

Inverse Kohn–Sham Density Functional Theory: Progress and Challenges

Yuming Shi* and Adam Wasserman*



Cite This: *J. Phys. Chem. Lett.* 2021, 12, 5308–5318



Read Online

ACCESS |



Metrics & More

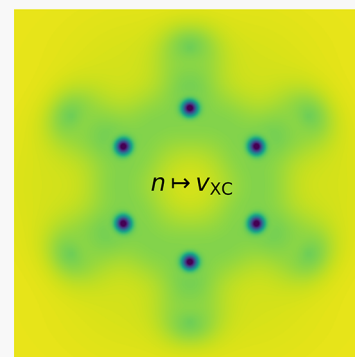


Article Recommendations



Supporting Information

ABSTRACT: Inverse Kohn–Sham (iKS) methods are needed to fully understand the one-to-one mapping between densities and potentials on which density functional theory is based. They can contribute to the construction of empirical exchange–correlation functionals and to the development of techniques for density-based embedding. Unlike the forward Kohn–Sham problems, numerical iKS problems are ill-posed and can be unstable. We discuss some of the fundamental and practical difficulties of iKS problems with constrained-optimization methods on finite basis sets. Various factors that affect the performance are systematically compared and discussed, both analytically and numerically, with a focus on two of the most practical methods: the Wu–Yang method (WY) and the partial differential equation constrained optimization (PDE-CO). Our analysis of the WY and PDE-CO highlights the limitation of finite basis sets. We introduce new ideas to make iKS problems more tractable, provide an overall strategy for performing numerical density-to-potential inversions, and discuss challenges and future directions.



In Principle versus in Practice. Kohn–Sham (KS) density functional theory (DFT)^{1,2} has long been the most widely used method for electronic-structure calculations in condensed matter physics and quantum chemistry computations.^{3,4} KS-DFT is formally exact in the sense that, given the exact exchange–correlation (XC) energy functional, a numerically exact solution of the self-consistent KS-DFT equations is guaranteed to yield the exact ground-state density $n(\mathbf{r})$ and energy for any system of N electrons in a time-independent external potential $v(\mathbf{r})$. The word “exact”, which has been used 4 times already in this introductory paragraph, is sometimes dismissed with scorn when confronted with results in practice. The results of actual DFT calculations are evidently not exact. Nevertheless, it is the proven existence of an *exact* one-to-one correspondence between ground-state densities and potentials which has given impetus to the development of approximations in DFT. A different one-to-one mapping exists for any choice of electron–electron interaction. Nature’s choice is the Coulomb interaction, but the correspondence can be established also for a fictitious system of noninteracting electrons. Calculations are simpler in this fictitious world, and the one-to-one maps allow one to connect the answers back to the real world, *in principle*. In practice, the exchange–correlation energy $E_{xc}[n]$, “nature’s glue”,⁵ needs to be approximated. This Perspective discusses a clash between “in principle” and “in practice”: The density-to-potential mappings are in principle one-to-one. For a given density $n(\mathbf{r})$ and choice of approximate $E_{xc}[n]$, there is in principle only one XC potential $v_{xc}(\mathbf{r})$ corresponding to $n(\mathbf{r})$. That potential is the functional derivative of $E_{xc}[n]$ with respect to the density, evaluated at that density. The potential $v_{xc}(\mathbf{r})$ should be calculable from the given density by solving an inverse

problem, in principle. In practice, unfortunately, the process is a numerical minefield. For a user’s choice of approximate $E_{xc}[n]$, most quantum-chemistry codes solve the *forward* KS-DFT problem; that is, self-consistently calculate $n(\mathbf{r})$ and, from it, the total energy for a given $v(\mathbf{r})$ and N . Another output of this calculation is the self-consistent $v_{xc}(\mathbf{r})$. Much less common are codes that can solve the inverse Kohn–Sham problem (iKS) to find the XC potential corresponding to a given density.^{6,7} But why would one want to do such calculations? We list four answers below:

