RESEARCH ARTICLE SUMMARY

DFT METHODS

Reproducibility in density functional theory calculations of solids

Kurt Lejaeghere,* Gustav Bihlmayer, Torbjörn Björkman, Peter Blaha, Stefan Blügel, Volker Blum, Damien Caliste, Ivano E. Castelli, Stewart J. Clark, Andrea Dal Corso, Stefano de Gironcoli, Thierry Deutsch, John Kay Dewhurst, Igor Di Marco, Claudia Draxl, Marcin Dułak, Olle Eriksson, José A. Flores-Livas, Kevin F. Garrity, Luigi Genovese, Paolo Giannozzi, Matteo Giantomassi, Stefan Goedecker, Xavier Gonze, Oscar Grånäs, E. K. U. Gross, Andris Gulans, François Gygi, D. R. Hamann, Phil J. Hasnip, N. A. W. Holzwarth, Diana Iuşan, Dominik B. Jochym, François Jollet, Daniel Jones, Georg Kresse, Klaus Koepernik, Emine Küçükbenli, Yaroslav O. Kvashnin, Inka L. M. Locht, Sven Lubeck, Martijn Marsman, Nicola Marzari, Ulrike Nitzsche, Lars Nordström, Taisuke Ozaki, Lorenzo Paulatto, Chris J. Pickard, Ward Poelmans, Matt I. J. Probert, Keith Refson, Manuel Richter, Gian-Marco Rignanese, Santanu Saha, Matthias Scheffler, Martin Schlipf, Karlheinz Schwarz, Sangeeta Sharma, Francesca Tavazza, Patrik Thunström, Alexandre Tkatchenko, Marc Torrent, David Vanderbilt, Michiel J. van Setten, Veronique Van Speybroeck, John M. Wills, Jonathan R. Yates, Guo-Xu Zhang, Stefaan Cottenier*

INTRODUCTION: The reproducibility of results is one of the underlying principles of science. An observation can only be accepted by the scientific community when it can be confirmed by independent studies. However, reproducibility does not come easily. Recent works have painfully exposed cases where previous conclusions were not upheld. The scrutiny of the scientific community has also turned to research involving computer programs, finding that reproducibility depends more strongly on implementation than commonly thought. These problems are especially relevant for property predictions of crystals and molecules, which hinge on precise computer implementations of the governing equation of quantum physics.

RATIONALE: This work focuses on density functional theory (DFT), a particularly popular quantum method for both academic and industrial applications. More than 15,000 DFT papers are published each year, and DFT is now increasingly used in an automated fashion to build large databases or apply multiscale techniques with limited human supervision. Therefore, the reproducibility of DFT results underlies the scientific credibility of a substantial fraction of current work in the natural and engineering sciences. A plethora of DFT computer codes are available, many of them differing considerably in their details of implementation, and each yielding a certain "precision" relative to other codes. How is one to decide for more than a few simple cases which code predicts the correct result, and which does not? We devised a procedure to assess the precision of DFT methods and used this to demonstrate reproducibility among many of the most widely used DFT codes. The essential part of this assessment is a pairwise comparison of a wide range of methods with respect to their predictions of the equations of state of the elemental crystals. This effort required the combined expertise of a large group of code developers and expert users.

RESULTS: We calculated equation-of-state data for four classes of DFT implementations, totaling 40 methods. Most codes agree very well, with pairwise differences that are comparable to those between different high-precision experiments. Even in the case of pseudization approaches, which largely depend on the atomic potentials used, a similar precision can be obtained as when using the full potential. The remaining deviations are due to subtle effects, such as specific numerical implementations or the treatment of relativistic terms.

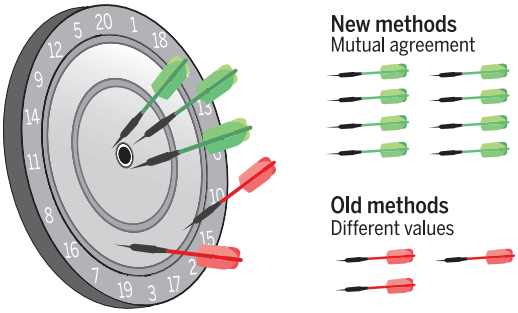
ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.aad3000

CONCLUSION: Our work demonstrates that the precision of DFT implementations can be determined, even in the absence of one absolute reference code. Although this was not the case 5 to 10 years ago, most of the commonly used codes and methods are now found to predict essentially identical results. The established precision of DFT codes not only ensures the reproducibility of DFT predictions but also puts several past and future developments on a firmer footing. Any newly developed methodology can now be tested against the benchmark to verify whether it reaches the same level of precision. New DFT applications can be shown to have used a sufficiently precise method. Moreover, high-precision DFT calculations are essential for developing improvements to DFT methodology, such as new density functionals, which may further increase the predictive power of the simulations. ■

The list of author affiliations is available in the full article online.
*Corresponding author. E-mail: kurt.lejaeghere@ugent.be (K.L.); stefaan.cottenier@ugent.be (S.C.)
Cite this article as K. Lejaeghere *et al.*, *Science* **351**, aad3000 (2016). DOI: 10.1126/science.aad3000

Recent DFT methods yield reproducible results. Whereas older DFT implementations predict different values (red darts), codes have now evolved to mutual agreement (green darts). The scoreboard illustrates the good pairwise agreement of four classes of DFT implementations (horizontal direction) with all-electron results (vertical direction). Each number reflects the average difference between the equations of state for a given pair of methods, with the green-to-red color scheme showing the range from the best to the poorest agreement.



Scorecard

	0.3	0.3	0.6	1.0	0.9	0.3	1.5	0.6	0.9	0.4	1.0	0.4	0.4	6.3	13.5	1.1	2.1	0.7	1.4
0.3		0.1	0.5	0.9	0.8	0.2	1.5	0.6	0.8	0.4	1.0	0.5	0.3	6.3	13.4	1.1	2.1	0.7	1.4
0.3	0.1		0.5	0.9	0.8	0.2	1.5	0.6	0.8	0.4	0.9	0.5	0.3	6.3	13.4	1.1	2.1	0.7	1.4
0.6	0.5	0.5		0.8	0.6	0.4	1.5	0.6	0.8	0.6	1.0	0.7	0.5	6.3	13.2	1.0	1.9	0.6	1.3
1.0	0.9	0.9	0.8		0.9	0.9	1.8	0.9	1.3	1.0	1.4	1.0	0.9	6.4	13.0	1.2	1.8	1.0	1.6
0.9	0.8	0.8	0.6	0.9		0.8	1.7	0.7	1.1	0.8	1.3	1.0	0.8	6.5	13.2	1.1	1.8	0.8	1.5
0.3	0.2	0.2	0.4	0.9	0.8		1.5	0.5	0.8	0.3	1.0	0.5	0.3	6.2	13.4	1.0	2.0	0.6	1.4

RESEARCH ARTICLE

DFT METHODS

Reproducibility in density functional theory calculations of solids

Kurt Lejaeghere,[1]* Gustav Bihlmayer,[2] Torbjörn Björkman,[3,4] Peter Blaha,[5] Stefan Blügel,[2] Volker Blum,[6] Damien Caliste,[7,8] Ivano E. Castelli,[9] Stewart J. Clark,[10] Andrea Dal Corso,[11] Stefano de Gironcoli,[11] Thierry Deutsch,[7,8] John Kay Dewhurst,[12] Igor Di Marco,[13] Claudia Draxl,[14,15] Marcin Dułak,[16] Olle Eriksson,[13] José A. Flores-Livas,[12] Kevin F. Garrity,[17] Luigi Genovese,[7,8] Paolo Giannozzi,[18] Matteo Giantomassi,[19] Stefan Goedecker,[20] Xavier Gonze,[19] Oscar Grånäs,[13,21] E. K. U. Gross,[12] Andris Gulans,[14,15] François Gygi,[22] D. R. Hamann,[23,24] Phil J. Hasnip,[25] N. A. W. Holzwarth,[26] Diana Iuşan,[13] Dominik B. Jochym,[27] François Jollet,[28] Daniel Jones,[29] Georg Kresse,[30] Klaus Koepernik,[31,32] Emine Küçükbenli,[9,11] Yaroslav O. Kvashnin,[13] Inka L. M. Locht,[13,33] Sven Lubeck,[14] Martijn Marsman,[30] Nicola Marzari,[9] Ulrike Nitzsche,[31] Lars Nordström,[13] Taisuke Ozaki,[34] Lorenzo Paulatto,[35] Chris J. Pickard,[36] Ward Poelmans,[1,37] Matt I. J. Probert,[25] Keith Refson,[38,39] Manuel Richter,[31,32] Gian-Marco Rignanese,[19] Santanu Saha,[20] Matthias Scheffler,[15,40] Martin Schlipf,[22] Karlheinz Schwarz,[5] Sangeeta Sharma,[12] Francesca Tavazza,[17] Patrik Thunström,[41] Alexandre Tkatchenko,[15,42] Marc Torrent,[28] David Vanderbilt,[23] Michiel J. van Setten,[19] Veronique Van Speybroeck,[1] John M. Wills,[43] Jonathan R. Yates,[29] Guo-Xu Zhang,[44] Stefaan Cottenier[1,45]*

