QUANTUM CHEMISTRY

Pushing the frontiers of density functionals by solving the fractional electron problem

James Kirkpatrick[1]*†, Brendan McMorrow[1]†, David H. P. Turban[1]†, Alexander L. Gaunt[1]†, James S. Spencer[1], Alexander G. D. G. Matthews[1], Annette Obika[1], Louis Thiry[2], Meire Fortunato[1], David Pfau[1], Lara Román Castellanos[1], Stig Petersen[1], Alexander W. R. Nelson[1], Pushmeet Kohli[1], Paula Mori-Sánchez[3], Demis Hassabis[1], Aron J. Cohen[1,4]*

Density functional theory describes matter at the quantum level, but all popular approximations suffer from systematic errors that arise from the violation of mathematical properties of the exact functional. We overcame this fundamental limitation by training a neural network on molecular data and on fictitious systems with fractional charge and spin. The resulting functional, DM21 (DeepMind 21), correctly describes typical examples of artificial charge delocalization and strong correlation and performs better than traditional functionals on thorough benchmarks for main-group atoms and molecules. DM21 accurately models complex systems such as hydrogen chains, charged DNA base pairs, and diradical transition states. More crucially for the field, because our methodology relies on data and constraints, which are continually improving, it represents a viable pathway toward the exact universal functional.

Computing electronic energies underpins theoretical chemistry and materials science, and density functional theory (DFT) (*1*, *2*) promises an exact and efficient approach. However, there is a conundrum at the heart of DFT: The exact functional—mapping electron density to energy—is proven to exist, but little practical guidance is given on its explicit form. Approximations to the exact functional have enabled numerous investigations of matter at a microscopic level and rank as some of the most impactful works in the whole of science (*3*). Nevertheless, despite their design and success, pathological errors persist in these approximations, and it has been known for over a decade (*4*) that the root cause of many of these errors is the violation of exact conditions for systems with fractional electrons. In this work, we used deep learning to train a functional that respects these conditions and thus has excellent performance across main-group chemistry.

Since the early days of DFT, it has been clear that approximations improve when they satisfy more of the mathematical constraints of the exact functional and fit more systems. Seventeen known exact constraints (but not the fractional constraints) are satisfied by the strongly constrained and appropriately normed (SCAN) functional (*5*), which achieves unprecedented accuracy and predictiveness for bonded systems among the functionals that are not fitted to any bonded system. Recent work has also begun to address the fractional constraints, of particular interest being a localized correction on the orbitals (*6*, *7*). In parallel, machine learning has emerged as a powerful tool at the level of molecular modeling in chemistry (*8*, *9*) and has been recently applied to functional development (*10*, *11*). Proof-of-principle studies have shown that neural networks (*12*–*16*) can be trained on molecular data, but to date, they are not competitive in accuracy with traditional hand-designed functionals.

In this work, we present a functional, DM21 (DeepMind 21) that is state of the art on thorough benchmark evaluation and has qualitatively improved properties because it obeys two classes of constraints on systems with fractional electrons. The types of fractional constraints considered were fractional charge (FC) systems, with a noninteger total charge, and fractional spin (FS) systems, with noninteger spin magnetization. In both cases, the exact energy is a linear interpolation of the energy of the neighboring integer systems (*17*, *18*). FC and FS systems are fictitious, but real charge densities can include regions that have FC or FS character, and therefore, correctly modeling these idealized problems helps to ensure that functionals behave correctly in a wide variety of molecules and materials. The FC and FS linearity conditions have shown to be challenging to address with manual design of the functional, but they are easy to illustrate as examples. This situation is ideally suited to a deep learning framework, in which the constraints can be expressed as data and a functional can be trained to obey them and to reproduce the energy of molecular systems.