- Exploring the Hohenberg–Kohn¹ and Runge–Gross⁸ mappings: Density-to-potential inversions are useful for calculating numerically exact XC potentials from numerically exact densities, i.e., a point-by-point exploration of the Hohenberg–Kohn and Runge–Gross mappings. In ground-state DFT, inversions can reveal features of XC potentials that encode the physics of strong correlations^{9,10} and are missed by approximate functionals. In time-dependent DFT, inversions can also shed light on properties of the exact time-dependent XC potentials needed to describe electronic dynamics^{11,12} and to propose approximations that go beyond the adiabatic limit,^{13–16} as needed for describing multiple and charge-

Received: March 8, 2021

Accepted: May 27, 2021

Published: June 1, 2021



- transfer excitations in molecules or excitonic transitions in solids.
- Analyzing errors of approximate XC functionals: iKS can be utilized for measuring density-driven errors in KS-DFT;¹⁷ iKS from numerically exact densities can assist in the development of explicit XC functionals. There are also recent examples of the use of iKS to improve XC functionals in the field of nuclear DFT.^{18–20} Furthermore, machine-learning approaches under development^{21–24} also take inverted XC potentials as input. The efficiency and accuracy of the iKS methods are critically important here.
 - Quantum embedding: Various modern variants of density-based subsystem and embedding approaches^{25,26} employ potential reconstruction techniques.^{27–29} Here, typically, a target density is updated many times in a self-consistent loop and the potential corresponding to that target density must be found at each step of the loop, implying a significant additional computational cost to calculations that do not make use of efficient inversion techniques.
 - Optimized effective potentials (OEP): Approximate expressions for $E_{xc}[n]$ can be either explicit or implicit functionals of the density. Explicit functionals are usually preferred for their simplicity and efficiency. However, implicit XC functionals³⁰ that depend on the KS orbitals are often used when higher accuracy is needed.^{31,32} In particular, OEP methods are being increasingly used for a variety of applications.^{33–35} The OEP method is often the bottleneck of the calculations. The direct OEP method developed by Yang and Wu³⁶ shares many of the same features of iKS discussed here.^{37–39} Improvements on inversion methods would potentially offer significant speed-ups allowing for more widespread adoption of OEP calculations.

iKS Methods. Several methods exist for solving the iKS problem through self-consistent density-based calculations.^{40–50} The connections between many of these methods have been cogently analyzed recently by Kumar et al.^{51,52} An entirely different approach that makes use of the wave function (instead of the density) has been developed by Ryabinkin et al. (mRKS) and has been shown to provide very accurate results.^{53,54} mRKS uses as input the one- and two-particle reduced density matrices. Its computational cost is thus out of reach for calculations on large systems. A detailed discussion explaining the success of wave function-based methods compared to pure density inversions (i.e., where only the densities are needed) was provided recently by Kumar et al.⁵⁵ We concentrate on *pure Kohn–Sham inversions*, referred to as iKS in this Perspective. Methods other than self-consistent calculations have also been designed, many of which feature constrained optimizations.^{50,52,56–60}

In spite of all the methods in hand, iKS problems are still difficult. First, different methods feature different capabilities regarding accuracy and efficiency. Many of them have been tested only on single atoms or diatomic molecules for illustration. Some accurate methods, such as mRKS, are difficult to apply to molecules with more than ~ 10 atoms. Kanungo et al. recently implemented the PDE-constrained optimization method (PDE-CO) on a systematically improvable finite-element basis that provided results for polyatomic systems with an accuracy comparable to that of mRKS using affordable