The widespread popularity of density functional theory has given rise to an extensive range of dedicated codes for predicting molecular and crystalline properties. However, each code implements the formalism in a different way, raising questions about the reproducibility of such predictions. We report the results of a community-wide effort that compared 15 solid-state codes, using 40 different potentials or basis set types, to assess the quality of the Perdew-Burke-Ernzerhof equations of state for 71 elemental crystals. We conclude that predictions from recent codes and pseudopotentials agree very well, with pairwise differences that are comparable to those between different high-precision experiments. Older methods, however, have less precise agreement. Our benchmark provides a framework for users and developers to document the precision of new applications and methodological improvements.

Scientific results are expected to be reproducible. When the same study is repeated independently, it should reach the same conclusions. Nevertheless, some recent articles have shown that reproducibility is not self-evident. A widely resounding *Science* article (*1*), for example, demonstrated a lack of reproducibility among published psychology experiments. Although the hard sciences are believed to perform better in this respect, concerns about reproducibility have emerged in these fields as well (*2*–*4*). The issue is of particular interest when computer programs are involved. Undocumented approximations or undetected bugs can lead to wrong conclusions (*5*). In areas where academic codes compete with commercial software, the unavailability of source code can hinder assessment of the relevance of conclusions (*6*, *7*).

Density functional theory (DFT) calculations (*8*, *9*) are a prominent example of an area that depends on the development and appropriate use of complex software. With rigorous foundations in the quantum theory of matter, DFT describes the structure and properties of molecules and solids at the atomic scale. Over the years, many academic groups have developed implementations of DFT in computer codes, and several of these have been adopted by large user communities. Commercial alternatives are entering this area as well. At present, more than 15,000 papers are published each year that make use of DFT codes (*10*), with applications varying from metallurgy to drug design. Moreover, DFT calculations are used nowadays to build large databases (*11*, *12*) and in multiscale calculations in which they serve as one part of the tool chain (*13*, *14*). The precision of DFT codes thus underlies the scientific credibility and reproducibility of a substantial fraction of current work in the natural and engineering sciences, and therefore it has implications that reach far beyond the traditional electronic-structure research community.

The main idea of DFT is to solve the intractable many-particle Schrödinger equation by replacing the complete electron wave function with the much simpler ground-state electron density as the fundamental variable. Although this reformulation is in principle exact, it is not fully known how the interaction between individual electrons should be transformed. As a result, the specific form of the unknown part of the interaction energy, the exchange-correlation functional, has been the focus of many investigations, leading to a plethora of available functionals in both solid-state physics (*15*–*19*) and quantum chemistry (*15*, *20*–*23*).

Once a particular exchange-correlation functional has been chosen, the mathematical problem is completely specified as a set of Kohn-Sham equations, whose solution yields orbitals and energies from which the total electronic energy can be evaluated. A variety of such numerical solution schemes have been implemented in different computer codes. Comparisons of their performance are much less frequent or extensive than those of exchange-correlation functionals, however (*21*, *24*–*29*). One might reasonably expect that because they solve the same equations, they all produce similar answers for a given crystal structure, but a glance at the literature shows that this assumption is by no means always true. Figure 1 demonstrates that even for a well-studied material such as silicon, deviations between predictions from different codes (the "precision") are of the same order of magnitude as the deviation from the 0 K experimental value (the "accuracy") (*26*, *30*). Because all of the codes shown in Fig. 1 treat silicon at the same level of theory, using the same exchange-correlation functional, they yield the same accuracy by definition. However, the particular predictions vary from one code to another because of approximations that are unrelated to the exchange-correlation functional. These approximations decrease the computational load but limit the precision.

What level of precision can we now achieve? Discussion of precision-related issues is uncommon in reports of solid-state DFT studies. The reproducibility of predictions is sometimes checked by cross-validation with other codes (*21*, *24*–*28*), but we are not aware of any systematic assessments of precision (also called "verification"), even though such studies would reinforce confidence in practical DFT calculations.

As a group of 69 code developers and expert users, we determined the error bar associated with energy-versus-volume [$E(V)$] predictions of elemental solids by running the same benchmark protocol with various DFT codes. Parameters of these equations of state (EOS), such as the lattice parameter or the bulk modulus, are commonly used for accuracy assessments (*15*–*19*). By considering elemental solids, we have established a broad and comprehensive test for precision. Elemental solids have a wide range of chemical environments and constitute a reasonable first approximation to sampling the broad compositional space of multicomponent systems. Our effort has resulted in 18,602 DFT calculations, which we aimed to execute with a rigorously determined precision. This exercise might seem simple, but each code tackles the Kohn-Sham equations and subsequent energy evaluation in its own way, requiring different approaches to

deal with difficulties in different parts of the computational procedure.

Kohn-Sham solution techniques

The Kohn-Sham equations describe a many-electron system in terms of a density built from single-particle wave functions. By expressing these wave functions as a linear combination of predefined basis functions, the Kohn-Sham equations reduce to a matrix equation, which can in principle be solved exactly. Their solutions should yield identical results, irrespective of the form of the basis functions, provided that the basis set is complete. However, achieving technical convergence of the complete Kohn-Sham problem is not feasible in practice. Consider silicon, whose electronic structure is schematically illustrated in Fig. 2. The Aufbau principle requires first populating the lowest energy level, which is the 1*s* band. This is much lower in energy than the valence and conduction bands, and the localization of the orbitals close to the nuclei demands high spatial resolution. These core electrons do not contribute directly to chemical bonding, so they can be separated out and represented using a different basis that is better suited to describe localized atomic-like states. Core orbitals may be either computed in an isolated atom environment, with their effect on valence transferred unaltered to the crystal, or relaxed self-consistently in the full crystal field. They can moreover be treated using a relativistic Hamiltonian, which is essential for core electrons in heavy atoms. Different relativistic schemes may lead to differences in the predicted $E(V)$ curves.

To stitch together a complete solution, the wave functions of the semi-core and valence electrons (2*s* 2*p* and 3*s* 3*p*, respectively, in the case of silicon) must be constructed to include the effect of orthogonality to the core electrons. This central problem can be solved in a number of different ways, depending on the choice of numerical method. For methods that are based on plane-wave expansions or uniform real-space grids, the oscillatory behavior near the nucleus cannot be accurately represented because of the limited spatial resolution. The need for unmanageably large basis sets can be mitigated by adding a carefully designed repulsive part to the Kohn-Sham potential, a so-called pseudopotential. This pseudopotential affects only a small region around the nuclei (gray zones in Fig. 2) and may conserve the core-region charge [norm-conserving pseudopotentials (*31*, *32*)], giving rise to an analytically straightforward formalism, or it may break norm conservation by including a compensating augmentation charge [ultrasoft pseudopotentials (*33*)], allowing for smoother wave functions and hence smaller basis sets. Alternatively, the projector-augmented wave (PAW) approach defines an explicit transformation between the all-electron and pseudopotential wave functions by means of additional partial-wave basis functions (*34*, *35*). This allows PAW codes to obtain good precision for small numbers of plane waves or large grid spacings, but choosing suitable partial-wave projectors is not trivial. Here we refer to both pseudopotential and PAW methods as pseudization approaches. In contrast to these approaches, all-electron methods explicitly construct basis functions that are restricted to a specific energy range [linearized augmented plane wave (LAPW) (*36*–*39*) and linear muffin-tin orbital (LMTO) (*40*) methods] or treat core and valence states on equal footing (e.g., by using numerical atomic-like orbitals) (*41*, *42*). Avoiding pseudization enables better precision but inevitably increases the computation time. In these codes, the ambiguity in solving the Kohn-Sham problem shifts from the choice of the pseudization scheme to the choice of the basis functions. This choice leads to a variety of methods as well, which, despite solving the same Kohn-Sham equations, differ in many other details. Because each all-electron or pseudization method has its own fundamental advantages, it is highly desirable to achieve high precision for all of them.

