Chapter 1

An Introduction to High Performance Computing and Its Intersection with Advances in Modeling Rare Earth Elements and Actinides

Deborah A. Penchoff,[*,1,2,3] Edward Valeev,[4] Heike Jagode,[1] Piotr Luszczek,[1] Anthony Danalis,[1] George Bosilca,[1] Robert J. Harrison,[5,6] Jack Dongarra,[1] and Theresa L. Windus[7,8]

[1]Innovative Computing Laboratory, University of Tennessee, 1122 Volunteer Boulevard, Knoxville, Tennessee 37996, United States

[2]Department of Nuclear Engineering, University of Tennessee, 1412 Circle Drive, Knoxville, Tennessee 37996, United States

[3]Howard H. Baker Jr. Center for Public Policy, University of Tennessee, 1640 Cumberland Avenue, Tennessee 37996, Unites States

[4]Department of Chemistry, Virginia Tech, 421G Davidson Hall, Blacksburg, Virginia, 24061, United States

[5]Institute for Advanced Computational Science, Stony Brook University, 100 Nicholls Road, Stony Brook, New York 11790, United States

[6]Computational Sciences Center, Brookhaven National Laboratory, Building 463B, Upton, New York 11973, United States

[7]Department of Chemistry, Iowa State University, 2415 Osborn Drive, 1605 Gilman Hall, Ames, Iowa 50011, United States

[8]Ames Laboratory, 125 Spedding Hall, Ames, Iowa 50011, United States

[*]Email: dpenchof@utk.edu

Computationally driven solutions in nuclear and radiochemistry heavily depend on efficient modeling of Rare Earth Elements (REEs) and actinides. Accurate modeling of REEs and actinides faces challenges stemming from limitations from an imbalanced hardware-software ecosystem and its implications on inefficient use of High Performance Computing (HPC). This chapter provides a historical perspective on the evolution of HPC hardware, its intersectionality with domain sciences, the importance of benchmarks for performance, and an overview of challenges and advances in modeling REEs and actinides. This chapter intends to provide an introduction for researchers at the intersection of scientific computing, software development for HPC, and applied computational modeling of REEs and actinides. The chapter is structured in five sections. First, the *Introduction* includes subsections focusing on the *Importance of REEs and Actinides* (1.1), *Hardware, Software, and the HPC Ecosystem* (1.2), and *Electronic Structure Modeling of REEs*

Penchoff et al.; Rare Earth Elements and Actinides: Progress in Computational Science Applications
ACS Symposium Series; American Chemical Society: Washington, DC, 2021.

and Actinides (1.3). Second, a section in *High Performance Computing* focuses on the *TOP500* (2.1), *HPC Performance* (2.2), *HPC Benchmarks: Processing, Bandwidth, and Latency* (2.3), and *HPC Benchmarks and their Relationship to Chemical Modeling* (2.4). Third, the *Software Challenges and Advances* focus on *NWChem/NWChemEx* (3.1), *MADNESS* (3.2), and *MPQC* (3.3). The fourth section provides a short overview of *Artificial Intelligence* in HPC applications relevant to nuclear and radiochemistry. The fifth section illustrates *A Protocol to Evaluate Complexation Preferences in Separations of REEs and Actinides through Computational Modeling*.

1. Introduction

High Performance Computing (HPC) has enabled advances in science and technology with direct benefit to society. Since the creation of the first supercomputers, the ability to apply computational models has increased dramatically. It has allowed discoveries and design of new medical treatments, more efficient engines, design of airplanes and rockets, development of telecommunication devices and infrastructure, and population and climate models. HPC has also facilitated advances in data science and artificial intelligence applications, which is setting a fast-paced trajectory to new findings and advances with global impact. HPC can significantly accelerate finding solutions for modeling Rare Earth Elements (REEs) and actinides; however, even with recent advances in supercomputing resources, the HPC hardware-software ecosystem still needs multidisciplinary synergy to provide robust and practical solutions for the specific challenges in modeling REEs and actinides. *[Note: Rare Earth Elements (REEs) is the name given to a set of 17 chemical elements, which include scandium, yttrium, and the lanthanide series (lanthanum, cerium, praseodymium, neodymium, promethium, samarium, europium, gadolinium, terbium, dysprosium, holmium, erbium, thulium, ytterbium, and lutetium). Actinides is the name given to a set of 15 chemical elements, which includes actinium, thorium, protactinium, uranium, neptunium, plutonium, americium, curium, berkelium, californium, einsteinium, fermium, mendelevium, nobelium, and lawrencium.]*

This chapter explores the importance of REEs and actinides, discusses challenges modeling REEs and actinides utilizing electronic structure descriptions, and provides an overview of the hardware, software, and the HPC ecosystem. It also includes an introduction to HPC, its metrics, and evaluations. The challenges and advances in modeling REEs and actinides discussed in this chapter focus mainly on applications relevant to molecular systems in gas and solvated phases. HPC modeling of extended phases relevant in areas including materials science, catalysis, and various energy applications are not specifically addressed in the chapter.

1.1. Importance of REEs and Actinides

REEs and actinides are essential to the growth and health of the U.S. economy and directly impact national security. REEs are critical in a variety of needs, including advanced technological applications such as electronics, cell phones, computers, satellites, lasers, energy production and storage, classical and quantum information processing, military applications, screens, wastewater treatment, radiation monitoring, clean energy, catalysis, and medical applications (*1*–*6*). The actinides are crucial in defense and energy applications and medical treatments (*7*–*12*).

Several reports to Congress (*2*, *3*) and other legislative materials (*13*–*39*) have highlighted the criticality of finding long-term solutions for the shortage of REEs and the implications of insufficient national reserves and stockpiles on society, the economy, and national defense (*40*). In 2011, the U.S. Department of Energy (DOE) built a criticality matrix in which 16 elements essential to clean energy were classified as critical, near-critical, and not critical with respect to availability until 2025 (*41*). From the 16 elements crucial to clean energy, nine are REEs, with five being declared as critical (*41*). Since 2010, the Department of Defense has provided several reports (*42*–*47*) from Congressional mandates focusing on REEs, highlighting the need to address REE shortages as a national security risk. According to the 2020 U.S. Geological Survey, the U.S. only possesses cerium, lanthanum, and REE magnet feedstock as potential acquisitions in its stockpile (*48*). At present, many countries, including the U.S., are mainly dependent on China for REEs (*49*). The U.S. imports 80% of its REEs from China, 6% from Estonia, 3% from Japan and Malaysia, and 8% from other parts of the world (*48*). High dependence on foreign sources for REEs can create socioeconomic and national security instability. In particular, importing most REE resources from China presents concerns due to China's ability for sudden restrictions of exports (*50*–*52*).

Optimization of separations of REEs and actinides is critical to guarantee a supply of REEs and actinides, and for purification processes needed for radiological waste management and environmental concerns. Selective separation of lanthanides is difficult due to the similar chemistries presented by these elements, the need for expert facilities to handle radiation, and safety concerns (*53*). The need for novel solutions to this challenge is clear. It can be addressed by accelerating the understanding and development of selective extracting processes for efficient separation from ores and recycling.

Trial-and-error experimental approaches to advance solutions for REEs and actinides present challenges due to limitations in the availability of materials and the need for expert facilities to handle radioactive elements. Therefore, computational modeling is essential to advance solutions for REEs and actinides. HPC-enabling capabilities are critical to provide application-ready tools to the community.

1.2. Hardware, Software, and the HPC Ecosystem

Advances in supercomputing often highlight hardware. There is much excitement waiting for reports such as the biannual TOP500 (*54*–*57*) to reveal the world's fastest supercomputers. Computational scientists often seek to utilize the fastest and newest supercomputers; however, computational modeling and simulation progress depends on a complex software ecosystem (*58*). Even though advances in hardware provide possibilities of massive scalability, the software needs to be efficiently parallelized and optimized for the simulation to take full advantage of powerful new hardware. Physical descriptions of systems of interest in applied scientific research are described algorithmically and further encoded in software packages. These packages not only require development, but they also need to be maintained, ported, and enhanced to be able to be utilized in varying architectures as new hardware is available (i.e., more recent "supercomputers"). The software often outlasts the hardware for which they were initially designed by years or decades (*58*). In addition, new theoretical and computational models are just as important as the hardware to advance the scientific and technological enterprise.

Software development commonly utilized to model REEs and actinides has not advanced at the same speed as hardware, which causes computational scientists in the community to be unable to take advantage of newer supercomputers' power. This causes computational scientists focusing on

nuclear- and radio-chemical modeling often to have to choose between qualitative and quantitative descriptions. A dynamic develops in which choosing methods to predict properties of choice is usually decided based on the system's size (such as the number of atoms in a molecular compound). Even when the hardware allows for hundreds of thousands of cores to be available to the user, if the software cannot effectively take advantage of the available hardware, the limiting factor for the user becomes the molecular or material system and not the supercomputing resources. When modeling REEs and actinides, intrinsic chemical challenges in accurate descriptions of the electronic structure is further complicated by the increasing number of electrons and relativistic considerations. For the specific need of utilizing HPC to optimize separations of REEs and actinides, the analysis often requires evaluating dozens to hundreds of chemical reactions including large molecular compounds and complex environments including varying solvents, temperatures, acidity, and interactions with extracting agents (such as in ion exchanges or solvent extractions).

Next-generation computational simulation and modeling demand creating and promulgating a new paradigm in software development in a highly multidisciplinary ecosystem (*58*). Recent efforts by the DOE in the Exascale Computing Project (DOE-ECP) (*59*) highlight the need for focusing on hardware design, software development, and hardware-software integration to provide computational scientists the ability to efficiently utilize resources at the exascale level. The DOE-ECP is based on a holistic approach that includes co-design and integration of application development, software technologies, and hardware technologies to achieve exascale capabilities. This is a $3.6 billion effort to help DOE advance needs in national, energy and economic security, scientific discoveries, and healthcare.

1.3. Electronic Structure Modeling of REEs and Actinides

The applications of REEs and actinides previously described underscores the critical need for understanding and controlling the electrons in REE- and actinide-containing compounds and materials. The emergence of practical and robust first-principles (beyond density functional theory) models of correlated electrons are starting to transform the role of computation across light-element chemistry and materials science, from rationalization to prediction (*60–62*).

Several attributes are essential to achieve chemically-predictive electronic structure simulation for REEs and actinides. These include:

- fully relativistic (Dirac) treatment of all electrons interacting via the classical relativistic electromagnetic field,
- infinite-order treatment of dynamical electron correlation, exemplified by the nonlinear ansatz of the coupled-cluster (CC) method,
- frequent need to use multideterminant references when describing even ground states, and
- robust control of the numerical (basis set) errors despite the high angular momenta of the occupied (bi)spinors, the vast range of energy/length-scales involved, and the inefficient modeling of dynamical correlation effects using spinor products (particularly, in the relativistic regime).

While the challenges of the first three items on this list are formidable, but well understood, the last item is relatively less explored, yet is as critical as the others. The understanding of the Gaussian atomic orbital (AO) basis set technology – namely, how to design systematically convergent families of basis sets and how to evaluate quantum operators in them efficiently and precisely – is less mature for the relativistic all-electron computations compared to the non-relativistic counterparts. This is

especially so for the many-body methods, where relativistic effects impart even slower convergence to the basis set limit than the already painfully-slow rate of the non-relativistic correlation energy.

However, recent work hints that the numerical challenges of the relativistic electronic structure can be finally addressed robustly by abandoning the Gaussian AO technology altogether:

- Harrison and co-workers demonstrated for the first time how Dirac-Hartree-Fock-Coulomb (DHFC) equations can be solved for general molecules by using an adaptively-refined real-space numerical representation of spinors, without any need for ad hoc numerical safety measures (i.e., the kinetic balance) and with the level of precision control unmatched by Gaussians (*63*).
- Valeev and co-workers demonstrated for the first time how many-body electronic structure methods can be efficiently formulated in real space for general molecules with dozens of atoms and with precision superior to Gaussians, by combining ideas of explicit correlation (F12) and pair-natural orbitals (PNO) (*64*).

These breakthroughs suggest that it should be possible to formulate accurate relativistic many-body methods with numerical errors robustly controlled. Such ability could, first, provide crucial reference benchmarks for molecular relativistic computations. Second, these advances hint that real-space numerical representations can provide a practical alternative to the Gaussian AO numerical technology that rectifies its known failings for computing highly-precise energies and properties.

Quantitatively accurate all-electron relativistic descriptions of electronic structure in REEs and actinides are going to be significantly more expensive than the non-relativistic and/or light-element counterparts, in part due to the physics and in part due to the sheer number of electrons that need to be described explicitly in such systems. To ensure practicality, it is therefore imperative to be able to deploy these proposed relativistic methods to the modern HPC platforms.

For relativistic all-electron theory, the motivation for efficiently exploiting modern HPC hardware (characterized by emphasis on data and operation parallelism and pervasive heterogeneity) goes beyond making optimal use of the available resources. It boils down to reducing time-to-solution to tighten the feedback loop of the computational experiments.

Modern HPC platforms nowadays are almost universally:

(1) heterogeneous, accelerated by inclusion of general purpose graphical processing units (GPGPUs) or custom-purpose hardware (e.g., tensor processing units), and
(2) massively-parallel, architected as clusters of nodes with one or more accelerators in them.

Such platforms have become synonymous with HPC, yet very few electronic structure codes can take advantage of them, primarily due to the low level of abstraction of the non-standard and non-portable programming models (CUDA and its close cousins) targeting the disjoint-address-space accelerated hardware.

2. High Performance Computing

Since 1993, trends in HPC have been reported by the TOP500 (*54*–*57*) project. The TOP500 provides a list of the 500 fastest supercomputers globally, and it is published twice a year. The order of the supercomputers on this list is determined by their performance when applying the LINPACK benchmark (*57*).

2.1. TOP500

The 56th edition of the TOP500 (*54–57*) list, published in November 2020 indicated the Fugaku supercomputer continues to top the chart. The list observed a flattening in performance over the top supercomputers globally. The list included the least number of new entries on the chart since it was initiated in 1993 (*54*).

The LINPACK benchmark revealed that the list's entry level increased from 1.23 petaflops in the 55th edition to 1.32 petaflops. Overall, the 500 supercomputers on the list increased from 2.22 exaflops in the previous list to 2.43 exaflops in the current list. The performance of the top (N=1) and bottom (N=500) system on the list and the aggregate (SUM) is shown in Figure 1 (*54*).

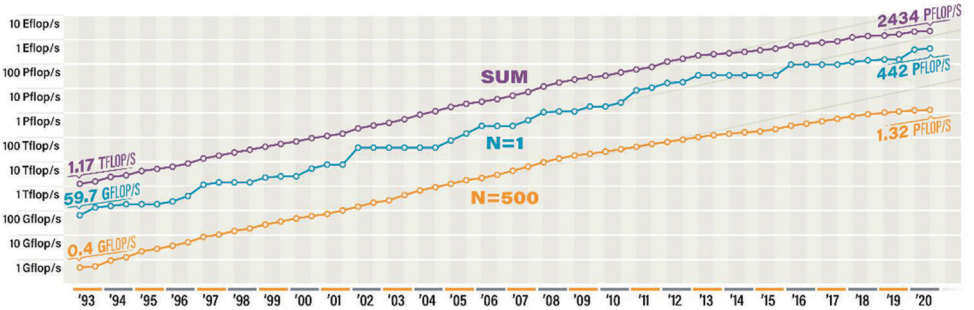

Figure 1. Performance development and projection of HPC between 1993 and 2020 from the TOP500. Data from references (54–57).

The 56th TOP500 list shows the first record of performance above one exaflop. The number one system on the list, Fugaku, increased its performance to 2.0 exaflops (calculated with the HPC-AI benchmark) (*54*). The top 10 systems on the list represent 42% of the total performance of the TOP500, and they are listed in Table 1.

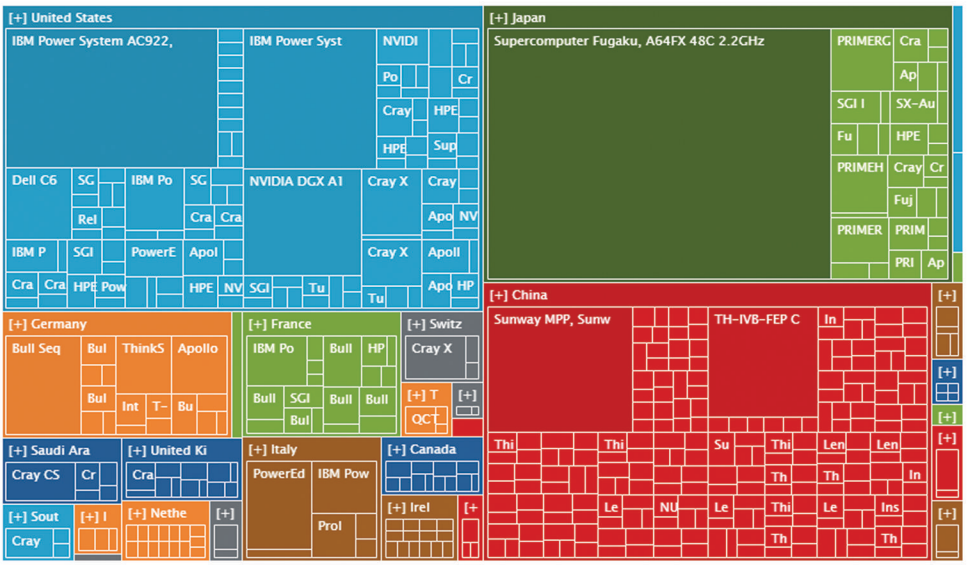

Figure 2. Treemap showing the total performance per country of the TOP500 list. Reproduced with permission from reference (54). Copyright 2020 Top500.

Table 1. Top 10 Systems in the 56th TOP500 List[a]

Rank	*Computer*	*Site*	*Country*	*Cores*	*Rmax (Petaflop/s)*	*Rpeak (Petaflop/s)*	*Power (MW)*
1	Supercomputer Fugaku	RIKEN Center for Computational Science	Japan	7,630,848	442	537	29
2	Summit	Oak Ridge National Laboratory	United States	2,414,592	148	200	10
3	Sierra	Lawrence Livermore National Laboratory	United States	1,572,480	94	125	7,4
4	Sunway TaihuLight	National Supercomputing Center in Wuxi	China	10,649,600	93	125	15
5	Selene	NVIDIA Corporation	United States	555,520	63	79	2,6
6	Tianhe-2A	National Super Computer Center in Guangzhou	China	4,981,760	61	100	18
7	JUWELS Booster Module	Forschungs-zentrum Juelich	Germany	449,280	44	70	1,7
8	HPC5	Eni S.pA.	Italy	669,760	35	51	2,2
9	Frontera	Texas Advanced Computing, Univ. of Texas	United States	448,448	23	38	
10	Dammam-7	Saudi Aramco	Saudi Arabia	672,520	22	55	

[a] Data from reference (*54*).

China has 42.8% of the systems in the 56th TOP500 list, followed by the United States with 22.6%, and Japan with 6.8%. The United States leads with 27.5% of the systems, Japan follows at 24.4%, and China has 23.3%. The treemap shown in Figure 2 shows the share of the 56th TOP500 list by country.

Between the first TOP500 list (published in June 1993) and the latest one (published in November 2020), the United States has had the number one machine 55% of the time (i.e., 31 times), Japan 25% (14 times), and China 20% (11 times), as shown in Figure 3 (*54*). The complete list of number one machines is included in Appendices 1 and 2.

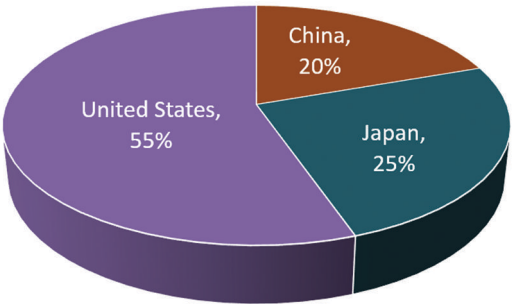

Figure 3. Percentage of number one machines in the TOP500 list between 1993 and 2020. Data from reference (*54*).