Our functional is illustrated in Fig. 1. Only the exchange-correlation term was learned and interfaced to a standard Kohn-Sham DFT code [PySCF (*19*)]. The functional was evaluated by integrating local energies computed by a multilayer perceptron (MLP), which took as input both local and nonlocal features of the occupied Kohn-Sham (KS) orbitals, and can be described as a local range-separated hybrid. To train the functional, the sum of two objective functions was used: a regression loss for learning the exchange-correlation energy itself and a gradient regularization term that ensured that the functional derivatives can be used in self-consistent field (SCF) calculations after training. For the regression loss, we used a dataset of fixed densities representing reactants and products for 2235 reactions, and the network was trained to map from these densities to high-accuracy reaction energies by means of a least-squares objective (Fig. 1B). Specifically, 1161 training reactions represented atomization, ionization, electron affinity, and intermolecular binding energies of small main-group, H-Kr molecules, and 1074 represented the crucial FC and FS densities only for the atoms H-Ar (supplementary materials, section 2.1). The fixed densities for the main-group molecules were obtained from a popular traditional functional [B3LYP (*20*)], and the energy labels were either obtained from literature (*21*–*25*) or based on in-house complete basis set CCSD(T) (coupled-cluster with single and double and perturbative triple excitations) calculations. More justification on the use of a fixed charge density is provided in the supplementary materials, section 4.3. For gradient training, perturbation theory gives the leading order change in energy, δE_{SCF}, after a single SCF iteration (supplementary materials, section 1.3.1). This energy change depends on the derivatives of the exchange-correlation functional (Fig. 1C), and adding $\delta E_{\mathrm{SCF}}^2$ to the training objective encourages the model to avoid making spuriously large orbital rotations away from reasonable orbitals during self-consistent iteration. This approach was considerably cheaper than supervising explicit self-consistent iterations during training (*26*) or Monte Carlo methods to supervise densities (*12*). Networks with gradients regularized in this way were able to run self-consistently on all reactions in large main-group benchmarks, and DM21 produced accurate molecular densities (supplementary materials, section 5).

After training, the behavior of the functional was analyzed, starting with the archetypal FC and FS systems shown in Fig. 2, A and B. We compared DM21 with SCAN and popular hybrid functionals B3LYP (*20*), M06-2X (*27*), and ωB97X (*28*), with all calculations carried out by using a modified version of PySCF (*19*). Generally, traditional functional approximations are convex with respect to the FC exact condition and concave with respect to the FS exact condition, with improved performance

[1]DeepMind, 6 Pancras Square, London N1C 4AG, UK. [2]Département d'informatique, ENS, CNRS, PSL University, Paris, France. [3]Departamento de Química and IFIMAC, UAM, 28049, Madrid, Spain. [4]Max Planck Institute for Solid State Research, 70569 Stuttgart, Germany.
*Corresponding author. Email: kirkpatrick@google.com (J.K.); aroncohen@google.com (A.J.C.)
†These authors contributed equally to this work.

on FC coming at the cost of a larger error in FS, and vice versa. DM21 stands out in comparison as close to the correct behavior for both FC and FS. The functional was trained only on the exact conditions for bare atoms, but correct behavior was also seen on fragments of molecules for both FC and FS, albeit with a somewhat larger error. This result shows that DM21 has not simply memorized the training examples but has found features in the charge density of the atom data that usefully generalize to molecular systems.

Additional limitations of current functionals associated with FC and FS errors are incorrect description of bond breaking for charged and closed-shell neutral molecules, respectively. When dissociating a charged molecule, functionals with a convex error for FC artificially lower the energy by delocalizing charge; as such, they predict that—even at infinite separation—a charged molecule is bound. This limitation is the essence of the well-known charge delocalization error in DFT, and DM21 achieves the correct asymptote as in Fig. 2C. Related discussion on eigenvalues is available in the supplementary materials, section 6. Traditional functionals also grossly overestimate the energy of a stretched closed-shell molecule, whereas DM21 yields correct asymptotes (Fig. 2D). This overestimation is often described in terms of static correlation error under the interpretation that at large separation, there is near degeneracy of bonding and antibonding states that cannot be represented by a single reference method.

Following previous studies (*4*, *29*), we revisited this interpretation and instead suggest that the error is due to the overestimation of the energy for spin delocalized solutions: Closed-shell orbitals are not capable of artificially breaking spin symmetry and localizing spins, giving asymptotes that are too high for functionals with FS error. Additionally, we made a quantitative evaluation of the advantage of DM21 for bond breaking by using an accurate Quantum Monte Carlo bond breaking benchmark (BBB) (supplementary materials, section 8.1). For neutral molecules at intermediate distances, DM21 could exhibit a "hump" in the energy. This feature, seen before with other methods such as the random phase approximation (*30*), can be corrected with an extension to fractional occupation of the closed-shell orbitals (*31*). Of the functionals presented, optimization of the orbital occupations lowered the hump energy only for DM21 (supplementary materials, section 3.2).