The Δ matrix

The case study for silicon (Fig. 1) demonstrates that different approaches to the potential or basis functions may lead to noticeably different predictions, even for straightforward properties such as the lattice parameter. There is no absolute reference against which to compare these methods; each approach has its own intricacies and approximations. To determine whether the same results can be obtained irrespectively of the code or (pseudo)potential, we instead present a large-scale pairwise code comparison using the Δ gauge. This criterion was formulated by Lejaeghere *et al.* (*26*) to quantify differences between DFT-predicted $E(V)$ profiles in an unequivocal way. That study proposed a benchmark set of 71 elemental crystals and defined, for every element i, the quantity Δ_i as the root-mean-square difference between the EOS of methods a and b over a ±6% interval around the equilibrium volume $V_{0,i}$. The calculated EOS are lined up with respect to their minimum energy and compared in an interval that is symmetrical around the average equilibrium volume (Fig. 3).

$$\Delta_i(a,b) = \sqrt{\frac{\int_{0.94V_{0,i}}^{1.06V_{0,i}} \left(E_{b,i}(V) - E_{a,i}(V)\right)^2 dV}{0.12V_{0,i}}} \quad (1)$$

A comparison of Δ_i values allows the expression of EOS differences as a single number, and a small Δ_i automatically implies small deviations between equilibrium volumes, bulk moduli, or any other EOS-derived observables. The overall difference Δ between methods a and b is

[1]Center for Molecular Modeling, Ghent University, Technologiepark 903, BE-9052 Zwijnaarde, Belgium. [2]Peter Grünberg Institute and Institute for Advanced Simulation, Forschungszentrum Jülich and JARA (Jülich Aachen Research Alliance), D-52425 Jülich, Germany. [3]Department of Physics, Åbo Akademi, FI-20500 Turku, Finland. [4]Centre of Excellence in Computational Nanoscience (COMP) and Department of Applied Physics, Aalto University School of Science, Post Office Box 11100, FI-00076 Aalto, Finland. [5]Institute of Materials Chemistry, Vienna University of Technology, Getreidemarkt 9/165-TC, A-1060 Vienna, Austria. [6]Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, USA. [7]Université Grenoble Alpes, Institut Nanosciences et Cryogénie–Modeling and Material Exploration Department (INAC-MEM), Laboratoire de Simulation Atomistique (L_Sim), F-38042 Grenoble, France. [8]Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA), INAC-MEM, L_Sim, F-38054 Grenoble, France. [9]Theory and Simulation of Materials (THEOS) and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. [10]Department of Physics, University of Durham, Durham DH1 3LE, UK. [11]International School for Advanced Studies (SISSA) and DEMOCRITOS, Consiglio Nazionale delle Ricerche–Istituto Officina dei Materiali (CNR-IOM), Via Bonomea 265, I-34136 Trieste, Italy. [12]Max-Planck-Institut für Mikrostrukturphysik, Weinberg 2, D-06120 Halle, Germany. [13]Department of Physics and Astronomy, Division of Materials Theory, Uppsala University, Post Office Box 516, SE-75120 Uppsala, Sweden. [14]Institut für Physik and Integrative Research Institute for the Sciences (IRIS)–Adlershof, Humboldt-Universität zu Berlin, Zum Großen Windkanal 6, D-12489 Berlin, Germany. [15]Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany. [16]Center for Atomic-Scale Materials Design, Department of Physics, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark. [17]Material Measurement Laboratory, National Institute of Standards and Technology, 100 Bureau Drive, Stop 8553, Gaithersburg, MD 20899, USA. [18]Department of Mathematics, Computer Science, and Physics, University of Udine, Via delle Scienze 206, I-33100 Udine, Italy. [19]Institute of Condensed Matter and Nanosciences–Nanoscopic Physics (NAPS), Université Catholique de Louvain, Chemin des Étoiles 8, BE-1348 Louvain-la-Neuve, Belgium. [20]Institut für Physik, Universität Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland. [21]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA. [22]Department of Computer Science, University of California–Davis, Davis, CA 95616, USA. [23]Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854-8019, USA. [24]Mat-Sim Research, Post Office Box 742, Murray Hill, NJ 07974, USA. [25]Department of Physics, University of York, Heslington, York YO10 5DD, UK. [26]Department of Physics, Wake Forest University, Winston-Salem, NC 27109, USA. [27]Scientific Computing Department, Science and Technology Facilities Council, Rutherford Appleton Laboratory, Didcot OX11 0QX, UK. [28]CEA, DAM, DIF, F-91297 Arpajon, France. [29]Department of Materials, University of Oxford, 16 Parks Road, Oxford OX1 3PH, UK. [30]Faculty of Physics and Center for Computational Materials Science, University of Vienna, Sensengasse 8/12, A-1090 Vienna, Austria. [31]Leibniz-Institut für Festkörper- und Werkstoffforschung (IFW) Dresden, Post Office Box 270 116, D-01171 Dresden, Germany. [32]Dresden Center for Computational Materials Science (DCMS), Technische Universität Dresden, D-01069 Dresden, Germany. [33]Institute for Molecules and Materials, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, Netherlands. [34]Institute for Solid State Physics, The University of Tokyo, Kashiwa 277-8581, Japan. [35]Institut de Minéralogie, de Physique des Matériaux, et de Cosmochimie (IMPMC), Sorbonne Universités–Pierre and Marie Curie University Paris 06, Centre National de la Recherche Scientifique (CNRS) Unité Mixte de Recherche (UMR) 7590, Muséum National d'Histoire Naturelle, Institut de Recherche pour le Développement (IRD) Unité de Recherche 206, 4 Place Jussieu, F-75005 Paris, France. [36]Department of Materials Science and Metallurgy, University of Cambridge, 27 Charles Babbage Road, Cambridge CB3 0FS, UK. [37]High Performance Computing Unit, Ghent University, Krijgslaan 281 S9, BE-9000 Ghent, Belgium. [38]Department of Physics, Royal Holloway, University of London, Egham TW20 0EX, UK. [39]ISIS Facility, Science and Technology Facilities Council, Rutherford Appleton Laboratory, Didcot OX11 0QX, UK. [40]Department of Chemistry and Biochemistry and Materials Department, University of California–Santa Barbara, Santa Barbara, CA 93106-5050, USA. [41]Institute for Solid State Physics, Vienna University of Technology, A-1040 Vienna, Austria. [42]Physics and Materials Science Research Unit, University of Luxembourg, L-1511 Luxembourg. [43]Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. [44]Institute of Theoretical and Simulational Chemistry, School of Chemistry and Chemical Engineering, Harbin Institute of Technology, Harbin 150001, People's Republic of China. [45]Department of Materials Science and Engineering, Ghent University, Technologiepark 903, BE-9052 Zwijnaarde, Belgium.
***Corresponding author. E-mail: kurt.lejaeghere@ugent.be (K.L.); stefaan.cottenier@ugent.be (S.C.)**

obtained by averaging Δ_i over all 71 crystals in the benchmark set. Alternative definitions of Δ essentially render the same information (*27*, *28*). In this work, we applied the original Δ protocol to 40 DFT implementations of the Perdew-Burke-Ernzerhof (PBE) functional (*43*). Appropriate numerical settings were determined separately for each method, ensuring converged results. In all calculations, valence and semi-core electrons were treated on a scalar-relativistic level, because not all codes support spin-orbit coupling. This is not a limitation, because the aim is to compare codes with each other rather than to experiment. We do not elaborate here on speed and memory requirements, for which we refer to the documentation of the respective codes.