2.2. HPC Benchmarks: Processing, Bandwidth, and Latency

HPC benchmarks focus on measuring the three primary metrics on modern supercomputers: information processing rate, data transfer rate, and the delay in signaling across the units of the machine be it the registers, memory hierarchy levels, sockets, or compute nodes. There are three main approaches to influence these metrics, namely improvement, optimization, and mitigation. Improvements may be achieved by raw increases in hardware speed including clock rate frequency, specialization of compute units for critical tasks, and abundant parallelism of functional units or the chipset pins and communication lanes. Optimization can also happen at the hardware level when the hardware vendors iterate on their product roadmap over past designs to refine the chip design based on the real-life usage scenarios of their processors. But a more important optimization work takes place on the software side as the scientific code is ported and then optimized for new hardware to better the strengths of new platforms while relying less on their weaknesses. Finally, mitigation is the last resort effort to overcome the inherent limitations of computing such as latency. One typical example here would be latency hiding that performs the waiting for completion of work or arrival of data in parallel with other helpful task. The balance between these three metrics has changed dramatically over the decades. The early Cray computer models were designed with overprovisioned memory bandwidth and a comparative scarcity of computational power. Comparing these early designs with the state-of-the-art in GPU hardware cannot offer a starker contrast. Some of the compute units measure their compute power in hundreds of tera-operations per second. At the same time, the extremely wide interface to the memory can deliver only on the order of 1 terabyte per second of data transfer rates at over 5,000 bits devoted to the main memory bus.

Computers, or more so supercomputers, may be considered tools of scientific discovery. This gave rise to the emergence of computational science complementing the scientific endeavor's theoretical and experimental branches. From the perspective of these scientific applications, a significant figure of merit is the discovery per unit of time, be it the number of gene sequence

matches per second or the simulated nano-seconds of molecular dynamics per hour of compute time – application benchmarks produce those kinds of results. They bundle the implementation and the input data to deliver information in the application context. Examples of these include the NAS Parallel Benchmarks (*65*–*68*) with multiple input data set classes labeled with letters A, B, C, and D, as well as the SPEC consortium that licenses a plethora of benchmarking suites for a variety of testing scenarios such as integer and floating-point computations called SPECint and SPECfp, respectively, and are collected in SPEC CPU2006 (*69*, *70*). The complexity of the implementation and the limited data set sizes lead to the development of efforts on simpler performance probes.

To avoid the complexities of complete applications, mini-apps have been developed to provide relevant representatives that are easier to manage, port, and run with the results remaining interpretable in the context of fully deployed applications. Mantevo (*71*) is an example collection of mini-apps that included representatives of applications of interest to the DOE. Due to a reduced footprint of the mini-apps' code and its functionality, the input data is also much reduced and is more manageable for runs at different scales. Mini-apps may be considered a modern replacement of the collection of representatives loops such as Livermore Loops (*72*, *73*).

Further reduction of the complexity may be achieved with kernel benchmarks. They do away with resemblance to the complete applications or even mini-apps. They simply focus on portions of the code that consume most of the resources, be it time, performance, or available bandwidth. There are many examples of kernel benchmarks due to their relative simplicity when compared to any application code. High Performance LINPACK (HPL) (*74*) is potentially one of the most influential codes that are used to rank the supercomputers reported biannually on the TOP500 (*54*–*57*) list based on their performance in solving a system of dense linear system of equations using Gaussian Elimination with partial pivoting in double precision arithmetic. The Green500 (*75*) list measures the power used when running HPL and ranks supercomputers in energy efficiency for every floating-precision operation. Arguably, a better representation for PDE solvers is the High Performance Conjugate Gradients (HPCG) benchmark (*76*) that solves systems of sparse linear equations arising from a simple PDE and its 27-point stencil discretization. Ranking based on HPL and HPCG often diverge and offers a complementary ranking of large-scale machines. To take advantage of modern hardware with multi-precision hardware and large numbers of low-precision floating-point units, the HPL-AI benchmark was created as a combination of dense factorization in lower precision and an iterative scheme based on GMRES (*77*) to attain higher precision accuracy of the solution. Focusing on isolated kernels radically reduces complexity, improves portability, and can even be performed on prototype systems with a barely functioning hardware/software stack. However, they may indicate artificially elevated performance levels since a kernel may be easily scaled to maximize the utilization of the hardware beyond what is possible in the actual application code.

Finally, synthetic benchmarks completely detach from application-based metrics of merit. They primarily represent the interests of the hardware engineers or library developers. They focus on synthetic hardware or software behavior measures, for example, peak performance of floating-point units or minimum latency of accessing memory in the prescribed pattern of data requests. While synthetic benchmarks might seemingly be of little interest to scientific application developers, they offer valuable information to develop support components of high-end computer systems and offer a comparison between competing products and implementations.

Performance benchmarks are not only valuable for rank new machines, but they are also essential in hardware development (as proxies for full-fledged applications) and use (to ensure proper operations) (*78*). The hardware designers feed their nascent designs with past data of user codes but these are always from the past and ran on old hardware. Having up-to-date information requires up-

to-date results that can be obtained from simple kernel benchmarks. These benchmarks work even on unsupported systems that are far from production-ready and may never even make it beyond the experimental lab. For such a system, the benchmark's simplicity is at the premium, which makes kernel-based testing very valuable.

Figure 4 shows the performance of each kernel benchmarks HPL, HPCG, and HPL-AI with the peak double precision performance for the machines from the latest TOP500 list (*54–57*, *79*) issued in November 2020. A somewhat counter-intuitive line representing peak performance uses 64-bit floating-point capability of the systems. With modern hardware, much larger hardware levels are possible when using low-precision floating-point arithmetic such as FP16 or BF16. These new formats and their corresponding performance are captured by the HPL-AI benchmark and are shown in the figure as the highest attained execution rate. Rmax and Rpeak refers to the metric that ranks the machines on the TOP500 list. Rmax is the floating-point rate when solving a system of dense linear equations in double precision floating-point arithmetic. Rpeak is the peak performance of the machine in double precision that considers all processing units available in the hardware.

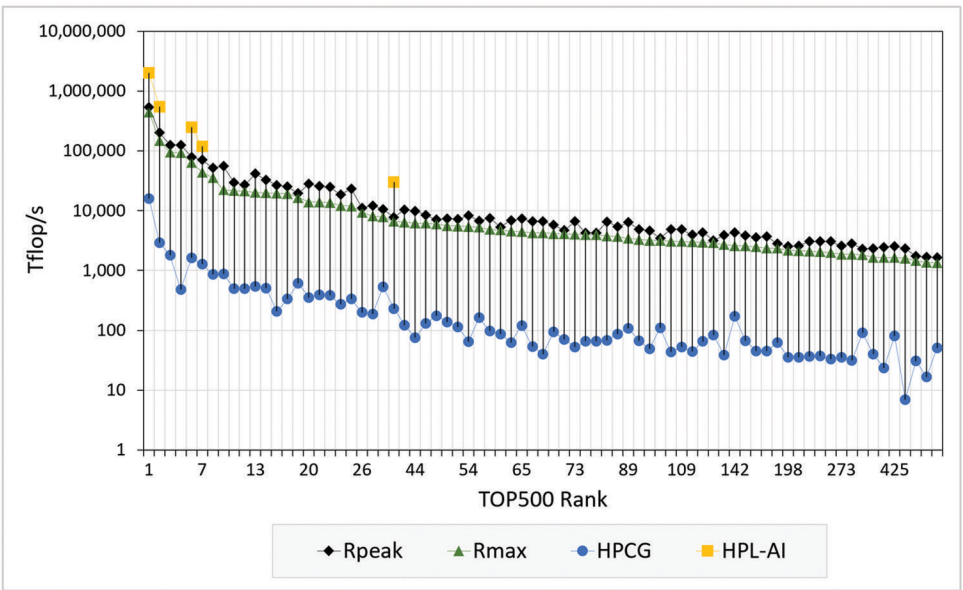

Figure 4. Performance of the supercomputers in the TOP500 56th list calculated with the HPCG and HPL-AI benchmarks. Data from reference (*79*).

2.3. HPC Performance

In the era of extreme scale platforms, understanding the performance characteristics of computational science applications is indispensable to identify and overcome the barriers to achieving performance goals. This becomes more important – albeit more and more cumbersome – as the architectures grow more complex. Particularly, with the profound changes in the hardware landscape, featuring increasing complexity and heterogeneity of today's and projected future high-performance computers, most of the computational chemistry and nuclear methods are unable to take full advantage of current computer resources at national supercomputer centers. As a result, these methods fail to sustain the expected performance scalability on these systems and are bound to drop back even more on next-generation systems. This section provides insights into the kind of performance data that can be measured on distributed, many-core, heterogeneous architectures

and how it can be used to effectively analyze, characterize, and optimize the performance of computational science applications.

Achieving performance goals does not merely target time-to-solution – although it is often used as a starting point – but implies the efficient use and utilization of the underlying hardware and software stack. If and how efficiently different hardware resources in heterogeneous extreme-scale systems are utilized can be detected with detailed measurements from a wide range of performance counters. The HPC community has relied on the Performance API (PAPI) monitoring library (*80*) to track low-level hardware operations for more than two decades. In addition to traditional CPU counters, such as floating point operations, retired instructions, cache access, and misses, PAPI's latest developments enable the monitoring of counters on GPUs (AMD, Intel, NVIDIA), communication networks, I/O systems, as well as energy monitoring with power capping features (*81*–*86*). The convenience of PAPI enables the collection of performance counter data from across the entire compute system with a single interface, and without putting the burden on the application developer to study, and incorporate various library primitives for accessing counters on different hardware.

In addition to providing hardware-specific performance metrics, PAPI has been extended with a standardizing layer for monitoring software-defined events (SDE) that expose the internal behavior of runtimes and libraries, such as communication and math libraries to the applications (*87*, *88*). The lack of information exchange between different libraries and blackbox-style "code entities" employed in an application can lead to suboptimal interaction between the software stack, which can lead to loss of performance. PAPI SDEs are meant to be used by the developers who write software libraries so that the internal behavior of their library can be better understood by application scientists who work with it. For example, a task scheduling runtime could export the number of available tasks at different points in time; and a climate-modeling code could export a performance metric such as "simulated years per second." PAPI enables monitoring of both hardware and software events through the same PAPI_start(), PAPI_read() interface, which allows tuning for more efficient use of heterogeneous hardware resources and presents a complete picture of the entire application performance.

In addition to using PAPI directly, application scientists can take advantage of performance frameworks that use PAPI for performance counter monitoring under the covers. Such tools employ hardware counter sampling, as well as call path and binary analysis. As a result, they can offer their users graphical profiles that attribute counter values and time spent to particular application code lines. Examples of such integrated performance frameworks are Arm MAP (*89*), HPCToolkit (*90*) and TAU (*91*).

Tracing is a different approach to analyzing application behavior. A trace shows a detailed timeline of performance metrics and the corresponding code that caused them. Vampir (*92*) is an integrated performance framework that utilizes tracing to display program behavior and allows users to view the trace at different zoom levels to acquire information at varying detail levels.

Suppose the target application is instrumented – manually or via provided compiler wrappers - using a tool such as Score-P (*93*), then information is collected at specific points in the application code, such as entries and exits of functions or other code regions of interest. This level of information acquisition, together with PAPI hardware and software metrics, gives an exact picture of the application behavior and can reveal rare bottlenecks missed by sampling since the latter is statistical. However, sampling has the benefit that it can be performed on unmodified binaries.

PAPI counters have been used with various quantum chemistry codes – primarily at the larger computer centers – either directly or with other optimization software such as TAU. In particular, the counters were extensively used at the Environmental Molecular Sciences Laboratory HPC system to

optimize NWChem and GAMESS performance. By examining regions of the code with low percent peak and high cache misses, data layout and localization as well as prefetching was used to improve the overall performance of the software.

2.4. HPC Performance, Benchmarks, and Their Relationship to Chemical Modeling

Since chemistry software often contains many different types of simulations, the full suite of benchmarks can help to inform the performance of the software on new architectures and possible optimizations in the software. For example, some parts of the code, such as the (T) part of the CCSD(T) will rely heavily on matrix multiplies that are often benchmarked on new systems. The HPL benchmarks can help to inform places in the software (such as the iterative coupled cluster and Hartree-Fock) since there are some similar data patterns. Benchmarks to understand the communication performance can aid in optimization of communication within parts of the code such as classical molecular dynamics and FFTs.

Most chemistry codes do not use % peak as a metric for performance, preferring to focus on time to solution. The major exception here is in the dense matrix multiplies that can reach very high % peak. However, there is often a correlation of the performance of the benchmarks between machines and the performance of the computational chemistry software on those machines. Of course, one must be careful to use appropriate benchmarks associated with the software, as described above.

The benchmarks discussed and the chemical modeling applications relevant to this discussion share computational patterns. For example, software discussed in the following section, such as MADNESS, operates on blocks with block-sparse representation and the benchmarks compute on blocks. Computing energy levels and wave function basis becomes an eigenvalue problem after linearization, and linear eigenvalue problems use matrix-matrix operations that are cache-friendly and exploit floating-point units effectively. This is also the case for the HPL benchmark. The real contrast is what percentage of peak each one achieves. As a generalization, HPL provides approximately 70% of peak, while symmetric eigenvalue problems are almost 50% (*94*).

Developing efficient methods for chemical modeling requires comparing and evaluating dataflow-based executions over coarse grain parallelism in terms of scalability, resource utilization, and programmability. For this, software-defined events (SDEs) are utilized for performance evaluation focusing on exposing performance-critical events. The PAPI library is often utilized for large-scale computational chemistry applications. It is essential to enable monitoring of both types of performance events—hardware- and software-related events—in a uniform way, through the same consistent PAPI interface.

To this day, Coarse Grained Parallelism (CGP) remains the most popular programming model for large scale scientific applications, where the software is commonly structured as serial code, and parallelism is achieved by injecting carefully implemented send and receive calls to the communication layer, and executing multiple parallel instances of the code. The execution of such applications on today's largest HPC systems shows that the CGP models already struggle to efficiently harness the parallelism available on systems with multiple cores, accelerators, and multiple non-uniform memory access (NUMA) hierarchies. As the community moves toward Exascale computing, with increasing parallelism, one of the challenges for the mainstream programming models is to sustain the expected performance scalability.

Further challenges emerge from the increasing complexity and heterogeneity of hardware architectures, where coarse-grained parallelism and fork-join expressions of parallelism limit performance portability of applications, and thereby sacrifice the productivity of scientific application developers.

Computational chemistry applications comprise one of the driving forces of HPC. In particular, computational methods—as they are present in quantum chemistry codes such as software resources discussed in the next section—are extremely compute intensive and consume significant computer resources at national supercomputer centers.

Hardware-Based Measurements

Analysis of hardware-based measurements is critical to identify sub-optimal parts of the software chain and unveil ways to improve performance and system utilization. This is essential to enable high-performance software in chemical modeling.

To illustrate these measurements, an example of performance profiles conducted with the Coupled-Cluster method (CCSD) in NWChem captured during the execution of two different data-flow implementations of CCSD is shown in Figure 5. These measurements evaluated network and CPU utilization. Specifically, the blue and purple boxes correspond to the execution of communication tasks, the red boxes correspond to the execution of computation tasks, and the gray area corresponds to idle time. The two implementations shown in Figure 5 differ in how the communication tasks are prioritized over the computation tasks. As a result, it is observable that when the priority of communication tasks is too high the network is overloaded with traffic leading to significant idle time at the beginning of the run. On the other hand, when task priorities allow for a more balanced execution of communication and computation tasks, the data transfers are overlapped with useful work and the hardware is utilized more efficiently.

Software-Based Measurements

The direct use of Software-defined Events (SDEs) in real-world applications can be highly beneficial for in-depth analysis or for quickly identifying performance and scalability bottlenecks.

Besides the PAPI's capabilities described, SDEs can be applied beyond libraries and performance tools. For example, PAPI's SDEs were applied in a study targeting the field of electronic structure theory focused on the implementation of SDEs to characterize the performance and the level of parallelism of the NWChem quantum chemistry application (*95*). Specifically, the Coupled Cluster Single Double (CCSD) methods (*96*) were evaluated as they are currently implemented in the Tensor Algebra for Many-body Methods (TAMM) library, which is expected to be part of the new C++ version of NWChemEx.

The findings of these software-based performance metrics for several CCSD kernels are discussed below in the "Usage Examples" section. Ultimately, the objective is to have computational chemistry experts, who are aware of the scientific function of different code segments, add SDEs that correspond to physically meaningful quantities that could reveal information, such as "computed electron potentials per second". Table 2 lists the SDE-registered NWChem events and their descriptions as they are available through PAPI. All four of the NWChem single-value SDEs are registered as 64-bit integer counters. As for the counting mode, DGEMM- and FlopCount are registered as delta counters---which means PAPI reports the difference between the counter value at PAPI_start() and PAPI_read()---while the Contraction_ID and the MaxTaskLength return instantaneous values—which means PAPI reports the absolute value of the counter every time PAPI_read() is called.

In addition to these single-value SDEs, support for two multi-values SDEs is included. For instance, instead of just monitoring how big the biggest CCSD task in the NWChem code is, which is accomplished via MaxTaskLength for each tensor contraction, with the addition of

LengthPerTask_RCRD, the length of all tasks per contraction and the distribution of task lengths can easily be recorded. Similarly, the FlopsPerTask_RCRD recorder produces important details about the number of floating-point operations (FLOPs) computed for each sequential task.

Usage Examples

The input data is a beta-carotene molecule ($C_{40}H_{56}$) in the 6-31G basis set, composed of 472 basis set functions. In these tests, all core electrons are kept frozen, and 296 electrons are correlated. The beta-carotene molecule serves as a strongly representative test case of large molecular complexes with a relatively small basis set. This is a fairly common use case for researchers. Thus, in that context, the computational resource requirements for beta-carotene, both from the memory and computational power perspective, represent a classic use-case scenario with large implications. Figure 6 shows the distribution of task lengths and FLOPs count for five contractions (presented by the different colors in the figure) with the highest workload in the case of beta-carotene. To calculate the load, the SDE counters that monitor the total number of floating-point operations for each CCSD contraction were used (e.g., FlopCount_I). For instance, the first contraction (t2_7_3) computes 327,680 DGEMMs (DgemmCount) and 32,557,729,032 FLOPs (32.5 GFLOPs) (FlopCount). The x-axis shows the various length of the tasks, meaning the number of sequential DGEMMs per tasks for each contraction. In the case of contraction (t2_7_3), there are two task lengths only, with 16 being the maximum number of sequential DGEMMs per task.

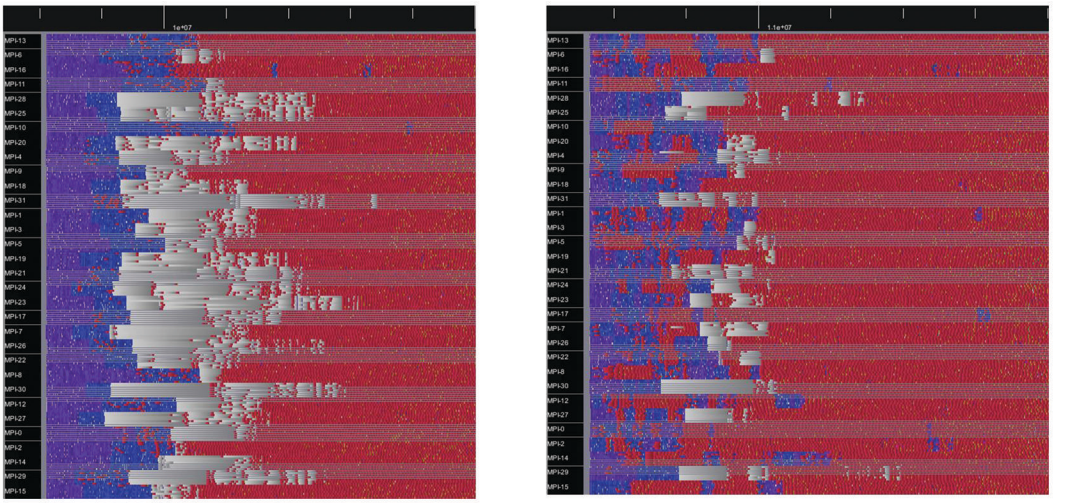

Figure 5. Trace of execution with idle time (left) and with better CPU utilization (right).

Data from the LengthPerTask_RCRD recorder is plotted by the bars, and the FlopsPerTask_RCRD data is shown by the labels on top of each bar. For instance, contraction t2_7_3, which happens to also be the most compute intensive portion of the code, has more than 24,000 tasks and approximately 2/3 of these tasks compute 16 sequential DGEMMs, while the other 8,000 compute eight sequential DGEMMs. Each of these 16,000 tasks performs over 20 GFLOPs and the remaining 8,000 perform about 10 GFLOPs.

