

PARALLEL IMPLEMENTATIONS OF PARAMETRIC IDENTIFICATION ALGORITHMS FOR THREE-DIMENSIONAL CRYSTAL LATTICES

D.V. Kirsh^{1,2}, A.V. Kupriyanov^{1,2}

¹Image Processing Systems Institute – Branch of the Federal Scientific Research Centre “Crystallography and Photonics” of Russian Academy of Sciences, Samara, Russia

²Samara National Research University, Samara, Russia

Abstract. The article describes parallel implementations of two parametric identification algorithms. The first algorithm is based on the estimation of Bravais unit cell parameters, and the second one estimates Wigner-Seitz cell volumes. The developed parallel implementations are based on the two-tier model of concurrency. An external tier of concurrency uses Message Passing Interface (MPI) technique to distribute parametric identification tasks between computing nodes. An internal tier uses Open Multi-Processing (OpenMP) technique to split single steps of the parametric identification algorithms into independent subtasks. Experimental results on multi-processor / multi-core systems have demonstrated almost linear speedup of the parallel implementations, confirmed the effectiveness of the two-tier concurrency model and proved the stability of the identification approach under lattice distortion.

Keywords: crystal lattice, unit cell, parametric identification, parallel algorithm, MPI, OpenMP, scalability.

Citation: Kirsh D, Kupriyanov A. Parallel implementations of parametric identification algorithms for three-dimensional crystal lattices. CEUR Workshop Proceedings, 2016; 1638: 451-459. DOI: 10.18287/1613-0073-2016-1638-451-459

Introduction

Structural identification of crystal lattices in three-dimensional space is one of the basic problems of X-ray diffraction analysis [1]. The main difficulties are the reconstruction of three-dimensional objects [2, 3] and the ambiguity of unit cell choice from which the ambiguity of lattice system identification arises [4, 5].

To solve the problem of structural identification in three dimensional space, we proposed two approaches: the first was based on parametric identification methods [6] and the second was based on application of fuzzy neural networks [7]. According to conducted experiments, the first approach proved to be significantly more universal than the second one. In the beginning, it estimates a number of parameters: edge

lengths and angle values of Bravais unit cells [6] and volumes of Wigner-Seitz cells [8]. Then the proposed measures calculate the similarity between estimated parameters and the set of reference ones. However, thorough research showed that structural identification of high accuracy (above 95%) requires a huge database of reference lattices [9].

Therefore, the reverse side of the proposed approach is its high computational complexity. To overcome this drawback, we decided to develop the parallel implementations of parametric identification algorithms that would use the computing power of modern multi-processor / multi-core systems with maximum efficiency.

The main requirement for the developed parallel implementations was scaling efficiency. Consequently, it was necessary to provide a large number of independent computing tasks that could be executed in parallel. To solve the problem of scaling, we proposed a two-tier concurrency model, described in the article.

External concurrency tier of parametric identification algorithms

The developed methods of crystal lattice parametric identification estimate and compare parameters of the unit cells that form the entire lattice. It is obvious that these methods have a high degree of locality: the initial data they require are coordinates of a few, the most closely located, lattice points (Fig. 1.).

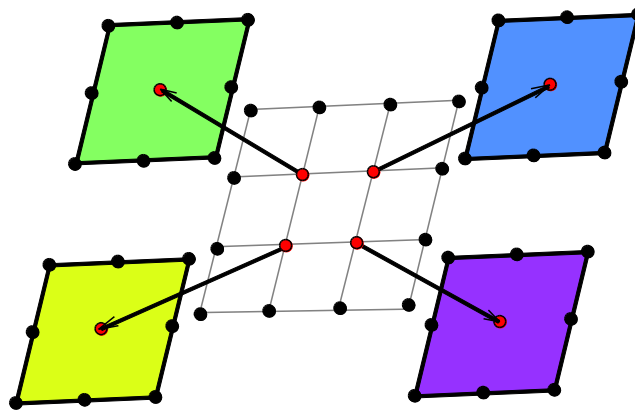


Fig. 1. Splitting a two-dimensional lattice into unit cells areas

It should be noted that the estimation of parameters for each unit cell can process independently. The necessity of data exchange appears only at the final stage, when the analysis of unit cell parameters and the calculation of the entire lattice parameters take place.

