e.,* with the distribution $|\psi^2|$. A FN DMC calculation is then performed by carrying out the following steps:

- Denote a walker's position at the start of an iteration by $\mathbf{R}'$.

- Displace each walker by $\delta\tau\mathbf{V}(\mathbf{R}')$.

- Displace each walker by a Gaussian random number with variance $\delta\tau$ so that the new position is $\mathbf{R}$.

- Duplicate or annihilate walkers such that the number of copies of a given walker that continue on to the next iteration is

$$N = \text{int}\left[G_B(\mathbf{R} \leftarrow \mathbf{R}', \delta\tau) + \eta\right] \ , \tag{1.25}$$

  where η is a uniform random number drawn from $[0, 1]$.

After a certain number of iterations, the excited state components of f will have died away. One can then average expectation values over walkers and time steps to obtain DMC estimates. In particular, the DMC energy is evaluated by the '*mixed estimator*', which is given by

$$\frac{\langle\Phi|\hat{H}|\psi\rangle}{\langle\Phi|\psi\rangle} = \lim_{\tau\to\infty} \frac{\int \mathrm{d}\mathbf{R} f(\mathbf{R},\tau) E_L(\mathbf{R})}{\int \mathrm{d}\mathbf{R} f(\mathbf{R},\tau)} \approx \frac{1}{M}\sum_{i=1}^{M} E_L(\mathbf{R}_i) \ . \tag{1.26}$$

The length of the equilibration period can be estimated using the formula given at the end of Sec. 1.1. In practice, there are a number of additional refinements to the algorithm that improve efficiency and reduce bias, such as an accept/reject step [25] and modifications to the Green's function [26, 27]. There also exist several other schemes for treating the statistical weights [2, 27]. Henceforth, 'FN DMC' will be referred to simply as 'DMC', since all of the DMC calculations performed for the work presented in this thesis used the FN approximation as described above.

It is clear that the drift velocity acts to carry walkers away from nodes. Sampling in this way effectively enforces the FN approximation because walkers almost never approach the nodal surface. Rejecting the few node-crossing moves that do occur has been found to be the least biased way of dealing with them [2]. One might erroneously assume that by preventing walkers from crossing nodes, the algorithm described above will fail to cause the walkers to properly explore phase space. Each walker will explore only a single nodal pocket (region of configuration space where the wave function is of a constant sign) in a simulation, and there are typically many more pockets than configurations. Fortunately, the tiling theorem dictates that one can generate any point in configuration space by permuting the labels of indistinguishable particles in a single nodal pocket. As a consequence, the FN approximation should not introduce ergodicity problems [28].

The DMC energy depends only on the nodes of ψ, but expectation values of operators that do not commute with the Hamiltonian depend on the trial wave function everywhere. For an operator $\hat{A}$ that does not commute with the Hamiltonian, we can remove the first order contribution from the error in the trial wave function by combining VMC and DMC results [29]. This is the '*extrapolated estimator*', which we use later in Chapters 3 and 4. Let us write the trial wave function ψ as its ground state component Φ plus a small amount ε of error Δ,

$$|\psi\rangle = |\Phi + \varepsilon\Delta\rangle \ . \tag{1.27}$$

Assuming that $\hat{H}$ has time-reversal symmetry, so that Φ, ψ and Δ may be chosen real, the variational estimator may be rewritten as

$$\begin{aligned} \langle A\rangle_{\rm var} &= \frac{\langle\psi|\hat{A}|\psi\rangle}{\langle\psi|\psi\rangle} \\ &= \frac{\langle\Phi+\varepsilon\Delta|\hat{A}|\Phi+\varepsilon\Delta\rangle}{\langle\Phi+\varepsilon\Delta|\Phi+\varepsilon\Delta\rangle} \\ &= \frac{\langle\Phi|\hat{A}|\Phi\rangle}{\langle\Phi|\Phi\rangle} + 2\varepsilon\left[\frac{\langle\Phi|\hat{A}|\Delta\rangle}{\langle\Phi|\Phi\rangle} - \frac{\langle\Phi|\hat{A}|\Phi\rangle\langle\Delta|\Phi\rangle}{\langle\Phi|\Phi\rangle^2}\right] + \mathcal{O}(\varepsilon^2) \ , \end{aligned} \tag{1.28}$$

and the mixed estimator is

$$\begin{aligned} \langle A\rangle_{\rm mix} &= \frac{\langle\Phi|\hat{A}|\psi\rangle}{\langle\Phi|\psi\rangle} \\ &= \frac{\langle\Phi|\hat{A}|\Phi+\varepsilon\Delta\rangle}{\langle\Phi|\Phi+\varepsilon\Delta\rangle} \\ &= \frac{\langle\Phi|\hat{A}|\Phi\rangle}{\langle\Phi|\Phi\rangle} + \varepsilon\left[\frac{\langle\Phi|\hat{A}|\Delta\rangle}{\langle\Phi|\Phi\rangle} - \frac{\langle\Phi|\hat{A}|\Phi\rangle\langle\Delta|\Phi\rangle}{\langle\Phi|\Phi\rangle^2}\right] + \mathcal{O}(\varepsilon^2) \ . \end{aligned} \tag{1.29}$$

It is clear that the $\mathcal{O}(\varepsilon)$ error in $\langle A\rangle_{\rm mix}$ is half that of $\langle A\rangle_{\rm var}$, so an appropriate estimator with the linear error removed is

$$\langle\hat{A}\rangle_{\rm ext} = 2\langle\hat{A}\rangle_{\rm mix} - \langle\hat{A}\rangle_{\rm var} = \frac{\langle\Phi|\hat{A}|\Phi\rangle}{\langle\Phi|\Phi\rangle} + \mathcal{O}(\varepsilon^2) \ . \tag{1.30}$$

We have made a short-time ($\delta\tau \to 0$) approximation to justify two parts of the algorithm described above. First, we wrote the Green's function as the product $G_D G_B$, giving an error of $\mathcal{O}(\delta\tau^3)$. Secondly, we assumed that $\mathbf{V}(\mathbf{R})$ is constant between $\mathbf{R}$ and $\mathbf{R}'$ for each step, since at each iteration we displace configurations by $\delta\tau\mathbf{V}(\mathbf{R}')$ — this results in an error in the Green's function of $\mathcal{O}(\delta\tau^2)$. In practice one must of course use a finite but small time step when running DMC calculations. A side effect of this is that consecutive configurations in imaginary time are serially correlated — in fact this also occurs in VMC where no short-time approximation is made, since each configuration is generated from the last. It follows that expectation value data also exhibit serial correlation. This has no effect on mean values but makes calculation of an accurate statistical error bar nontrivial. We discuss in Chapter 2 methods for evaluating accurate error bars and the implications of noisy estimates.

To remove time step bias from DMC expectation values, one typically extrapolates data to zero time step. We describe in Sec. 1.5 a scheme for distributing computational effort that minimises the statistical error in the extrapolate.

1.5 Maximising the efficiency of DMC time step extrapolation

As described above, DMC is only accurate in the limit of a small time step $\delta\tau$. However, the computational effort required to achieve a given error bar goes as $1/\delta\tau$, ruling out the use of infinitesimal time steps in practice. Hence, where high accuracy is required, two or more finite time steps $\{\delta\tau_i\}$ are generally used and the ground-state energy is obtained by extrapolating to $\delta\tau = 0$ [1,3]. Here we explain how the statistical error in a zero-time-step extrapolate may be minimised by a judicious choice of time steps $\{\delta\tau_i\}$, and the sensible deployment of a limited total computing time between those time steps.

For sufficiently small $\delta\tau$, the DMC energy scales linearly with the time step as $E(\delta\tau) = E_0 + \kappa\delta\tau$, where κ is a constant and E_0 is the energy at zero time step. Suppose we calculate $E(\delta\tau)$ at R different time steps $\{\delta\tau_i\}$ in the linear-bias regime, where each $E(\delta\tau_i)$ has an associated statistical uncertainty Δ_i. The error bars fall off with the time step $\delta\tau_i$ and the CPU time devoted to the calculation T_i as $\Delta_i = C/\sqrt{\delta\tau_i T_i}$, where C is a constant. To determine the ground-state energy at zero time step E_0, we minimise the χ^2 error of the linear fit,

$$\begin{aligned}\chi^2 &= \sum_{i=1}^{R} \frac{[E(\delta\tau_i) - E_0 - \kappa\delta\tau_i]^2}{\Delta_i^2} \\ &= \frac{1}{C^2}\sum_{i=1}^{R} T_i\delta\tau_i[E(\delta\tau_i) - E_0 - \kappa\delta\tau_i]^2 \end{aligned} \tag{1.31}$$

with respect to κ and E_0. Setting $\partial\chi^2/\partial\kappa = \partial\chi^2/\partial E_0 = 0$, we obtain

$$E_0 = \frac{2\sum_{i=1}^{R}\sum_{j=1}^{R} E(\delta\tau_i)T_iT_j\delta\tau_i\delta\tau_j^2(\delta\tau_j - \delta\tau_i)}{\sum_{i=1}^{R}\sum_{j=1}^{R} T_iT_j\delta\tau_i\delta\tau_j(\delta\tau_j - \delta\tau_i)^2} \,. \tag{1.32}$$

Assuming the data are Gaussian-distributed, the square of the standard error in the extrapolate E_0 is

$$\begin{aligned}\Delta_0^2 &\approx \sum_{k=1}^{R} \Delta_k^2 \left[\frac{\partial E_0}{\partial E(\delta\tau_k)}\right]^2 \\ &= 4C^2\sum_{k=1}^{R} T_k\delta\tau_k \left[\frac{\sum_{j=1}^{R} T_j\delta\tau_j^2(\delta\tau_j - \delta\tau_k)}{\sum_{i=1}^{R}\sum_{j=1}^{R} T_iT_j\delta\tau_i\delta\tau_j(\delta\tau_j - \delta\tau_i)^2}\right]^2 \,. \end{aligned} \tag{1.33}$$

As expected the standard error falls off as the time steps $\{\delta\tau_i\}$ are increased and as more time $\{T_i\}$ is dedicated to the calculations. However, since we are considering a linear fit, $\delta\tau$ should not be increased beyond $\delta\tau_{\text{max}}$, the limit of the region in which the bias is linear. The effort allocated to the calculations also cannot be increased indefinitely because one is constrained by the total time $T = \sum_{i=1}^{R} T_i$ for all of the simulations. We now minimise Δ_0^2 subject to the constraint that T is fixed.

Let us first suppose that we are to perform just $R = 2$ simulations. We start by fixing the time steps $\delta\tau_1$ and $\delta\tau_2$, and minimising Δ_0^2 with respect to the run lengths in the presence of a Lagrange multiplier to constrain the total run time T. This yields the optimal simulation durations

$$\frac{T_1}{T} = \frac{\delta\tau_2^{3/2}}{\delta\tau_1^{3/2} + \delta\tau_2^{3/2}} \tag{1.34}$$

and

$$\frac{T_2}{T} = \frac{\delta\tau_1^{3/2}}{\delta\tau_1^{3/2} + \delta\tau_2^{3/2}} \,, \tag{1.35}$$

which attempts to reduce the error bar on the calculation with the smaller time step beyond the distribution of effort $T_1/T = \delta\tau_2/(\delta\tau_1 + \delta\tau_2)$ that would aim for error bars of

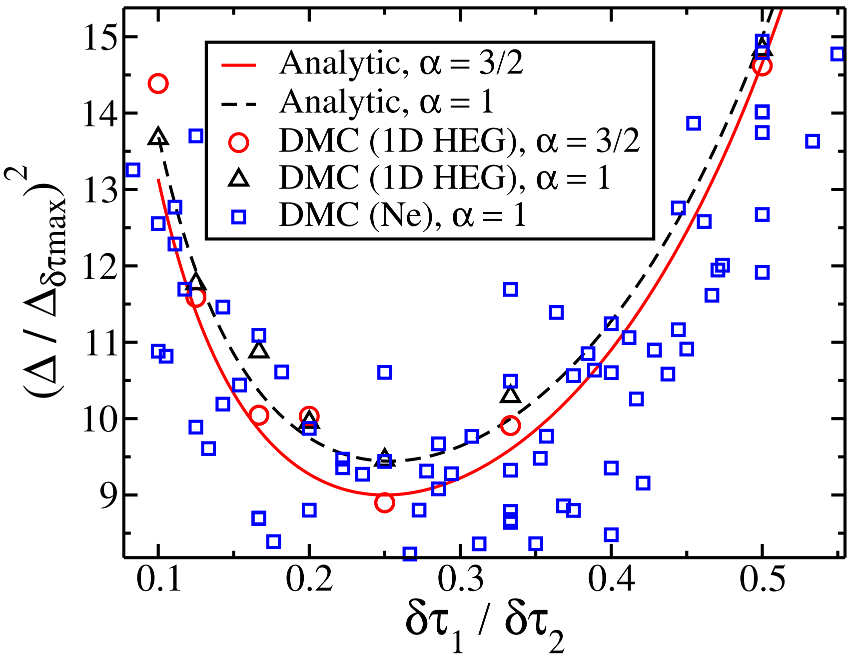

Figure 1.1: The uncertainty in the extrapolated DMC energy Δ against relative step size, $\delta\tau_1/\delta\tau_2$. The distribution of effort for each data set was $[T_1/T_2 = (\delta\tau_2/\delta\tau_1)^{\alpha}]$, where $\alpha = 3/2$ is optimal and $\alpha = 1$ is a common choice. The error bars are normalised by $\Delta_{\delta\tau\mathrm{max}}$, the error bar of a DMC run at the upper time step $\delta\tau_2$ if all of the computational resources $(T_1 + T_2)$ were dedicated to it.

equal size. The extrapolate is intuitively more sensitive to the error bar on the run with the smaller time step than the larger. Without loss of generality, we now assume that $\delta\tau_2 > \delta\tau_1$, with $\delta\tau_2 = \delta\tau_{\mathrm{max}}$ pinned near the boundary of the linear regime, and we search for the optimal time step $\delta\tau_1$. Using the optimal durations T_1 and T_2, minimisation of Δ_0^2 reveals that the optimal choice of time step is $\delta\tau_1 = \delta\tau_2/4$. The corresponding optimal physical run times are therefore $T_1 = 8T/9$ and $T_2 = T/9$. The full dependence of the final error upon the relative time step $\delta\tau_1/\delta\tau_2$ is shown in the Figs. 1.1 and 1.2.

Now suppose that more than two time steps are used to perform the extrapolation. We find that Δ_0^2 is minimised when the computational effort is preferentially dedicated to the two points that are nearest to having a relative time step of 4 and have the largest maximum value of $\delta\tau$. Computational effort should therefore be focused solely on that optimal pair as long as the linear regime is well-defined. Clearly, in an idealised model where we wish to fit a straight line to the data, there no advantage to using more than

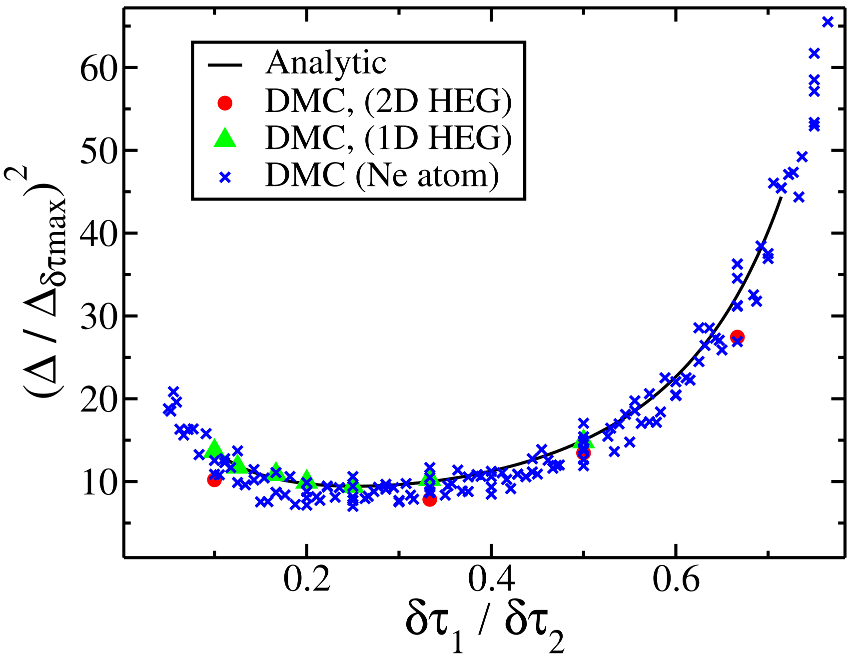

Figure 1.2: The uncertainty in the extrapolated DMC energy against relative step size, $\delta\tau_1/\delta\tau_2$ over a larger range. The effort was distributed according to $(T_1/T_2 = \delta\tau_2/\delta\tau_1)$.

$R = 2$ data points.

This scheme is the optimal extrapolation procedure when the extent of the linear regime is known. The strategy is thus highly applicable to studies of many similar systems where the linear regime can be assumed to be the same for multiple runs. For systems where the behaviour of the time step bias has not been established, one has no alternative but to perform multiple runs over a wide domain of time steps and determine where the spectrum first increases superlinearly. In such cases, one can use the RMS distance (as discussed in Sec. 1.1) diffused by an electron over a single step as an initial order-of-magnitude estimate for where the linear regime begins. For all-electron atomic systems, for example, one would expect the linear regime to occur for time steps less than of the order $\delta\tau = 1/(3Z^2)$, where Z is the largest atomic number occurring in the system. This choice of time step ensures that the RMS distance diffused is equal to one Bohr radius of the largest atom under study. For a homogeneous electron gas, where the only physically-significant length scale is defined by the density, the equivalent time step would be $\delta\tau = (r_s^2)/d$, where r_s is the radius of the sphere (circle in 2D) that contains one electron on average, and d is the dimensionality. Time step bias is reduced when the modifications of Ref. [27] are made to the DMC Green's function, and also when higher-quality wave functions are used.

If one has accumulated a significant set of results for $\delta\tau < \delta\tau_{\rm max}$ in determining the extent of the linear regime, the prescription for minimising the error in the extrapolate has the potential to differ from the two-run procedure. If one has a large amount of computing time remaining after determining $\delta\tau_{\rm max}$, the two-run approach is unchanged. In the event that little computing time remains after determining $\delta\tau_{\rm max}$, one should devote the remaining time to the run whose contribution falls the quickest with computer time, *i.e.,* the run i with the most negative value of $\partial\Delta_0/\partial T_i$, which may be found by differentiating Eq. (1.33).

The most robust strategy is to avoid higher-order fitting functions and use only data from within the linear regime for the extrapolation. Though the formalism here can be extended to study higher-order fitting functions, finding the appropriate regimes for higher-order terms would require a larger amount of computational effort and there is a danger of numerical stability and branching problems affecting calculations for very large $\delta\tau$. Linear extrapolation is always an option since the leading-order term in the bias is known to be $O(\delta\tau)$.

As emphasised above, the two-run procedure is useful for managing a limited computing resource in a situation where a number of similar calculations (where the regime of linear bias is transferable) is to be performed. One should exercise caution when attempting to extrapolate $\delta\tau$ to zero for isolated runs, because ideally the extent of the linear regime should be determined to the same accuracy as that desired for the final result. It would be very unreliable to determine $\delta\tau_{\rm max}$ using results with error bars of order x and to then extrapolate linearly through two points in this region with error bars an order-of-magnitude smaller than x.

We highlight the benefits of the two-run extrapolation procedure with some example calculations on the 1D and 2D HEGs, and the Ne atom. For the HEG systems, once the maximum allowed time step $\delta\tau_{\rm max}$ in the linear regime had been determined, pairs of runs were performed at $\delta\tau_2 = \delta\tau_{\rm max}$ and incrementally smaller time steps $\delta\tau_1$. The pairs of runs were each performed using the same total amount of computing time. The time was distributed either to ensure equal-sized error bars or according to the prescription

$T_1/T_2 = (\delta\tau_2/\delta\tau_1)^{3/2}$ to guarantee minimal final extrapolated error. The simulation times were sufficient to ensure that the data could be reblocked for accurate error estimates — the reblocking method is explained in Sec. 2.1. The final extrapolated energy estimates all agreed to within the expected uncertainty, consistent with the assertion that all of the time steps are within the linear regime. The results shown in Figs. 1.1 and 1.2 highlight that, for the range of $\delta\tau_2/\delta\tau_1$ tested, there is strong agreement between the analytical prediction and the DMC results. In particular, the error bar on the extrapolate with the optimal distribution of effort is clearly minimised by the choice $\delta\tau_2/\delta\tau_1 = 4$. The distribution of effort according to $T_1/T_2 = (\delta\tau_2/\delta\tau_1)^{3/2}$ yields a modest computational advantage over the choice $T_1/T_2 = \delta\tau_2/\delta\tau_1$.

For the Ne atom, we performed DMC calculations at 20 different time steps, all inside the linear regime to within the desired accuracy. We then extrapolated through all possible pairs of points, giving 190 extrapolates. Since the computational effort varied between different pairs, it was necessary to normalise the errors in the extrapolate by $\Delta_{\tau\max}$, the error in the energy at the larger time step if all of the resources for that pair had been devoted to it. Figs. 1.1 and 1.2 show the comparison between the data and the analytical prediction. There is good agreement, although the neon data in particular are noisy — this is because the runs for that system were relatively short so that the random errors themselves were subject to uncertainty.

In summary, to minimise the statistical error bar on the DMC energy extrapolated to zero time step, one should perform one DMC calculation at the largest time step $\delta\tau_{\max}$ for which the bias is still linear in the time step and a second DMC calculation with time step $\delta\tau_{\max}/4$. Eight times as much computational effort should be devoted to the latter calculation as to the former. One could use a similar approach to efficiently extrapolate expectation values to other limits within QMC, *e.g.,* to infinite population or to infinite system size.

In this chapter we have summarised the VMC and DMC methods and proposed an efficient scheme for extrapolating the DMC time step to zero. In Chapter 2, we discuss in more detail the calculation of statistical errors from QMC data and examine the effect of uncertainty in the error on the distribution of results.

Chapter 2

Random errors in QMC

In this section, we investigate the frequency with which “outliers” occur in QMC results. We define an outlier as a result located more than a given number of *estimated* error bars from the underlying mean value. For example, one may fit a straight line to DMC energies at small τ. If there are sufficient data points, the linear fit is a good estimate of the underlying mean; one would usually expect, by the central limit theorem (CLT), a fraction 0.32 of the points to deviate from the fitted function by more than a single error bar. Here we address the observation that QMC estimates can lie outside statistical error bars of the underlying mean more often than one would expect were the error bars correctly describing the width of an underlying Gaussian distribution. We will demonstrate that uncertainty in the estimated correlation length is largely responsible for the effect.

We begin with a description of how errors are calculated from QMC calculations, outlining reblocking [30] and direct computation of the correlation length. We then directly observe the numbers of outliers for two systems, the C atom and the Si crystal. By performing a large number of short VMC calculations for each system, we count directly the number of energies occurring more than Q error bars from the underlying mean (as a function of Q), where the error is estimated separately for each run. Each estimate of the statistical error is also implicitly an estimate of the correlation length, as we describe in the next section.

We then derive an analytic expression for the fraction of points expected to lie more than Q error bars from the mean under the assumption that the distribution of local energies is Gaussian. The resulting expression depends on the distribution of estimated correlation lengths. Finally, we compare the expected result from this purely Gaussian model process with that found earlier from VMC, forming conclusions about the validity of the Gaussian assumption and the origins of outliers.

2.1 Reblocking

As described in Secs. 1.1 and 1.4, configurations and expectation values generated from a QMC calculation usually show some degree of serial correlation, making it necessary to account for serial correlation when calculating the statistical uncertainties.

Ideally we would estimate the energy as the average over a set of ν_0 independent and identically-distributed local values. Suppose that the variance of the distribution of local energies, a quantity that is often estimated in order to gauge the quality of a trial wave function, is given by σ_0^2. Let us denote the sample variance by σ^2. The accurate statistical error bar Δ is then $\Delta = \sigma_0/\sqrt{\nu_0}$. In practice, such a calculation will require $n > \nu_0$ local energies due to serial correlation, where $n_{\mathrm{corr}} = n/\nu_0$ is the correlation length, and σ^2 is accessible but σ_0^2 is not. Let us define ν as the estimate for ν_0, the effective number of steps or independent samples. The estimated error is

$$\Delta = \frac{\sigma}{\sqrt{\nu}} = \eta_{\mathrm{err}} \frac{\sigma}{\sqrt{n}} \ , \tag{2.1}$$

where η_{err} is the error factor, equal to the square root of the estimated correlation length. For clarity, let us briefly summarise the variables that we have just introduced; n is the number of local energies generated by a QMC calculation, ν_0 is the effective number of energies, ν is the random estimate for ν_0, n_{corr} is the (exact) correlation length, and η_{err} is the error factor, which is a random estimate of the square root of the correlation length.

The reason for writing the estimated error in the form of Eq. (2.1) is that it makes it clear that the reblocking method of calculating the statistical error is also estimating the correlation length. A serially-correlated data set of length n is reblocked by grouping data points into contiguous blocks of length B and forming a new data set of n/B block averages. Taking weights and non-integer block lengths into account, the reblocked error is the standard deviation of the block averages divided by the square root of the number of blocks *i.e.,* the block averages are treated as a new dataset of independent samples. Qualitatively, when B is larger than the correlation length then the block averages will be

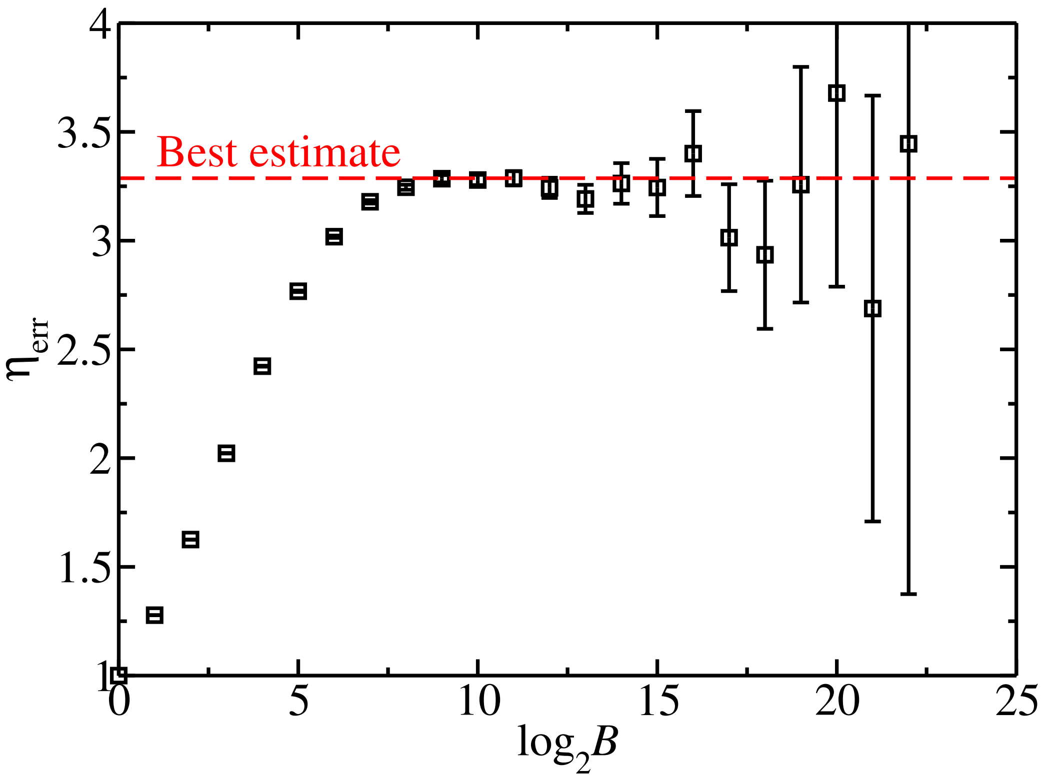

Figure 2.1: Plot of the error factor η_{err} against the block length B. The error factor is the reblocked error divided by the error with $B = 1$ (*i.e.*, the unreblocked estimate). The plot shows VMC data for the all-electron carbon atom and was constructed using 10^7 local energies. The dashed line shows the error estimate obtained by choosing the minimum B for which $B^3 > 2n\eta_{\text{err}}^4$ holds.

approximately statistically-independent and the error will no longer be an underestimate due to serial correlation.

Historically, the value of B was chosen heuristically by examining a plot of B against Δ [or, equivalently, B against η_{err} since $\eta_{\text{err}}(B) = \Delta(B)/\Delta(1)$]. An example is shown in Fig. 2.1. The correlation-corrected estimate of Δ is found by choosing B where the plateau in the error bar occurs. The reasoning behind this is that increasing B from a value of 1 will initially increase the error bar monotonically as serial correlation is removed from the data. When the data are roughly independent, increasing B merely has the effect of reducing the number of points, so that rather than increase with B the error fluctuates around a constant. When B is comparable to the size of the data set, the estimate of the error becomes very noisy, as seen for $\log(B) \gtrsim 16$ in Fig. 2.1.

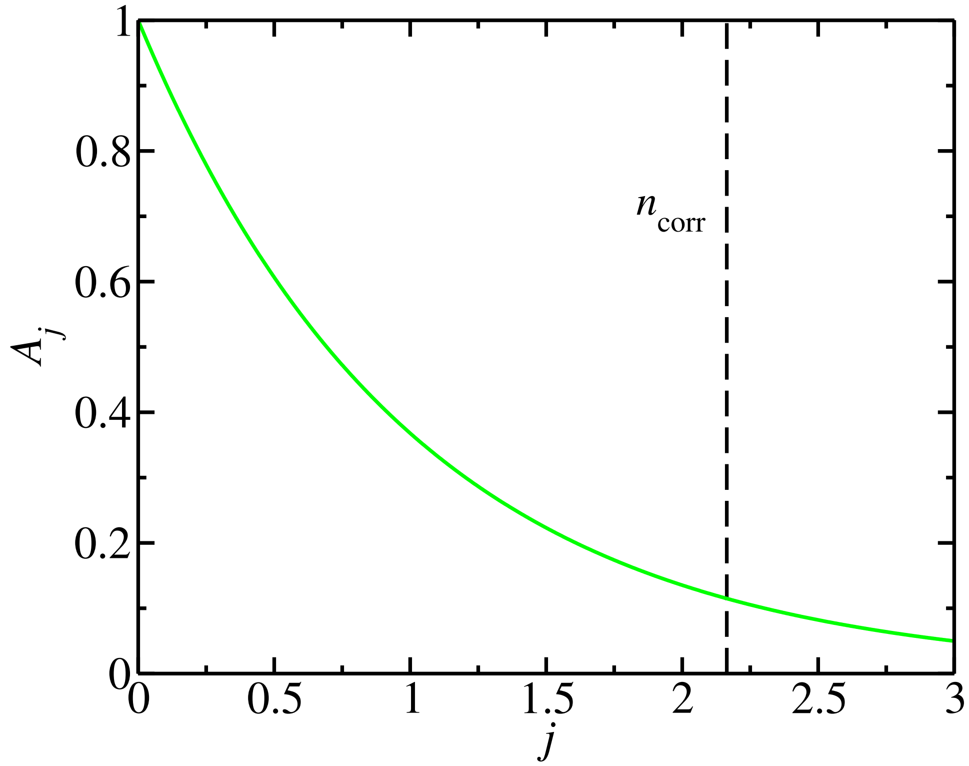

Figure 2.2: Plot of an example autocorrelation function $A_j = \exp(-j)$. The correlation length [from Eq. (2.2)] is 2.16.

Imagining the reblocking procedure to have removed all serial correlation is an over-simplification, however, and misunderstands the nature of serial correlation. One can calculate $n_{\rm corr}$ for a data set with $n \to \infty$ as

$$\begin{aligned} n_{\rm corr} &= 1 + 2\sum_{j=1}^{\infty} A_j \ , \\ A_j &= \frac{1}{\sigma_0^2}\langle (E_i - \langle E \rangle)(E_{i+j} - \langle E \rangle)\rangle_i \ , \end{aligned} \tag{2.2}$$

where E_i is the i-th local energy and A_j is the autocorrelation function. The correlation length is thus a characteristic scale of A_j, but correlations by no means cease for data points separated by more than $n_{\rm corr}$. This means that a systematic error is introduced by choosing a finite B in the reblocking procedure. Figure 2.2 illustrates the simple example $A_j = \exp(-\frac{j}{\alpha})$, for which we obtain the result $n_{\rm corr} = \coth\left(\frac{1}{2\alpha}\right)$.

We can easily convince ourselves that Eq. (2.2) gives a suitable measure of the correlation length. Let us consider the model example discussed by López Ríos [6], although the analysis we perform here is new. Consider a set of independent energies $E_1, E_2, E_3, \ldots$.

Let us now artificially introduce serial correlation by including p repeats of each energy, *e.g.,* $E_1, E_1, E_2, E_2, E_3, E_3, \ldots$ corresponds to $p = 2$. One can easily show that

$$\begin{aligned} A_j &= 1 - \tfrac{j}{p} \,, & j < p \,, \\ A_j &= 0 \,, & j \geq p \,, \end{aligned} \tag{2.3}$$

leading to

$$n_{\mathrm{corr}} = 1 + 2 \sum_{j=1}^{p-1} \left(1 - \frac{j}{p} \right) = 1 + 2 \left(\frac{p-1}{2} \right) = p \,, \tag{2.4}$$

which is clearly an appropriate measure of the length over which correlations are important.

There exist previous quantitative studies into the best choice of B. In Ref. [31], the error in the random error is shown to have statistical and systematic contributions of order $\sqrt{2B/n}$ and n_{corr}/B, respectively. One can define the optimal block size B_{opt} as that which minimises the sum of these two terms,

$$B_{\mathrm{opt}} = \sqrt[3]{2nn_{\mathrm{corr}}^2} \,, \tag{2.5}$$

so that for $B < B_{\mathrm{opt}}$ the statistical noise in the error bar is small compared to the systematic uncertainty originating from the truncation of A_j — the distribution of error bars is narrow but its peak is far from the accurate value. For $B > B_{\mathrm{opt}}$, the statistical error in the error dominates due to the small number of block averages — in this case the distribution of errors has a broad peak whose centre is close to the accurate value.

In the presence of a reliable estimate for n_{corr}, Eq. (2.5) is trivial to evaluate. Otherwise, one must estimate n_{corr} from the data themselves. This can be done self-consistently by reducing B from the largest possible value and choosing the last block length encountered before $B^3 > 2n\eta_{\mathrm{err}}^4$ is violated. In practice we consider only powers of two for B, since it is logarithmically-distributed [31]. It is preferable to overestimate the statistical error and have a small systematic component, since studies usually quote the statistical error bar and assume the systematic error to be small. The line labelled 'Best estimate' in Fig. 2.1 shows the result of following this procedure. Figure 2.3 shows the comparison between B^3 and $2n\eta_{\mathrm{err}}^4$, demonstrating how the estimate in Fig. 2.1 was determined.

The plot of Fig. 2.1 represents a long run of 10^7 local energies; the plateau is clear and choosing the peak by hand agrees closely with the method described above. For shorter

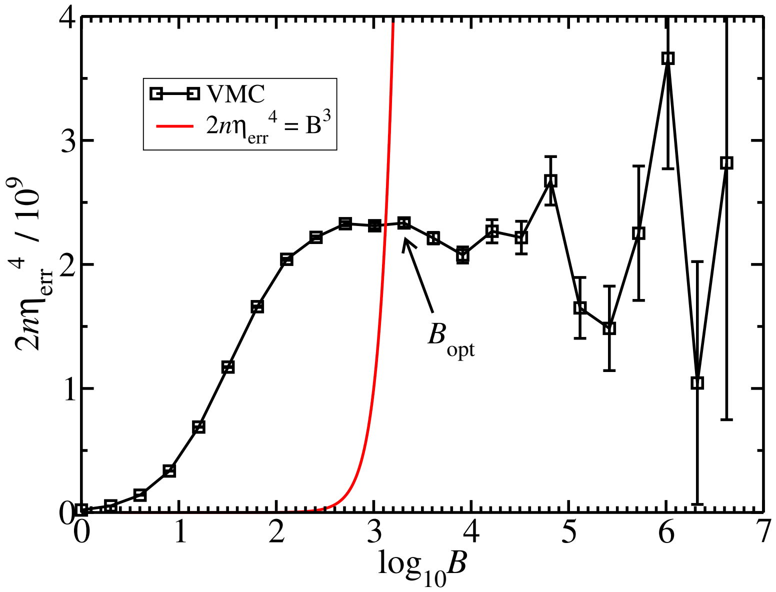

Figure 2.3: Plot showing where the criterion $B^3 > 2n\eta_{\rm err}^4$ for the choice of block length is satisfied. If B is reduced from its maximum value, the last block length for which the inequality is satisfied is $B_{\rm opt} = 2048 = 2^{11}$.

