Offer #2022-04982

PhD Position F/M Scaling the solving of Ordinary Differential Equation for Computational Biology (and Deep Learning)

Contract type : Fixed-term contract

Level of qualifications required : Graduate degree or equivalent

Fonction : PhD Position

About the research centre or Inria department

The Inria research centre in Lyon (previously the Lyon branch of the Inria centre in Grenoble) is the 9th Inria research centre, formally created in December 2021. It brings together approximately 270 people (including 110 Inria employees) in 15 research teams and research support services.

Its staff are distributed at this stage on 2 campuses: in Villeurbanne La Doua (Centre / INSA Lyon / UCBL) on the one hand, and Lyon Gerland (ENS de Lyon) on the other. A third site should be opened in the course of 2022. The teams are mainly hosted with our partners. The centre's teams work closely with research and higher education institutions (ENS de Lyon, UCBL, INSA Lyon, etc.), their laboratories, and other research organisations in Lyon (CNRS, INRAE, competitiveness clusters, etc.), but also with Lyon and regional economic players. Many international collaborations are also underway.

The Lyon centre is active in the fields of software, distributed and high-performance computing, embedded systems, quantum computing and privacy in the digital world, but also in digital health and computational biology.

Context

In biology, the vast majority of systems can be modeled as ordinary differential equations (ODEs). Modeling more finely biological objects leads to increase the number of equations. Simulating ever larger systems also leads to increasing the number of equations. Therefore, we observe a large increase in the size of the ODE systems to be solved. A major lock is the limitation of ODE numerical resolution software (ODE solver) to a few thousand equations due to prohibitive calculation time. The AEx ExODE tackles this lock via 1) the introduction of new numerical methods that will take advantage of the mixed precision that mixes several floating number precisions within numerical methods, 2) the adaptation of these new methods for next generation highly hierarchical and heterogeneous computers composed of a large number of CPUs and GPUs. For the past year, a new approach to Deep Learning has been proposed to replace the Recurrent Neural Network (RNN) with ODE systems. The numerical and parallel methods of ExODE will be evaluated and adapted in this framework in order to improve the performance and accuracy of these new approaches.

Description of the thesis

Ordinary differential equations (ODEs) are a major modeling formalism in biology (see below for more details). These examples often result in systems of N strongly connected equations, which gives time this computation in $O(N^2)$:

$$y0(t) = f(t, y(t)), t > 0, y(0) = y0, y(t) \ \square\ RN, f : R \times RN \rightarrow RN.$$

These models are often limited to a few thousand equations because of the computing time and memory usage required for their resolution. The ability of ODE solver to solve in a reasonable time becomes critical because of the need for scalability ie a significant increase in the size of the modeled systems: 100 to 1000 times more equations (an increase in complexity from 10^4 to 10^6).

ODE solvers have characteristics of precision and stability. A scheme may be stable, but not accurate (eg implicit Euler) or accurate but not stable (high order Runge-Kutta). For the intended biological applications, the stability criterion is more important than accuracy, which would tend to use low order schemes.

In order to make the use of deep learning ever more efficient, mixed precision as well as computational schemes using them (e.g. convolution) have been introduced in hardware architectures for several years eg Intel VNNI, Intel Nervana, NVidia Tensor-Core, Google TPU. The idea of mixed precision is to combine the use of half-precision (16-bit) and single-precision (32-bit) to reduce memory usage and increase compute density. If the hardware has been developed for classical deep learning, the use of mixed precision is also considered in linear algebra [12, 4] or for solving linear problems that occur in the discretization of partial differential equations [2]. The impact of precision on computational time and

the stability of results has been studied in the context of ODEs [13]. But these studies for methods mixing several precisions have been made only for high order methods and relatively small system sizes [11].

Of course, the transition to mixed precision can limit the accuracy of current ODE solvers, and possibly induce numerical instabilities. However, researchers use in practice the default settings of ODE software, with accuracies of the order of 3 to 6 decimals (for example the packages of Solve de R or scipy.integrate.ode in python are based on ODEPACK [6]). With semi-precision floats, which have 3-4 precision decimals, this accuracy is largely attainable. Predictor-corrector methods are well suited for mixed precision calculations, with reduced prediction and extended precision correction. It will also be useful to test these methods in reduced precision when one will only be interested in the statistical properties of the solutions. In addition, one of the exploratory aspects of this thesis is to mix several precisions even within a calculation (e.g. entries in half-precision and accumulator in simple precision) in order to limit the recourse to methods too expensive of prediction-correction.

In order to make the most of the resources of the computing platforms, it is necessary to take into account the parallelization of the ODE solvers. This is a field studied for many years [15, 7], there are 3 main axes of parallelization: method, time and system. Most approaches have focused on the first two as they apply to relatively small systems. The systems we consider are large and therefore allow enough parallelism to be exposed at the system level. Consequently, ODE optimization methods such as loop transformation [14, 10] are therefore not relevant in our case. It should be noted that the different precisions, methods of resolution, corrections and parallelism bring compromises performance and precision that will not be the same depending on the systems. It is therefore important to study these compromises and to propose solutions that automate this choice as much as possible. Several methods [9, 8] have been proposed in this context but without taking into account the semi-precision / correction. In addition, the proposed solutions are generalist whereas we propose to study a precise field of application by integrating their specificities from the design stage.