Having established the improved FC and FS behavior of DM21 on textbook systems, how this behavior leads to improved description of subtle electronic structure in systems of scientific interest is illustrated in Fig. 3. Three systems from across the sciences were considered: charge delocalization in a DNA base pair, magnetic properties of a compressed hydrogen chain, and reaction barrier heights for a ring-opening intermediate with diradical character. Charge transport in DNA is a subject of considerable experimental and theoretical interest (*32*), and the distribution of the charge of an ionized base pair (adenine

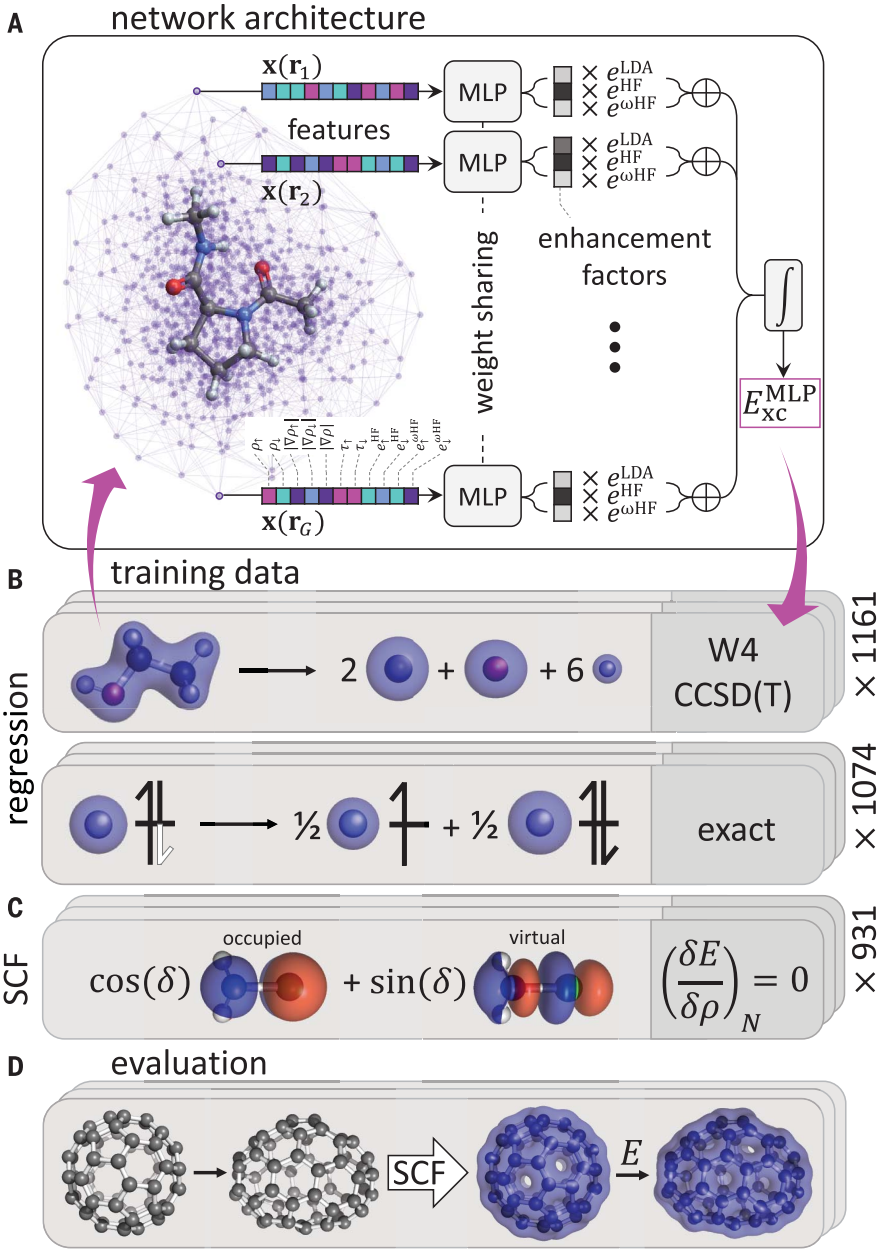


Fig. 1. Overview of the functional architecture and training. (**A**) Features of the electron density computed from KS orbitals are sampled on an atom-centered quadrature grid. Specifically, the input features are the spin-indexed charge density ρ, the norm of its gradient $|\nabla\rho|$, the kinetic energy density τ, and the (range-separated) local Hartree Fock exchange energy densities $e^{\omega HF}$ and e^{HF}. These are fed through a shared MLP that predicts local enhancement factors for local density approximation and Hartree-Fock contributions to the local exchange-correlation energy density, which is then integrated over all space. A dispersion correction is then added to the functional. (**B**) The network is trained by using a dataset of KS input densities and high-accuracy energy labels for molecules and exact mathematical constraints. (**C**) The gradient of the learned functional at fixed electron number (*N*) is supervised by requesting that the supplied orbitals are a stationary point of the total energy with respect to unitary rotation of occupied and virtual orbitals (illustrated by angle δ). (**D**) Once trained, the functional can be deployed in self-consistent calculations. Numbers on the right indicate dataset sizes (excluding grid augmentations) for the DM21 functional.