Hence, the proposed allocation of parameter estimation tasks called the external tier of concurrency is universal for all parametric identification method. It can be described by the following algorithm:

1. Find all lattice points that do not belong to any lattice face – the step executes on the master node.
2. Scatter uniformly the found lattice points in conjunction with the nearest points that lies inside the sphere of a given radius among the computing nodes.
3. Calculate parameters for each set of lattice points – the step executes by all computing nodes in parallel.
4. Gather all calculated parameters and estimate final ones for the entire lattice – the step executes on the master node.

It is obvious that the described algorithm of parallel implementation corresponds to the MPI technique [10]: computations are local, communications between tasks are minimal and appear only at the initial and final stages.

The proposed approach allows to avoid unbalanced load distribution (the number of tasks on each computing nodes will not differ by more than one task).

Internal concurrency tier of parametric identification algorithms

However, in case of using a large amount of multi-core processors, the number of independent tasks that calculate unit cell parameters may be insufficient to load all available computing cores. Due to the lack of independent tasks on the external concurrency tier, we decided to organize an additional internal tier of concurrency for both parametric identification method by parallelizing the steps of parameter estimations.

Parallel implementation of the Bravais unit cell parameter estimation

An algorithm of the Bravais unit cell parameter estimation is a sequence of steps that must be carried out strictly one by one [6]. Consequently, the concurrency can be organized only within the steps.

The computational complexity of the algorithm steps for a lattice with N lattice points are as follows:

1. Select the first nearest node – $O(N)$.
2. Exclude nodes that lie on the same line with the selected node – $O(N)$.
3. Select the second nearest node – $O(N)$.
4. Exclude nodes that lie in the same plane with the two selected nodes – $O(N)$.
5. Select the third nearest node – $O(N)$.
6. Calculate the parameters of the Bravais unit cell – $O(1)$.

The first five steps have the maximum computational complexity. Consequently, the parallel execution of these steps will allow to achieve the maximum scalability. Figure 2 demonstrates the scheme of the proposed parallel implementation.

Parallel implementation of the Wigner-Seitz unit cell volume estimation

An algorithm of the Wigner-Seitz unit cell volume estimation is a sequence of steps that must be carried out strictly one by one [8]. As for the previous algorithm, the

concurrency can be organized only within the steps. Figure 3 demonstrates the scheme of the proposed parallel implementation.

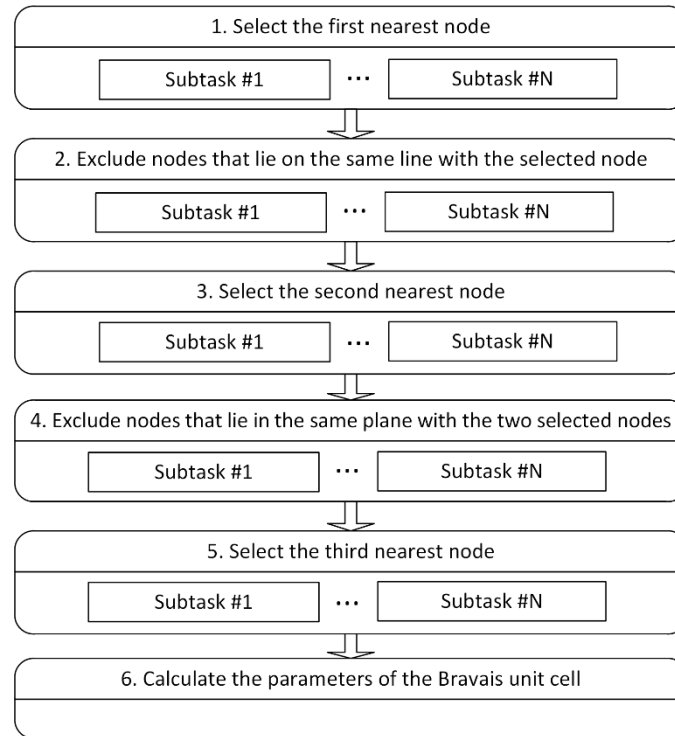


Fig. 2. Scheme of the parallel implementation of the Bravais unit cell parameter estimation

The computational complexity of the algorithm steps for a lattice with N lattice points and L scattered points that is used to calculate the cell volume are as follows

1. Calculate normal vectors to the boundary planes – $O(N)$.
2. Generate L values of uniformly distributed vector in 3D – $O(L)$.
3. Count the number of points that fell inside the boundary planes – $O(L \times N)$.
4. Calculate the volume of the Wigner-Seitz unit cell – $O(1)$.