In summary, these tasks are very expensive, which ultimately limits the scalability of CCSD because they execute serially—so they cannot take advantage of additional cores. It would be beneficial to break down the computation of the CCSD methods into more fine-grained tasks, so that

the serialization imposed by the traditional, linear algorithms can be transformed into parallelism, allowing the overall computation to scale to larger computational resources.

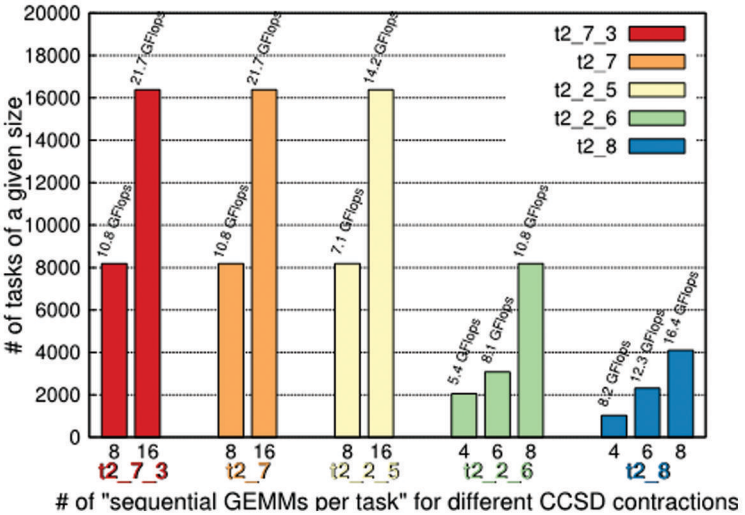

Figure 6. PAPI SDE-Recorder logging task lengths and FLOPs count for NWChem CCSD.

Table 2. Registered SDEs in TAMM/NWChem to Enable Users to Gather CCSD-Specific Details though PAPI

SDE Name (prefixed with sde:::TAMM::)	***SDE Description***
ContractionID_I	ID of CCSD contraction (I=integer)
DgemmCount_I::ContrID=46	Total number of DGEMMs for CCSD ContrID=46 (I=integer)
FlopCount_I::ContrID=46	Total number of floating-point operations for CCSD ContrID=46 (I=integer)
MaxTaskLength_I::ContrID=46	Maximum number of sequential DGEMMs per task for CCSD ContrID=46 (I=integer)
LengthPerTask_RCRD_I::ContrID=46	Array of the number of sequential DGEMMs per task for CCSD ContrID=46 (RCRD=recorder) (I=integer)
FlopsPerTask_RCRD_I::ContrID=46	Array of the number of FLOPs per task for CCSD ContrID=46 (RCRD=recorder) (I=integer)

The PAPI SDEs can be leveraged in an equivalent way for REEs to expose performance metrics that are specific to the chemical modeling of separations of adjacent REEs. Exporting these REE-specific SDEs through a standard middleware, as established as PAPI, creates the opportunity for the instrumentation to be reused by other developers and, more importantly, enables higher level toolkits (e.g., TAU , Scalasca , Vampir , Score-P, HPCToolkit) to take advantage of it and include it in their powerful analysis and visualization solutions without implementing a custom interface for each library. The convenience of PAPI enables the collection of REE software-related and HPC hardware-related performance counter data from across the entire compute system with a single interface, and without putting the burden on the computational scientist to study and incorporate

various library primitives for accessing counters on different systems. Usage examples on REE- and actinide-containing systems are future items to be explored for exascale implementation.

3. Software Challenges and Advances

There are many software packages often utilized to model chemical compounds and materials. Three software – NWChem/NWChemEx, MADNESS, and MPQC are discussed due to their ability to model REEs and actinides and their active development to provide enabling capabilities at the exascale level. Amongst the many software options, the authors discuss these three given their widespread utilization across chemical modeling scope and their scalability. These software status are highlighted with respect to modeling REEs and actinides.

3.1. NWChem/NWChemEx

NWChem (*97*) was designed and developed in the early 1990s to enable the use of high performance computers to solve grand challenge problems, especially those of importance to the DOE. At the time, the goals were to run efficiently and to scale well on gigaflop to teraflop computers. Over the years, the architecture of NWChem was adjusted to enable computations on petaflop computers. However, with the soon to be realized advent of exascale computers, it was realized that the old architecture of NWChem had significant issues that would make it difficult to use heterogeneous computers at scale. In particular, NWChem was not well suited to address the increased scale of parallelism, the greater heterogeneity of the processors, the complex memory-hierarchy, the limited ability to hide latency, and the necessary data abstraction and complexity for new reduced scaling methods.

Fortunately, DOE initiated the Exascale Computing Project (ECP) (*59*) to ensure that scalable scientific codes were available when the exascale computers came online. The NWChem team made the difficult decision to design and implement an updated version of NWChem, NWChemEx, to address the previously noted challenges. NWChemEx is designed with a focus on performance and flexibility by decoupling concerns as much as possible through the use of modules and object-oriented and functional programming styles. Each module is a self-contained, opaque, callable object adhering to one of many possible standardized application programming interfaces (APIs). Toward that end, a simulation development environment that enables separation of concerns and facilitates incorporation of new methods, libraries, and methods has been developed. The use of C++ instead of Fortran was seen as a positive step forward to enable this programming model and to allow for the use of many of the other mathematical and computer science capabilities that were being developed within ECP. The full redesign of NWChemEx allows us to flexibly address the issues discussed above that will be paramount when dealing with exascale computations.

Another important aspect of the software development is the use of software engineering practices that help to enable revision control, continuous integration testing, code coverage, code quality tools, documentation generators, software distribution, peer review, and software deployment. NWChemEx uses various tools to enable these capabilities with the core being GitHub, Catch2 for unit testing, Gcov, Gcovr, Codecov, TravisCI, Doxygen, Sphinx, CMake, and Docker. These tools all enable and support our code peer review that helps to provide internal mechanisms for software improvements.

While having much larger computers is important to perform more realistic simulations, new physical models are just as important. In this context, NWChemEx will be enabling sparse, reduced-scaling methods as well as the conventional methods. For example, resolutions of identity, Cholesky

decomposition, and domain-local pair natural orbital methods are being used at all levels of theory to reduce the scaling of the computations. The goal is to enable the computation of on the order of thousands of atoms using accurate, high-level quantum mechanical methods and one hundred thousand to one million atoms using less accurate methods such as density functional methods. This capability, combined with embedding techniques to combine the two regimes, will allow for the realistic and accurate simulation of chemical systems and their energies and properties. While the initial science capability is aimed at catalytic processes in zeolites, the capabilities within NWChemEx will provide ubiquitous capabilities in many science domains. At this point, NWChemEx does not have all of the infrastructure for REE and actinide chemistry, but the design is in place to enable this capability future implementations. In particular, NWChemEx currently does not handle effective core potentials, relativistic corrections such as Douglas-Kroll, and implicit solvent methods required for more realistic simulations of the experimental conditions.

3.2. Multiresolution Adaptive Numerical Environment for Scientific Simulation (MADNESS)

MADNESS (*98*) started as an environment for fast and accurate numerical simulation in chemistry (*99*) but rapidly expanded to include applications in nuclear physics (*100*), boundary value problems (*101*), solid state physics, and atomic physics in intense laser fields (*102*). Other projects such as MPQC and NWChemEx employ the MADNESS parallel runtime.

Applications to electronic structure (*99*) have primarily focused on single-particle methods (HF and DFT) for molecules (*103*, *104*). The MADNESS molecular DFT code includes energies, gradients, solvation, and linear response for LDA, HF, GGA, and hybrid functionals, using localized orbitals to reduce scaling of the computational cost with system size (including linear-scaling Hartree-Fock exchange) (*105*). Computations beyond 3-D include time evolution in an intense laser pulse of H_2^+ in 4-D (3 electronic and 1 vibrational) (*102*), and in 6-D with the first ever numerical computation of the MP2 energy of a nonlinear molecule (*106*, *107*). The most ambitious such project is due to Bischoff (Humbolt University, Berlin) who has made steady progress towards the goal of implementing full coupled-cluster energy and response properties in 6D on top of MADNESS (*108–110*). Independent implementations have also been demonstrated (*111*, *112*).

MADNESS employs adaptive multi-resolution algorithms for fast computation with guaranteed precision and separated representations for efficient computation in many dimensions (*99*). Multi-resolution analysis (MRA) allows a compact representation of sufficiently smooth functions with arbitrary guaranteed precision, sparse representation of many differential and integral operators, and fast algorithms for their applications, thus, realizing a basis-free fast computational integro-differential calculus (*113–115*).

There exist only two widely-used numerical representations for solving the many-body Dirac equation for all electrons in general molecules or periodic solids:

- Linear combination of atomic orbitals (LCAO), in which solutions are expanded in terms of pre-optimized fixed sets of atomic-centered orbitals. Gaussian AOs (*116*) are used almost exclusively, but Slater (*117*, *118*) and numerical orbitals (*119–121*) can also be used. (Note that only the basis used to expand the orbitals matters here; e.g., use of real-space representations or plane waves for intermediate quantities to accelerate LCAO computations, such as in the pseudospectral method, still fall under the LCAO umbrella.)
- Muffin-tin-like approaches which utilize separate representations in the interatomic (interstitial) region, typically described by plane waves, and the spherical intraatomic

regions, which are described by a variety of analytic and numerical forms. These include full-potential linear muffin-tin orbital method (LMTO) (*122*), linearized augmented plane waves (LAPW) (*123*) APW+lo (*124*, *125*) and other full-potential cousins of the original APW approach (*118*). Other related methods (*126*) are in principle also usable for all-electron relativistic computations.

Unlike the conventional LCAO representation, where the basis is chosen *a priori*, the MRA representation has no predefined basis. The basis to represent a given function (e.g., a molecular orbital) is extended adaptively (where needed) until the user-defined precision threshold ε is reached. The abilities to compute solutions with guaranteed precision and speed (*O(N)* in the basis size) are the crucial advantages of the MRA representation over LCAO. To guarantee precision, every function ($O(10^4)$ in a large calculation) has an independent and dynamically refined mesh, and composing functions or applying operators can change the mesh refinement. These meshes are represented as k-d-trees – i.e., trees in either 3 or 6 spatial dimensions for one or two electron problems, respectively. Typically, these trees are poorly balanced and change very dynamically and, hence, pose significant challenges for computation on a parallel machine. For instance, in simulating high-harmonic generation in an intense infra-red laser pulse the electron is literally ripped hundreds of atomic units of distance from the atom only to be brought back to rescatter off the nucleus.

Recently Harrison and co-workers demonstrated (*127*) the first converged numerical solution of the Dirac-Fock equations for general non-linear molecules using the MRA representation. Our novel approach extended the integral (Lippmann-Schwinger) equation approach employed in the non-relativistic setting by using the relationship between the Dirac and Schrodinger free-electron Green's functions (*128*). The Dirac Green's function defines a 4x4 matrix of integral operators that can be accurately and efficiently applied in the MRA representation. Crucially, in subsequent work, Beylkin proved that the iteration employing this Green's function to solve the Diract equation is globally and robustly convergent to the sought electronic states, without any need to impose *ad hoc* kinetic balance conditions upon the basis set to avoid collapse to negative-energy (positronic) states. This new relativistic capability has the same robust error control and scaling of computational cost as demonstrated for the Schrodinger equation. The MRA representation is capable of depicting the finest scale details of the molecular spinors; for some spinors as many as 18 levels of refinement were needed in order to describe accurately the nuclear penetration of the spinors. Although the current code is only a pilot implementation of the method, since it is missing many optimizations as well as necessary features such as treating open shells, its superior precision relative to Gaussian AOs is motivating its further development that will enable calculations free of basis set error on systems containing lanthanides and actinides.

From the outset, the twin goals of both performance and productivity were paramount. Performance means enabling efficient computation on systems spanning from laptop to exascale supercomputer. To this end, MADNESS introduced its own parallel runtime that is fully asynchronous and employs task-based parallelism to utilize multi-core processors.

The MADNESS runtime has evolved into a powerful environment for the composition of a wide range of parallel algorithms on many distributed data structures including trees in MADNESS and the sparse tensors in TiledArray. The central elements of the parallel runtime are a) futures for hiding latency and managing dependencies, b) global namespaces with one-sided access, c) remote method invocation in objects in global namespaces, and d) dynamic load balancing and data redistribution. An application in the MADNESS runtime can be viewed as a dynamically constructed DAG, with futures as edges. Moreover, the structure of the discontinuous, high-order, spectral-element basis

results in key computational steps being phrased as small matrix-matrix operations that are very well suited to modern processors that emphasize cache locality and use of SIMD functional units.

Productivity means addressing the semantic gap that inflates a few pages of equations into millions of lines of software (the size of a typical quantum chemistry package). At the lowest level, this is addressed in part by the MADNESS parallel runtime strongly separating the responsibilities of the programmer expressing parallelism and the runtime managing data and scheduling computation. At the highest level a prototype domain specific language, which maps nearly one-to-one onto the C++ API, means that application codes are written in terms of the functions and integro-differential operators of the physics equations being implemented, rather than in terms of matrix elements and coefficients.

The limitation to these choices is due to the need to describe small (high-energy) states of core electrons explicitly, i.e., without a pseudopotential; plane waves alone or traditional real-space numerical representations (grids, finite elements) cannot economically represent such tightly-localized states. Furthermore, if accurate ab initio (post-DFT) treatment of electron correlation is desired, LCAO becomes the primary choice. This is largely due to the fact that compact sets of nonvalence unoccupied states, the most convenient building block for dynamical correlation, can be constructed a priori from inexpensive atomic computations and fixed once and for all; MT-like representations are more verbose, i.e., use significantly more basis functions than the AO based counterparts. Computation of operators in the LCAO representation is also simpler – due to the single functional form of the basis – than in the muffin-tin-like approaches which combine multiple types of bases.

3.3. Massively Parallel Quantum Chemistry (MPQC) Program

MPQC (*129*) is a large-scale research platform for electronic structure molecules and materials. Its development began at Sandia National Laboratories in the early 1990s, motivated by the need to deploy electronic structure methods on shared and distributed memory machines that became available to the community at the time. To reach these objectives the MPQC team recognized the need to employ abstractions that went beyond simple message passing; thus both hybrid (intranode tasking + message passing) and partitioned global address space (PGAS) abstractions were prominently used throughout the code. MPQC was the first large-scale molecular electronic structure package to be implemented in C++, utilizing its support for object oriented and generic (templates) programming styles. The modular design of MPQC facilitated code reuse within the package and composition with external packages.

In the 2010s, the MPQC package was redesigned almost from scratch to address several pressing challenges, which included:

- the rapidly growing complexity of HPC platforms that nowadays feature 100s of execution units within the node, the extreme size of high-end distributed memory HPC platforms (spanning 10s and 100s of thousands of nodes), and the growing importance of heterogeneous hardware and the associated asynchronous programming models;
- critical importance of reduced-complexity algorithms for modern accurate electronic structure methods; and
- increasing costs of maintaining legacy C++ codebase and the need to exploit the full power of modern C++.

Today's MPQC is based on the multiparadigm MADNESS parallel runtime (*98*) that supports traditional (message-passing, threads) and advanced (active messages, distributed task graphs, distributed object model) techniques for writing parallel algorithms, allowing to hide latencies across the entire memory hierarchy (intranode and internode) and interoperate naturally with the asynchronous heterogeneous programming models (e.g., CUDA). The TiledArray block-sparse tensor framework (*130*) allows facile high-level composition of conventional and reduced-scaling mean-field and many-body algorithms in MPQC. The latest MPQC fully leverages the power of modern C++ (the 2017 ISO standard) and can exploit heterogeneous (CUDA-based) platforms for a key subset of its functionality (*131*).

TiledArray (*132*), a modern C++ framework (*130*) for parallel data-sparse tensor algebra translates sparse tensor algebra into a dataflow-style depiction of operators. Supporting block-sparse and other non-standard tensor structures is a significant structural challenge that needs to be exposed to the algorithm in a comprehensive way. Dense matrix and tensor algebra have well established strong-scaling algorithms with theoretically proven bounds for data movements, but extension of these algorithms to unstructured sparse tensor computation is far from straightforward. To address these challenges and deal with the ramifications of sparsity, TiledArray extends upon the capabilities of the task-based MADNESS while putting forward several features that distinguish TiledArray from most other tensor algebra frameworks, such as the Tensor Contraction Engine (*133*): 1) TiledArray is domain-neutral, composable via a high-level DSL and is broadly applicable in chemistry, condensed matter physics, quantum information theory, and applied mathematics; 2) it is a standalone generic framework, not part of a monolithic package, and is designed to be easily reusable and customizable to a science domain; and 3) it is designed to support dataflow-style concurrent execution of dense and data-sparse tensor algebra.

While not specifically a Domain Specific Language, TiledArray supports multiple levels of composition, from the low-level interface dealing with individual tiles and tasks to the top-level end-user interface for composing complex tensor expressions, e.g., the right-hand side of a simple linear tensor equation of one of the simplest many-body electronic structure method. Recently TiledArray has been extended to support multiple underlying task-based runtimes (such as PaRSEC and MADNESS), and to support rising hybrid architectures, multi-GPU multi-node platforms while delegating the onus of resource management and scheduling to the underlying runtime system.

To address the pressing needs of the REE and actinide chemistry community the MPQC team is developing support for all-electron relativistic Hamiltonians. Accurate treatment of realistic large models of experimental systems will be possible by the use of several existing types of reduced-scaling coupled-cluster formalisms. The basis set error of many-body methods can be robustly reduced by the use of explicitly correlated formalisms, also already available in the package. Lastly, the needs for robust treatment of strong correlation will be addressed by the ongoing development of sparse CI and other electronic structure models.

4. Artificial Intelligence

Advances in supercomputing have enabled HPC capabilities in Artificial Intelligence (AI) to accelerate findings across disciplines. Unlike the challenges with efficient utilization of HPC resources from traditional software packages to model chemical systems, advanced data science stemming from AI techniques can readily take advantage of increased HPC resources. However, applications of AI to modeling REEs and actinides in applications relevant to optimization of separations, binding selectivity, and recycling – key areas to solve REEs and actinides challenges – are in its infancy.

AI encompasses various advanced computational techniques, including machine learning, deep learning, symbolic AI, Bayesian networks, and support vector machine, amongst others. AI techniques rely on previously collected data sets to train algorithms to generate predictive capabilities to provide insight into properties of interest. Predicting properties of REE- and actinide-containing systems relevant to solutions to guarantee a healthy supply of these elements (from extraction to purification) and those needed for waste management and environmental concerns requires thermochemical data. These data are challenging when needing to obtain strong, reliable, and complete data sets to train algorithms for AI applications. Populating these data sets require a large amount of experimentally obtained results or theoretical values from predictive modeling. As discussed in previous sections in this chapter, accurate modeling of REE- and actinide-containing compounds faces great challenges due to the limitations in the software's capability to make efficient use of current HPC resources. Therefore, obtaining a reliable data set to model properties needed for binding and separation optimization for REEs and actinides is highly limited.

For example, AI techniques including machine learning and deep learning have been applied to predict properties of materials, including chemical and spectroscopic characteristics for various systems (*134–142*), but applications to thermochemistry of REEs and actinides remain scarce. A recent review of learning-based techniques in nuclear science and engineering discusses the utilization of machine learning in nuclear technologies, highlighting detection through gamma spectroscopy and reactor health monitoring, radiation protection, environmental monitoring, and materials, amongst others (*143*). Other AI applications (including machine learning and deep learning) in radiochemistry have included detection of gamma-ray sources, and radiotherapy to enhance cancer treatment (*144–150*). Machine learning has also been utilized in nuclear forensics to determine dispersion of airborne particles, and monitoring and detection for identification of radioactive particles (*151*, *152*). Modeling of nuclear reactors anomalies detection and sensor training in nuclear plants has included deep learning approaches (*153*), amongst others.

Developing software such as NWChemEX, MADNESS, and MPQC to make efficient use of exascale resources will accelerate opportunities for creating data sets for AI applications. This will provide valuable advances to predict chemical reactions critical in optimization of separations and selective binding to REEs and actinides.