computation resources.⁶⁰ Moreover, KS inversion is usually ill-posed. A problem is well-posed if a solution exists, is unique, and continuously changes with the input as defined by Hadamard⁶¹ (if a problem is not well-posed, it is then ill-posed). Analogous inverse problems in many fields are generally known for their instabilities. In the case of iKS, the uniqueness is guaranteed by the Hohenberg–Kohn theorem. The existence is known as the ν -representability problem. In discretized systems, densities are ensemble ν -representable.^{62–64} The existence is usually assumed to be true, and the continuity was proven for discrete densities.⁶⁵ However, when the problem comes into practice with a finite basis set and rounding-errors, the errors/variations to the densities are usually arbitrary, the smoothness condition breaks down, and the iKS problem becomes ill-posed.⁵⁹ Because the kinetic operator in the KS equation plays a role of regulator, different XC potentials can reproduce very similar densities.^{22,66,67} In practical KS inversion calculations on finite basis sets, many factors can lead to unphysical oscillations/overfitting to the final $\nu_{xc}(\mathbf{r})$. Therefore, special consideration and regularization is often essential for reasonable results. Simple tricks can often greatly improve an inversion when using one particular inversion method, but those same tricks may be totally unhelpful when used in combination with other methods. Many of these tricks depend on error cancellations of some form,^{38,39,54,60,68–70} making it extremely difficult to predict when they will work or fail.

Organization of this Perspective. We focus on two of the most efficient constrained-optimization methods for iKS, the Wu–Yang method⁶⁹ (WY) and PDE-CO.^{58–60} First, we review the theory behind both methods and their implementation on finite basis sets, including the first implementation of PDE-CO on finite Gaussian potential basis sets. Except for the well-known drawbacks of the finite potential basis set (PBS), PBS significantly improves the efficiency and helps to control problems that one would encounter in a general PDE-CO problem.^{59,60,71} Different factors that influence the stability of the inversion are systematically discussed and compared, both analytically and numerically, including finite basis sets, regularization/corrections, optimization methods, and guide potentials.

There is no way of avoiding the use of many acronyms for the methods, algorithms, and basis sets employed here. Table 1 compiles the acronyms we use most. We highlight our use of “CX/CY” when cc-pCVXZ is being used as the basis set to expand the orbitals and cc-pCVYZ as the basis set to expand the potentials. The input density is CCSD unless further specified. The exact XC potential data comes from quantum Monte Carlo calculations,^{72,73} as in many other discussions on iKS.^{48,49,51–54,69} All the calculations are implemented on Psi4.^{74,75} Atomic units are used throughout.

The Wu–Yang Method. The central idea of the WY method follows directly Levy’s³¹ and Lieb’s⁷⁶ definition of the universal functional. The noninteracting kinetic energy is optimized under the constraint that the density matches a target. Unlike many other methods that depend on some sort of self-consistent calculation, the WY method relies on gradient and Hessian-based optimizations. Thus, the WY method can be very easily implemented with a standard general optimizer and is robust for most systems. On the other hand, given the ill-posed nature of iKS and the ill-conditioning of the Hessian matrices, there can be numerical problems, and regularization is usually essential. The Lagrangian for the WY constrained optimization is

Table 1. Acronym List

CCSD	coupled-cluster singles-and-doubles
CX	cc-pCVXZ
CX/CY	The OBS/PBS is CX/CY.
FA	Fermi–Amaldi
iKS	inverse Kohn–Sham
KS	Kohn–Sham
LDA	local density approximation
mRKS	modified Ryabinkin–Kohut–Staroverov
NSC	null-space correction
OBS	atomic orbital basis set
PBS	potential basis set
PBE	Perdew–Burke–Ernzerhof
PDE	partial differential equation
PDE-CO	PDE-constrained optimization
QMC	quantum Monte Carlo
TSVD	truncated singular value decomposition
WY	the Wu–Yang method
XC	exchange–correlation
ZMP	Zhao–Morrison–Parr

$$W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}] = T_s[\Psi_{\text{det}}] + \int d\mathbf{r} \nu_{\text{KS}}(\mathbf{r}) \{n(\mathbf{r}) - n_{\text{in}}(\mathbf{r})\} \quad (1)$$

