Psi-k workshop/summer school report

Quantum Monte Carlo in the Apuan Alps VIII

and

Quantum Monte Carlo and the CASINO program VIII

27th July - 11th August 2013,
Apuan Alps Centre for Physics, Vallico Sotto, Italy

M.D. Towler
University of Cambridge and University College, London

1 Summary

Every year since 2005 I have organized an international workshop *Quantum Monte Carlo in the Apuan Alps* and a summer school *Quantum Monte Carlo and the CASINO program* in my private institute - the 'Apuan Alps Centre for Physics' in the village of Vallico Sotto in Italy. The aim of the workshops is to bring together the greatest experts in the field (with associated students and young researchers) to discuss the development and application of the continuum quantum Monte Carlo (QMC) method in condensed matter physics and quantum chemistry. Those who have attended the workshops each year have indicated that they found them to be an intensely valuable experience; indeed, they have essentially been adopted by much of the international community as their 'Annual General Meeting'. Each year the workshop has been immediately followed by a summer school, the purpose of which is to provide students with a thorough working knowledge of the quantum Monte Carlo electronic structure method as currently used in modern research, and to show him or her how to use the Cambridge CASINO QMC program for serious calculations. The course consists of around 20 hours of lectures by recognized experts in the field and a series of practical exercises in using the CASINO program led by its authors.

For summer 2013 we hosted an eighth workshop and school. The workshop was attended by 40 participants (with 5 family members) from 8 countries, and the school by 25 students (with 9 family members and 4 staff) from 12 countries. Both events were generally reckoned to be highly successful.

2 Description of the scientific content of and discussion at the event

QMC is a highly accurate method for calculating the electronic structure of atoms, molecules and materials whose computational cost - unlike most accurate quantum chemistry methods - scales reasonably with the system size (at best, as N^2). It has the further advantage that it can be made to scale essentially linearly with the number of processors on a parallel machine, allowing the efficient utilization of the largest computational facilities available (the CASINO program itself has been run on more than half a million processor cores of the Japanese 'K computer' with essentially perfect efficiency). In computational electronic structure theory it is vitally important to have a 'benchmark method' applicable to real

systems where the results in individual cases can be relied upon, which is not really the case with the cheaper and more commonly used density functional theory (DFT) technique. This is particularly so in condensed-matter systems where most other accurate quantum chemistry methods cannot be easily employed.

QMC deals directly with the full many-body Schrödinger wave function $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$. Despite the additional complexity introduced by the dependence on $3N$ variables, we use this object because we know what equation to solve to get $\Psi(\mathbf{R})$ – the Schrödinger equation. In DFT the ground-state density depends only on three variables but the equation which it satisfies is, by contrast, completely unknown (efforts to find better approximations are vital and ongoing). As a consequence, QMC is capable of giving by far the more accurate energetics. There are two QMC techniques in widespread use – variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) – and there are a number of other more specialized variants. In VMC, we take the wave function as a given analytic form and use Monte Carlo numerical integration to evaluate the energy directly as the expectation value of the Hamiltonian: $\langle\Psi|H|\Psi\rangle/\langle\Psi|\Psi\rangle$. Monte Carlo integration is used as it is the only method that works for high-dimensional integrands (such as the many-body wave function). If necessary we change the shape of the wave function by varying the parameters which define it until the energy or the variance is minimized. In such calculations, Ψ is typically chosen to be the product of a Slater determinant of single-particle orbitals (e.g. from a DFT or Hartree-Fock calculation) multiplied by a *Jastrow factor* which is dependent on the inter-particle distances and allows efficient inclusion of both long- and short-range correlation effects (including van der Waals etc.). The Jastrow also imposes the exact cusp condition, that is, it introduces the correct divergence in the kinetic energy to compensate for the divergence in the potential energy as two electrons coalesce. Usually determinants containing single-particle orbitals from DFT or Hartree-Fock calculations produce quite acceptable trial wave functions.

