Task 2

The purpose of the paper is to introduce a new massively parallel version of CRYSTAL, a periodic ab initio code that allows for efficient treatment of very large unit cells for crystalline systems on HPC architectures. The aim is to provide end-users with criteria for deciding the amount of resources to be allocated for a calculation and to facilitate tailoring of in-house computing facilities based on commodity PC-based clusters as a function of the system size of interest for a given research group.

The paper reports on the parallel performance of the CRYSTAL ab initio periodic program when running on HPC architectures for the study of complex materials of great importance. It is also the most complete memory usage and performance analysis of the program to date.

The structure of the CRYSTAL code permits the study of systems periodic in one-dimension (1D) (polymers), 2D (slabs), and 3D (crystals). A Gaussian type basis set is used, and both density functional theory (DFT) and Hartree–Fock (HF) approximations are available. The expansion coefficients of the Bloch functions are calculated by solving the usual matrix equation. Four-center bielectronic (as well as monoelectronic) integrals are evaluated analytically in a very efficient way.

The new massively parallel version of CRYSTAL is very well-behaved as far as scalability with the number of cores at fixed system size is concerned. The paper shows that systems as large as X14 (a 1x1x14 MCM-41 supercell) can be run on three different MPP architectures, without memory overflows or any other technical problem, with CPU costs that remain reasonable when a thousand-processor machine is available.

The main limitation of the CRYSTAL code remains the quadratic scaling with the system size that emerges with cells containing more than 200 irreducible atoms. To solve this problem, major interventions on the code are needed, particularly revisiting the loop structure in the main routines.

No standard libraries are available to perform a distributed sparse matrix by matrix multiplication, a key operation in any quantum mechanical program. This last point would require a large amount of human work to reach the final target of linear scalability with respect to both the system size and the number of cores.

The new massively parallel version of CRYSTAL opens up new possibilities for studying complex solid systems. With recent improvements, density functional calculations of low-symmetry systems containing up to 100,000 atomic orbitals and 8000 atoms are now feasible on the most advanced HPC architectures available to European researchers today. However, the limitations of the code, particularly the quadratic scaling with the system size, need to be addressed to achieve linear scalability with respect to both the system size and the number of cores.