Figure 4 presents an overview of the most important Δ values, categorized by method: all-electron, PAW, ultrasoft pseudopotentials, and norm-conserving pseudopotentials. Approaches with a similar intrinsic precision are clustered together in this way. Both the full results and the most important numerical settings are included in tables S3 to S42. A complete specification would have to include code defaults and hard-coded values, so a reasonable compromise was chosen. A full specification could be realized by recent endeavors in full-output databases (*44*, *45*) or workflow scripting (*46*, *47*), but this capacity is not yet available for several of the codes used in this study. We have, however, tried to provide generation scripts for as many methods as possible (*48*), and we emphasize the need for such tools as an important future direction.

Comparing all-electron methods

Although the definition of Δ does not favor a particular reference, it is instructive to first examine the Δ values with respect to all-electron methods (Fig. 4). They generally come at a computationally higher cost, but all-electron approaches are often considered to be a standard for DFT calculations, because they implement the potential without pseudization. By comparing pseudopotential or PAW methods with all-electron codes, we can therefore get an idea of the error bar associated with each pseudization scheme. The Δ values between different all-electron methods reflect the remaining discrepancies, such as a different treatment of the scalar-relativistic terms or small differences in numerical methods.

To gain some insight into typical values of Δ, we should first establish which values for Δ can be qualified as "small," so that we know which results can be considered equivalent. A first indication comes from converting differences between high-precision measurements of EOS parameters into a Δ format. Comparing the high-quality experimental data of Holzapfel *et al.* for Cu, Ag, and Au (*49*) with those of Kittel (*50*) and Knittle (*51*), for example, shows a small difference Δ_{exp} of 1.0 meV per atom. Because the average all-electron Δ for these materials is only 0.8 meV per atom, this implies that the precision of many DFT codes outperforms experimental precision.

Secondly, we also considered the differences between codes in terms of commonly reported EOS parameters. The 1.0 meV-per-atom maximum Δ among all-electron codes (Fig. 4, top) corresponds to an average volume deviation of 0.14 Å^3 per atom (0.38%) or a median deviation of 0.05 Å^3 per atom (0.24%) over the entire 71-element test set. For the bulk modulus, the average deviation is 1.6 GPa (4.0%), and the median deviation 0.8 GPa (1.6%). Compared with the scatter on experimental values, which can amount to up to 35% for the bulk moduli of the rare earth metals [for instance, see (*52*)], these values are very small. The difference between EOS obtained by independent all-electron codes is hence smaller than the spread between independent experimental EOS. We conclude that, unless some elements deviate substantially from the overall trend, codes with a

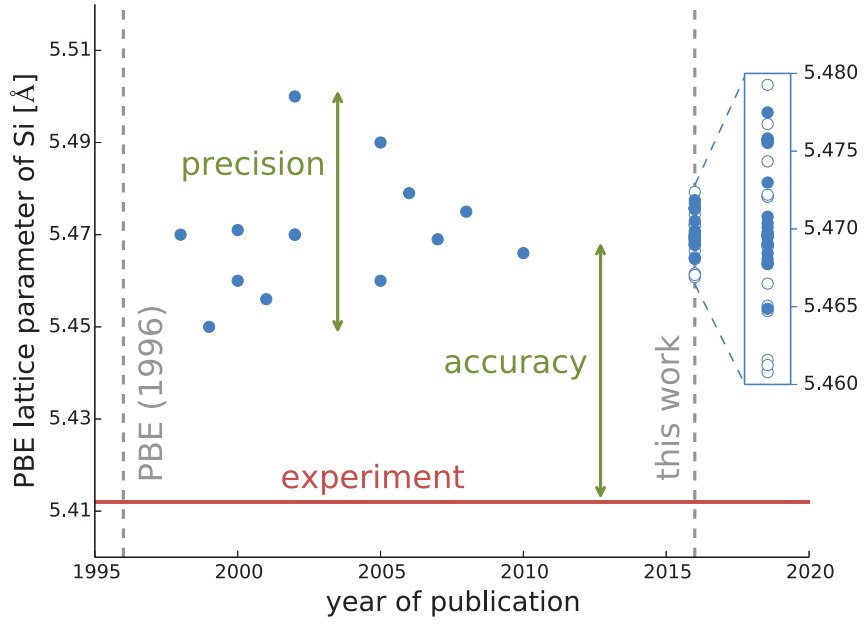


Fig. 1. Historical evolution of the predicted equilibrium lattice parameter for silicon. All data points represent calculations within the DFT-PBE framework. Values from literature (data points before 2016) (*15*, *16*, *18*, *56*–*62*, *63*–*65*) are compared with (i) predictions from the different codes used in this study (2016 data points, magnified in the inset; open circles indicate data produced by older methods or calculations with lower numerical settings) and (ii) the experimental value, extrapolated to 0 K and corrected for zero-point effects (red line) (*26*). The concepts of precision and accuracy are illustrated graphically.

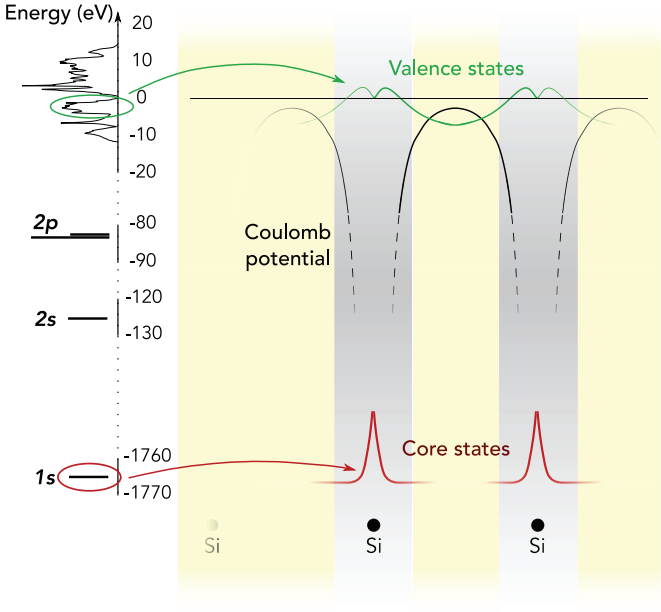


Fig. 2. Electronic states in solid silicon. The valence states are delocalized over the solid (green line), because the wave functions overlap from one atom to the next. The lowest-energy 1*s* state (red) is at an energy two orders of magnitude lower than the valence states and is strongly localized near the nucleus, with no overlap between the atoms. The gray regions around the atoms indicate approximately where the wave function, density, and potential are smoothed in pseudized methods.

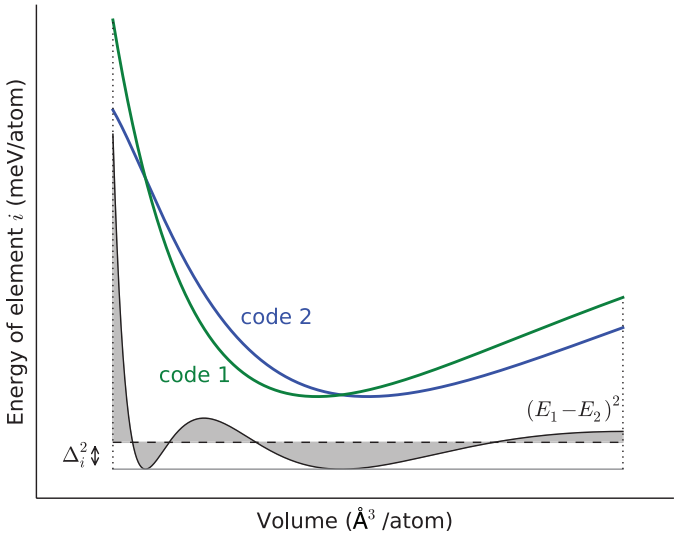


Fig. 3. Graphical representation of the Δ gauge. The black curve depicts the quadratic energy difference between two EOS [$(E_1 - E_2)^2$, where the subscripts correspond to the two codes shown], and Δ_i corresponds to the root-mean-square average. This is demonstrated by the shaded area, which is equally large above and below the Δ_i^2 line.

