

Inhalt



Editorial	2
Iterationsschleife	6
Multigrid implementation for Lattice QCD	8
DLR und LRZ unterzeichnen	
Kooperationsvertrag „Terra-Byte“	15
Seventh BGCE Student Paper Prize:	
Elizabeth Qian and Zakia Zainib visit Bavaria	21
Three of a kind: SCCS version	26
Cambridge's Tea Party	30
Neue Version von LIKWID	34
Marvell/Cavium ThunderX2 am RRZE	35
Neues vom KONWIHR	37
Notiz*Notiz*Notiz	38

Das Quartl erhalten Sie online unter <http://www5.in.tum.de/quartl/>



Das Quartl ist das offizielle Mitteilungsblatt des *Kompetenznetzwerks für Technisch-Wissenschaftliches Hoch- und Höchstleistungsrechnen in Bayern* (KONWIHR) und der *Bavarian Graduate School of Computational Engineering* (BGCE)

Editorial

Seit einiger Zeit geistert immer wieder die Idee durch die Universitäten, national wie international, die inzwischen nicht mehr ganz so unreifen Produkte von Plagiats-Software standardmäßig bei der Einreichung von Dissertationen einzusetzen. Manche Unis wie die TUM überlegen das aktuell, andere wie die DTU in Dänemark stellen diese Praxis nach einer Erprobungsphase wieder ein. Die Protagonisten preisen dies als weitere Segnung der Digitalisierung und als wichtigen Schritt in Sachen Qualitätskontrolle, die anderen wehren sich gegen den ihrer Ansicht nach zugrunde liegenden Generalverdacht gegen alle. Und die Promovierenden sind verunsichert. Ich rede dabei nicht von den echt schwarzen Schafen, sondern von denen ohne böse Absicht, die die Gefahr sehen, zu unrecht beschuldigt zu werden. Doch der Reihe nach.

Szenario 1: Die Universität erwirbt Lizenzen einer mehr oder minder passenden Software und unterzieht anschließend standardmäßig jede eingereichte Doktorarbeit einer Prüfung – getreu dem Motto „Wer nichts zu verbergen hat, kann nichts dagegen haben“. So weit, so gut. Die entscheidende Frage ist, was mit dem Prüfergebnis passiert. Das Horrorszenario wäre wohl ein Automatismus, der z.B. ab X Prozent Markierungen Alarm schlägt und letztlich zur Ablehnung der Arbeit führt. Aber das fordern aktuell nur sehr wenige. Die meisten sehen vielmehr einen kleinen Ausschuss gefordert, der mit Betreuer bzw. Betreuerin sowie Kandidat bzw. Kandidatin redet und sich ein Bild von der Angelegenheit macht. Hinsichtlich Ergebnis ist dann das gesamte Spektrum möglich – von „nichts passiert“ bis zur Ablehnung samt Einleitung der üblichen Maßnahmen bei schweren Verstößen gegen die gute wissenschaftliche Praxis.

Das Unbehagen dabei betrifft nicht die Schurken – mit denen hat niemand Mitleid (abgesehen von den Schurken selbst natürlich) – sondern diejenigen ohne böse Absicht. Entspricht es wirklich dem Selbstverständnis der Wissenschaft, hinter jedem Baum einen Betrüger zu vermuten und sich von dieser Vermutung die eigene Handlung vorschreiben zu lassen?

Aus dieser Überlegung heraus resultiert Szenario 2: Die Universität erwirbt

Lizenzen und stellt diese allen Mitgliedern – also in die Betreuung, Bewertung oder Anfertigung einer Arbeit involvierte Personen – zur Verfügung. Jeder und jede mögen also bitte vorab selbst prüfen, was für Farbkleckse erzeugt werden, warum dies so ist, und was man an der Dissertation anders machen müsste, um dies zu vermeiden. Kritiker werfen jetzt ein, die Polizei gebe die Routen ihrer Streifen auch nicht potenziellen Einbrechern bekannt. Da ist sie schon wieder, diese ärgerliche „Schuldvermutung“. Nein, diese Analogie passt ganz und gar nicht. Wer bewusstes Fremdkopieren absichtlich kaschieren möchte, kommt schon selbst auf die Idee, mal einen Plagiatscheck anzuwerfen ... Wenn schon Analogie, dann schon eher die der angekündigten Radarkontrollen, die bekanntlich ja durchaus einen positiven Effekt auf die Verkehrssicherheit haben.

Als wir das nun neulich an der Fakultät diskutierten, erregte dies die Gemüter sehr. Schnell war klar: Nicht mögliche Plagiate sind das Problem, sondern die Gefahr von Selbstplagiaten, wenn man seine eigenen Publikationen verwende. Aber Moment, dafür gibt es doch die publikationsbasierte Dissertation, und dort werden eigene Publikationen doch explizit eingebunden – da kann dann kein Vorwurf eines Selbstplagiats drohen. Darauf hieß es dann, eine publikationsbasierte Dissertation wollten die meisten wegen der hohen Hürden (z.B. die erforderliche explizite Zustimmung aller Co-Autoren aller Publikationen – was bei den in manchen Fächern nicht unüblichen langen Autorenlisten schon aufwändig werden kann; oder das in der Regel geforderte „final angenommen“, das auch dauern kann) ja gerade nicht. Um es kurz zu machen: Das Ganze endete in einer Diskussion, was denn eigentlich eine Monographie ausmache.

Also schauen wir uns das mal an – und sprechen gleich eine Warnung aus: Ich scheine da wieder mal ziemlich altmodisch zu ticken ... Für mich ist die klassische Dissertation in Form einer Monographie von der Konzeption her ein einheitliches Werk, quasi ein Buch. Auch wenn dieses in Kapitel unterteilt ist, in denen i.W. die Ergebnisse aus eigenen Publikationen stehen, so ist es doch ein Ganzes. Ein Roman ist ja auch etwas anderes als eine Sammlung von Kurzgeschichten. Während ein Paper typischerweise ein Detail beleuchtet, steht in der Dissertation auch die Einordnung in den Kontext im Vordergrund. Außerdem richten sich Zeitschriftenartikel und Tagungsbeiträge im Allge-

meinen an ein sehr enges Fachpublikum, das vergleichbar gut oder vielleicht sogar besser mit der Materie vertraut ist und daher mit einem kompakten Schreibstil und einer – sagen wir sparsamen – Argumentationskette gut umgehen kann. Dissertationen dagegen werden ja schon mal Bachelor- oder Masterstudierenden in die Hand gedrückt, zwecks Einlesens. Letzteres ist immer mein Hauptargument, dass schon rein stilistisch ein Paper textlich nicht direkt in eine Dissertation einfließen könne.

Was ich inzwischen aus zahlreichen Gesprächen mit Kolleginnen und Kollegen aus verschiedenen Fächern gelernt habe, beruhigt nicht wirklich. Die Praxis, die im Laufe der Promotionszeit angefertigten Publikationen mehr oder weniger direkt in eine Monographie einfließen zu lassen, ist nicht so ganz unverbreit. Man tut das erstens, weil man so – im Gegensatz zur publikationsbasierten Dissertation – nicht warten muss, bis die Arbeit angenommen ist (und daher auch mal Impact-technisch sehr hoch greifen kann (in den Life Sciences würde man sagen „Nature“), ohne Risiko quasi), seine Co-Autoren nicht um Erlaubnis fragen und auch nicht die Frage nach dem eigenen Anteil beantworten muss. Und man tut das zweitens, weil in einem „Paraphrasieren“ (vulgo bloßen Umformulieren) kein Mehrwert, sondern nur Zeitverschwendung gesehen wird – wie viele weitere hochwertige Publikationen könne man in der klassischen Phase des Zusammenschreibens schreiben. Dieses Vorgehen hat bisher wohl kaum jemanden gestört – allenfalls, wenn’s zu offensichtlich war. Ein automatischer Plagiatscheck dagegen würde hier anschlagen: Textuelle Übereinstimmungen größeren Umfangs mit (eigenen) Arbeiten würden festgestellt und, da es keine publikationsbasierte Dissertation ist, auch als „Selbstplagiat“ bewertet. Und das schreckt jetzt natürlich alle auf.