Due to the fact that the second and the third steps have the greatly higher computational complexity (the minimum number of scattered points to ensure a sufficient level of volume calculation accuracy is 100 000 points), the parallel execution should be organized within these steps, leaving other steps sequential.

Common features of the developed parallel implementations of the both algorithms are data dependence and a large number of communications at each step and between the steps. Thus, the most appropriate model for the parallel implementation is an OpenMP model [9].

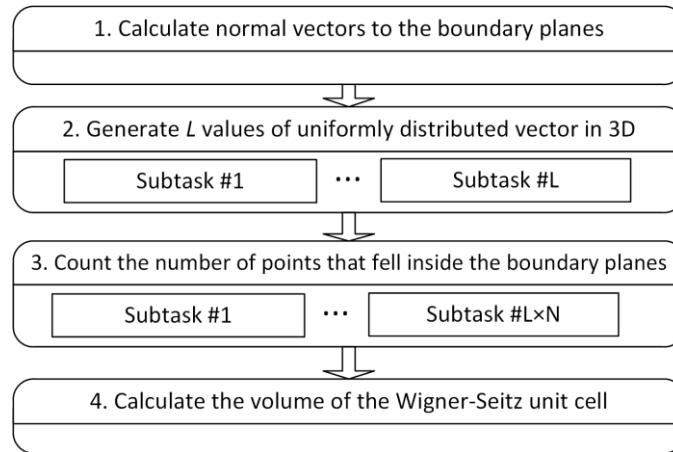


Fig. 3. Scheme of the parallel implementation of the Wigner-Seitz unit cell volume estimation

Experimental results

The aim of this experiment was an investigation of the dependence between the speedup of the developed parallel implementations and the number of computing nodes / identification tasks.

The speedup obtained using a parallel implementation for p cores compared with a sequential algorithm was determined by the following equation [11]:

$$S_p = \frac{T_1}{T_p}, \quad (1)$$

where T_1 – the execution time of a sequential algorithm;

T_p – the execution time of a parallel implementation on p cores.

According to the equation (1), the best value of the speedup is a linear speedup when the execution time of parallel implementation grows linearly with the increase of core number.

A two triclinic lattices were used as an analyzed lattices: small (s) and big (b). Lattice sizes were chosen in order to provide the following time of parameter estimation computing on a single core: about 1 second for the small lattice; about 5 minutes for the big one. This selection allows to estimate speedup at a small number of concurrent tasks when the overhead of parallel computing prevails, as well as predict the speedup growth with the increase of the lattice size, the number of cores and the task complexity.

Investigation of the speedup on multi-core personal computer

The experiment was conducted on a desktop computer based on a quad-core processor Intel Core i5 3470 (base frequency 3.8 GHz).

Lattice size for each parametric identification method were as follows:

- the method of Bravais unit cell parameters estimation – 24 and 51 translations in each direction;
- the method of Wigner-Seitz unit cell volume estimation – 8 and 17 translations in each direction.

The obtained data are presented in the form of then diagram showing the speedup of the parallel implementations for different numbers of cores (Fig. 4).

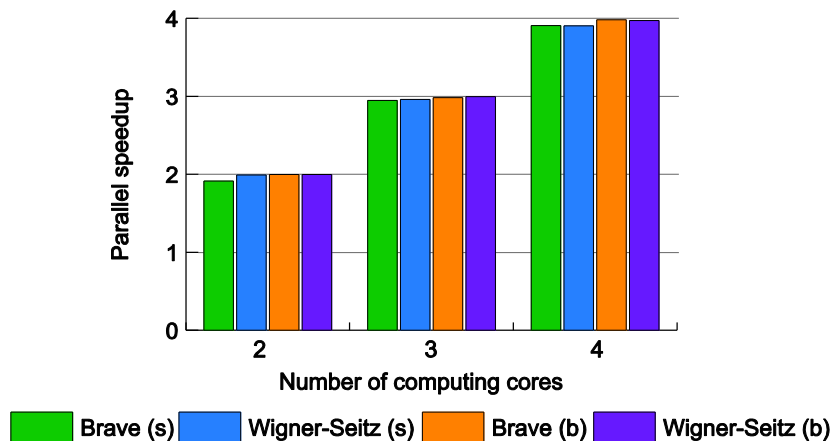


Fig. 4. Speedup on the multi-core personal computer

The results confirm that the parallel implementations of both parametric identification methods provide almost linear speedup over the whole range of used lattice sizes and core configurations. The difference between the obtained values of speedup for small and big lattice is less than 5%.