Table 1. Agreement between osmium crystal predictions at nearly identical settings. The top group includes Δ_i values for the osmium crystal (in millielectron volts per atom) produced by four APW+lo calculations that tried to mimic the same settings as well as possible. These settings are therefore different from the ones used for Fig. 4 and reported in tables S3, S4, S8, and S15. The bottom group includes the corresponding equilibrium volumes V_0, bulk moduli B_0, and bulk modulus derivatives B_1.

	Elk	FLEUR	WIEN2K	exciting
Δ(Elk)	–	0.03	0.02	0.20
Δ(FLEUR)	0.03	–	0.04	0.22
Δ(WIEN2K)	0.02	0.04	–	0.18
Δ(exciting)	0.20	0.22	0.18	–
V_0 (Å^3 per atom)	14.276	14.276	14.276	14.274
B_0 (GPa)	397.5	397.9	397.6	397.4
B_1 (unitless)	4.86	4.89	4.83	4.82

Table 2. Precision evolution of PAW and pseudopotential sets over time. The Δ values are expressed as an average over the all-electron methods (in millielectron volts per atom) and are listed chronologically per code. The corresponding code settings and the DFT-predicted EOS parameters are listed in tables S17, S19 to S26, S30, S31, and S33. The most recent potentials are the ones used to generate the data shown in Fig. 4.

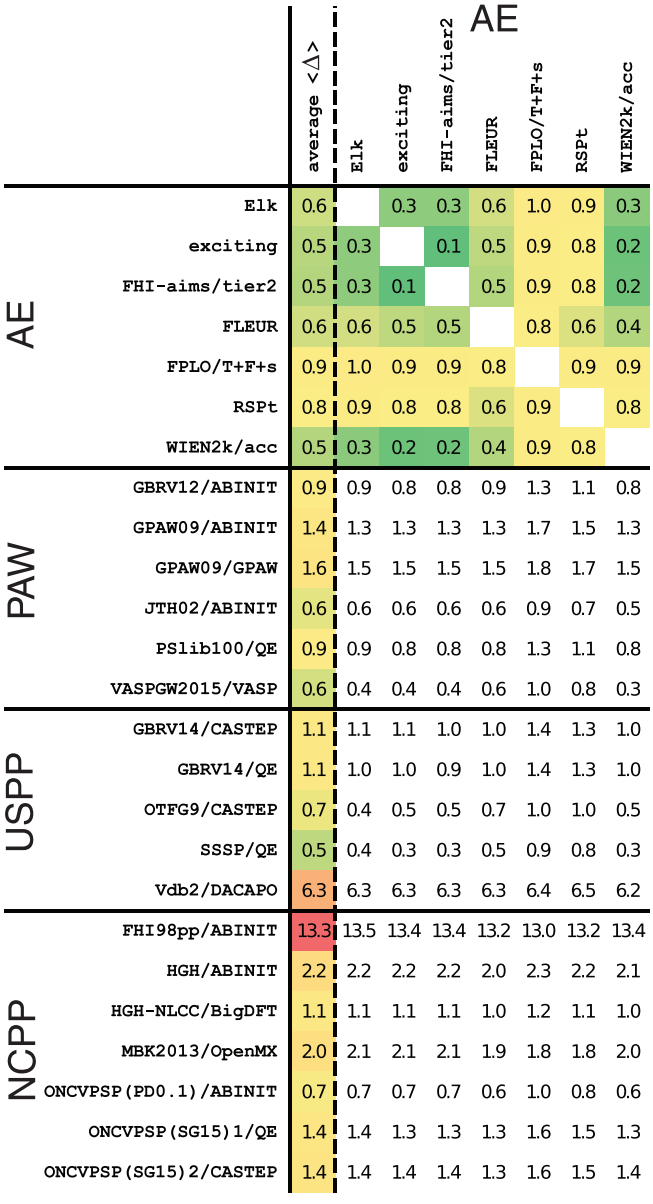
	Year	⟨Δ⟩versus AE
JTH01/ABINIT	2013	1.1
JTH02/ABINIT	2014	0.6
Vdb/CASTEP	1998	6.5
OTFG7/CASTEP	2013	2.6
OTFG9/CASTEP	2015	0.7
GPAW06/GPAW	2010	3.6
GPAW09/GPAW	2012	1.6
PSlib031/QE	2013	1.7
PSlib100/QE	2013	1.0
VASP2007/VASP	2007	2.0
VASP2012/VASP	2012	0.8
VASPGW2015/VASP	2015	0.6

mutual Δ of 1 or even 2 meV per atom can be deemed to yield indistinguishable EOS for all practical purposes.

The above-mentioned differences correspond to the best attainable precision for each all-electron code, using highly converged or "ultimate" computational settings. However, particular choices for these settings may still slightly change the Δ values. It is not always necessary to set such stringent requirements, because efficient codes are able to perform well with less-than-perfect settings. Nevertheless, the difference between default- and ultimate-precision EOS may sometimes reach a few millielectron volts per atom (table S2). To eliminate the effect of numerical convergence altogether, we used the osmium crystal to test whether it is possible to obtain exactly the same result with different codes. Rather than aiming for the best representation of the ideal PBE results, as in the rest of this work, the goal in this case was to choose input settings as consistently as possible (using the same basis functions, grids, and other parameters). Comparing four APW+lo (augmented plane waves plus local orbitals) calculations in this way yielded the results in Table 1. Whereas numerical noise in various subroutines gives rise to fluctuations of only 0.02 to 0.04 meV per atom, the larger deviation of ~0.2 meV per atom in comparisons involving the code known as "exciting" can partly be attributed to a different scalar-relativistic treatment of the valence electrons in this code. There is no single universal method to account for the relativistic change of the electron mass in the kinetic energy. The "exciting" code uses the infinite-order regular approximation (*53*), whereas the other three APW+lo codes use the Koelling-Harmon scheme (*54*). A third possibility is to use the atomic zero-order regular approximation, as was done in the FHI-aims code package (tables S5 to S7) (*42*, *55*).

Comparing (pseudo)potential libraries

In comparison with all-electron codes, pseudization approaches are generally faster, because fewer states are considered, and explicit construction and diagonalization of the Hamiltonian matrix is avoided. Among these, PAW and ultrasoft pseudopotentials require fewer basis functions than the norm-conserving variety, but advanced features such as linear response theory or hybrid functionals sometimes may not be available because of the increased complexity of the implementation. However, pseudization approaches all perform very well in terms of precision when compared with all-electron results (Fig. 4). For EOS, the precision of current potentials is able to compete with that of all-electron methods, yielding Δ values of about 1 meV per atom, with a low approaching 0.3 meV per atom. This has not always been the case. As suggested by the example of silicon (Fig. 1), the available potentials have improved considerably over time. In Table 2, it can be seen that for several codes, the Δ value is smaller for newer potential sets. Moreover, older potentials such as the Troullier-Martins FHI98pp norm-conserving set in ABINIT or the Vanderbilt-type ultrasoft sets in Dacapo and CASTEP all have a substantially larger Δ (Fig. 4). This evolution is evidence of internal quality-control mechanisms used by developers of potentials in the past, as well as of additional, more recent efforts based on the Δ gauge [e.g., the Jollet-Torrent-Holzwarth (JTH) and Standard Solid-State Pseudopotentials