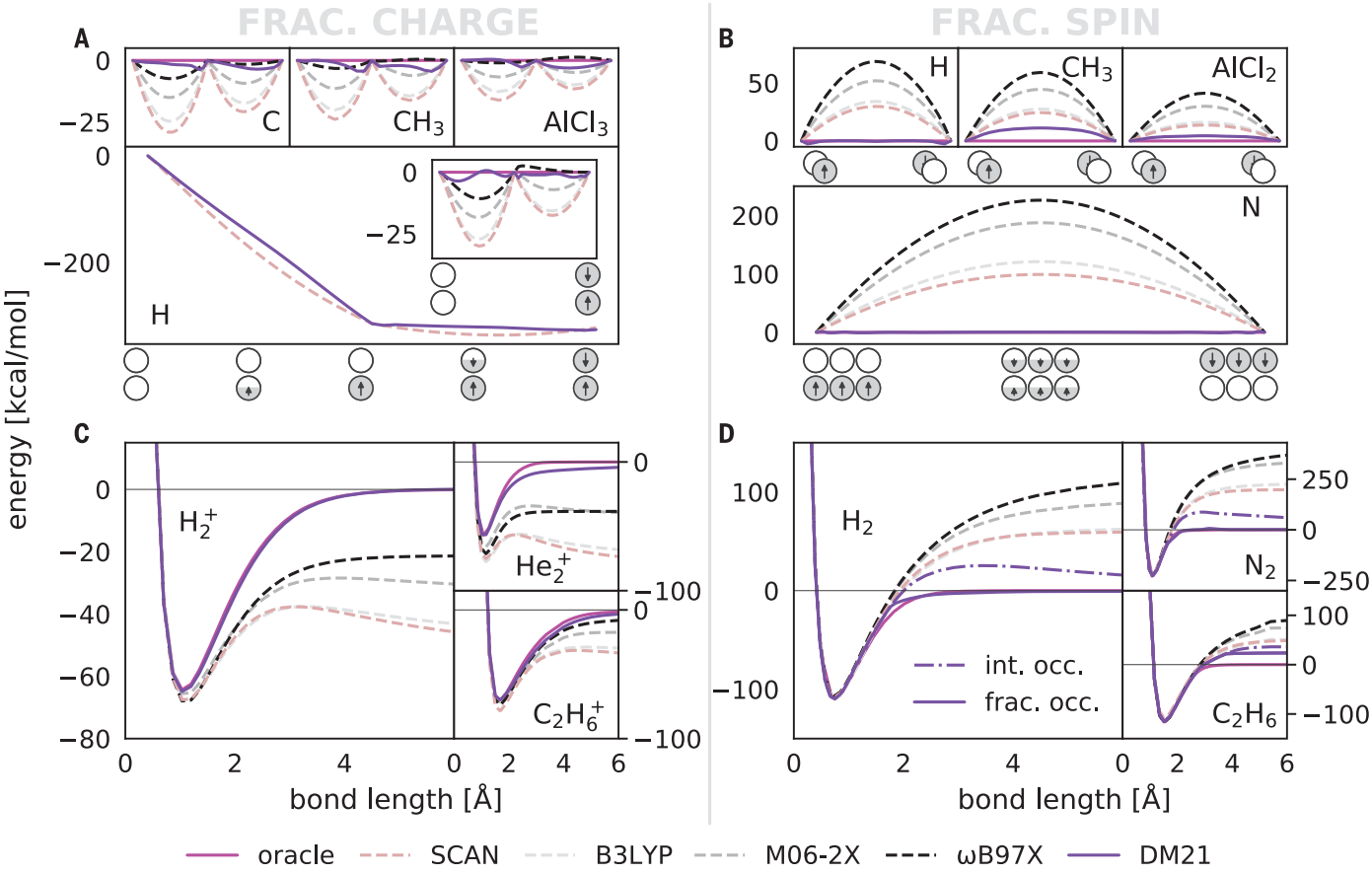


Fig. 2. Training on fractional electron constraints solves charge and spin localization and delocalization errors. (A) and (C) relate to the FC constraint, and (B) and (D) relate to the FS constraint. (**A**) DM21 correctly captures the piecewise linear energy of a H atom as the electron number is continuously varied. (Insets) Deviation from linear behavior for simple atoms (H and C), and small molecules. (**B**) DM21 correctly captures the constancy condition of energy upon interpolating between spin flipped solutions. Shown are the results for a quadruplet (N) and some doublets (H, CH_3, and $AlCl_2$). (**C**) Correct handling of the fractionally charged states generalizes to improved cation binding curves for DM21. The oracle is HF for H_2^+ and UCCSD(T) for He_2^+ and $C_2H_6^+$. (**D**) Improved performance on closed-shell bond breaking. DM21 gives the correct stretched limit but shows a bump at intermediate distances, which is corrected in a restricted optimization that allows fractional occupation of the highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO). Oracles for these curves come from FermiNet QMC calculations, except for C_2H_6, which used UCCSD(T) at the basis set limit.

and thymine) is shown in Fig. 3A. A popular functional such as B3LYP predicts charge density delocalized over both base pairs, but this prediction is an artefact driven by the violations of FC conditions for the individual bases. Conversely, DM21 is much closer to the correct FC behavior, and charge is localized on the adenine unit alone. The difficulties for traditional functionals to localize charges are well understood, but artificial spin localization errors associated with FS violation are less well studied.