Investigation of the speedup on multiprocessor / multicore cluster system

The experiment was conducted on 16 cluster nodes of supercomputer “Sergey Korolev”. Each node contained two quad-core Intel Xeon X5560 processor (base frequency 2.80 GHz).

Lattice size for each parametric identification method were as follows:

- the method of Bravais unit cell parameters estimation – 19 and 37 translations in each direction;
- the method of Wigner-Seitz cell volume estimation – 4 and 13 translations in each direction.

It should be noted that the lattice of 4 translations cannot provide a sufficient number of parameter calculation tasks that is necessary for the uniform loading of all computing cores of the cluster (only 27 parallel tasks). In this case, the internal concurrency tier begins to play the leading role balancing the processing load on each core.

The obtained data are presented in the form of the diagram showing the speedup of parallel implementations for different numbers of computing nodes (Fig. 5).

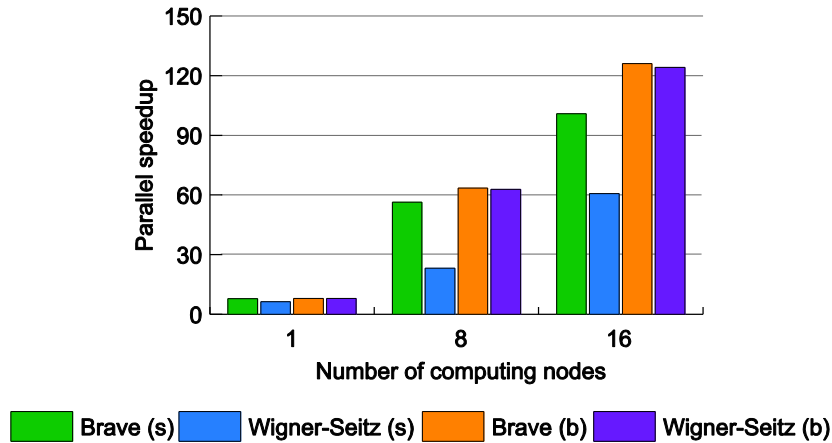


Fig. 5. Speedup on the multiprocessor / multicore cluster system

From the presented diagram it can be seen, the parallel implementations provide almost linear speedup on a single octa-core node of the cluster. These results indirectly confirm the results from the previous section. As expected for the small lattices, the worst speedup was shown by the parallel implementation of the Wigner-Seitz cell volume estimation. However, even in the case of small lattice, the parallel implementation provides the speedup significantly higher than the number of the cluster nodes. This result confirms the effectiveness of the two-tier model of concurrency.

Investigation of the stability under lattice distortion

An ideal lattice is only a mathematical simplification, since real lattices always contain some dislocations and impurities in their structure. This fact can crucially affect the accuracy of parameter estimation, especially if we analyze not the whole lattice structure but only a single unit cell. The example of a distorted lattice and incorrectly estimated parameters of Brave and Wigner-Seitz unit cells are shown in Figure 6.

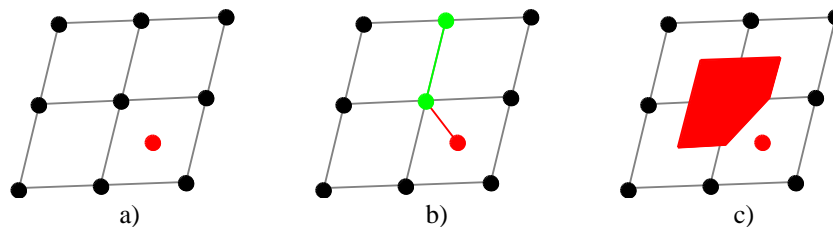


Fig. 6. Effect of a lattice distortion on the parameter estimation:
 a) lattice with random impurity in the structure; b) incorrect estimation of translational vectors;
 c) incorrect construction of a Wigner-Seitz cell

In the last experiment, we investigated the stability of the proposed parameter estimation algorithms under lattice distortion implemented by randomly added/removed lattice points in the lattice structure. A cubic lattice contained 1000 lattice points (10 translations in each direction) was a sample. The number of randomly added/removed lattice points varied from 10 to 100 points. The proposed parallel algorithms was used to estimate unit cell parameters whereas the similarity measures determined the accuracy of corresponding parameter estimations. Results of the experiment are presented in Figure 7.