		average <Δ>	AE: Elk	AE: exciting	AE: FHI-aims/tier2	AE: FLEUR	AE: FPLO/T+F+s	AE: RSPt	AE: WIEN2k/acc
AE	Elk	0.6		0.3	0.3	0.6	1.0	0.9	0.3
	exciting	0.5	0.3		0.1	0.5	0.9	0.8	0.2
	FHI-aims/tier2	0.5	0.3	0.1		0.5	0.9	0.8	0.2
	FLEUR	0.6	0.6	0.5	0.5		0.8	0.6	0.4
	FPLO/T+F+s	0.9	1.0	0.9	0.9	0.8		0.9	0.9
	RSPt	0.8	0.9	0.8	0.8	0.6	0.9		0.8
	WIEN2k/acc	0.5	0.3	0.2	0.2	0.4	0.9	0.8	
PAW	GBRV12/ABINIT	0.9	0.9	0.8	0.8	0.9	1.3	1.1	0.8
	GPAW09/ABINIT	1.4	1.3	1.3	1.3	1.3	1.7	1.5	1.3
	GPAW09/GPAW	1.6	1.5	1.5	1.5	1.5	1.8	1.7	1.5
	JTH02/ABINIT	0.6	0.6	0.6	0.6	0.6	0.9	0.7	0.5
	PSlib100/QE	0.9	0.9	0.8	0.8	0.8	1.3	1.1	0.8
	VASPGW2015/VASP	0.6	0.4	0.4	0.4	0.6	1.0	0.8	0.3
USPP	GBRV14/CASTEP	1.1	1.1	1.1	1.0	1.0	1.4	1.3	1.0
	GBRV14/QE	1.1	1.0	1.0	0.9	1.0	1.4	1.3	1.0
	OTFG9/CASTEP	0.7	0.4	0.5	0.5	0.7	1.0	1.0	0.5
	SSSP/QE	0.5	0.4	0.3	0.3	0.5	0.9	0.8	0.3
	Vdb2/DACAPO	6.3	6.3	6.3	6.3	6.3	6.4	6.5	6.2
NCPP	FHI98pp/ABINIT	13.3	13.5	13.4	13.4	13.2	13.0	13.2	13.4
	HGH/ABINIT	2.2	2.2	2.2	2.2	2.0	2.3	2.2	2.1
	HGH-NLCC/BigDFT	1.1	1.1	1.1	1.1	1.0	1.2	1.1	1.0
	MBK2013/OpenMX	2.0	2.1	2.1	2.1	1.9	1.8	1.8	2.0
	ONCVPSP(PD0.1)/ABINIT	0.7	0.7	0.7	0.7	0.6	1.0	0.8	0.6
	ONCVPSP(SG15)1/QE	1.4	1.4	1.3	1.3	1.3	1.6	1.5	1.3
	ONCVPSP(SG15)2/CASTEP	1.4	1.4	1.4	1.4	1.3	1.6	1.5	1.4

Fig. 4. Δ values for comparisons between the most important DFT methods considered (in millielectron volts per atom). Shown are comparisons of all-electron (AE), PAW, ultrasoft (USPP), and norm-conserving pseudopotential (NCPP) results with all-electron results (methods are listed in alphabetical order in each category). The labels for each method stand for code, code/specification (AE), or potential set/code (PAW, USPP, and NCPP) and are explained in full in tables S3 to S42. The color coding illustrates the range from small (green) to large (red) Δ values. The mixed potential set SSSP was added to the ultrasoft category, in agreement with its prevalent potential type. Both the code settings and the DFT-predicted EOS parameters behind these numbers are included in tables S3 to S42, and fig. S1 provides a full Δ matrix for all methods mentioned in this article.

(SSSP) libraries]. The considerable difference in the older potentials, even for the predefined structures in this relatively simple test set, provides a compelling argument to use only the most recent potential files of a given code.

In addition to the comparison with all-electron codes, it is also interesting to assess how the same PAW or pseudopotential recipes are implemented in different ways. When both the GPAW and ABINIT codes use the GPAW 0.9 PAW set, for example, they agree to within a Δ of 0.6 meV per atom. A similar correspondence is found for the Schlipf-Gygi 2015-01-24 optimized norm-conserving Vanderbilt pseudopotentials (ONCVPSP) (0.3 meV per atom between Quantum ESPRESSO and CASTEP), the Garrity-Bennett-Rabe-Vanderbilt (GBRV) 1.4 ultrasoft pseudopotentials (0.3 meV per atom between Quantum ESPRESSO and CASTEP) and the GBRV 1.2 set (0.7 meV per atom between PAW potentials in ABINIT and ultrasoft potentials in Quantum ESPRESSO). In this case, too, the small Δ values indicate a good agreement between codes. This agreement moreover encompasses varying degrees of numerical convergence, differences in the numerical implementation of the particular potentials, and computational differences beyond the pseudization scheme, most of which are expected to be of the same order of magnitude or smaller than the differences among all-electron codes (1 meV per atom at most).

Conclusions and outlook

Solid-state DFT codes have evolved considerably. The change from small and personalized codes to widespread general-purpose packages has pushed developers to aim for the best possible precision. Whereas past DFT-PBE literature on the lattice parameter of silicon indicated a spread of 0.05 Å, the most recent versions of the implementations discussed here agree on this value within 0.01 Å (Fig. 1 and tables S3 to S42). By comparing codes on a more detailed level using the Δ gauge, we have found the most recent methods to yield nearly indistinguishable EOS, with the associated error bar comparable to that between different high-precision experiments. This underpins the validity of recent DFT EOS results and confirms that correctly converged calculations yield reliable predictions. The implications are moreover relevant throughout the multidisciplinary set of fields that build upon DFT results, ranging from the physical to the biological sciences.

In spite of the absence of one absolute reference code, we were able to improve and demonstrate the reproducibility of DFT results by means of a pairwise comparison of a wide range of codes and methods. It is now possible to verify whether any newly developed methodology can reach the same precision described here, and new DFT applications can be shown to have used a method and/or potentials that were screened in this way. The data generated in this study serve as a crucial enabler for such a reproducibility-driven paradigm shift, and future updates of available Δ values will be presented at http://molmod.ugent.be/deltacodesdft. The reproducibility of reported results also provides a sound basis for further improvement to the accuracy of DFT, particularly in the investigation of new DFT functionals, or for the development of new computational approaches. This work might therefore substantially accelerate methodological advances in solid-state DFT.

Future work can examine the reproducibility of different codes even further. Such work might involve larger benchmark sets (describing different atomic environments per element), other functionals, an exhaustive comparison of different relativistic treatments, and/or a more detailed account of computational differences (using databases or scripts, for example). The precision of band gaps, magnetic anisotropies, and other non-EOS properties would also be of interest. However, the current investigation of EOS parameters provides the most important pass-fail test of the quality of different implementations of Kohn-Sham theory. A method that is not able to reach

an acceptable precision with respect to the EOS of the elemental crystals will probably not fulfill even more stringent demands.

Methods summary

This study relied on the collective efforts of a large group of developers and expert users to make pairwise comparisons of widely used DFT codes. We compared 40 DFT methods in terms of Δ, which expresses the root-mean-square difference between the EOS of two codes, averaged over a benchmark set of 71 elemental crystals (Eq. 1). Our approach, including details about the codes used, is described further in the supplementary materials. The reported settings yield highly converged results but may not be necessary for typical DFT applications. In particular, the use of sometimes very small electronic smearing widths requires much higher numbers of k-points than routine DFT calculations warrant.

