Quantum Hall effect and high performance computing

Short Visit Grant Reference 2129:
Gunnar Möller, University of Cambridge
to visit
Nicolas Regnault, Ecole Normale Supérieure Paris
26.11.2007–30.11.2007

Report

Purpose of the visit

The core purpose of this visit consisted in discussing a roadmap for the future evolution of a numerical library, *DiagHam*, that the above collaborators maintain. With the ongoing trend towards a stronger focus on parallel architectures in high-performance computing facilities, our existing approaches to the core task of this library, diagonalising Hamiltonians, needed a review. Underlying motivations arise from the development of computing hardware: the number of available CPU's is rapidly increasing and new potential for processor power arises from the use of Graphics Processor Units. However, the accessibility of memory remains a limiting factor, in particular for the type of calculation *DiagHam* is designed to perform. Below, I outline the strategy we devised during my visit of how to best adapt our software to these new imperatives of parallel computer architectures.

Work carried out during the visit

Minimizing memory requirements The core component of *DiagHam* relies on an implementation of the Lanczos algorithm which is well suited to find the lowest-lying eigenvalues of large matrices and relies on the repeated multiplication of that matrix onto a sequence of vectors. In the current parallelisation strategy of DiagHam, state vectors are distributed in pieces of equal length to perform multiplication on multiple processors. The result of this multiplication is in principle a vector of length of the full Hilbert-space dimension, and each sub-process currently requires a copy of such a vector to prevent concurrential memory-access. For very large matrices, the storage of the state-vectors in this form exhaust the available memory. We designed the following approach to prevent this problem: Each sub-process should hold a virtual output vector in a sparse representation, which is really a buffer which stores the operations performed on the vector. A single master process could then be implemented to collect all the modifications, and perform them on a single copy of the vector. In order to minimise possible delays from communications, a dual buffer can be implemented that allows a process to write one part of it while the master process can read the other.

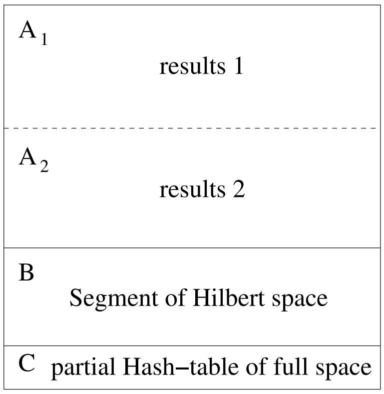

Figure 1: Memory allocation on the graphics card: Two results buffers avoid problems of concurrential access, each result buffer is 1.5 times the initial data in the segment of Hilbert-space analyzed.

Calculation on Graphics Cards A tremendous opportunity for numerical calculation lies in the recent progress of the abilities of graphics processor units (GPU), as these provide up to tenfold computing power per dollar compared to CPU's. However, their instruction set is limited, and GPU's are ideally suited to performing a given operation repeatedly on many instances of data. For large scale exact diagonalization calculations, the actual Hamiltonian matrix cannot be stored. Instead, its elements have to be recalculated at each iteration. We designed an algorithm for the implementation of this calculation aided by GPU's. Quantum states are represented in the second quantized notation $|\alpha\rangle = \prod_k (a_k^\dagger)^{n_k(\alpha)}|vac\rangle$, and two particle interactions can be expressed as sums of operators $a_{m1}^\dagger a_{m2}^\dagger a_{m3} a_{m4}$. A graphics card would be charged to apply this operator to a sequence of basis-states $|\alpha\rangle$ corresponding to a fraction of the Hilbert-space. The memory of the graphics card would be distributed as shown in Fig. 1. The calculation would proceed as follows:

1. Communicate segment of Hilbert-space and an equilibrated partial Hash-table of the full Hilbert-space to the graphics card
2. replicate data of Hilbert-space segment, interlacing states with 32 bit integers to hold results.
3. calculate application of the given operator onto the basis states (parallelised over multiple GPU's present on graphics card), test if resulting vector satisfies constraints of Hilbert-space, the sign from operator permutations, and eventually a partial search of the resulting state using the Hash-table.

4. When all states have been processed, mark first buffer for reading proceed calculation of next operator using second buffer.

Discrete symmetries The use of discrete symmetries in the representation of states is yet another tool that can be used to reduce the size of memory required for exact diagonalisations. We worked on the implementation of such symmetries for the analysis of Fermions on the sphere with spin, in particular focusing on total spin reversal. Similar techniques can also be used for the torus.

Physics of rotating Fermi gases While the above aspects of our work focused on technical points of the implementation of numerical calculations, this visit also gave us the opportunity to discuss the results of a specific project on Fermions with spin 1/2 that is based on these tools. In particular, we discussed the nature of incompressible states present in a simple model with s- and p-wave interactions of Fermions.

Main results obtained

The main achievement of this visit is the roadmap for future extensions of *DiagHam*. An in-depth discussion of these points has given us a clearer vision of how to proceed. While we outlined the mechanism for more efficient parallelisation of exact diagonalisations, implementation of these ideas will require further work.

In our discussion about incompressible quantum liquids of Fermions with spin, we developed a new candidate wavefunction for a spin-singlet state found at filling factor $\nu = 2/3$.

Future collaboration with host institution

As mentioned above, our project entails a large quantity of work to implement the ideas developed during my visit to Paris, which will feed into my ongoing collaboration with N. Regnault. At the occasion of this visit, I also had the opportunity to discuss with Th. Jolicoeur, M. Goerbig, W. Krauth, S. Ouvry, and K. Wiese, which strengthened my links to the physics community in Paris, including laboratories at ENS as well as Université Paris-Sud. It was envisioned that B. Chung, Ph.D. student of Th. Jolicoeur could be more involved into our efforts and contribute to new collaborative projects.

Projected publications

Our results on incompressible spin-singlet states in rapidly rotating cold Fermi gases will be published in the near future.