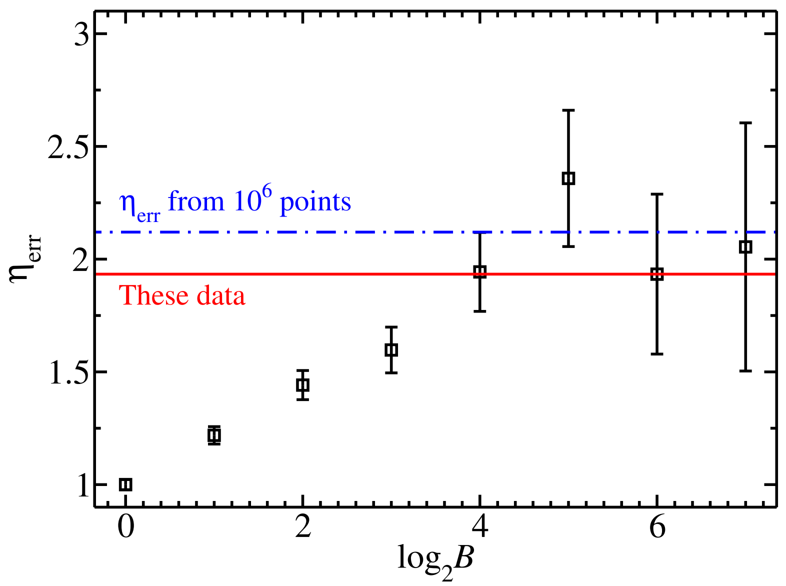

Figure 2.4: Reblocking analysis for a run of 10^3 VMC steps with the CH_4 molecule. The solid line labelled 'These data' shows the error obtained from demanding that $B^3 > 2n\eta_{\rm err}^4$, and corresponds to using $B = 64$ with the 10^3 local energies. The dash-dot line labelled '$\eta_{\rm err}$ from 10^6 points' shows the error factor that one obtains from performing a much longer run.

runs this is not the case and noise in the plot can make it difficult to discern where Δ plateaus. Furthermore, even calculations of $B_{\rm opt}$ are subject to noise in the estimated correlation length. Figure 2.4 shows an example of this; the calculation was 10^3 VMC steps with the CH_4 molecule. Separately, we also performed a calculation on the same system with 10^6 steps, allowing a much more accurate estimate of $n_{\rm corr}$. We construct a more accurate estimate of the error bar for the shorter run using the estimated correlation length from the longer run — this is also shown in the plot.

It is clear that uncertainty in the correlation length is equivalent to uncertainty in the error bar. We now investigate how this affects the distribution of results, focusing in particular on the numbers of outliers.

2.2 VMC calculations

Before performing any sort of analysis using assumed PDFs, we directly access the distribution of results from an ensemble of VMC calculations. We have performed VMC calculations for two typical examples of systems currently studied; the all-electron carbon atom and a periodic crystalline silicon system. For the C atom we performed 5×10^4, 2×10^4 and 10^4 calculations of length 200, 500 and 1000 steps, respectively. The Si system used a periodic simulation cell containing 54 silicon atoms, where the $1s^2 2s^2 2p^6$ electrons were described by pseudopotentials. For the Si system, we performed 1.5×10^5, 7.5×10^4 and 3×10^4 calculations of length 100, 200 and 500 steps, respectively.

Each short calculation yields an energy and estimated error. The errors were obtained by reblocking each short run individually using the criterion of Eq. (2.5). From the data we estimate the probability $\mathrm{P}\left(\delta\bar{E} > Q\Delta\right)$ of observing a VMC energy $\bar{E}$ at a position more than $Q\Delta$ from the true mean E_0, where $\delta\bar{E} = |\bar{E} - E_0|$ and Δ is the *estimated* error bar, itself also a random variable. The underlying mean E_0 is calculated accurately using a much longer run or by averaging over all of the shorter runs. If the error bars exactly described the width of an underlying Gaussian distribution, one would expect $\mathrm{P}\left(\delta\bar{E} > Q\Delta\right) = \mathrm{erfc}(Q/\sqrt{2})$. Figure 2.5 shows VMC energies and error bars for the C atom, and the symbols in Figs. 2.8 and 2.9 show the deviation of the VMC results from the ideal Gaussian case.

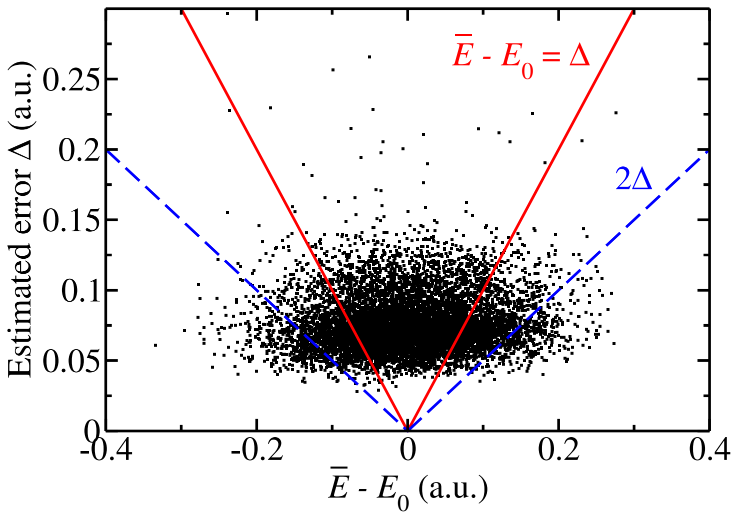

Figure 2.5: Distribution of average energies and estimated error bars from performing 10^4 all-electron VMC calculations for the C atom. Each calculation consisted of 1000 steps and the error bars were obtained by reblocking. A numerical estimate of $\mathrm{corr}(|\bar{E} - E_0|, \Delta)$ yields a correlation coefficient of 5.65×10^{-2}.

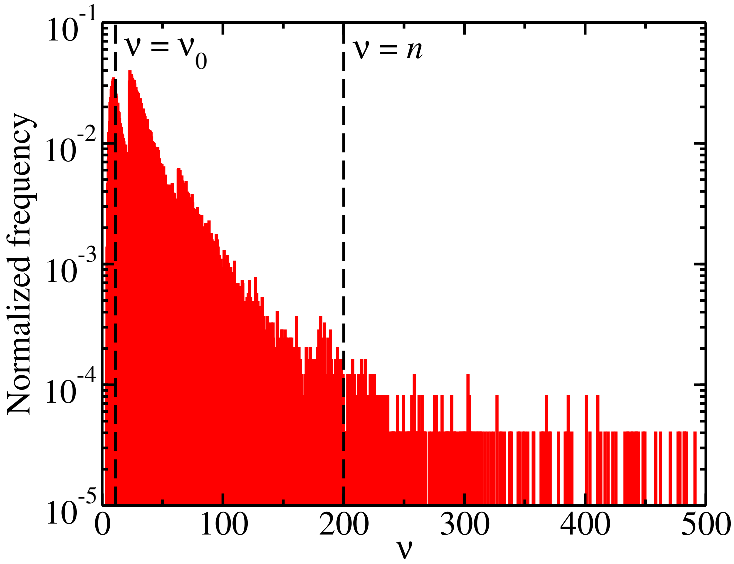

Figure 2.6: Distribution of $\nu = n/\eta_{\mathrm{err}}^2$ from performing 5×10^4 all-electron VMC calculations for the C atom. Each calculation consisted of $n = 200$ steps and the error factors were obtained by reblocking. The dashed lines show the accurate effective number of steps, $\nu_0 = 18.1$, and the effective number of steps corresponding to no serial correlation, $\nu = n$.

By estimating the statistical error bar for each run (and also computing the uncorrected error in each case), we are able to estimate $\mathrm{p_{ind}}$, which is the distribution of the estimated effective number of steps $\nu = n/\eta_{\mathrm{err}}^2$, where n is the number of VMC steps and η_{err} is the error factor of Eq. (2.1). An example histogram of $\mathrm{p_{ind}}$ is shown in Fig. (2.6); one can see that ν is occasionally estimated to be larger than n. This is clearly unphysical, stemming from noise in the estimate of the correlation length, and results in underestimation of the statistical error bar. The distribution $\mathrm{p_{ind}}$ appears to decay at large ν as ν^{-A}, where A is between 4.5 and 6.5.

2.3 Gaussian model

We now replace VMC sampling with an ideal process where the underyling distributions are Gaussian. Our starting point is the distribution of local energies, $\mathrm{p_{loc}}$, from which energies are drawn at successive points along the random walk in configuration space. The quantity of interest is again the probability $\mathrm{P}\left(\delta\bar{E} > Q\Delta\right)$ of observing a sample mean energy $\bar{E}$ at a position more than $Q\Delta$ from the true mean E_0.

Let us assume that the distribution of local energies is Gaussian,

$$\mathrm{p_{loc}}(E_{\mathrm{L}}) = \frac{1}{\sqrt{2\pi}\sigma_0}\exp\left(\frac{-(E_{\mathrm{L}}-E_0)^2}{2\sigma_0^2}\right) , \tag{2.6}$$

where σ_0^2 is the variance of the distribution. Consider drawing n samples $\{E_i\}$ from the probability density function (PDF) of Eq. (2.6) using the Metropolis algorithm; as described above this yields $\nu_0 \leq n$ independent samples due to serial correlation. For this simple case the sample mean, $\bar{E} = (1/n)\sum_{i=1}^{n} E_i$, has the distribution

$$\mathrm{p_{ave}}(\bar{E}) = \sqrt{\frac{\nu_0}{2\pi\sigma_0^2}}\exp\left(\frac{-(\bar{E}-E_0)^2}{2\sigma_0^2/\nu_0}\right) , \tag{2.7}$$

which shows the well-known result that the width of the Gaussian mean-energy PDF reduces as $n^{-1/2}$ with the number of steps.

The statistical error bar on $\bar{E}$ is calculated from the same set of local energies as the estimate itself. However, since estimates of the correlation length are subject to noise, there is uncertainty in the effective number of independent samples. Although this leaves $\bar{E}$ unaffected, it does influence the estimated error. As before, we define ν as the random estimate of ν_0 and again refer to the PDF $\mathrm{p_{ind}}$ from which ν is drawn.

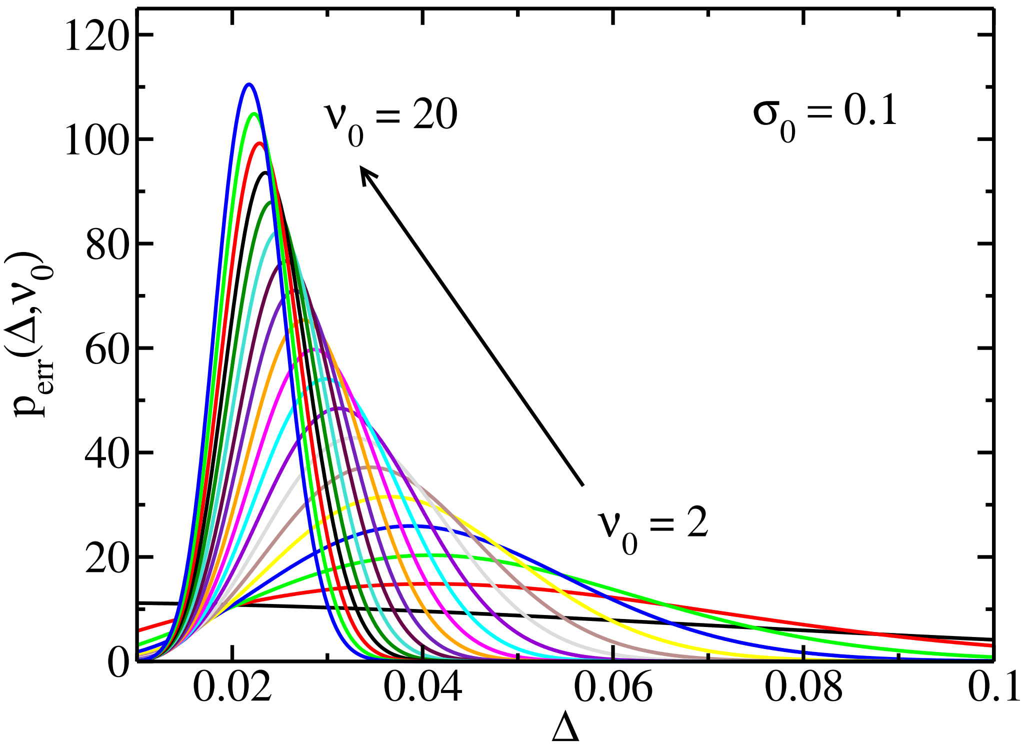

Figure 2.7: Plot of the distribution $\mathrm{p_{err}}(\Delta, \nu_0)$ of error bars, Δ, with ν_0 independent samples and $\mathrm{p_{ind}}(\nu) = \delta(\nu - \nu_0)$. For this plot, $\sigma_0 = 0.1$ was used — the peak in the distribution tends to $\sigma_0/\sqrt{\nu_0}$ as ν_0 increases.

It is well-known that a sum of squares of normally-distributed random numbers follows the chi-square distribution [32]. Since the error bar Δ is related to the sample variance through Eq. (2.1), we can write down the bivariate PDF $\mathrm{p_{err}}$ for Δ and ν,

$$\mathrm{p_{err}}(\Delta, \nu) = \frac{\Delta^{\nu-2} \exp\left[-\frac{\nu(\nu-1)\Delta^2}{2\sigma_0^2}\right] \mathrm{p_{ind}}(\nu)}{\left(\frac{\nu(\nu-1)}{\sigma_0^2}\right)^{\frac{1-\nu}{2}} 2^{\frac{\nu-3}{2}}\, \Gamma\left(\frac{\nu-1}{2}\right)}\,, \tag{2.8}$$

where Δ is only allowed to take positive values and Γ is the Gamma function.

Let us pause to observe the behaviour of the estimated error bar implied by Eq. (2.8), which is plotted in Fig. 2.7. If for a moment we assume $\mathrm{p_{ind}}(\nu) = \delta(\nu - \nu_0)$, *i.e.,* we have an exact value of the correlation length, and set the derivative of Eq. (2.8) with respect to Δ to zero, we find that the modal error is

$$\Delta_{\mathrm{mode}} = \sigma_0 \sqrt{\frac{2}{\nu_0} - \frac{1}{\nu_0 - 1}}\,, \tag{2.9}$$

which leads directly to the well-known $\sigma_0/\sqrt{\nu_0}$ error bar for large ν_0.

It is straightforward to find analytically the probability of observing an energy more than Q error bars from the mean as a function of Q and Δ. This is done by integrating Eq. (2.7),

$$2\int_{E_0+Q\Delta}^{\infty} \mathrm{d}\bar{E}\ \mathrm{p}_{\mathrm{ave}}(\bar{E}) = \mathrm{erfc}\left(\frac{Q\Delta}{\sigma_0}\sqrt{\frac{\nu_0}{2}}\right)\ . \tag{2.10}$$

To find the desired probability, $\mathrm{P}\left(\delta\bar{E} > Q\Delta\right)$, we evaluate the expectation value of Eq. (2.10) with respect to the distribution of Δ and ν,

$$\mathrm{P}\left(\delta\bar{E} > Q\Delta\right) = \int_{2}^{\infty} \mathrm{d}\nu \int_{0}^{\infty} \mathrm{d}\Delta\ \mathrm{p}_{\mathrm{err}}(\Delta,\nu)\ \mathrm{erfc}\left(\frac{Q\Delta}{\sigma_0}\sqrt{\frac{\nu_0}{2}}\right)\ , \tag{2.11}$$

where we have used the fact that the sample mean and sample variance are independent for Gaussian distributed random variables [33,34]. To evaluate the integral of Eq. (2.11), we require the distribution $\mathrm{p}_{\mathrm{ind}}$ and an accurate estimate of the true effective number of steps, ν_0. We will take these quantities from the VMC results of Sec. 2.2, so that the integral of Eq. (2.11) represents an ideal Gaussian process accompanied by the uncertainty in the number of independent samples (and thus the correlation length) that we observe in VMC. The integral of Eq. (2.11) may then be evaluated numerically.

2.4 Results

Figures 2.8 and 2.9 show the actual fractions of outliers from the VMC calculations compared with those predicted by Eq. (2.11), which used $\mathrm{p}_{\mathrm{ind}}$ and ν_0 from the VMC calculations but otherwise assumed a model Gaussian process. The fraction of points occuring more than Q error bars from the mean has been offset by $\mathrm{erfc}(Q/\sqrt{2})$ in the figures to highlight the deviation from the result when the correlation length is known exactly, *i.e.*, $\mathrm{p}_{\mathrm{ind}}(\nu) = \delta(\nu - \nu_0)$.

When n takes smaller values, the uncertainty in the correlation length is greater and the fraction of points which may be classified as outliers is larger. A poor trial wave function could also contribute to the effect by reducing the sampling efficiency, although for the systems studied here we have used the same wave function throughout. In the case of the C atom, instead of the 0.13 probability of observing an energy more than 1.5 error bars from the mean that one would expect on the basis of Gaussian statistics, the VMC results are consistent with a 0.25 probability (for runs of 200 local energies). For the other runs and the silicon system, the effect is of the same order of magnitude.

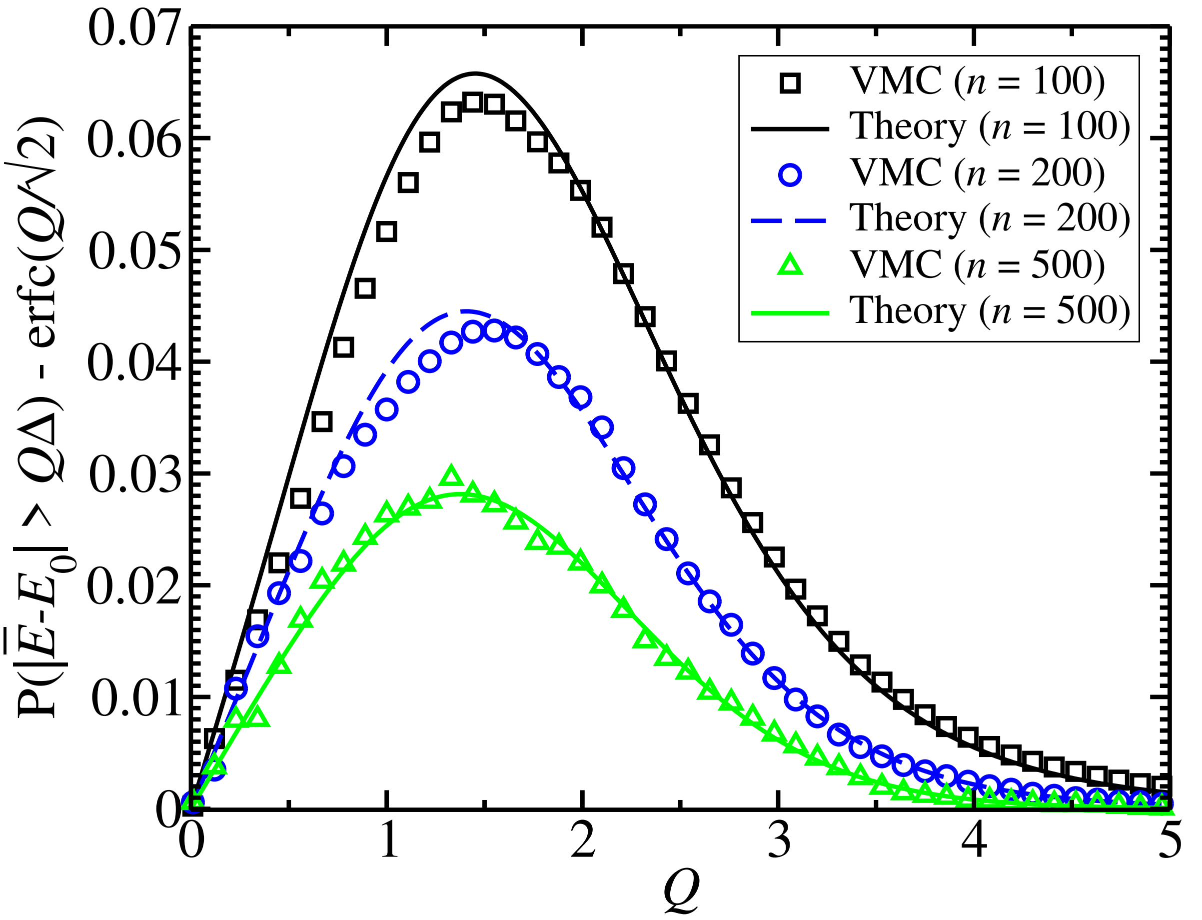

Figure 2.8: Enhancement of the probability of observing an energy more than Q error bars from the mean for 54-atom (216-electron) bulk Si. The square, circular and triangular symbols show the results of VMC calculations of $n = 100$, 200 and 500 local energies, respectively. The number of calculations for each set was $(1.5 \times 10^7)/n$. The lines show the results of evaluating the integral of Eq. (2.11), where ν_0 and $\mathrm{p_{ind}}$ were determined from the VMC data.

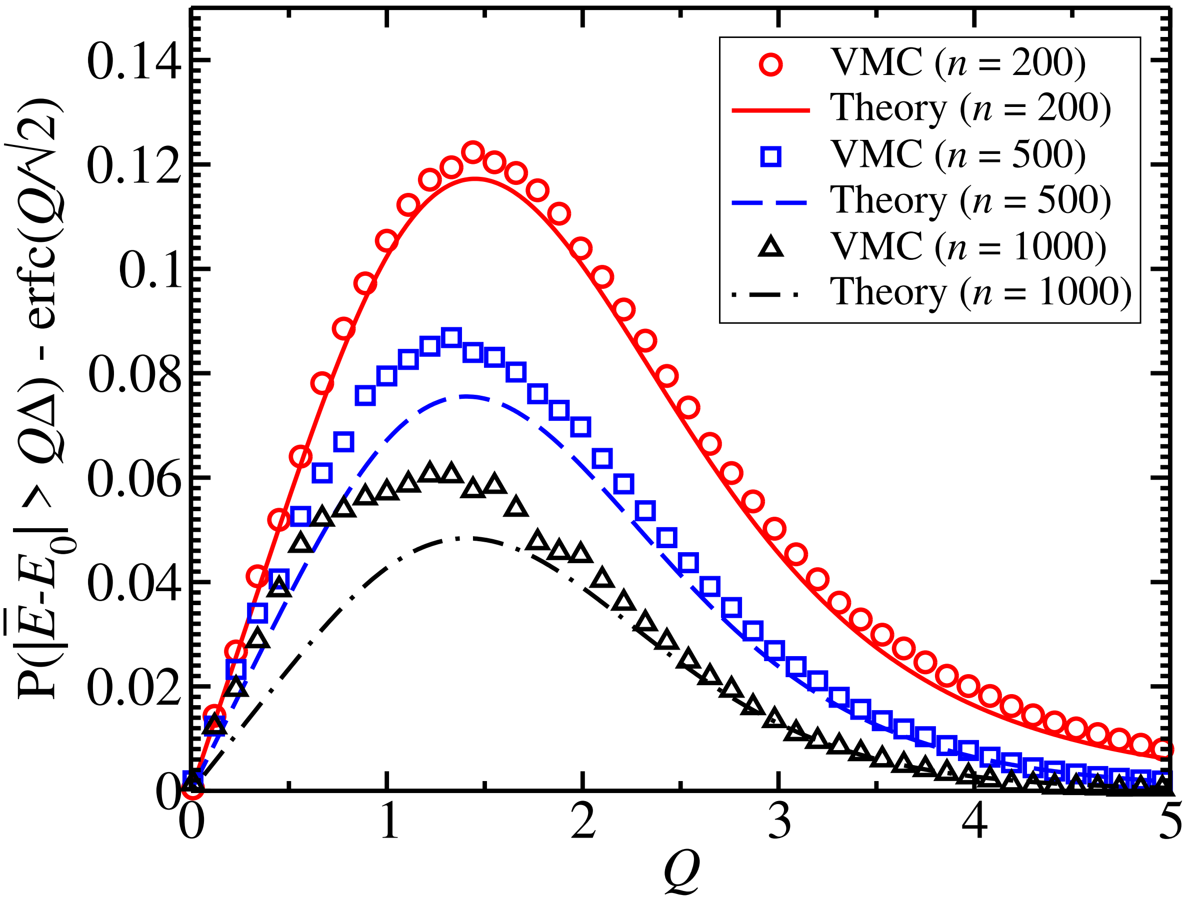

Figure 2.9: Enhancement of the probability of observing an energy more than Q error bars from the mean for the C atom. The circles, squares and triangles represent all-electron VMC results with n = 200, 500, and 1000 local energies, respectively. The number of calculations for each set was $10^7/n$. The lines represent the results of evaluating the integral of Eq. (2.11), where ν_0 and $\mathrm{p_{ind}}$ were determined from the VMC data.

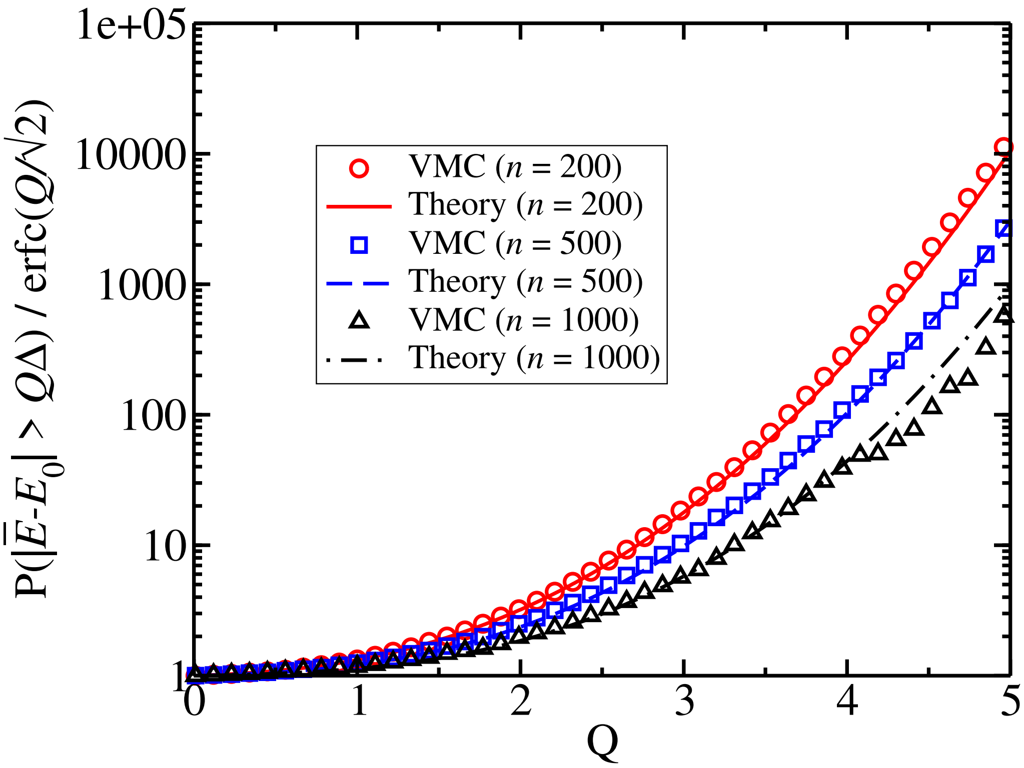

Figure 2.10: Plot of the probability of observing an energy more than Q error bars from the mean as a multiple of the result expected when the correlation length is known exactly. The calculations are for the all-electron carbon atom.

Uncertainty in the correlation length has a rather dramatic effect on the probability of observing extreme outliers. Figure 2.10 shows that for the runs of length $n = 200$, we observe 10^4 times more energies lying more than 5 error bars from the mean than expected. In fact, these data are from 5×10^4 VMC runs, and since $\mathrm{erfc}(5/\sqrt{2}) = 6 \times 10^{-7}$, we do not expect to observe any energies at all more than 5 error bars from the mean — we actually observe 385 such results in the VMC data. Similarly, as shown in Fig. 2.10, for the runs of length $n = 1000$ we find that a factor of 10^3 more energies occur more than 5 error bars from the mean than expected. These are of course very small absolute probabilities, but the enhancement over the ideal case is striking.

We have highlighted that outliers can occur much more frequently than expected in QMC results. The effect depends on the system and run length. In certain situations, there appear to be ways of alleviating the problem. One can produce a return to the expected distribution of results by using a correlation length estimate from a much longer

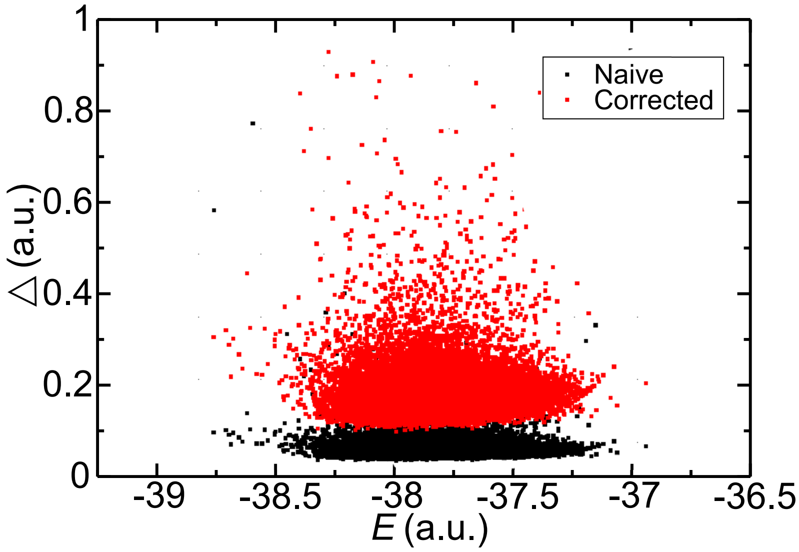

Figure 2.11: Plot of the energies and error bars obtained from VMC with the C atom and $n = 200$. The points labelled 'naive' are from assuming the data to be independent, and the points labelled 'corrected' are from applying a single estimate of the correlation length, $\eta_{\rm err}^2 = 11.1$, to all of the shorter runs (multiplying the 'naive' results by $\sqrt{11.1}$).

calculation. For the C and Si systems, estimating the error bars for each short run using a single more accurate estimate of the correlation length (from a longer run of 10^7 steps or by averaging over the shorter runs), results in a return to $\mathrm{P}\left(\delta\bar{E} > Q\Delta\right) = \mathrm{erfc}(Q/\sqrt{2})$. This is perhaps unsurprising. Figures 2.11 and 2.12 illustrate the point; Fig. 2.11 shows the distribution of energies and error bars from assuming that the local energies are independent (this is equivalent to reblocking with $B = 1$). Also shown is the distribution of error bars obtained by multiplying the errors in the first dataset by $\eta_{\rm err}$, which was calculated from a much longer run, thus making them much more accurate. Figure 2.12 shows the probability of observing an energy more than Q error bars from the mean for the C atom with $n = 200$ using several methods to estimate the errors; the first is the naive (data assumed to be independent) estimate, the second is from reblocking each short run separately (introducing noise into the estimated errors), and the third is from calculating the error factor accurately once (from a longer calculation) and applying it to all of the short runs. Using an accurate estimate of the correlation length to estimate the error bars for all of the shorter calculations essentially returns the frequencies of outliers to the expected values.

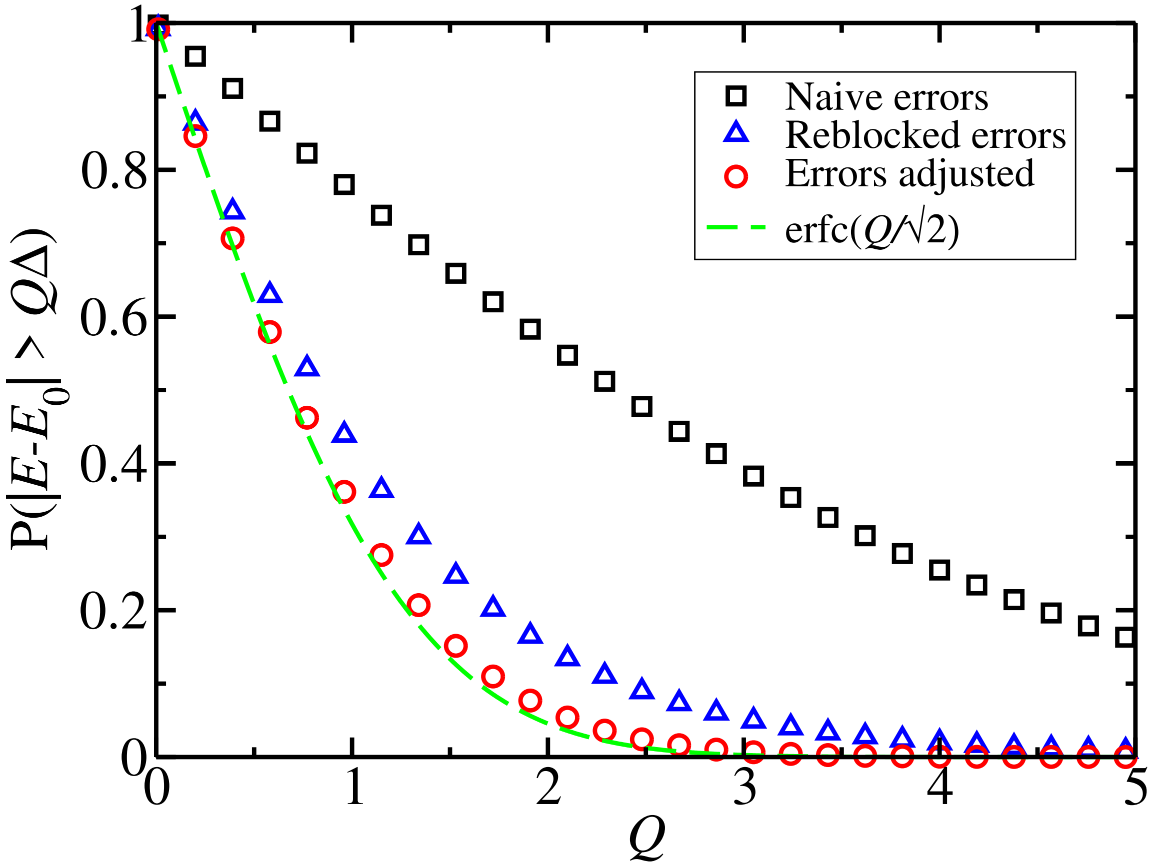

Figure 2.12: Plot of the probability of observing an energy more than Q error bars from the mean for the C atom with $n = 200$. The black squares are from assuming the initial data to be independent. The blue triangles show the result from reblocking each shorter run separately. The red circles show the result of using a single accurate estimate of the correlation length in all of the estimates of the errors for the shorter runs. The result for exact correlation lengths and Gaussian statistics, $\mathrm{erfc}(Q/\sqrt{2})$, is also shown.

For systems exhibiting singularities in the local energy, such as those with approximate wave function nodes or unfulfilled cusp conditions, the distribution of local energies possesses tails $\propto |E_L - E_0|^4$, which one might expect to play a role in determining the frequency with which outliers are observed [35]. This is a commonly encountered situation and is indeed the case for the two systems that we have considered.

We find that the contribution from the non-Gaussian parts of the energy PDF towards the frequency of outliers is statistically insignificant. There is considerable evidence for this statement; first, the integrals based on a purely Gaussian $\mathrm{p}_{\mathrm{loc}}$ agree very well with the VMC data, suggesting that uncertainty in the correlation length is almost solely responsible for the effect. Secondly, attempting to fit a function with power law tails (of the form suggested in Ref. [35]) to the VMC energies yields very small values for the weight under the tails (usually within error bars of zero), even though the distribution of local energies is itself manifestly non-Gaussian. For example, we form a biased estimate for the weight of the power-law tails by fitting Eq. (48) of Ref. [35] to the distribution of energies obtained from 10^4 VMC runs, each of 1000 steps. We find $\lambda_3 = 1.1(8)$ and $\lambda_3 = 0.2(4)$ for the C atom and the bulk Si system, respectively. The χ^2 error in the fit was 0.95 per data point for the C atom and 1.03 per data point for the bulk Si system. In addition, in Appendix A.4 we repeat the analysis of this Chapter using the more complicated heavy-tailed distributions of Ref. [35], finding that the results are largely unchanged.

Furthermore, a numerical estimate of the linear correlation of the average energy and the estimated error bar,

$$\mathrm{corr}(\delta\bar{E},\Delta) = \frac{\sum_i \left(\delta\bar{E}_i - \langle\delta\bar{E}\rangle\right)\left(\Delta_i - \langle\Delta\rangle\right)}{\sqrt{\sum_i \left(\delta\bar{E}_i - \langle\delta\bar{E}\rangle\right)^2}\sqrt{\sum_i \left(\Delta_i - \langle\Delta\rangle\right)^2}}\,, \tag{2.12}$$

gives small values, roughly consistent with the Gaussian case, in which $\delta\bar{E}$ and Δ are independent. For example, estimating $\mathrm{corr}(\delta\bar{E},\Delta)$ using 10^4 VMC energy estimates for the C atom, each an average over 1000 local energies, yields a correlation coefficient of 5.65×10^{-2}. Similarly, estimating the correlation coefficient from 7.5×10^4 VMC calculations for the bulk Si system, each comprising 200 local energies, gives a value of 1.97×10^{-2}.

For isolated calculations where there are no external estimates of n_{corr} available, the problem of the increased frequency of outliers is one of gathering sufficient data for an accurate estimate of the correlation length. This highlights the importance of exploring ways of improving the efficiency of QMC. Where dependence upon several parameters is being investigated for large systems, one should calculate accurately the correlation length with a single long run. Assuming that this value is transferable, the accurate estimate of the correlation length can then be used to calculate the error bars on related shorter calculations in two ways: using Eq. (2.1), replacing η_{err} with the square root of the accurately-estimated correlation length, or using Eq. (2.5), allowing the accurate estimate of n_{corr} to guide the choice of B_{opt}. Agreement of the two methods provides a consistency check.