Zur Beruhigung der Gemüter sei gesagt, dass die Welt (wieder mal) nicht binär ist. Seitenweise Text aus Papers zu übernehmen geht natürlich nicht, auch wenn die Quelle angegeben wird. Aber ebenso natürlich kann man Ergebnisse aus Papers und deren Aufbereitung (Graphiken, Tabellen inkl. der erläuternden Unterschriften) schon übernehmen, ohne den Pfad der Tugend zu verlassen. Aber wer eine klassische Dissertation in Form einer Monographie schreiben will, der muss schon ein eigenes Werk verfassen. Ich kann schließlich auch nicht meine 90 Quartl-Editorials aneinanderkleben und dies

als meinen Erstlingsroman anpreisen.

Dessen ungeachtet kämpfen wir gegen flächendeckende Plagiats- oder Eigenplagiatstests bei Dissertation. Es darf keine generelle Krimialisierung stattfinden. Gleichwohl mag ein solcher Test, selbst angewandt, dem einen oder der anderen bei der Interpretation dessen helfen, was denn eine klassische Dissertation ausmacht.

Die gesamte Quartl-Redaktion wünscht Ihnen, liebe Leserinnen und Leser, einen schönen Sommer, hoffentlich mit etwas Urlaub, zunächst aber natürlich viel Spaß bei der Lektüre der neusten Ausgabe Ihres Quartls ... (Hab ich das nicht schon mal gelesen resp. geschrieben? ist das am Ende ein Selbstplagiat??)

Hans-Joachim Bungartz.

Iterationsschleife

N=31

11. Juni 2019

Mikroorganismen sind kleinste Lebewesen, die entweder als Einzeller oder auch als Mehrzeller die Welt bevölkern und unter anderem auch den Menschen begleiten. Die Zahl der Mikroorganismen auf unserem Körper übersteigt wohl die Zahl der Zellen in unserem Körper. Mikroorganismen haben darüber hinaus aber auch die Eigenschaft, unser Verhalten zu verändern. Thomas Bosch kommt gemeinsam mit Tobias Rees und Angela Douglas zu dem Schluss, dass die Existenz und Bedeutung von Mikroorganismen am und im menschlichen Körper die Frage aufwirft, ob wir von einem „Selbst“ in der bisherigen Form noch sprechen können^a.

Thomas Bosch bezieht sich auf eine biologische Definition des Selbst, die sich das Immunsystem, das Gehirn und das Genom als Grundlage des Selbst zu eigen macht. Im Sinne einer allgemeinen Definition des Selbst kann man an dieser Definition zweifeln, doch ist die Herausforderung an das menschliche Selbst nicht so einfach von der Hand zu weisen. Tatsächlich verändern Mikroorganismen auch unsere Denkprozesse und damit das, was umgangssprachlich als unser Selbst angesehen wird. Ob und inwieweit Mikroorganismen unseren „Charakter“ beeinflussen wird damit zu einer wissenschaftlich relevanten Frage.

Eine weitere Frage tut sich auf. Wenn das Selbst im Ansturm der Mikroorganismen ins Wanken gerät ist zu fragen, ob das Selbst selbst eine relative Größe ist, die nicht absolut gesetzt wird, sondern von den Mikroorganismen determiniert wird. Damit sänke der westliche autonome Mensch zurück in den Strudel der Biologie, aus dem ihn die wissenschaftliche Forschung über Jahrhunderte scheinbar befreit hat.

Eine Parallele zum Konzept der Information drängt sich auf. Auch der Begriff der Information ist relativ und setzt ein Interpretationssystem voraus, in dessen Kontext aus Daten oder Wahrnehmungen erst Informationen werden. So ist etwa die Inschrift auf dem Diskos von Phaistos aktuell keine Information mehr. Uns fehlt der Interpretationsrahmen, der es uns erlauben würde, die mögliche Information zu extrahieren.

^aTobias Rees, Thomas Bosch, Angela E. Douglas, How the microbiome challenges our concept of self, <https://journals.plos.org/plosbiology/article/authors?id=10.1371/journal.pbio.2005358>

Das Wesentliche an der Information ist also zunächst nicht die Information selbst, sondern das System innerhalb dessen die Daten interpretiert werden. Das biologische Selbst wäre auch erst innerhalb des Rahmens, der von Mikroorganismen gesetzt wird, ein Selbst. Unser Selbst wäre also relativ und es stellt sich die Frage ob der Begriff des Selbst dann sinnvoll ist oder ob von einem möglichen Selbst unter vielen zu sprechen wäre, das wir aktuell als unser Selbst betrachten können ohne Anspruch auf Dauerhaftigkeit. Das traditionelle Menschenbild vom Individuum – wie es vor allem im Westen vorherrscht – wäre damit beinahe ad absurdum geführt.

Eine weitere Herausforderung an das Menschenbild finden wir in der „künstlichen Intelligenz“. Während allenthalben die Angst vor der Überwachung oder dem Verlust von Arbeitsplätzen diskutiert wird, ist ein Aspekt in der Diskussion meist weniger relevant. Der denkende Mensch, der sich über seine Fähigkeit zu denken definiert und für den in der westlichen Welt das Konzept des „rationalen Handelns“ erarbeitet wurde, das ihn über die Tiere erhebt, wird herausgefordert von einer Technik, die den Anspruch erhebt, tatsächlich rational zu agieren und damit den Menschen zurück in die gemeinsame biologische Ecke der „Lebewesen“ zu schicken. Künstliche Intelligenz kann vermeintlich nicht nur schneller entscheiden, sondern auch rationaler und damit im eigentlich Sinn unseres Verständnisses vom Menschen, menschlicher. Künstliche Intelligenz wäre also logischerweise als menschlicher anzusehen als der Mensch selbst.

Der Mensch des 21ten Jahrhunderts sieht sich aufgrund der wissenschaftlichen Entwicklungen, die er selbst vorantreibt, um sein Verständnis der Welt zu verbessern, zahlreichen Zweifeln an seiner selbstdefinierten Rolle ausgesetzt. Die konkrete Untersuchung nachvollziehbarer und logisch erklärbarer Prozesse in Biologie und Informatik führt dazu, dass konkrete Begriffe relativiert werden müssen und sich die Frage stellt, ob diese Begriffe wissenschaftlich noch Sinn machen.

Die Wissenschaft richtet sich damit nicht gegen den Menschen aber sie richtet sich gegen ein Menschenbild, das noch immer einem Geist des Positivismus anhängt wie ihn Auguste Comte definiert hat^a. Damit zwingt uns die Wissenschaft dazu, auch in unserem Selbstbild endlich im 21ten Jahrhundert anzukommen.

M. Resch

^a Auguste Comte, Rede über den Geist des Positivismus, Meiner, 1994 (ursprünglich: Discours sur l'ensemble du positivisme, 1848)

Multigrid implementation for Lattice QCD

Quantum Chromodynamics (QCD) is one of the fundamental theories in particle physics. In numerical simulations of QCD on a discrete space-time lattice, the computational cost is dominated by the solution of a linear system involving the so-called Dirac matrix. In state-of-the-art projects, the dimension of this matrix is $O(10^9)$. The good news is that it is sparse. The bad news is that its condition number diverges in the limit of physical interest. In this situation, the inversion of the Dirac matrix requires an adaptive multigrid algorithm. We describe our implementation of such an algorithm on parallel machines based on SIMD processors.

1 Lattice QCD

Lattice QCD is one of the early adopters of High-Performance Computing (HPC), with the theory being developed in 1974 [1] and the first simulations carried out in 1980 [2]. Lattice QCD is complicated mathematically since it involves an $SU(3)$ gauge theory as well as fermions. However, from a computational point of view it is rather simple since it is formulated on a fixed hypercubic lattice and on a parallel computer the communication patterns are regular and predictable. The field is now at a point where the quality of the computational results approaches the precision of large-scale accelerator experiments at CERN and elsewhere so that Lattice QCD can be used as a discovery tool in the search for new physics beyond the standard model. To get to this point required an interplay of hardware developments, algorithmic inventions, and optimized implementations. In fact, a number of special-purpose machines to simulate Lattice QCD were developed over the decades, including the QPACE series of machines in the framework of SFB/TRR-55 led by the Regensburg group [3] [4] [5].