5. Protocol to Evaluate Complexation Preferences in Separations of REEs and Actinides through Computational Modeling

There are many areas that can benefit from modeling REEs and actinides, including applications in synthesis of REE- and actinide-containing compounds, characterization and detection driven by spectroscopic analysis, and other uses. Although beyond the scope of the discussion in this section, excellent examples of traditional computational modeling in these areas are provided in works by Vallet, Bagus, Knope, Peterson, Batista, Yang, Gagliardi, Peterson, Dixon, Kaltsoyannis, and others (*154–180*). In this section, highlights are provided for a systematic protocol to evaluate complexation preferences of specific interest in separations through applications of computational modeling.

Optimization of separations of REEs and actinides is critical in many areas, including efforts to guarantee a supply of REEs, purification processes needed for radiotherapy, to manage radiological waste, and for environmental and energy concerns. Selective separation of REEs is challenging due to the similar chemistries presented by these elements, and the need for expert facilities to handle radiation and safety concerns. The need for novel solutions to this challenge is clear, and it can be addressed by accelerating the understanding and development of selective extracting processes for

efficient separation from ores and recycling. Trial-and-error experimental approaches to advance solutions for REEs and actinides present challenges due to limitations in the availability of materials and the need for expert facilities to handle radioactive elements. Therefore, computational modeling is essential to advance solutions for REEs and actinides. HPC-enabling capabilities are critical to provide application-ready tools to the community. However, as previously discussed, the current HPC ecosystem is not optimum to facilitate the computational modeling needed. Given available software/hardware integration, some useful qualitative descriptions can be provided to aid selective separations of REEs and actinides.

Current efforts in modelling REE and actinide separations include a large variation in choices of levels of theory (methods, basis sets, etc.) often providing qualitative non-systematic descriptions (i.e., appropriate for one REE or actinide but not another). There have been great strides in applying currently available HPC resources (*181–195*). These efforts highlight the limitations in accurate descriptions of separations thermochemistry in this area.

Previous studies of $[Ln(NO_3)]^{+2}$ and $[An(NO_3)]^{+2}$ in gas and aqueous phases for the entire lanthanide and actinide series illustrated the need to apply predictive many-body methods to such systems, underscored by often drastic differences between binding energies (ordinary and differential) calculated with CCSD(T)/CBS and DFT (with even more variations within functional and basis sets of choice) (*196*, *197*). For example, the differential in some contiguous actinides such as in Pu/Am binding energies were predicted to be +24.47 kcal/mol with CCSD(T)/CBS and -20.00 kcal/mol with B972/6-311++G** (*196*, *197*). Of course, one of the challenges is that the conventional CCSD(T) method scales as $O(N^7)$ while DFT scales as $O(N^3)$, where N is the number of basis functions in the system. This makes it difficult to scale the canonical high accuracy methods to large systems without improved methodologies and HPC. Recently, binding data was obtained for actinium and lanthanum with formic and acetic acids using pseudopotential and all-electron CASPT2, CCSD(T), MP2, and DFT with various functionals and basis sets (*198*, *199*) (shown in Figure 8 and described in the Appendix). Yet again, strong variations of the differential binding energies are observed within DFT results obtained with different functionals and basis sets, between DFT and reliable many-body predictions, and between pseudopotential and all-electron computations.

Advancing efficiency in separations of REEs and actinides by addressing binding selectivity through computations can extend beyond accurate predictive thermochemistry. Some of these properties include electron withdrawing and donation effects, structural characteristics, electronic contribution to binding orbitals, sterics, speciation, and competition in solution. There are many variables present when modeling binding environments in ion-exchange and solvent extraction mechanisms for separations of REEs and actinides, including solvation effects, ligands' characteristics, backbone variations, alkyl substitutions, solvents, ions in solution, speciation, temperature, and pH.

Thermochemical calculations are often utilized when seeking quantification. This involves calculations of Gibbs free energies of reaction (including enthalpic and entropic effects) which relate to the reaction constants and distribution coefficient driving the reactions in separations. However, as previously discussed, given the limitations in utilization of high level theory methods with the current software in the unbalanced hardware-software ecosystem available, highly accurate quantitative predictions are often not feasible for real-size fully complexed REE- and actinide-containing systems of interest in separations. Nonetheless, there is much qualitative information that can aid in optimization of separations by providing insight into complexation preferences, which, if adjusted, can accelerate targeted binding in selective separations. A protocol to provide qualitative information

including five evaluations steps is described below. These steps take in consideration stages in the complexation and binding preference between the time in which the REE or actinide is fully hydrated (i.e., complexed with only water molecules) to a state in which it is fully complexed with only ligands and no water molecules (solvent extraction), or ligands and waters (ion-exchange). These steps consist of an evaluation of: (1) coordination number, (2) coordination geometry, (3) complexation mechanism, (4) preferential binding, and (5) other binding characteristics.

- **Coordination number**: The coordination number of REEs and actinides show variations (*10*, *53*). For example, lanthanides present a coordination number preference of 8 and 9, for lighter and heavier lanthanides, respectively, with actinides allowing for a wider range. Depending on the environment (ions in solution, ligands' chelation characteristics, and other factors) the coordination number may change. To illustrate this first step, Figure 7 shows a plot which aids in the evaluation of complexation by providing the researcher with a simple visual which shows the coordinating atoms. In the example illustrated in Figure 7, einsteinium is complexed with 4-hydroxy-5-iodo-2,3-dimethoxy-6-methylbenzoic acid. The result shows a clear limit of nine coordinating oxygen atoms from six ligands in an $EsR_3(HR)_3$ configuration. The coordination shell includes coordinating oxygen atoms between 2.38 and 2.70Å, with an outer coordination shell starting at 3.89Å.
- **Coordination geometry**: The coordination analysis provides structural information about the compound. In the example shown in Figure 7, it can be seen that the coordination shell is composed by nine oxygen atoms from which six are from the carboxylate groups (deprotonated ligands) binding with the metal in a bidentate form, and three oxygen atoms from carboxyl groups in a monodentate binding. All other oxygen atoms are clearly outside the coordination shell (with Es-O interatomic distances larger than 3.89Å. This type of evaluation provides fundamental information about the coordination preferences, solvation shells, and often chelation competition from ions in solution, which can provide critical information to tailor separation environments to optimize selective binding.
- **Complexation mechanism**: In this step, the complexation mechanism is evaluated, which includes step-wise reactions and systematic additions. Understanding how the complexation occurs is a critical step to design optimized processes for selective binding. Experimentally, one can measure/characterize when the REE is extracted into the organic phase (or separated) after initially being in in the aqueous phase. However, the process in which the ligands complex with the REE cannot be easily evaluated in a step-wise fashion experimentally. This requires proposed step-wise addition of ligands while removing water molecules from the coordination shell (for a separation process in which the initial stage is an aqueous phase). This computational visualization is similar to that shown in Figure 7, but also includes competing ions and water molecules (*200*).
- **Preferential binding**: Preferential binding in complexation of REEs and actinides can be qualitatively addressed to determine preferences not just inferred from thermodynamic calculations, but focusing on structural characteristics that can drive binding preferences (for example, as it was observed in actinyls bound with ligands producing delocalization in chelation) (*201*). Creating relative binding preference plots calculated for various separation environments to evaluate binding preferences can also contribute to selective separations. When the size of the molecular system is suitable for high level accurate

calculations, thermochemical predictions can predict the binding preference; however, as previously discussed, for fully complexed systems, this is often not feasible with current computing resources. In this case it can be helpful to create maps that survey thermochemical properties, such as Gibbs free energies of reaction, with lower-level methods (feasible to be calculated with currently available computing resources) against target high-level methods. Maps of this nature have been built for lanthanides and actinides across the series bound with nitrates, including targets with CCSD(T)/CBS (*196*, *197*). Evaluations including Gibbs free energies of reactions calculated with multireference methods have been pursued in efforts including separation of actinium and lanthanum (as shown in Figure 8) (*198*, *199*). Including relativistic effects from spin-orbit properties can also be challenging to calculate for fully complexed systems. Spin-orbit contributions calculated using the X2C-Dirac-Hartree-Fock showed that the contribution to the Gibbs free energy of reaction in a study of lanthanides bound with nitrates was less than 0.5 kcal/mol (except for Pr, which results shows 3.8 kcal/mol) (*196*). This type of analysis can aid in determining efficient methods to predict qualitative preferential binding.

- **Binding characteristics**: In this step, other properties that can drive binding preferences, such as electron withdrawing/donation effects, and delocalization in chelation are evaluated. Previous work in which binding strength is correlated directly to structural characteristic and electron donation/withdrawing effects in actinides bound with cyclic imide dioximes (*201*) and nitrates (*197*), revealed strong qualitative predictive insights.

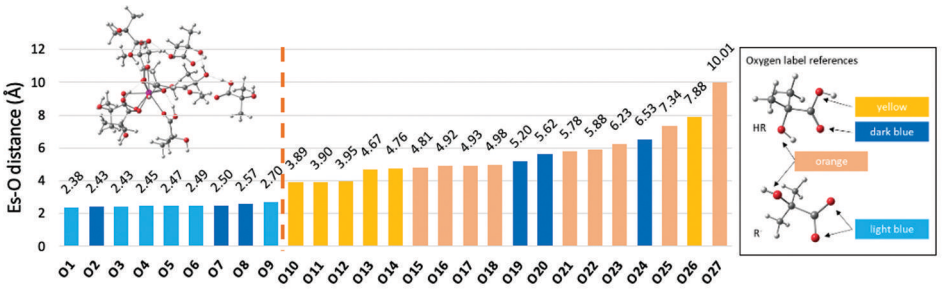

Figure 7. Einsteinium coordination with 4-hydroxy-5-iodo-2,3-dimethoxy-6-methylbenzoic acid in $EsR_3(HR)_3$.

Conclusion

Accelerating solutions to guarantee a healthy supply of REEs and efficient handling of actinides to fulfill national and global needs requires accurate computational modeling. Computational techniques ranging from electronic structure modeling to applications of AI require HPC. Traditional modeling involving electronic structure theory through software packages, such as NWChemEX, MADNESS, and MPQC needs advances in hardware-software integration to enable software to make efficient use of HPC resources. The great increase observed in hardware architectures, as shown in the global TOP500 list compiled since 1993, indicates immense promise on continuing advances in supercomputers. Recent investments from the U.S. DOE into the Exascale Computing Project provide encouraging efforts in creating a more balanced hardware-software ecosystem in which not only the hardware sees advances, but also the software are designed to make efficient use of the most advanced HPC architectures. As the community moves onto the HPC exascale era, redesigned and newly developed software will allow for accurate predictive capabilities to optimize separations

and binding preferences of critical materials. This will be possible from accurate predictive thermochemistry utilized in direct chemical analysis, and for creating robust data sets for AI applications. This will enable finding solutions for REEs and actinides in their myriad of applications in energy, defense, medical treatments, and environmental needs.

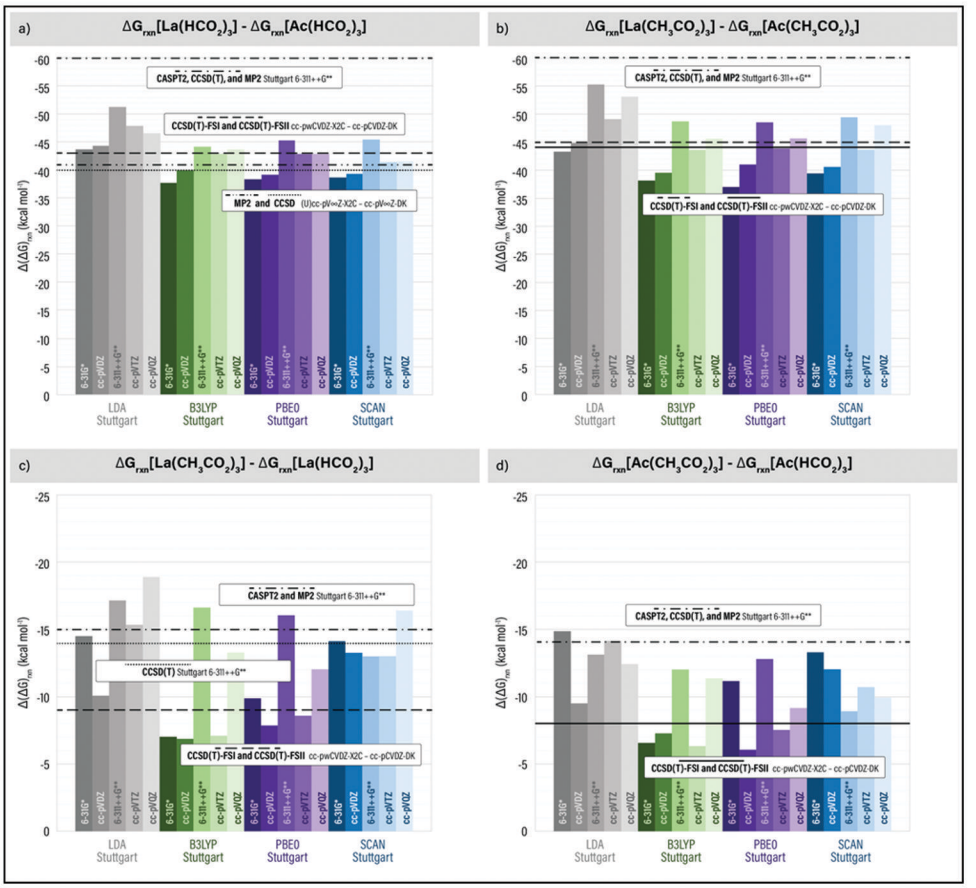

Figure 8. Relative Gibbs free energies of reaction calculated for $La(HCO_2)_3$, $Ac(HCO_2)_3$, $La(CH_3CO_2)_3$, and $Ac(CH_3CO_2)_3$, as described in the Appendix.

Associated Content

Table A1 and Table A2 include a list of the number one supercomputer in the TOP500 list from 1993 through 2020. The Appendix text includes computational details relevant to section 5.

Acknowledgments

The authors gratefully acknowledge Dr. Charles C. Peterson for sharing results from his work on *ab initio* calculations on lanthanum- and actinium-containing compounds. The authors also acknowledge Dr. Robert J. Hinde, Dr. George K. Schweitzer, Dr. Rose A. Boll, Dr. Laetitia Delmau, Dr. Paul D. Benny, and Dr. Howard L. Hall for useful discussions, and David Rogers for graphics design. The authors thank Eleigha M. Wrancher and Shelby Kemp for their contributions during their internships. Part of this research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National

Nuclear Security Administration. The authors acknowledge support from the NWChemEx Exascale Computing Project from the U.S. Department of Energy (DOE) at Ames Laboratory under contract No. DE-AC02-07CH11358 and Brookhaven National Laboratory under Contract No. DE-SC0012704. The EPEXA project is supported by the National Science Foundation under grants OAC-1931387 at Stony Brook University, OAC1931347 at Virginia Tech, and ACI-1450300 at the University of Tennessee. Additionally, some of this material is based upon work supported in part by the National Science Foundation (NSF) under grant 1450429 "SI2-SSI: Collaborative Proposal: Performance Application Programming Interface for Extreme-scale Environments (PAPI-Ex)", and grant 1642440 "SI2-SSE: PAPI Unifying Layer for Software-Defined Events (PULSE)". This material is partially based upon work supported by the U.S. Department of Energy, under award number DE-AC05-00OR22725 "Sparse Iterative Solver Benchmarking and the SparseBench Project." This work was also supported in part by NVIDIA Corporation.

Appendix

Table A1. Number 1 Supercomputer in TOP500 List from 1993 to 2020—Specs, Site, and Vendor[a]

Year	*List*	*Computer*	*Country*	*Site*	*Specs*	*Vendor*
2020	Nov	Supercomputer Fugaku	Japan	RIKEN Center for Computational Science	A64FX, 48C, 2.2 GHz, Tofu Interconnect D	Fujitsu
	June	Supercomputer Fugaku	Japan	RIKEN Center for Computational Science	A64FX, 48C, 2.2 GHz, Tofu Interconnect D	Fujitsu
2019	Nov	Summit	United States	Oak Ridge National Laboratory	AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	IBM
	June	Summit	United States	Oak Ridge National Laboratory	AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	IBM

Table A1. (Continued). Number 1 Supercomputer in TOP500 List from 1993 to 2020—Specs, Site, and Vendor[a]

Year	*List*	*Computer*	*Country*	*Site*	*Specs*	*Vendor*
2018	Nov	Summit	United States	Oak Ridge National Laboratory	AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	IBM
	June	Summit	United States	Oak Ridge National Laboratory	AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband	IBM
2017	Nov	Sunway TaihuLight	China	National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz	NRCPP
	June	Sunway TaihuLight	China	National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz	NRCPP
2016	Nov	Sunway TaihuLight	China	National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz	NRCPP
	June	Sunway TaihuLight	China	National Supercomputing Center in Wuxi	Sunway MPP, Sunway SW26010 260C 1.45GHz	NRCPP
2015	Nov	Tianhe-2A	China	National Supercomputing Center in Guangzhou	TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P	NUDT
	June	Tianhe-2A	China	National Supercomputing Center in Guangzhou	TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P	NUDT

Table A1. (Continued). Number 1 Supercomputer in TOP500 List from 1993 to 2020—Specs, Site, and Vendor[a]

Year	*List*	*Computer*	*Country*	*Site*	*Specs*	*Vendor*
2014	Nov	Tianhe-2A	China	National Supercomputing Center in Guangzhou	TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P	NUDT
	June	Tianhe-2A	China	National Supercomputing Center in Guangzhou	TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P	NUDT
2013	Nov	Tianhe-2A	China	National Supercomputing Center in Guangzhou	TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P	NUDT
	June	Tianhe-2A	China	National Supercomputing Center in Guangzhou	TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P	NUDT
2012	Nov	Titan	United States	Oak Ridge National Laboratory	Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x	Cray HPE
	June	Sequoia	United States	Lawrence Livermore National Laboratory	BlueGene/Q, Power BQC 16C 1.60 GHz, Custom	IBM
2011	Nov	K computer	Japan	RIKEN Advanced Institute for Computational Science (AICS)	SPARCC64 VIIIfx 2.0GHz, Tofu interconnect	Fujitsu
	June	K computer	Japan	RIKEN Advanced Institute for Computational Science (AICS)	SPARCC64 VIIIfx 2.0GHz, Tofu interconnect	Fujitsu

Table A1. (Continued). Number 1 Supercomputer in TOP500 List from 1993 to 2020—Specs, Site, and Vendor[a]

Year	*List*	*Computer*	*Country*	*Site*	*Specs*	*Vendor*
2010	Nov	Tianhe-1A	China	National Supercomputing Center in Tianjin	NUDT TH MPP, X5670 2.93GHz 6C, NVIDIA GPU, FT-1000 8C	NUDT
	June	Jaguar	United States	Oak Ridge National Laboratory	Cray XT5-HE Opteron 6-core 2.6 GHz	Cray HPE
2009	Nov	Jaguar	United States	Oak Ridge National Laboratory	Cray XT5-HE Opteron 6-core 2.6 GHz	Cray HPE
	June	Roadrunner	United States	Los Alamos National Laboratory	Bladecenter QS22/ LS21 Cluster, PowerXCell 8i 3.2 GHz / Opteron DC 1.8 GHz, Voltaier Infiniband	IBM
2008	Nov	Roadrunner	United States	Los Alamos National Laboratory	Bladecenter QS22/ LS21 Cluster, PowerXCell 8i 3.2 GHz / Opteron DC 1.8 GHz, Voltaier Infiniband	IBM
	June	Roadrunner	United States	Los Alamos National Laboratory	Bladecenter QS22/ LS21 Cluster, PowerXCell 8i 3.2 GHz / Opteron DC 1.8 GHz, Voltaier Infiniband	IBM
2007	Nov	BlueGene/L	United States	Lawrence Livermore National Laboratory	eServer Blue Gene Solution	IBM
	June	BlueGene/L	United States	Lawrence Livermore National Laboratory	eServer Blue Gene Solution	IBM
2006	Nov	BlueGene/L	United States	Lawrence Livermore National Laboratory	eServer Blue Gene Solution	IBM
	June	BlueGene/L	United States	Lawrence Livermore National Laboratory	eServer Blue Gene Solution	IBM

Table A1. (Continued). Number 1 Supercomputer in TOP500 List from 1993 to 2020—Specs, Site, and Vendor[a]

Year	*List*	*Computer*	*Country*	*Site*	*Specs*	*Vendor*
2005	Nov	BlueGene/L	United States	Lawrence Livermore National Laboratory	eServer Blue Gene Solution	IBM
	June	BlueGene/L	United States	Lawrence Livermore National Laboratory	eServer Blue Gene Solution	IBM
2004	Nov	BlueGene/L	United States	Department of Energy	BlueGene/L DD2 beta-System (0.7 GHz PowerPC 440)	IBM
	June	Earth-Simulator	Japan	Japan Agency for Marine-Earth Science and Technology		NEC
2003	Nov	Earth-Simulator	Japan	Japan Agency for Marine-Earth Science and Technology		NEC
	June	Earth-Simulator	Japan	Japan Agency for Marine-Earth Science and Technology		NEC
2002	Nov	Earth-Simulator	Japan	Japan Agency for Marine-Earth Science and Technology		NEC
	June	Earth-Simulator	Japan	Japan Agency for Marine-Earth Science and Technology		NEC
2001	Nov	ASCI White	United States	Lawrence Livermore National Laboratory	SP Power3 375 MHz	IBM
	June	ASCI White	United States	Lawrence Livermore National Laboratory	SP Power3 375 MHz	IBM
2000	Nov	ASCI White	United States	Lawrence Livermore National Laboratory	SP Power3 375 MHz	IBM

Table A1. (Continued). Number 1 Supercomputer in TOP500 List from 1993 to 2020—Specs, Site, and Vendor[a]

Year	*List*	*Computer*	*Country*	*Site*	*Specs*	*Vendor*
2000	June	ASCI Red	United States	Sandia National Laboratories		Intel
1999	Nov	ASCI Red	United States	Sandia National Laboratories		Intel
	June	ASCI Red	United States	Sandia National Laboratories		Intel
1998	Nov	ASCI Red	United States	Sandia National Laboratories		Intel
	June	ASCI Red	United States	Sandia National Laboratories		Intel
1997	Nov	ASCI Red	United States	Sandia National Laboratories		Intel
	June	ASCI Red	United States	Sandia National Laboratories		Intel
1996	Nov	CP-PACS/2048	Japan	Center for Computational Sciences, University of Tsukuba		Hitachi / Tsukuba
	June	SR2201/1024	Japan	University of Tokyo		Hitachi
1995	Dec	Numerical Wind Tunnel	Japan	National Aerospace Laboratory of Japan		Fujitsu
	June	Numerical Wind Tunnel	Japan	National Aerospace Laboratory of Japan		Fujitsu
1994	Nov	Numerical Wind Tunnel	Japan	National Aerospace Laboratory of Japan		Fujitsu
	June	XP/S140	United States	Sandia National Laboratories		Intel
1993	Nov	XP/S140	United States	Sandia National Laboratories		Intel
	June	CM-5/1024	United States	Los Alamos National Laboratory		Thinking Machines Corporation

[a] This table was built from information available online, and it provides a unique view of the data unavailable in other sources (*54*–*57*).