The accuracy of VMC is ultimately limited by the analytic form chosen, which generally turns out not to have enough variational freedom to represent the true wave function. We thus turn to DMC, which is the ultra-accurate alternative used in most modern calculations. In DMC, Ψ is represented using a *non-analytic* method (the distribution in time and configuration space of an ensemble of diffusing particles). The algorithm encourages the particles to distribute themselves as in the true ground state through the use of a projection technique involving evolution in imaginary time. Once the particles have distributed themselves appropriately, we numerically integrate by sampling the wave function configuration space as before. In principle, this yields the exact ground-state energy.

The DMC technique exploits the similarity of the kinetic part of the Schrödinger equation to a diffusion equation, as in the theory of Brownian particles, while the potential part can be treated as a source or sink of the Brownian particles. With this equivalence, DMC simulates the diffusion, birth and death of these particle configurations or 'walkers', and the approximate Ψ from VMC is used to guide their diffusion, in a manner related to importance sampling. In practice, for many-body fermionic systems, convergence to the lowest energy *antisymmetric* state requires the use of the 'fixed-node approximation' and the method ceases to be exact. Nevertheless, we know from very many comparisons with experiment and quantum chemistry that fixed-node error is often extremely small, and we routinely test this in a variety of calculations. The only other significant approximation comes from the use of pseudopotentials, but again these errors can be estimated.

DMC has major advantages over other possible techniques for high-accuracy calculations: it is ideal for large parallel machines; with blip-function basis sets, convergence to the basis-set limit is rapid and automatic; scaling with system size is mild ($\sim N^3$ or less); and it can be used equally well for molecules in free space or for periodically repeated systems.

Despite such capabilities the technology of QMC remains under intense development; its application to arbitrary finite and periodic systems, particularly those containing heavier atoms, is still far from routine and there are still many open methodological and algorithmic problems to interest the computational electronic structure theorist. Many of these topics were widely discussed at the workshop.

Workshop

The following 40 participants attended the workshop:

Dario Alfè	UCL, London, U.K.
Alberto Ambrosetti	Fritz-Haber-Institut, Berlin, Germany
Sam Azadi	Imperial College, London, U.K.
Anouar Benali	Argonne National Laboratory, U.S.A.
Dario Bressanini	Università dell'Insubria, Italy
Pascal Bugnion	University of Cambridge, U.K.
Mauro Causà	University of Napoli, Italy
Gareth Conduit	University of Cambridge, U.K.
Mike Deible	University of Pittsburgh, U.S.A.
Andrea Droghetti	Trinity College, Dublin, Ireland
Neil Drummond	Lancaster University, U.K.
Edgar Engel	University of Cambridge, U.K.
Elif Ertekin	University of Illinois, U.S.A.
Matthew Foulkes	Imperial College, London, U.K.
Mike Gillan	University College, London, U.K.
Richard Hennig	Cornell University, U.S.A.
Kenta Hongo	JAIST, Japan
Ken Jordan	University of Pittsburgh, U.S.A.
Martin Krupicka	Max Planck Institute, Mulheim, Germany
Jonathan Lloyd-Williams	University of Cambridge, U.K.
Pierre-Francois Loos	Australian National University, Australia
Matthew Lyle	University of Cambridge, U.K.
Pablo López Ríos	Cambridge University, U.K.
Ryo Maezono	JAIST, Japan
Natalia Matveeva	University of Trento, Italy
Bartomeu Monserrat-Sanchez	University of Cambridge, U.K.
Elaheh Mostaani	University of Lancaster, U.K.
Tim Mueller	John Hopkins University, Baltimore, U.S.A.
Richard Needs	University of Cambridge, U.K.
Joshua Schiller	University of Illinois, U.S.A.
Luke Shulenburger	Sandia National Laboratory, U.S.A.
James Spencer	Imperial College, U.K.
Jian Sun	University of Cambridge, U.K.
Alex Thom	Imperial College, U.K.
Alexandre Tkatchenko	Fritz-Haber-Institut, Berlin, Germany
Mike Towler	Cambridge University/UCL, U.K.
Peter Townsend	University of Cambridge, U.K.
Tack Uyeda	JAIST, Japan
Will Vigor	Imperial College, London, U.K.
Lucas Wagner	University of Illinois, U.S.A.