Chapter 3

QMC with biexcitons

3.1 Introduction

We begin this chapter with an introduction to excitonic systems and an overview of previous work.

Semiconductors are materials that at 0K have a full valence band and an empty conduction band, with a band gap of ≈ 1 eV. Shining light on a semiconductor can excite an electron from the valence band into the conduction band, generating a hole state in the valence band. The energy of such an excitation is lower than the sum of the energies of an isolated electron and an isolated hole because of electron-hole binding. A bound electron-hole pair is called an *exciton*.

In closed-shell materials, such as ionic compounds and inert gases, excitons tend to be small compared to the interatomic spacing; these are called *Frenkel excitons* [36]. In materials where the bonding is more covalent in character, such as silicon and gallium arsenide, excitons have a much larger extent in space [37]. In these cases, excitons are of the *Mott-Wannier* type [38]. We focus here on Mott-Wannier excitons for several reasons. First, the greater radius of a Mott-Wannier exciton implies that its experimental lifetime will be greater. Secondly, the ability to use a theoretical picture of the underlying semiconductor that is averaged over many atomic sites (*i.e.,* the material is accounted for by an appropriate dielectric constant) is very convenient. Thirdly, Mott-Wannier excitons are generally more mobile than Frenkel excitons, making them a vehicle for transporting energy without transporting charge and giving them a large number of applications.

The creation and recombination of excitons is one of the principal mechanisms by which light interacts with semiconductors and is important for a large number of devices and physical phenomena. For example, the functioning of some organic solar cells depends on excitonic properties. In a dispersed-heterojunction solar cell, where the active region is a mixture of two organic semiconductors, an incident photon can create an exciton which then moves diffusively through the material. When the exciton reaches an interface between the electron donor and acceptor materials, it is likely to ionise, allowing the electron and the hole to contribute separately towards the current [39]. Excitons are also instrumental in the functioning of LEDs and photographic devices, and have recently been suggested as providing a possible route to the realisation of a quantum computer [40]. As a result, excitonic systems have been the subject of numerous experimental [41–45] and theoretical [46–49] studies in recent years.

However, it is the phenomenon of Bose-Einstein condensation (BEC) that has motivated the development of the experimental geometry that we investigate here. The observation of BEC of ultracold atoms in the 1990s was the culmination of decades of efforts [50, 51]. One can very easily gain a qualitative understanding of the experimental difficulties. The thermal de Broglie wave length λ is

$$\lambda = \frac{h}{\sqrt{2\pi m k_B T}} , \tag{3.1}$$

where m is the particle mass and T is the temperature. The onset of BEC is expected when λ is comparable to the interparticle separation, which is proportional to $n^{-1/3}$ in 3D, where n is the density. This leads to the relation

$$T_c \propto n^{2/3} m^{-1} , \tag{3.2}$$

where T_c is the condensation temperature. Some of the difficulties of producing a BEC are encapsulated in Eq. (3.2); the density n must be sufficiently small for the composite boson picture to be valid, while at the same time the density must be sufficiently high for condensation to occur at a given temperature. It is also clear that BEC of lighter particles is easier to achieve than that of heavy particles.

In a similar way to atoms with integer spins, excitons may be regarded as composite bosons in the low density limit [43, 52]. Inserting reasonable estimates for the exciton mass and density [43] into Eq. (3.2) yields a condensation temperature of ≈ 1 K, which is much higher than T_c for atoms. However, there are some difficulties associated with observing BEC of excitons rather than atoms. Measurements of excitonic properties are necessarily external probes of semiconducting systems, where there are many processes that have the potential to obscure evidence of BEC [53–55]. Furthermore, the electrons and holes can have a tendency to form an electron-hole plasma rather than an excitonic state.

Several experimental studies of bulk Cu_2O in the 1980s and 1990s claimed to have observed BEC of excitons [56–58]. Initially, $T_c \approx 2$K was expected, which was easily reached in the experiments [53]. However, later studies found that the Auger recombination rate had been underestimated in the previous work by as much as two orders of magnitude [59]. This meant that the exciton densities that were achieved in the experiments were much lower than first thought. Furthermore, the Auger process was acting to heat the excitons, slowing down thermalisation [41]. As a result, BEC was effectively ruled out as an explanation for the experimental observations. That the earlier papers claiming to have observed BEC were largely unchallenged is an indicator of the complexity of the problem. However, as experimental techniques and theoretical understanding improve, there is one particular physical limitation that hinders the creation of an exciton BEC in bulk materials.

The greatest obstacle to the observation of BEC of excitons in bulk semiconductors is that excitons recombine on very short time scales — on the order of nanoseconds. This is typically too little time for thermalisation to occur. The bilayer, or coupled-quantum-well (CQW), geometry has been shown to both greatly extend exciton lifetimes and to speed up the rate at which they cool [41, 42, 45, 60, 61].

The bilayer system comprises alternating thin layers of (typically two) different semiconductors with an electric field applied in the growth direction. The bilayer system thus confines a 2D electron gas and a 2D hole gas in parallel planes separated by a thin layer of a different material. This can give rise to *indirect excitons*, which consist of electron-hole pairs bound across the two layers. Figure 3.1 shows a schematic of the system.

The bilayer geometry possesses several clear advantages over bulk systems for achieving BEC. Indirect excitons have longer average lifetimes because of the barrier between the electrons and the holes, and this barrier can be controlled by the growth process. The

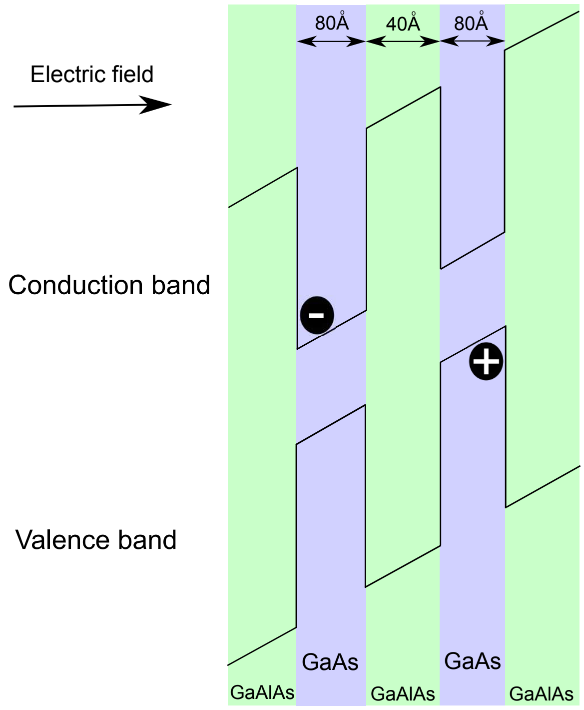

Figure 3.1: A schematic of the bilayer system. The layer thicknesses are typical of experimental values and are taken from Ref. [42].

phase diagram is also reasonably well-understood, allowing choices of density and layer separation that should yield excitonic phases [49]. Finally, it has been found empirically that exciton-phonon coupling, the mechanism by which excitons thermalise, is up to three orders of magnitude greater in bilayer systems than in the bulk, so that thermal equilibrium is much more accessible [60].

Despite careful study of the bilayer phase diagram [49], our understanding of the exciton-exciton interaction in bilayer systems is limited. On the one hand there is a repulsive electrostatic interaction between excitons. For example, if the layer separation is nonzero then the excitons have parallel dipole moments, giving an asymptotically dominant repulsive interaction. Furthermore, the static charge distribution of each exciton has a permanent quadrupole moment in general (even at zero layer separation, provided the electron and hole masses differ), giving another repulsive interaction term [48]. On the other hand, fluctuating dipole (van der Waals) forces result in an attraction between excitons at short range. Because of the existence of the van der Waals forces, it is sometimes possible for biexcitons (bound states of pairs of excitons) to form.

Biexciton formation would inhibit exciton condensation and make interpretation of experimental data more problematic. One could conceive of a condensate of biexcitons [62, 63], but the analysis is more complicated and efforts have largely concentrated on producing condensates of excitons, neglecting the possibility of biexciton formation. Knowledge of the stability of biexcitons in different geometries is thus vital. Furthermore, a better understanding of the interaction between excitons in CQWs will aid the determination of the exciton densities achieved in experiments.

The dependence of exciton and biexciton binding energies on the layer separation has been investigated by Tan *et al.*, who found that, while the exciton binding energy decays slowly as the inverse of the layer separation, the biexciton binding energy decays extremely rapidly [47]. Recent studies of the exciton-exciton interaction using a heavy-hole approximation have found there to be a critical layer separation for each electron/hole mass ratio, beyond which biexcitons become unstable with respect to dissociation into two separate excitons [48, 64]. Here, we report QMC calculations of the binding energies of biexcitons in bilayer systems and exciton-exciton interaction potentials.

This chapter is arranged as follows. In Sec. 3.2 we describe our calculations of the binding energies of biexcitons and investigate the range of layer separations and mass ratios for which biexcitons are stable. In Sec. 3.3 we present our data for the exciton-exciton interaction potential. In Sec. 3.4 we report radial distribution functions (RDFs) for biexcitons. Finally, we draw our conclusions in Sec. 3.5. We report final energies in exciton Rydbergs ($R_y^* = \mu e^4/[2(4\pi\epsilon_0\epsilon)^2\hbar^2]$, where $\mu = m_e m_h/(m_e + m_h)$ is the reduced mass of an exciton and m_e and m_h are the electron and hole masses) and lengths in terms of exciton Bohr radii [$a_B^* = 4\pi\epsilon_0\epsilon\hbar^2/(\mu e^2)$].

3.2 Biexciton binding energies

3.2.1 Model system

We have modelled the coupled-quantum-well system by an idealised two-dimensional (2D) bilayer, in which the electrons and holes are confined to two parallel planes, and the effective mass tensors of the electrons and holes are isotropic. In reality, electrons and holes are free to move within quantum wells of finite width (*e.g.*, one experimental setup [43]

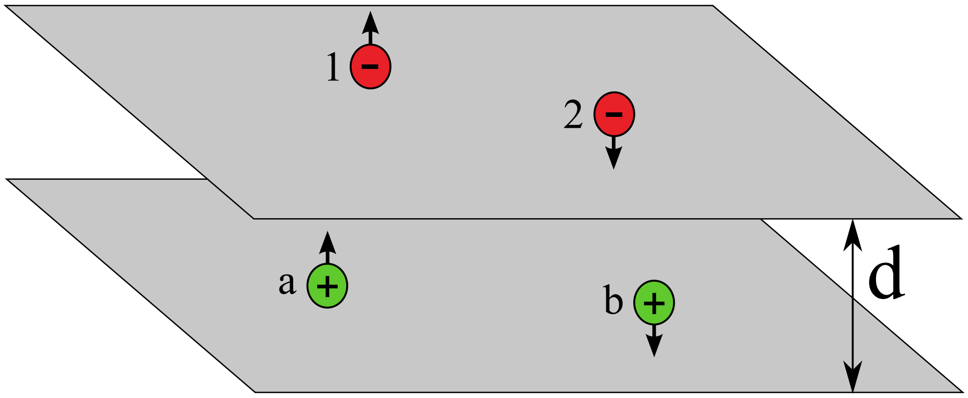

Figure 3.2: The model bilayer system. The electrons are labelled '1' and '2', and the holes are labelled 'a' and 'b'.

has well widths of 8 nm and a well separation of 4 nm), although the Coulomb attraction between electrons and holes should confine the particles to the inner edges of their respective wells. The model we consider here has two electrons in one layer and two holes in the other. We have also restricted our attention to biexciton systems in which the two electrons have opposite spins, as do the two holes, because this is the ground-state spin configuration. This system is ideal for a QMC approach because of the small number of particles and the lack of wave function nodes — it is a fairly simple task to obtain exact energies for the model using DMC.

The biexciton Hamiltonian is

$$\begin{aligned}\hat{H} &= -\frac{1}{2m_e}(\nabla_1^2+\nabla_2^2)-\frac{1}{2m_h}(\nabla_a^2+\nabla_b^2) \\ &\quad +\frac{1}{r_{12}}+\frac{1}{r_{ab}}-\frac{1}{\sqrt{r_{1a}^2+d^2}}-\frac{1}{\sqrt{r_{1b}^2+d^2}}-\frac{1}{\sqrt{r_{2a}^2+d^2}}-\frac{1}{\sqrt{r_{2b}^2+d^2}}\,,\end{aligned} \tag{3.3}$$

where 1 and 2 denote the electron coordinates, a and b denote the hole coordinates, and $r_{12}=|\mathbf{r}_1-\mathbf{r}_2|$, $r_{1a}=|\mathbf{r}_1-\mathbf{r}_a|$, etc. are the in-plane separations. Figure 3.2 provides an illustration of the system. We first investigate the biexciton binding at a range of layer separations d and electron/hole mass ratios $\sigma=m_e/m_h$, The trial wave function we use is

$$\begin{aligned}
\psi &= \exp[J]\Psi_{ee}\Psi_{hh}\Psi_{eh} \\
\Psi_{ee} &= \exp\left(\frac{c_1 r_{12}}{1+c_2 r_{12}}\right) \\
\Psi_{hh} &= \exp\left(\frac{c_3 r_{ab}}{1+c_4 r_{ab}}\right) \\
\Psi_{eh} &= \exp\left(\frac{c_5 r_{1a}+c_6 r_{1a}^2}{1+c_7 r_{1a}}+\frac{c_5 r_{2b}+c_6 r_{2b}^2}{1+c_7 r_{2b}}+\frac{c_5 r_{2a}+c_8 r_{2a}^2}{1+c_9 r_{2a}}+\frac{c_5 r_{1b}+c_8 r_{1b}^2}{1+c_9 r_{1b}}\right) \\
&+ \exp\left(\frac{c_5 r_{1a}+c_8 r_{1a}^2}{1+c_9 r_{1a}}+\frac{c_5 r_{2b}+c_8 r_{2b}^2}{1+c_9 r_{2b}}+\frac{c_5 r_{2a}+c_6 r_{2a}^2}{1+c_7 r_{2a}}+\frac{c_5 r_{1b}+c_6 r_{1b}^2}{1+c_7 r_{1b}}\right) \ , \qquad (3.4)
\end{aligned}$$

where c_1–c_9 are parameters and $\exp[J]$ is a Jastrow factor. As described in Ref. [47], some of the parameter values are fixed by known physical behaviour. We enforce $c_2, c_4, c_7, c_9 > 0$ to prevent singularities in the wave function, and $c_6, c_8 < 0$ to ensure that the wave function is small where electron-hole separations are large. The cusp conditions [65, 66] are enforced by fixing the values of c_1 and c_3, and additionally fixing c_5 when $d = 0$. For $d \neq 0$, we set $c_5 = 0$.

The $\Psi_{ee}\Psi_{hh}\Psi_{eh}$ part of the wave function is of the same form as that used by Tan *et al.* [47], although our wave function is more flexible because we also multiply by a Jastrow factor [16]. The Jastrow function J takes the form

$$\begin{aligned}
J &= (r_{12}-L_{ee})^3\Theta(L_{ee}-r_{12})\sum_{l=1}^{N_{ee}} u_l r_{12}^l \\
&+ (r_{ab}-L_{hh})^3\Theta(L_{hh}-r_{ab})\sum_{l=1}^{N_{hh}} v_l r_{ab}^l \\
&+ \sum_{l=1}^{N_{eh}} w_l \left\{ r_{1a}^l (r_{1a}-L_{eh})^3\Theta(L_{eh}-r_{1a}) + r_{1b}^l (r_{1b}-L_{eh})^3\Theta(L_{eh}-r_{1b}) \right. \\
&\qquad \left. + r_{2a}^l (r_{2a}-L_{eh})^3\Theta(L_{eh}-r_{2a}) + r_{2b}^l (r_{2b}-L_{eh})^3\Theta(L_{eh}-r_{2b}) \right\} \ , \qquad (3.5)
\end{aligned}$$

where $\{u_i\}$, $\{v_i\}$ and $\{w_i\}$ are sets of optimisable parameters and the cutoffs L_{eh}, L_{ee}, and L_{hh} were also optimised. The first, second and third sums describe electron-electron, hole-hole and electron-hole correlation, respectively. The functions are truncated smoothly in Eq. (3.5) so that the kinetic energy is well-behaved when an interparticle separation moves through the relevant cutoff. Excluding cutoffs, the Jastrow factor contained 24 parameters (since $N_{ee} = N_{hh} = N_{eh} = 8$) for $\sigma \neq 1$, but for $\sigma = 1$ the electron-electron correlation is equivalent to the hole-hole correlation by symmetry and the number of parameters is reduced to 16. There are also fewer parameters in Eq. (3.4) when $\sigma = 1$ because $c_1 = c_3$ and $c_2 = c_4$ by the same logic.

The most important feature of Eq. (3.4) is that it can describe states where all four particles are bound and states comprising two well-separated excitons. Generally, two separate excitons emerge when one of either c_6 or c_8 goes to zero.

Our calculations of the binding energy also used a three-body Jastrow factor [3]. For a typical case where $E_{\text{VMC}} - E_{\text{DMC}} = 3 \times 10^{-4} R_y^*$, the reduction in the VMC energy from the inclusion of a three-body term was $10^{-4} R_y^*$. Although the DMC energy is independent of the trial wave function, the statistical efficiency of the method is increased and biases are reduced when the wave function is improved. Obtaining the best possible trial wave function was especially important for the RDF calculations described in Sec. 3.4, where the extrapolated estimator of Eq. (1.30) was used to reduce the effect of the error in the wave function.

We optimised the free parameters of Eq. (3.4) by unreweighted variance minimisation [67–69] and linear-least-squares energy minimisation [14]. The trial wave function can describe the dissociated system more accurately than it can describe the bound system; hence energy minimisation is the more sensible choice for investigating binding, although this depends upon initial parameters and configurations. In other words, the minimum of the energy variance is sometimes a state with two well-separated excitons even when the ground state is actually a bound biexciton. In principle, performing DMC with an initial distribution describing the dissociated system could result in a nonzero binding energy, but in practice the repulsive tails of the exciton-exciton interaction keep the two quasiparticles apart. There were several cases, especially around the critical layer separation, where energy minimisation yielded a bound biexciton at the VMC level, and variance minimisation gave two separate excitons (with a biexciton binding energy of zero). We also found that simply fixing the wave function parameters to keep the particles close together and then allowing the ground state to emerge at the DMC level (*i.e.,* using an inaccurate localised wave function) often allows one to produce the correct binding, although this is extremely inefficient and likely to suffer from increased bias. However, the issue of bias is simply one of computation convenience as long as the biexciton binding is qualitatively correct; one can remove bias by extrapolating to zero time step and using a larger population.

3.2.2 Results

For each layer separation d and electron/hole mass ratio $\sigma = m_e/m_h$, the biexciton binding energy was calculated as $E_b = 2E_X - E_{XX}$, where E_X is the energy of a single exciton and E_{XX} is the energy of the four-body biexciton system. The exciton energy E_X was obtained using a numerically exact Runge-Kutta integration technique, (converged to much greater accuracy than the biexciton QMC results) as described in Ref. [47], while DMC was used to calculate the biexciton energy E_{XX}. The DMC energies were converged with respect to time step and population size; any remaining bias is much smaller than the statistical error bars.

Biexciton binding energies for $\sigma = 0.3$, 0.5, and 1 are shown in Figs. 3.3, 3.4, and 3.5, respectively. It can be seen that our results are close to those of Tan *et al.* [47], the difference arising from our use of exact single-exciton energies. Tan *et al.* used Eq. (3) of Ref. [47] (a rational functional fitted to the exact results) to generate E_X values, introducing a small, systematic error. Removing this error reveals that our DMC data are in statistical agreement with those of Tan *et al.* The random errors in our data are much smaller, so we can locate the layer separation at which the biexciton ceases to be bound. Tan *et al.* also fitted an exponential form to their binding energy data, which resulted in the erroneous conclusion that biexciton binding persists to infinite layer separation. The exponential fits are also shown in Figs. 3.4 and 3.5.

Our DMC results show some deviation from the binding energies obtained by Schindler and Zimmermann [48], especially when $d \to 0$, when $E_b \to 0$, and when the mass ratio is close to 1, because we have performed a full simulation of all four particles in the biexciton, whereas they simulated a pair of excitons interacting via a model potential. In fact, Schindler and Zimmermann calculated an effective exciton-exciton interaction potential using infinite hole masses (and an analogue of the Born-Oppenheimer approximation), and then solved the Schrödinger equation for two (finite-mass) particles interacting via their potential. The deviation of our binding energies from those of Schindler and Zimmermann is approximately $4 \times 10^{-3} R_y^*$ where $E_b \to 0$. At intermediate d the agreement is much better, but below $d \approx 0.1 a_B^*$ we find larger differences, reaching a maximum of almost

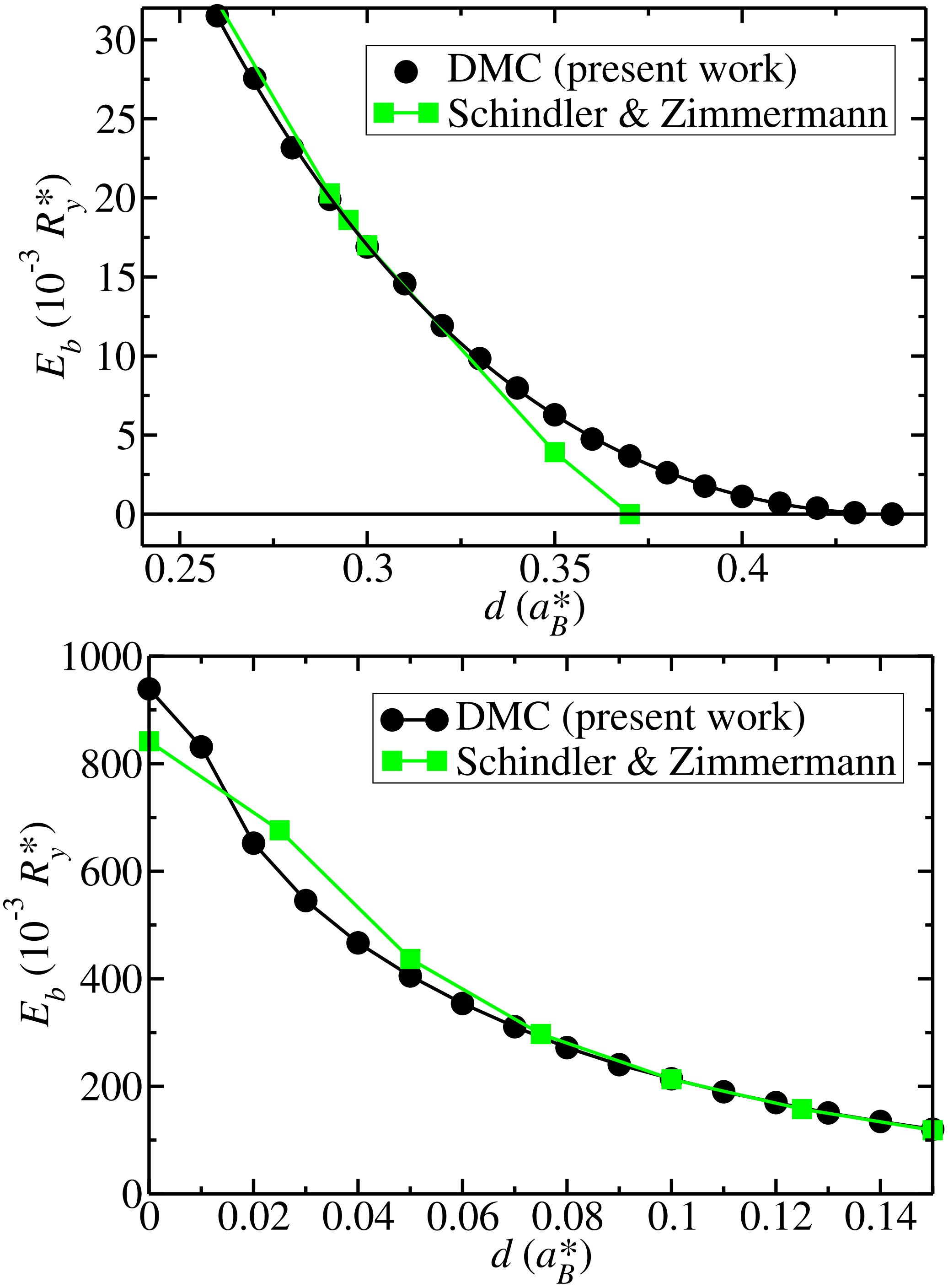

Figure 3.3: Biexciton binding energy E_b as a function of layer separation d for electron/hole mass ratio $\sigma = 0.3$. The upper panel shows the binding energy for layer separations close to the critical separation; the lower panel shows the binding energy for small layer separations.

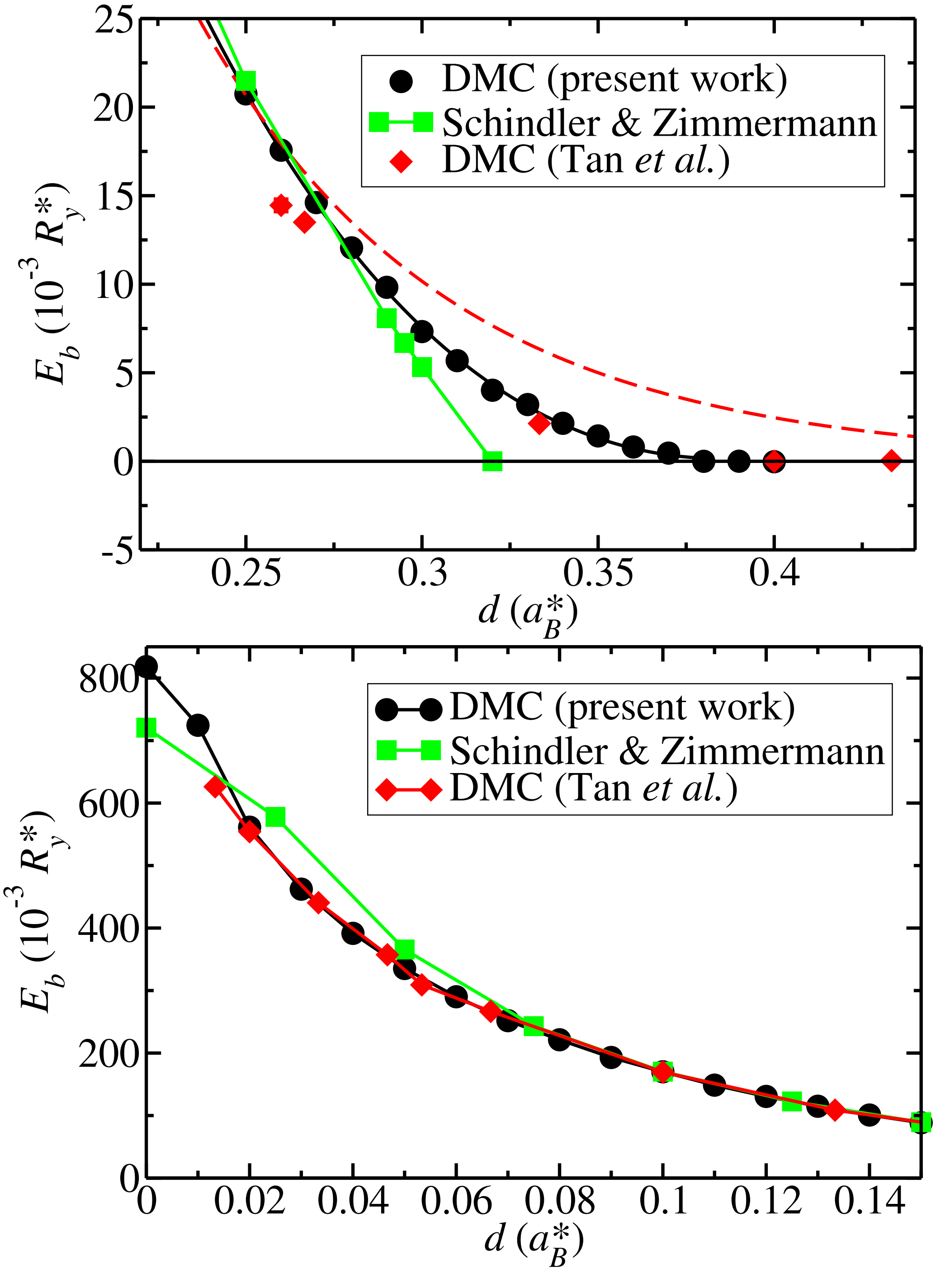

Figure 3.4: Biexciton binding energy E_b as a function of layer separation d for electron/hole mass ratio $\sigma = 0.5$. The upper panel shows the binding energy for layer separations close to the critical separation; the lower panel shows the binding energy for small layer separations. The dashed line shows the exponential fit of Ref. [47].

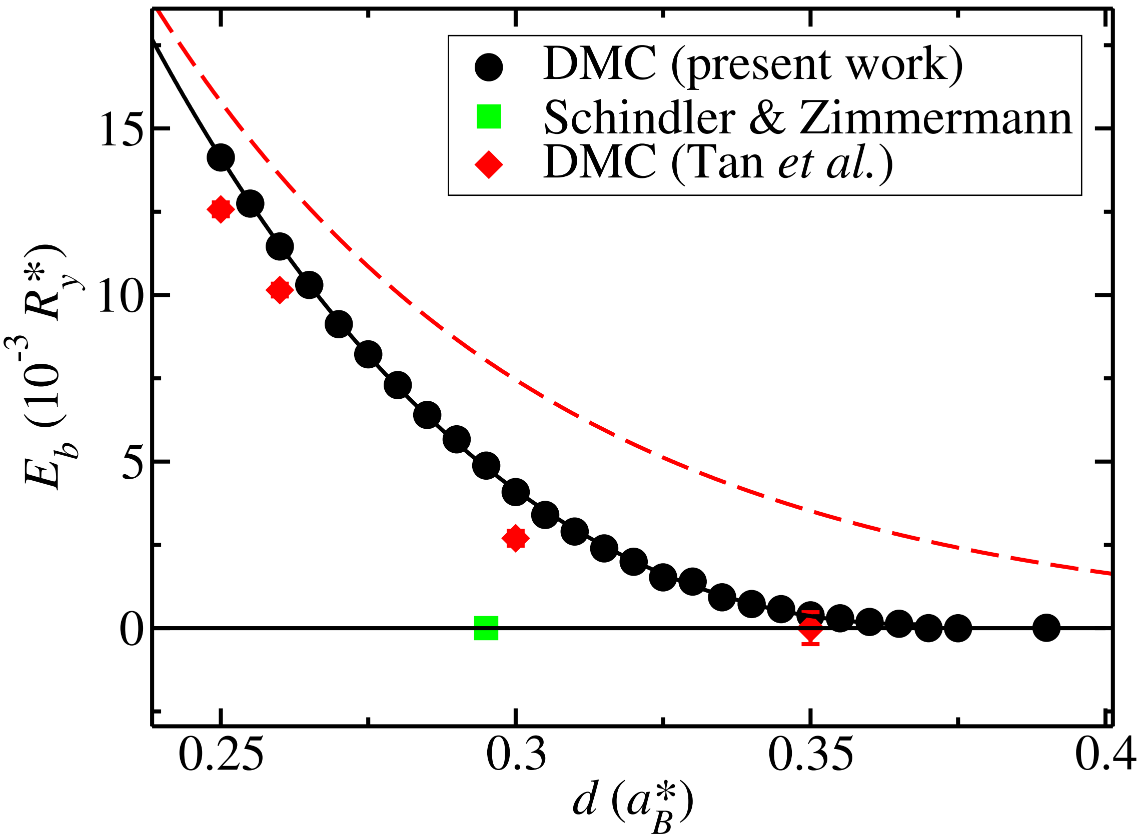

Figure 3.5: Biexciton binding energy E_b as a function of layer separation d for equal electron and hole masses ($\sigma = 1$). The square shows Schindler and Zimmermann's estimate of the critical point at which the biexciton ceases to be bound [48]. The dashed line shows the exponential fit of Ref. [47].

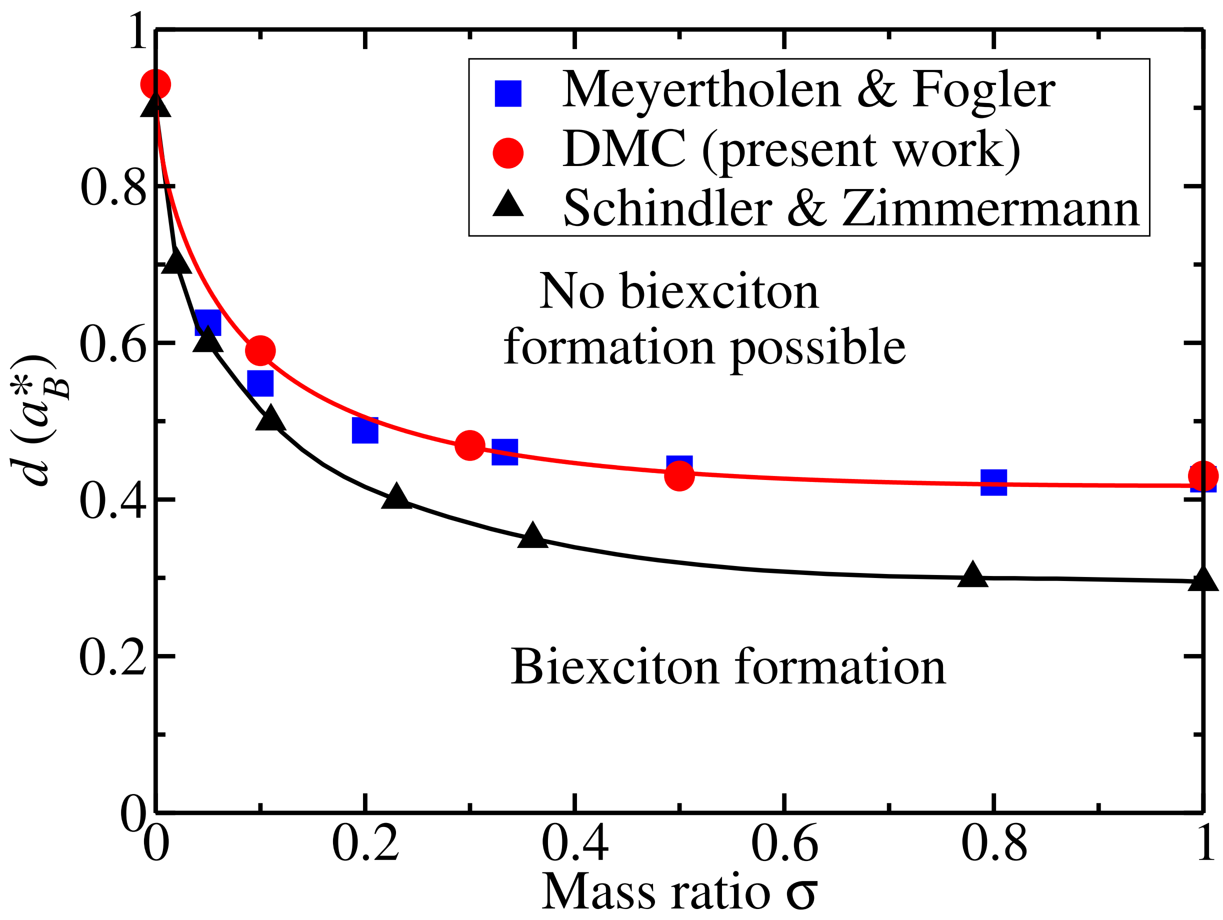

Figure 3.6: The region of biexciton stability from DMC calculations compared with that found by Schindler and Zimmermann [48] and by Meyertholen and Fogler [70]. The critical points were found by extrapolating the biexciton binding energies to zero using the fitting form set out in Ref. [70]. The statistical errors are comparable to the size of the symbols.

$0.1R_y^*$ at $d = 0$, as shown in Figs. 3.3 and 3.4.

As can be seen in Fig. 3.6, which shows the range of σ and d over which the biexciton is stable, we find a somewhat larger region of stability for the biexciton than Schindler and Zimmermann. Let $d_{\rm crit}(\sigma)$ be the critical layer separation, beyond which the biexciton is unbound. As $\sigma \to 0$, the heavy-hole approximation made by Schindler and Zimmermann becomes increasingly accurate, and our results for $d_{\rm crit}(0)$ agree with theirs. On the other hand, for $\sigma = 1$ their interaction potential is less accurate and our value of $d_{\rm crit}(1)$ is therefore significantly higher than theirs.

Meyertholen and Fogler used the stochastic variational method (SVM), which is a highly accurate variational method in which the energy of a trial wave function (constructed from a basis set of correlated Gaussians) is minimised in a similar way to VMC. Our data are mostly in excellent agreement with those of Meyertholen and Fogler [70], although at small σ we find a slightly larger region of biexciton stability. This is not an artifact of the extrapolation, which followed the scheme set out in Ref. [70], for we were able to find points with nonzero binding energies outside the region of stability defined by Meyertholen and Fogler. This is consistent with the variational principle that applies to their results.