REFERENCES AND NOTES

1. Open Science Collaboration, Estimating the reproducibility of psychological science. *Science* **349**, aac4716 (2015). doi: 10.1126/science.aac4716; pmid: 26315443
2. A checklist for photovoltaic research. *Nat. Mater.* **14**, 1073 (2015). pmid: 25613704
3. W. F. van Gunsteren, The seven sins in academic behavior in the natural sciences. *Angew. Chem. Int. Ed.* **52**, 118–122 (2013). doi: 10.1002/anie.201204076; pmid: 23203439
4. C. Drummond, N. Japkowicz, W. Klement, S. Macskassy, "Replicability is not reproducibility: Nor is it good science," in *Proceedings, Twenty-Sixth International Conference on Machine Learning*, L. Bottou, M. Littman, Eds. (2009); www.csi.uottawa.ca/~cdrummon/pubs/ICMLws09.pdf.
5. G. Miller, A scientist's nightmare: Software problem leads to five retractions. *Science* **314**, 1856–1857 (2006). doi: 10.1126/science.314.5807.1856; pmid: 17185570
6. M. Vihinen, No more hidden solutions in bioinformatics. *Nature* **521**, 261 (2015). doi: 10.1038/521261a; pmid: 25993922
7. D. C. Ince, L. Hatton, J. Graham-Cumming, The case for open computer programs. *Nature* **482**, 485–488 (2012). doi: 10.1038/nature10836; pmid: 22358837
8. P. Hohenberg, W. Kohn, Inhomogeneous electron gas. *Phys. Rev.* **136**, B864–B871 (1964). doi: 10.1103/PhysRev.136.B864
9. W. Kohn, L. J. Sham, Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965). doi: 10.1103/PhysRev.140.A1133
10. R. O. Jones, Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.* **87**, 897–923 (2015). doi: 10.1103/RevModPhys.87.897
11. S. Curtarolo *et al.*, AFLOW: An automatic framework for high-throughput materials discovery. *Comput. Mater. Sci.* **58**, 218–226 (2012). doi: 10.1016/j.commatsci.2012.02.005
12. A. Jain *et al.*, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **1**, 011002 (2013). doi: 10.1063/1.4812323
13. C.-C. Fu, J. Dalla Torre, F. Willaime, J.-L. Bocquet, A. Barbu, Multiscale modelling of defect kinetics in irradiated iron. *Nat. Mater.* **4**, 68–74 (2005). doi: 10.1038/nmat1286
14. M. Friák *et al.*, Methodological challenges in combining quantum-mechanical and continuum approaches for materials science applications. *Eur. Phys. J. Plus* **126**, 101 (2011). doi: 10.1140/epjp/i2011-11101-2
15. S. Kurth, J. P. Perdew, P. Blaha, Molecular and solid-state tests of density functional approximations: LSD, GGAs, and Meta-GGAs. *Int. J. Quantum Chem.* **75**, 889–909 (1999). doi: 10.1002/(SICI)1097-461X(1999)75:4/5<889::AID-QUA54>3.0.CO;2-8
16. V. N. Staroverov, G. E. Scuseria, J. Tao, J. P. Perdew, Tests of a ladder of density functionals for bulk solids and surfaces. *Phys. Rev. B* **69**, 075102 (2004). doi: 10.1103/PhysRevB.69.075102
17. P. Haas, F. Tran, P. Blaha, Calculation of the lattice constant of solids with semilocal functionals. *Phys. Rev. B* **79**, 085104 (2009). doi: 10.1103/PhysRevB.79.085104
18. G. I. Csonka *et al.*, Assessing the performance of recent density functionals for bulk solids. *Phys. Rev. B* **79**, 155107 (2009). doi: 10.1103/PhysRevB.79.155107
19. P. Pernot, B. Civalleri, D. Presti, A. Savin, Prediction uncertainty of density functional approximations for properties of crystals with cubic symmetry. *J. Phys. Chem. A* **119**, 5288–5304 (2015). doi: 10.1021/jp509980w; pmid: 25626469
20. L. A. Curtiss, K. Raghavachari, P. C. Redfern, J. A. Pople, Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation. *J. Chem. Phys.* **106**, 1063 (1997). doi: 10.1063/1.473182
21. J. Paier, R. Hirschl, M. Marsman, G. Kresse, The Perdew-Burke-Ernzerhof exchange-correlation functional applied to the G2-1 test set using a plane-wave basis set. *J. Chem. Phys.* **122**, 234102 (2005). doi: 10.1063/1.1926272; pmid: 16008425
22. Y. Zhao, D. G. Truhlar, Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **41**, 157–167 (2008). doi: 10.1021/ar700111a; pmid: 18186612
23. L. Goerigk, S. Grimme, A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.* **13**, 6670–6688 (2011). doi: 10.1039/c0cp02984j pmid: 21384027
24. A. Kiejna *et al.*, Comparison of the full-potential and frozen-core approximation approaches to density-functional calculations of surfaces. *Phys. Rev. B* **73**, 035404 (2006). doi: 10.1103/PhysRevB.73.035404
25. B. Grabowski, T. Hickel, J. Neugebauer, *Ab initio* study of the thermodynamic properties of nonmagnetic elementary fcc metals: Exchange-correlation-related error bars and chemical trends. *Phys. Rev. B* **76**, 024309 (2007). doi: 10.1103/PhysRevB.76.024309
26. K. Lejaeghere, V. Van Speybroeck, G. Van Oost, S. Cottenier, Error estimates for solid-state density-functional theory predictions: An overview by means of the ground-state elemental crystals. *Crit. Rev. Solid State* **39**, 1–24 (2014). doi: 10.1080/10408436.2013.772503
27. F. Jollet, M. Torrent, N. Holzwarth, Generation of Projector Augmented-Wave atomic data: A 71 element validated table in the XML format. *Comput. Phys. Commun.* **185**, 1246–1254 (2014). doi: 10.1016/j.cpc.2013.12.023
28. E. Küçükbenli *et al.*, http://arxiv.org/abs/1404.3015 (2014).
29. S. Poncé *et al.*, Verification of first-principles codes: Comparison of total energies, phonon frequencies, electron-phonon coupling and zero-point motion correction to the gap between ABINIT and QE/Yambo. *Comput. Mater. Sci.* **83**, 341–348 (2014). doi: 10.1016/j.commatsci.2013.11.031
30. The studies of accuracy and precision for DFT calculations are often referred to as validation and verification (V&V), respectively.
31. D. R. Hamann, M. Schlüter, C. Chiang, Norm-conserving pseudopotentials. *Phys. Rev. Lett.* **43**, 1494–1497 (1979). doi: 10.1103/PhysRevLett.43.1494
32. L. Kleinman, D. M. Bylander, Efficacious form for model pseudopotentials. *Phys. Rev. Lett.* **48**, 1425–1428 (1982). doi: 10.1103/PhysRevLett.48.1425
33. D. Vanderbilt, Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **41**, 7892–7895 (1990). doi: 10.1103/PhysRevB.41.7892; pmid: 9993096
34. P. E. Blöchl, Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994). doi: 10.1103/PhysRevB.50.17953; pmid: 9976227
35. G. Kresse, D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999). doi: 10.1103/PhysRevB.59.1758
36. J. C. Slater, Wave functions in a periodic potential. *Phys. Rev.* **51**, 846–851 (1937). doi: 10.1103/PhysRev.51.846
37. O. K. Andersen, Linear methods in band theory. *Phys. Rev. B* **12**, 3060–3083 (1975). doi: 10.1103/PhysRevB.12.3060
38. E. Sjöstedt, L. Nordström, D. J. Singh, An alternative way of linearizing the augmented plane-wave method. *Solid State Commun.* **114**, 15–20 (2000). doi: 10.1016/S0038-1098(99)00577-3
39. G. K. H. Madsen, P. Blaha, K. Schwarz, E. Sjöstedt, L. Nordström, Efficient linearization of the augmented plane-wave method. *Phys. Rev. B* **64**, 195134 (2001). doi: 10.1103/PhysRevB.64.195134
40. J. M. Wills *et al.*, *Full-Potential Electronic Structure Method*, vol. 167 of *Springer Series in Solid-State Sciences* (Springer-Verlag, 2010).
41. K. Koepernik, H. Eschrig, Full-potential nonorthogonal local-orbital minimum-basis band-structure scheme. *Phys. Rev. B* **59**, 1743–1757 (1999). doi: 10.1103/PhysRevB.59.1743
42. V. Blum *et al.*, *Ab initio* molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **180**, 2175–2196 (2009). doi: 10.1016/j.cpc.2009.06.022
43. J. P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996). doi: 10.1103/PhysRevLett.77.3865; pmid: 10062328
44. G. Yuan, F. Gygi, ESTEST: A framework for the validation and verification of electronic structure codes. *Comput. Sci. Discov.* **3**, 015004 (2010). doi: 10.1088/1749-4699/3/1/015004
45. The NoMaD Repository, http://nomad-repository.eu.
46. S. R. Bahn, K. W. Jacobsen, An object-oriented scripting interface to a legacy electronic structure code. *Comput. Sci. Eng.* **4**, 56–66 (2002). doi: 10.1109/5992.998641
47. G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, B. Kozinsky, AiiDA: Automated interactive infrastructure and database for computational science. *Comput. Mater. Sci.* **111**, 218–230 (2016). doi: 10.1016/j.commatsci.2015.09.013
48. Scripts are available at https://cmr.fysik.dtu.dk/dcdft/dcdft.html.
49. W. B. Holzapfel, M. Hartwig, W. Sievers, Equations of state for Cu, Ag, and Au for wide ranges in temperature and pressure up to 500 GPa and above. *J. Phys. Chem. Ref. Data* **30**, 515 (2001). doi: 10.1063/1.1370170
50. C. Kittel, *Introduction to Solid State Physics* (Wiley, ed. 8, 2005).
51. E. Knittle, in *Mineral Physics and Crystallography: A Handbook of Physical Constants*, T. Ahrens, Ed. (American Geophysical Union, 1995), pp. 98–142.
52. W. A. Grosshans, W. B. Holzapfel, Atomic volumes of rare-earth metals under pressures to 40 GPa and above. *Phys. Rev. B* **45**, 5171–5178 (1992). doi: 10.1103/PhysRevB.45.5171; pmid: 10000231
53. K. G. Dyall, E. van Lenthe, Relativistic regular approximations revisited: An infinite-order relativistic approximation. *J. Chem. Phys.* **111**, 1366 (1999). doi: 10.1063/1.479395
54. D. D. Koelling, B. N. Harmon, A technique for relativistic spin-polarized calculations. *J. Phys. C Solid State Phys.* **10**, 3107–3114 (1977). doi: 10.1088/0022-3719/10/16/019
55. E. van Lenthe, E. Baerends, J. Snijders, Relativistic total energy using regular approximations. *J. Chem. Phys.* **101**, 9783 (1994). doi: 10.1063/1.467943
56. I.-H. Lee, R. M. Martin, Applications of the generalized-gradient approximation to atoms, clusters, and solids. *Phys. Rev. B* **56**, 7197–7205 (1997). doi: 10.1103/PhysRevB.56.7197
57. T. Miyake, T. Ogitsu, S. Tsuneyuki, Quantum distributions of muonium and hydrogen in crystalline silicon. *Phys. Rev. Lett.* **81**, 1873–1876 (1998). doi: 10.1103/PhysRevLett.81.1873
58. W. Windl, M. M. Bunea, R. Stumpf, S. T. Dunham, M. P. Masquelier, First-principles study of boron diffusion in silicon. *Phys. Rev. Lett.* **83**, 4345–4348 (1999). doi: 10.1103/PhysRevLett.83.4345
59. L. Vitos, B. Johansson, J. Kollár, H. Skriver, Exchange energy in the local Airy gas approximation. *Phys. Rev. B* **62**, 10046–10050 (2000). doi: 10.1103/PhysRevB.62.10046
60. R. Miotto, G. P. Srivastava, A. C. Ferraz, Dissociative adsorption of PH_3 on the Si(001) surface. *Phys. Rev. B* **63**, 125321 (2001). doi: 10.1103/PhysRevB.63.125321
61. L. E. Ramos *et al.*, Structural, electronic, and effective-mass properties of silicon and zinc-blende group-III nitride semiconductor compounds. *Phys. Rev. B* **63**, 165210 (2001). doi: 10.1103/PhysRevB.63.165210
62. L. V. C. Assali, W. V. M. Machado, J. F. Justo, Titanium impurities in silicon, diamond, and silicon carbide. *Braz. J. Phys.* **34**, 602–604 (2004). doi: 10.1590/S0103-97332004000400016
63. J. Heyd, J. E. Peralta, G. E. Scuseria, R. L. Martin, Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional. *J. Chem. Phys.* **123**, 174101 (2005). doi: 10.1063/1.2085170; pmid: 16375511
64. J. Paier *et al.*, Screened hybrid density functionals applied to solids. *J. Chem. Phys.* **124**, 154709 (2006). doi: 10.1063/1.2187006; pmid: 16674253
65. F. Tran, R. Laskowski, P. Blaha, K. Schwarz, Performance on molecules, surfaces, and solids of the Wu-Cohen GGA exchange-correlation energy functional. *Phys. Rev. B* **75**, 115131 (2007). doi: 10.1103/PhysRevB.75.115131