The theory is simulated on a discrete and finite lattice with 3 spatial dimensions and one temporal dimension (for experts: after a Wick rotation to Euclidean space). There are at least three important parameters: the lattice

spacing a (we assume the same spacing in all 4 dimensions), the volume V , and the mass m_π of the lightest particle, the pion (m_π is related to the mass m of the light quarks as $m_\pi \sim \sqrt{m}$). In a particular simulation, a is nonzero and V is finite. Therefore extrapolations are required to the physically interesting limits: the continuum limit $a \rightarrow 0$ and the thermodynamic limit $V \rightarrow \infty$. Furthermore, as we shall see in the next paragraph, it is expensive to work at the physical pion mass.

The execution time of a Lattice QCD simulation is dominated by the solution of the linear system

$$Dx = b, \quad (1)$$

where D is the so-called Dirac matrix (which is sparse and non-Hermitian), b is a known vector, and x is the solution vector. The dimension of these objects is 12 times the lattice volume. State-of-the-art lattices can be of size, e.g., $96^3 \times 192$, implying a dimension of $2 \cdot 10^9$. Equation (1) is solved using an iterative Krylov-space method. The condition number of D scales with an inverse power of ma , i.e., it becomes increasingly difficult to invert D as we approach the continuum limit $a \rightarrow 0$ and the physical quark or pion mass. Figure 1 gives an impression of how bad the situation is. Note the logarithmic scale on the vertical axis and the dramatic improvement from an already quite sophisticated algorithm (BiCGStab) to the multigrid algorithm, which is called AMG (adaptive algebraic multigrid based on domain decomposition) and was developed by the group of Andreas Frommer in Wuppertal, our collaborators in SFB/TRR-55.

2 Adaptive multigrid

AMG is used as a preconditioner for D . For details on the algorithm we refer to [6]. One main point is that the error in the iterative solution of (1) receives contributions from the low and high eigenmodes of D . Multigrid means that we go back and forth between the original “fine” grid (or lattice) and one or more levels of “coarse” grids. On the fine grid we reduce the

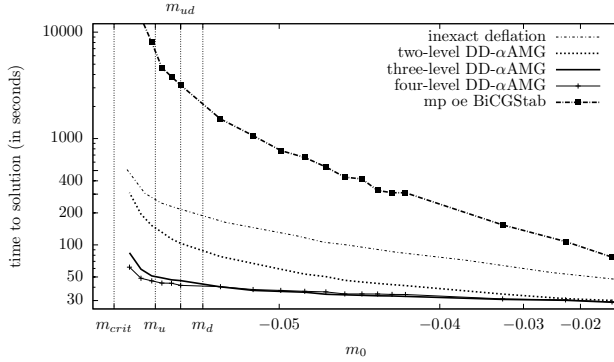


Figure 1: Time to solution of Eq. (1) for various algorithms on a 64^4 lattice as the physical masses m_u and m_d of the up and down quarks are approached, with $m_{ud} = (m_u + m_d)/2$. BiCGStab is the previous state of the art, AMG is an adaptive multigrid algorithm. Taken from [6].

contributions of the high modes to the error, while on the coarse grid we reduce the contributions of the low modes.

We go from the fine to the coarse grid by an operation R (“restriction”) and from the coarse back to the fine grid by an operation P (“prolongation”). The coarsened representation of D reads $D_c = RDP$. Note that $P = R^\dagger$ and that the matrix representations are rectangular. On the coarse grid we solve the analog of (1) with D_c instead of D approximately to low precision (e.g., using FGMRES). This is called coarse-grid correction (CGC). On the fine grid we solve (1) approximately with the result from the CGC as a starting guess. This is called smoother. Here, we employ a domain-decomposition method (more precisely, the Schwarz Alternating Procedure [7]), which is a communication-avoiding algorithm and thus advantageous on a parallel machine.

The adaptive multigrid method consists of two parts: “setup” and “solve”. The latter is simply the application of the multigrid preconditioner. The former is more subtle and defines what is meant by “adaptive”. Its aim is the construction of a set of test vectors that approximately span the near-null space of D . These test vectors are then used to construct the operators P ,

R , and D_c . The setup is quite expensive, but it only has to be done once for a given D , and its result can then be used for different right-hand sides (RHS) b in (1). If there are many RHS, the cost of the setup will be amortized. However, the problem space of Lattice QCD also contains situations with only a single RHS. Therefore it is important to also optimize the setup [8].

3 Implementation details and results

Since Lattice QCD has a byte-to-flop ratio of about 1 (the actual number depends on some choices one can make) it is memory-bandwidth bound on all current architectures. Therefore it is essential to optimize the data layout to obtain streaming memory access in the performance-critical parts. Note that the optimal data layout can be different in different parts of the algorithm. We change the data layout accordingly unless the cost of the layout change is higher than the resulting performance gain. Furthermore, we run the multigrid method in single precision to reduce the working set and thus the requirements on cache capacity and memory bandwidth. Since the multigrid method is run as a preconditioner to an outer iterative method run in double precision, this does not impact the convergence of the solver.

Our current machines QPACE 2 and 3 (among others) are all based on the Intel Xeon Phi series, i.e., processors with SIMD units with a width of 512 bits. Hence it is important to choose a good vectorization strategy, which can be different in different parts of the algorithm. In the smoother we vectorize by site-fusing, i.e., several lattice sites are combined in a SIMD register. (More precisely, the 24 floating-point components of x and b corresponding to one lattice site are stored in 24 separate registers.) In the other parts of AMG it is simpler and leads to higher performance if we vectorize over the test vectors created in the setup. Finally, when multiple RHS are present we obtain the highest performance if we vectorize over them. In all cases, real and imaginary parts are not mixed in the same SIMD register.

The parallelization over cores is done using OpenMP. We decompose the lattice into pieces that are assigned to individual threads. We use persistent threads with synchronization points since this turns out to give better

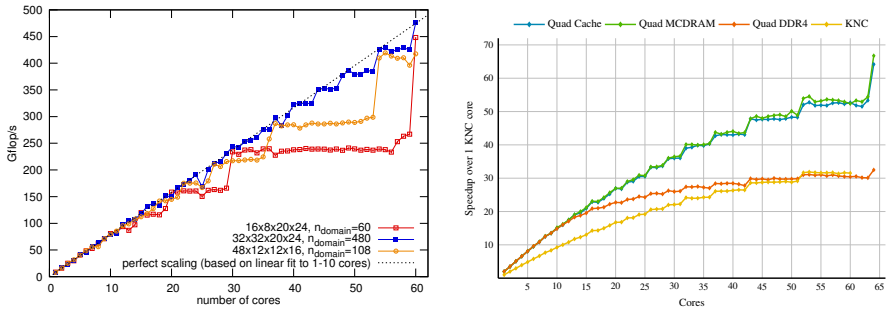


Figure 2: Left: On-chip scaling of the DD smoother on the KNC [9]. Right: Speedup of the entire multigrid preconditioner on KNL over KNC in different KNL memory configuration modes [5].

performance than fork-join.

The parallelization over nodes is done using MPI. Having each thread issue its own MPI calls is not a good idea since (a) the overhead of an MPI call is high and (b) the message sizes would be too small for an efficient network utilization. In the smoother, we therefore combine the surface data of all domains and communicate them using a single thread. Furthermore, we partially hide the communication latency by starting to send the boundary data when half of them are ready, as explained in detail in [9]. If multiple RHS are present we combine them to obtain fewer and larger messages.

On QPACE 2, which is based on Intel Xeon Phi 7120X (Knights Corner or KNC) with in-order cores, additional tricks were necessary to obtain optimal performance: (a) manually tuned L1 and L2 prefetches using intrinsics, and (b) the storage of some data in half precision to further increase the benefits of the reduced-precision effects discussed above. (a) is not necessary on out-of-order architectures, and the overhead associated with (b) turned out to give lower overall performance on the Xeon Phi 7210 (Knights Landing or KNL). The code discussed above is available at [10]. Some performance examples are shown in Fig. 2.

Intel’s discontinuation of the Xeon Phi product series forces us to turn towards other architectures in the future. Porting the Xeon Phi codebase

described above is not a valid option as it was specifically written for these kinds of processors. Hence we are currently integrating [11] [12] this multigrid solver algorithm into the modern “Grid” library for Lattice QCD [13], which is an open-source data-parallel library with major emphasis on performance portability. In its first release, it featured support for all x86 CPUs and Xeon Phi processors. Since then, support for ARM CPUs has been added by the Regensburg group [14]. Further work to include support for NVIDIA GPUs and NEC Aurora accelerators is ongoing. This project will enable us to run the optimal inverter for our problem, i.e., multigrid, on nearly all HPC architectures that are currently on the market.