One may parameterise the boundary of the region of biexciton stability in Fig. 3.6. Expressing $d_{\rm crit}$ in terms of $\sigma+\sigma^{-1}$ ensures that the correct behaviour is observed upon exchanging the electron and hole masses [i.e., $d_{\rm crit}(\sigma^{-1}) = d_{\rm crit}(\sigma)$]. A suitable fitting function is

$$d_{\rm crit}(\sigma) = \frac{F}{\sqrt{\sigma+\sigma^{-1}}} \tanh\left[G\sqrt{\sigma+\sigma^{-1}}\right] + 0.93\,, \tag{3.6}$$

where the parameter values $F = 1.19(5)$ and $G = -0.50(4)$ give a χ^2 error of 0.4 per data point. The functional form of Eq. (3.6) satisfies most of the conditions derived in Ref. [70]: $d'_{\rm crit}(0)$ is infinite, $d'_{\rm crit}(1) = 0$ and $d_{\rm crit}(0) - d_{\rm crit}(\sigma) \propto \sqrt{\sigma}$ for $\sigma \ll 1$.

Finally, Tan *et al.*, also estimated the biexciton binding energy expected in the experimental setup of Butov *et al.* [42], which is described by Fig. 3.1. The well-width was estimated to be $d = 0.64a_B^*$, while the electron and hole masses in GaAs are known to be $0.067m_0$ and $0.45m_0$ respectively, although these are bulk values. Tan *et al.*, used a hole mass of $0.134m_0$ to give $\sigma = 0.5$, arguing that the bulk value should be reduced due to confinement — this gave them a binding energy of $E_b = 8.2 \times 10^{-5} R_y^*$. Even taking the unmodified bulk values for the effective masses, we find that these parameters correspond to the region of Fig. 3.6 where biexcitons are unstable with respect to separation into two excitons; for $\sigma = 0.15$ we find $d_{\rm crit} = 0.54a_B^*$, and for $\sigma = 0.5$ we find $d_{\rm crit} = 0.43a_B^*$, both of which are below the experimental estimate.

3.3 Exciton-exciton interaction

In the previous section we simulated all four particles to find the ground state, yielding information about the stability of biexcitons. Now, we constrain two excitons to lie a certain distance apart in order to probe the exciton-exciton interaction. This is of interest to experimentalists, for example, who might wish to assume a certain form for the interaction in order to estimate exciton densities.

The exciton-exciton interaction potential $E_I(R)$ at separation R is defined to be the energy of a biexciton system in which the centres of mass of the two excitons are constrained to be a distance R apart, minus the energies of two isolated excitons. The Hamiltonian for the constrained biexciton system may be written as

$$\begin{aligned} \hat{H} &= -\frac{1}{2\mu}\left(\nabla_1^2+\nabla_2^2\right) - \frac{1}{\sqrt{r_1^2+d^2}} - \frac{1}{\sqrt{r_2^2+d^2}} \\ &\quad + \left|\mathbf{R}+\frac{\mu}{m_e}(-\mathbf{r}_2+\mathbf{r}_1)\right|^{-1} + \left|\mathbf{R}+\frac{\mu}{m_h}(-\mathbf{r}_1+\mathbf{r}_2)\right|^{-1} \\ &\quad - \left[\left|\mathbf{R}-\frac{\mu}{m_h}\mathbf{r}_1-\frac{\mu}{m_e}\mathbf{r}_2\right|^2+d^2\right]^{-1/2} - \left[\left|\mathbf{R}+\frac{\mu}{m_e}\mathbf{r}_1+\frac{\mu}{m_h}\mathbf{r}_2\right|^2+d^2\right]^{-1/2}, \end{aligned} \tag{3.7}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the in-plane electron-hole separations within the two excitons. The first two potential terms represent the intra-exciton electron-hole potentials, followed by the hole-hole, electron-electron, and finally the two inter-exciton electron-hole terms. DMC calculations can then be performed for an effective two-particle system, with in-plane coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$. Figure 3.7 illustrates the relationship between the position vectors in the effective two-particle system and the unconstrained four-particle system of Sec. 3.2.

The form of trial wave function was the same as that used in Sec. 3.2, but with the electron and hole coordinates being re-expressed in terms of $\mathbf{r}_1$, $\mathbf{r}_2$, and the fixed vector $\mathbf{R}$.

The centre-of-mass constraint may not be used to calculate the interaction potential at very small exciton-exciton separations for $\sigma \neq 0$, because in that limit the repulsion becomes strong enough to dissociate the two individual excitons. This means that although the ground state of Eq. (3.7) at very small R is well-defined, it corresponds to our model

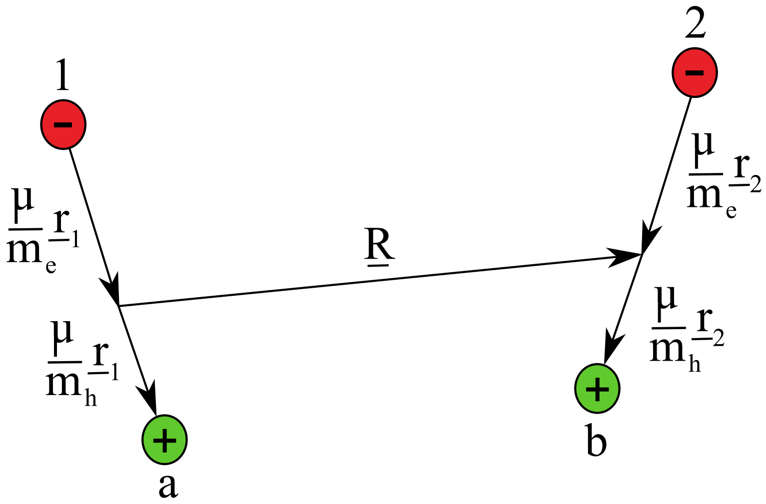

Figure 3.7: Diagram showing the coordinates used in the constrained exciton system. The vector normal to the layers is coming out of the page. $\mathbf{r}_1$ is the in-plane position of particle 1 (corresponding to the electron-hole separation vector of the first exciton), $\mathbf{r}_2$ is the in-plane position of particle 2, and $\mathbf{R}$ is the centre-of-mass separation.

breaking down. In other words, the ground state of Eq. (3.7) in the limit $\mathbf{R} \to 0$ has average intra-exciton separations $\langle|\mathbf{r}_i|\rangle \gg R$; the hole in the first exciton effectively binds with the electron in the second exciton. We have calculated the exciton-exciton potential only at separations R for which the excitons remain bound. Figure 3.8 demonstrates this effect, exhibiting a potential which decreases at small R to physically unreasonable values.

Our DMC calculations yield a smooth exciton-exciton potential. Assuming for a moment that each exciton contains an electron and a hole directly opposite each other in the bilayer, the dipole moment of an exciton is $\mathbf{p} = (0, 0, d)$ in Cartesian coordinates. Since the z-component of the exciton-exciton separation is zero, the dipole-dipole interaction takes the form

$$E_I(R) = \frac{2\mathbf{p}\cdot\mathbf{p}}{R^3} = \frac{2d^2}{R^3}\,. \tag{3.8}$$

The interaction energies shown in Fig. 3.8 do not deviate from Eq. (3.8) by more than $3.5 \times 10^{-5} R_y^*$ above an exciton-exciton separation of $\approx 7a_B^*$. The fits to the interaction potential data are shown in App. A.1. The repulsive tails of the interaction ($R > 10a_B^*$) calculated for pairs of excitons with $0.1 < \sigma < 1$ all collapse onto a single curve for each value of d when scaled into excitonic units, showing a maximum deviation from each other

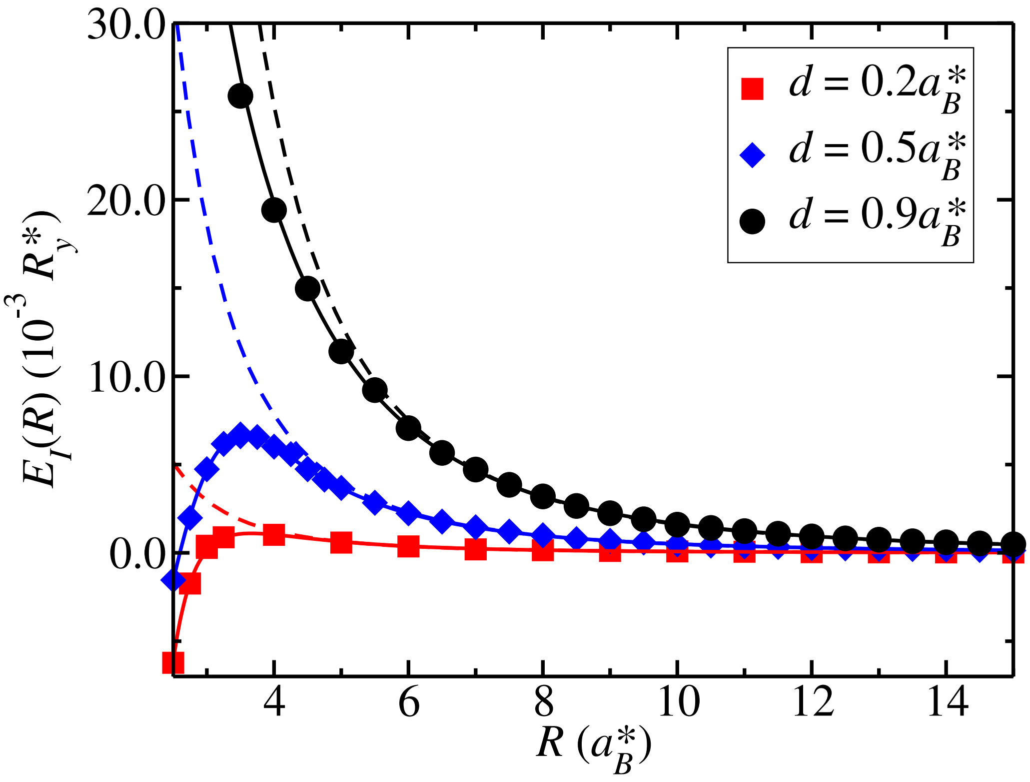

Figure 3.8: Exciton-exciton interaction potential $E_I(R)$ as a function of centre-of-mass separation R with $\sigma = 1$. The solid lines show the fit to the DMC data. Dashed lines show the dipole-dipole interaction energy [Eq. (3.8)].

and Eq. (3.8) of $8 \times 10^{-5} R_y^*$.

For an electron/hole mass ratio of $\sigma = 0$, our results should reduce to the exciton-exciton interaction under the heavy-hole approximation [48]. Figure 3.9 demonstrates the agreement with the interaction potential calculated by Schindler and Zimmermann. Equation (A.1) (in App. A.1) shows a functional form suitable for fitting to our data.

For large layer separations d the interaction is purely repulsive, whereas for smaller d the interaction is attractive at short range. The critical point in the binding occurs near the layer separation for which the minimum in the exciton-exciton interaction potential disappears. Schindler and Zimmermann's approach uses a model exciton-exciton interaction potential which depends on the layer separation d but not the mass ratio σ [48]. Constraining the centre of mass rather than the hole positions allows us to observe the interaction potential for different mass ratios and layer separations, so that we do not need to apply the interaction potential obtained in one system to another with different

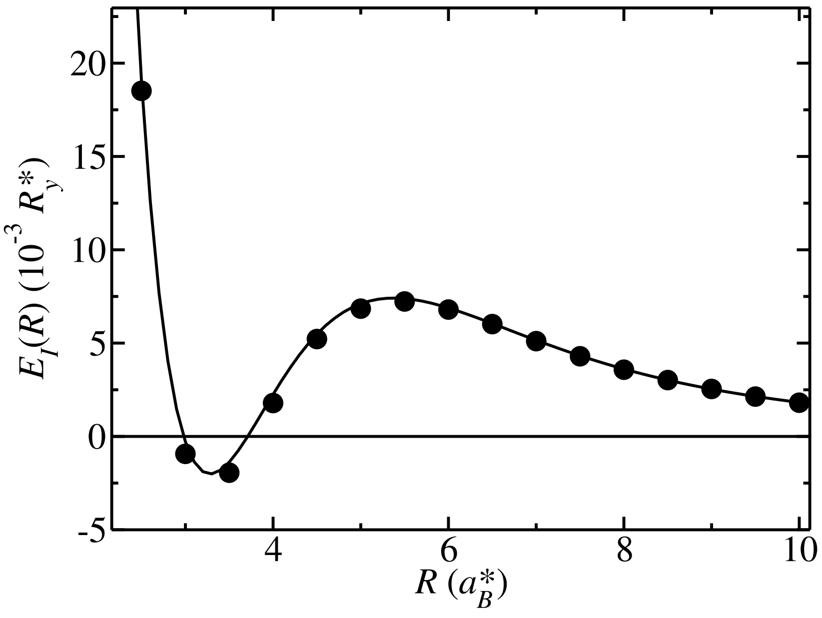

Figure 3.9: Exciton-exciton interaction potential $E_I(R)$ as a function of hole-hole separation R for the heavy-hole case ($\sigma = 0$) with $d = 0.9a_\mathrm{B}^*$. The solid line shows the interaction potential from Ref. [48].

parameters. The dependence of the interaction potential upon σ is clear from Fig. 3.10, and is consistent with the results shown in Fig. 3.6, in which biexcitons are stable at $d = 0.9a_B^*$ for $\sigma = 0$ but not $\sigma = 1$.

For the strictly two-dimensional case ($d = 0$), we can compare our values of the Haynes factor, $f_H = E_b/E_X$, with those of previous work. Usukura *et al.* performed numerically exact variational calculations, finding $f_H = 0.665$ for $\sigma = 0$ and $f_H = 0.193$ for $\sigma = 1$ [71]. These data agree well with our values of $f_H = 0.670(3)$ and $0.19287(2)$ for $\sigma = 0$ and 1, respectively.

3.4 RDFs in biexcitons

We now return to the full four-particle system of Sec. 3.2 and examine the distribution of interparticle separations in our simulations. The RDFs of electrons and holes in biexcitons reveal important information about the physics of biexciton binding. The

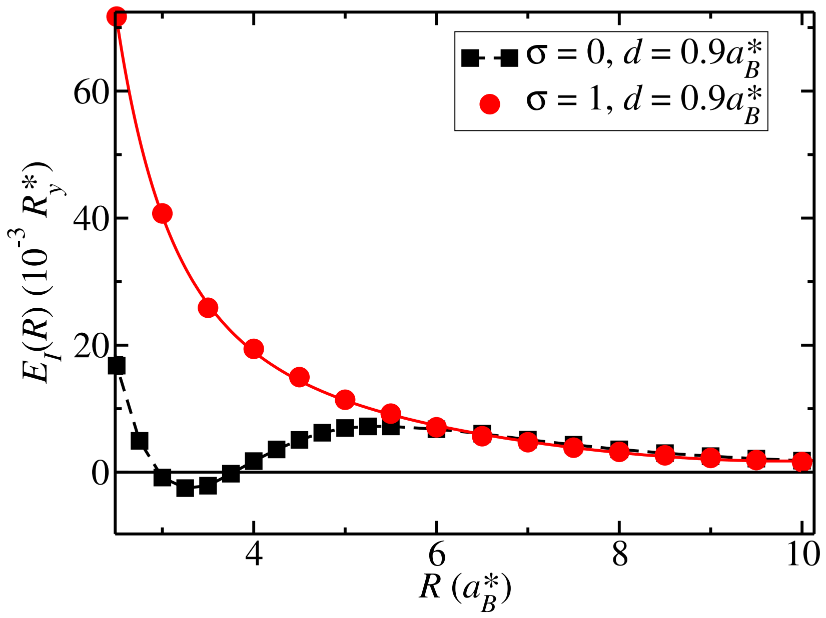

Figure 3.10: Exciton-exciton interaction potential $E_I(R)$ as a function of the (constrained) centre-of-mass separation R for $d = 0.9a_B^*$ and $\sigma = 0$ and 1.

electron-electron RDF is defined as

$$g_{\mathrm{ee}}(r) = \frac{1}{2\pi r} \langle \delta(r_{12} - r) \rangle \,, \tag{3.9}$$

where the angled brackets denote the average over sets of electron and hole coordinates distributed as the square of the ground-state wave function. The hole-hole RDF is defined in a similar fashion. The electron-hole RDF is defined to be

$$g_{\mathrm{eh}}(r) = \frac{1}{8\pi r} \langle \delta(r_{1a} - r) + \delta(r_{2a} - r) + \delta(r_{1b} - r) + \delta(r_{2b} - r) \rangle \;, \tag{3.10}$$

remembering that r_{1a}, r_{2a}, etc. are the in-plane electron-hole separations. The RDFs may be accumulated within QMC by binning interparticle distances. The errors in the VMC and DMC estimates of the RDF [$g^{\mathrm{VMC}}(r)$ and $g^{\mathrm{DMC}}(r)$] are linear in the error in the trial wave function; we present here the extrapolated estimator of Eq. (1.30). Our VMC and DMC RDFs are very close to one another, so the errors in our extrapolated estimates are small. The RDFs presented here have been normalised such that

$$\int_0^\infty 2\pi r g^{\mathrm{ext}}(r) \, dr = 1. \tag{3.11}$$

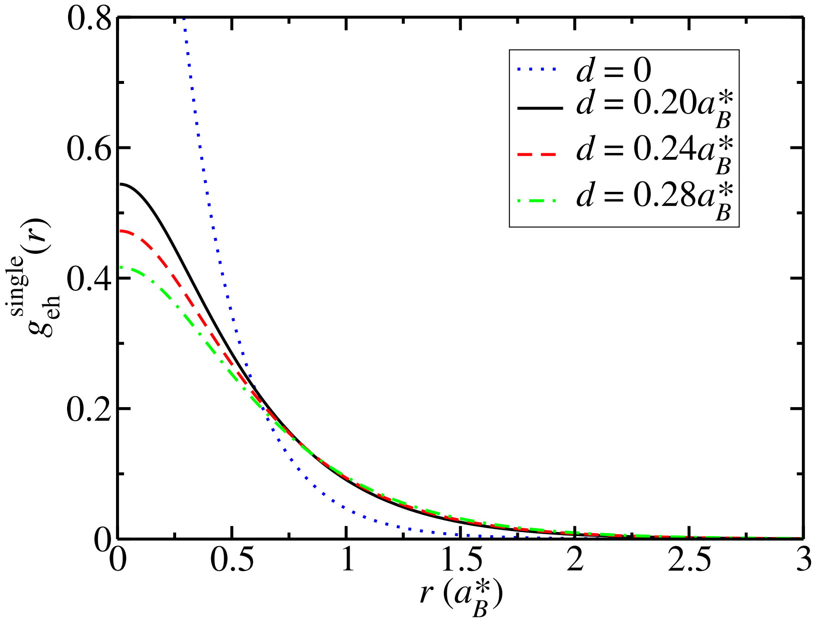

Figure 3.11: RDF $g_{\text{eh}}^{\text{single}}(r)$ for an isolated electron-hole pair from the exact solution of Eq. (2) in Ref. [47], shown at several layer separations for $\sigma = 1$.

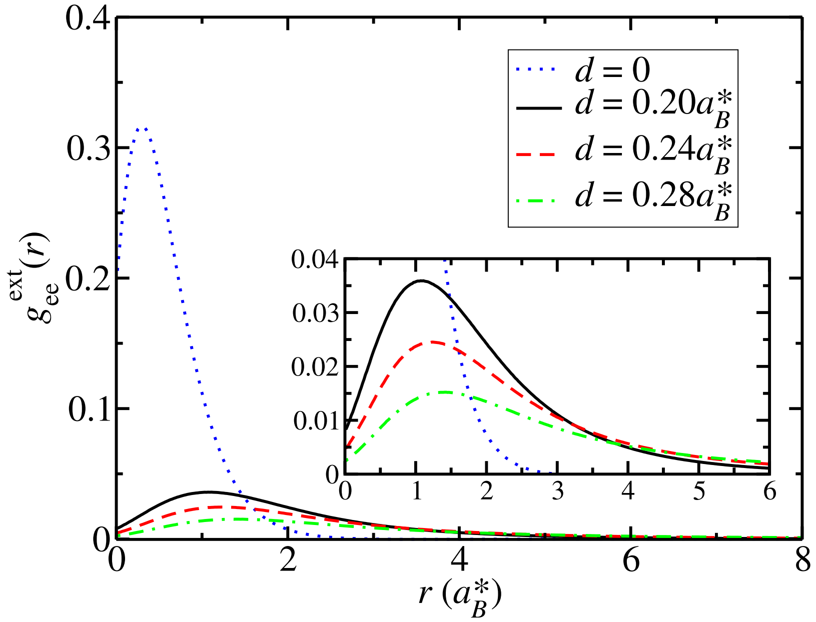

Figure 3.12: Extrapolated electron-electron RDF $g_{\text{ee}}^{\text{ext}}(r)$ for bound biexcitons with $\sigma = 1$. The hole-hole and electron-electron RDFs are identical for equal electron and hole masses. The inset shows the maximum of the RDF in greater detail.

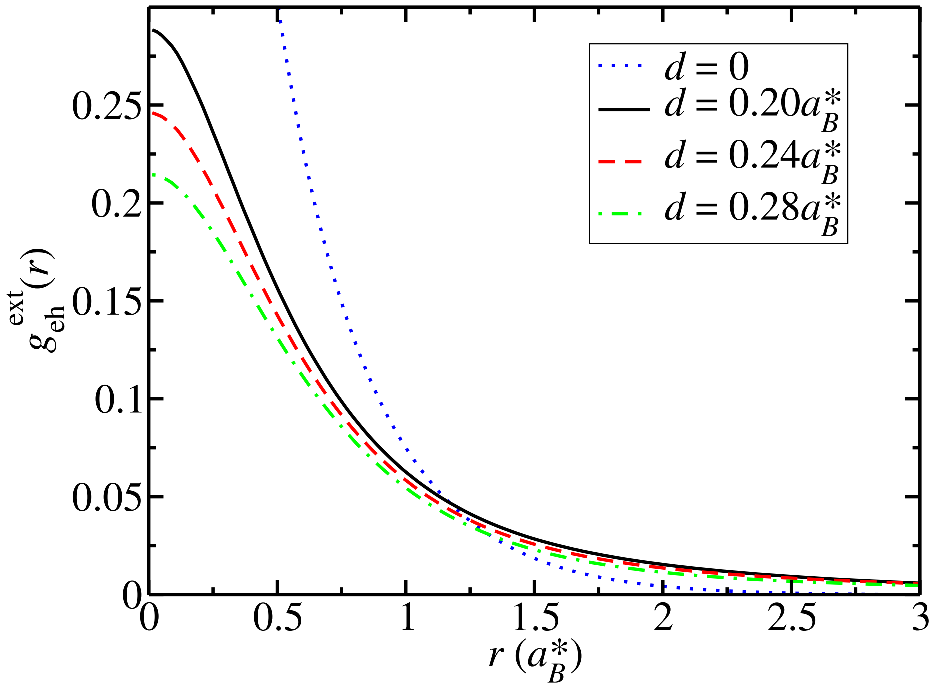

Figure 3.13: Extrapolated electron-hole RDF $g_{\rm eh}^{\rm ext}(r)$ for the biexciton system with $\sigma = 1$ and several bilayer separations d.

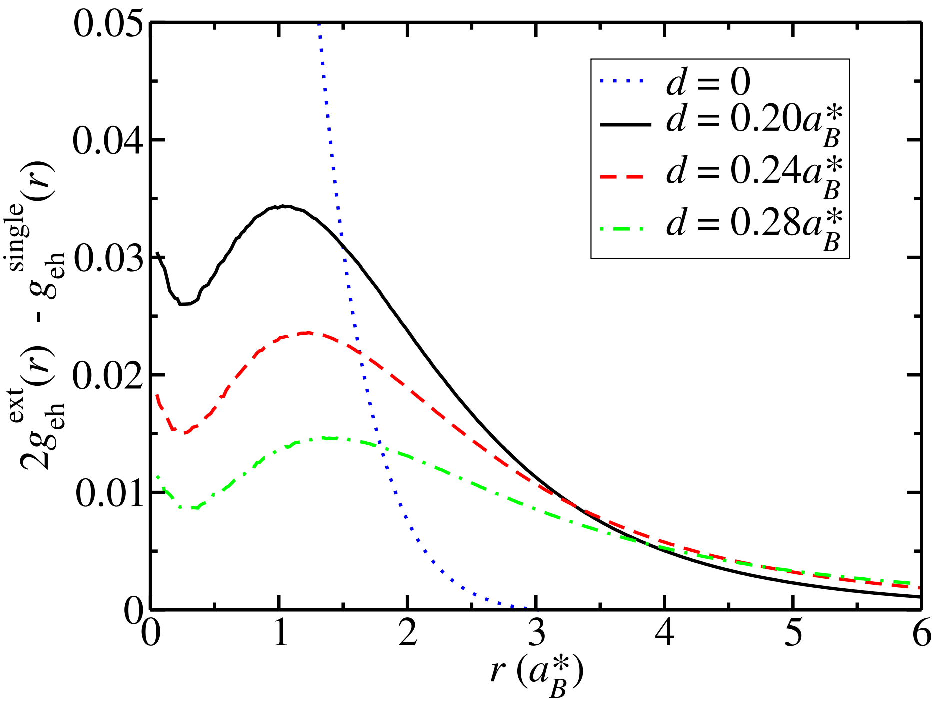

Figure 3.14: Biexciton electron-hole RDF relative to the single-exciton RDF, $2g_{\rm eh}^{\rm ext}(r) - g_{\rm eh}^{\rm single}(r)$, at $\sigma = 1$ and several bilayer separations d.

Figure 3.11 shows the electron-hole RDF for a single exciton, $g_{\rm eh}^{\rm single}$, obtained from the exact numerical solution to Eq. (2) of Ref. [47]. Figures 3.12 and 3.13 show electron-electron and electron-hole RDFs, respectively, for the biexciton system with $\sigma = 1$. At smaller layer separations the electron-hole RDF exhibits a larger peak at zero interparticle separation, and decays more rapidly with interparticle distance.

The size of the biexciton is most easily judged by examining the electron-electron RDF (which is identical to the hole-hole RDF for $\sigma = 1$). The size of the biexciton diverges as the critical layer separation ($d_{\rm crit} = 0.43a_B^*$ for $\sigma = 1$) is approached. At zero layer separation, the electron-electron RDF is negligible for interparticle distances larger than $3a_B^*$ and has a maximum at $0.3a_B^*$.

Although a second peak cannot be discerned in Fig. 3.13, the quantity $2g_{\rm eh}^{\rm ext}(r) - g_{\rm eh}^{\rm single}(r)$ plotted in Fig. 3.14 allows one to see the inter-exciton electron-hole RDF superimposed on the change in the intra-exciton RDF due to the presence of the other exciton. This has the consequence of approximately removing the intra-exciton contribution from the biexciton RDF. The peaks in Fig. 3.14 occur at the same separation as those in Fig. 3.12, confirming that excitons retain their identity in bound biexcitons for large layer separations, even when electrons and holes have equal masses. For zero layer separation there is no discernable peak, however, and the function rises sharply to a maximum at zero interparticle separation. This may be due to the large change in the single-exciton RDF arising from the presence of the other exciton which swamps the inter-exciton electron-hole RDF. We are thus unable to conclude with certainty that excitons retain their identities in bound biexcitons throughout the region of biexciton stability. Attempts to describe biexciton properties by using effective exciton-exciton potentials are expected to be more successful when the layer separation is large.

3.5 Conclusions

We have carried out a QMC study of the interaction between pairs of excitons in bilayer systems. The exciton-exciton interaction potential was calculated by constraining the centre-of-mass separation, which we believe gives a more accurate pair potential at short range than the potential calculated by assuming the holes to be infinitely heavy [48].

We find that for large layer separations, excitons retain their identity when they bind to form a biexciton, suggesting that treating excitons as individual particles is a reasonable approximation. However, by solving the Schrödinger equation for all four particles in a biexciton, we find that the range of layer separations and mass ratios over which biexcitons are stable is somewhat larger than the region of stability predicted using exciton-exciton pair potentials.

Chapter 4

QMC with the one-dimensional electron liquid

4.1 Introduction

Landau's theory of Fermi liquids has proven tremendously successful at describing a wide range of systems of interacting fermions. In particular, the theory legitimises the free electron model by casting fermionic systems in terms of weakly-interacting quasiparticles. Systems of electrons in 1D provide an intriguing example of departure from the Landau Fermi liquid paradigm, exhibiting non-Fermi-liquid behaviour for any finite strength of the electron-electron interaction [72]. Perhaps the simplest model of electrons in one dimension is the one-dimensional (1D) homogeneous electron gas (HEG), which comprises electrons on a uniform positively-charged background.

The strong correlation occurring in 1D results in excitations which are collective in nature rather than electron-like quasiparticles. The theory of Tomonaga and Luttinger gives an appropriate description of the low-energy spectrum of the 1D HEG [73–75]. There are several experimental signatures of the Tomonaga-Luttinger (TL) liquid which distinguish it from the normal Fermi liquid; these are largely accessible to transport and tunnelling experiments. For example, the conductivity of a 1D channel as a function of temperature is expected to vary logarithmically in the presence of weak disorder for the Fermi liquid, and as a power law for the TL liquid [76, 77]. Analogous relations hold for the differential conductivity and the optical conductivity. Spin-charge separation, whereby spin and charge excitations propagate at different characteristic velocities, is also associated with the lack of quasiparticles in the TL liquid [75, 78, 79].

One-dimensional models are easy to envisage, but experimental observation of 1D behaviour is problematic. Low-dimensional systems are never entirely independent of their 3D environment, leading to effects that can obscure the 1D behaviour. Furthermore, the presence of impurities has been shown to alter drastically the behaviour of a TL liquid [72, 80]. However, even in manifestly 3D systems, behaviour unambiguously characteristic of electrons in 1D arises surprisingly frequently. Features associated with the Luttinger model have been observed in organic conductors (e.g., tetrathiafulvalene-tetracyanoquinodimethane and the Bechgaard salts) [81–86], transition metal oxides [87, 88], carbon nanotubes [77,89–92], edge states in quantum Hall liquids [93–95], semiconductor heterostructures [96–100], confined atomic gases [101–103], and atomic nanowires [104]. Theoretical work on electrons in 1D thus has a large region of potential applicability.

The exactly-solvable Luttinger model describes electrons moving in one dimension with short-range interactions and linear dispersion. Studies with long-range interactions have found that the exponents and excitation velocities are nontrivially altered [105]. One thus expects to be able to describe the 1D HEG within the Luttinger model, but the exact behaviour of the parameters of the model is largely unclear. The interactions that we study here are long-ranged, possessing a $1/|x|$ Coulombic tail. This is most applicable to systems where screening is a small effect, such as isolated metallic carbon nanotubes and semiconductor structures where there is negligible coupling to the substrate.

The 1D HEG has been studied with a variety of theoretical and computational approaches. The principal distinction between various studies is the choice of electron-electron interaction. The bare Coulomb interaction, $1/|x|$, which describes an infinitely-thin wire, is perhaps conceptually the simplest choice, although it is largely avoided in the literature [106] in its original form due to the divergence at $x = 0$. Instead, many previous authors have removed the singularity while retaining the long-range behaviour by investigating interaction potentials of the form $V(x) \propto (x^2 + d^2)^{-1/2}$, where d is a parameter related to the width of the wire. This interaction has been studied analytically [105, 107] and numerically [108].

Otherwise, one can derive an effective 1D interaction by factorising the wave function into longitudinal and transverse parts and assuming that the transverse component is the (2D) single-particle ground state of the confining potential. The 1D interaction is then the matrix element of the 3D Coulomb interaction with respect to the transverse eigenfunctions [109, 110]. An example of this is the harmonic wire, in which the transverse confinement is provided by a parabolic potential, leading to a Gaussian density profile in the transverse plane. The full derivation of the harmonic wire interaction is given in App. A.2. The harmonic wire has been studied with QMC [110–112], variants of the Singwi-Tosi-Land-Sjölander approach [113–116], and the Fermi hypernetted-chain approximation [117].

We have studied both the infinitely-thin wire and the harmonic wire using QMC. In this chapter we present QMC calculations of the momentum density (MD), energy, pair-correlation function (PCF), and static structure factor (SSF) of the infinitely-thin wire at a variety of densities and system sizes. The MD results in particular show the non-Fermi-liquid character of the system and allow us to recover one of the parameters of the TL model. The total energy data that we provide are essentially exact and may be regarded as a benchmark for future work. We also present calculations of the MD for the harmonic wire, again extracting one of the TL parameters.

The rest of this chapter is structured as follows: the models we use are described in Sec. 4.2. In Sec. 4.3 we outline details of the calculations that are specific to the 1D HEG. We report the ground state energies of both models in Sec. 4.4.1 and describe the PCFs in Sec. 4.4.2. In Sec. 4.4.3 we give the SSFs that we find for the infinitely-thin wire and in Sec. 4.4.5 we give the MDs for both models. We describe the procedure for estimating a parameter of the TL model in Sec. 4.4.6. Finally, we draw our conclusions in Sec. 4.5.

4.2 Models

4.2.1 Hamiltonian

The Hamiltonians for both of the models studied may be written as

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{N}\frac{\partial^2}{\partial x_i^2} + \sum_{i<j} V(x_{ij}) + \frac{N}{2}V_{\mathrm{Mad}} \ , \tag{4.1}$$

where V_{Mad} is the Madelung energy (the interaction of a particle with its own background and periodic images), $x_{ij} = |x_i - x_j|$ is the distance between electron i and electron j, and $V(x_{ij})$ is the Ewald interaction; this is the interaction of an electron at x_i with another electron at x_j, all of electron j's periodic images, and $1/N$-th of the uniform positive background. The two models studied here differ in the $V(x_{ij})$ and V_{Mad} terms.

4.2.2 Infinitely-thin wire

The Ewald interaction for the infinitely-thin wire may be written

$$V(x_{ij}) = \sum_{n=-\infty}^{\infty} \left(\frac{1}{|x_{ij} + nL|} - \frac{1}{L} \int_{-L/2}^{L/2} \frac{dy}{|x_{ij} + nL - y|} \right), \tag{4.2}$$

which is calculated in practice using an accurate approximation based on the Euler-Maclaurin summation formula; see Eq. (4.8) of Ref. [118] for details. The Madelung term [119] for the infinitely-thin wire is

$$V_{\mathrm{Mad}} = \lim_{x \to 0} \left\{ V(x) - \frac{1}{x} \right\} . \tag{4.3}$$

The interaction of Eq. (4.2) diverges as $1/x_{ij}$ when $x_{ij} \to 0$. In higher dimensions, the divergence in the interaction energy is cancelled by an equal and opposite divergence in the kinetic energy, so that nodes do not necessarily occur where two antiparallel spins occupy the same position [65]. In the infinitely-thin 1D system, the curvature of the wave function is unable to compensate for the divergence in the interaction potential, so the trial wave function has nodes at all of the coalescence points (both parallel and antiparallel spin pairs). The ground state energy is then independent of the spin-polarisation and depends only on the density. In other words, the Lieb-Mattis theorem [120] does not apply and the paramagnetic and ferromagnetic states are degenerate for the interaction of Eq. (4.2). We have examined only the fully spin-polarised case for the infinitely-thin wire. Wires can of course be artificially spin-polarised through the application of an external magnetic field.

4.2.3 Harmonic wire

The second model we have studied describes electrons in a 2D confinement potential given by

$$V_{\perp}(r_{\perp}) = r_{\perp}^2 / 8b^4, \tag{4.4}$$

where b is the width parameter and $r_\perp$ is the magnitude of the projection of the electron position onto the plane perpendicular to the axis of the wire. The Ewald-like interaction for this model may be written as [110, 121]

$$\begin{aligned} V(x_{ij}) &= \sum_{m=-\infty}^{\infty} \left\{ \frac{\sqrt{\pi}}{2b} \exp\left\{[(X_{ij})_m]^2\right\} \mathrm{erfc}\left[(X_{ij})_m\right] - \frac{1}{2b(X_{ij})_m} \mathrm{erf}\left[(X_{ij})_m\right] \right\} \\ &+ \frac{2}{L} \sum_{n=1}^{\infty} E_1\left[(bGn)^2\right] \cos(Gnx_{ij}) , \end{aligned} \tag{4.5}$$

where $G = 2\pi/L$ and $(X_{ij})_m = |x_{ij} - mL|/(2b)$. The Madelung term for the harmonic wire is

$$V_{\mathrm{Mad}} = \lim_{x\to 0} \left\{ V(x) - \frac{\sqrt{\pi}}{2b} \right\} , \tag{4.6}$$

since the real space part of the interaction at the origin is

$$\lim_{x\to 0} \left\{ \frac{\sqrt{\pi}}{2b} \exp\left(\frac{x^2}{4b^2}\right) \mathrm{erfc}\left(\frac{|x|}{2b}\right) \right\} = \frac{\sqrt{\pi}}{2b} . \tag{4.7}$$

Equation (4.5) possesses a long-range Coulomb tail and is finite at $x_{ij} = 0$. A derivation of Eq. (4.5) is given in App. A.2. The paramagnetic and ferromagnetic states are not in general degenerate, so we have probed different polarisations, $\zeta = |N_\uparrow - N_\downarrow|/N$.

At the start of a calculation, we evaluate Eq. (4.5) on a regular grid covering the range $0 < x_{ij} < L/2$ and store the resulting values. Later evaluations of the local energy then use cubic spline interpolation to accurately and efficiently obtain the interaction potential. This makes evaluation of Eq. (4.5) several orders of magnitude faster.