In a highly compressed hydrogen chain, unrestricted DFT calculations with traditional functionals localizes spin on what appear to be antiferromagnetic domains (Fig. 3B). This observation has been used to bolster the evidence for a magnetic phase transition (*33*). Conversely, high-level wave function methods did not yield spin polarized solutions, suggesting that the symmetry breaking might be driven by errors in traditional functionals, and DM21 predicts a ground state with no spin symmetry breaking. In the example of ring opening in bicyclobutane (C_4H_6), the energy of the disrotatory transition state is highly overestimated in unrestricted calculations with functionals such as M06-2X and ωB97X but is correctly predicted by DM21. This is again linked to FS error; although the transition state is a singlet, DFT predicts partial localization of the spin with an intermediate $\langle S^2 \rangle$ value between 0 and 1. This means that functionals with incorrect FS behavior give large errors. This result is highly reminiscent of the way that energy is overestimated for closed-shell bond breaking. For C_4H_6, high-level reference calculations are available to verify that DM21 behaves correctly (*34*), but we also observed that hybrid functionals tend to overestimate barrier heights for transition states with intermediate levels of spin polarization in other sets of reactions (supplementary materials, section 7.2), and this phenomeneon has also been observed in the literature for other cycloaddition reactions (*35*).

Last, beyond the treatment of FC and FS, analysis of DM21 is extended to consider broader classes of main-group chemistry contained in large benchmark sets. Shown in Fig. 4 is the summary performance of DM21 compared with existing functionals on the GMTKN55 benchmark (*36*), a set of subbenchmarks used to probe the behavior of functionals for several important chemical tasks that require extrapolation to systems very distinct from the training set. GMTKN55 includes systems that contain heavy atoms beyond Kr that were never seen during training and that therefore we would not normally recommend for DM21 [we evaluated these using pseudo-potentials following the method in (*36*)]. We calculated the mean absolute error for each subbenchmark and

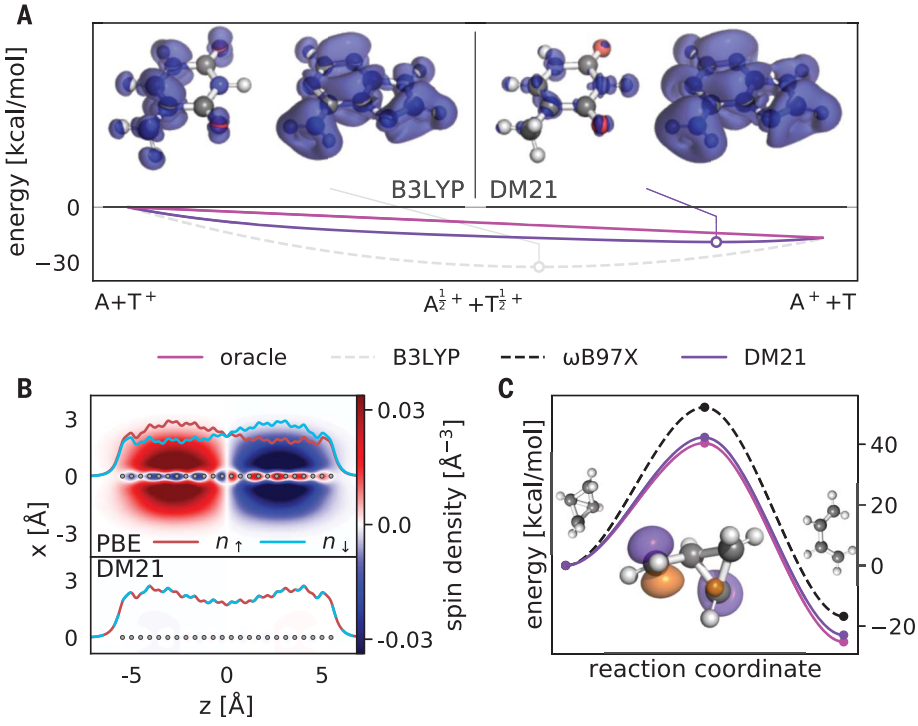
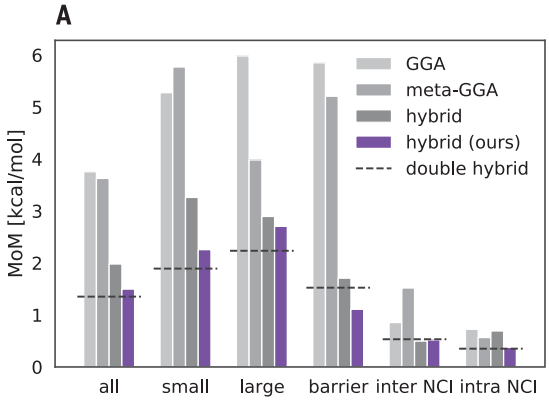