Table A2. Number 1 Supercomputer in TOP500 List from 1993 to 2020—Cores, Rmax, Rpeak, and Power[a]

Year	*List*	*Computer*	*Cores*	*Rmax (Teraflops/s)*	*Rpeak (Teraflops/s)*	*Power (kW)*
2020	November	Supercomputer Fugaku	7,630,848.0	442,010.0	537,212.0	29,899
	June	Supercomputer Fugaku	7,299,072.0	415,530.0	513,854.7	28,335
2019	November	Summit	2,414,592.0	148,600.0	200,794.9	10,096
	June	Summit	2,414,592.0	148,600.0	200,794.9	10,096
2018	November	Summit	2,397,824.0	143,500.0	200,794.9	9,783
	June	Summit	2,282,544.0	122,300.0	187,659.3	8,806
2017	November	Sunway TaihuLight	10,649,600.0	93,014.6	125,435.9	15,371
	June	Sunway TaihuLight	10,649,600.0	93,014.6	125,435.9	15,371
2016	November	Sunway TaihuLight	10,649,600.0	93,014.6	125,435.9	15,371
	June	Sunway TaihuLight	10,649,600.0	93,014.6	125,435.9	15,371
2015	November	Tianhe-2A	3,120,000.0	33,862.7	54,902.4	17,808
	June	Tianhe-2A	3,120,000.0	33,862.7	54,902.4	17,808
2014	November	Tianhe-2A	3,120,000.0	33,862.7	54,902.4	17,808
	June	Tianhe-2A	3,120,000.0	33,862.7	54,902.4	17,808
2013	November	Tianhe-2A	3,120,000.0	33,862.7	54,902.4	17,808
	June	Tianhe-2A	3,120,000.0	33,862.7	54,902.4	17,808
2012	November	Titan	560,640.0	17,590.0	27,112.5	8,209
	June	Sequoia	1,572,864.0	16,324.8	20,132.7	7,890
2011	November	K computer	705,024.0	10,510.0	11,280.4	12,660
	June	K computer	548,352.0	8,162.0	8,773.6	9,899
2010	November	Tianhe-1A	186,368.0	2,566.0	4,701.0	4,040
	June	Jaguar	224,162.0	1,759.0	2,331.0	6,950
2009	November	Jaguar	224,162.0	1,759.0	2,331.0	6,950
	June	Roadrunner	129,600.0	1,105.0	1,456.7	2,483
2008	November	Roadrunner	129,600.0	1,105.0	1,456.7	2,483
	June	Roadrunner	122,400.0	1,026.0	1,375.8	2,345
2007	November	BlueGene/L	212,992.0	478.2	596.4	2,329
	June	BlueGene/L	131,072.0	280.6	367.0	1,433

Table A2. (Continued). Number 1 Supercomputer in TOP500 List from 1993 to 2020—Cores, Rmax, Rpeak, and Power[a]

Year	*List*	*Computer*	*Cores*	*Rmax (Teraflops/s)*	*Rpeak (Teraflops/s)*	*Power (kW)*
2006	November	BlueGene/L	131,072.0	280.6	367.0	1,433
	June	BlueGene/L	131,072.0	280.6	367.0	1,433
2005	November	BlueGene/L	131,072.0	280.6	367.0	1,433
	June	BlueGene/L	65,536.0	136.8	183.5	716
2004	November	BlueGene/L	32,768.0	70,720.0	91,750.0	
	June	Earth-Simulator	5,120.0	35,860.0	40,960.0	3,200
2003	November	Earth-Simulator	5,120.0	35,860.0	40,960.0	3,200
	June	Earth-Simulator	5,120.0	35,860.0	40,960.0	3,200
2002	November	Earth-Simulator	5,120.0	35,860.0	40,960.0	3,200
	June	Earth-Simulator	5,120.0	35,860.0	40,960.0	3,200
2001	November	ASCI White	8,192.0	7,226.0	12,288.0	
	June	ASCI White	8,192.0	7,226.0	12,288.0	
2000	November	ASCI White	8,192.0	4,938.0	12,288.0	
	June	ASCI Red	9,632.0	2,379.0	3,207.0	
1999	November	ASCI Red	9,632.0	2,379.0	3,207.0	
	June	ASCI Red	9,472.0	2,121.0	3,154.0	
1998	November	ASCI Red	9,152.0	1,338.0	1,830.0	
	June	ASCI Red	9,152.0	1,338.0	1,830.0	
1997	November	ASCI Red	9,152.0	1,338.0	1,830.0	
	June	ASCI Red	7,264.0	1,068.0	1,453.0	
1996	November	CP-PACS/2048	2,048.0	368.2	614.4	
	June	SR2201/1024	1,024.0	220.4	307.2	
1995	December	Numerical Wind Tunnel	140.0	170.0	235.8	
	June	Numerical Wind Tunnel	140.0	170.0	235.8	
1994	November	Numerical Wind Tunnel	140.0	170.0	235.8	
	June	XP/S140	3,680.0	143.4	184.0	

Table A2. (Continued). Number 1 Supercomputer in TOP500 List from 1993 to 2020—Cores, Rmax, Rpeak, and Power[a]

Year	*List*	*Computer*	*Cores*	*Rmax (Teraflops/s)*	*Rpeak (Teraflops/s)*	*Power (kW)*
1993	November	XP/S140	3,680.0	143.4	184.0	
	June	CM-5/1024	1,024.9/13/ 21 3:13:00 AM0	59.7	131.0	

[a] This table was built from information available online, and it provides a unique view of the data unavailable in other sources (*54*–*57*).

Methods Utilized to Calculate the Gibbs Free Energies of Reaction of Lanthanum- and Actinium-Containing Compounds

The Gibbs free energies of reaction for $La(HCO_2)_3$, $Ac(HCO_2)_3$, $La(CH_3CO_2)_3$, and $Ac(CH_3CO_2)_3$ (shown in Figure 8) were calculated with optimizations in an aqueous phase utilizing the Local Density Approximation (LDA) (*202*), the hybrid-GGA B3LYP (*203*) and PBE0 (*204*) functionals, and the Strongly Constrained and Appropriately Normed Semilocal (SCAN) (*205*) functional. The Stuttgart RSC Segmented ECP and associated basis set for lanthanum (*206*), and the Stuttgart RSC 1997 ECP and associated basis set for actinium are utilized. The 6-31G* (*207*), cc-pVDZ (*208*), 6-311++G** (*207*), cc-pVTZ (*208*), cc-pVQZ (*208*) basis sets are utilized for carbon, oxygen, and hydrogen atoms. Gibbs free energies of reaction calculated as single point calculations are obtained with spin unrestricted coupled cluster with single and double excitations and with perturbative triples (CCSD(T)) and without (CCSD) (*209*, *210*), second-order Møller-Plesset perturbation (MP2) (*211*), and Complete Active Space Second-Order Perturbation Theory (CASPT2) (*212*). Basis sets utilized for La and Ac in single point calculations include the all-electron eXact 2-Component (X2C) based correlation consistent basis set with (cc-p(wC)VnZ-X2C where n=D,T,Q,∞ (with ∞=Complete Basis Set extrapolation (CBS)) along with uncontracted versions (denoted as (U)cc-pVnZ-X2C) (*213*, *214*). Single point calculations include the 6-31G* and 6-311++G** basis sets and the correlation consistent basis sets (cc-pVnZ and cc-pVnZ-DK where n=D,T,Q,∞) for the C, O, and H atoms.

For HF Complete Basis Set extrapolations, the formula by Karton Martin (*215*) shown in Eq. 1 is used;

$$E_n = E_{CBS} + A(n+1)e^{-6.57\sqrt{n}} \quad (1)$$

while the correlation energy is calculated with the extrapolation by Feller and co-workers, shown in Eq. 2 (*216*)

$$E_n = E_{CBS} + B(n+0.5)^{-4} \quad (2)$$

Energies calculated with CCSD, CCSD(T), and MP2 include two correlation spaces (denoted FSI and FSII). The correlation space for the FSI level includes the 6s6p5f electrons of the Ac atom, the 5s5p4f electrons of the La atom, the 2s2p electrons of the C and O atoms, and the 1s electrons of the H atom. The FSII level adds outer-core correlation by incorporating the 5s5p5d electrons of the Ac atom, 4s4p4d electrons of the La atom, and the 1s electrons of C and O to the correlation space.

Multireference calculations for the compounds are performed with CASPT2 at the FSI level of correlation. An active space of 18 electrons and 16 orbitals that includes the π-bonding and non-bonding lone pairs of the acetate/formate ligand and the f orbitals of the actinium and lanthanum atoms are considered.

Spin-orbit relativistic effects are calculated with Dirac-Hartree-Fock with basis sets developed by Dyall at a double-ζ level (*217*, *218*). The effect of spin-orbit coupling (SO) in the proposed systems is determined by the difference between the total energy using Dirac-Hartree-Fock and the total energy using Spin-Free-Dirac-Hartree-Fock (shown in Eq. 3) as described by Dyall and co-workers (*219*) and previously implemented by Peterson in basis set development of lanthanides (*220*) and actinides (*214*, *221*). This technique has also been previously applied to Gibbs free energies of reaction of lanthanide (*196*) and actinide (*197*) nitrates. Gibbs free energies were also calculated using Dirac-DFT methods with the eXact 2-Component (X2C) Hamiltonian with basis sets developed by Dyall et. al. (Dyall-nZ where n= D,T) denoted as DB3LYP and DPBE0 along with DB3LYPSF and DPBE0 for Spin-Free X2C Hamiltonian (*222*, *223*).

$$\Delta E_{SO} = E_{Dirac-Hartree-Fock} - E_{Spin-Free-Dirac-Hartree-Fock} \tag{3}$$

The proposed binding reaction for $La(HCO_2)_3$, $La(CH_3CO_2)_3$, $Ac(HCO_2)_3$ and $Ac(CH_3CO_2)_3$ is indicated in Eq. 4 and 5, describing binding in the gas phase (g) and in aqueous environment (aq), respectively. [Note: Although La is not an actinide, for simplicity, the An general notation is utilized throughout this appendix to indicate La and Ac.]

$$An^{3+}{}_{(g)} + 3\ R^{-}{}_{(g)} \rightarrow AnR_{3(g)} \tag{4}$$

$$An^{3+}{}_{(aq)} + 3\ R^{-}{}_{(aq)} \rightarrow AnR_{3(aq)} \tag{5}$$

with An = La and Ac, and $R^- = HCO_2^-$ and $CH_3CO_2^-$.

The Gibbs free energy of reaction for Eq. 5 and 6 is calculated as shown in Eq. 6 and 7, respectively.

$$\Delta G_{rxn\ AnR_3\,(g)} = \Delta G\left[AnR_{3(g)}\right] - \Delta G\left[An^{3+}{}_{(g)}\right] - 3\,\Delta G\left[R^{-}{}_{(g)}\right] \tag{6}$$

$$\Delta G_{rxn\ AnR_3\,(aq)} = \Delta G\left[AnR_{3(aq)}\right] - \Delta G\left[An^{3+}{}_{(aq)}\right] - 3\,\Delta G\left[R^{-}{}_{(aq)}\right] \tag{7}$$

Relative Gibbs free energies of reaction, $\Delta(\Delta G)_{rxn}$, between La- and Ac-containing compounds (maintaining the ligand constant) in the gas and aqueous phases are calculated as indicated in Eq. 8 and Eq. 9, respectively.

$$\Delta(\Delta G)_{rxn(g)} = \Delta G_{rxn\ AcR_{3(g)}} - \Delta G_{rxn\ LaR_3\,(g)} \tag{8}$$

$$\Delta(\Delta G)_{rxn(aq)} = \Delta G_{rxn\ AcR_{3(aq)}} - \Delta G_{rxn\ LaR_3\,(aq)} \tag{9}$$

with $R^- = HCO_2^-$ and $CH_3CO_2^-$.

The geometry optimizations were obtained without symmetry constraints to avoid pre-conceived symmetry assumptions, with tight tolerances, and extra fine grid. Some complex frequencies were obtained for the gas phase structures, which were removed almost entirely by applying the solvation model. DFT geometry optimizations, and harmonic vibrational frequency calculations are obtained with the NWChem 6.8 package (*224*). The Natural Bond Orbital 7.0

(NBO7) program (*225*) was used to obtain NBO (*226*) population analysis. CCSD, CCSD(T) and MP2 single point calculations utilized Molpro2019 (*227*) default convergence criteria. OpenMOLCAS version 18.0 (*228*) was used for the CASPT2 calculations. Spin-orbit relativistic effects were calculated with Dirac-Hartree-Fock and Dirac-DFT with the DIRAC18 (*229*) package. Partial charges, electron withdrawing effects, and orbital occupancies were calculated with Natural Bond Orbital (NBO) (*226*). Aqueous phase calculations were obtained with the COSMO (*230*) solvation model implemented in NWChem 6.8 to account for implicit solvation effects. Basis sets and effective core potentials were obtained from the Basis Set Exchange (*231–233*).

The results stemming from the calculations described in this appendix were presented at ACS symposia (*198*, *199*).

References

1. Atwood, D. A. Sustainability of Rare Earth Resources. In *The Rare Earth Elements: Fundamentals and Applications*; Atwood, D. A., Ed.; John Wiley & Sons, Incorporated, 2012; pp 21–27.
2. Humphries, M. *Rare Earth Elements: The Global Supply Chain*; 2013.
3. Bailey Grasso, V. *Report 115-255 - Rare Earth Elements in National Defense: Background, Oversight Issues, and Options for Congress*; 2013.
4. Mamula, N.; Bridges, A. *Ground Breaking*; TERRYDYNAMICS Corporation, 2018.
5. Jowitt, S. M. Introduction to a Resource Special Issue on Criticality of the Rare Earth Elements: Current and Future Sources and Recycling. *Criticality of the Rare Earth Elements: Current and Future Sources and Recycling*; 2018; pp 1–4.
6. Cotton, S. A.; Harrowfield, J. M. Lanthanides in Living Systems. In *The Rare Earth Elements: Fundamentals and Applications*; Atwood, D. A. , Ed.; John Wiley & Sons, Inc., 2012; pp 65–72.
7. Grenthe, I.; Drozdzynski, J.; Fujino, T.; Buck, E. C.; Albrecht-Schmitt, T. E.; Wolf, S. F. Uranium. *The Chemistry of the Actinides and Transactinide Elements*; Springer: Netherlands, 2011; pp 253–698. https://doi.org/10.1007/978-94-007-0211-0.
8. Runde, W. H.; Schulz, W. W. Americium. *The Chemistry of the Actinide and Transactinide Elements*; Springer: Netherlands, 2011; pp 1265–1395. https://doi.org/10.1007/978-94-007-0211-0.
9. Loveland, W. D.; Morrissey, D. J.; Seaborg, G. T. *Modern Nuclear Chemistry*; 2017.
10. Cotton, S. *Lanthanide and Actinide Chemistry*; John Wiley & Sons Ltd.: West Sussex, England, 2007.
11. Deblonde, G. J. P.; Abergel, R. J. Active Actinium. *Nat. Chem.* **2016**, *8* (11), 1084. https://doi.org/10.1038/nchem.2653.
12. Miederer, M.; Scheinberg, D. A.; McDevitt, M. R. Realizing the Potential of the Actinium-225 Radionuclide Generator in Targeted Alpha Particle Therapy Applications. *Adv. Drug Deliv. Rev.* **2008**, *60* (12), 1371–1382. https://doi.org/10.1016/j.addr.2008.04.009.
13. "H.R. 1960, National Defense Authorization Act (NDAA) for FY2014." 2013.
14. "S. 1197, NDAA for FY2014." Armed Services Committee. 2013.
15. "S. 1600, Critical Minerals Policy Act of 2013." Energy and Natural Resources Committee. 2013.
16. "H.R. 761, National Strategic and Critical Minerals Production Act of 2013." Committee on Natural Resources. 2013.

17. "H.R. 981, RARE Act of 2013." Subcommittee on Energy and Minerals Resources. 2013.
18. "H.R. 1063, National Strategic and Critical Minerals Policy Act of 2013." Subcommittee on Energy and Mineral Resources. 2013.
19. "P.L. 112-239 (H.R. 3310, 112th Congress), the National Defense Authorization Act for 2013." 2012.
20. "P.L. 111-393, the Ike Skelton National Defense Authorization Act for FY2011." 2010.
21. "H.R. 4402, the National Strategic and Critical Minerals Production Act." 2012.
22. Ryan, P. H.R. 3449, Defense Supply Chain and Industrial Base Security Act. 2011.
23. Coffman, M. H.R. 2184, Rare Earth Policy Task Force and Materials Act. 2011.
24. Hultgren, R. "H.R. 2090, Energy Critical Elements Advancement Act of 2011." Subcommittee on Energy and the Environment. 2011.
25. Lamborn, D. "H.R. 2011, National Strategic and Critical Minerals Policy Act of 2011." House Committee on Natural Resources. 2011.
26. Green, G. "H.R. 2284, Responsible Electronic Recycling Act." Subcommittee of Energy and the Environment. 2011.
27. Ciciline, D. "H.R. 1875, Building Our Clean Energy Future Now Act of 2011." House Committees on Ways and Means, Transportation and Infrastructure, Energy and Commerce, and Science, Space, and Technology. 2011.
28. Coffman, M. "H.R. 1388, the Rare Earths Supply Chain Technology and Resources Transformation Act of 2011." House Committee on Science, Space, and Technology, Subcommittee on Energy on the Environment, and the Committees on Natural Resources and Armed Services. 2011.
29. McKeon, H. "H.R. 1540, the National Defense Authorization Act for FY2012." 2011.
30. Stabenow, D. "S. 734, the Advanced Vehicle Technology Act of 2011." Committee on Natural Resources. 2011.
31. Peters, G. "H.R. 1367, the Advanced Vehicle Technology Act of 2011." Committee on Science, Space and Technology. 2011.
32. Johnson, H. "H.R. 1314, the Resource Assessment of Rare Earths (RARE) Act of 2011." House Natural Resources Committee, Subcommittee on Energy and Mineral Resources. 2011.
33. Miller, B. "H.R. 952, the Energy Critical Elements Renewal Act of 2011." Committee on Science, Space, and Technology. 2011.
34. Udall, M. "S. 383, the Critical Minerals and Materials Promotion Act of 2011." Committe on Energy and Natural Resources, Subcommittee on Energy. 2011.
35. Boswell, L. "H.R. 618, the Rare Earths and Critical Materials Revitalization Act of 2011." Committee on Science, Space, and Technology. 2011.
36. Murkowski, L. "S. 1113, the Critical Minerals Policy Act of 2011." Committee on Energy and Natural Resources. 2011.
37. Coffman, M. "H.R. 4866, the Rare Earths Supply-Chain Technology and Resources Transformation Act of 2010 (RESTART)." House Armed Services Committee, House Ways and Means Subcommittee on Trade, House of Financial Services Committee. 2010.
38. Dahlkemper, K. "H.R. 6160, the Rare Earths and Critical Materials Revitalization Act of 2010." National Materials and Minerals Policy. 2010.