where $n_{\text{in}}(\mathbf{r})$ is the target input density and $n(\mathbf{r})$ is

$$n(\mathbf{r}) = N \int d\mathbf{r}_2 \dots d\mathbf{r}_N |\Psi_{\text{det}}(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2 \quad (2a)$$

$$= 2 \sum_i^{N/2} |\psi_i(\mathbf{r})|^2 \quad (2b)$$

The Ψ_{det} in eqs 1 and 2 is the KS Slater determinant consisting of $N/2$ doubly occupied orthonormal KS orbitals $\{\psi_i\}$. Our discussions will be limited to spin-unpolarized systems for convenience. $\nu_{\text{KS}}(\mathbf{r})$ in eq 1 is the KS potential

$$\left[-\frac{1}{2} \nabla^2 + \nu_{\text{KS}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (3)$$

but appears in eq 1 as a Lagrange multiplier because of the necessary condition on the stationary point:

$$\frac{\delta W[\Psi_{\text{det}}, \nu_{\text{KS}}]}{\delta n(\mathbf{r})} = 0 \quad (4)$$

Here, a Lagrangian dual problem of the ZMP⁵⁶ problem is solved and $\nu_{\text{KS}}(\mathbf{r})$ is the dual variable of the density.^{52,77} Therefore, eq 1 needs to be maximized. This will also be proven later through features of the Hessian matrices.

From eq 1

$$\frac{\delta W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]}{\delta \nu_{\text{KS}}(\mathbf{r})} = n(\mathbf{r}) - n_{\text{in}}(\mathbf{r}) \quad (5)$$

and

$$\begin{aligned} \frac{\delta^2 W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]}{\delta \nu_{\text{KS}}(\mathbf{r}) \delta \nu_{\text{KS}}(\mathbf{r}')} &= \frac{\delta n(\mathbf{r})}{\delta \nu_{\text{KS}}(\mathbf{r}')} \\ &= 2 \sum_i^{\text{occ.}} \sum_a^{\text{unocc.}} \frac{\psi_i^*(\mathbf{r}) \psi_a(\mathbf{r}) \psi_i(\mathbf{r}') \psi_a^*(\mathbf{r}')}{\epsilon_i - \epsilon_a} \end{aligned} \quad (6)$$

PDE-Constrained Optimization (PDE-CO). Considering the Hohenberg–Kohn theorem¹ and assuming nondegeneracy, one

would expect from an “exact” input density the exact KS potential. It is generally assumed that, at least near the “exact” density, a more accurate potential is expected to give a more accurate density and vice versa. To find the best $\nu_{\text{KS}}(\mathbf{r})$ given the limitation of a basis set or a grid, searching for a density that is closest to the “exact” input density is usually a good idea. To do this, one could define a density error function to minimize

$$\text{argmin}_{\nu_{\text{KS}}(\mathbf{r})} \int d\mathbf{r} u(\mathbf{r}) |n(\mathbf{r}) - n_{\text{in}}(\mathbf{r})|^w \quad (7)$$

subject to constraints (eq 3) and

$$\int |\psi_i(\mathbf{r})|^2 d\mathbf{r} = 1 \quad (8)$$

$u(\mathbf{r})$ in eq 7 can be a positive weight function that should not change the final convergence theoretically. $u(\mathbf{r}) = \frac{1}{n(\mathbf{r})^p}$ for $p \in [1, 2)$ can be helpful for the asymptotic region when $\nu_{\text{KS}}(\mathbf{r})$ is calculated on a grid. Properties of the weight function $u(\mathbf{r})$ when calculations are performed on a grid can be found in the recent work by Kanungo et al.⁶⁰ Because in our basis-set implementation $\nu_{\text{KS}}(\mathbf{r})$ is expanded as in eq 11, the asymptotic behavior is determined by $\nu_0(\mathbf{r})$ (see below), and we use $u(\mathbf{r}) = 1$ and $w = 2$ in what follows. Note that eq 7 is just the simplest form of an error function. The main reason to choose this form is to derive the gradient analytically, as we will do. However, with the help of automatic differentiation,⁷⁸ more sophisticated error functions can be designed.