In advance of the workshop I suggested that the following topics would be of particular interest:

(1) Further development of quantum Monte Carlo algorithms in terms of efficiency, speed, functionality, ease of use, scaling behaviour with system size, atomic number and number of processors.

(2) Application of the method to systems and situations where standard less accurate techniques often fail, such as in weakly-interacting systems, strongly-correlated materials, metal-insulator transitions, magnetic properties and biological systems.

(3) The development of better many-electron wave functions and methods of optimization (particularly of parameters which change the nodal surface and thus have some potential to bypass the fermion sign problem).

(4) The calculation of QMC forces and their implementation in dynamical simulation methods. How to exploit the power of QMC in cheaper density functional theory (DFT) and classical-force-field molecular dynamics calculations (in the context of 'embedding methods' or otherwise).

(5) The relationship between QMC and density functional theory (DFT). In particular, I hoped to look towards obtaining a better understanding of the relative benefits of QMC/DFT in systems and situations where DFT often fails I stressed that we needed to look into the better integration of DFT and QMC codes. In industrial applications, where users typically require ease-of-use and a graphical user interface, how can we move towards a situation where the user may do e.g. pre-screening of candidate structures with the faster DFT, then press a 'Make the Answer Better button' which feeds the relevant wave function files into a QMC programme and spits out reliably accurate energetics for the selected problems?

(6) How to use QMC to develop better exchange-correlation functionals for use in DFT.

Given the special nature of the meeting, I therefore warmly encouraged applications to attend from DFT people who had never done any QMC calculations themselves but were nevertheless interested in the topic.

In the end, the participants gave 33 diverse presentations (listed in section 4) on a wide variety of topics, including much of the above.

Summer school

The following 25 students attended the summer school:

Yasmine Al-Hamdani	University College London, U.K.
Tadeusz Andruniów	Politechnika Wroclawska, Poland
Can Ataca	MIT, U.S.A.
Samaneh Ataei	University of Tehran, Iran
Salvatore Cardamone	University of Manchester, U.K.
Ambesh Dixit	Indian Institute of Technology Jodhpur, India
Guillaume Ferlat	University of Paris, France
Laura Giacopetti	University of Cagliari, Italy
Yury Gladush	Institute for Spectroscopy RAS, Troitsk, Russia
Cecilia Goyenola	Linköping University, Sweden
Blazej Jaworowski	Wroclaw University of Technology, Poland
Manuel Perez Jigato	Centre de Recherche G. Lippmann, Luxembourg
Peter Korir	University of Eldoret, Kenya
Matthew Malcomson	University of Lancaster, U.K.
Jana Mathauserová	Charles University in Prague, Czech Republic
Marcos Menéndez	University of Oviedo, Spain
Elaheh Mostaani	University of Lancaster, U.K.
Minh Tam Nguyen	KU Leuven, Belgium
Antonio Noto	University of Palermo, Italy
Eduardo Menendez Proupin	University of Chile, Santiago, Chile
Samuel Ridgway	University of St. Andrews
Kayahan Saritas	MIT, U.S.A.
Pedro Silva	Universidade Fernando Pessoa, Porto, Portugal
Kyrylo Snizhko	University of Lancaster, U.K.
Ludmila Szulakowska	Wroclaw University of Technology, Poland

Over the last few decades we in the Cambridge group have made a great effort to develop a general-purpose computer program - CASINO - and one of the main points of the school is to instruct the students in the use of this software. The code is capable of performing QMC calculations on a wide variety of systems, which may be of finite extent (atoms or molecules) or may obey periodic boundary conditions in one, two or three dimensions, modelling respectively polymers, slabs (or surfaces) and crystalline solids. Following recent modifications this software is ideally placed to exploit the 'petascale computing' hardware with a million or more processors that is now becoming available. Systems containing thousands of atoms per molecule or per cell are thus now within reach of the QMC method using CASINO.