ACKNOWLEDGMENTS

This research benefited from financial support from the Research Board of Ghent University; the Fond de la Recherche Scientifique de Belgique (FRS-FNRS), through Projet de Recherches (PDR) grants T.0238.13-AIXPHO and T.1031.14-HiT4FiT; the Communauté Française de Belgique, through the BATTAB project (grant ARC

14/19-057); the U.S. NSF (grant DMR-14-08838); the Swedish Research Council; the Knut and Alice Wallenberg Foundation (grants 2013.0020 and 2012.0031); the Fund for Scientific Research–Flanders (FWO) (project no. G0E0116N); and the U.S. Department of Energy (grant DOE-BES DE-SC0008938). N.A.W.H. was supported by U.S. NSF grant DMR-1105485. J.A.F.-L. acknowledges financial support from the European Union's 7th Framework Marie-Curie Scholarship Program within the ExMaMa Project (project no. 329386). I.D.M., O.E., O.G., D.I., Y.O.K., I.L.M.L., and L.N. acknowledge support from eSSENCE. T.B. was supported by the Academy of Finland (grant 263416) and the COMP Centre of Excellence. C.D., A.G., and S.L. acknowledge support from the Deutsche Forschungsgemeinschaft (DFG) and the Einstein Foundation, Berlin. M.Sche. and C.D. received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 676580 with The Novel Materials Discovery (NOMAD) Laboratory, a European Center of Excellence. A.D.C., S.d.G., and E.K. acknowledge support from the Italian Ministry of Education, Universities, and Research (MIUR) through PRIN (Projects of National Interest) 2010–2011 (registration no. 20105ZZTSE_005). P.J.H., D.B.J., and M.I.J.P. are grateful for financial support by the Engineering and Physical Sciences Research Council (EPSRC) under UK Car-Parrinello (UKCP) grant EP/K013564/1. C.J.P. and J.R.Y. acknowledge support from the Collaborative Computational Project for NMR Crystallography under EPSRC grant EP/J010510/1. W.P. acknowledges funding by FWO. D.J. is grateful for financial support by EPSRC under grant EP/J017639/1. S.Sa. acknowledges support from the Swiss National Science Foundation (SNSF). G.-M.R. is thankful for personal financial support from FRS-FNRS. The work by I.E.C. and N.M. was supported by the SNSF's National Centre of Competence in Research MARVEL. G.K. and P.B. acknowledge support by the Austrian Science Fund, project SFB-F41 (ViCoM). S.C. acknowledges financial support from OCAS NV by an OCAS-endowed chair at Ghent University. Computational resources were as follows: The Ghent University contributors used the Stevin Supercomputer Infrastructure at Ghent University, which is funded by Ghent University, FWO, and the Flemish Government (Economy, Science, and Innovation Department). The Université Catholique de Louvain contributors used the Tier-1 supercomputer of the Fédération Wallonie-Bruxelles (funded by the Walloon Region under grant agreement no. 1117545), the Centre de Calcul Intensif et de Stockage de Masse–Université Catholique de Louvain supercomputing facilities, and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie-Bruxelles (CÉCI) (funded by the FRS-FNRS under convention 2.5020.11). The Science and Technology Facilities Council, Scientific Computing Department's SCARF (Scientific Computing Application Resource for Facilities) cluster was used for the CASTEP calculations. The Basel University and École Polytechnique Fédérale de Lausanne contributors used the Swiss National Supercomputing Center in Lugano. Finland's IT Centre for Science was used for the RSPt calculations. K.L. and F.T. thank C. Becker for instructive discussions on the comparison of atomic-scale simulations. K.L. and S.C. thank W. Dewitte for drafting the summary figure. S.J.C., P.J.H., C.J.P., M.I.J.P., K.R., and J.R.Y. declare the receipt of income from commercial sales of CASTEP by Biovia. N.M. and M.Sche. are members of the Board of Trustees of the Psi-k Electronic Structure Network. P.G. is director of the Quantum ESPRESSO Foundation, and N.M. is a representative member. X.G., D.R.H., M.T., D.C., F.J., and G.-M.R. are members of the Advisory Board of ABINIT, an organization that develops and publishes open-source software related to this article. Commercial software is identified to specify procedures. Such identification does not imply recommendation by the National Institute of Standards and Technology. Atomic Simulation Environment scripts (*46*) for several of the codes are available online (*48*). All data are listed in tables S3 to S42.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/351/6280/aad3000/suppl/DC1
Materials and Methods
Fig. S1
Tables S1 to S42
References (*66–115*)

27 August 2015; accepted 19 February 2016
10.1126/science.aad3000