The ZMP method⁵⁶ has a connection to PDE-CO in the sense that the penalty terms in ZMP (as defined by Kumar and Harbola⁵²) can be seen as an error function, while ZMP can in turn be regarded as a special method of optimization.

The Lagrangian can be written as

$$\begin{aligned} L[\nu_{\text{KS}}, \{\psi_i\}, \{\epsilon_i\}, \{p_i\}, \{\mu_i\}] \\ = \int d\mathbf{r} (n(\mathbf{r}) - n_{\text{in}}(\mathbf{r}))^2 \\ + \sum_{i=1}^{N/2} \int d\mathbf{r} p_i(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 + \nu_{\text{KS}} - \epsilon_i \right) \psi_i(\mathbf{r}) \\ + \sum_{i=1}^{N/2} \mu_i \left(\int |\psi_i(\mathbf{r})|^2 d\mathbf{r} - 1 \right) \end{aligned} \quad (9)$$

where $p_i(\mathbf{r})$ and μ_i are the Lagrange multipliers introduced for eqs 3 and 8. Variations with respect to the $\psi_i(\mathbf{r})$, ϵ_i , and $\nu_{\text{KS}}(\mathbf{r})$ yield

$$\begin{aligned} \left(-\frac{1}{2} \nabla^2 + \nu_{\text{KS}}(\mathbf{r}) - \epsilon_i \right) p_i(\mathbf{r}) \\ = 8 \times (n_{\text{in}}(\mathbf{r}) - n(\mathbf{r})) \psi_i(\mathbf{r}) - 2\mu_i \psi_i(\mathbf{r}) \end{aligned} \quad (10a)$$

$$\int p_i(\mathbf{r}) \psi_i(\mathbf{r}) d\mathbf{r} = 0 \quad (10b)$$

$$\frac{\delta L}{\delta \nu_{\text{KS}}(\mathbf{r})} = \sum_{i=1}^{N/2} p_i(\mathbf{r}) \psi_i(\mathbf{r}) \quad (10c)$$

The optimization strategy generally adopted is to solve for the $p_i(\mathbf{r})$ and the $\psi_i(\mathbf{r})$ from eqs 3, 10a, and 10b and build $\frac{\delta L}{\delta \nu_{\text{KS}}}$ from eq 10c for each step. The shape of $L[\nu_{\text{KS}}, \{\psi_i\}, \{\epsilon_i\}, \{p_i\}, \{\mu_i\}]$ can be arbitrary (not necessarily convex or concave). Thus, a

Table 2. Wu–Yang: Finite Difference Tests

basis set	Wu–Yang Gradient Test ^a					
	CD/CD		CD/CQ		CQ/CQ	
	before ^c	after	before	after	before	after
Be	5.6×10^{-6}	4.1×10^{-3}	3.9×10^{-6}	0.5	1.8×10^{-5}	1.3×10^{-2}
Ne	5.5×10^{-6}	0.06	5.8×10^{-6}	2.0×10^{-2}	5.3×10^{-5}	2.4×10^{-2}
Ar	1.6×10^{-5}	0.2	3.9×10^{-5}	0.24	5.9×10^{-5}	0.2

basis set	Wu–Yang Hessian Test ^b					
	CD/CD		CD/CQ		CQ/CQ	
	before	after	before	after	before	after
Be	2.3×10^{-7}	8.1×10^{-8}	1.5×10^{-7}	7.0×10^{-8}	1.5×10^{-6}	1.8×10^{-7}
Ne	2.2×10^{-7}	3.2×10^{-8}	3.1×10^{-7}	6.3×10^{-8}	2.6×10^{-6}	1.5×10^{-7}
Ar	4.3×10^{-7}	8.8×10^{-8}	6.6×10^{-7}	1.0×10^{-7}	2.7×10^{-6}	3.2×10^{-7}