39. Murkowski, L. "S. 3521, the Rare Earth Supply Technology and Resources Transformation Act of 2010." Committee on Energy and Natural Resources, Subcommittee on Energy. 2010.
40. Strategic and Critical Materials 2013. Report on Stockpile Requirements. 2013.
41. Critical Materials Strategy. U. S. Department of Energy. 2011.
42. "Rare Earth Materials in Defense Application." Report to Congress, U.S. Department of Defense. 2012.
43. "Assessment of Feasibility and Advisability of Establishment of Rare Earth Material Inventory." Report to Congress, U.S. Department of Defense. 2012.
44. "Feasibility and Desirability of Recycling, Recovery, and Reprocessing of Rare Earth Elements." Report to Congress, U.S. Department of Defense. 2012.
45. "Recovery of Rare Earth Elements from Fluorescent Lighting Materials." Report to Congress, U.S. Department of Defense. 2013.
46. "Diversification of Supply Chain and Reclamation Activities Related to Rare Earths." Report to Congress, U.S. Department of Defense. 2014.
47. "Implementation of Rare Earth Elements Strategy in the Joint Strike Fighter Program." Report to Congress, U.S. Department of Defense. 2014.
48. U.S. Geological Survey. *Mineral Commodity Summaries 2020*; 2020.
49. U. S. Geological Survey. *Rare Earths*; 2019.
50. Haxel, G. B.; Hedrick, J. B.; Orris, G. J. "Rare Earth Elements - Critical Resources for High Technology." U.S. Geological Survey, 2002, USGS Fact Sheet 087-02. 2002.
51. Pitts, M. Let's Take Better Care of Our Rare Earth Elements. *New Sci.* **2011**, 2799.
52. Bernhardt, D.; Reilly, J. F., II *Mineral Commodity Summaries 2019*; Restin, Virginia, 2019.
53. Schweitzer, G. K.; Pesterfield, L. L. *The Aqueous Chemistry of the Elements*; Oxford University Press: New York, 2010.
54. Meuer, H. W.; Strohmeir, E.; Dongarra, J.; Simon, H.; Meuer, M. *The TOP500 List*; 2020.
55. Strohmaier, E.; Meuer, H. W.; Dongarra, J.; Simon, H. D. The TOP500 List and Progress in High-Performance Computing. *IEEE Comput.* **2015**, *48* (11), 42–49. https://doi.org/10.1109/MC.2015.338.
56. Meuer, H.; Strohmaier, E.; Dongarra, J.; Simon, H. *Top500 Supercomputer Sites*, 17th ed.; 2001.
57. Dongarra, J.; Luszczek, P. TOP500. In *Encyclopedia of Parallel Computing*; Padua, D. , Ed.; Boston, MA, 2011.
58. Dongarra, J. A Not so Simple Matter of Software. *Sci. Comput.* **2005**, 14–16.
59. Alexander, F.; Almgren, A.; Bell, J.; Bhattacharjee, A.; Chen, J.; Colella, P.; Daniel, D.; DeSlippe, J.; Diachin, L.; Draeger, E.; Dubey, A.; Dunning, T.; Evans, T.; Foster, I.; Francois, M.; Germann, T.; Gordon, M.; Habib, S.; Halappanavar, M.; Hamilton, S.; Hart, W.; Huang, Z.; Hungerford, A.; Kasen, D.; Kent, P. R. C.; Kolev, T.; Kothe, D. B.; Kronfeld, A.; Luo, Y.; Mackenzie, P.; McCallen, D.; Messer, B.; Mniszewski, S.; Oehmen, C.; Perazzo, A.; Perez, D.; Richards, D.; Rider, W. J.; Rieben, R.; Roche, K.; Siegel, A.; Sprague, M.; Steefel, C.; Stevens, R.; Syamlal, M.; Taylor, M.; Turner, J.; Vay, J. L.; Voter, A. F.; Windus, T. L.; Yelick, K. Exascale Applications: Skin in the Game. *Philos. Trans. R. Soc. A* **2020**, *378*, 20190056. https://doi.org/10.1098/rsta.2019.0056.

60. Tajti, A.; Szalay, P. G.; Csaszar, A. G.; Kallay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vazquez, J.; Stanton, J. F. HEAT: High Accuracy Extrapolated Abinitio Thermochemistry. *J. Chem. Phys.* **2004**, *121* (23), 11599–11613.
61. Yang, J.; Hu, W.; Usvyat, D.; Matthews, D.; Schutz, M.; Chan, G. K. Ab Initio Determination of the Crystalline Benzene Lattice Energy to Sub-Kilojoule/Mole Accuracy. *Science* **2014**, *345* (6197), 640–643.
62. Booth, G. H.; Gruneis, A.; Kresse, G.; Alavi, A. Towards an Exact Descriptionof Electronic Wavefunctions in Real Solids. *Nature* **2013**, *493* (7432), 365–370.
63. Anderson, J.; Sundahl, B.; Harrison, R. J.; Beylkin, G. Dirac-Fock Calculations Onmolecules in an Adaptive Multiwavelet Basis. *J. Chem. Phys.* **2019**, *151* (23), 234112.
64. Kottmann, J.; Bischoff, F. A.; Valeev, E. F. Direct Determination of Optimalpair-Natural Orbitals in a Real-Space Representation: The Second-Order Moller–Plesset Energy. *J. Chem. Phys.* **2020**, *152* (7), 074105.
65. Bailey, D. H.; Barszcz, E.; Barton, J. T.; Browning, D. S.; Carter, R. L.; Dagum, L.; Fatoohi, R. A.; Frederickson, P. O.; Lasinski, T. A.; Schreiber, R. S. The NAS Parallel Benchmarks. *Int. J. High Perform. Comput. Appl.* **1991**, *5* (3), 63–73.
66. Bailey, D.; Barscz, E.; Barton, J.; Browning, D.; Carter, R.; Dagum, L.; Fatoohi, R.; Fineberg, S.; Frederickson, P.; Lasinski, T.; Schreiber, R.; Simon, H.; Venkatakrishnan, V.; Weeratunga, S. *The NAS Parallel Benchmarks. Technical Report NAS Technical Report RNR-94-007*; Moffett Field, CA, 1994.
67. Bailey, D. H.; Harris, T.; Saphir, W. C.; Van der Wijngaart, R. F.; Woo, A. C.; Yarrow, M. *The NAS Parallel Benchmarks 2.0. Technical Report NAS-95-020*; Moffett Field, CA, 1995.
68. Van der Wijngaart, R. F. *NAS Parallel Benchmarks Version 2.4. NAS Technical Report NAS-02-007*; Moffett Field, CA, 2002.
69. Henning, J. L. SPEC CPU2006 Benchmark Descriptions. In *SPEC CPU Subcommittee, and Performance Engineer*; 2006.
70. Ortega, D.; Martel, I.; Krishnan, V.; Ayguade, E.; Valero, M. Quantifying the Benefits of SPECint Distant Parallelism in Simultaneous Multithreading Architectures. In *1999 International Conference on Parallel Architectures and Compilation Techniques*; 1999; p (Cat. No.PR00425) Access restricted.
71. Heroux, M. A.; Doerfler, D. W.; Crozier, P. S.; Willenbring, J. M.; Edwards, H. C.; Williams, A.; Rajan, M.; Keiter, E. R.; Thornquist, H. K.; Numrich, R. W. *Improving Performance via Mini-Applications. Technical Report SAND2009-5574*; 2009.
72. Vetter, J. Building Analytical Models into an Interactive Performance Prediction Tool. *Supercomputing '89*; Dept. of Comput. Sci., Indiana Univ., 1989; pp 521–530.
73. Event-Based Performance Perturbation: A Case Study. *SIGPLAN Notices (Third ACM SIGPLAN Symp. Principles and Practice of Parallel Programming)*; Center for Supercomput. Res. & Dev., Illinois Univ.: Urbana, IL, 1991.
74. Dongarra, J. J.; Luszczek, P.; Petitet, A. The LINPACK Benchmark: Past, Present, and Future. *Concurr. Comput. Pract. Exp.* **2003**, *15* (9), 803–820.
75. Feng, W. C.; Cameron, K. W. The Green500 List: Encouraging Sustainable Supercomputing. *IEEE Comput.* **2007**, *40* (12), 50–55.

76. Dongarra, J.; Heroux, M. A.; Luszczek, P. High-Performance Conjugate-Gradient Benchmark: A New Metric for Ranking High-Performance Computing Systems. *The International Journal of High Performance Computing Applications*; 2016. https://doi.org/10.1177/1094342015593158.
77. Saad, Y.; Schultz, M. H. GMRES: A Generalized Minimal Residual Algorithm for Solving Non-Symmetric Linear Systems. *SIAM J. Sci. Stat. Comput.* **1986**, *7*, 856–869.
78. Dongarra, J.; Heroux, M. A.; Luszczek, P. A New Metric for Ranking High-Performance Computing Systems. *Natl. Sci. Rev.* **2016**, *3* (1), 30–35. https://doi.org/10.1093/nsr/nwv084.
79. Dongarra, J. J. An Overview of High Performance Computing and Challenges for the Future. *2020 IEEE/ACM International Conference for High Performance Computing, Networking, Storage, and Analysis*; 2020.
80. Terpstra, D.; Jagode, H.; You, H.; Dongarra, J. Collecting Performance Data with PAPI-C. *Tools for High Performance Computing*; 2010; pp 157–173.
81. Haidar, A.; Jagode, H.; Vaccaro, P.; YarKhan, A.; Tomov, S.; Dongarra, J. Investigating Power Capping toward Energy-Efficient Scientific Applications. *Concurr. Comput. Pract. Exp. CCPE Spec. Issue Power-Aware Comput.* **2018** (e4485), 1–14.
82. Haidar, A.; Jagode, H.; YarKhan, A.; Vaccaro, P.; Tomov, S.; Dongarra, J. Power-Aware Computing: Measurement, Control, and Performance Analysis for Intel Xeon Phi. In *2017 IEEE High Performance Extreme Computing Conference (HPEC '17)*; 2017.
83. Jagode, H.; YarKhan, A.; Danalis, A.; Dongarra, J. Power Management and Event Verification in PAPI. *Tools for High Performance Computing 2015, 9th International Workshop on Parallel Tools for High Performance Computing*; Springer International Publishing: Dresden, Germany, 2016; pp 41–51.
84. Jagode-McCraw, H.; Ralph, J.; Danalis, A.; Dongarra, J. Power Monitoring with PAPI for Extreme Scale Architectures and Dataflow-Based Programming Models. *Workshop on Monitoring and Analysis for High Performance Computing Systems Plus Applications (HPCMASPA 2014), IEEE Cluster 2014*; Madrid, Spain, 2014; pp 385–391.
85. Jagode-McCraw, H.; Terpstra, D.; Dongarra, J.; Davis, K. M. R. Beyond the CPU: Hardware Performance Counter Monitoring on Blue Gene/Q. In *International Supercomputing Conference 2013 (ISC'13), Leipzig, Germany, June 16−20, 2013*; Kunkel, J. M., Ludwig, T., Meuer, H. , Eds.; Springer, 2013.
86. Malony, A.; Biersdorff, S.; Shende, S.; Jagode, H.; Tomov, S.; Juckeland, G.; Dietrich, R.; Poole, D.; Lamb, C. Parallel Performance Measurement of Heterogeneous Parallel Systems with GPUs. *ICPP 2011*; Taipei, Taiwan, 2011.
87. Jagode, H.; Danalis, A.; Anzt, H.; Dongarra, J. PAPI Software-Defined Events for in-Depth Performance Analysis. *Int. J. High Perform. Comput. Appl.* **2019**, *33* (6), 1113–1127.
88. Danalis, A.; Jagode, H.; Herault, T.; Luszczek, P.; Dongarra, J. Software-Defined Events through PAPI. *24th International Workshop on High-Level Parallel Programming Models and Supportive Environments (HIPS), in conjunction with 33rd IEEE International Parallel & Distributed Processing Symposium (IPDPS)*; Rio de Janeiro, Brazil, 2019; pp 363–372.
89. ARM Map. https://developer.arm.com/tools-and-software/server-and-hpc/debug-and-profile/arm-forge%0A.

90. Adhianto, L.; Banerjee, S.; Fagan, M.; Krentel, M.; Marin, G.; Mellor-Crummey, J.; Tallent, N. HPCToolkit: Tools for Performance Analysis of Optimized Parallel Programs. *Concurr. Comput. Pract. Exp.* **2010**, *22* (6), 685–701.
91. Shende, S.; Malony, A. D. The TAU Parallel Performance System. *SAGE Publ. Int. J. High Perform. Comput. Appl.* **2006**, *20* (2), 287–311.
92. Knüpfer, A.; Brunst, H.; Doleschal, J.; Jurenz, M.; Lieber, M.; Mickler, H.; Müller, M. S.; Nagel, W. E. The Vampir Performance Analysis Tool Set. *Tools for High Performance Computing*; 2008; pp 139–155.
93. Knüpfer, A.; Rossel, C.; Mey, D.; Biersdorff, S.; Diethelm, K.; Eschweiler, D.; Geimer, M.; Gerndt, M.; Lorenz, D.; Malony, A.; Nagel, W.; Oleynik, Y.; Philippen, P.; Saviankou, P.; Schmidl, D.; Shende, S.; Tschuter, R.; Wagner, M.; Wesarg, B.; Wolf, F. Score-P: A Joint Performance Measurement Run-Time Infrastructure for Periscope, Scalasca, TAU, and Vampir. *Tools for High Performance Computing*; Springer: Berlin Heidelberg, 2011; pp 79–91.
94. Abdelfattah, A.; Anzt, H.; Boman, E. G.; Carson, E.; Cojean, T.; Dongarra, J.; Fox, A.; Gates, M.; Higham, N. J.; Li, X. S.; Loe, J.; Luszczek, P.; Pranesh, S.; Rajamanickam, S.; Ribizel, T.; Smith, B. F.; Swirydowicz, K.; Thomas, S.; Tomov, S.; Tsai, Y. M.; Meier Yang, U. A Survey of Numerical Linear Algebra Methods Utilizing Mixed-Precision Arithmetic. *Int. J. High Perform. Comput.* **2021**. https://doi.org/10.1177/10943420211003313.
95. Jagode, H.; Danalis, A.; Anzt, H.; Dongarra, J. PAPI Software-Defined Events for in-Depth Performance Analysis. *Int. J. High Perform. Comput. Appl.* **2019**, *33* (6), 1113–1127.
96. Kowalski, K.; Krishnamoorthy, S.; Olson, R. M.; Tipparaju, V.; Apra, E. Scalable Implementations of Accurate Excited-State Coupled Cluster Theories: Application of High-Level Methods to Porphyrin-Based Systems. *International Conference for High Performance Computing, Networking, Storage and Analysis*; ACM: New York, NY, p 72:1−72:10. https://doi.org/10.1145/2063384.2063481.
97. Aprà, E.; Bylaska, E. J.; De Jong, W. A.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Valiev, M.; Van Dam, H. J. J.; Alexeev, Y.; Anchell, J.; Anisimov, V.; Aquino, F. W.; Atta-Fynn, R.; Autschbach, J.; Bauman, N. P.; Becca, J. C.; Bernholdt, D. E.; Bhaskaran-Nair, K.; Bogatko, S.; Borowski, P.; Boschen, J.; Brabec, J.; Bruner, A.; Cauët, E.; Chen, Y.; Chuev, G. N.; Cramer, C. J.; Daily, J.; Deegan, M. J. O.; Dunning, T. H.; Dupuis, M.; Dyall, K. G.; Fann, G. I.; Fischer, S. A.; Fonari, A.; Früchtl, H.; Gagliardi, L.; Garza, J.; Gawande, N.; Ghosh, S.; Glaesemann, K.; Götz, A. W.; Hammond, J.; Helms, V.; Hermes, E. D.; Hirao, K.; Hirata, S.; Jacquelin, M.; Jensen, L.; Johnson, B. G.; Jónsson, H.; Kendall, R. A.; Klemm, M.; Kobayashi, R.; Konkov, V.; Krishnamoorthy, S.; Krishnan, M.; Lin, Z.; Lins, R. D.; Littlefield, R. J.; Logsdail, A. J.; Lopata, K.; Ma, W.; Marenich, A. V.; Martin Del Campo, J.; Mejia-Rodriguez, D.; Moore, J. E.; Mullin, J. M.; Nakajima, T.; Nascimento, D. R.; Nichols, J. A.; Nichols, P. J.; Nieplocha, J.; Otero-De-La-Roza, A.; Palmer, B.; Panyala, A.; Pirojsirikul, T.; Peng, B.; Peverati, R.; Pittner, J.; Pollack, L.; Richard, R. M.; Sadayappan, P.; Schatz, G. C.; Shelton, W. A.; Silverstein, D. W.; Smith, D. M. A.; Soares, T. A.; Song, D.; Swart, M.; Taylor, H. L.; Thomas, G. S.; Tipparaju, V.; Truhlar, D. G.; Tsemekhman, K.; Van Voorhis, T.; Vázquez-Mayagoitia, A.; Verma, P.; Villa, O.; Vishnu, A.; Vogiatzis, K. D.; Wang, D.; Weare, J. H.; Williamson, M. J.; Windus, T. L.; Woliński, K.; Wong, A. T.; Wu, Q.; Yang, C.; Yu, Q.;

Zacharias, M.; Zhang, Z.; Zhao, Y.; Harrison, R. J. NWChem: Past, Present, and Future. *J. Chem. Phys.* **2020**, *152* (18), 184102. https://doi.org/10.1063/5.0004997.