Figure 4.1 demonstrates that our implementation of the interaction of Eq. (4.5) is in agreement with an existing code, and that the accuracy may be chosen by truncating the sums in Eq. (4.5) at different points. The figure shows the calculated energy of an evenly-spaced array of electrons in different simulation cell sizes for a given density. Since the interaction of Eq. (4.5) includes periodic images, the energy per electron in each test system should be the same. We truncated the sums in Eq. (4.5) and chose the spline grid density such that the error in the interaction potential was $\mathcal{O}(10^{-8})$ Ha.

Another test of the implementation of the finite-width interaction is to ensure that performing VMC using only a Slater trial wave function (*i.e.,* without backflow or a Jastrow factor) produces results in agreement with HF calculations. We evaluate the

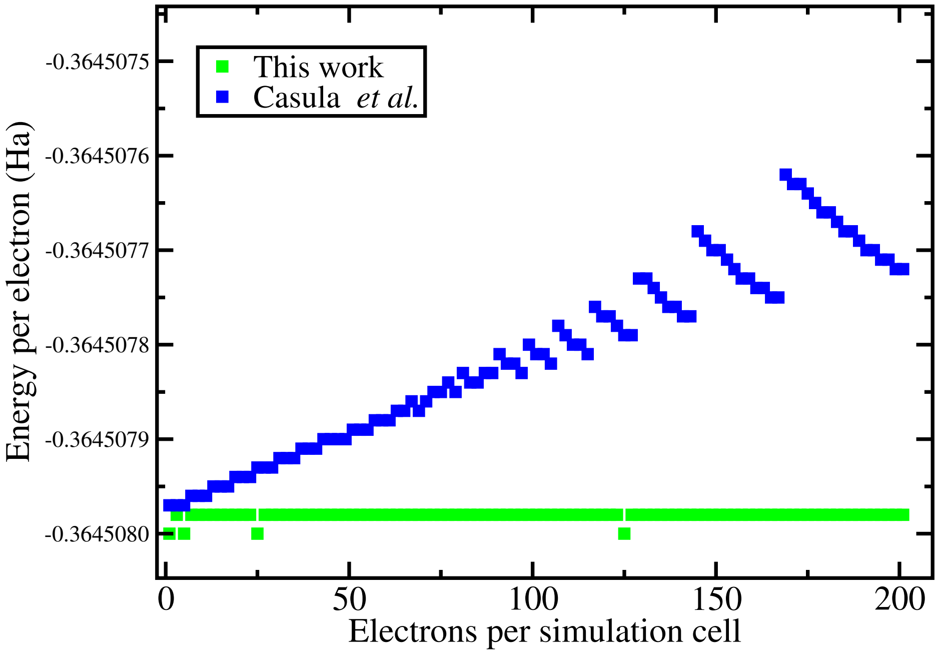

Figure 4.1: Plot of the energy of an infinite array of evenly-spaced electrons calculated using different simulation cell sizes. The green squares show our implementation of Eq. (4.5) and the blue squares were evaluated using the code of Casula, which was used in Ref. [110]. It was stated in that paper that the interaction potential was accurate to $\mathcal{O}(10^{-6})$ Ha, which is consistent with the results of our tests with their code.

Hartree Fock energy [122] for the 1D HEG using the expression

$$E_{\mathrm{HF}} = \frac{1}{2}\sum_k \theta(k)\left[k^2 - \frac{1}{L}\sum_q \tilde{V}(q)\theta(k+q)\right] + \frac{V_{\mathrm{Mad}}}{2}\,, \tag{4.8}$$

where $\tilde{V}$ is the Fourier transform of the Ewald interaction and $\theta(k) = 1$ if k is the wave vector of an occupied state, and $\theta(k) = 0$ otherwise. Figure 4.2 shows that the results of HF and VMC with a Slater wave function are in statistical agreement for a wide range of densities.

4.3 Details of calculations

The positions of the wave function nodes are known for the 1D HEG, so DMC is, in principle, exact. For the infinitely-thin wire, nodes lie wherever two electrons coincide. For the harmonic wire, the nodes occur where parallel spins coincide.

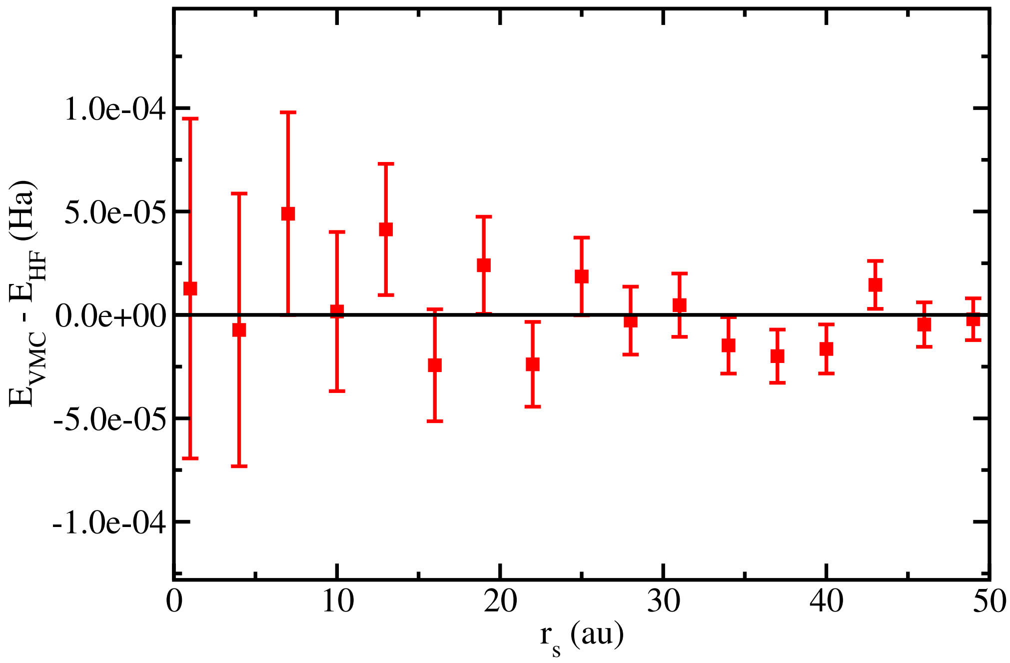

Figure 4.2: Plot of the difference between the VMC energy with a single Slater determinant wave function, and the Hartree Fock energy. The results shown are for $N = 37$.

We used a Slater-Jastrow-backflow wave function [16] for both systems, where the Jastrow function J is given by

$$J(\mathbf{R}) = \sum_{i \neq j}^{N} \left[\sum_{A=1}^{N_p} a_A \cos\left(\frac{2\pi A}{L} x_{ij} \right) + (x_{ij} - L_u)^3 \Theta(L_u - x_{ij}) \sum_{r=0}^{N_u} \alpha_r x_{ij}^r \right] , \qquad (4.9)$$

where $\{a_A\}$ and $\{\alpha_r\}$ are optimisable parameters, L_u is the cutoff, and N_p and N_u are chosen to achieve a compromise between speed, accuracy, and reliability. The use of very large numbers of parameters can make it difficult for the optimiser to find the appropriate minimum and risks introducing spurious features into the Jastrow function.

The orbitals in the Slater determinants were plane waves with wave vectors up to $k_F = \pi/(4r_s)$ for the paramagnetic systems and $k_F = \pi/(2r_s)$ for the ferromagnetic systems. The orbitals were evaluated at quasiparticle coordinates related to the actual coordinates by a backflow transformation [123]. Backflow provides an efficient way of describing three-body correlations in the 1D HEG, but leaves the exact nodal surface unchanged.

One method for assessing the wave function quality is to examine the fraction of the

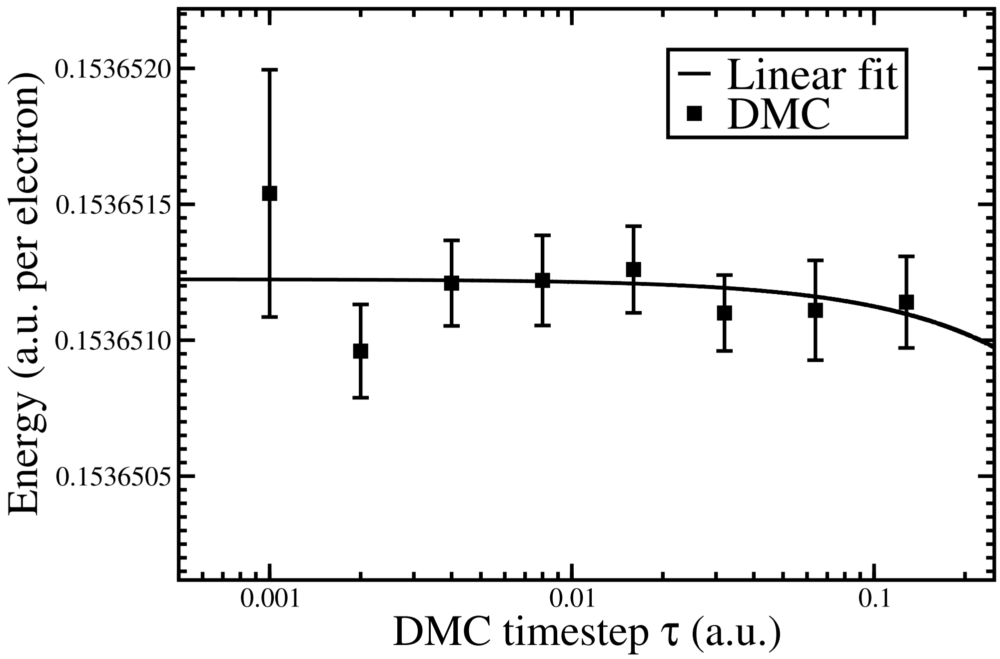

Figure 4.3: DMC energy of the infinitely-thin wire for several different timesteps and a linear fit to the data. The plot is for $r_s = 1$ a.u. and $N = 37$ with 1000 configurations.

correlation energy retrieved, $(E_{\rm HF} - E_{\rm VMC})/(E_{\rm HF} - E_{\rm DMC})$, where $E_{\rm HF}$ is the Hartree-Fock energy, and $E_{\rm DMC}$ and $E_{\rm VMC}$ are the DMC and VMC energies, respectively. We tested several types of wave function for the infinitely-thin wire with $r_s = 15$ a.u., $N = 15$, and $\zeta = 1$; our VMC calculations retrieved 99.9989(9)% of the correlation energy when we used a two-body Jastrow factor and backflow transformations [the error bars were $\mathcal{O}(10^{-8})$ a.u.], which is the type of wave function we use throughout this chapter. The inclusion of three-body terms in addition to backflow did not improve the wave function quality. Using a three-body term in the absence of backflow transformations allowed the recovery of the backflow result in most cases. However, backflow appears to provide a more robust description of three-body correlations for this system. While it is indeed the case that DMC is formally exact for the 1D HEG, the quality of the trial wave function is important for the statistical efficiency of the DMC method and the accuracy of the extrapolated estimator of Eq. (1.30).

The DMC energy did not change beyond statistical error upon varying the number of walkers between 640 and 2000, so we used ~ 1000 walkers in our calculations and assumed population control bias to be negligible. The dependence of the energy upon the DMC timestep τ was also investigated; Fig. 4.3 shows that for small τ the energy is constant. We performed our calculations at a single timestep given by $\tau = 0.008\, r_s^2$. This fairly conservative choice was made to ensure that time step bias is entirely negligible. The RMS distance diffused by each electron in a single step was thus slightly less than $r_s/10$.

For the infinitely-thin wire, we used simulation cells containing 37, 55, 73, and 99 particles subject to periodic boundary conditions for our calculations of the energy, PCF, and SSF. Our MD calculations for the infinitely-thin wire also used a much larger cell with $N = 255$, so that the grid of wave vectors on which the MD is defined was reasonably dense — this was important for the fitting procedure described in Sec. 4.4.6. For the harmonic wire, we used cells with $N = 123$, 155, and occasionally 255 for the $\zeta = 1$ systems and cells with $N = 22$ and 102 for the $\zeta = 0$ systems.

Previous work encountered difficulties in sampling different spin configurations of the harmonic wire for $\zeta \neq 1$ due to the presence of "pseudo-nodes" at the antiparallel coalescence points [112], although these problems were largely overcome by the use of lattice-regularised DMC (LRDMC) in Ref. [110]. The problem occurred because for strong, repulsive interactions the wave function becomes small when two antiparallel spins approach one another. Combined with a small time step this can lead to simulations where opposite spins exchange positions infrequently and the space of spin configurations is explored very inefficiently. Use of a small time step is a necessary part of the algorithm of projector methods like DMC. We have avoided ergodicity problems by using VMC to study the harmonic wire; as described in Sec. 1.3, in the VMC method there is no restriction other than ergodicity on the transition PDF and one may propose moves however one wishes provided that the acceptance probability is modified accordingly. We use electron-by-electron sampling with the transition PDF given by a Gaussian centred on the initial electron position. The VMC "time step" in fact bears no relation to real time and is simply the variance of the transition PDF. In practice, the unmodified time steps (chosen to achieve a 50% acceptance ratio) used in VMC are usually large enough to eliminate ergodicity problems in the 1D HEG, although we found some cases where it was necessary to enforce a lower limit on the width of the transition PDF. Table 4.1 shows the frequency with which electrons changed positions in our simulations for both high and low density systems with strong and weak confinement. Figure 4.4 shows the

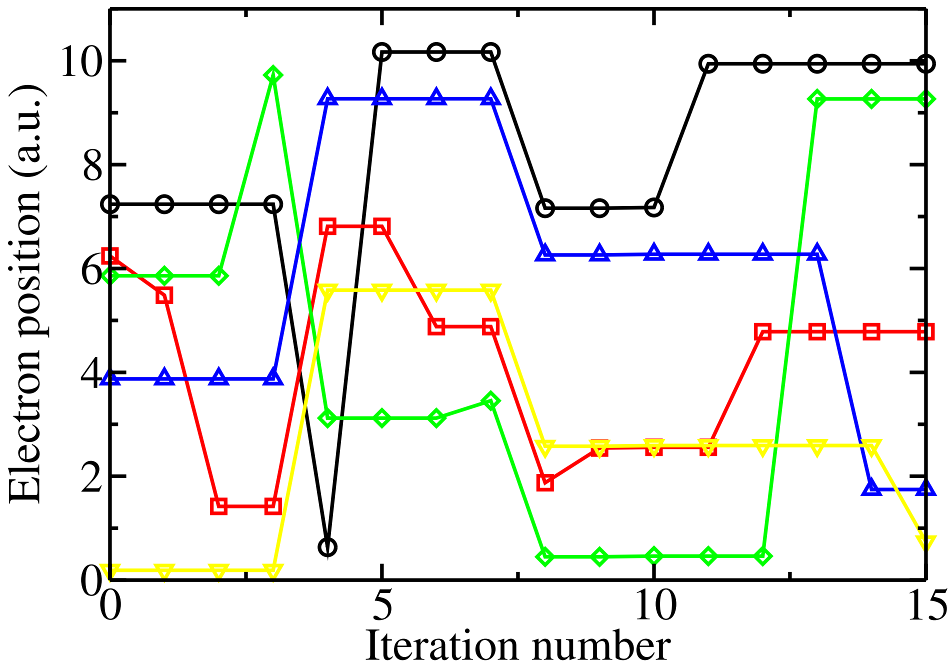

Figure 4.4: Plot of the electron positions during part of a test run to count the number of spin swaps. All of the electron positions are folded into a single simulation cell and the ordering is checked at every iteration (here defined as a proposed single-electron move). The figure shows a section of the run in which 10 iterations resulted in spin exchanges. The coordinates are relative to an arbitrarily-chosen electron, *i.e.*, there is an electron at the origin in the figure — at iterations 3 and 7 all electron positions change because the electron defining the origin has moved.

electron positions in a short portion of a run with a small system.

4.4 Results

4.4.1 Energy

For the infinitely-thin wire, we used DMC to calculate the exact ground state energy since there is no ergodicity problem. Table 4.2 shows the DMC energies obtained for $r_s = 1$, 2, 5, 10, 15, and 20 a.u., with $N = 37$, 55, 73, and 99 particles. We use the form $E(N) = E_\infty + AN^{-2}$, where A and E_∞ are fitting parameters, to extrapolate the energy $E(N)$ to the thermodynamic limit E_∞. Figure 4.5 demonstrates that this form fits the

r_s (a.u.)	b (a.u.)	$s_{\rm exch}$
1	0.1	0.051(1)
1	4	0.160(2)
15	0.1	0.0016(2)
15	4	0.0020(3)

Table 4.1: Frequency with which electrons' paths cross in VMC simulations of the harmonic wire. The quantity $s_{\rm exch}$ is the proportion of proposed single-electron moves that result in a change in the ordering of the particles. A typical calculation comprises between 10^7 and 10^8 proposed single-electron moves. The data shown are for $N = 22$.

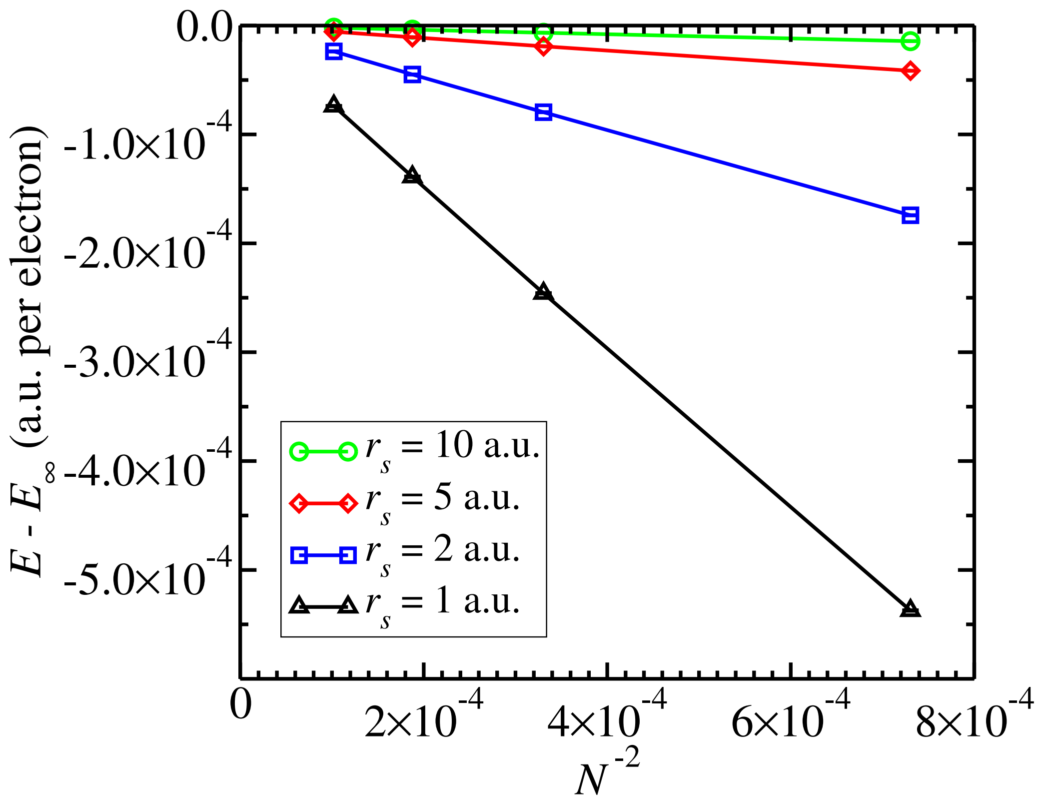

Figure 4.5: Plot of the DMC energy against the reciprocal of the square of the system size for the infinitely-thin wire. The energy has been offset by the extrapolate E_∞ obtained using the form $E(N) = E_\infty + BN^{-2}$.

r_s (a.u.)	N	E_{DMC} (a.u. / elec.)
1	37	0.1536513(3)
1	55	0.1539427(2)
1	73	0.1540497(3)
1	99	0.1541147(2)
2	37	−0.20637509(9)
2	55	−0.20628042(7)
2	73	−0.20624573(6)
2	99	−0.20622457(9)
5	37	−0.20397386(3)
5	55	−0.20395138(2)
5	73	−0.20394308(2)
5	99	−0.20393799(2)
10	37	−0.14288342(1)
10	55	−0.14287568(1)
10	73	−0.14287284(1)
10	99	−0.142871058(9)
15	37	−0.110474492(5)
15	55	−0.110470307(4)
15	73	−0.110468755(4)
15	99	−0.110467811(5)
20	37	−0.090782764(5)
20	55	−0.090780068(2)
20	73	−0.090779064(2)
20	99	−0.090778454(2)

Table 4.2: DMC energies for the infinitely-thin wire.

r_s (a.u.)	E_∞ (a.u. / elec.)
1	0.1541886(2)
2	−0.20620084(7)
5	−0.20393235(2)
10	−0.142869097(9)
15	−0.110466761(4)
20	−0.090777768(2)

Table 4.3: The DMC energies for the infinitely-thin wire extrapolated to the thermodynamic limit using the form $E(N) = E_\infty + BN^{-2}$.

r_s (a.u.)	E_{VMC} (a.u. / elec.)	E_{LRDMC} (a.u. / elec.)
1	0.0901489(7)	0.09014(1)
2	−0.1631207(8)	−0.16311(2)
10	−0.1231560(3)	−0.123157(3)
15	−0.0971194(1)	−0.097120(2)
20	−0.0807160(2)	−0.080717(1)

Table 4.4: Comparison of our VMC energies for the harmonic wire ($b = 1$ a.u., $\zeta = 1$) with those of Ref. [110], calculated using the LRDMC method. For both sets of results the energies were extrapolated to the thermodynamic limit using the functional form $E(N) = E_\infty + BN^{-1} + CN^{-2}$, where E_∞, B, and C are fitting parameters.

data well, and Table 4.3 shows the extrapolated energies E_∞. We discuss in Sec. 4.4.4 why this extrapolation function is suitable.

For the harmonic wire, the trial wave functions in our calculations are of sufficient quality that the variational energies obtained are in statistical agreement with exact results in the literature [110]; Table 4.4 shows the comparison.

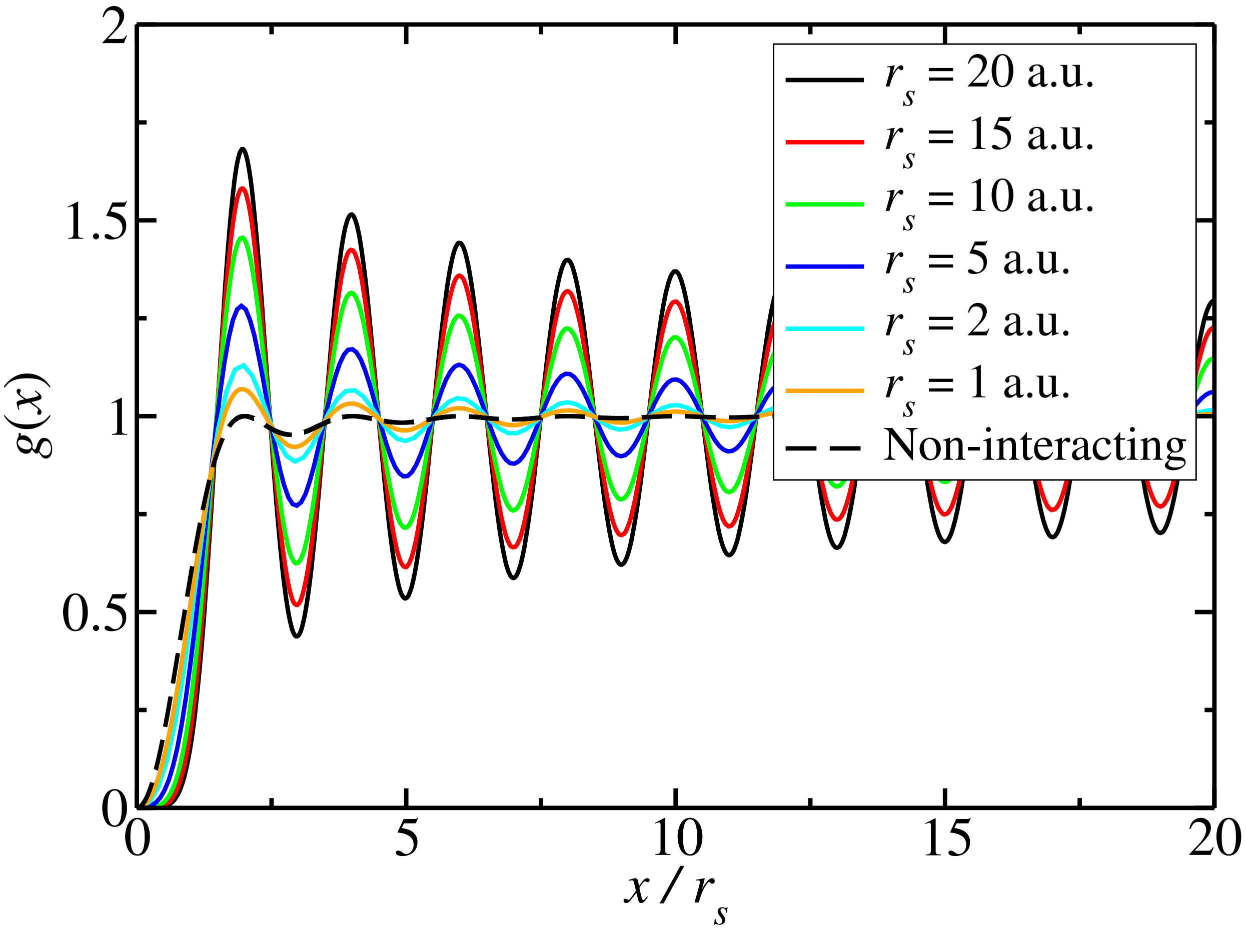

Figure 4.6: PCF of the infinitely-thin wire at five densities. The data shown are for $N = 99$ and are extrapolated estimates $[2g_{\rm DMC}(x) - g_{\rm VMC}(x)]$.

4.4.2 Pair-correlation function

The PCF $g(x)$ is the normalised probability of finding an electron at the position x given that there is one at the origin. For homogeneous systems like the 1D HEG, the PCF is a function of a single variable. The function contains a vast amount of information about the system; the *exchange-correlation hole*, which is the region around an electron in which the density is below average, may be directly accessed through the PCF, and the PCF also contains information about the phase of the system and the potential energy.

The PCF is accumulated in QMC simply by binning the interparticle distances throughout the simulation. The parallel-spin PCF is

$$g_{\uparrow\uparrow}(x) = \frac{1}{\rho_\uparrow^2}\left\langle \sum_{i>j}^{N_\uparrow} \delta(|x_{i,\uparrow} - x_{j,\uparrow}| - x) \right\rangle , \qquad (4.10)$$

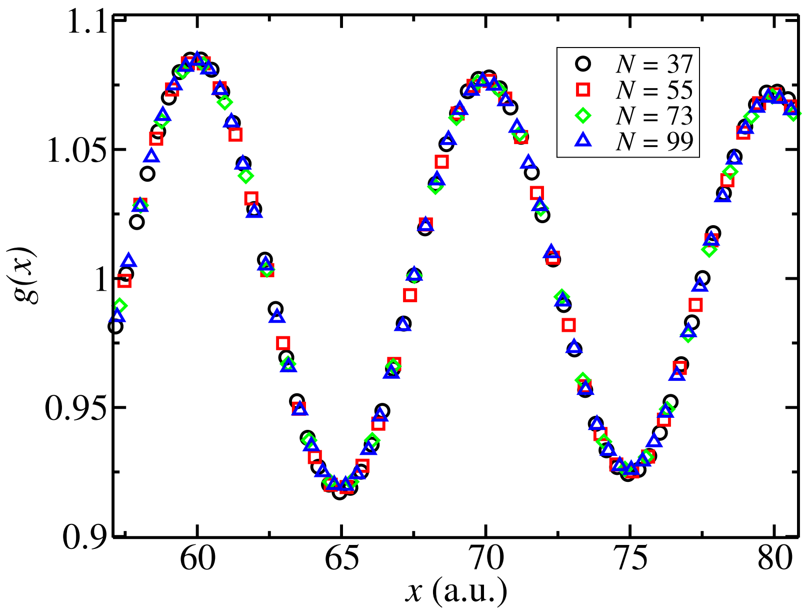

Figure 4.7: PCF of the infinitely-thin wire with $r_s = 5$ a.u. The data shown are extrapolated estimates $[2g_{\mathrm{DMC}}(x) - g_{\mathrm{VMC}}(x)]$.

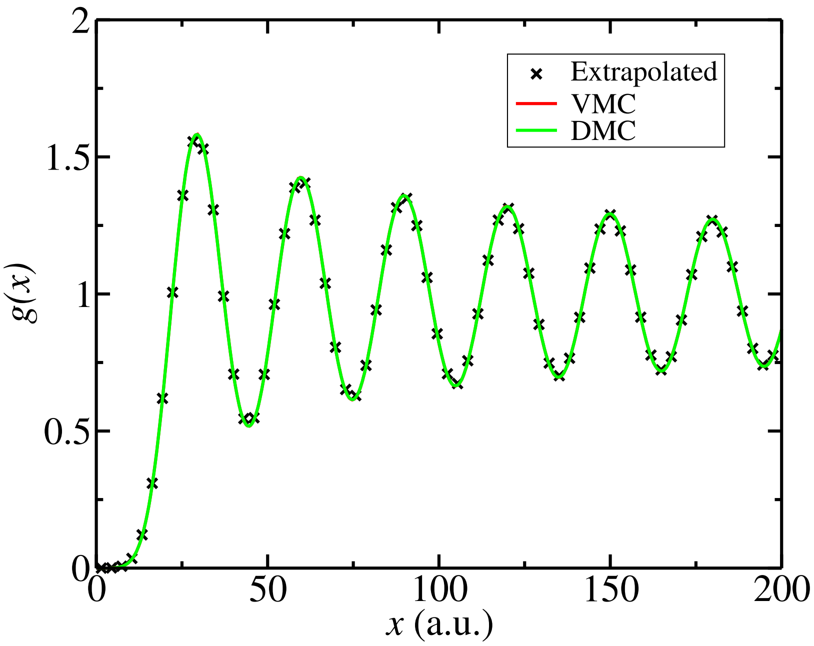

Figure 4.8: PCF of the infinitely-thin wire with $r_s = 15$ a.u. calculated with VMC and DMC, and the extrapolated estimate $[2g_{\mathrm{DMC}}(x) - g_{\mathrm{VMC}}(x)]$. The VMC and DMC data lie on top of one another.

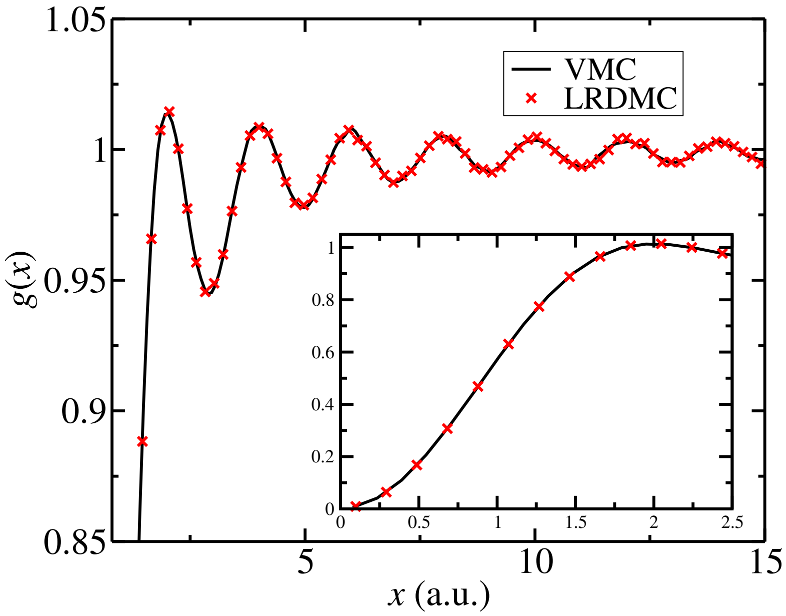

Figure 4.9: PCF of the harmonic wire with $r_s = 1$ a.u., $N = 39$, $b = 1$ a.u., and $\zeta = 1$. The solid line shows our VMC results and the symbols show the LRDMC results of Ref. [110]. The inset shows the same data at the origin.

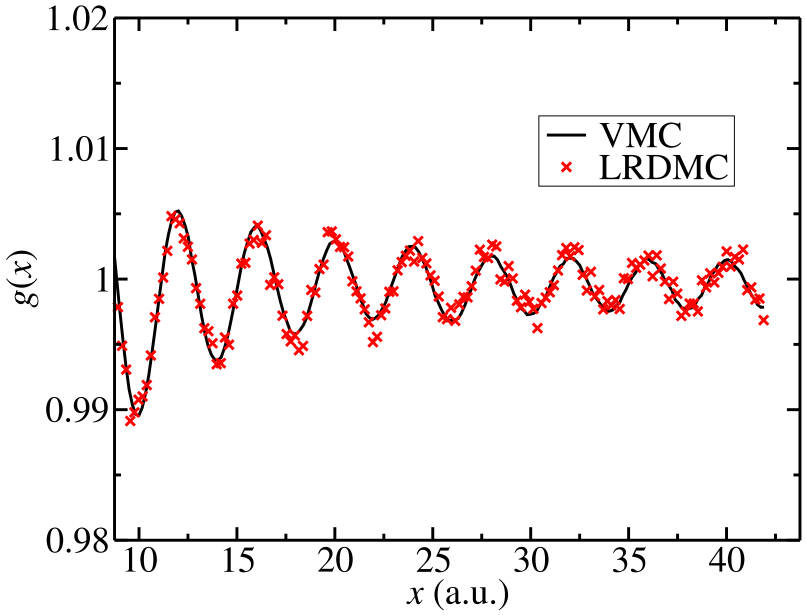

Figure 4.10: PCF of the harmonic wire with $r_s = 1$ a.u., $N = 42$, $b = 1$ a.u., and $\zeta = 0$. The solid line shows our VMC results and the symbols show the LRDMC results of Ref. [110]. The function plotted is $[g_{\uparrow\uparrow}(x) + g_{\uparrow\downarrow}(x)]/2$.

where ρ_σ is the average density of electrons with spin σ, $x_{i,\sigma}$ is the position of the ith electron with spin σ and the angular brackets denote an average over the configurations generated by the QMC algorithms. The antiparallel-spin PCF may be written as

$$g_{\uparrow\downarrow}(x) = \frac{1}{\rho_\uparrow \rho_\downarrow} \left\langle \sum_i^{N_\uparrow} \sum_j^{N_\downarrow} \delta(|x_{i,\uparrow} - x_{j,\downarrow}| - x) \right\rangle . \tag{4.11}$$

The PCF for the harmonic wire was calculated for different confinements and system sizes by Casula *et al.* using the lattice-regularised DMC method [110]. Figures 4.9 and 4.10 show the agreement of the present work with the LRDMC results. Figure 4.6 shows the PCF for the infinitely-thin wire at several values of r_s. Figures 4.7 and 4.8 show the convergence with respect to system size and the agreement of the PCF data between VMC and DMC, respectively.

4.4.3 Static structure factor

The SSF of the 1D HEG is defined as [72]

$$S(k) = 1 + \frac{N}{L} \int [g(x) - 1] e^{-ikx} \, dx \ , \tag{4.12}$$

and the SSFs that we present here are for the ferromagnetic infinitely-thin wire. As explained in the introduction, the antiferromagnetic and ferromagnetic phases are degenerate for the infinitely-thin wire, so we do not violate the Lieb-Mattis theorem with our choice of system.

The SSF, like the PCF, contains information about the phase of the system. Inserting Eq. (4.10) into Eq. (4.12) shows that the SSF is a measure of the average squared amplitude of density fluctuations with wave vector k [72]. The behaviour of the SSF at $k = 2k_\mathrm{F}$ is particularly interesting because this corresponds to fluctuations with period $2r_s$, which is the average inter-electron spacing [note that we use $k_\mathrm{F} = \pi/(2r_s)$ for the ferromagnetic infinitely-thin wire]. In the liquid phase, the SSF is expected to be roughly independent of system size at $k = 2k_\mathrm{F}$. However, for a Wigner crystal we would expect the SSF at $k = 2k_\mathrm{F}$ to scale linearly with the simulation cell length L as the ordered phase extends throughout the cell [111].