^aRelative errors for gradient ($|\text{grad} - \text{grad}_{\text{approximation}}|/|\text{grad}|$); $|\cdot|$ is the L^2 vector norm. ^bRelative errors for Hessian ($|\text{hess} - \text{hess}_{\text{approximation}}|/|\text{hess}|$); ^c“Before” means result before the optimization; “after” means result after the optimization.

good initial guess for $\nu_{\text{KS}}(\mathbf{r})$ is important to find the stationary point of the Lagrangian.

Potential Basis Sets (PBS). There is a general trade-off between the accuracy and efficiency in iKS. Using PBS improves the efficiency significantly. The original WY method⁶⁹ represented $\nu_{\text{KS}}(\mathbf{r})$ as

$$\nu_{\text{KS}}(\mathbf{r}) = \nu_{\text{ext}}(\mathbf{r}) + \nu_0(\mathbf{r}) + \nu_{\text{PBS}}(\mathbf{r}) \quad (11)$$

where $\nu_{\text{ext}}(\mathbf{r})$ is the external potential due to the nuclei and $\nu_0(\mathbf{r})$ is a guide potential, whose role is discussed below. The rest is expanded on a finite potential basis set (PBS) $\{\phi_i\}$

$$\nu_{\text{PBS}}(\mathbf{r}) = \sum_t b_t \phi_t(\mathbf{r}) \quad (12)$$

Compared to a fine mesh, this choice is usually limited at resolving fine features of the XC potentials, but it greatly improves the efficiency. The PDE-CO method has been implemented on systematically improvable basis sets providing remarkable accuracy on polyatomic systems, including the “strongly-correlated” C_6H_4 .⁶⁰ Here we implement PDE-CO for the first time on finite PBS.

In WY, the gradient vector of $W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]$ with respect to the $\{b_i\}$ is

$$\frac{\partial W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]}{\partial b_i} \quad (13a)$$

$$= \int d\mathbf{r}' \frac{\delta W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]}{\delta \nu_{\text{KS}}(\mathbf{r}')} \frac{\partial \nu_{\text{KS}}(\mathbf{r}')}{\partial b_i} \quad (13b)$$

$$= \int d\mathbf{r}' (n(\mathbf{r}') - n_{\text{in}}(\mathbf{r}')) \phi_i(\mathbf{r}') \quad (13c)$$

The Hessian matrix from first-order perturbation theory is

$$\frac{\partial^2 W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]}{\partial b_i \partial b_p} \quad (14a)$$

$$= \int d\mathbf{r}' d\mathbf{r}'' \left(\frac{\delta^2 W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]}{\delta \nu_{\text{KS}}(\mathbf{r}') \delta \nu_{\text{KS}}(\mathbf{r}'')} \frac{\partial \nu_{\text{KS}}(\mathbf{r}')}{\partial b_i} \frac{\partial \nu_{\text{KS}}(\mathbf{r}'')}{\partial b_p} \right) \quad (14b)$$

$$= 2 \sum_i^{\text{occ.}} \sum_a^{\text{unocc.}} \frac{\langle \phi_i | \phi_a \rangle \langle \phi_a | \phi_p \rangle}{\epsilon_i - \epsilon_a} \quad (14c)$$

It can be easily proven that the Hessian matrices are negative definite, which means that $W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]$ is concave for any given PBS $\{\phi_i\}$ as discussed above.

Table 2 shows that both gradient and Hessian, calculated through a finite-difference approximation, are numerically accurate. The tests reported in Table 2 also show where the influence of numerical errors comes in. It can be seen that, before the optimization, the relative errors for the gradient are always small. However, after convergence has been achieved and the analytical gradients are small, the relative gradient errors increase because the absolute numerical errors remain.