Daniel Richtmann, Tilo Wettig

References

- [1] K.G. Wilson, Phys. Rev. D10 (1974) 2445 <https://journals.aps.org/prd/abstract/10.1103/PhysRevD.10.2445>
- [2] M. Creutz, Phys. Rev. D21 (1980) 2308 <https://journals.aps.org/prd/abstract/10.1103/PhysRevD.21.2308>
- [3] G. Goldrian et al., Comput. Sci. Eng. 10 (2008) 46 <https://arxiv.org/abs/1904.03927>
- [4] P. Arts et al., PoS LATTICE2014 (2015) 021 <https://pos.sissa.it/214/021>
- [5] P. Georg, D. Richtmann an T. Wettig, EPJ Web Conf. 175 (2018) 02007 [dx.doi.org/10.1051/epjconf/2018175020071710.07041](https://doi.org/10.1051/epjconf/2018175020071710.07041) <https://arxiv.org/abs/1710.07041>
- [6] A. Frommer, K. Kahl, S. Krieg, B. Leder and M. Rottmann, SIAM J. Sci. Comput. 36 (2014) A1581 <https://arxiv.org/abs/1710.070411303.1377> <https://arxiv.org/abs/1303.1377>

- [7] M. Lscher, *Comput. Phys. Commun.* 156 (2004) 209 <https://arxiv.org/abs/1303.1377> [hep-lat/0310048](https://arxiv.org/abs/hep-lat/0310048) <https://arxiv.org/abs/hep-lat/0310048>
- [8] Dr. Richtmann, S. Heybrock and T. Wettig, *PoS LATTICE2015* (2016) 035 <https://pos.sissa.it/251/035> 1601.03184 <https://arxiv.org/abs/1601.03184>
- [9] S. Heybrock, B. Jo, D.D. Kalamkar, M. Smelyanskiy, K. Vaidyanathan, T. Wettig et al., *Lattice QCD with Domain Decomposition on Intel Xeon Phi Co-Processors*, in *The International Conference for High Performance Computing, Networking, Storage and Analysis: SC14: HPC matters (SC)* New Orleans, LA, USA, November 16-21, 2014, 2014, 1412.2629 <https://arxiv.org/abs/1412.2629>, DOI <https://ieeexplore.ieee.org/document/7012993>
- [10] S.Heybrock, <https://rqcd.ur.de:8443/hes10653/mic-qcd-solver>
- [11] D. Richtmann, <https://github.com/DanielRichtmann/Grid>
- [12] D. Richtmann, P. A. Boyle and T. Wettig, <https://pos.sissa.it/334/032> <https://arxiv.org/abs/1904.08678>
- [13] P. A. Boyle, G. Cossu, A. Yamaguchi and A. Portelli, *PoS LATTICE2015* (2016)023<https://pos.sissa.it/251/023> 1512.03487 <https://arxiv.org/abs/1512.03487>
- [14] N. Meyer, D. Pleiter, S. Solbrig and T. Wettig, *PoS LATTICE2018* (2019)316<https://pos.sissa.it/334/316> 1904.03927<https://arxiv.org/abs/1904.03927>

DLR und LRZ unterzeichnen Kooperationsvertrag „Terra-Byte“



Das Deutsche Zentrum für Luft- und Raumfahrt (DLR) und das LRZ kooperieren in Zukunft unter Namen „Terra-Byte“ bei der Auswertung der enormen Datenmengen, die Erdbeobachtungssatelliten im Zusammenspiel mit weiteren Datenquellen täglich über den Zustand unseres Planeten erfassen. Am 27. Mai 2019 wurde der Kooperationsvertrag zwischen DLR, BAdW und LRZ sowie dem Bayerischen Staatsminister für Wissenschaft und Kunst, Bernd Sibler, unterzeichnet.



Abbildung 1: Bei der Unterzeichnung der Kooperationsvereinbarung (v.l.n.r): Prof. Dr. Dieter Kranzlmüller (LRZ), Prof. Dr. Thomas O. Höllmann (BAdW), Bernd Sibler (Staatsminister für Wissenschaft und Kunst), Prof. Dr. Hansjörg Dittus (DLR), Prof. Dr. Stefan Dech (DLR); (C)Alessandro-Podo.de



Abbildung 2: Vor dem Höchstleistungsrechner SuperMUC-NG (v.l.n.r.): Prof. Dr. Stefan Dech, Direktor des Deutschen Fernerkundungsdatenzentrums des DLR, Prof. Dr. Thomas O. Höllmann, Präsident der Bayerischen Akademie der Wissenschaften, Bernd Sibler, Bayerischer Staatsminister für Wissenschaft und Kunst, Prof. Dr. Dieter Kranzlmüller, Vorsitzender des Direktoriums des Leibniz-Rechenzentrums, Prof. Dr. Hansjörg Dittus, DLR-Vorstand Raumfahrtforschung und –technologie, (C)Alessandro-Podo.de

Urbanisierung, Landwirtschaft und schmelzende Gletscher - Höchste Rechenleistung zur Erforschung des Globalen Wandels in der Kooperation „Terra-Byte“

Eines der größten europäischen Höchstleistungsrechenzentren – das Leibniz-Rechenzentrum (LRZ) der Bayerischen Akademie der Wissenschaften – und die größte Raumfahrtforschungseinrichtung Europas – das Deutsche Zentrum für Luft- und Raumfahrt (DLR) – kooperieren in Zukunft bei der Auswertung der enormen Datenmengen, die Erdbeobachtungssatelliten im Zusammenspiel mit weiteren globalen Datenquellen wie den sozialen Netzwerken

täglich über den Zustand unseres Planeten erfassen.

„Die Zusammenarbeit ist ein Meilenstein für die Münchner Wissenschaftslandschaft und den Forschungsstandort Bayern“, betonte Bayerns Wissenschaftsminister Bernd Sibler anlässlich der Unterzeichnung eines Kooperationsvertrages der beteiligten Partnereinrichtungen am 27. Mai 2019 in Garching bei München. „Diese Kooperation verdeutlicht, wie groß die Potenziale unserer Forschungseinrichtungen in Garching und Oberpfaffenhofen sind! Beide Institute leisten herausragende, international beachtete Arbeit. Ihr gemeinsames Projekt eröffnet nun völlig neue Möglichkeiten, um den globalen Wandel und seine Auswirkungen auf unsere Region umfassend zu erforschen. Für Politik und Gesellschaft ist das eine bedeutende Grundlage, um geeignete Rahmenbedingungen für eine lebenswerte Zukunft auf unserer Erde zu schaffen. Luft- und Raumfahrt, Supercomputing und Künstliche Intelligenz greifen dabei ineinander. Hier wird die Technik von morgen entwickelt!“

Mit der Zusammenarbeit bündeln DLR und LRZ ihre große Expertise in den Bereichen satellitengestützte Erdbeobachtung und Supercomputing.

„Wenn wir Prozesse des Globalen Wandels verstehen und seine Entwicklung nachvollziehen wollen, müssen wir in der Lage sein, die Daten unserer Satelliten im All bestmöglich auszuwerten“, betonte Prof. Hansjörg Dittus, DLR-Vorstand für Raumfahrtforschung und -technologie. „Die Kooperation zwischen DLR und LRZ wird es uns künftig ermöglichen, riesige Datenmengen mit neuesten Methoden hoch effizient und unabhängig zum Verständnis globaler Trends und deren Folgen zu analysieren. Beispiele dafür sind die international zunehmende Urbanisierung und Ausweitung der landwirtschaftlichen Nutzung zu Lasten naturbelassener Ökosysteme oder die rasanten Veränderungen in den polaren Gebieten der Erde und in der Atmosphäre, die ohne jede Frage nicht ohne Folgen auch für die Menschen sind. Für die Analyse bringen wir unsere Entwicklungen und Technologien der Raumfahrtforschung sowie eigene Sensordaten mit ein.“

Prof. Dieter Kranzlmüller, Leiter des LRZ, betont: „Mit der Kooperation dieser beiden führenden Forschungseinrichtungen schließen sich zwei Partner zusammen, die sich in ihren Kompetenzen ideal ergänzen und entsprechende Expertisen, Mittel und Forschungsthemen gemeinsam einbringen.