The PCF can only be directly measured in QMC for $x < L/2$ due to the finite extent of the simulation cell. Figure 4.11 shows the scaling of the SSF peak at $k = 2k_\mathrm{F}$ with system size. The height of the $k = 2k_\mathrm{F}$ peak in the finite-cell SSFs does not scale as N (and so L)

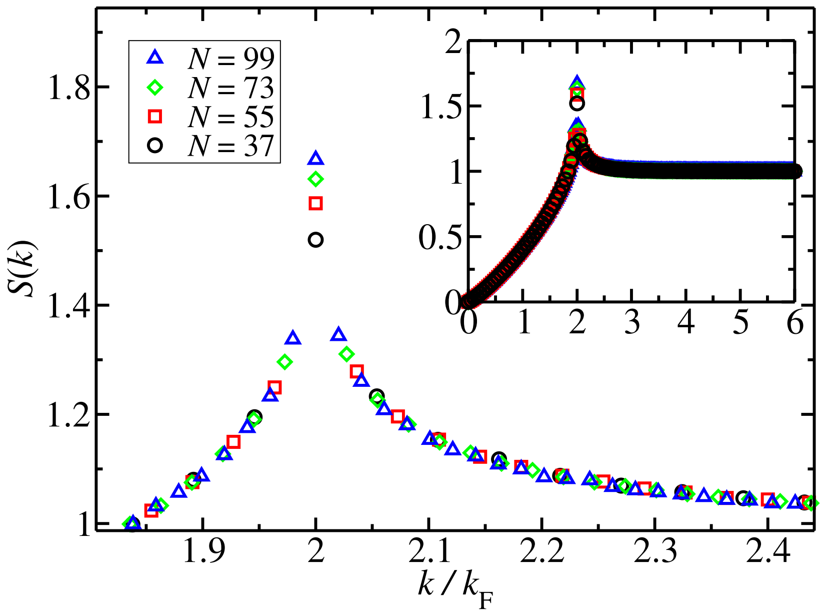

Figure 4.11: SSF of the infinitely-thin wire at four system sizes. The data shown are extrapolated estimates $[2S_{\text{DMC}}(k) - S_{\text{VMC}}(k)]$ for $r_s = 2$ a.u.. The main plot shows the behaviour at the peak and the inset shows a large-scale view. The PCF was not extended beyond $L/2$.

to any single power but appears to be sub-linear. This is consistent with the well-known picture of quasi long range order, which is the 1D analogue of a Wigner crystal. Quasi long range order describes phases where the oscillations in the charge-charge correlation function decay non-algebraically, *i.e.,* slower than any power law [124]. At k away from the peak the SSFs appear to agree very well for different cell sizes. Figure 4.12 shows the deviation of the $N = 37$, 55, and 73 SSFs from the $N = 99$ result for $r_s = 2$ a.u. The most significant deviations occur around the peak.

We further investigated finite-size effects by performing a fit to the oscillatory tails of the PCF and using the fitted function to extend the PCF far beyond $L/2$ before using Eq. (4.12) to calculate the SSF. After testing a number of functional forms, we found that a good-quality and simple fit to the oscillatory tails of the PCF is form [105, 110]

$$g(x) - 1 = A\cos(2k_{\text{F}}x)\exp(-B\sqrt{\ln x}) \;, \tag{4.13}$$

where A and B are treated as fitting parameters. The choice of Eq. (4.13) is motivated by the charge-charge correlation function of Ref. [105]. The parameters we obtained when fitting Eq. (4.13) to our results are given in Table A.3 in Appendix A.3.

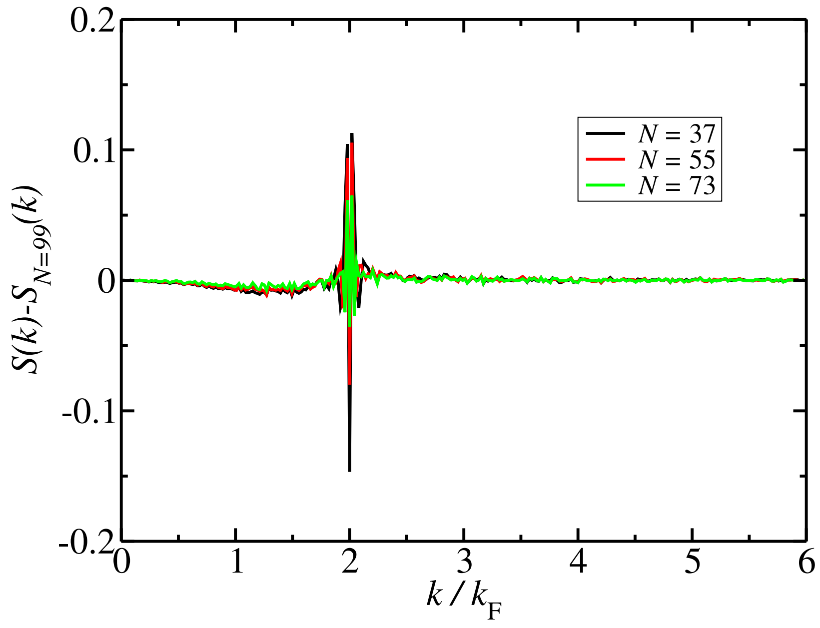

Figure 4.12: Plot of the difference between the SSFs at different system sizes and that at $N = 99$ for $r_s = 2$ a.u.

We fitted Eq. (4.13) to the PCF data for $6r_s < x < L/2 - 6r_s$, although we found that the results were not very sensitive to the region of data included in the fit. The data close to the origin were not included in the fit since Eq. (4.13) is only a good fit for long-range correlations. The data at the edge of the cell were excluded because that is the region midway between the electron at the origin and its next periodic image, and might be expected to be a region where the PCF suffers particularly badly from finite-size effects.

We then formed the extended PCF by reinstating all of the original PCF data up to $L/2 - 6r_s$ and appending a tail for $x > L/2 - 6r_s$ using Eq. (4.13) and the fitted parameters. Performing the Fourier transform of Eq. (4.12) numerically on the extended PCF results in a SSF (for $r_s \leq 20$ a.u.) with a greatly-enhanced peak at $k = 2k_{\rm F}$, but that agrees very well with the finite-cell SSF everywhere else. Figure 4.13 shows the difference between the SSFs obtained from the finite-cell and the extended PCFs. Under the extension scheme, the peak at $2k_{\rm F}$ appears to be susceptible to noise; in particular, the density of k-points at which the SSF is calculated heavily affects the apparent height. Figures 4.16 and 4.17 show that the fitting function of Eq. (4.13) possesses a peak at $k = 2k_{\rm F}$ in Fourier space and smoothly decays away elsewhere. The fact that extending

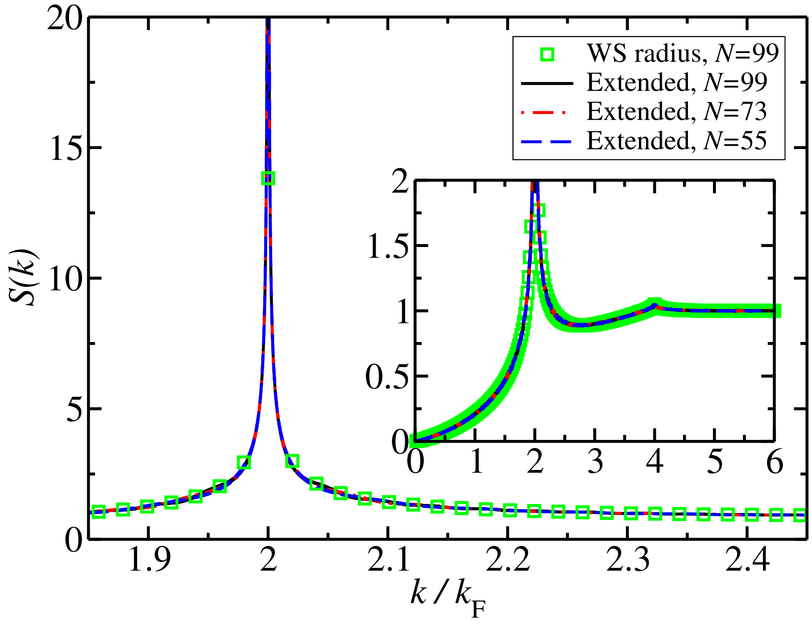

Figure 4.13: Effect of extending the PCF before performing the transformation of Eq. (4.12). The square symbols (labeled "WS radius") show the SSF obtained from the finite-cell PCF. The solid, dash-dot, and dashed lines (which all lie on top of one another) are from the $N = 99$, 73, and 55 PCFs, respectively, where in each case the PCF has been extended out to many simulation cell lengths using the fitting form of Eq. (4.13). The data shown are for $r_s = 20$ a.u.

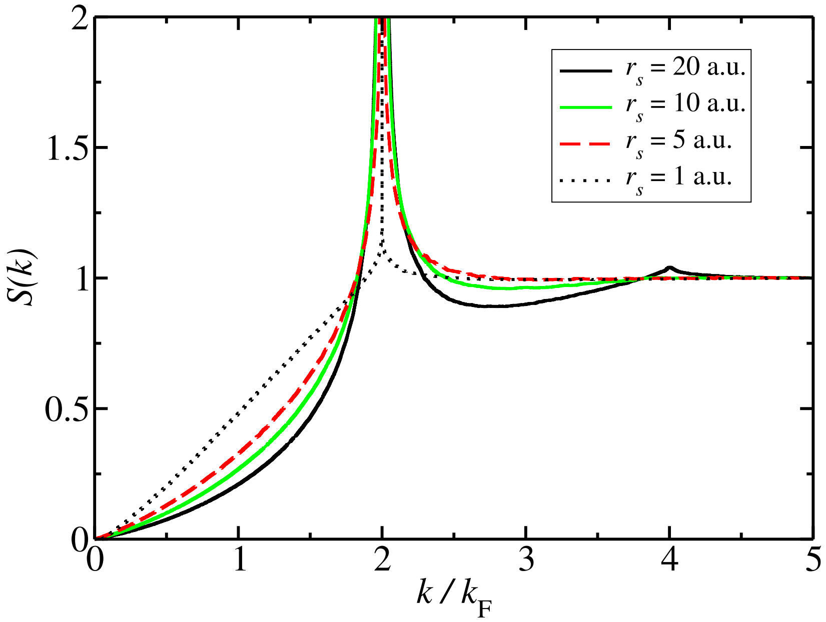

Figure 4.14: SSFs of infinitely thin wires obtained from Eq. (4.12) and the extended $N = 99$ PCFs.

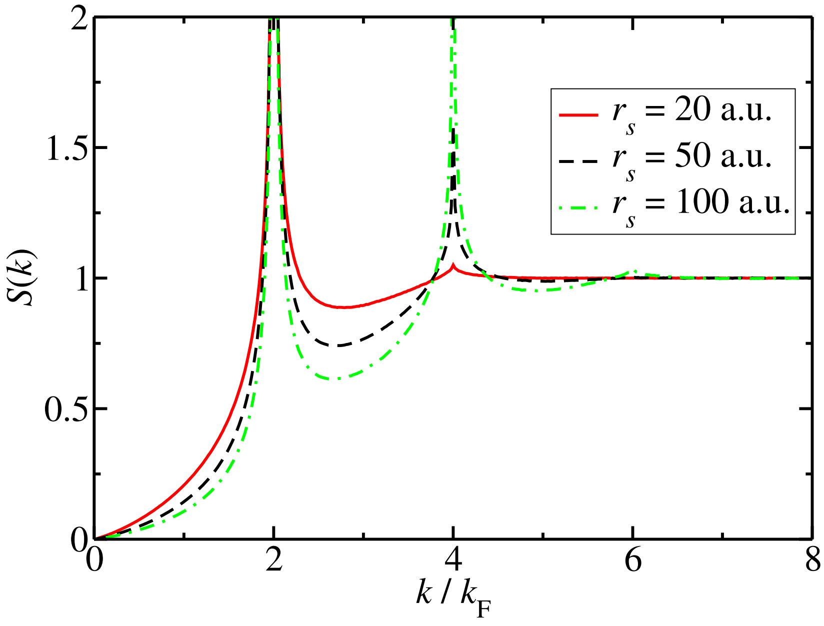

Figure 4.15: SSF of the infinitely-thin wire at very low densities. The results shown are for $N = 99$. The finite-cell VMC PCFs were used to generate the SSFs in the plot.

the PCFs using Eq. (4.13) has no effect other than to increase the size of the $k = 2k_{\mathrm{F}}$ peak in the SSFs, and the agreement between different system sizes, suggest that the electronic correlation is well-described in our calculations even for the smaller simulation cells that we have used.

The asymptotically-correct charge-charge correlation function of Schulz [105] that inspired Eq. (4.13) also includes higher order terms containing oscillations at wave numbers given by even multiples of k_{F}. For $r_s < 15$ a.u. we find no discernable features at larger k. However, a small feature at $4k_{\mathrm{F}}$ starts to develop at $r_s \approx 15$ a.u., and for $r_s = 20$ a.u. we observe a clear peak, visible in Figs. 4.13, 4.14, and 4.15. We performed short VMC calculations at extremely low densities, $r_s = 50$ a.u. and 100 a.u., where the electron-electron coupling is very large, to search for more noticeable features at $k > 2k_{\mathrm{F}}$. We find that peaks in the SSF do indeed appear at even multiples of k_{F} for these systems, as shown in Fig. 4.15. The SSF of the $r_s = 100$ a.u. system has clear peaks at $k = 2k_{\mathrm{F}}$, $4k_{\mathrm{F}}$, and $6k_{\mathrm{F}}$. This suggests that one could add higher-order terms to the fit of Eq. (4.13) for the low density systems and perform the extension scheme again, although this seems unlikely to produce any interesting new behaviour. The behaviour of the $r_s = 50$ a.u. and 100 a.u. systems is largely academic since these densities are currently out of experimental reach.

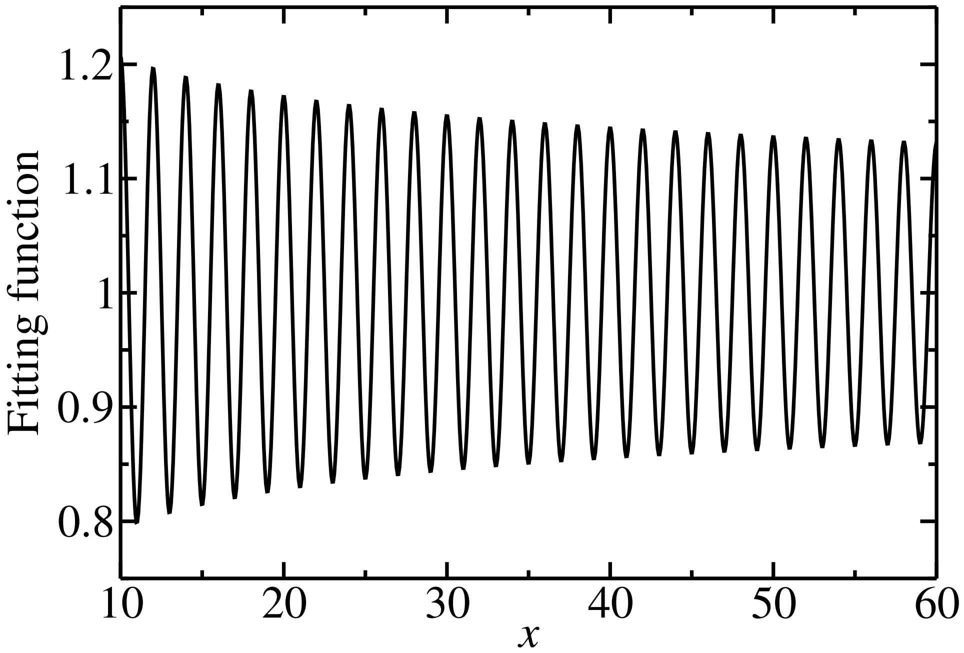

Figure 4.16: Plot of the fitting function of Eq. (4.13) for arbitrary parameter values.

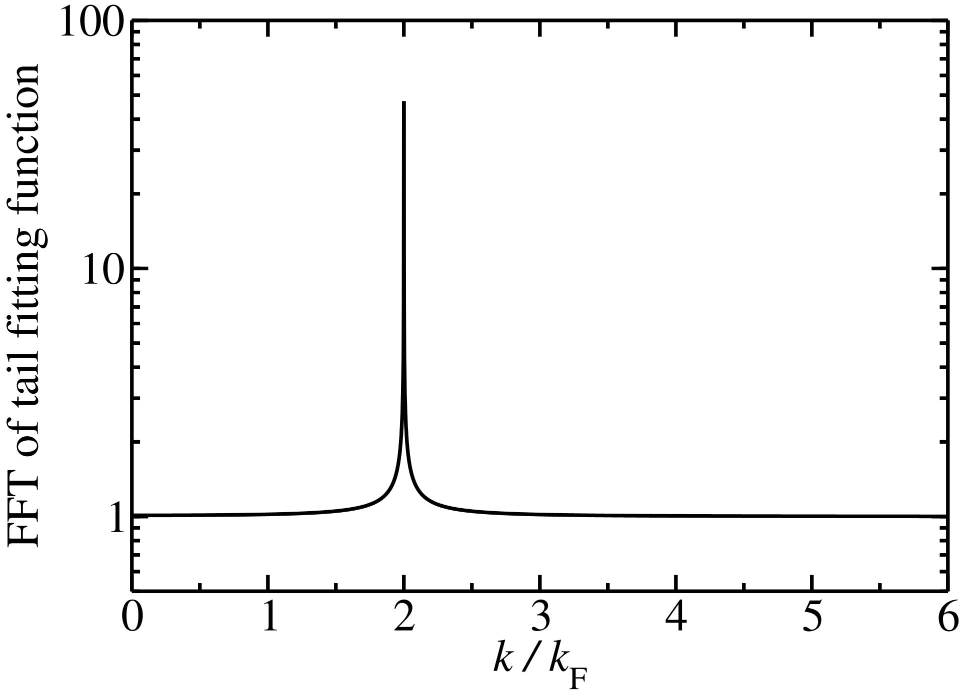

Figure 4.17: The function obtained from performing the Fourier transform of Eq. (4.12) on the fitting function of Eq. (4.13) for arbitrary parameter values. The function possesses a peak at $k = 2k_{\mathrm{F}}$ and smoothly decays everywhere else.

4.4.4 Finite-size effects

We have described our results for the energy, PCF, and SSF. We now make use of some of those findings to describe finite-size effects in the 1D HEG. We are interested in the properties of the 1D HEG in the thermodynamic limit, which we define as where $N \to \infty$ and $L \to \infty$ such that the density N/L is constant. Finite-size effects, such as compression of the exchange-correlation hole and shell-filling effects, mean that the properties calculated with a finite periodic simulation cell are nontrivially related to those at the infinite-system limit. We concern ourselves in this section only with the infinitely-thin wire, since the expectation values described above for the harmonic wire have already been thoroughly investigated in Ref. [110].

A technique often employed to reduce single-particle finite-size effects in QMC calculations is twist averaging [125]. We will explain why this is unhelpful for the infinitely-thin wire. The many-body Bloch theorem states that the wave function ψ_T satisfies [126]

$$\psi_T(x_1, \ldots, x_j + L, \ldots, x_N) = e^{ik_sL}\psi_T(x_1, \ldots, x_N) \ , \tag{4.14}$$

where k_s is the simulation cell Bloch wave number. Averaging over k_s in the irreducible Brillouin zone (BZ) has been shown to reduce greatly single-particle finite-size effects in two and three dimensions [125, 127, 128]. Figure 4.18 shows how nonzero k_s can alter the distribution of occupied single-particle states in 2D. The energy as a function of k_s in the first BZ is piecewise parabolic with discontinuities in the gradient occuring where the occupation numbers change.

In 1D, however, use of a nonzero k_s does not result in reoccupation of the orbitals because the arrangement of single-particle states cannot be changed by shifting the grid to the left or right. Applying a twist therefore changes the phase of the wave function but does not alter the distribution of configurations. The potential energy is thus independent of k_s. The kinetic energy, however, does depend on the phase.

Define $\psi(k_s)$ as the twist-averaged wave function, and ψ_0 as the wave function without

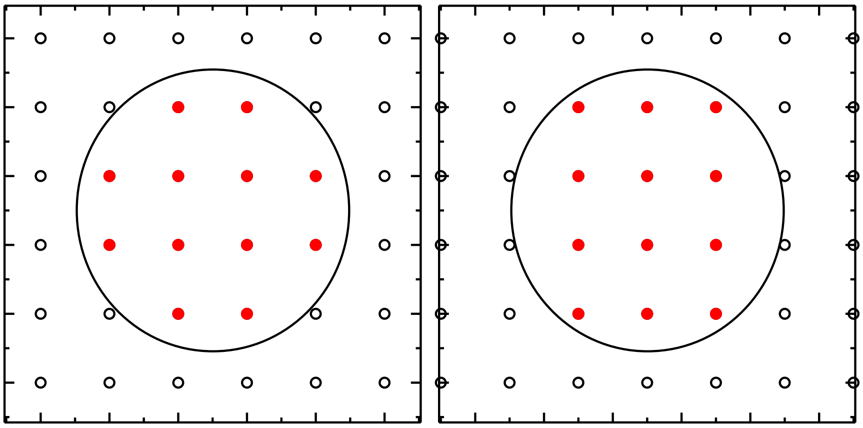

Figure 4.18: Schematic showing how the application of a twist can result in reoccupation of the orbitals in 2D. The filled and empty symbols represent occupied and empty states, respectively. The large circle shows the Fermi energy. In the left panel, no twist has been applied ($k_s = 0$). In the right panel, the grid of single-particle states has been offset and the distribution of occupied states has changed.

twist averaging, *i.e.*, $\psi(0) = \psi_0$. The kinetic energy estimator T is

$$\begin{aligned} T &= \left\langle -\sum_{j=1}^{N} \frac{\nabla_j^2 \psi(k_s)}{2\psi(k_s)} \right\rangle_{k_s} \\ &= \left\langle -\frac{1}{2} \sum_{j=1}^{N} \left(-k_s^2 + \frac{2ik_s e^{ik_s(x_1+\ldots+x_N)} \nabla_j \psi_0}{\psi(k_s)} + \frac{e^{ik_s(x_1+\ldots+x_N)} \nabla_j^2 \psi_0}{\psi(k_s)} \right) \right\rangle_{k_s} \\ &= \left\langle \frac{Nk_s^2}{2} \right\rangle_{k_s} + \left\langle -\frac{1}{2} \sum_{j=1}^{N} \frac{\nabla_j^2 \psi_0}{\psi_0} \right\rangle_{k_s} , \end{aligned} \tag{4.15}$$

where the angular brackets denote averaging over k_s in the first BZ. The last term on the third line of Eq. (4.15) is the kinetic energy estimator for the system with $k_s = 0$. It is easy to show that the extra term $\langle Nk_s^2/2 \rangle_{k_s}$ contributes $\pi^2/(24 r_s^2 N^2)$ to the energy per particle. One could thus simply add $\pi^2/(24 r_s^2 N^2)$ to the energy instead of performing twist averaging. However, as mentioned in Sec. 4.4.1, we extrapolate the energy $E(N)$ to the thermodynamic limit E_∞ using the form $E(N) = E_\infty + A/N^2$, where A and E_∞ are determined by the fitting procedure. The correction of Eq. (4.15) is thus unnecessary because it leaves the extrapolated energies unchanged and does not make extrapolation any easier. Figure 4.20 demonstrates that this analysis is correct; the DMC energy varies smoothly with the twist as described by Eq. (4.15).

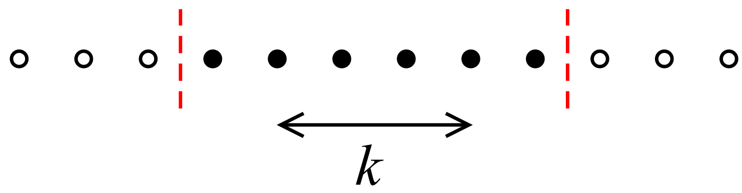

Figure 4.19: Illustration of the single-particle states in the first Brillouin zone (BZ). Filled circles represent occupied states, empy circles represent unoccupied states, and the dashed lines show the boundaries of the BZ. Moving the grid of states to the left or right, which corresponds to using a nonzero twist, does not alter the occupation of states.

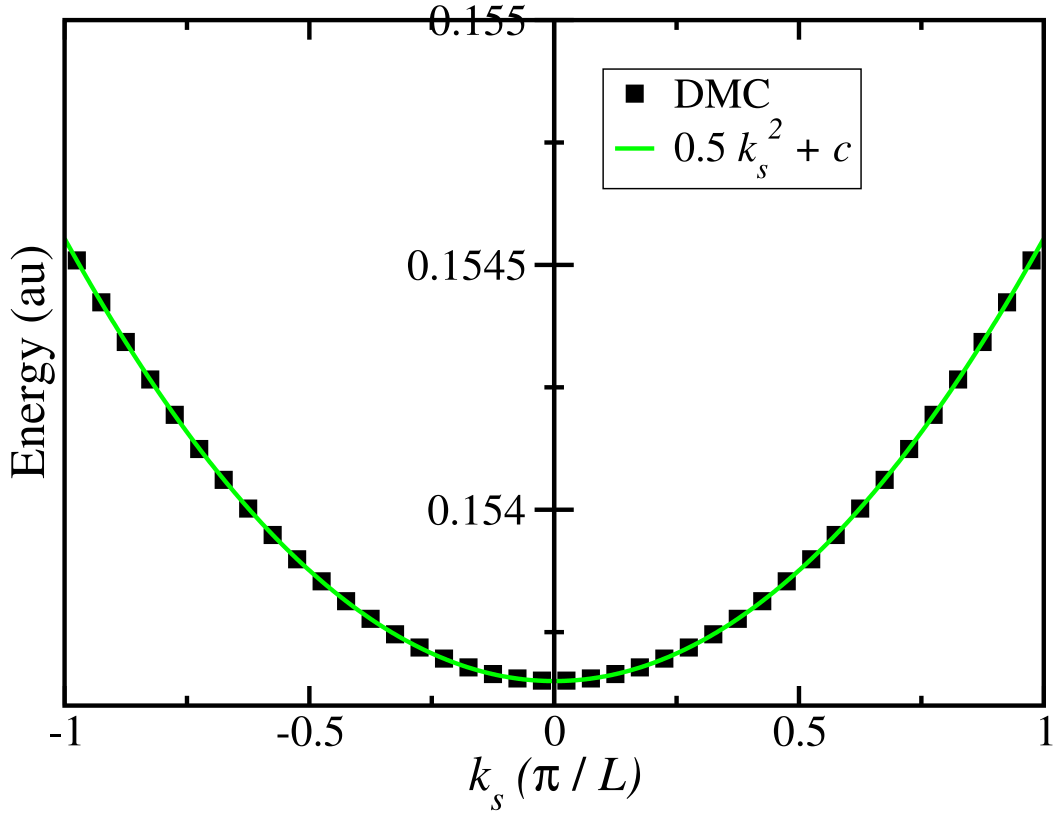

Figure 4.20: Plot of the variation of the energy with the twist angle. The smooth parabolic shape is due to the lack of orbital reoccupation. The solid line is the correction term of Eq. (4.15).

In order to see why $E(N) = E_\infty + A/N^2$ is a suitable function for extrapolation of the energy of the infinitely-thin wire to the thermodynamic limit, we reproduce the first part of the finite-size error analysis of Ref. [129], which was performed for 3D systems. We start with the Hamiltonian of Eq. (4.1). Since the interaction potential includes a sum over simulation cell lattice vectors, the Poisson summation formula [130] may be used to rewrite the interaction term as a sum over reciprocal lattice vectors, allowing us to restate the Hamiltonian as

$$\hat{\mathcal{H}} = -\frac{1}{2}\sum_{j=1}^{N} \frac{\partial^2}{\partial x_j^2} + \frac{1}{2L}\sum_{m\neq 0}^{\infty} \tilde{V}\left(k_m\right)\left[\rho_{k_m}\rho_{-k_m} - N\right] + \frac{N}{2}V_{\mathrm{Mad}} \ , \tag{4.16}$$

where $\tilde{V}$ is the Fourier transform of the interaction potential, $\rho_k = \sum_{j=1}^{N}\exp(ikx_j)$ is the density operator in Fourier space, and $k_m = 2\pi m/L$. The reader is directed to Ref. [72] for a full discussion of Eq. (4.16) — the $k_m = 0$ term has been dropped from the interaction potential as it as been cancelled with electron-background and background-background terms to yield a Hamiltonian that is well-defined in the thermodynamic limit. The Madelung term for the infinitely-thin wire is given by Eq. (4.3).

Inserting the definition of the PCF, Eq. (4.10), into the definition of the SSF, Eq. (4.12), gives

$$S(k) = \frac{\langle \rho_k \rho_{-k} \rangle}{N} \ , \tag{4.17}$$

where the angular brackets denote an average over configurations distributed according to the square of the ground state wave function. Equation (4.17) allows us to write the average potential energy per electron,

$$\left\langle \frac{1}{2NL}\sum_{m\neq 0} \tilde{V}\left(k_m\right)\left[\rho_{k_m}\rho_{-k_m} - N\right] \right\rangle = \frac{1}{2L}\sum_{m\neq 0} \tilde{V}\left(k_m\right)\left[S(k_m) - 1\right] \ . \tag{4.18}$$

As the simulation cell increases in size, the grid of $\{k_m\}$ becomes finer, eventually leading to the sum of Eq. (4.18) being well-approximated by an integral. The exact expression for the potential energy in the limit of infinite system-size is

$$V_\infty = \frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}k \ \tilde{V}(k)\left[S_\infty(k) - 1\right] \ , \tag{4.19}$$

where $S_\infty(k)$ is the exact SSF in the thermodynamic limit. The finite-size correction ΔV_∞ to the potential energy is given by the difference between Eqs. (4.19) and (4.18),

$$\Delta V_\infty = \frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}k \ \tilde{V}(k)\left[S_\infty(k) - 1\right] - \frac{1}{2L}\sum_{m\neq 0} \tilde{V}\left(k_m\right)\left[S(k_m) - 1\right] \ . \tag{4.20}$$

The observations of Ref. [129] may be summarised as follows; first, the -1 terms inside the square brackets are responsible for a large fraction of ΔV_∞. In fact, the contribution due to these terms is

$$-\frac{1}{4\pi}\int_{-\infty}^{\infty} \mathrm{d}k\ \tilde{V}(k) + \frac{1}{2L}\sum_{m\neq 0}\tilde{V}\left(k_m\right) = \frac{V_{\mathrm{Mad}}}{2}\ , \tag{4.21}$$

which is proved rigorously in Ref. [119]. Since we have explicitly calculated and incorporated the Madelung term into the Hamiltonian [see Eq. (4.1)], this contribution to the finite-size error is effectively already corrected for. We thus move on to the next largest source of error.

The second relevant point in Ref. [129] is that the next correction comes from one of two potential sources of error (whichever is larger), the first is the difference between $S_\infty(k)$ and $S(k)$, and the second is effectively an integration error; the contribution from the first Brillouin zone in the integral over k in Eq. (4.19) is not even approximately described by the sum of Eq. (4.18) since the $k = 0$ term is missing. Our results, shown in Fig. 4.21, indicate that the SSF around the origin converges quickly to the thermodynamic limit, so the difference $S_\infty(k) - S(k)$ is unlikely to be a significant source of error. We thus pursue the integration error, leading us to evaluate ΔV, where $\Delta V_\infty = V_{\mathrm{Mad}}/2 + \Delta V +$(higher order terms), giving the correction

$$\Delta V \propto \int_{-\pi/L}^{\pi/L} \mathrm{d}k\ \tilde{V}(k)S(k)\ . \tag{4.22}$$

To calculate the integral of Eq. (4.22) for the infinitely-thin 1D HEG, we make use of the (empirical) form of $S(k)$ at small k, and the explicit form of $\tilde{V}(k)$,

$$\lim_{k\to 0}\{S(k)\} \quad \propto \quad |k|\ , \tag{4.23}$$

$$\tilde{V}(k) \quad \approx \quad -\log\left(\frac{k^2}{4}\right) - 2\gamma\ , \tag{4.24}$$

where γ is the Euler-Mascheroni constant. The linear form of $S(k)$ for small k was observed in our QMC calculations for all of the densities and system sizes studied. Equation (4.24) is derived from the infinitely-thin wire Ewald interaction, described in Ref. [118].

Inserting Eqs. (4.23) and (4.24) into Eq. (4.22), we find

$$\begin{aligned} \Delta V \quad &\propto \quad \int_{-\pi/L}^{\pi/L} \mathrm{d}k\ \left[-\log\left(k^2/4\right) - 2\gamma\right]|k|\ , \\ \Delta V \quad &\propto \quad L^{-2}\left[A + \log\left(L^{-2}\right)\right]\ , \end{aligned} \tag{4.25}$$

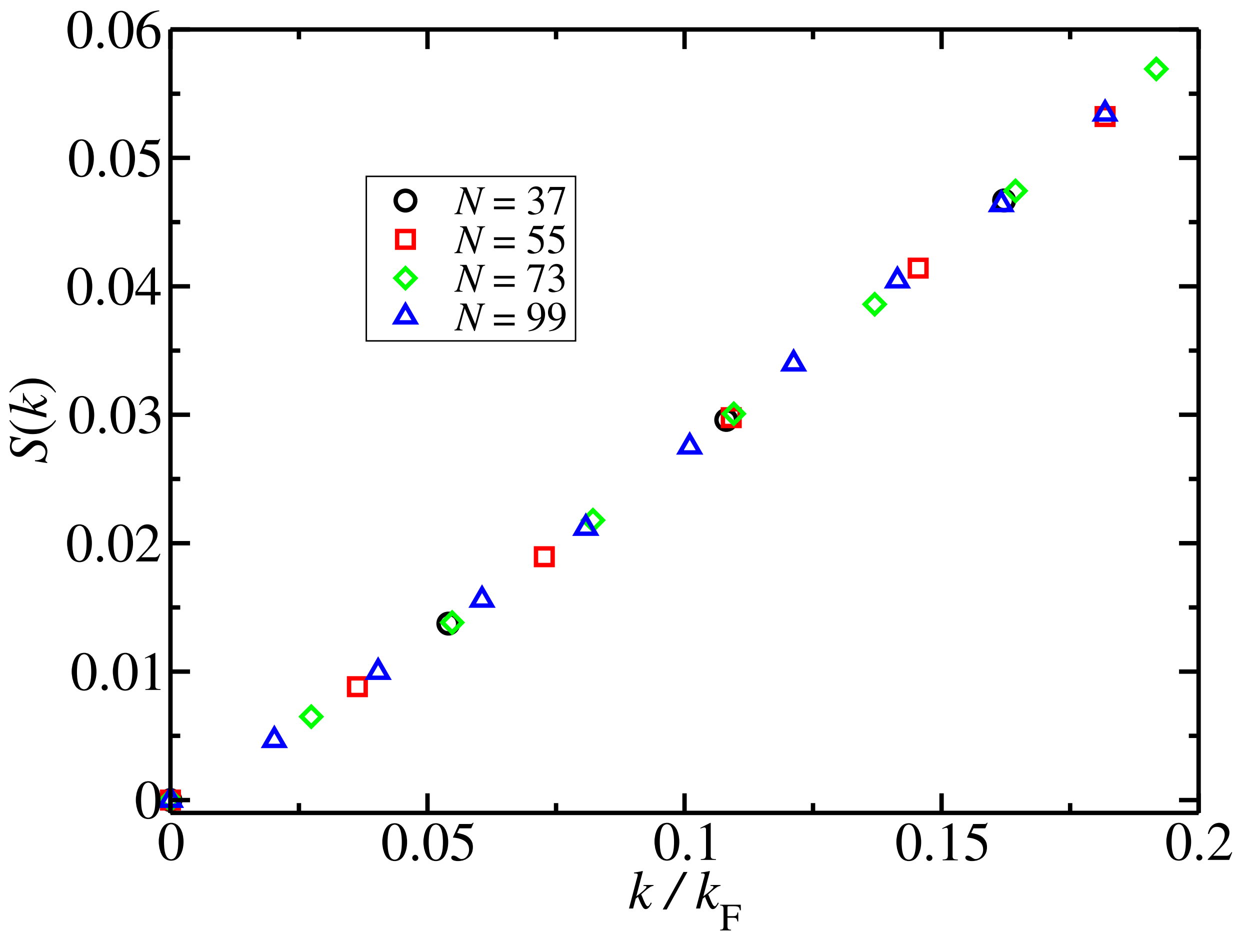

Figure 4.21: Plot of the small k behaviour of the SSF for $r_s = 2$ a.u. The SSFs converge very quickly to the thermodynamic limit and are linear at small k.

where A is a constant. We neglect the log term in Eq. (4.25), so that the correction becomes $\Delta V \propto L^{-2}$.