During the mornings, the students attended over 20 hours of lectures given by Neil Drummond, Pablo López Ríos, and Mike Towler. After lunch, they participated in extensive technical exercises involving the use of the CASINO code. Through these practical sessions they learnt (1) how to setup and compile the code on essentially any machine, (2) how to do variational Monte Carlo, diffusion Monte Carlo, and wave function optimization calculations, (3) how to generate trial wave functions from a variety of third-party codes, (4) various statistical analysis techniques, (5) how to work with pseudopotentials, and (5) how to deal with 'finite-size effects'.

On the final day, the students took a technical examination aimed at testing their newly-acquired knowledge of quantum Monte Carlo. A prize was awarded for the highest mark, and this was ultimately achieved by Blazej Jaworowski of Wroclaw University of Technology in Poland. This was followed by an 'advice session' where each student was asked to outline why he or she had attended the school and to explain their particular interest in quantum Monte Carlo stating what kinds of systems they wished to study. In most cases, the lecturers were able to give detailed recommendations as to the best way to perform these

calculations. An opportunity was also provided for students to give short talks about their research.

3 Assessment of the results and impact of the event on the future direction of the field

Workshop

The vitality of research in this field was confirmed by the extremely wide variety of applications discussed, which included high-pressure solid hydrogen, rare gas crystals, biomolecules, ultracold atomic gases, water clusters, water-acene complexes, the design of metal alloys, hexagonal boron nitride, point defects in semiconductors and wide-band gap oxides, molecular crystals such as para-diiodobenzene, proton disorder in ice, nanocrystalline alumina, electron-hole systems, the metallization of solid helium, the binding energy of bilayer graphene, correlation-bound anions and organic diradicals, methane binding in ice clathrates, and (for the first time) the high-T_c superconducting cuprates. Wagner's talk on this latter subject - in which he showed highly-accurate DMC calculations of the spin-lattice coupling in the *real material* (as opposed to a model system) for the first time - was particularly interesting. He found that some lattice degrees of freedom depended strongly on the magnetic state, and that the spin-lattice coupling was removed with 25% doping. Considerable attention was also paid to computing weak interactions such as van der Waals in QMC; talks by Tkatchenko, Benali, Ambrosetti, Deible, Gillan, Jordan, Mostaani were all devoted to this topic in one or another, and showed the very considerable superiority of QMC results over DFT in this area. As the power of available computers and the sophistication of the algorithms and software increases, it is clear that the size and complexity of the systems that can be treated within QMC continues to rise rapidly. A number of researchers, including Gillan, Wagner, Maezono, Towler, and Alfè are particularly active in pushing the boundaries of what calculations can be done for real systems on large computers.

A number of more technical questions were addressed. These included forces and correlated sampling, finite-size effects, direct incorporation of light nuclei at the quantum level in DMC calculations, maximum probability domains in crystals, and the use of CASINO on systems containing GPUs. A number of speakers such as Tim Mueller stressed the need for using QMC to generate publically-available databases of highly accurate results - for examples of formation energies - and the subsequent requirement for developing methods for doing 'high-throughput QMC'. QMC is ideal for such databases - which need only be created once and would be useful for a very long time - not only because it produces highly accurate energetics, but because it scales well with processor number and system size and because it works for 'everything' (molecules, metals, insulators, semiconductors..). Mueller's talk had the interesting and surprising conclusion that - given suitable recipes and workflow improvements -'*by 2016-ish we should be able to calculate QMC energies for every known inorganic material on a single supercomputer in (roughly) about a week*'.