Fig. 3. Exact constraints improve performance on challenging chemistry. (**A**) Charge density for a singly ionized adenine-thymine base pair. B3LYP unphysically delocalizes charge on both base pairs despite adenine having a deeper ionization potential. Conversely, DM21 displays localization of charge on the adenine only. (**B**) Spin density for a compressed chain of 24 hydrogen atoms at 0.48 Å separation. The line density n for each spin channel is overlaid (supplementary materials, section 3.3). The PBE functional (*41*) breaks spin symmetry and leads to an apparent magnetization along the chain. This effect is also predicted by other functionals but is absent in DM21. (**C**) The conrotatory pathways of bicyclobutane isomerization. The HOMO of a single spin channel in an unrestricted calculation is shown for the transition states. Spin is delocalized across two atoms in the conrotatory path, requiring satisfaction of FS for accurate modeling. The oracle is diffusion Monte Carlo from (*42*).



B

	GMTKN55 MoM	BBB	QM9
SCAN(:D3$_{BJ}$)	3.63	52.90	3.57
ωB97X(-V)	2.45	70.28	2.13
M06-2X	2.08	64.50	2.20
PW6B95:D3$_0$	1.98	57.81	2.24
DM21	**1.50**	**13.20**	**1.66**

Fig. 4. State-of-the-art performance by DM21 on benchmarks. All errors are in kcal/mol. (**A**) The MoM error metric in each class of reactions from GMTKN55. More details are available in the supplementary materials, section 8.2. DM21 is compared with functionals at rungs two to five, with strong GMTKN55 metrics from (*43*): revPBE:D3$_{BJ}$ (*44*), SCAN:D3$_{BJ}$, and PW6B95:D3$_0$ (*45*). The dashed black line indicates the performance of the double-hybrid functional DSD-PBEP86:D3$_{BJ}$ (*46*). (**B**) Performance of DM21 compared with the SCAN functional and the three best performing hybrid functionals on three benchmark sets: BBB (supplementary materials, section 8.1), GMTKN55, and QM9. The BBB benchmark measures mean absolute errors for selected first- and second-row diatomics relative to high-level UCCSD(T) (cation) and QMC (neutral) calculations. The neutral dimer DFT calculations use a restricted ansatz with integer occupation. QM9 errors are taken relative to high-level G4(MP2) theory. (:D3$_{BJ}$) indicates that the best of SCAN with or without D3 correction is reported on each metric and similarly for VV10 variants (*47*) of ωB97X.

report the mean of these means (MoM), with additional reweighted scores presented in the supplementary materials, section 8.2. Overall, the behavior of DM21 is better than the best hybrid functional and approaches the performance of the much more expensive double-hybrid functionals. In particular, DM21 excels at the description of barrier heights, which is unexpected given that no transition states were present in the training data. Ablation of the training data revealed that the improved performance on this class stems from the FC and FS training data (supplementary materials, section 8.2). Additionally, DM21 considerably outperforms existing functionals on the mindless benchmark subset (MB16-43) of GMTKN55, a set of atomization energies for randomly generated geometries of atoms that was designed to test performance on out-of-distribution exotic geometries. Otherwise well-performing functionals have large errors for this dataset, such as ωB97-X (>30 kcal/mol error), as well as non-empirical functionals, such as SCAN (15 kcal/mol error), but DM21 had an error of <5 kcal/mol. This dataset is as far from the training and validation set as is possible, showing that our functional performs well even when applied to out-of-distribution generalization examples. To further stress this out-of-distribution generalization, benchmarking on the QM9 (*37*, *38*) dataset is shown in Fig. 4B. This is a collection of 133,857 enumerated isomers of

organic molecules with up to nine heavy atoms, and again DM21 displays state-of-the-art performance when compared with that of existing functionals.