One can employ a similar approach for the kinetic energy, again expressing the correction as the difference between a sum and an integral over k. The correction is then [129]

$$\Delta T \propto \int_0^{\pi/L} \mathrm{d}k \; k^2 \tilde{u}(k) \;, \tag{4.26}$$

where $\tilde{u}(k)$ is the Fourier transform of the two-body Jastrow factor. We find the small-k behaviour of $\tilde{u}(k)$ by performing a Fourier transform on a fully-optimised two-body Jastrow factor from our calculations and directly observing the result. Figure 4.22 shows a respresentative example; our results are closely-consistent with

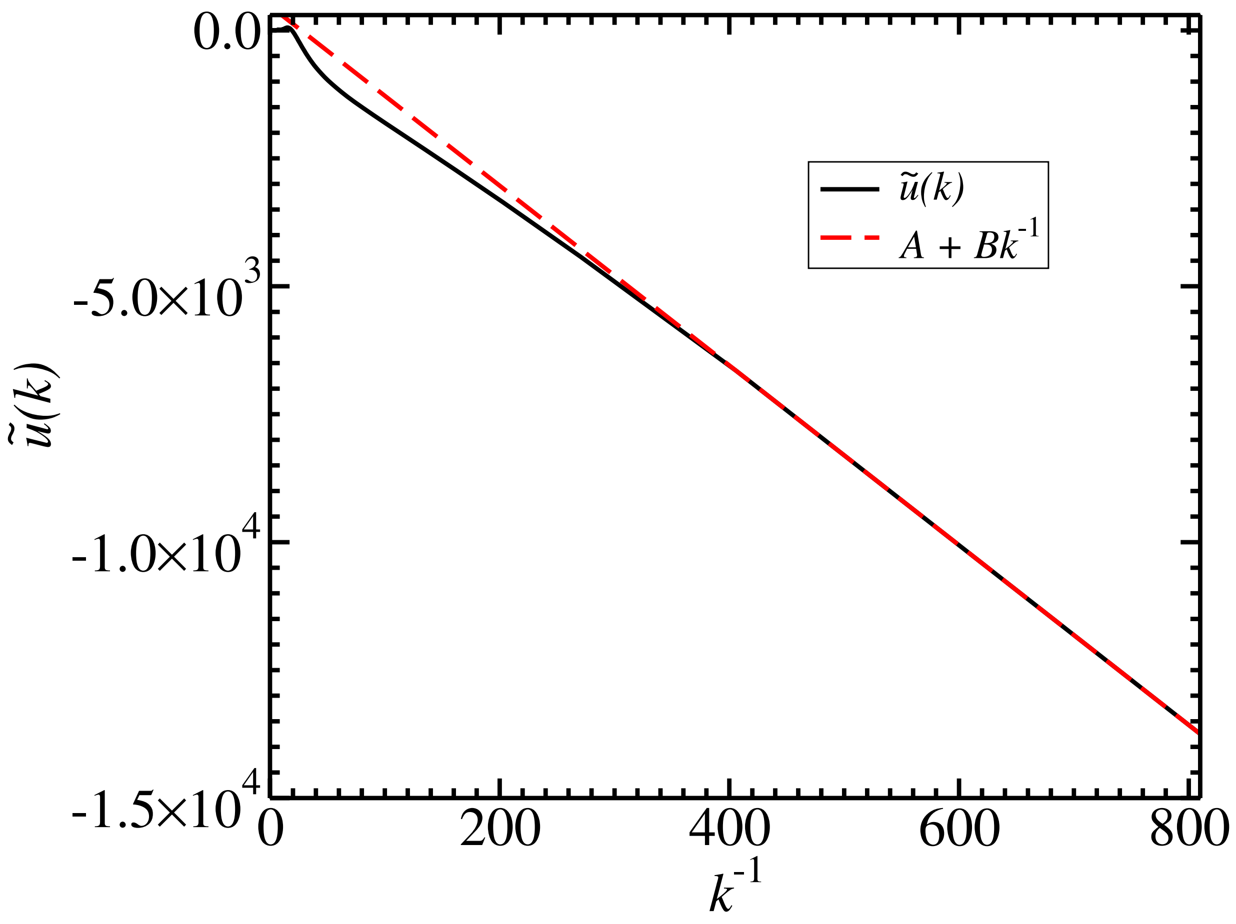

Figure 4.22: Plot of the Fourier transform of the optimised two-body Jastrow factor from a calculation on the infinitely-thin wire with $r_s = 20$ a.u. and $N = 99$. The dashed line shows a function of the form $\tilde{u}(k) = A + B/k$, where A and B are constants, to guide the eye.

$$\lim_{k\to 0}\{\tilde{u}(k)\} \propto k^{-1} \;, \tag{4.27}$$

leading very simply to the result

$$\Delta T \propto L^{-2} \; . \tag{4.28}$$

Equations (4.25) and (4.28), and the fact that $L = 2r_sN$, show that the leading-order finite-size errors in both the potential and kinetic energies go as N^{-2}. We thus conclude that the function

$$E(N) = E_\infty + \alpha N^{-2} \;, \tag{4.29}$$

where α is a fitting parameter, provides a suitable form for extrapolation of the energy $E(N)$ to the thermodynamic limit E_∞. The plot of Fig. (4.5) shows that our DMC energies for the infinitely-thin wire fit very well to Eq. (4.29).

4.4.5 Momentum density

The MD is accumulated in QMC as

$$n(k) = \left\langle \frac{1}{2\pi} \int \frac{\psi(r)}{\psi(x_1)} \exp[ik(x_1 - r)] \, dr \right\rangle \;, \tag{4.30}$$

where $\psi(r)$ is the trial wave function evaluated at $(r, x_2, \ldots, x_n)$ and angular brackets denote an average over configurations. The MD is the integral of the spectral function from minus infinity up to the chemical potential [72]. The spectral function is the probability density for changing the energy by an amount between ϵ and $\epsilon + \mathrm{d}\epsilon$ by adding or removing a particle in state α. For our purposes, the state α is the wave vector of a plane wave state. For example, in the non-interacting case the spectral function is $\delta(\epsilon - k^2/2)$, since injecting an electron with wave vector k leaves the system in a well-defined state, and equates to simply occupying an additional single-particle state. The MD exhibits a drop at $k = k_\mathrm{F}$ because that is where the peak in the spectral function reaches the chemical potential. If the peak in the spectral function is a δ-function at $k = k_\mathrm{F}$ (*i.e.,* the spectral function possesses a quasiparticle peak) then the MD is discontinuous at the Fermi edge. However, in 1D we expect the excitations to be collective rather than single-particle-like. The 1D systems should thus have MDs that are continuous at k_F, although TL liquid theory predicts that the gradient will be singular [105].

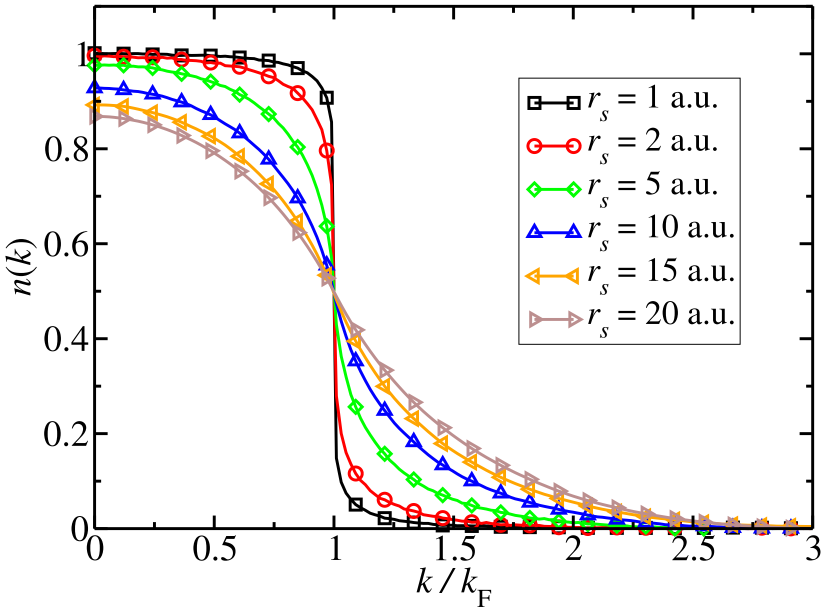

Figure 4.23: MD of the infinitely-thin wire at several densities. The data shown are for $N = 99$ and are extrapolated estimates $[2n_{\rm DMC}(k) - n_{\rm VMC}(k)]$. The statistical error bars are much smaller than the symbols and some symbols have been omitted for clarity.

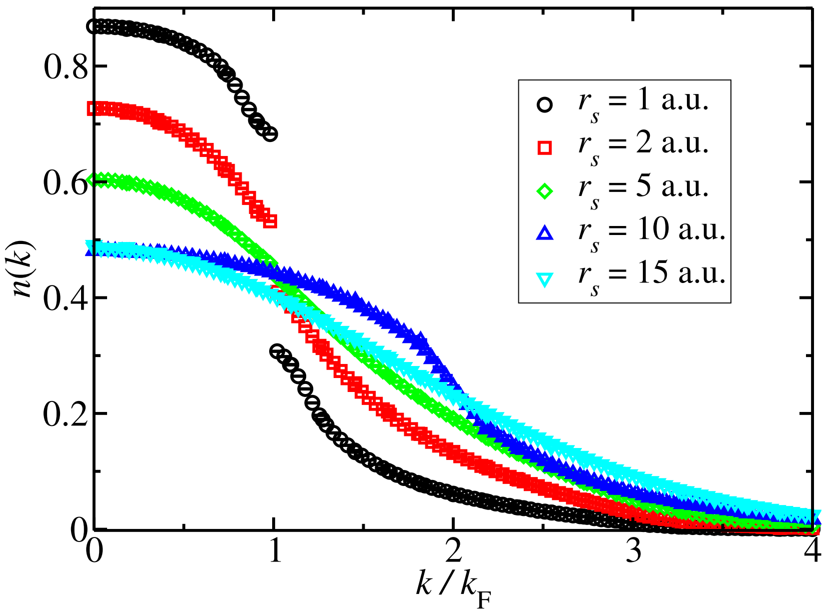

Figure 4.24: VMC MD of the harmonic wire with $b = 0.1$ a.u. and $\zeta = 0$ at several densities. The data shown for each density are for $N = 22$ and 102 (joined to form one data set). The statistical error bars are smaller than the symbols.

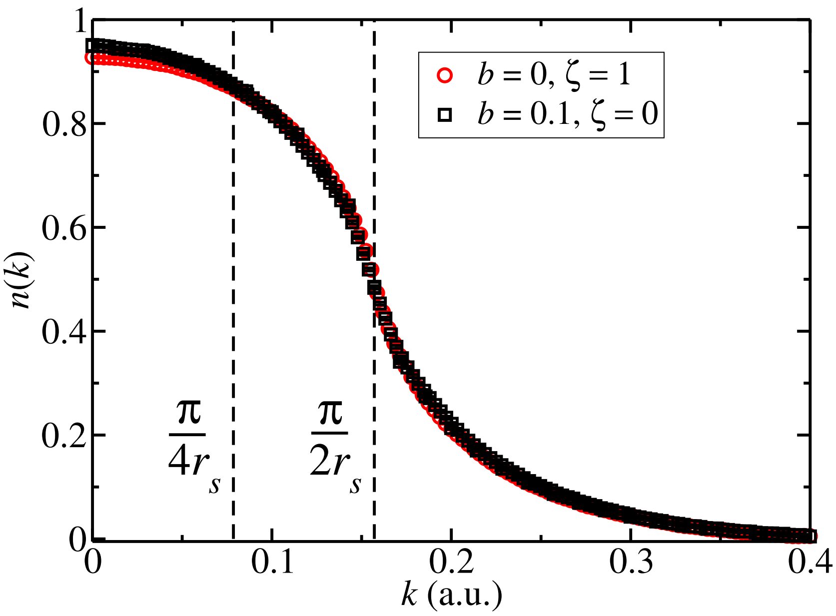

Figure 4.25: Comparison of the $r_s = 10$ a.u. MD for the infinitely-thin wire ($b = 0$, $\zeta = 1$) with that obtained for the harmonic wire with $b = 0.1$ a.u. and $\zeta = 0$. The statistical error bars are similar in size to the symbols. The dashed lines show the values of k at which Eq. (4.31) was fitted to the data for the calculations of the exponent α.

For the systems with $\zeta = 1$, we have used $k_{\rm F} = \pi/(2r_s)$, whereas for the systems with $\zeta = 0$, we have used $k_{\rm F} = \pi/(4r_s)$. Figure 4.23 shows the MDs obtained by evaluating the extrapolated estimator $2n_{\rm DMC}(k) - n_{\rm VMC}(k)$ for the infinitely-thin wire. The VMC and DMC results differed by a maximum of ≈ 2 error bars, so that evaluating the extrapolated estimator changed the results very little. Note that evaluating the extrapolated estimator of Eq. (1.30) does not remove the linear error in the MD in the same way as for other expectation values, since Eq. (4.30) includes wave function ratios. We still use the extrapolated estimator as a measure of the error, however, since the degree of agreement between VMC and DMC results is an indicator of accuracy. Figure 4.24 shows the MD for the harmonic wire with $b = 0.1$ a.u. and $\zeta = 0$.

A particularly interesting feature of the paramagnetic harmonic wire MD is that as r_s is increased and b is decreased, much of the weight of the function shifts to larger k, and $n(0)$ reduces to values around 0.5. This is a direct manifestation of the harmonic wire becoming more like the ferromagnetic infinitely-thin system. One can in some cases see a feature resembling the gradient discontinuity appearing at $\pi/(2r_s)$, *i.e.,* at twice the paramagnetic Fermi wave vector. In particular, for $r_s = 10$ a.u. and $b = 0.1$ a.u. the MD possesses a feature at $\pi/(2r_s)$. Upon closer inspection we find that the MD for the unpolarized system with $b = 0.1$ a.u. agrees very well with that of the infinitely-thin wire ($b = 0$ and $\zeta = 1$). Figure 4.25 illustrates this comparison. It thus appears possible to in some sense tune the effective Fermi wave vector by adjusting the strength of the confinement (and the density). A dense paramagnetic system with very weak confinement shows significant occupation of momentum states up to approximately $\pi/(4r_s)$. Increasing the effective interaction strength increases this value of k until it eventually saturates at the ferromagnetic k_F. This reflects the fact that in the limit $r_s \to \infty$ the pseudo-nodes at antiparallel-spin coalescence points become true nodes.

4.4.6 Tomonaga-Luttinger liquid parameters

Close to the Fermi wave vector, TL liquid theory suggests that the MD should take the form [74, 131]

$$n(k) = n(k_{\rm F}) + A[\text{sign}(k - k_{\rm F})]|k - k_{\rm F}|^{\alpha} \,, \tag{4.31}$$

which we have fitted to our results treating $n(k_{\rm F})$, A, and α as fitting parameters. Note that within TL liquid theory the exponent α is related to the TL liquid parameter [132] K_ρ by

$$\alpha = \frac{1}{4}\left(K_\rho + \frac{1}{K_\rho} - 2\right) \,. \tag{4.32}$$

If the range of data included in the fit is described by $|k - k_{\rm F}| < \varepsilon k_{\rm F}$, the choice of ε can present some difficulties. Ideally, one would choose $\varepsilon \to 0$ since Eq. (4.31) is valid for $k \to k_{\rm F}$, and indeed using the entire range of MD results yields rather poor fits. However, the estimate of α becomes noisy when ε is small, and at the extreme where just two data points are included, one can of course obtain any value for α. This leads us to include fits constructed using a larger range of k values. In practice, we chose to perform a linear extrapolation to $\varepsilon = 0$ excluding fits where $\varepsilon < 0.05$ for the $\zeta = 1$ systems, as shown in Fig. 4.26. For the $\zeta = 0$ systems, we excluded fits for which $\varepsilon \lesssim 0.25$. The trend that we observe in the exponent with respect to ε is similar to that found in Ref. [133].

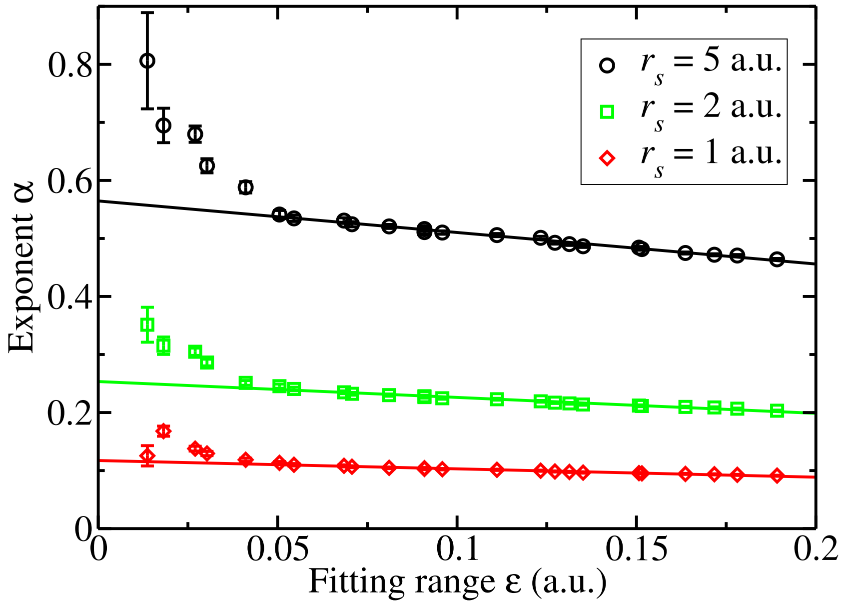

Figure 4.26: Exponent α [found from fitting Eq. (4.31) to the MD] against the range of data included in the fit. The range of data is described by $|k - k_{\rm F}| < \varepsilon k_{\rm F}$. The symbols show the fitted exponent values and the solid lines are linear fits to the exponents in the region $\varepsilon > 0.05$. The data shown are for the infinitely-thin wire.

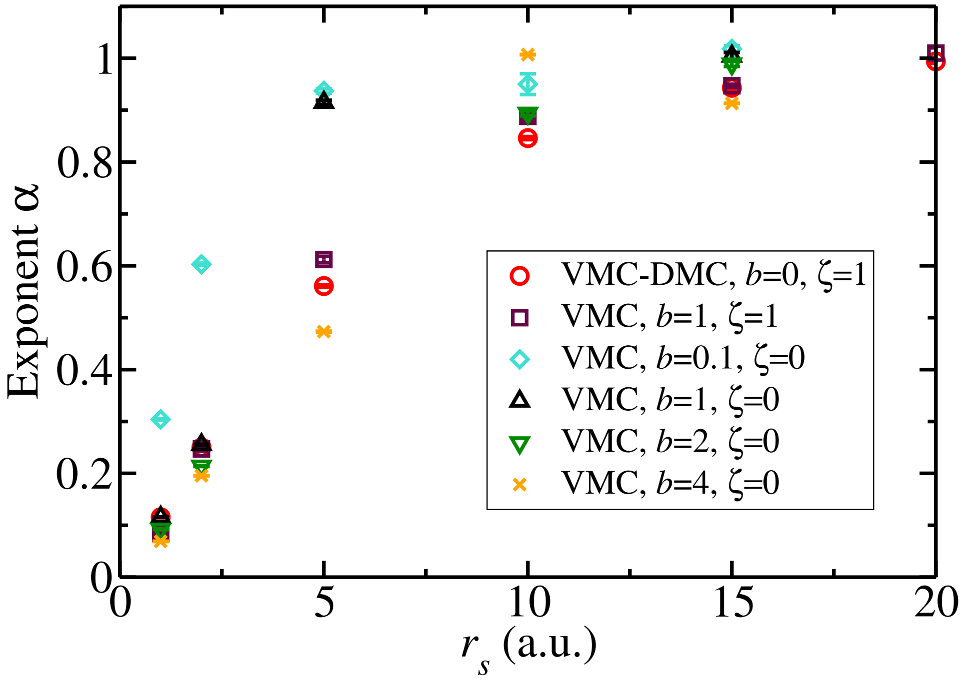

Figure 4.27: Exponent α found from fitting Eq. (4.31) to the MD around $k = \pi/(2r_s)$ for the $\zeta = 1$ systems and $k = \pi/(4r_s)$ for the $\zeta = 0$ systems.

Figure 4.27 shows the exponents α obtained for several densities, polarisations, and confinements. All of the systems show the same general trend; α tends to 0 in the high-density limit and to 1 in the low-density limit. As mentioned earlier, it is important to note that for the systems with $\zeta = 1$ we fitted Eq. (4.31) to the MD at $\pi/(2r_s)$, whereas for $\zeta = 0$ we used $\pi/(4r_s)$. The change in shape of the MD upon varying the interaction strength that we noted in Sec. 4.4.5, and the apparent shift in the 'effective' k_{F}, suggest that one could also extract a relevant exponent from fitting to other values of k. For example, we showed in Fig. 4.25 the similarity between the MD with $r_s = 10$ a.u., $b = 0.1$ a.u., and $\zeta = 0$, and that with $r_s = 10$ a.u., $b = 0$, and $\zeta = 1$. Despite the similarity of the MDs for the two systems, the fits used to extract α from each system were performed at different values of k — a factor of two apart in fact. The result is that the exponent for the paramagnetic wire is larger, since the Fermi edge for that system has apparently shifted to k above the fitting region. There is little to guide one in choosing a different region in which Eq. (4.31) may be fitted to the data, so this appears to be a flaw in the method for the $\zeta \neq 1$ systems. Presumably, the exponent measured in experiments is that at the effective Fermi wave vector, so we believe that our results for the $\zeta = 1$ systems are more representative of physical systems than those of the $\zeta = 0$ systems.

With this in mind, Fig. 4.28 shows the $\zeta = 1$ results alone, since for the ferromagnetic systems one can clearly and reliably state that $k_{\mathrm{F}} = \pi/(2r_s)$ for the whole range of densities. The exponent α for the infinitely-thin wire is reasonably well-approximated by the function

$$\alpha = \tanh(r_s/8) \ , \tag{4.33}$$

which gives a maximum deviation of 0.011(3) from the $b = 0$ QMC results, which occurs at $r_s = 15$ a.u. The exponents for the harmonic wire with $b = 1$ a.u. and $\zeta = 1$ show a maximum deviation from Eq. (4.33) of 0.057(6), which we find at $r_s = 5$ a.u.

The exponent α has been reported in previous theoretical and experimental studies. Reference [121] gives the exponents for $b = 0.1$, 1, and 4 a.u. (with $r_s = 1$ a.u. and $\zeta = 0$) from VMC calculations. In Fig. 4.29 we have shown how the results given there compare with ours. It appears that the principal difference between the two studies is the procedure for deciding upon a fitting region; Ref. [121] does not give details of any

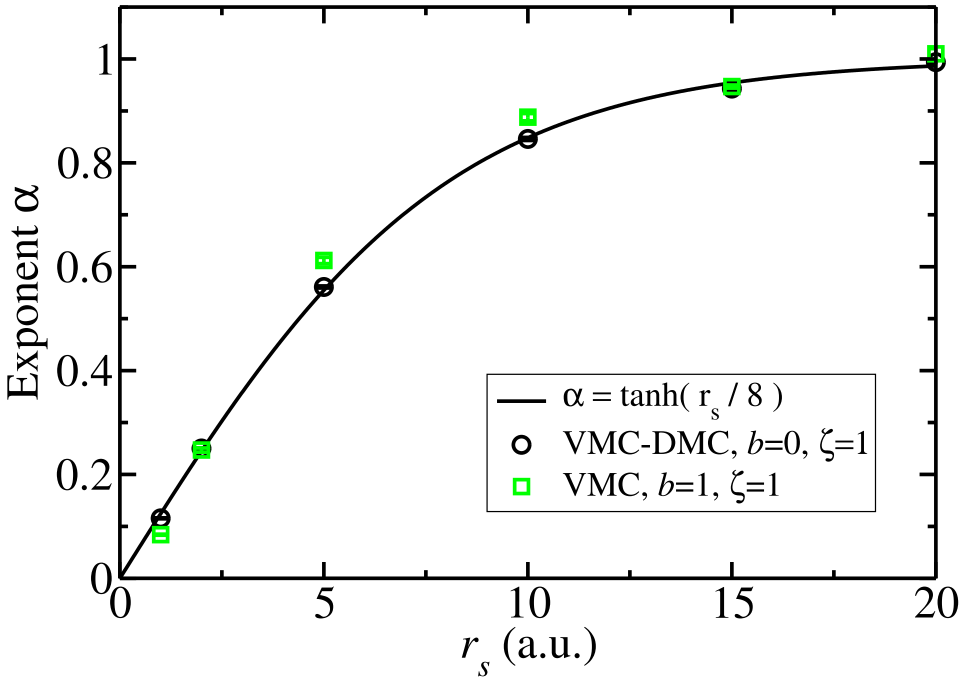

Figure 4.28: Exponent α found from fitting Eq. (4.31) to the MDs of the ferromagnetic systems.

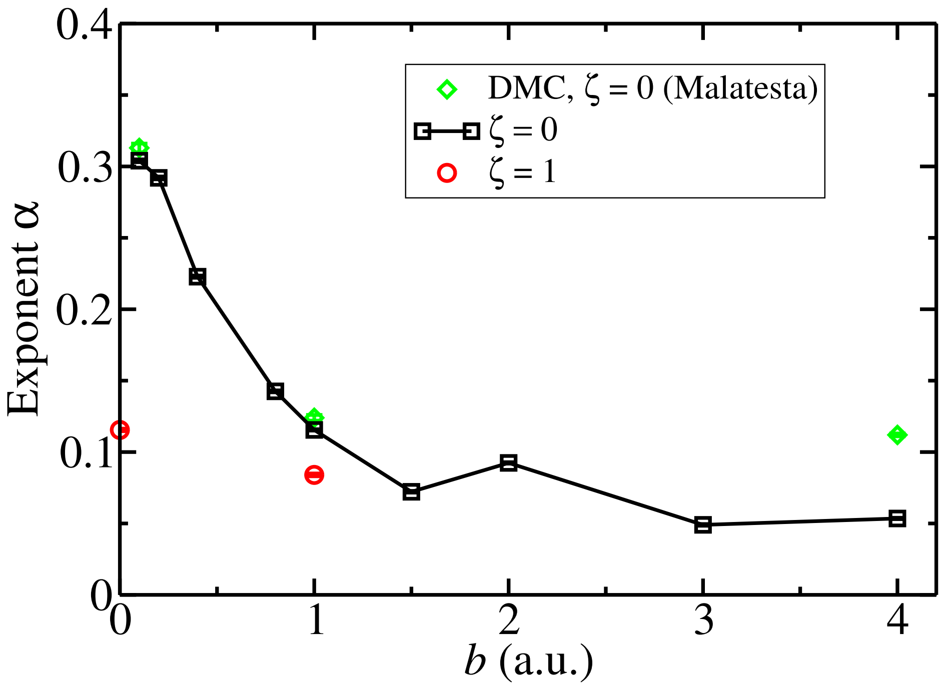

Figure 4.29: Exponent α found from fitting Eq. (4.31) to the MD of systems with $r_s = 1$ a.u.

extrapolation to $\varepsilon = 0$ and it appears that the whole range of $n(k)$ was included in the fit. Figure 4.29 also includes the exponent we find for the infinitely-thin wire (from VMC and DMC estimates of the MD) at $b = 0$.

The exponent α has also been reported from experiments, mostly through measurements accessing the single-particle density of states near the Fermi edge. The exponent for carbon nanotubes ranges between 0.2 and 0.4, although it is difficult to map the behaviour of electrons in these systems onto our model since the electronic properties depend on the folding geometry [77, 89–92]. For the Bechgaard salts, which have a 1D carrier density of $r_s \approx 6.9$ a.u., exponents between 0.5625 and 0.8 have been reported [82, 134–136], which agree with all of our $\zeta = 1$ results, and also the $\zeta = 0$, $b = 4$ results.

4.5 Conclusions

We have presented calculations of the ground state energy, PCF, SSF, and MD of the infinitely-thin 1D HEG model using VMC and DMC. We observe the development of peaks at increasingly-large even-integer-multiples of k_{F} in the SSF as the density is lowered, consistent with the predictions of Schulz [105].

For the harmonic wire model, we have reported ground-state MDs and TL parameters for a range of densities and confinements. We used VMC to produce these results; comparison of our PCFs, SSFs, and ground state energies with LRDMC results [110] where available indicates that our results are extremely accurate.

The MDs of the $\zeta = 0$ systems tend towards the MDs of the infinitely-thin wire and ferromagnetic harmonic wire as b is decreased and as r_s is increased, both of which have the effect of increasing the electron-electron coupling. One interpretation for this is that correlation is dominating over kinetic confinement, so that antiparallel spin pairs are avoiding one another almost as much as parallel spin pairs.

The TL parameters calculated for the $b = 1$ a.u., $\zeta = 1$ system show reasonable agreement with the infinitely-thin wire results; the maximum deviation of α between the two systems is 0.051(6), which occurs at $r_s = 5$ a.u. The exponent α, which describes the behaviour of the MD at k_{F}, takes values between 0 and 1. The exponent for the $\zeta = 0$ systems shows the same general trend, although the value of α is typically higher than for the $\zeta = 1$ systems. This seems to be largely a consequence of the shift of the weight in the MD (including the singularity in the gradient) to larger k as the coupling is increased.

Chapter 5

Conclusions

In this thesis we have proposed a scheme for efficiently extrapolating DMC results to zero time step, performed a short investigation into statistical errors, and applied QMC to two model systems.

The analysis of Sec. 1.5 yielded clear recommendations. When extrapolating DMC results to zero time step, one should perform two calculations; one at a time step of $\delta\tau$ and another at $\delta\tau/4$, spending 1/9 of the computational effort on the former and 8/9 on the latter. $\delta\tau$ should be fixed at the largest value for which the time step bias is linear and a linear fit should be used for the extrapolation. The scheme is expected to be most useful when one can safely assume the extent of the linear regime, which is the case for studies of many similar systems.

We investigated statistical errors in QMC, focusing specifically on outliers and uncertainty in the correlation length. First, we outlined a self-consistent method for choosing the optimal block length B when reblocking QMC data. One should reduce B from the largest possible value and choose the last block length encountered before $B^3 > 2n\eta_{\mathrm{err}}^4$ is violated, where n is the number of steps and η_{err} is the error factor. Secondly, we performed a number of VMC calculations with the carbon atom and crystalline silicon using different random number sequences. This gave us direct access to the distribution of average energies, error bars, and deviations from the underlying mean as a fraction of the estimated error. Relative to the result expected for exact correlation lengths and Gaussian statistics, we observed a significant increase in the frequencies of outliers for both systems, with the effect appearing much more severe for shorter runs. The probability of observing an energy more than 5 error bars from the underlying mean in VMC calculations on the all-electron carbon atom was enhanced by a factor of between 10^3 and 10^4 for runs of length $200 < n < 1000$. One can return the distribution of results to that expected for Gaussian PDFs and exact correlation lengths by using an accurate estimate of the correlation length obtained from a longer run (or by averaging over the available data).

There are several similarities between the biexciton system of Chapter 3 and the one-dimensional electron gas of Chapter 4. Both models replace the microscopic detail of the environment with a homogeneous medium. Both models also involve confining the quantum particles so that we can restrict the computational approach to a reduced number of dimensions. Perhaps most importantly, both model wave functions have nodal structures that are known exactly. As a result, the DMC approach is ideally-placed to calculate exact energies. In practice, the trial wave functions that we have been able to construct are of sufficient quality that our VMC results have also been extremely accurate and, in some cases, essentially exact. The following paragraphs summarise our key findings.

In our study of biexcitons in bilayer systems, we performed exact DMC calculations of the biexciton binding energy as a function of mass ratio $\sigma = m_e/m_h$ and layer separation d. When d passes some critical value $d_{\rm crit}(\sigma)$, biexcitons dissociate into two separate excitons. Defining $x = \sqrt{\sigma + \sigma^{-1}}$ to account for symmetry, We find that $d_{\rm crit}(x) = (1.19/x)\tanh(-0.5x) + 0.93$ accurately describes the region of biexciton stability. We also performed calculations where the exciton centres-of-mass were constrained, so that the problem was reduced to 5 spatial degrees of freedom instead of the original 8. This yielded exciton-exciton interaction potentials which we were able to parameterise and which closely followed the expected dipole-dipole repulsion form at large exciton-exciton separations. Finally, we examined radial distribution functions (RDFs) in biexcitons to find the spatial extent of bound biexcitons. In terms of size, bound biexcitons are very similar to isolated excitons. We examined the difference between the isolated exciton and bound biexciton electron-hole RDFs to determine approximately whether individual excitons retain their identities when the system is bound. At large layer separations (but of course still satisfying $d < d_{\rm crit}$), we find that bringing two excitons close together gives a peak in the inter-exciton electron-hole RDF superimposed on the change in the intra-exciton RDF due to the presence of the other exciton that is at the same separation as that of the electron-electron RDF. This suggests that, at least for large layer separations, excitons retain their identities in bound biexcitons, and implies that methods based on pair-potentials should be accurate.

In our investigation into the ground state properties of the one-dimensional homogeneous electron gas (HEG) we examined two models for the electron-electron interaction. The first was the infinitely-thin wire, where the electrons are coupled by the bare Coulomb interaction. There is relatively little literature covering the expectation values of the infinitely-thin wire. We reported calculations of the energy, pair correlation function (PCF), static structure factor (SSF), and momentum density (MD). Our calculations of the energy are exact, and comparison of VMC and DMC results indicates that other expectation values are also extremely accurate. We find that the finite-size error on the energy per particle decays as N^{-2} and that twist averaging is ineffective for the infinitely-thin wire. We are able to resolve peaks in the SSF at even-integer-multiples of k_F, consistent with theoretical predictions and indicative of quasi-long-ranged order. By performing a fit to the MD around $k = k_F$, we extracted the Luttinger exponent. Comparing our exponents to those found experimentally for carbon nanotubes, we find that the results coincide for $2 < r_s < 4$. Our results for the ferromagnetic harmonic and infinitely-thin wires and for the harmonic wire with $b = 4$ and $\zeta = 0$ showed agreement with experimental results for the Bechgaard salts.

For the harmonic wire, we reported VMC energies, PCFs and SSFs, demonstrating that our variational calculations are in statistical agreement with exact results in the literature. We then presented MDs and Luttinger exponents for the harmonic wire, noting agreement between the ferromagnetic harmonic wire and the infinitely-thin wire. An interesting effect, which also presented a problem for the determination of the Luttinger exponent, was that the effective Fermi wave vector appears to change depending on the density and confinement of the paramagnetic harmonic wire. When the electron-electron coupling is weak (*i.e.,* when b is large and r_s is small), the Fermi wave vector is $k_F = \pi/(4r_s)$ as expected. As the coupling is increased, the effective Fermi wave vector (the location of the singularity in the gradient of the MD) gradually increases to $\pi/(2r_s)$.

This is consistent with the development of ‘pseudo-nodes’ in the wave function at the antiparallel coalescence points, increasing the system’s resemblance to the infinitely-thin wire, where real nodes occur at these points.

Appendix A

Appendix

A.1 Fit to the exciton-exciton potential

The exciton-exciton potential curves with $\sigma = 0$ in Fig. 3.9 may be fitted to a function of the form

$$\begin{aligned} E_I &= \left(p_1 + \frac{1000}{R} + \frac{p_2}{R^2} + \frac{p_3}{R^4}\right) \exp\left(-\frac{p_4 R^3}{1000}\right) \\ &+ \left(\frac{2d^2}{R^3} + \frac{p_5}{R^5} + \frac{p_6}{R^6}\right) \left[1 - \exp\left(-p_7 R^3\right)\right], \end{aligned} \tag{A.1}$$

where d is the layer separation and $p_1, \ldots, p_7$ are the fitting parameters. The function has the correct long-range behaviour, $E_I \propto 2d^2/R^3$ for $R \to \infty$. The fitting parameter values are shown in Table A.1.

The interaction potentials in Fig. 3.8 with $\sigma = 1$ may be fitted to a function similar to Eq. (A.1). This time the form is

$$E_I = \left(\frac{2d^2}{R^3} + \frac{p_1}{R^5} + \frac{p_2}{R^6}\right) \left[1 - \exp\left(-\frac{p_3 R^{p_4}}{1000}\right)\right], \tag{A.2}$$

where the long range behaviour is once again reproduced correctly and each of the terms in the first bracket has a physical interpretation. The $1/R^5$ term may be associated with quadrupole-quadrupole repulsion and the $1/R^6$ term with van der Waals attraction. The signs of the fitting parameters are consistent with this interpretation for $d = 0.2$ and $0.5a_B^*$. The parameter values are shown in Table A.2.

Parameter	$d = 0.9a_B^*$	$d = 1.0a_B^*$	$d = 1.1a_B^*$
p_1	-70.18	-66.25	-61.89
p_2	-4243	-4538	-4804
p_3	4296	8422	12560
p_4	8.086	7.319	6.420
p_5	21520	30740	40610
p_6	-15100	-54120	-98510
p_7	0.1284	0.1451	0.2424

Table A.1: Coefficients appearing in Eq. (A.1) allowing the reproduction of fits to the points shown in Fig. 3.9. Performing the fits using data with $R \geq 3a_B^*$ yields χ^2 errors of 0.79, 1.1 and 1.4 per data point for $d = 0.9$, 1.0, and $1.1a_B^*$, respectively.

Parameter	$d = 0.2a_B^*$	$d = 0.5a_B^*$	$d = 0.9a_B^*$
p_1	2302	1463	-6797
p_2	-8947	-12580	73200
p_3	1316	5.813	24.94
p_4	0.1123	4.703	2.465

Table A.2: Coefficients appearing in Eq. (A.2) allowing the reproduction of fits to the DMC results in Fig. 3.8. Performing the fits using data with $R \geq 2.5a_B^*$ yields χ^2 errors of 1.55, and 1.07 per data point for $d = 0.2$ and $0.5a_B^*$, respectively. The χ^2 error is larger for $d = 0.9a_B^*$, the purely repulsive curve, but the maximum deviation from the data points is only $1.2 \times 10^{-3} R_y^*$.

A.2 Derivation of the quasi-1D interaction

One may derive Eq. (4.5) from first principles. Suppose that the Schödinger equation is separable and we may write the wave function as a product $\theta(\mathbf{r}_\perp)\psi(x)$, where x is the projection of the electron position onto the axis of the wire and $\mathbf{r}_\perp$ is the transverse position.