Das Leibniz-Rechenzentrum verfügt über eine ausgewiesene Erfahrung als innovativer IT-Dienstleister und Höchstleistungsrechenzentrum und ist zuverlässiger und leistungsstarker Partner für die bayerischen Hochschulen und künftig auch für das DLR mit seinen Instituten in Oberpfaffenhofen.“

Riesige Datenmengen der Erdbeobachtung

Täglich liefern Erdbeobachtungssatelliten riesige Datenmengen in so hoher Auflösung, dass konventionelle Auswerteverfahren längst an ihre Grenzen gekommen sind. „Nur eine Kombination aus Online-Verfügbarkeit verschiedenster historischer und aktueller Datenbestände in Verbindung mit modernsten Höchstleistungsrechensystemen versetzt unsere Forscher in die Lage, globale Informationen in höchster Auflösung abzuleiten, die uns Aussagen über die Entwicklung des Planeten Erde ermöglichen. Bei der vollautomatischen Analyse spielen Verfahren der Künstlichen Intelligenz eine zunehmend entscheidende Rolle. Somit können wir Phänomene und Entwicklungen erkennen, wie es auf konventionellem Weg nur schwer möglich wäre“, sagt Prof. Stefan Dech, Direktor des Deutschen Fernerkundungsdatenzentrums des DLR. „Für die an der satellitengestützten Erdbeobachtung forschenden Institute des DLR in Oberpfaffenhofen besitzt diese Kooperation entscheidende Bedeutung. Wir können nun eine Vielfalt an methodischen und geowissenschaftlichen globalen Analysen durchführen, die bislang aufgrund der Datenmengen und limitierter Rechnerleistung nur exemplarisch möglich war. Besonders bedeutsam ist das gemeinsam zwischen DLR und LRZ entwickelte technologische Datenkonzept, das das LRZ mit dem Deutschen Satellitendatenarchiv des DLR in Oberpfaffenhofen verbinden wird und neben Online-Verfügbarkeit aktueller weltweiter Datenbestände auch historische Daten aus unserem Archiv sowie DLR-eigene Daten verbindet“, so Dech weiter.

Herausforderung für die Datenanalyse

Schon jetzt haben beispielsweise die Daten des europäischen Erdbeobachtungsprogramms Copernicus die Schwelle von 10 Petabytes überschritten. 1 Petabyte entspricht dem Inhalt von etwa 223.000 DVDs – einer Menge, die ca. 3,5 Tonnen wiegen würde. Bis 2024 werden die Sentinel-Satelliten des

Copernicus-Programms schon mehr als 40 Petabytes an Daten erzeugt haben. Diese werden ergänzt durch weitere Petabytes der nationalen Erdbeobachtungsmissionen wie die DLR-Radarsatelliten TerraSAR-X und TanDEM-X oder die amerikanischen Landsat-Daten. Doch nicht nur die großen Datenmengen der Satellitenmissionen stellen die Wissenschaftler derzeit vor Herausforderungen, auch Daten zum Globalen Wandel, die in den sozialen Netzwerken veröffentlicht werden, sind wertvolle Quellen. Diese bringen allerdings neue Herausforderungen mit sich, da diese Daten sehr heterogen sind, ihr Wahrheitsgehalt nicht eindeutig ist und sie nicht in unbegrenzter Dauer zur Verfügung stehen.

Forscherinnen und Forscher des DLR setzen daher zunehmend Methoden der Künstlichen Intelligenz (KI) und des maschinellen Lernens ein, um Trends des Globalen Wandels und Analysen von Naturkatastrophen und ökologischen Zusammenhängen in globalen und regionalen Zeitreihen zu finden, die sich über Dekaden hinweg erstrecken.

Diese Methoden erfordern es jedoch, dass die dazu notwendigen Daten auf hoch performanten Datenanalyse-Plattformen (HPDA = High Performance Data Analytics) online zur Verfügung stehen. Das technische Ziel der Kooperation ist deshalb der Aufbau einer solchen Plattform, die über das Deutsche Satellitendatenarchiv (D-SDA) des DLR in Oberpfaffenhofen und Daten-Verteilpunkte verschiedener Anbieter frei verfügbarer Satellitendaten Zugriff auf alle für die Forschung notwendigen Erdbeobachtungsdaten hat.

Das Deutsche Fernerkundungsdatenzentrum (DFD) des DLR koordiniert die Aktivitäten der Kooperation für die beteiligten DLR-Institute: Neben dem DFD sind das Institut für Methodik der Fernerkundung, das Institut für Physik der Atmosphäre sowie das Institut für Hochfrequenztechnik und Radarsysteme in Oberpfaffenhofen beteiligt. Bei der Umsetzung der Technologie sind außerdem das Institut für Datenwissenschaften in Jena sowie die Einrichtung Simulations- und Softwaretechnik in Köln beteiligt.

Kooperation für den Globalen Wandel

In der Kooperation befasst sich das DLR mit Forschungsfragestellungen rund um das Thema Umweltveränderungen und Globaler Wandel, methodische und algorithmische Verfahrensentwicklungen im Bereich physikalische

Modellierung und Künstliche Intelligenz (KI), dem Management von Langzeitarchiven und der Prozessierung großer Datenbestände.

Das LRZ fokussiert in der Kooperation auf die Forschung und Umsetzung von operationellen, skalierbaren, sicheren und zuverlässigen IT-Diensten und Technologien, Optimierung von Prozessen und Verfahren, dem Supercomputing und Cloud Computing sowie der Nutzung von Verfahren der Künstlichen Intelligenz (KI) und Big Data. Dabei kommen auch die vorhandenen IT-Systeme des LRZ (u.a. der Höchstleistungsrechner SuperMUC-NG) und die Erfahrungen im energieeffizienten Höchstleistungsrechnen zum Einsatz.

Geplant ist die Realisierung der Online-Verfügbarkeit von etwa 40 Petabytes für Tausende von Rechenkernen. Durch gemeinsame Investitionen des DLR und des LRZ ist geplant, die erste Stufe des Ausbaus bis Ende 2020 zu realisieren. Die neue HPDA-Plattform wird in die bestehende Infrastruktur des LRZ in Garching integriert. Auf die größtenteils freien und offenen Daten der Plattform werden aber auch die Wissenschaftler der Bayerischen Universitäten und Hochschulen Zugriff haben.

Ludger Palm

7.th BGCE Student Paper Prize: Elizabeth Qian & Zakia Zainib visit Bavaria



For the seventh time, the Bavarian Graduate School of Computational Engineering (BGCE) of the Elite Network of Bavaria (ENB) awarded the “BGCE Student Prize” for the best paper at the SIAM Conference on Computational Science and Engineering. For the second time only since 2007, 2 prizes were given - to Elizabeth Qian and to Zakia Zainib. The prize consists of a one-week visit to the universities involved in BGCE: FAU Erlangen-Nürnberg and TU München, during which the award winners had the opportunity to get to know the two universities, get in contact with BGCE’s educational and research program, and presented their own research work in more detail.



Figure 1: Prize winners Elizabeth Qian (left) and Zakia Zainib (right) with Hans Bungartz

Out of 12 high-quality contributions from 12 universities and research institutions of eight countries, two Ph.D candidates - Elizabeth Qian from MIT and Zakia Zainib from SISSA were selected as the prize winners by an international jury at the SIAM CS&E Conference 2019.

Elizabeth is currently a 3rd year Ph.D. candidate in the research group of Karen Willcox, Ph.D., MNZM at the Massachusetts Institute of Technology. She holds an M.S. in Aerospace, Aeronautical and Astronautical Engineering and an B.S. in Aerospace Engineering from MIT. Her research interests are in the areas of computational methods for optimization, model reduction, and uncertainty quantification.

Zakia is the last year Ph.D. candidate in SISSA (International School for Advanced Studies - Trieste, Italy), supervised by Prof. Dr. Gianluigi Rozza. She is looking forward to the completion and defense of her dissertation in September 2019. Zakia received her M.Phil. in Computational Mathematics and B.S. with Honours in Mathematics from University of the Punjab, Lahore. Her research interests broadly fall in the field of numerical analysis. Specifically, she is interested in improvement and applications of existing numerical techniques, as well as in the extension and application of reduced order methods for parametric space reduction, optimal flow control and uncertainty quantification.