98. Harrison, R. J.; Beylkin, G.; Bischoff, F. A.; Justus, A.; Fann, G. I.; Fosso-tande, J.; Galindo, D.; Jeff, R.; Hartman-baker, R.; Hill, J. C.; Jia, J. U. N.; Jakob, S.; Ou, M. Y.; Pei, J.; Ratcliff, L. E.; Reuter, M. G.; Richie-halford, A. C.; Romero, N. A.; Sekino, H.; Shelton, W. A.; Sundahl, B. E.; Scott, W.; Al, H. E. T. MADNESS: A Multiresolution, Adaptive Numerical Environment for Scientific Simulation. *Soc. Ind. Appl. Math.* **2016**, *38* (5), 123–142.
99. Harrison, R. J.; Fann, G. I.; Yanai, T.; Gan, Z.; Beylkin, G. Multiresolution Quantum Chemistry: Basic Theory and Initial Applications. *J. Chem. Phys.* **2004**, *121* (23), 11587–11598.
100. Pei, J. C.; Fann, G. I.; Nazarewicz, W.; Harrison, R. J.; Hill, J.; Galindo, D.; Jia, J. Coordinate-Space Hartree-Fock-Bogoliubov for Superfluid Fermi Systems in Large Boxes. *J. Phys. Conf. Ser.* **2012**, *402*.
101. Reuter, M.; Hill, J.; Harrison, R. J. Solving PDEs in Irregular Geometries with Multiresolution Methods I: Embedded Dirichlet Boundary Conditions. *Comput. Phys. Commun.* **2012**, *183*, 1–7.
102. Vence, N.; Harrison, R.; Krstic, P. Attosecond Electron Dynamics: A Multiresolution Approach. *Phys. Rev. A* **2012**, *85* (3), 33403. https://doi.org/10.1103/PhysRevA.85.033403.
103. Yanai, T.; Fann, G. I.; Gan, Z.; Harrison, R. J.; Beylkin, G. Multiresolution Quantum Chemistry in Multiwavelet Bases: Analytic Derivatives for Hartre Fock and Density Functional Theory. *J. Chem. Phys.* **2004**, *121* (7), 2866. https://doi.org/10.1063/1.1768161.
104. Yanai, T.; Fann, G. I.; Gan, Z.; Harrison, R. J.; Beylkin, G. Multiresolution Quantum Chemistry in Multiwavelet Bases: Hartree Exchange. *J. Chem. Phys.* **2004**, *121* (14), 6680. https://doi.org/10.1063/1.1790931.
105. Yanai, T.; Fann, G. I.; Gan, Z.; Harrison, R. J.; Beylkin, G. Multiresolution Quantum Chemistry: Hartree-Fock Exchange. *J. Chem. Phys.* **2004**, *121* (14)
106. Jia, J.; Hill, J. C.; Beylkin, G.; Harrison, R. J. MULTIRESOLUTION FAST METHODS FOR A PERIODIC 3-D NAVIER STOKES SOLVER. *Proceedings of the 8th International Symposium on Distributed Computing and Applications to Business, Engineering and Science*; 2009; pp 13–16.
107. Bischoff, F. A.; Harrison, R. J.; Valeev, E. F. Computing Many-Body Wave Functions with Guaranteed Precision: The First-Order Moller-Plesset Wave Function for the Ground State of Helium Atom. *J. Chem. Phys.* **2012**, *137*, 104103.
108. Bischoff, F. A.; Valeev, E. F. Computing Molecular Correlation Energies with Guaranteed Precision. *J. Chem. Phys.* **2013**, *139* (11), 114106. https://doi.org/10.1063/1.4820404.
109. Kottmann, J. S.; Bischoff, F. A. Coupled-Cluster in Real Space. 1. CC2 Ground State Energies Using Multiresolution Analysis. *J. Chem. Theory Comput.* **2017**, *13* (12), 5945–5955. https://doi.org/10.1021/acs.jctc.7b00694.
110. Kottmann, J. S.; Bischoff, F. A. Coupled-Cluster in Real Space. 2. CC2 Excited States Using Multiresolution Analysis. *J. Chem. Theory Comput.* **2017**, *13* (12), 5956–5965. https://doi.org/10.1021/acs.jctc.7b00695.
111. Bischoff, F. A.; Valeev, E. F. Low-Order Tensor Approximations for Electronic Wave Functions: Hartree-Fock Method with Guaranteed Precision. *J. Chem. Phys.* **2011**, *134* (10), 104104. https://doi.org/10.1063/1.3560091.

112. Jensen, S. R.; Flå, T.; Jonsson, D.; Monstad, R. S.; Ruud, K.; Frediani, L. Magnetic Properties with Multiwavelets and DFT: The Complete Basis Set Limit Achieved. *Phys. Chem. Chem. Phys.* **2016**, *18* (31), 21145–21161. https://doi.org/10.1039/C6CP01294A.
113. Beylkin, G.; Coifman, R.; Rokhlin, V. Fast Wavelet Transforms and Numerical Algorithms I. *Commun. Pure Appl. Math.* **1991**, *44* (2), 141–183. https://doi.org/10.1017/CBO9780511662294.012.
114. Alpert, B. K. A Class of Bases in L2 for the Sparse Representation of Integral Operators. *SIAM J. Math. Anal.* **1993**, *24* (1), 246–262. https://doi.org/10.1137/0524016.
115. Beylkin, G.; Cheruvu, V.; Perez, F. Fast Adaptive Algorithms in the Non-Standard Form for Multidimensional Problems. *Appl. Comput. Harmon Anal.* **2008**, *24* (3), 354–377. https://doi.org/10.1016/j.acha.2007.08.001.
116. Boys, S. F.; Egerton, A. C. Electronic Wave Functions - I. A General Method of Calculation for the Stationary States of Any Molecular System. *Proc. R. Soc. London Ser. Math. Phys. Sci.* **1950**, *200* (1063), 542–554. https://doi.org/10.1098/rspa.1950.0036.
117. Baerends, E. J.; Ellis, D. E.; Ros, P. Self-Consistent Molecular Hartree—Fock—Slater Calculations I. The Computational Procedure. *Chem. Phys.* **1973**, *2* (1), 41–51. https://doi.org/10.1016/0301-0104(73)80059-X.
118. Watson, M. A.; Handy, N. C.; Cohen, A. J. Density Functional Calculations, Using Slater Basis Sets, with Exact Exchange. *J. Chem. Phys.* **2003**, *119* (13), 6475–6481. https://doi.org/10.1063/1.1604371.
119. Averill, F. W.; Ellis, D. E. An Efficient Numerical Multicenter Basis Set for Molecular Orbital Calculations: Application to FeCl4. *J. Chem. Phys.* **1973**, *59* (12), 6412–6418. https://doi.org/10.1063/1.1680020.
120. Delley, B. An All-electron Numerical Method for Solving the Local Density Functional for Polyatomic Molecules. *J. Chem. Phys.* **1990**, *92* (1), 508–517. https://doi.org/10.1063/1.458452.
121. Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab Initio Molecular Simulations with Numeric Atom-Centered Orbitals. *Comput. Phys. Commun.* **2009**, *180* (11), 2175–2196. https://doi.org/10.1016/j.cpc.2009.06.022.
122. Wills, J. M.; Cooper, B. R. Synthesis of Band and Model Hamiltonian Theory for Hybridizing Cerium Systems. *Phys. Rev. B* **1987**, *36* (7), 3809–3823. https://doi.org/10.1103/PhysRevB.36.3809.
123. Wimmer, E.; Krakauer, H.; Weinert, M.; Freeman, A. J. Full-Potential Self-Consistent Linearized-Augmented-Plane-Wave Method for Calculating the Electronic Structure of Molecules and Surfaces: O_2 Molecule. *Phys. Rev. B* **1981**, *24* (2), 864–875. https://doi.org/10.1103/PhysRevB.24.864.
124. Madsen, G. K. H.; Blaha, P.; Schwarz, K.; Sjöstedt, E.; Nordström, L. Efficient Linearization of the Augmented Plane-Wave Method. *Phys. Rev. B* **2001**, *64* (19), 195134. https://doi.org/10.1103/PhysRevB.64.195134.
125. Michalicek, G.; Betzinger, M.; Friedrich, C.; Blügel, S. Elimination of the Linearization Error and Improved Basis-Set Convergence within the FLAPW Method. *Comput. Phys. Commun.* **2013**, *184* (12), 2670–2679. https://doi.org/10.1016/j.cpc.2013.07.002.
126. Pahl, F. A.; Handy, N. C. Plane Waves and Radial Polynomials: A New Mixed Basis. *Mol. Phys.* **2002**, *100* (20), 3199–3224. https://doi.org/10.1080/00268970210133206.

127. Anderson, J.; Sundahl, B.; Harrison, R.; Beylkin, G. Dirac-Fock Calculations on Molecules in an Adaptive Multiwavelet Basis. *J. Chem. Phys.* **2019**, *151* (23), 234112. https://doi.org/10.1063/1.5128908.
128. Blackledge, J.; Babajanov, B. On the Dirac Scattering Problems. *Math. Aeterna* **2013**, *3* (7), 535–544. https://doi.org/10.21427/D7JP6B.
129. Peng, C.; Lewis, C. A.; Wang, X.; Clement, M. C.; Pierce, K.; Rishi, V.; Pavošević, F.; Slattery, S.; Zhang, J.; Teke, N.; Kumar, A.; Masteran, C.; Asadchev, A.; Calvin, J. A.; Valeeva, E. F. Massively Parallel Quantum Chemistry: A High-Performance Research Platform for Electronic Structure. *J. Chem. Phys.* **2020**, *153*, 044120.
130. Calvin, J. A.; Lewis, C. A.; Valeev, E. F. Scalable Task-Based Algorithm for Multiplication of Block-Rank-Sparse Matrices. *IA3 15 Proc. 5th Workshop Irregul. Appl. Archit. Algorithms* **2015**, *4*, 1–8.
131. Peng, C.; Calvin, J. A.; Valeev, E. F. Coupled-cluster Singles, Doubles and Perturbative Triples with Density Fitting Approximation for Massively Parallel Heterogeneous Platforms. *Int. J. Quantum Chem.* **2019**, *119* (12), e25894.
132. Calvin, J. A.; Valeeva, E. F. TiledArray: A general-purpose scalable block-sparse tensor framework. https://github.com/valeevgroup/tiledarray.
133. Auer, A. A.; Baumgartner, G.; Bernholdt, D. E.; Bibireata, A.; Choppella, V.; Cociorva, D.; Gao, X.; Harrison, R. J.; Krishnamoorthy, S.; Krishnan, S.; Lam, C.; Lu, Q.; Nooijen, M.; Pitzer, R.; Ramanujam, J.; Sadayappan, P.; Sibiryakov, A. Automatic Code Generation for Many-Body Electronic Structure Methods: The Tensor Contraction Engine. *Mol. Phys.* **2006**, *104* (2), 211–228.
134. Hansen, K.; Montavon, G.; Biegler, F.; Fazli, S.; Rupp, M.; Scheffler, M.; Von Lilienfeld, O. A.; Tkatchenko, A.; Müller, K. R. Assessment and Validation of Machine Learning Methods for Predicting Molecular Atomization Energies. *J. Chem. Theory Comput.* **2013**, *9* (8), 3404–3419. https://doi.org/10.1021/ct400195d.
135. Pilania, G.; Mannodi-Kanakkithodi, A.; Uberuaga, B. P.; Ramprasad, R.; Gubernatis, J. E.; Lookman, T. Machine Learning Bandgaps of Double Perovskites. *Sci. Rep.* **2016**, *6* (January), 1–10. https://doi.org/10.1038/srep19375.
136. Schütt, K. T.; Glawe, H.; Brockherde, F.; Sanna, A.; Müller, K. R.; Gross, E. K. U. How to Represent Crystal Structures for Machine Learning: Towards Fast Prediction of Electronic Properties. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2014**, *89* (20), 1–5. https://doi.org/10.1103/PhysRevB.89.205118.
137. Huang, B.; Von Lilienfeld, O. A. Communication: Understanding Molecular Representations in Machine Learning: The Role of Uniqueness and Target Similarity. *J. Chem. Phys.* **2016**, *145* (16), 161102. https://doi.org/10.1063/1.4964627.
138. Ghosh, K.; Stuke, A.; Todorović, M.; Jørgensen, P. B.; Schmidt, M. N.; Vehtari, A.; Rinke, P. Deep Learning Spectroscopy: Neural Networks for Molecular Excitation Spectra. *Adv. Sci.* **2019**, *6* (9), 1801367. https://doi.org/10.1002/advs.201801367.
139. Selzer, P.; Gasteiger, J.; Thomas, H.; Salzer, R. Rapid Access to Infrared Reference Spectra of Arbitrary Organic Compounds: Scope and Limitations of an Approach to the Simulation of Infrared Spectra by Neural Networks. *Chem.—Eur. J.* **2000**, *6* (5), 920–927. https://doi.org/10.1002/(SICI)1521-3765(20000303)6:5<920::AID-CHEM920>3.0.CO;2-W.

140. Qu, X.; Huang, Y.; Lu, H.; Qiu, T.; Guo, D.; Agback, T.; Orekhov, V.; Chen, Z. *Accelerated Nuclear Magnetic Resonance Spectroscopy with Deep Learning*; 2019.
141. Mater, A. C.; Coote, M. L. Deep Learning in Chemistry. *J. Chem. Inf. Model.* **2019**, *59*, 2545–2559. https://doi.org/10.1021/acs.jcim.9b00266.
142. Montavon, G.; Rupp, M.; Gobre, V.; Vazquez-Mayagoitia, A.; Hansen, K.; Tkatchenko, A.; Müller, K. R.; Anatole Von Lilienfeld, O. Machine Learning of Molecular Electronic Properties in Chemical Compound Space. *New J. Phys.* **2013**, *15*, 095003. https://doi.org/10.1088/1367-2630/15/9/095003.
143. Gomez-Fernandez, M.; Higley, K.; Tokuhiro, A.; Welter, K.; Wong, W. K.; Yang, H. Status of Research and Development of Learning-Based Approaches in Nuclear Science and Engineering: A Review. *Nucl. Eng. Des.* **2020**, *359* (October 2019), 110479. https://doi.org/10.1016/j.nucengdes.2019.110479.
144. Fagan, D. K.; Robinson, S. M.; Runkle, R. C. Statistical Methods Applied to Gamma-Ray Spectroscopy Algorithms in Nuclear Security Missions. *Applied Radiation and Isotopes*; Pergamon: October 2012; pp 2428–2439. https://doi.org/10.1016/j.apradiso.2012.06.016.
145. Hata, H.; Yokoyama, K.; Ishimori, Y.; Ohara, Y.; Tanaka, Y.; Sugitsue, N. Application of Support Vector Machine to Rapid Classification of Uranium Waste Drums Using Low-Resolution γ-Ray Spectra. *Appl. Radiat. Isot.* **2015**, *104*, 143–146. https://doi.org/10.1016/j.apradiso.2015.06.030.
146. Kamuda, M.; Zhao, J.; Huff, K. A Comparison of Machine Learning Methods for Automated Gamma-Ray Spectroscopy. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.* **2020**, *954*, 161385. https://doi.org/10.1016/j.nima.2018.10.063.
147. Abdel-Aal, R. E.; Al-Haddad, M. N. Determination of Radioisotopes in Gamma-Ray Spectroscopy Using Abductive Machine Learning. *Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spectrometers Detect. Assoc. Equip.* **1997**, *391* (2), 275–288. https://doi.org/10.1016/S0168-9002(97)00391-4.
148. Boldrini, L.; Bibault, J. E.; Masciocchi, C.; Shen, Y.; Bittner, M. I. Deep Learning: A Review for the Radiation Oncologist. *Front. Oncol.* **2019**, *9* (October) https://doi.org/10.3389/fonc.2019.00977.
149. Lou, B.; Doken, S.; Zhuang, T.; Wingerter, D.; Gidwani, M.; Mistry, N.; Ladic, L.; Kamen, A.; Abazeed, M. E. An Image-Based Deep Learning Framework for Individualising Radiotherapy Dose: A Retrospective Analysis of Outcome Prediction. *Lancet Digit. Health* **2019**, *1* (3), e136–e147. https://doi.org/10.1016/S2589-7500(19)30058-5.
150. Pucci, C.; Martinelli, C.; Ciofani, G. Innovative Approaches for Cancer Treatment: Current Perspectives and New Challenges. *Ecancermedicalscience* **2019**, *13*, 1–26. https://doi.org/10.3332/ecancer.2019.961.
151. Varley, A.; Tyler, A.; Smith, L.; Dale, P.; Davies, M. Remediating Radium Contaminated Legacy Sites: Advances Made through Machine Learning in Routine Monitoring of "Hot" Particles. *Sci. Total Environ.* **2015**, *521–522*, 270–279. https://doi.org/10.1016/j.scitotenv.2015.03.131.
152. Varley, A.; Tyler, A.; Smith, L.; Dale, P. Development of a Neural Network Approach to Characterise 226Ra Contamination at Legacy Sites Using Gamma-Ray Spectra Taken from Boreholes. *J. Environ. Radioact.* **2015**, *140*, 130–140. https://doi.org/10.1016/j.jenvrad.2014.11.011.

153. Calivá, F.; De Ribeiro, F. S.; Mylonakis, A.; Demazirere, C.; Vinai, P.; Leontidis, G.; Kollias, S. A Deep Learning Approach to Anomaly Detection in Nuclear Reactors. *Proc. Int. Jt. Conf. Neural Netw.* **2018**, *2018* (July) https://doi.org/10.1109/IJCNN.2018.8489130.
154. Vallet, V.; Grenthe, I. Structure and Bonding in Uranyl(VI) Peroxide and Crown Ether Complexes; Comparison of Quantum Chemical and Experimental Data. *Inorg. Chem.* **2017**, *56* (24), 15231–15240. https://doi.org/10.1021/acs.inorgchem.7b02584.
155. Kervazo, S.; Réal, F.; Virot, F.; Severo Pereira Gomes, A.; Vallet, V. Accurate Predictions of Volatile Plutonium Thermodynamic Properties. *Inorg. Chem.* **2019**, *58* (21), 14507–14521. https://doi.org/10.1021/acs.inorgchem.9b02096.
156. Lindqvist-Reis, P.; Réal, F.; Janicki, R.; Vallet, V. Unraveling the Ground State and Excited State Structures and Dynamics of Hydrated Ce^{3+} Ions by Experiment and Theory. *Inorg. Chem.* **2018**, *57* (16), 10111–10121. https://doi.org/10.1021/acs.inorgchem.8b01224.
157. Oher, H.; Réal, F.; Vercouter, T.; Vallet, V. Investigation of the Luminescence of $[UO_2X_4]^{2-}$ (X = Cl, Br) Complexes in the Organic Phase Using Time-Resolved Laser-Induced Fluorescence Spectroscopy and Quantum Chemical Simulations. *Inorg. Chem.* **2020**, *59* (9), 5896–5906. https://doi.org/10.1021/acs.inorgchem.9b03614.
158. Ferrier, M. G.; Batista, E. R.; Berg, J. M.; Birnbaum, E. R.; Cross, J. N.; Engle, J. W.; La Pierre, H. S.; Kozimor, S. A.; Lezama Pacheco, J. S.; Stein, B. W.; Stieber, S. C. E.; Wilson, J. J. Spectroscopic and Computational Investigation of Actinium Coordination Chemistry. *Nat. Commun.* **2016**, *7*, 12312. https://doi.org/10.1038/ncomms12312.
159. Wacker, J. N.; Vasiliu, M.; Huang, K.; Baumbach, R. E.; Bertke, J. A.; Dixon, D. A.; Knope, K. E. Uranium(IV) Chloride Complexes: UCl_6^{2-} and an Unprecedented $U(H_2O)_4Cl_4$ Structural Unit. *Inorg. Chem.* **2017**, *56* (16), 9772–9780. https://doi.org/10.1021/acs.inorgchem.7b01293.
160. Wacker, J. N.; Vasiliu, M.; Huang, K.; Baumbach, R. E.; Bertke, J. A.; Dixon, D. A.; Knope, K. E. Uranium(IV) Chloride Complexes: UCl_6^{2-} and an Unprecedented $U(H_2O)_4Cl_4$ Structural Unit. *Inorg. Chem.* **2017**, *56* (16), 9772–9780. https://doi.org/10.1021/acs.inorgchem.7b01293.
161. Cox, R. M.; Kafle, A.; Armentrout, P. B.; Peterson, K. A. Bond Energy of ThN^+: A Guided Ion Beam and Quantum Chemical Investigation of the Reactions of Thorium Cation with N_2 and NO. *J. Chem. Phys.* **2019**, *151* (3), 034304. https://doi.org/10.1063/1.5111534.
162. Feng, R.; Glendening, E. D.; Peterson, K. A. Actinyl Cation-Cation Interactions in the Gas Phase: An Accurate Thermochemical Study. *Phys. Chem. Chem. Phys.* **2019**, *21* (15), 7953–7964. https://doi.org/10.1039/c9cp00760a.
163. Cross, J. N.; Su, J.; Batista, E. R.; Cary, S. K.; Evans, W. J.; Kozimor, S. A.; Mocko, V.; Scott, B. L.; Stein, B. W.; Windorff, C. J.; Yang, P. Covalency in Americium(III) Hexachloride. *J. Am. Chem. Soc.* **2017**, *139* (25), 8667–8677. https://doi.org/10.1021/jacs.7b03755.
164. Daly, S. R.; Keith, J. M.; Batista, E. R.; Boland, K. S.; Clark, D. L.; Kozimor, S. A.; Martin, R. L. Sulfur K-Edge X-Ray Absorption Spectroscopy and Time-Dependent Density Functional Theory of Dithiophosphinate Extractants: Minor Actinide Selectivity and Electronic Structure Correlations. *J. Am. Chem. Soc.* **2012**, *134* (35), 14408–14422. https://doi.org/10.1021/ja303999q.