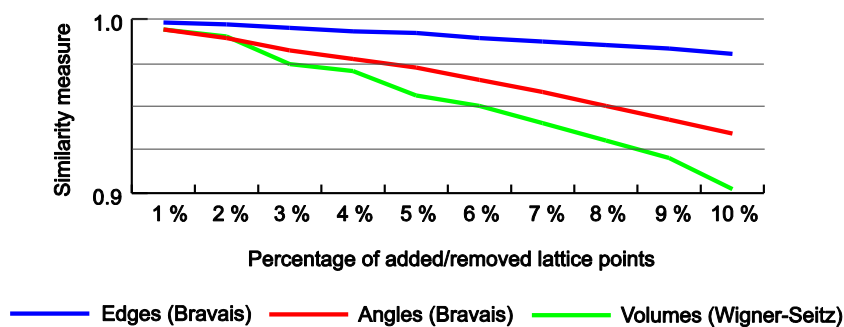


Fig. 7. Stability of the developed parameter estimation algorithms under lattice distortion

According to the graph, the edge similarity showed the maximum stability. Its value did not fall below 0.98 at 10 % of added/removed lattice points. On the contrary, the volumes similarity measure demonstrated the lowest stability among other measures. However, even at the maximum level of distortion, its value was more than 0.9. The obtained results prove high stability of the developed parameter estimation algorithms to random distortion of the lattice structure.

Conclusion

The developed parallel implementations of parametric identification algorithms for crystal lattices in three-dimensional space effectively combine the two basic approaches to parallel programming: MPI at the external concurrency tier and OpenMP at the internal concurrency tier.

The experiments have shown that the developed parallel implementations achieves almost linear speedup on both a desktop computer and a cluster. Therefore, the use of a two-tier concurrency model allows to load the computing powers of multiprocessor / multicore systems with maximum efficiency.

In addition, the developed parallel implementation helped to prove the stability of the proposed approach of crystal lattice parameter estimation to random distortion of the lattice structure.

Acknowledgements

This work was partially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of SSAU among the world's leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 14-07-97040, # 15-29-03823, # 15-29-07077, # 16-57-48006); by the ONIT RAS program # 6 "Bioinformatics, modern information technologies and mathematical methods in medicine" 2016.

References

1. Tilley R. Crystals and crystal structure. West Sussex: John Wiley & Sons Ltd, 2006: 17-32.
2. Fursov VA, Goshin YeV. Information technology for digital terrain model reconstruction from stereo images. *Computer Optics*, 2014; 38(2): 335-342. [in Russian]
3. Kotov AP, Fursov VA, Goshin YeV. Technology for fast 3d-scene reconstruction from stereo images. *Computer Optics*, 2015; 39(4): 600-605. [in Russian] DOI: 10.18287/0134-2452-2015-39-4-600-605.
4. Shirokanev AS, Kirsh DV, Kupriyanov AV. Researching methods of reconstruction of three-dimensional crystal lattice from images of projections. *CEUR Workshop Proceedings*, 2015; 1490: 290-297. DOI: 10.18287/1613-0073-2015-1490-290-297.
5. Kirsh DV, Kupriyanov AV. Modeling and Identification of Centered Crystal Lattices in Three-Dimensional Space. *CEUR Workshop Proceedings*, 2015; 1490: 162-170. DOI: 10.18287/1613-0073-2015-1490-162-170.
6. Kupriyanov AV, Kirsh DV. Estimation of the Crystal Lattice Similarity Measure by Three-Dimensional Coordinates of Lattice Nodes. *Optical Memory & Neural Networks (Information Optics)*, 2015; 24(2): 145-151.
7. Soldatova OP, Lezin IA, Lezin IV, Kupriyanov AV, Kirsh DV. Application of fuzzy neural networks for defining crystal lattice types in nanoscale images. *Computer Optics*, 2015; 39(5): 787-795. DOI: 10.18287/0134-2452-2015-39-5-787-794.
8. Kirsh DV, Kupriyanov AV. Identification of Three-Dimensional Crystal Lattices by Estimation of Their Unit Cell Parameters. *CEUR Workshop Proceedings*, 2015; 1452: 40-45.
9. Kirsh DV, Kupriyanov AV. Crystal lattice identification by coordinates of their nodes in three dimensional space. *Pattern recognition and image analysis*, 2015; 25(3): 456-460.
10. Antonov AS. Techniques of parallel programming MPI and OpenMP. Moscow: Publisher of Moscow State University, 2012. 344 p. [in Russian]
11. Gergel VP. Theory and Practice of Parallel Programming. Moscow: Internet University of Information Technologies "BINOM", 2007: 423 p. [in Russian]