FAU and TUM finally hosted Elizabeth and Zakia in Erlangen and Munich from May 19 to May 26. The focus of this visit was two-fold: First, our guests got in close contact with the hosting universities and involved persons. Second, they presented their (and their group's) fascinating research works both at FAU and TUM, what led to interesting and fruitful discussions with the corresponding research groups. In particular, Elizabeth gave a talk on "Transform & Learn: A data-driven approach to nonlinear model reduction", while Zakia presented her work on "Reduced order methods for parametric optimal flow control in coronary bypass grafts, towards patient-specific data assimilation". Moreover, we managed to organize one of BGCE's traditional Research Days (i.e. the meeting of all persons involved in BGCE once per semester at every of the three locations) this time in Munich such that they both were able to join in, contribute with their above mentioned talks and have detailed discussions with the BGCE students afterwards.

For those interested in details, here is a brief abstract of Elizabeth’s current research work: *The development of cheap and accurate approximate models is a critical enabling technology for optimization, uncertainty quantification, and control of nonlinear systems, where traditional numerical solution methods are expensive. In this talk, we present Transform & Learn, a new surrogate modeling method that unites the perspectives of model reduction and machine learning. The proposed method uses insight from the physics of the problem — in the form of partial differential equation (PDE) models — to derive a state transformation in which the system admits a quadratic representation. Snapshot data from a high-fidelity model simulation are transformed to the new state representation and subsequently are projected onto a low-dimensional basis. The quadratic reduced model is then learned via a least-squares-based operator inference procedure. The state transformation thus plays two key roles in the proposed method: it allows the task of nonlinear model reduction to be reformulated as a structured model learning problem, and it parametrizes the machine learning problem in a way that recovers efficient, generalizable models. We demonstrate the proposed method on the Euler fluid dynamics equations and the FitzHugh-Nagumo prototypical neuron activation model.*

As well as a brief abstract of Zakia’s talk: *Bypass grafting is a common surgical treatment to circumvent blood flow blockage caused in congenital heart diseases such as coronary artery disease (CAD). This talk is aimed at presenting a complete numerical pipeline from clinical images to parametrized optimal control on blood flow in patient-specific coronary artery bypass grafts (CABGs), with the aim to track patient-specific physiological data. Minimization of cost functional in such problems ensures bridging the gap between simulations and clinical measurements, whereas unknown control provides an opportunity to obtain trustworthy and meaningful boundary conditions. Additionally, the framework is accompanied with reduced order methods (ROMs), such as proper orthogonal decomposition (POD)-Galerkin, to achieve reliable solution along with cost efficiency in parameter dependent context.*

To explain what the BGCE prize looks like, here comes a short chronological overview of Elizabeth's and Zakia's visit, with special focus on the "Munich part" of the visit: After Elizabeth and Zakia arrived in Munich they took a train to Erlangen, and from Monday morning until Wednesday were in the safe hands of our colleagues from FAU. There, they got introduced to the research group and campus facilities, and were shown the charming university town of Erlangen. Besides being talented researchers, both Elizabeth and Zakia are very sporty and curious ladies, therefore a 20 kilometers walking tour through the inner city of Nuremberg, rich in historical landmarks, was not challenging at all. They moved from Erlangen to Munich on Wednesday, May 22. On their first day in Munich we took them on a tour around the fast-developing campus in Garching and introduced them to colleagues from the Scientific Computing in Computer Science (SCCS) research group. The following day, we organized a guided city tour in Munich, where historical facts together with funny stories and urban legends made Bavaria's capital even more amusing. On Thursday afternoon, the above mentioned Research Day took place. After the talks it was time for the pizza-get-together part of the event where interesting scientific discussions were continued in a more relaxed atmosphere. Spontaneously, our guests and a couple of colleagues from the SCCS chair continued the pleasant evening at famous Chinesischen Turm beer garden, this turned out to be a nice end of the intense day. On Friday, May 24, our guest joined a 2 hour tour in the LRZ facility in Garching, one of the largest supercomputing facilities in Germany and home of SuperMUC-NG – one of the top 10 most powerful supercomputers in the world. After the LRZ tour our guests had some vigorous and prolific scientific discussions with other Ph.D students from the SCCS chair. In smaller groups they presented various aspects of their work in more detail and got a better insight into the ongoing projects at the chair. The Friday afternoon for Elizabeth and Zakia was reserved for a half day excursion around Upper Bavaria together with their hosts Hans Bungartz and Ivana Jovanovic. Perfect weather conditions and beautiful of the landscapes around the lakes Starnbergersee and Ammersee allured all of us. Moreover, musical rehearsals for the upcoming wedding in flamboyant baroque Andechs church added to the fairy-tale like atmosphere. We finished the excursion with a traditional

Bavarian dinner and taste of Andechs beer made in a nearby brewery. In this pleasant atmosphere plans for future collaborations were created. Finally, on their last day in Munich, Saturday 26, both Zakia and Elizabeth visited the Green City Street-Life Festival, which was a nice occasion to say goodbye to the green capital of Bavaria and bound even further newly established relationships with their colleagues from Munich.

Summarizing, Elizabeth's and Zakia's visit led to fruitful discussions and new insight and ideas at both our institutes. They both proved to be very bright and talented researchers. We are very happy that we had a chance to spend time together and to establish connections (we may even say friendship :-)) and are looking forward to future cooperations with both of them, as well as with their groups in Boston and Trieste. We hope that our guests enjoyed their time in Bavaria and are looking forward to meeting them again soon!

Ivana Jovanovic

Three of a kind: SCCS version

“3 is larger than 2, even for large values 2”, or so begin some numerical calculus textbooks. Or in poker terms, three of a kind are better than pairs of two. Or, even better, three-in-a-row Best Poster Awards is better than two-in-a-row awards! The latter version of $3 > 2$ is the Chair’s of Scientific Computing in Computer Science favourite.



Figure 1: Spokane at the end of February, 2019.

<https://twitter.com/search?f=images&vertical=default&q=%23SIAMCSE19&src=typd>

Between February 25 and March 01, 2019, the bi-annual conference on Computational Science and Engineering organized by the Society for Industrial and Applied Mathematics (SIAM) was held in Spokane, Washington. Spokane, you say? Spokane's population is second only to Seattle in the state of Washington (on the same note, Buffalo is the second largest city in the New York state ...). End of February is a great time to go to Spokane if one loves heavy winters, snow, frost, cold, delayed flights, giant breakfast pancakes, but also friendly locals and a convention center capable to host a conference with about 2000 participants.

In this probably-most-important conference for scientific computing, one of the major highlights (if not the major!) is the poster session. Therein, graduate students, postdocs or other scientific computing enthusiasts present their research on a giant piece of paper called a poster. At this year's event, more than 200 posters were presented, covering topics from software for scientific computing, to machine learning + X or computational plasma physics and other real-world applications. Moreover, there were also fewer than 20 poster minisymposiums, that is, groups of posters on the same theme. Most minisymposiums covered software for scientific computing, but some did also tackle other topics, such as multifidelity and multilevel methods. Due to the large number of participating poster presenters, the dedicated session was split in two: the first part was held on Tuesday, February 26 and the second part took place on Wednesday, February 27. In each of the two sessions, a jury consisting of SIAM selected scientists were among the crowd. And based on several rounds of evaluations in which the presenters were asked by the jury to summarize their work, novelty and impact of their poster in computational science, a total of four Best Poster or best Minisymposium Awards were given at the end of each of the two days.

In the Tuesday session, we were among the presenters with our poster titled "Sensitivity-driven Dimension-adaptive Sparse Stochastic Approximations in Liner Gyrokinetics", authored by myself and Dr. Tobias Görler and Prof. Frank Jenko from the Max Planck Institute for Plasma Physics in Garching, and by Dr. Tobias Neckel and Prof. Hans-Joachim Bungartz from the Scientific Computing in Computer Science chair from the Technical University of Munich. In a nutshell, the research outlined on our poster

is concerned with a novel algorithmic technique to quantify uncertainty in complex, real-world problems, such as plasma microturbulence analysis. To this end, we made use of adaptive sparse grid approximations and proposed a novel adaptive technique that preferentially refines the directions rendered important from a stochastic perspective. We tested the proposed approach in two plasma microturbulence test cases, (i) a modified benchmark problem and (ii) a real-world application based on the geometry and experimental data from the ASDEX Upgrade reactor¹ in Garching.