165. Ferrier, M. G.; Stein, B. W.; Batista, E. R.; Berg, J. M.; Birnbaum, E. R.; Engle, J. W.; John, K. D.; Kozimor, S. A.; Pacheco, J. S. L.; Redman, L. N. Synthesis and Characterization of the Actinium Aquo Ion. *ACS Cent. Sci.* **2017**, *3* (3), 176–185. https://doi.org/10.1021/acscentsci.6b00356.
166. Infante, I.; Kovacs, A.; MacChia, G. L.; Shahi, A. R. M.; Gibson, J. K.; Gagliardi, L. Ionization Energies for the Actinide Mono- and Dioxides Series, from Th to Cm: Theory versus Experiment. *J. Phys. Chem. A* **2010**, *114* (19), 6007–6015. https://doi.org/10.1021/jp1016328.
167. Roos, B. O.; Malmqvist, P. Å.; Gagliardi, L. Exploring the Actinide-Actinide Bond: Theoretical Studies of the Chemical Bond in Ac2, Th2, Pa2, and U2. *J. Am. Chem. Soc.* **2006**, *128* (51), 17000–17006. https://doi.org/10.1021/ja066615z.
168. Sharma, P.; Pahls, D. R.; Ramirez, B. L.; Lu, C. C.; Gagliardi, L. Multiple Bonds in Uranium–Transition Metal Complexes. *Inorg. Chem.* **2019**, *58* (15), 10139–10147. https://doi.org/10.1021/acs.inorgchem.9b01264.
169. Spezia, R.; Beuchat, C.; Vuilleumier, R.; D'Angelo, P.; Gagliardi, L. Unravelling the Hydration Structure of ThX_4 (X = Br, Cl) Water Solutions by Molecular Dynamics Simulations and X-Ray Absorption Spectroscopy. *J. Phys. Chem. B* **2012**, *116* (22), 6465–6475. https://doi.org/10.1021/jp210350b.
170. Hastings, A. M.; Ray, D.; Jeong, W.; Gagliardi, L.; Farha, O. K.; Hixon, A. E. Advancement of Actinide Metal–Organic Framework Chemistry via Synthesis of Pu-UiO-66. *J. Am. Chem. Soc.* **2020**, *142* (20), 9363–9371. https://doi.org/10.1021/jacs.0c01895.
171. Bagus, P. S.; Nelin, C. J.; Hrovat, D. A.; Ilton, E. S. Covalent Bonding in Heavy Metal Oxides. *J. Chem. Phys.* **2017**, *146* (13), 134706. https://doi.org/10.1063/1.4979018.
172. Bagus, P. S. Covalent Interactions in Oxides. *J. Electron Spectrosc. Relat. Phenom.* **2014**, 8.
173. Ilton, E. S.; Bagus, P. S. Ligand Field Effects on the Multiplet Structure of the U4f XPS of UO_2. *Surf. Sci.* **2008**, *602* (5), 1114–1121. https://doi.org/10.1016/j.susc.2008.01.010.
174. Bagus, P. S.; Nelin, C. J.; Al-Salik, Y.; Ilton, E. S.; Idriss, H. Multiplet Splitting for the XPS of Heavy Elements: Dependence on Oxidation State. *Surf. Sci.* **2016**, *643*, 142–149. https://doi.org/10.1016/j.susc.2015.06.002.
175. Ilton, E. S.; Du, Y.; Stubbs, J. E.; Eng, P. J.; Chaka, A. M.; Bargar, J. R.; Nelin, C. J.; Bagus, P. S. Quantifying Small Changes in Uranium Oxidation States Using XPS of a Shallow Core Level. *Phys. Chem. Chem. Phys.* **2017**, *19* (45), 30473–30480. https://doi.org/10.1039/C7CP05805E.
176. Bagus, P. S.; Ilton, E. S.; Martin, R. L.; Jensen, H. J. Aa.; Knecht, S. Spin–Orbit Coupling in Actinide Cations. *Chem. Phys. Lett.* **2012**, *546*, 58–62. https://doi.org/10.1016/j.cplett.2012.07.035.
177. Bagus, P. S.; Nelin, C. J.; Ilton, E. S.; Baron, M.; Abbott, H.; Primorac, E.; Kuhlenbeck, H.; Shaikhutdinov, S.; Freund, H.-J. The Complex Core Level Spectra of CeO2: An Analysis in Terms of Atomic and Charge Transfer Effects. *Chem. Phys. Lett.* **2010**, *487* (4–6), 237–240. https://doi.org/10.1016/j.cplett.2010.01.041.
178. Bagus, P. S.; Nelin, C. J.; Ilton, E. S. Theoretical Modeling of the Uranium 4f XPS for U(VI) and U(IV) Oxides. *J. Chem. Phys.* **2013**, *139* (24), 244704. https://doi.org/10.1063/1.4846135.

179. Nelin, C. J.; Bagus, P. S.; Ilton, E. S. Theoretical Analysis of the U L3-Edge NEXAFS in U Oxides. *RSC Adv.* **2014**, *4* (14), 7148–7153. https://doi.org/10.1039/C3RA46738D.
180. Bagus, P. S.; Ilton, E. S. Theory for the XPS of Actinides. *Top. Catal.* **2013**, *56* (12), 1121–1128. https://doi.org/10.1007/s11244-013-0078-2.
181. Stamberga, D.; Healy, M. R.; Bryantsev, V. S.; Albisser, C.; Karslyan, Y.; Reinhart, B.; Paulenova, A.; Foster, M.; Popovs, I.; Lyon, K.; Moyer, B. A.; Jansone-Popova, S. Structure Activity Relationship Approach toward the Improved Separation of Rare-Earth Elements Using Diglycolamides. *Inorg. Chem.* **2020**, *59* (23), 17620–17630.
182. Chapleski, R. C.; Chowdhury, A. U.; Wanhala, A. K.; Bocharova, V.; Roy, S.; Keller, P. C.; Everly, D.; Jansone-Popova, S.; Kisliuk, A.; Sacci, R. L.; Stack, A. G.; Anderson, C. G.; Doughty, B.; Bryantsev, V. S. A Molecular-Scale Approach to Rare-Earth Beneficiation: Thinking Small to Avoid Large Losses. *iScience* **2020**, *23* (9), 101435.
183. Roy, S.; Wu, L.; Srinivasan, S. G.; Stack, A. G.; Navrotsky, A.; Bryantsev, V. S. Hydration Structure and Water Exchange Kinetics at Xenotime–Water Interfaces: Implications for Rare Earth Minerals Separation. *Phys. Chem. Chem. Phys.* **2020**, *22* (15), 7719–7727. https://doi.org/10.1039/D0CP00087F.
184. Sutton, J. E.; Roy, S.; Chowdhury, A. U.; Wu, L.; Wanhala, A. K.; De Silva, N.; Jansone-Popova, S.; Hay, B. P.; Cheshire, M. C.; Windus, T. L.; Stack, A. G.; Navrotsky, A.; Moyer, B. A.; Doughty, B.; Bryantsev, V. Molecular Recognition at Mineral Interfaces: Implications for the Beneficiation of Rare Earth Ores. *ACS Appl. Mater. Interfaces* **2020**, *12* (14), 16327–16341.
185. Wanhala, A. K.; Doughty, B.; Bryantsev, V. S.; Wu, L.; Mahurin, S. M.; Jansone-Popova, S.; Cheshire, M. C.; Navrotsky, A.; Stack, A. G. Adsorption Mechanism of Alkyl Hydroxamic Acid onto Bastnäsite: Fundamental Steps toward Rational Collector Design for Rare Earth Elements. *J. Colloid Interface Sci.* **2019**, *553*, 210–219. https://doi.org/10.1016/j.jcis.2019.06.025.
186. Healy, M. R.; Ivanov, A. S.; Karslyan, Y.; Bryantsev, V. S.; Moyer, B. A.; Jansone-Popova, S. Efficient Separation of Light Lanthanides(III) by Using Bis-Lactam Phenanthroline Ligands. *Chem.—Eur. J.* **2019**, *25* (25), 6326–6331. https://doi.org/10.1002/chem.201806443.
187. Stack, A. G.; Stubbs, J. E.; Srinivasan, S. G.; Roy, S.; Bryantsev, V. S.; Custalcean, R.; Gordon, A. D.; Hexel, C. R. Mineral–Water Interface Structure of Xenotime (YPO_4) {100}. *J. Phys. Chem. C* **2018**, *122* (35)
188. Baldwin, A. G.; Ivanov, A. S.; Williams, N. J.; Ellis, R. J.; Moyer, B. A.; Bryantsev, V. S.; Shafer, J. C. Outer-Sphere Water Clusters Tune the Lanthanide Selectivity of Diglycolamides. *ACS Cent. Sci.* **2018**, *4* (6), 739–747.
189. Vo, M. N.; Bryantsev, V. S.; Johnson, J. K.; Keith, J. A. Quantum Chemistry Benchmarking of Binding and Selectivity for Lanthanide Extractants. *Int. J. Quantum Chem.* **2018**, *118* (7), e25516. https://doi.org/10.1002/qua.25516.
190. Srinivasan, S. G.; Shivaramaiah, R.; Kent, P. C.; Stack, A.; Navrotsky, A.; Riman, R. E.; Anderko, A.; Bryantsev, V. S. Crystal Structures, Surface Stability, and Water Adsorption Energies of La-Bastnäsite via Density Functional Theory and Experimental Studies. *J. Phys. Chem. C* **2016**, *120* (30), 16767–16781.
191. Williams, N. J.; Dehaudt, J.; Bryantsev, V. S.; Luo, H.; Abney, C. W.; Dai, S. Selective Separation of Americium from Europium Using 2,9-Bis(Triazine)-1,10-Phenanthrolines in Ionic Liquids: A New Twist on an Old Story. *Chem. Commun.* **2017**, *53* (18), 2744–2747. https://doi.org/10.1039/C6CC09823A.

192. Ellis, R. J.; Brigham, D. M.; Delmau, L.; Ivanov, A. S.; Williams, N. J.; Nguyen Vo, M.; Reinhart, B.; Moyer, B. A.; Bryantsev, V. S. "Straining" to Separate the Rare Earths: How the Lanthanide Contraction Impacts Chelation by Diglycolamide Ligands. *Inorg. Chem.* **2017**, *56* (3), 1152–1160.
193. Srinivasan, S. G.; Shivaramaiah, R.; Kent, P. R. C.; Stack, A. G.; Riman, R.; Anderko, A.; Navrotsky, A.; Bryantsev, V. S. A Comparative Study of Surface Energies and Water Adsorption on Ce-Bastnäsite, La-Bastnäsite, and Calcite via Density Functional Theory and Water Adsorption Calorimetry. *Phys. Chem. Chem. Phys.* **2017**, *19* (11), 7820–7832. https://doi.org/10.1039/C7CP00811B.
194. Ivanov, A. S.; Bryantsev, V. S. A Computational Approach to Predicting Ligand Selectivity for the Size-Based Separation of Trivalent Lanthanides. *Eur. J. Inorg. Chem.* **2016**, *2016* (21), 3474–3479. https://doi.org/10.1002/ejic.201600319.
195. Casey, K. C.; Brown, A. M.; Robinson, J. R. Computer-Aided Molecular Design of Bis-Phosphine Oxide Lanthanide Extractants. *Inorg. Chem. Front.* **2021**, *8*.
196. Peterson, C. C.; Penchoff, D. A.; Auxier, J. D.; Hall, H. L. Establishing Cost-Effective Computational Models for the Prediction of Lanthanoid Binding in $[Ln(NO_3)]^{2+}$ (with Ln = La to Lu). *ACS Omega* **2019**, *4* (1), 1375–1385. https://doi.org/10.1021/acsomega.8b02403.
197. Penchoff, D. A.; Peterson, C. C.; Quint, M. S.; Auxier, J. D.; Schweitzer, G. K.; Jenkins, D. M.; Harrison, R. J.; Hall, H. L. Structural Characteristics, Population Analysis, and Binding Energies of $[An(NO_3)]^{2+}$ [with An = Ac - Lr]. *ACS Omega* **2018**, *3*, 14127–14143. http://dx.doi.org/10.1021/acsomega.8b01800.
198. Peterson, C. C.; Penchoff, D. A. Computational Protocol for Binding Selectivity of Lanthanide and Actinide Systems. In *258th ACS National Meeting and Exposition*; American Chemical Society: San Diego, CA, 2019; p NUCL-0023.
199. Peterson, C. C.; Penchoff, D. A.; Hall, H. L.; Harrison, R. J. Understanding Selectivity of Lanthanide and Actinide Compounds by Computational Techniques. *257th ACS National Meeting & Exposition*; American Chemical Society: Orlando, FL, 2019; p NUCL-0004.
200. Penchoff, D. A.; Schweitzer, G. K.; Harrison, R. J. Theoretical Studies of Binding Preferences in Separations of Lanthanides and Actinides. *Abstracts of Papers, 249th ACS National Meeting & Exposition*; American Chemical Society: Denver, CO, 2015; p INOR-850.
201. Penchoff, D. A.; Peterson, C. C.; Camden, J. P.; Bradshaw, J. A.; Auxier, J. D.; Schweitzer, G. K.; Jenkins, D. M.; Harrison, R. J.; Hall, H. L. Structural Analysis of the Complexation of Uranyl, Neptunyl, Plutonyl, and Americyl with Cyclic Imide Dioximes. *ACS Omega* **2018**, *3* (10), 13984–13993. https://doi.org/10.1021/acsomega.8b02068.
202. Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58* (8), 1200–1211. https://doi.org/10.1139/p80-159.
203. Becke, A. D. A New Mixing of Hartree–Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372–1377.
204. Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158–6170. https://doi.org/10.1063/1.478522.

205. Sun, J.; Ruzsinszky, A.; Perdew, J. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115* (3), 1–6. https://doi.org/10.1103/PhysRevLett.115.036402.
206. Cao, X.; Dolg, M. Segmented Contraction Scheme for Small-Core Lanthanide Pseudopotential Basis Sets. *J. Mol. Struct. Theochem* **2002**, *581*, 139–147.
207. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions. *J. Chem. Phys.* **1980**, *72* (1), 650–654. https://doi.org/10.1063/1.438955.
208. Dunning, T. H. J. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023. https://doi.org/10.1063/1.456153.
209. Watts, J. D.; Gauss, J.; Bartlett, R. J. Coupled-cluster Methods with Noniterative Triple Excitations for Restricted Open-shell Hartree–Fock and Other General Single Determinant Reference Functions. Energies and Analytical Gradients. *J. Chem. Phys.* **1993**, *98*, 8718–8733.
210. Knowles, P. J. Coupled Cluster Theory for High Spin, Open Shell Reference Wave Functions. *J. Chem. Phys.* **1993**, *99*, 5219–5227.
211. Hampel, C.; Peterson, K.; Werner, H.-J. A Comparison of the Efficiency and Accuracy of the Quadratic Configuration Interaction (QCISD), Coupled Cluster (CCSD), and Brueckner Coupled Cluster (BCCD) Methods. *Chem. Phys. Lett.* **1992**, *190*, 1.
212. Werner, H.-J. Third-Order Multireference Perturbation Theory The CASPT3 Method. *Mol. Phys.* **1996**, *89*, 645–661.
213. Lu, Q.; Peterson, K. A. Correlation Consistent Basis Sets for Lanthanides: The Atoms La–Lu. *J. Chem. Phys.* **2016**, *145*, 054111.
214. Feng, R.; Peterson, K. A. Correlation Consistent Basis Sets for Actinides. II. The Atoms Ac and Np - Lr. *J. Chem. Phys.* **2017**, *147*, 084108.
215. Karton, A.; Martin, J. M. L. Comment on: "Estimating the Hartree–Fock Limit from Finite Basis Set Calculations" [Jensen F (2005) Theor Chem Acc 113:267]. *Theor. Chem. Acc.* **2006**, *115*, 330–333.
216. Feller, D.; Peterson, K. A.; Grant Hill, J. On the Effectiveness of CCSD(T) Complete Basis Set Extrapolations for Atomization Energies. *J. Chem. Phys.* **2011**, *135*, 044102/1-18. https://doi.org/10.1063/1.3613639.
217. Gomes, A. S. P.; Visscher, L.; Dyall, K. G. Relativistic Double-Zeta, Triple-Zeta, and Quadruple-Zeta Basis Sets for the Lanthanides La–Lu. *Theor. Chem. Acc.* **2010**, *127*, 369–382.
218. Dyall, K. G. Relativistic Double-Zeta, Triple-Zeta, and Quadruple-Zeta Basis Sets for the Actinides Ac–Lr. *Theor. Chem. Acc.* **2007**, *117*, 491–500.
219. Dyall, K. G. An Exact Separation of the Spin-Free and Spin-Dependent Terms of the Dirac–Coulomb–Breit Hamiltonian. *J. Chem. Phys.* **1994**, *100* (3), 2118.
220. Lu, Q.; Peterson, K. A. Correlation Consistent Basis Sets for Lanthanides: The Atoms La-Lu. *J. Chem. Phys.* **2016**, *145* (5), 054111. https://doi.org/10.1063/1.4959280.
221. Peterson, K. A. Correlation Consistent Basis Sets for Actinides. I. The Th and U Atoms. *J. Chem. Phys.* **2015**, *142* (7), 074105(1-14). https://doi.org/10.1063/1.4907596.

222. Dyall, K. G. Relativistic Double-Zeta, Triple-Zeta, and Quadruple-Zeta Basis Sets for the Actinides Ac-Lr. *Theor. Chem. Acc.* **2007**, *90*, 491–500. https://doi.org/10.1007/s00214-009-0725-7.
223. Gomes, A. S. P.; Dyall, K. G.; Visscher, L. Relativistic Double-Zeta, Triple-Zeta, and Quadruple-Zeta Basis Sets for the Lanthanides La–Lu. *Theor. Chem. Acc.* **2010**, *127*, 369–381. https://doi.org/10.1007/s00214-009-0725-7.
224. Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; De Jong, W. A. NWChem: A Comprehensive and Scalable Open-Source Solution for Large Scale Molecular Simulations. *Comput. Phys. Commun.* **2010**, *181* (9), 1477–1489. https://doi.org/10.1016/j.cpc.2010.04.018.
225. Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Karafiloglou, P.; Landis, C. R.; Weinhold, F. *NBO 7.0*; Theoretical Chemistry Institute, University of Wisconsin: Madison, WI, 2018.
226. Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural Population Analysis. *J. Chem. Phys.* **1985**, *83* (2), 735–746. https://doi.org/10.1063/1.449486.
227. Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General Purpose Quantum Chemistry Package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242–253.
228. Aquilante, F.; Autschbach, J.; Carlson, R. K.; Chibotaru, L. F.; Delcey, M. G.; Vico, L. D.; Galván, I. F.; Ferré, N.; Frutos, L. M.; Gagliardi, L.; Garavelli, M.; Giussani, A.; Hoyer, C. E.; Manni, G. L.; Lischka, H.; Ma, D.; Malmqvist, P. Å.; Müller, T.; Nenov, A.; Olivucci, M.; Pedersen, T. B.; Peng, D.; Plasser, F.; Pritchard, B.; Reiher, M.; Rivalta, I.; Schapiro, I.; Segarra-Martí, J.; Stenrup, M.; Truhlar, D. G.; Ungur, L.; Valentini, A.; Vancoillie, S.; Veryazov, V.; Vysotskiy, V. P.; Weingart, O.; Zapata, F.; Lindh, R. MOLCAS 8: New Capabilities for Multiconfigurational Quantum Chemical Calculations across the Periodic Table. *J. Comput. Chem.* **2016**, *37*, 506–541.
229. Saue, T.; Visscher, L.; Jensen, H. J. Aa.; Bast, R. *DIRAC, a Relativistic Ab Initio Electronic Structure Program*, Release DIRAC18; 2018. https://doi.org/10.5281/zenodo.2253986.
230. Klamt, A.; Schüürmann, G. COSMO: A New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805. https://doi.org/10.1039/P29930000799.
231. Feller, D. The Role of Databases in Support of Computational Chemistry Calculations. *J. Comput. Chem.* **1996**, *17* (13), 1571–1586.
232. Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorthi, V.; Chase, J.; Li, J.; Windus, T. L. Basis Set Exchange: A Community Database for Computational Sciences. *J. Chem. Inf. Model.* **2007**, *47* (3), 1045–1052. https://doi.org/10.1021/ci600510j.
233. Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibson, T. D.; Windus, T. L. A New Basis Set Exchange: An Open, Up-to-Date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* **2019**, *59* (11), 4814–4820. https://doi.org/doi:10.1021/acs.jcim.9b00725.