This work has successfully demonstrated that deep learning provides a framework for the development of improved functionals with new properties. Specifically, development of DM21 combined highly accurate chemical data and fractional electron constraints to address major shortcomings in prior functionals. This combination led to a better description of the quantum mechanical interaction of electrons, from simple atomization energies to complicated reaction barriers and exotic compressed hydrogen chains. This work has focused on main-group chemistry, but the methodology can easily be extended to incorporate new data and constraints that will allow even better functionals to be trained. To illustrate this flexibility, results from adding the uniform electron gas condition to the functional are provided in the supplementary materials, section 4.2. Many natural phenomena, from charge transfer excitations to the stripe phase in superconducting cuprates (*39*), rely on subtle effects dependent on the motion of charge and spin polarization, and we believe that the functional presented here, and the approach that we suggest, are central to improving our understanding of these and other properties of materials.

REFERENCES AND NOTES

1. P. Hohenberg, W. Kohn, *Phys. Rev.* **136** (3B), B864–B871 (1964).
2. W. Kohn, L. J. Sham, *Phys. Rev.* **140**, A1133–A1138 (1965).
3. R. Van Noorden, B. Maher, R. Nuzzo, *Nature* **514**, 550–553 (2014).
4. A. J. Cohen, P. Mori-Sánchez, W. Yang, *Science* **321**, 792–794 (2008).
5. J. Sun, A. Ruzsinszky, J. P. Perdew, *Phys. Rev. Lett.* **115**, 036402 (2015).
6. C. Li, X. Zheng, N. Q. Su, W. Yang, *Natl. Sci. Rev.* **5**, 203–215 (2018).
7. N. Q. Su, C. Li, W. Yang, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 9678–9683 (2018).
8. R. Ramakrishnan, P. O. Dral, M. Rupp, O. A. von Lilienfeld, *J. Chem. Theory Comput.* **11**, 2087–2096 (2015).
9. O. T. Unke *et al*., *Chem. Rev.* **121**, 10142–10186 (2021).
10. J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, K. Burke, *Phys. Rev. Lett.* **108**, 253002 (2012).
11. L. Li *et al*., *Int. J. Quantum Chem.* **116**, 819–833 (2016).
12. R. Nagai, R. Akashi, O. Sugino, *npj Comput. Mater.* **6**, 43 (2020).
13. A. V. Sinitskiy, V. S. Pande, Physical machine learning outperforms "human learning" in quantum chemistry. arXiv:1908.00971 [physics.chem-ph] (2019).
14. K. Ryczko, D. A. Strubbe, I. Tamblyn, *Phys. Rev. A* **100**, 022512 (2019).
15. Y. Chen, L. Zhang, H. Wang, E. Weinan, *J. Chem. Theory Comput.* **17**, 170–181 (2021).
16. S. Dick, M. Fernandez-Serra, *Nat. Commun.* **11**, 3509 (2020).
17. J. P. Perdew, R. G. Parr, M. Levy, J. L. Balduz Jr., *Phys. Rev. Lett.* **49**, 1691–1694 (1982).
18. A. J. Cohen, P. Mori-Sánchez, W. Yang, *J. Chem. Phys.* **129**, 121104 (2008).
19. Q. Sun *et al*., *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1340 (2018).
20. A. D. Becke, *J. Chem. Phys.* **98**, 5648–5652 (1993).
21. A. Karton, N. Sylvetsky, J. M. L. Martin, *J. Comput. Chem.* **38**, 2063–2075 (2017).
22. A. Karton, S. Daon, J. M. L. Martin, *Chem. Phys. Lett.* **510**, 165–178 (2011).
23. A. Karton, A. Tarnopolsky, J. M.L. Martin, *Mol. Phys.* **107**, 977–990 (2009).
24. B. Brauer, M. K. Kesharwani, S. Kozuch, J. M. L. Martin, *Phys. Chem. Chem. Phys.* **18**, 20905–20925 (2016).
25. M. K. Kesharwani, D. Manna, N. Sylvetsky, J. M. L. Martin, *J. Phys. Chem. A* **122**, 2184–2197 (2018).
26. L. Li, S. Hoyer, R. Pederson, R. Sun, E. D. Cubuk, P. Riley, K. Burke, Kohn-Sham equations as regularizer: building prior knowledge into machine-learned physics. arXiv:2009.08551 [physics.comp-ph] (2020).
27. Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **120**, 215–241 (2008).
28. J.-D. Chai, M. Head-Gordon, *J. Chem. Phys.* **128**, 084106 (2008).
29. A. J. Cohen, P. Mori-Sánchez, W. Yang, *Chem. Rev.* **112**, 289–320 (2012).
30. M. Fuchs, Y.-M. Niquet, X. Gonze, K. Burke, *J. Chem. Phys.* **122**, 094116 (2005).
31. A. D. Becke, *Density Functionals*, E. R. Johnson, Ed. (Springer, 2014), vol. 365, pp. 175–186.
32. J. C. Genereux, J. K. Barton, *Chem. Rev.* **110**, 1642–1662 (2010).
33. M. Motta *et al*., *Phys. Rev. X* **7**, 031059 (2017).
34. N. P. Bauman, J. Shen, P. Piecuch, *Mol. Phys.* **115**, 2860–2891 (2017).
35. H. V. Pham, K. N. Houk, *J. Org. Chem.* **79**, 8968–8976 (2014).
36. L. Goerigk, S. Grimme, *J. Chem. Theory Comput.* **6**, 107–126 (2010).
37. R. Ramakrishnan, P. O. Dral, M. Rupp, O. A. von Lilienfeld, *Sci. Data* **1**, 140022 (2014).
38. H. Kim, J. Y. Park, S. Choi, *Sci. Data* **6**, 109 (2019).
39. Y. Zhang *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 68–72 (2020).
40. J. Kirkpatrick *et al*., Zenodo (2021); doi:10.5281/zenodo.5482370.
41. J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
42. R. Berner, A. Lüchow, *J. Phys. Chem. A* **114**, 13222–13227 (2010).
43. L. Goerigk *et al*., *Phys. Chem. Chem. Phys.* **19**, 32184–32215 (2017).
44. Y. Zhang, W. Yang, *Phys. Rev. Lett.* **80**, 890 (1998).
45. Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **109**, 5656–5667 (2005).
46. S. Kozuch, J. M. L. Martin, *Phys. Chem. Chem. Phys.* **13**, 20104–20107 (2011).
47. N. Mardirossian, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **16**, 9904–9924 (2014).