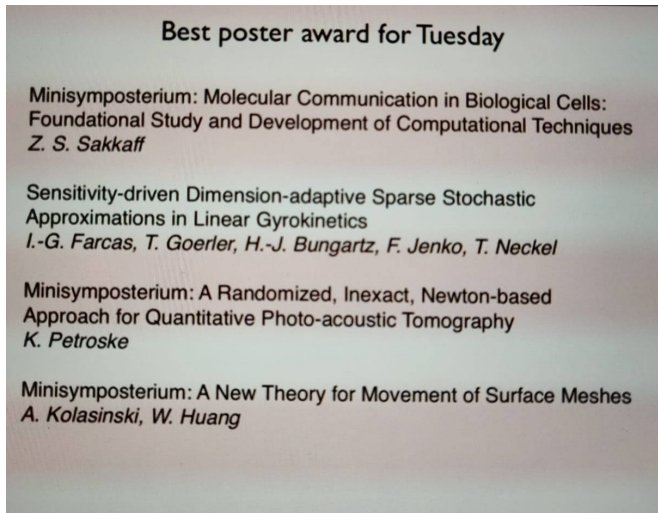


Figure 2: Best poster award for Tuesday, February 26, 2019.

Our research sparked a great deal of interest not only to the crowd, but also to the jury members, who came in increasing numbers in each new evaluation session. At the end of Tuesday, I had a good feeling about the poster session, so I almost forgot how tired I was due to the 18 hours of flying a day before or that I had my invited minisymposium talk in the following day as well. After

¹<https://www.ipp.mpg.de/16195/asdex>

the second session of poster presentations was over on Wednesday evening, the prizes winners were due to be announced. On Thursday early morning, at around 2 AM, I received the e-mail from the organizers announcing that we won a Best Poster Award for the Tuesday session! Moreover, among the four recipients for that day, ours was the only poster! The award ceremony took place on Thursday, February 28 , in which the poster winners took central stage early morning. Besides recognition, the award came also with a prize of \$500, proving that one can make (some) money with (good) scientific ideas. Although SIAM made the check delivery a bit complicated...

For me, personally, winning this poster award is an important vote of confidence that my research ideas are not bad. In addition, this is the second conference in which I win a Best Poster Award, although in the previous edition held in Atlanta in 2017 I could not, unfortunately, be there in person to present the winning poster. Receiving this type of recognition gives me great confidence and motivates me to improve my research ideas and to tackle even more complex and challenging computational problems. For the chair of Scientific Computing in Computer Science, this was the third Best Poster Award in a row, therefore, as a message for the participants in the 2021 conference, no pressure, but you have a legacy to carry on!

Ionut-Gabriel Farcas

Cambridge's Tea Party

No, it is not about the UK's Cambridge, home of the famous university with the same namesake, of Newton's Institute of Mathematical Sciences, English tea, or of the famous English weather. It is about Cambridge, Massachusetts, USA, host of the famous Ivy League university, Harvard, and of the similarly well-known and prestigious science and engineering school, Massachusetts Institute of Technology (MIT). These two exponents of academic grandeur are within two miles of each other on Massachusetts avenue, making Cambridge, Massachusetts one of the very top places in the world in terms of academia. Cambridge is essentially across the bridge from Boston, the largest city in the state and well known for many things, including the famous Tea Party from 1773.

From March 01 to April 01, 2019, I was fortunate enough to conduct a third research stay at the famous Institute of Technology. The beginning was not so easy. On March 1st, on a Friday, I ended a few hours earlier a successful SIAM Conference on Computational Science and Engineering, which took place across the country in Spokane, Washington. From there, very early in the morning, on a freezing, Canadian-like cold, I started my airborne journey towards Boston International. Bad weather was all over the North side of the States, and Boston was no exception. Throughout my first Sunday in Boston, there was a heavy snow storm. heavy enough to make it subject for the news, but not heavy enough to prevent the Bostonians to carry about their life. The aftermath of the storm did not prevent me to cross the Charles river via the Harvard bridge (yes, one way to get to MIT from Boston is to cross the Harvard bridge) and have my first day as visiting researcher at MIT on March 03.

Without exaggerating, MIT is a unique, mesmerizing place. Even if we leave aside its history and the impressive number of Nobel and Fields Medal laureates, MIT feels, from the moment you enter in any of the buildings, like a place which changes the world. Therein, I visited the Center for Computational Engineering led my Professor Youssef Marzouk, which resides in Building 37, in the Kavli Institute for Astrophysics and Space research. Prof.

Marzouk and his group carry out very diverse research in Computational Science and Applied Mathematics, including Bayesian inference, uncertainty propagation, optimization under uncertainty, or optimal experimental design.

My interest in visiting the group of Youssef Marzouk was mainly due to his expertise in so-called transport maps. Prof. Marzouk was the first researchers to employ techniques from optimal transport in Bayesian inference with the goal of efficiently addressing complex and computationally expensive inference problems. Bayesian inference is, broadly speaking, a way of probabilistic thinking in which a priori information is updated in the presence of new (and useful) knowledge, in the sense that the initial uncertainty about the quantity to be inferred is reduced (and quantified) as one acquires more information. In Scientific Computing, the quantities we wish to infer are parameters of physics or engineering models, while the new knowledge is observational data, usually stemming from experiments. The Bayesian framework then updates a prior probability measure on the input parameters, stemming from expert opinion, a priori available information, or the “best available guess” to a posterior measure, the solution of the inference problem, via the observational data. Inverse problems are notoriously difficult due to several aspects, one being the exploration of the posterior solution. Standard techniques are based on sampling, but their cost/accuracy ratio deteriorates significantly in high-dimensional settings. Transport maps alleviate this issue by trading off the sampling procedure with an optimization problem in which a map which transports samples from a known measure to the target, posterior measure, is sought.

One of the projects of my doctoral work is the development of a sampling-free, deterministic, multilevel, adaptive methodology based on sparse grid approximations for Bayesian inference in higher-dimensional, complex settings. The power and usefulness of the proposed approach was demonstrated in several Bayesian inference problems. However, our algorithm was not designed for problems in which the posterior density has more than one mode, i.e., in which the posterior has several local minima and/or maxima. These situations are, of course, relevant in practical applications, in which often the observational data and hence the corresponding posterior are multimodal. To this end, we wanted to extend our current approach to handle arbitrary

nonlinear, multimodal inverse problems. This is where the expertise of Prof. Marzouk in optimal transport comes into play. Our preliminary results are very promising and we are now able to handle arbitrarily nonlinear, unimodal inference problems. To handle arbitrary multimodal Bayesian inversion, our current efforts are invested in extending the transport map framework to include parametrizations in terms of bases that can capture the multimodality of the posterior. Moreover, in terms of practical applications, we wish to apply the proposed framework in plasma microturbulence analysis: starting from experimental data from tokamak experiments, we want to apply our framework for Bayesian inversion to obtain a probabilistic description of physical parameters characterizing the plasma particle species. Afterwards, we employ the Bayesian posterior solution in forward simulations with the goal of obtaining predictions about the microturbulence.



Figure 1: Myself, a puddle of mud in Killian Court, and the famous Dome’s entrance at MIT.

Overall, my stay at MIT was a great experience, not only scientifically, but also personally. MIT is a place in which it is difficult not to be or not to get inspired. It is a place full of creative and innovative people, in which, everyday something important is happening. Personally, it was great that the MIT environment felt like home and that my research ideas were embraced as being valuable and promising. With this visit, I started a collaboration with a great research group and I am keen to see where our research collaboration leads us. Last but not least, these research visits are also an opportunity to see that students are scientists from the Technical University of Munich are well respected around the world and that we are, essentially, at a similar level with the best!

Ionut-Gabriel Farcas

Neue Version von LIKWID



Die Werkzeugsammlung für performance-orientierte Programmierer und Anwender LIKWID („Like I Knew What I’m Doing“) ist in Version 4.3.4 erschienen². LIKWID bietet, neben Informationen wie beispielsweise der Systemtopologie und CPU Taktraten, Unterstützung für sogenannte Hardware Performance Counter zum Abfragen von Zählern auf Hardwareebene. Weitere Werkzeuge erlauben die Manipulation von Hardwareeigenschaften wie Taktraten und Prefetcher und die Kontrolle der Platzierung von Applikationsprozessen auf Prozessorkernen. Die Funktionalität steht sowohl als Kommandozeilenapplikation als auch zum Einbetten in eigene Programme als Bibliothek zur Verfügung. Die aktuelle Version unterstützt die letzten Intel Prozessoren mit der Mikroarchitektur „Cascadelake SP“ und die aktuellen AMD Prozessoren (Epyc, Threadripper).