For the PDE-CO method, we represent all the terms on finite basis sets. The KS-orbitals $\psi_i(\mathbf{r})$ are represented on a finite OBS, and $\nu_{\text{KS}}(\mathbf{r})$ is represented on PBS as discussed above. We use $\{\phi\}$ to denote the PBS and $\{\phi'\}$ for the OBS. The $p_i(\mathbf{r})$ are expanded on the OBS, just like $\psi_i(\mathbf{r})$:

$$p_i(\mathbf{r}) = \sum_k c_{ik} \phi'_k(\mathbf{r}) \quad (15)$$

To solve for the coefficients c_{ik} , we multiply ϕ'_k on both sides of eq 10a and integrate:

$$(\mathbf{F} - \epsilon_i \mathbf{S}) \mathbf{c}_i = \mathbf{g}_i \quad (16)$$

where $\mathbf{F}$ is the Fock matrix and $\mathbf{S}$ is the overlap matrix $S_{ij} = \langle \phi'_i | \phi'_j \rangle$. The vectors $\mathbf{g}_i$

$$\{\mathbf{g}_i\}_k = \int d\mathbf{r} \phi'_k(\mathbf{r}) [8(n_{\text{in}}(\mathbf{r}) - n(\mathbf{r}))\psi_i(\mathbf{r}) - 2\mu_i\psi_i(\mathbf{r})] \quad (17)$$

are derived from eq 10a. The gradient is derived from eq 10c:

$$\begin{aligned} \frac{dL}{db_i} &= \int d\mathbf{r} \frac{\delta L}{\delta \nu_{\text{KS}}(\mathbf{r})} \phi_i(\mathbf{r}) \\ &= \sum_{j=1}^{N/2} \int d\mathbf{r} p_j(\mathbf{r}) \psi_j(\mathbf{r}) \phi_i(\mathbf{r}) \end{aligned} \quad (18)$$

The well-known limitations of finite basis sets in connection to the ill-posed nature of iKS have stimulated the development of many methods^{38,39,68–70} to improve the results of density-to-potential inversions, though some of them have overlapped effects and they all have shortcomings. To date, there is no clear and straightforward strategy that works reliably. In the following, we explain the theoretical effects of basis sets, regularization, guide potentials, and use of different optimizers; compare the

results; and provide guidance on how to utilize them. In the end, considering both accuracy and efficiency, we provide a recommendation that is most robust (though not perfect) according to our experience and could be directly adopted for practical calculations.

We stress that there are two key aspects for numerical iKS: the accuracy of the densities and the completeness of the PBS. In a scenario in which both are commensurate, one can expect a well-behaved solution for most inversion methods. Unfortunately, this is not easy to achieve as discussed next.

Orbital Basis Sets (OBS). The limitations of finite OBS in density-to-potential inversion are well-documented.^{79–81} Moreover, because atomic orbitals are specifically designed to describe molecular orbitals, there is no guarantee that they can be satisfying for PBS, which is often true for inversion in embedding methods, whose embedding potentials usually have different spatial features compared to molecular orbitals. In practice, the results of both WY and PDE-CO methods can be highly sensitive to both the OBS and the PBS, but there is usually very little freedom in choosing them. There are two main reasons finite OBS can lead to unphysical oscillations in the potential. First, for a given finite OBS $\{\phi_\mu\}$, all the XC potentials that produce the same Fock matrices

$$V_{\mu\nu} = \int \phi_\mu^*(\mathbf{r}) v(\mathbf{r}) \phi_\nu(\mathbf{r}) d\mathbf{r} \quad (19)$$

will lead to the same density. This many-to-one problem is at the root of the relation between the exact external potential and “effective” external potentials discussed by Gaiduk et al.⁶⁸ and Kumar et al.⁵⁵ Second, the electron density cannot be represented exactly on finite Gaussian basis sets.⁸² Small input errors will lead to large oscillations given the ill-posed nature of iKS.^{59,80}