To sum up, the subject of the thesis is to integrate new approaches of mixed precision in the resolution of ODE on modern (and future) computing platforms for computational biology (and potentially for deep learning based on from ODE, see below for further explanation). Let us add that beyond the use cases in biology that are at the heart of the thesis, the ODE resolution is a major technological lock for other models in biology. Beyond biology, the resolution of ODE is at the heart of a new current of deep learning in order to be able to respond to the limitations of current methods [1]. In fact, the mechanisms currently used discretize the search spaces, and in the case where the data used are based on continuous variables, the usability of these methods is severely limited. In [1], the authors propose to replace the discrete method which is at the heart of the current deep learning ie the Recurrent Neural Network (RNN) which are the basic building blocks of convolution neural networks (CNN) [5] which are the cornerstone of the current explosion of the usability of deep learning. The authors of [1] as well as several recent works [3] propose to replace the RNNs by methods based on ODE which make it possible to take into account these continuous variables, to learn from better models but also and especially which allow to to build on several decades of theoretical research on ODEs. However, to be effective, these approaches require resolution a very large number of large ODE systems (from several hundreds of thousands to tens of millions of parameters). It is therefore necessary to have a powerful mechanism for solving ODE. Based on these similarities, we will explore the possibility that the approaches developed for computational biology will have a major impact on these new deep learning approaches in terms of performance and accuracy as well as scaling up these methods.

References

[1] Tian Qi Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. In Advances in Neural Information Processing Systems, pages 6571-6583, 2018.

[2] Dominik Göddeke, Robert Strzodka, and Stefan Turek. Performance and accuracy of hardware-oriented native-, Emulated-and mixed-precision solvers in fem simulations. International Journal of Parallel, Emergent and Distributed Systems, 22(4) :221-256, 2007.

[3] Will Grathwohl, Ricky TQ Chen, Jesse Betterncourt, Ilya Sutskever, and David Duvenaud. Ffjord : Free-form Continuous dynamics for scalable reversible generative models. arXiv preprint arXiv :1810.01367, 2018.

[4] Azzam Haidar, Stanimire Tomov, Jack Dongarra, and Nicholas J. Higham. Harnessing gpu tensor cores for fast fp16 arithmetic to speed up mixed-precision iterative refinement solvers. In Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis, SC '18, pages 47 :1-47 :11, Piscataway, NJ, USA, 2018. IEEE Press.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.

[6] Alan C Hindmarsh. Serial fortran solvers for ode initial value problems. URL: https://computation.llnl.gov/casc/odepack/ [cited June 6, 2019], 2002.

[7] P.J. Van Der Houwen and B.P. Sommeijer. Parallel iteration of high-order runge-kutta methods with stepsize control. Journal of Computational and Applied Mathematics, 29(1) :111 - 127, 1990.

[8] Natalia Kalinnik, Matthias Korch, and Thomas Rauber. An efficient time-step-based self-adaptive algorithm for predictor-corrector methods of runge-kutta type. Journal of Computational and Applied Mathematics, 236(3) :394 - 410, 2011. Aspects of Numerical Algorithms, Parallelization and Applications.

[9] Natalia Kalinnik, Matthias Korch, and Thomas Rauber. Online auto-tuning for the time-step-based parallel solution of odes on shared-memory systems. Journal of Parallel and Distributed Computing, 74(8) :2722 - 2744, 2014.

[10] Matthias Korch and Tim Werner. Accelerating explicit ode methods on gpus by kernel fusion. Concurrency and Computation : Practice and Experience, 30(18) :e4470, 2018. e4470 cpe.4470.

[11] Tomonori Kouya. Practical implementation of high-order multiple precision fully implicit runge-kutta methods with step size control using embedded formula. arXiv preprint arXiv :1306.2392, 2013.

[12] Xiaoye S. Li, James W. Demmel, David H. Bailey, Greg Henry, Yozo Hida, Jimmy Iskandar, William

Kahan, Suh Y. Kang, Anil Kapur, Michael C. Martin, Brandon J. Thompson, Teresa Tung, and Daniel J. Yoo. Design, Implementation and testing of extended and mixed precision blas. ACM Trans. Math. Softw., 28(2) :152-205, June 2002.

[13] L. Murray. Gpu acceleration of runge-kutta integrators. IEEE Transactions on Parallel and Distributed Systems, 23(1) :94-101, Jan 2012.

[14] T. Rauber and G. Rünger. How do loop transformations affect the energy consumption of multi-threaded runge-kutta methods ? In 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP), pages 499-507, March 2018.

[15] Thomas Rauber and Gudula Rünger. Parallel execution of embedded and iterated runge⬚kutta methods. Concurrency : Practice and Experience, 11(7) :367-385, 1999.

Assignment

The PhD student will contribute to the research activities of the exploratory action ExODE (Scaling the Resolution of Ordinary Differential Equation for Computational Biology) on the following subjects: numerical and parallel methods for the resolution of ordinary differential equation (ODE), ODE resolution scheme in mixed precision, numerical method adaptation for the ODE resolution applied to computational biology (and deep learning) and more generally to participate in all the scientific activities of the exploratory action ExODE.

Main activities

Activities:

- Analyze requirements
- Propose and analyze solutions
- Design experimental platforms
- Write articles
- Write the reports
- Present the work

Skills

Technical skills and level required :

Research master, Applied mathemathics (and/or), Computational sciences (and/or), high performance computing

Languages : English

Benefits package

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Remuneration

1st and 2nd year: 1 982 euros gross salary /month

3rd year: 2 085 euros gross salary / month

General Information

- **Theme/Domain :** Computational Biology
  Scientific computing (BAP E)
- **Town/city :** VILLEURBANNE
- **Inria Center :** Centre Inria de Lyon
- **Starting date :** 2022-09-01
- **Duration of contract :** 3 years
- **Deadline to apply :** 2022-07-16

Contacts

- **Inria Team :** BEAGLE
- **PhD Supervisor :**

Rouzaud-cornabas Jonathan / jonathan.rouzaud-cornabas@inria.fr

About Inria

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

> **Warning** : you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.

Instruction to apply

Defence Security :
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST).Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy :
As part of its diversity policy, all Inria positions are accessible to people with disabilities.