In der Roadmap von LIKWID ist ein Major Release im Herbst geplant und soll die von LIKWID unterstützten Architekturen um ARMv8 und POWER9 CPUs erweitern. Als größte Neuerung sollen NVIDIA GPUs unterstützt werden. Alle drei Erweiterungen befinden sich bereits in der Entwicklung und stehen bereits für Experimente zur Verfügung.

Thomas Gruber

²<https://github.com/RRZE-HPC/likwid>

Marvell/Cavium ThunderX2 am RRZE



Vor nicht allzu langer Zeit waren Prozessoren auf Basis der ARM-Architektur das lang ersehnte Licht am Ende des Tunnels beim energieeffizienten Supercomputing: losgelöst von all dem staubigen x86-Ballast konnte man angeblich mit viel weniger Transistoren und in ein paar Jahren das erreichen, wozu Intel Jahrzehnte brauchte und auf dem Weg sage und schreibe drei teure Prozessorprojekte³ einstampfen musste. Gut sichtbare Aktivitäten wie das Mont-Blanc-Projekt und unzählige Publikationen mit unverstandenen Performance- und Energiemessungen ließen hoffen. Und endlich, nach Jahren des gespannten Wartens, kam die kleine Firma Cavium (jetzt Marvell) aus dem sonnigen San Jose und gab uns den ThunderX2 SoC. Nicht nur, aber auch weil dieses Rechenherz nun in diversen Top500-Maschinen weltweit schlägt und auch von Cray in ihrer XC50-Linie angeboten wird, hat das RRZE schon vor einigen Monaten einen Knoten mit zwei Prozessoren des Typs „ThunderX2 9980-2200“ beschafft:

- ARM v8.1 ISA mit NEON SIMD-Erweiterungen (16 Byte Registerbreite)
- 2.2 GHz Taktfrequenz, 32 Cores pro Sockel
- 4 Hardware-Threads pro Core
- 8 Flops (2 FMAs) pro Taktzyklus in doppelter Genauigkeit
- 32 KiB L1-Cache und 256 KiB L2-Cache pro Core
- 32 MiB gemeinsamer L3-Cache pro Chip
- 64 GiB RAM pro Sockel an acht DDR4-2666 Speicherkanälen
- Thermal Design Power 180 W pro Chip

³Falls es jemanden interessiert: iAPX 432, i860 und Itanic. Pardon, Itanium. Den Xeon Phi zählen wir jetzt mal nicht mit, denn der Schmerz ist noch zu frisch und sitzt tief.

Bei einer gemessenen Speicherbandbreite von 127 GB/s pro Sockel (das sind ca. 75% vom theoretischen Maximum) kommt man damit auf eine Maschinenbalance von 0.23 Byte/Flop; gar nicht so schlecht gemessen an aktuellen Intel-Boliden, die so bei 0.06 Byte/Flop landen. Das liegt natürlich an der eher schmalen SIMD-Breite, aber auch die absolute Speicherbandbreite kann sich sehen lassen. Das Ganze kommt freilich zu einem gewissen Preis – nein, nicht der Kaufpreis, der ist tatsächlich konkurrenzfähig, sondern der Energieverbrauch: 180 Watt ist schon ein Wort, da purzeln die Joules schneller aus dem Server als man „Energieeffizienz“ sagen kann. Wenn man oben mitmischen will, muss man offenbar Kompromisse machen, aber zumindest gibt es jetzt außer AMD noch einen weiteren ernstzunehmenden Mitbewerber im Geschäft.

Wie sieht es mit den Tools aus? Bei den Compilern gibt es eine gute Auswahl: GCC, Clang und eine kommerzielle Toolsuite von ARM auf Basis von Clang sind verfügbar. Erste Tests legen nahe, dass sich alle Compiler noch schwer tun mit hocheffizientem numerischem Code, speziell wenn SIMD beteiligt ist. Auch bekommt man die maximale Speicherbandbreite nur unter Verwendung von Prefetch-Instruktionen, aber das ist nichts, was der Compiler von selbst tut. Ergo ist noch ein gewisser Weg zu gehen, bis man auf dem Niveau des Intel-Compilers angelangt ist. ARM hat dem RRZE freundlicherweise eine kostenlose Testlizenz zur Verfügung gestellt, und auf der nahen ISC 2019 wird es einen intensiven Austausch geben.

Die LIKWID-Suite⁴, am RRZE entwickelt und gepflegt und mittlerweile weltweit im Einsatz, wird natürlich ebenfalls auf den Chip portiert. Fans der „Bleeding Edge“ können den aktuellen Stand aus dem Master-Zweig auf github auschecken und auf Herz und Nieren testen. Ein neues Release mit offiziellem Support für ARM-Architekturen (auch jenseits TX2) ist vor der Supercomputing in Denver geplant.

Das RRZE stellt interessierten Forschern die Maschine gerne zum Testen zur Verfügung.

Georg Hager

⁴<https://github.com/RRZE-HPC/likwid>

Neues vom KONWIHR



Mit der zum 1.3.2019 abgelaufenen Antragsrunde gab es auch beim Förderprogramm eine Neuerung. Um die KONWIHR Familie breiter aufzustellen und die bayerische Vernetzung im HPC zu verbessern, wurden erstmalig sog. „Basisprojekte“ ausgeschrieben. Mit einer einmaligen Förderung von € 10.000 sollen bayerische Hochschulen in die Lage versetzt werden, zentrale Ansprechpartner zu den Themen HPC und KONWIHR zu etablieren. Diese sollen quasi als HPC-Wegweiser für Wissenschaftler mit hohem Rechenzeitbedarf dienen und

- etwa bei Fragen zu verfügbaren HPC-Ressourcen (in Bayern und auch darüber hinaus),
- HPC-Förderung (KONWIHR und auch darüber hinaus) oder
- HPC-Weiterbildung (in Bayern) Hilfestellung geben.

Erfreulicherweise wurden bereits beim ersten „Call“ drei Anträge eingereicht, die positiv begutachtet wurden und dem Ministerium zur Förderung vorliegen. Es sieht also gut aus, dass sich bald neben den KONWIHR Tankern in München und Erlangen, neue und schnittige Yachten in bayerischen HPC-Gewässern tummeln.

Weitere Informationen zum neuen Förderprogramm sind im neuen Informationsflyer auf der Webseite des KONWIHRs zu finden: www.konwihr.de

Gerhard Wellein

*** Notiz * Notiz * Notiz ***

Termine 2019

- **Society for Industrial and Applied Mathematics:**
SIAM Conference on Applied Algebraic Geometry (AG19) –
SIAM AG19 in Bern, Switzerland:
09.07.-13.07.2019 <https://www.siam.org/Conferences/CM/Main/ag19>
- **Supercomputing 2019:**
The International Conference for High Performance Computing, Networking, Storage and Analysis (SC19) –
SC 19 in Colorado, Denver, USA:
17.11.-22.11.2019 <https://sc19.supercomputing.org/>
- **Society for Industrial and Applied Mathematics:**
SIAM Conference on Uncertainty Quantification (UQ20) –
SIAM UQ20 in Garching, Germany:
24.03.-27.03.2020 <https://www.siam.org/Conferences/CM/Conference/uq20>

Quartl^{*} - Impressum

Herausgeber:

Prof. Dr. A. Bode, Prof. Dr. H.-J. Bungartz, Prof. Dr. U. Rüde

Redaktion:

S. Herrmann, S. Seckler, Dr. S. Zimmer

Technische Universität München, Fakultät für Informatik

Boltzmannstr. 3, 85748 Garching b. München

Tel./Fax: ++49-89-289 18611 / 18607

e-mail: herrmasa@in.tum.de, **www:** <http://www5.in.tum.de/quartl>

Redaktionsschluss für die nächste Ausgabe: **01.09.2019**

* **Quartel**: früheres bayerisches Flüssigkeitsmaß,

→ das **Quart**: $1/4$ Kanne = 0.27 l

(Brockhaus Enzyklopädie 1972)