A significant number of people in groups in London and Cambridge are involved with stochastic Monte-Carlo approaches to classic quantum chemistry techniques such as CI and coupled cluster, and these efforts were the subject of a number of interesting talks by Spencer, Thom, and Vigor. The FCI-QMC technique can reproduce the results of full CI calculations (for a given finite basis) essentially exactly but in considerably less time, and it is now possible to calculate energies for systems very much larger than was possible before. The stochastic coupled cluster technique discussed by Thom can reproduce large

coupled cluster calculations much more quickly; it is considerably simpler to implement than deterministic coupled-cluster, and is feasible on workstations and very parallelizable.

Quantum Monte Carlo in the Apuan Alps VIII was a fascinating meeting showcasing the health and vigour of the field, and the talent and dynamism of many of the people working in it.

Summer school

The Vallico Sotto summer school programme has evolved over many years and is designed to address the problem of lack of experienced users of quantum Monte Carlo in general and of the CASINO program in particular. Over the years we have trained over 200 people in the use of this important technique.

Comments received from almost all the participants after the school made it clear that on the whole the students greatly enjoyed the course and found it extremely beneficial. In particular they expressed the opinion that the course had the right balance between formal lectures, technical exercises, and social interaction (the students are strongly encouraged to take time to develop relationships which might form the basis of future collaborations).

On departure the students were provided with a copy of the CASINO code, and in the last couple of months I (MDT) have been in continuous email contact with a good proportion of the students who are now engaged in their first formal research projects in quantum Monte Carlo.

The interest shown in the school and the total number of students exceeded expectations from former years; interest in QMC is clearly becoming more widespread and that is in no small measure due to the assistance provided by Psi-k/ESF in these endeavours..

4 Final programme of the workshop

The slides for almost all the talks presented have been made freely available on the Institute web site: `www.vallico.net/tti/qmcitaa_13`. In participant alphabetical order, the talks given were:

Dario Alfe (`d.alfe at ucl.ac.uk`)
'*Graphene on Ir(111): growth and thermodynamics from combined experimental and theoretical methods*'

Alberto Ambrosetti (`ambrosetti at fhi-berlin.mpg.de`)
'*Long range correlation energy from isotropically damped quantum harmonic oscillators*'

Sam Azadi (`s.azadi at imperial.ac.uk`)
'*Quantum Monte Carlo study of high-pressure solid hydrogen*'

Anouar Benali (`abenali at alcf.anl.gov`)
'*Quantum Monte Carlo calculations of many-body van der Waals forces in rare gas crystals and biomolecules*'

Pascal Bugnion (`pob24 at cam.ac.uk`)
'*Spins, superfluidity, and ultracold atomic gases*'

Mauro Causà (`mauro.causa at unina.it`)
'*Correlated maximum probability domains in crystals*'

Gareth Conduit (`gjc29 at cam.ac.uk`)

'*Concurrent materials design*'

Mike Deible (mjd87 at pitt.edu)
'*Quantum Monte Carlo studies of water clusters and water-acene complexes*'

Andrea Droghetti (drogheta at tcd.ie)
'*A DFT+model Hamiltonian approach to zero-bias transport in nanostructures: work in progress*'

Neil Drummond (n.drummond at lancaster.ac.uk)
'*Electronic structure of two-dimensional crystals of hexagonal boron nitride*'

Elif Ertekin (ertekin at illinois.edu)
'*QMC for point defects in semiconductors and wide-band gap oxides*'

Matthew Foulkes (wmc.foulkes at imperial.ac.uk)
'*Forces and correlated sampling in DMC*'

Mike Gillan (m.gillan at ucl.ac.uk)
'*Quantum Monte Carlo benchmarks for weak non-covalent interactions*'

Richard Hennig (rhennig at cornell.edu)
'*Computational discovery and design of materials for energy technologies and electronic devices*'

Kenta Hongo (kenta_hongo at mac.com)
'*Finite-size effects in diffusion Monte Carlo simulations of para-diiodobenzene*'