ACKNOWLEDGMENTS

We thank C. Donner for technical assistance in running PySCF at scale, T. Back for help with project inception, K. Kavukcuoglu and H. Maclean for useful discussions, D. Rezende and O. Vinyals for reviewing the manuscript, and the rest of the DeepMind team for their support. **Funding:** This research received no external funding. **Author contributions:** J.K., B.M., D.H.P.T., A.L.G., and A. J.C. designed and built the DM21 network architecture and training datasets, with advice from P.M.-S.; J.K., B.M., D.H.P.T., A.L.G., and A.G.D.G.M. trained the functionals. J.K., B.M., D.H.P.T., A.L.G., J.S.S., A.G.D.G.M., P.M.-S., and A.J.C. evaluated the functionals on chemical systems and interpreted the results. J.S.S. and D.P. obtained the FermiNet oracles. J.K., D.H.P.T., L.T., M.F., A.W.R.N., A.J.C., P.K., and D.H. initiated the project. B.M., D.H.P.T., J.S.S., and S.P. contributed to software engineering. J.K., A.O., P.K., and D.H. managed the project. J.K., A.L.G., P.M.-S., and A.J.C. wrote the manuscript, with help from L.R.C. **Competing interests:** There is a pending patent application, US Provisional Application 63/135,223, on this work. The remaining authors declare no competing interests. **Data and materials availability:** The code and learned network weights for running our functionals on new systems using the open source PySCF package is available at https://github.com/deepmind/deepmind-research/tree/master/density_functional_approximation_dm21 and (*40*). The BBB benchmark, errors for DM21 on all reactions in GMTKN55, and all data underlying the figures are also available as digital materials (data files S1 to S3) in the supplementary materials. All other data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abj6511
Materials and Methods
Supplementary Text
Figs. S1 to S9
Tables S1 to S8
References (*48*–*73*)
Data Files S1 to S3

24 May 2021; accepted 16 September 2021
10.1126/science.abj6511

Science

Pushing the frontiers of density functionals by solving the fractional electron problem

James KirkpatrickBrendan McMorrowDavid H. P. TurbanAlexander L. GauntJames S. SpencerAlexander G. D. G. MatthewsAnnette ObikaLouis ThiryMeire FortunatoDavid PfauLara Román CastellanosStig PetersenAlexander W. R. NelsonPushmeet KohliPaula Mori-SánchezDemis HassabisAron J. Cohen

Science, 374 (6573), • DOI: 10.1126/science.abj6511

Improving DFT with deep learning

In the past 30 years, density functional theory (DFT) has emerged as the most widely used electronic structure method to predict the properties of various systems in chemistry, biology, and materials science. Despite a long history of successes, state-of-the-art DFT functionals have crucial limitations. In particular, significant systematic errors are observed for charge densities involving mobile charges and spins. Kirkpatrick *et al*. developed a framework to train a deep neural network on accurate chemical data and fractional electron constraints (see the Perspective by Perdew). The resulting functional outperforms traditional functionals on thorough benchmarks for main-group atoms and molecules. The present work offers a solution to a long-standing critical problem in DFT and demonstrates the success of combining DFT with the modern machine-learning methodology. —YS