Variational density functional theory using the JAX deep-learning differentiable framework.

Stephen G. Dale[1], Tianbo Li[2], Zekun Shi[2,3], Giovanni Vignale[1], A. H. Castro Neto[1,4], Kostya S. Novoselov[1]
[1] Institute for Functional Intelligent Materials, National University of Singapore, Block S9, Level 9, 4 Science Drive 2, Singapore 117544
[2] Sea AL Labs, 1 Fusionopolis Place, #17-10, Galaxis, Singapore 138522
[3]School of Computing, National University of Singapore, COM1, 13 Computing Dr, Singapore, 117417
[4]Centre for Advanced 2D Materials National University of Singapore 6 Science Drive 2, Singapore 117546

sdale@nus.edu.sg

Variational density functional theory (VDFT) is a state-of the-art method for computing the electronic structure of materials in which the traditional self-consistent field (SCF) approach (i.e. solving the Kohn-Sham equation to self-consistency) is converted to a direct-gradient-descent minimization of the total energy with respect to one-electron orbitals and an occupation function, subject to orthogonality constraints. The main advantages of this approach are:

1. the optimization can be executed within a deep-learning differentiable framework.
2. convergence properties are expected to improve. The new optimization methods available through deep-learning will not be susceptible to the same instability present in SCF methods, and will be resistant to becoming stuck in saddle-points as in standard gradient-descent methods. An example of this enhanced stability is presented in Figure 1.
3. nonlocal (hybrid) functionals can be implemented along with semi-local functionals.
4. Kohn-Sham eigenvalues and eigenfunctions can be obtained by a *single* diagonalization of the Kohn-Sham Hamiltonian matrix.

At the heart of our method is a novel reparameterization of the orthogonality constraint by QR decomposition.[1] Our programs are written using Google's JAX deep-learning framework and are designed to be end-to-end differentiable to provide additional tools that are essential for discovery and design of advanced materials. Ultimately any available input variable can be targeted by the direct-gradient-descent optimization function, enabling; alchemical analysis by making nuclear charge a variable; on-the-fly adjustment of density functional parameters; and incorporation of neural networks to train a wide range of solutions.

Experiments are carried out to demonstrate the advantages of our approach in terms of stability. We deliberately choose chemical systems that are known to be challenging for SCF methods and show that our method can reliably converge these. We construct a gradient-only based approach to geometry convergence that can simultaneously converge the electronic and the atomic structure. We will show that we can reliably predict the band structure and the potential energy surfaces of crystalline materials.[2] Finally, we discuss some of the cutting-edge applications we are actively working towards.

References

[1] Tianbo Li, Min Lin, Zheyuan Hua, Kunhao Zhenga, Giovanni Vignale, Kenji Kawaguchib, A. H. Castro Netoc, Kostya S. Novoselov, Shuicheng Yan, ICLR, (2023).
[2] Tianbo Li, Stephen Gregory Dale, Zekun Shi, Jingshu Li, Giovanni Vignale, A. H. Castro Neto, Kostya S. Novoselov, OpenReview, (2023).

Figures

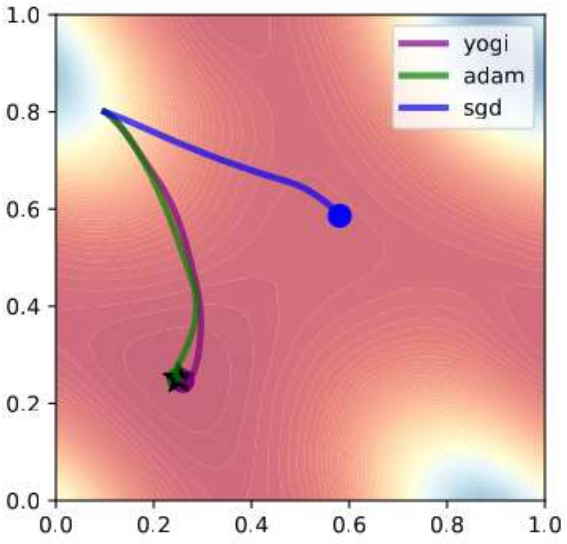

Figure 1. The optimization path taken by three different gradient descent optimizers from an identical starting point to the final (large dot) solution mapped onto a 2D potential energy surface, with the correct final solution represented by the black star. The system is a distorted diamond crystal shown in crystallographic coordinates. Notice that standard direct-gradient-descent method (sgd) gets stuck in saddle-points, newer convergence methods (yogi/adam) that include momentum and variable step sizes are more reliable when searching for local minima.