Ken Jordan (jordan at imap.pitt.edu)
'(1) *Quantum Monte Carlo studies of correlation-bound anions and of organic diradicals*'
'(2) *Proton disorder in ice*'

Pierre-Francois Loos (loos at rsc.anu.edu.au)
'*Uniform electron gases: electrons on a ring*'

Matthew Lyle (mjl78 at cam.ac.uk)
'*Low density nanocrystalline alumina*'

Ryo Maezono (rmaezono at mac.com)
'*Studies of electron-hole systems using DMC*'

Natalia Matveeva (matveeva.na at gmail.com)
'*Localization of an impurity in a bilayer system of dipoles*'

Bartomeu Monserrat-Sanchez (bm418 at cam.ac.uk)
'*White dwarf cooling: electron-phonon coupling and the metallization of solid helium*'

Elaheh Mostaani (emostaani at gmail.com)
'*Binding energy of bilayer graphene*'

Tim Mueller (tmueller at jhu.edu)
'*Quantum Monte Carlo for materials design*'

Richard Needs (rn11 at cam.ac.uk)
'*Decomposition and terapascal phases of water ice*'

Luke Shulenburger (lshulen at sandia.gov)
'*Status of DMC for condensed phases*'

James Spencer (j.spencer at imperial.ac.uk)

'*Full configuration interaction quantum Monte Carlo and coupled cluster Monte Carlo: a framework for stochastic quantum chemistry.*'

Alex Thom (`a.thom at imperial.ac.uk`)
'*Linked stochastic coupled cluster theory*'

Alexandre Tkatchenko (tkatchen at fhi-berlin.mpg.de)
'*Explicit many-body van der Waals corrections to DFT, and how QMC can help to develop them*'

Tack Uyeda (`tueeeda at jaist.ac.jp`)
'*Ohmic contact on diamond semiconductors*'

William Vigor (`w.vigor11 at imperial.ac.uk`)
'*Accelerating Full Configuration Interaction Quantum Monte Carlo*'

Anatole von Lilienfeld (`anatole at alcf.anl.gov`)
'*Preaching on first principles views on chemical compound space, atom-centered potentials, and statistical learning*'

Lucas Wagner (`lkwagner at illinois.edu`)
'*Can we understand the high-Tc superconducting cuprates from first principles?*'

5 Final programme of the summer school

The summer school featured the following lecture topics, in addition to practical afternoon exercises:

Mike Towler (University of Cambridge, U.K.)

- '*Quantum Monte Carlo : a practical solution to the correlation problem in electronic structure calculations*'

- '*The* CASINO *program : a basic introduction to functionality and input/output*'

- '*Three QMC scaling problems: many atoms, many protons, many processors*'

- '*Forces and dynamics. Expectation values other than the energy*'

Neil Drummond (University of Lancaster, U.K.)

- '*Diffusion Monte Carlo*'

- '*Optimization of many-electron wave functions*'

- '*Ewald interactions and finite size errors*'

- '*Some recent applications of quantum Monte Carlo simulation*'

Pablo López Ríos (University of Cambridge. U.K.)

- '*Statistics in quantum Monte Carlo*'

- '*Wave functions beyond Slater-Jastrow for QMC*'

- '*Pseudopotentials for QMC*'

Exercises

- '*Distribution, setup and compilation of the* CASINO *program*'

- '*Basic use of the* CASINO *program - simple VMC, DMC calculations*'

- '*Wave function optimization with* CASINO'

- '*Trial wave function generation with other programs (*CRYSTAL/PWSCF/CASTEP *etc.)*'

- '*Advanced use of the* CASINO *program*'

- '*General* CASINO *applications*'

More informal details of both events, together with talk slides, photographs, and comments by participants can be found on the Institute web site (`www.vallico.net/tti/qmcitaa_13` and `www.vallico.net/tti/qmcatcp_13`).