If the electrons are sufficiently confined in the transverse plane, one may obtain the 1D interaction $v(x)$ by integrating over the transverse part of the wave function,

$$v(x) = \int \frac{|\theta(\mathbf{r}_\perp)|^2 \, |\theta(\mathbf{r}'_\perp)|^2}{\left[x^2 + |\mathbf{r}_\perp - \mathbf{r}'_\perp|^2\right]^{1/2}} \, d\mathbf{r}_\perp d\mathbf{r}'_\perp \; . \tag{A.3}$$

The confining potential for the harmonic wire is $r_\perp^2/8b^4$, where b is a parameter. If $r_s \gg \pi b/4$, one may make the assumption that the electrons occupy only the lowest sub-band, which is given by

$$\theta(r_\perp) = \frac{1}{\sqrt{2\pi b^2}} \exp\left(-\frac{r_\perp^2}{4b^2}\right) \; . \tag{A.4}$$

Substituting Eq. (A.4) into Eq. (A.3) yields [113]

$$v(x) = \frac{\sqrt{\pi}}{2b} \exp\left(\frac{x^2}{4b^2}\right) \operatorname{erfc}\left(\frac{|x|}{2b}\right) \; , \tag{A.5}$$

which is finite at $x = 0$ but retains a long-range $1/|x|$ tail. The Fourier transform of Eq. (A.5) is

$$\tilde{v}(k) = E_1(b^2k^2) \exp(b^2k^2) \; , \tag{A.6}$$

where E_1 is the exponential integral function.

Having found the real and reciprocal space representations of the 1D interaction in a harmonic wire, one must perform an Ewald-like sum to enable calculations with periodic systems. We follow a route similar to that of Ref. [121].

The interaction of an electron at the origin with another at position x, all of that electron's periodic images, and its background is given by

$$\phi(x) = \sum_{m=-\infty}^{\infty} \left\{ v(x - mL) - \frac{1}{L} \int_{-L/2}^{L/2} dy \; v(x - mL - y) \right\} \; , \tag{A.7}$$

where L is the length of the simulation cell. The objective is to write Eq. (A.7) in terms of quickly converging discrete sums. The first step is to write Eq. (A.7) in the more useful form

$$\phi(x) = \gamma_0(x) - \frac{1}{L}\int_{-\infty}^{\infty} dy\ v(x-y)\ , \tag{A.8}$$

where

$$\gamma_0(x) = \sum_{m=-\infty}^{\infty} v(x-mL)\ . \tag{A.9}$$

Equation (A.9) is already in a form that is quick and easy to evaluate, so we turn our attention to reformulating the integral in the second term of Eq. (A.8). We first perform the trick of both adding and subtracting a Gaussian term $p(y)$, giving

$$-\frac{1}{L}\int_{-\infty}^{\infty} dy\ v(x-y) = \gamma_1(x) + \gamma_2(x)\ , \tag{A.10}$$

where

$$\begin{aligned}
\gamma_1(x) &= -\int_{-\infty}^{\infty} dy\ v(x-y)p(y)\ , && \text{(A.11)}\\
\gamma_2(x) &= \int_{-\infty}^{\infty} dy\ v(x-y)\left[p(y) - \frac{1}{L}\right]\ , && \text{(A.12)}
\end{aligned}$$

and the term that we have added and subtracted is

$$p(y) = \sum_{m=-\infty}^{\infty} \frac{1}{2b\sqrt{\pi}} \exp\left(-\frac{1}{4b^2}(y-mL)^2\right)\ . \tag{A.13}$$

It is clear that $\phi(x)$ may now be written simply as

$$\phi(x) = \gamma_0(x) + \gamma_1(x) + \gamma_2(x)\ . \tag{A.14}$$

We first inspect $\gamma_1(x)$, finding that it may be integrated directly to give

$$\gamma_1(x) = \sum_{m=-\infty}^{\infty} \left\{ -\frac{1}{|x-mL|}\ \text{erf}\left(\frac{|x-mL|}{2b}\right)\right\}\ , \tag{A.15}$$

which is a form suitable for numerical evaluation.

One may take the first step towards simplifying $\gamma_2(x)$ by performing a Poisson summation on $p(y)$,

$$p(y) = \frac{1}{L}\left[1 + 2\sum_{n=1}^{\infty} e^{-(bGn)^2}\cos(Gny)\right]\ , \tag{A.16}$$

where $G = 2\pi/L$. Putting Eq. (A.16) into Eq. (A.12) gives

$$\gamma_2(x) = \frac{2}{L}\sum_{n=1}^{\infty} e^{-(bGn)^2}\int_{-\infty}^{\infty} dy\ v(x-y)\cos(Gny)\ , \tag{A.17}$$

which may straightforwardly be rewritten in its final form,

$$\gamma_2(x) = \frac{2\sqrt{2\pi}}{L} \sum_{n=1}^{\infty} \tilde{v}(Gn) \cos(Gnx)\, e^{-(bGn)^2} \, , \tag{A.18}$$

where we have used the result

$$\int_{-\infty}^{\infty} dy\, v(x-y) \cos(Gny) = \sqrt{2\pi}\, \tilde{v}(Gn) \cos(Gnx) \, . \tag{A.19}$$

Finally, putting the expressions for the γ functions, Eqs. (A.9), (A.15), and (A.18), into Eq. (A.14) and remembering that $\tilde{v}(k)$ is given by Eq. (A.6), we obtain the more computationally convenient form

$$\begin{aligned} \phi(x) \;&=\; \sum_{m=-\infty}^{\infty} \left\{ \frac{\sqrt{\pi}}{2b} e^{(x-mL)^2/(4b^2)} \mathrm{erfc}\left(\frac{|x-mL|}{2b} \right) \right. \\ &\quad - \left. \frac{1}{|x-mL|} \,\mathrm{erf}\left(\frac{|x-mL|}{2b} \right) \right\} \\ &\quad + \frac{2}{L} \sum_{n=1}^{\infty} E_1 \left[(bGn)^2 \right] \cos(Gnx) \, . \end{aligned} \tag{A.20}$$

It should be noted that in Ref. [110], Rydberg rather than Hartree units were used so that the potentials given there differ from Eq. (A.20) by a factor of 2.

A.3 Pair correlation function fitting parameters

Table A.3 shows the parameters that we obtained when fitting Eq. (4.13) to the extrapolated estimates of the PCF for the infinitely-thin wire. We performed fits for $r_s = 1$, 2, 5, 10, 15, and 20 a.u. for systems containing $N = 37$, 55, 73, and 99 particles. The PCF data in the range $6r_s < x < L/2 - 6r_s$ were included.

r_s (a.u.)	N	A	B (a.u.)
1	37	1.908	3.291
1	55	3.090	3.683
1	73	1.940	3.446
1	99	2.113	3.671
2	37	4.851	2.952
2	55	4.573	2.979
2	73	6.047	3.069
2	99	8.545	3.273
5	37	8.029	2.237
5	55	8.310	2.258
5	73	10.061	2.359
5	99	9.262	2.320
10	37	8.465	1.735
10	55	9.349	1.788
10	73	9.066	1.780
10	99	10.206	1.839
15	37	8.520	1.502
15	55	8.363	1.501
15	73	8.918	1.534
15	99	9.788	1.580
20	37	7.754	1.320
20	55	8.361	1.359
20	73	8.625	1.377
20	99	8.895	1.396

Table A.3: Table showing the fitting parameters A and B from Eq. (4.13) obtained from fitting to the extrapolated estimates of the PCF for the infinitely-thin wire. The fits were to PCF data in the range $6r_s < x < L/2 - 6r_s$.

A.4 Outliers from distributions with heavy tails

In this appendix we repeat some of the analysis of Chapter 2 starting from the heavy-tailed distribution of energies given by Trail in Ref. [35].

The notation is the same as before — due to the large number of variables we give all of their definitions below:

- n — number of local energies;
- ν_0 — exact effective number of local energies;
- ν — estimated effective number of local energies;
- σ_0^2 — exact variance of distribution of local energies;
- $\bar{E}$ — estimated mean energy;
- E_0 — exact mean energy;
- n_{corr} — exact correlation length;
- η_{err}^2 — estimated correlation length;
- Δ — estimated error bar.

Where a variable is described as 'exact', it is a parameter of the system and is a single number, whereas other quantities are drawn from distributions.

As discussed in Chapter 2, singularities in the local energy occur when using approximate wave function nodes (so that in principle, even where the nodes are correctly placed, the behaviour of the wave function has the potential to produce singularities in E_L), and when the cusp conditions are violated. The wave function nodes for the C atom and the bulk Si systems that we studied in Chapter 2 are unknown, so we might expect these singularities to play an important role in determining the frequency with which outliers occur. Trail [35] showed that in such cases, the distribution of local energies develops heavy tails $\propto |E_L - E_0|^4$, and derived asymptotic expressions for the PDFs. Here, we use those distributions to repeat the analysis of Chapter 2 to see if we can recover the distribution of outliers observed directly within VMC. As we were already able to reproduce the VMC results using purely Gaussian energy PDFs with relatively simple analysis, we might anticipate that using heavy-tailed distributions will not significantly change the outcome.

Let us define $y = (\bar{E} - E_0)\sqrt{\nu_0}/\sigma_0$ to bring us in line with the notation of Ref. [35], where a series expansion of the PDF p_y of y is derived. The PDF is

$$\begin{aligned} \mathrm{p}_y(y) &= \frac{1}{\sqrt{2\pi}}\left[1 + \frac{\eta_3}{\sqrt{\nu_0}}\frac{d^3}{dy^3} + O\left(\frac{1}{\nu_0}\right)\right]e^{-y^2/2} \\ &- \left[\frac{\lambda_3}{3\pi\sqrt{\nu_0}}\frac{d^3}{dy^3} + \mathcal{O}\left(\frac{1}{\nu_0}\right)\right]D\left(\frac{y}{\sqrt{2}}\right) , \end{aligned} \tag{A.21}$$

where η_3 and λ_3 are system-dependent; η_3 describes the skew of the distribution and λ_3 describes the weight under the tails and the sharpness of the peak (the *kurtosis* of the PDF). In the limit $\lambda_3 \to 0$ and $\eta_3 \to 0$, we should recover the result obtained for Gaussian distributions. Dawson's integral $D(x)$ is given by

$$D(x) = e^{-x^2}\int_0^x e^{t^2}\,dt , \tag{A.22}$$

which may be straightforwardly differentiated. We evaluate Eq. (A.22) numerically using the relation

$$D(x) = \cfrac{x}{1 + 2x^2 - \cfrac{4x^2}{3+2x^2 - \cfrac{4x^2}{5+2x^2 - \cfrac{4x^2}{7+2x^2-\ldots}}}} . \tag{A.23}$$

Computing the derivatives in Eq. A.21 gives

$$\mathrm{p}_y(y) = \frac{1}{\sqrt{2\pi}}e^{-y^2/2} - \frac{\lambda_3}{3\pi\sqrt{\nu_0}}\left((3y - y^3)D\left(\frac{y}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}}(y^2 - 2)\right) + \mathcal{O}\left(\frac{1}{\nu_0}\right) , \tag{A.24}$$

where we have set $\eta_3 = 0$ so that we only consider symmetric PDFs – this seems to be an acceptable assumption for real data. Note that p_y is related to the distribution $\mathrm{p}_{\bar{E}}$ that we used in Chapter 2 by

$$\mathrm{p}_{\bar{E}}(\bar{E}) = \frac{\sqrt{\nu_0}}{\sigma_0}\mathrm{p}_y\left[\frac{(\bar{E} - E_0)\sqrt{\nu_0}}{\sigma_0}\right] . \tag{A.25}$$

Figure (A.1) shows the distribution of Eq. (A.24). Henceforth we drop the $\mathcal{O}(\nu_0^{-1})$ terms.

To find the probability of observing a mean energy more than Q error bars Δ from the underlying mean E_0, we take the same route as in Chapter 2 of first finding the probability

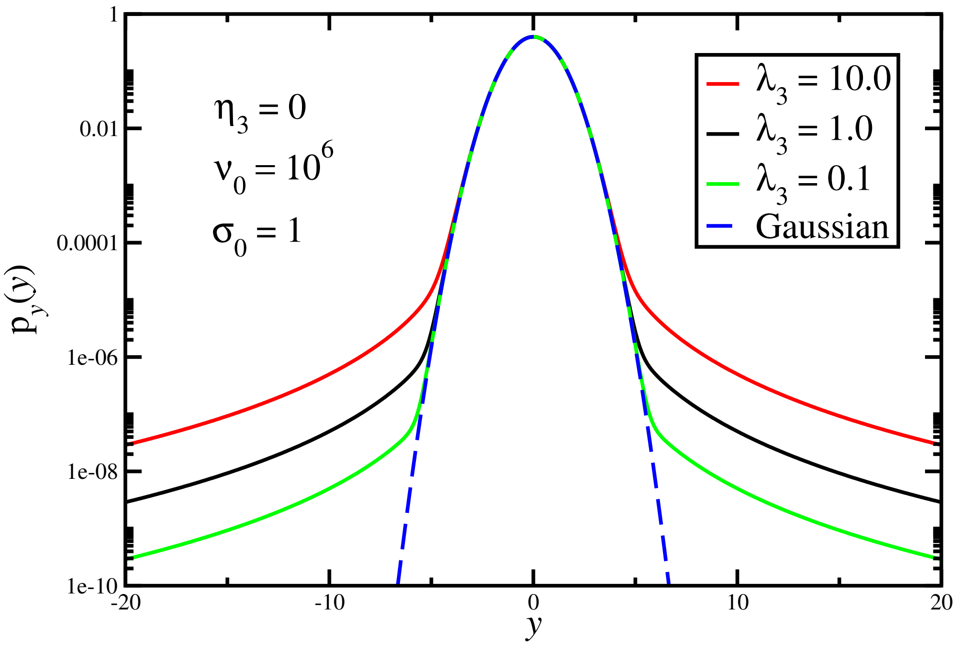

Figure A.1: Plot of Eq. (A.24) with $\eta_3 = 0$, $\nu_0 = 10^6$, $\sigma_0 = 1$, and a number of values of λ_3.

as a function of Q and Δ,

$$\begin{aligned} 2\int_{\frac{Q\Delta\sqrt{\nu_0}}{\sigma_0}}^{\infty} \mathrm{d}y\ \mathrm{p}_y(y) &= \mathrm{erfc}\left(\frac{Q\Delta\sqrt{\nu_0}}{\sqrt{2}\sigma_0}\right) - \frac{\sqrt{2}\lambda_3 Q\Delta}{3\pi\sigma_0} \\ &+ \frac{\lambda_3}{3\pi\sqrt{\nu_0}}\left[2\left(1-\left(\frac{Q\Delta\sqrt{\nu_0}}{\sigma_0}\right)^2\right) D\left(\frac{Q\Delta\sqrt{\nu_0}}{\sqrt{2}\sigma_0}\right)\right], \end{aligned} \tag{A.26}$$

so that finding the expectation value of Eq. (A.26) with respect to ν and Δ will give the desired probability $\mathrm{P}\left(\delta\bar{E} > Q\Delta\right)$. Note that when $\lambda_3 = 0$ only the first term of Eq. (A.26) remains, which is equal to Eq. (2.10) in Chapter 2.

Helpfully, Ref. [35] also provides the distribution of errors that arises when the PDF of local energies is described by Eq. (A.21). Translating the result in Trail's paper into our language of Δ and ν, we obtain

$$\mathrm{p}_{\mathrm{err}}(\Delta,\nu) = \frac{\sqrt{3}\nu\Delta}{\pi\gamma}\chi^2\exp(\chi^3)\left[-\mathrm{sgn}[\chi]K_{1/3}\left(|\chi|^3\right) + K_{2/3}\left(|\chi|^3\right)\right]\mathrm{p}_{\mathrm{ind}}(\nu)\ , \tag{A.27}$$

where K_n is the Bessel function of the nth kind and we have defined

$$\chi = \frac{\nu\Delta^2 - \sigma_0^2}{2\gamma} \tag{A.28}$$

and

$$\gamma = \sigma_0^2 \left[\frac{6\lambda_3^2}{\pi\nu} \right]^{1/3} , \tag{A.29}$$

so that γ is a scale parameter controlling the width of the distribution (although the variance of $\mathrm{p_{err}}$ with respect to Δ is undefined). The PDF $\mathrm{p_{ind}}(\nu)$ is again the distribution of the estimated effective number of steps.

The above analysis puts us in a position to construct the integral that yields the desired probability, $\mathrm{P}\left(\delta\bar{E} > Q\Delta\right)$. The expectation value of Eq. (A.26) with respect to Eq. (A.27) gives

$$\mathrm{P}\left(\delta\bar{E} > Q\Delta\right) = 2\int_2^\infty \mathrm{d}\nu \int_0^\infty \mathrm{d}\Delta \int_{\frac{Q\Delta\sqrt{\nu_0}}{\sigma_0}}^\infty \mathrm{d}y \; \mathrm{p}_y(y)\mathrm{p_{err}}(\Delta,\nu) , \tag{A.30}$$

which we perform numerically.

As before, there are some inputs to Eq. (A.30) for which we require VMC results. These are λ_3, $\mathrm{p_{ind}}(\nu)$, ν_0 and σ_0, although Eq. (A.30) is very insensitive to the value of σ_0.

Since we have already reproduced the VMC results using purely Gaussian distributions, we do not expect any new behaviour from Eq. (A.30), and consider it almost as a consistency check. Let us examine the same example as before (see Chapter 2), the C atom, using the new PDFs. We find from the VMC data that the values $\sigma_0 = 0.637$ and $\nu_0 = 18.07$ are appropriate for the $n = 200$ runs, and also take the PDF $\mathrm{p_{ind}}$ from VMC results. As mentioned in Chapter 2, λ_3 is problematic to estimate. Fitting Eq. (A.24) to a histogram of the distribution of energies from VMC yields $\lambda_3 = 1.1(8)$ — the value is unfortunately noisy because it is dependent on the frequency of samples in the tail regions of the energy PDF, which are difficult to sample properly. As a result, we have computed the integral of Eq. (A.30) for several values of λ_3 around the value $\lambda_3 = 1.1$; Fig. A.2 shows the results.

It is clear from Fig. A.2 that the integrals over the heavy-tailed distributions are capable of reproducing both the VMC results and those from assuming Gaussian energy PDFs. The agreement is almost perfect at $\lambda_3 = 0.5$, and the $\lambda_3 = 1.1$ result is still close to the VMC data. The value $\lambda_3 = 0.5$ is within error bars of our earlier estimate of $\lambda_3 = 1.1(8)$. One should note that the distribution of Eq. (A.27) that was derived by Trail is strictly-speaking only valid in the limit $\nu_0 \to \infty$. However, it was shown in

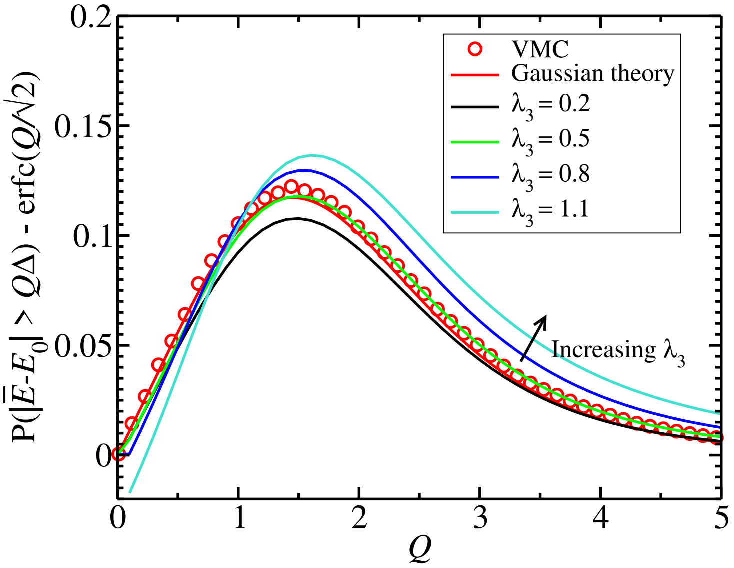

Figure A.2: The result of numerical evaluation of Eq. (A.24) for several values of λ_3. Also shown are the VMC results and the result obtained in Chapter 2 using purely Gaussian generating distributions. The results are for the C atom with $n = 200$. The parameters were $\sigma_0 = 0.637$ and $\nu_0 = 18.07$, and $\mathrm{p_{ind}}$ was taken from the VMC results.

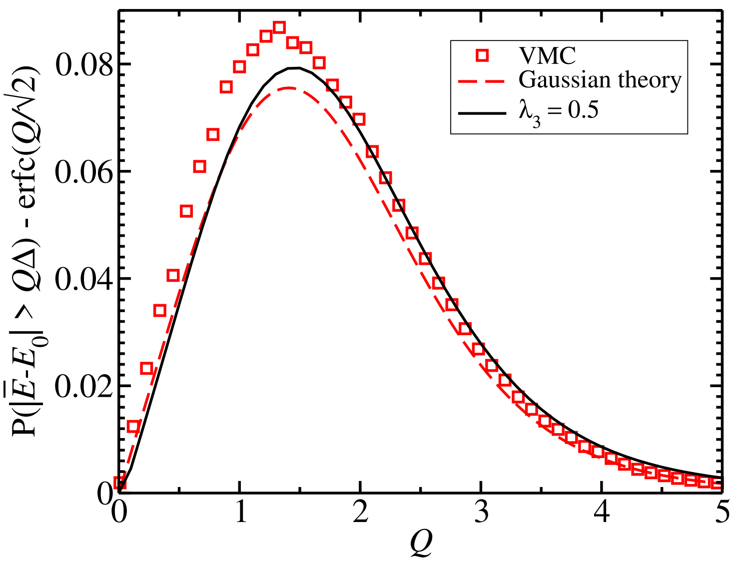

Figure A.3: The result of numerical evaluation of Eq. (A.24) for the C atom with $n = 500$. Also shown are the VMC results and the result obtained in Chapter 2 using purely Gaussian generating distributions. The parameters were $\sigma_0 = 0.637$ and $\nu_0 = 45.2$, and $\mathrm{p_{ind}}$ was estimated from the VMC results.

Ref. [35] that this limit is reached surprisingly quickly, and the agreement of our results for some values of λ_3 suggests that Eq. (A.27) is at least a good approximation to the actual underlying PDF. The reason that the heavy-tailed PDF results do not reduce to the Gaussian results in the limit $\lambda_3 \to 0$ is that the PDF of Eq. (A.27) does not reduce to Eq. (2.8), which was the corresponding Gaussian PDF, in the same limit. This is again due to Eq. (A.27) being only asymptotically accurate. In contrast with the analysis performed in Chapter 2, the assumption that the sample variance and the sample mean are independent is approximate when the PDFs possess heavy tails. The results also suggest that $\lambda_3 \approx 0.5$ is perhaps a more accurate estimate of λ_3.

Figure A.3 demonstrates the transferability of the estimate of λ_3. The estimate $\lambda_3 = 0.5$ yields good agreement between the theory and the $n = 500$ VMC results. This is reassuring, and confirms that $\lambda_3 = 0.5$ is a reasonable estimate for the C system.

The main result of the analysis presented above is that one can reproduce the distribution of outliers observed in VMC calculations using either Gaussian or heavy-tailed energy PDFs, suggesting that the heavy tails in the PDFs are relatively unimportant. The key ingredient appears to be the distribution of the estimated effective number of steps, which is related to the distribution of estimated correlation lengths. We conclude from this that it is uncertainty in the correlation length that is responsible for the increased frequency of outliers that we observe for the systems studied. It is of course conceivable that a system with particularly strong singularities in the local energy (*i.e.,* systems for which λ_3 is very large) could give results dominated by the heavy tails in the PDFs.

Bibliography

[1] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[2] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos, *CASINO version 2.1 User Manual* (University of Cambridge, Cambridge, 2007).

[3] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos, J. Phys.: Condens. Mat. **22**, 023201 (2010).

[4] M. D. Brown, *Energy Minimisation in Variational Quantum Monte Carlo*, PhD thesis, University of Cambridge, Cambridge, 2007.

[5] N. D. Drummond, *Application of quantum Monte Carlo methods to electronic systems*, PhD thesis, University of Cambridge, Cambridge, 2004.

[6] P. López Ríos, *Backflow and pairing wave function for quantum Monte Carlo methods*, PhD thesis, University of Cambridge, Cambridge, 2006.

[7] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in Fortran 77* (Cambridge University Press, Cambridge, 2007).

[8] P. E. Gill and G. F. Miller, J. Comp. **15**, 80 (1972).

[9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[10] M. D. Towler, Physica Status Solidi (b) **243**, 2573 (2006).

[11] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).

[12] N. D. Drummond and R. J. Needs, Phys. Rev. B **72**, 085124 (2005).

[13] P. R. C. Kent, R. J. Needs, and G. Rajagopal, Phys. Rev. B **59**, 12344 (1999).

[14] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007).

[15] A. J. Morris, *A stochastic approach to condensed matter physics*, PhD thesis, University of Cambridge, Cambridge, 2009.

[16] N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. B **70**, 235119 (2004).

[17] M. Gillan, M. Towler, and D. Alfè, *Psi-k Highlight of the Month* (February, 2011).

[18] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[19] G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009).

[20] D. Cleland, G. H. Booth, and A. Alavi, J. Chem. Phys. **132**, 041103 (2010).

[21] G. H. Booth and A. Alavi, J. Chem. Phys. **132**, 174104 (2010).

[22] M. Kalos, L. Colletti, and F. Pederiva, J. Low Temp. Phys. **138**, 747 (2005).

[23] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).

[24] J. B. Anderson, Int. Rev. Phys. Chem. **14**, 85 (1995).

[25] D. Ceperley, M. H. Kalos, and J. L. Lebowitz, Macromolecules **14**, 1472 (1981).

[26] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, J. Chem. Phys. **77**, 5593 (1982).

[27] C. J. Umrigar, M. P. Nightingale, and K. J. Runge, J. Chem. Phys. **99**, 2865 (1993).

[28] D. M. Ceperley, J. Stat. Phys. **63**, 1237 (1991).

[29] D. M. Ceperley and M. H. Kalos, Quantum many body problems, in *Monte Carlo Methods in Statistical Physics, 2nd ed.*, edited by K. Binder, Springer-Verlag, 1986.

[30] H. Flyvbjerg and H. G. Petersen, J. Chem. Phys. **91**, 461 (1989).

[31] U. Wolff, Comput. Phys. Commun. **156**, 143 (2004).

[32] W. Cochran, Math. Proc. Cambridge **30**, 178 (1934).

[33] R. C. Geary, Suppl. J. Royal Stat. Soc. **3**, 178 (1936).

[34] D. D. Boos and J. M. Hughes-Oliver, Am. Stat. **52**, 218 (1998).

[35] J. R. Trail, Phys. Rev. E **77**, 016703 (2008).

[36] J. Frenkel, Phys. Rev. **37**, 17 (1931).

[37] W. Liang, Phys. Educ. **5**, 226 (1970).

[38] G. H. Wannier, Phys. Rev. **52**, 191 (1937).

[39] H. Hoppe and N. Sariciftci, J. Mater. Res. **19**, 1925 (2004).

[40] E. Poem *et al.*, Nat. Phys. **6**, 993 (2010).

[41] D. Snoke, Science **298**, 1368 (2002).

[42] L. V. Butov, A. C. Gossard, and D. S. Chemla, Nature **418**, 751 (2002).

[43] L. V. Butov, J. Phys.: Condens. Mat. **16**, R1577 (2004).

[44] R. Rapaport *et al.*, Phys. Rev. Lett. **92**, 117405 (2004).

[45] A. T. Hammack *et al.*, Phys. Rev. Lett. **96**, 227402 (2006).

[46] M. H. Szymanska and P. B. Littlewood, Phys. Rev. B **67**, 193305 (2003).

[47] M. Y. J. Tan, N. D. Drummond, and R. J. Needs, Phys. Rev. B **71**, 33303 (2005).

[48] C. Schindler and R. Zimmermann, Phys. Rev. B **78**, 45313 (2008).

[49] S. De Palo, F. Rapisarda, and G. Senatore, Phys. Rev. Lett. **88**, 206401 (2002).

[50] E. Cornell and C. Wieman, Rev. Mod. Phys. **74**, 875 (2002).

[51] W. Ketterle, Rev. Mod. Phys. **74**, 1131 (2002).

[52] L. Keldysh and A. Kozlov, Sov. J. Exp. Theor. Phys. **27**, 521 (1968).

[53] L. V. Butov, Solid State Commun. **127**, 89 (2003).

[54] D. W. Snoke, Physica Status Solidi (b) **238**, 389 (2003).

[55] J. Eisenstein and A. MacDonald, Nature **432**, 691 (2004).

[56] D. Snoke, J. Wolfe, and A. Mysyrowicz, Phys. Rev. Lett. **59**, 827 (1987).

[57] D. Snoke, J. Wolfe, and A. Mysyrowicz, Phys. Rev. Lett. **64**, 2543 (1990).

[58] J. Lin and J. Wolfe, Phys. Rev. Lett. **71**, 1222 (1993).

[59] K. E. O'Hara, L. Ó Súilleabháin, and J. P. Wolfe, Phys. Rev. B **60**, 10565 (1999).

[60] L. Butov *et al.*, Phys. Rev. Lett. **86**, 5608 (2001).

[61] P. B. Littlewood and X. Zhu, Phys. Scripta **1996**, 56 (1996).

[62] L. L. Chase, N. Peyghambarian, G. Grynberg, and A. Mysyrowicz, Phys. Rev. Lett. **42**, 1231 (1979).

[63] S. Moskalenko and D. Snoke, *Bose-Einstein condensation of excitons and biexcitons: and coherent nonlinear optics with excitons* (Cambridge University Press, 2000).

[64] R. Zimmermann and C. Schindler, Solid State Commun. **144**, 395 (2007).

[65] T. Kato, Pure Appl. Math **10**, 151 (1957).

[66] R. T. Pack and W. B. Brown, J. Chem. Phys. **45**, 556 (1966).

[67] C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).

[68] P. R. C. Kent, R. J. Needs, and G. Rajagopal, Phys. Rev. B **59**, 12344 (1999).

[69] N. D. Drummond and R. J. Needs, Phys. Rev. B **72**, 085124 (2005).

[70] A. D. Meyertholen and M. M. Fogler, Phys. Rev. B **78**, 235307 (2008).

[71] J. Usukura, Y. Suzuki, and K. Varga, Phys. Rev. B **59**, 5652 (1999).

[72] G. Giuliani and G. Vignale, *Quantum theory of the electron liquid* (Cambridge University Press, 2005).

[73] S. Tomonaga, Prog. Theor. Phys. **5**, 544 (1950).

[74] J. M. Luttinger, J. Math. Phys. **4**, 1154 (1963).

[75] F. Haldane, J. Phys. C: Solid State **14**, 2585 (1981).

[76] V. Mitin, A. Sergeev, M. Bell, J. Bird, and A. Verevkin, J. Phys.: Conf. Series **193**, 012116 (2009).

[77] M. Bockrath *et al.*, Nature **397**, 598 (1999).

[78] J. Voit, Rep. Prog. Phys. **58**, 977 (1995).

[79] H. Schulz, G. Cuniberti, and P. Pieri, Fermi liquids and luttinger liquids, in *Field Theories for Low-Dimensional Condensed Matter Systems*, edited by G. Morandi, P. Sodano, A. Tagliacozzo, and V. Tognetti, Springer-Verlag, 2000.

[80] V. Meden, W. Metzner, U. Schollwöck, and K. Schönhammer, J. Low Temp. Phys. **126**, 1147 (2002).

[81] T. Ito *et al.*, Phys. Rev. Lett. **95**, 246402 (2005).

[82] M. Dressel, K. Petukhov, B. Salameh, P. Zornoza, and T. Giamarchi, Phys. Rev. B **71**, 075104 (2005).

[83] T. Lorenz *et al.*, Nature **418**, 614 (2002).

[84] A. Schwartz *et al.*, Phys. Rev. B **58**, 1261 (1998).

[85] V. Vescoli *et al.*, Eur. Phys. J. B **13**, 503 (2000).

[86] M. Dressel *et al.*, Synthetic Met. **120**, 719 (2001).

[87] S. Maekawa, *Physics of transition metal oxides,* Springer series in solid-state sciences (Springer, 2004).

[88] Z. Hu *et al.*, Eur. Phys. J. B **26**, 449 (2002).

[89] M. Dresselhaus, G. Dresselhaus, and P. Avouris, *Carbon nanotubes: synthesis, structure, properties, and applications,* Topics in applied physics (Springer, 2001).

[90] R. Egger, Phys. Rev. Lett. **83**, 5547 (1999).

[91] H. Ishii *et al.*, Nature **426**, 540 (2003).

[92] M. Shiraishi and M. Ata, Solid State Commun. **127**, 215 (2003).

[93] F. Milliken, C. Umbach, and R. Webb, Solid State Commun. **97**, 309 (1996).

[94] A. Chang, Rev. Mod. Phys. **75**, 1449 (2003).

[95] S. Mandal and J. Jain, Solid State Commun. **118**, 503 (2001).

[96] H. Steinberg *et al.*, Phys. Rev. B **73**, 113307 (2006).

[97] S. V. Zaitsev-Zotov, Y. A. Kumzerov, Y. A. Firsov, and P. Monceau, J. Phys.: Condens. Mat. **12**, L303 (2000).

[98] F. Liu *et al.*, Appl. Phys. Lett. **86**, 213101 (2005).

[99] A. Goni *et al.*, Phys. Rev. Lett. **67**, 3298 (1991).

[100] O. Auslaender *et al.*, Phys. Rev. Lett. **84**, 1764 (2000).

[101] H. Moritz, T. Stöferle, K. Günter, M. Köhl, and T. Esslinger, Phys. Rev. Lett. **94**, 210401 (2005).

[102] A. Recati, P. Fedichev, W. Zwerger, and P. Zoller, J. Opt. B-Quantum S. O. **5**, S55 (2003).

[103] H. Monien, M. Linn, and N. Elstner, Phys. Rev. A **58**, 3395 (1998).

[104] J. Schäfer, C. Blumenstein, S. Meyer, M. Wisniewski, and R. Claessen, Phys. Rev. Lett. **101**, 236802 (2008).

[105] H. J. Schulz, Phys. Rev. Lett. **71**, 1864 (1993).

[106] G. Astrakharchik and M. Girardeau, Arxiv preprint arXiv:1101.0103 (2010).

[107] M. M. Fogler, Phys. Rev. Lett. **94**, 56405 (2005).

[108] C. E. Creffield, W. Häusler, and A. H. MacDonald, Europhys. Lett. **53**, 221 (2001).

[109] M. Fabrizio, A. Gogolin, and S. Scheidl, Phys. Rev. Lett. **72**, 2235 (1994).

[110] M. Casula, S. Sorella, and G. Senatore, Phys. Rev. B **74**, 245427 (2006).

[111] L. Shulenburger, M. Casula, G. Senatore, and R. M. Martin, Phys. Rev. B **78**, 165303 (2008).

[112] A. Malatesta and G. Senatore, J. Phys.-Colloques **10**, 341 (2000).

[113] W. I. Friesen and B. Bergersen, J. Phys. C: Solid State **13**, 6627 (1980).

[114] L. Calmels and A. Gold, Phys. Rev. B **56**, 1762 (1997).

[115] M. Taş and M. Tomak, Phys. Rev. B **67**, 235314 (2003).

[116] V. Garg, R. K. Moudgil, K. Kumar, and P. K. Ahluwalia, Phys. Rev. B **78**, 045406 (2008).

[117] R. Asgari, Solid State Commun. **141**, 563 (2007).

[118] V. R. Saunders, C. Freyria-Fava, R. Dovesi, and C. Roetti, Comput. Phys. Commun. **84**, 156 (1994).

[119] N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes, Phys. Rev. B **78**, 125106 (2008).

[120] E. Lieb and D. Mattis, Phys. Rev. **125**, 164 (1962).

[121] A. Malatesta, *Quantum Monte Carlo study of a model one-dimensional electron gas*, PhD thesis, University of Trieste, Trieste, 1999.

[122] N. March, *Electron correlation in the Solid State* (Imperial College Press, 1999).

[123] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. E **74**, 66701 (2006).

[124] T. Giamarchi, *Quantum physics in one dimension*, International series of monographs on physics (Clarendon, 2004).

[125] C. Lin, F. Zong, and D. Ceperley, Phys. Rev. E **64**, 16702 (2001).

[126] G. Rajagopal, R. J. Needs, A. James, S. D. Kenny, and W. M. C. Foulkes, Phys. Rev. B **51**, 10591 (1995).

[127] N. D. Drummond and R. J. Needs, Phys. Rev. Lett. **102**, 126402 (2009).

[128] M. Holzmann, B. Bernu, V. Olevano, R. Martin, and D. Ceperley, Phys. Rev. B **79**, 41308 (2009).

[129] S. Chiesa, D. M. Ceperley, R. M. Martin, and M. Holzmann, Phys. Rev. Lett. **97**, 076404 (2006).

[130] K. F. Riley, M. P. Hobson, and S. J. Bence, *Mathematical Methods for Physics and Engineering*, 3rd ed. (Cambridge University Press, Cambridge, 2006).

[131] D. C. Mattis and E. H. Lieb, J. Math. Phys. **6**, 304 (1965).

[132] H. Schulz, Phys. Rev. Lett. **64**, 2831 (1990).

[133] D. W. Wang, A. J. Millis, and S. Das Sarma, Phys. Rev. B **64**, 193307 (2001).

[134] Z. Popovic *et al.*, Synthetic Met. **124**, 421 (2001).

[135] V. Vescoli *et al.*, Science **281**, 1181 (1998).

[136] A. Georges, T. Giamarchi, and N. Sandler, Phys. Rev. B **61**, 16393 (2000).