To address the first point above, we find that using similar basis sets, even small ones, for orbital and potential is typically a good choice. There is often error around the nuclei (Figure 1), where both OBS and PBS are insufficiently accurate. On the other hand, even though large basis sets for both the orbital and potential are preferred, this is impractical. A well-adopted strategy is to use a relatively small OBS to reduce the computational cost and to use a large and carefully chosen

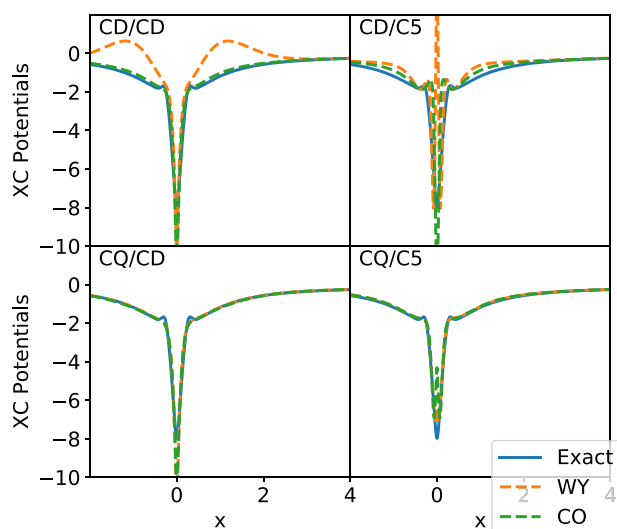


Figure 1. Neon atom XC potential from different combinations of OBS and PBS. The optimizer BFGS⁸³ runs for 30 iterations for all the results.

PBS to have a good representation for the potential. Following this strategy often introduces unphysical oscillations, and thus, regularization methods are required.

To better illustrate the limitations of finite basis sets and the sensitivity of the resulting potentials, we compare in Figure 1 the XC potentials obtained for the Ne atom from WY and PDE-CO methods using different combinations of basis sets. The results are very sensitive to the choice of PBS and OBS. One could imagine that the unphysical features of the resulting XC potentials would depend strongly on the method used to perform the iKS. However, Figure 2 confirms that they are not.

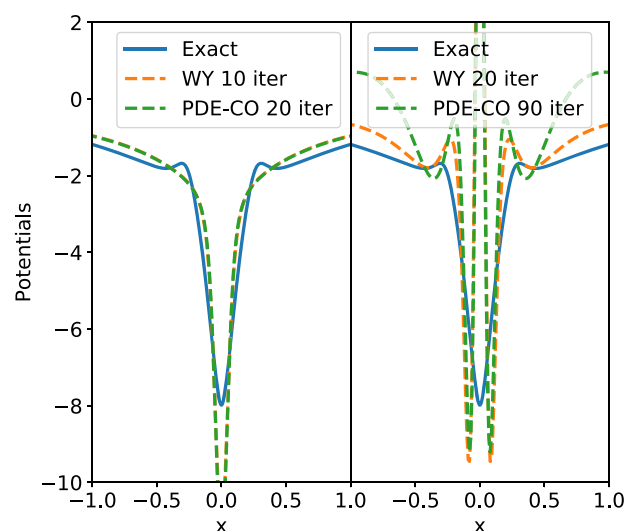


Figure 2. Ne atom XC potentials on CD/C5. BFGS is used with the different number of iterations specified. The same initial point ($\{b_i\} = 0$ and $v_0 = v_{FA}$) is used.

Even though the WY and PDE-CO methods are based on different principles, with optimizers traveling along different paths in the space spanned by the $\{\phi_t\}$, the unphysical features of the resulting XC potentials are largely the same, confirming the dominant role played by the finite basis sets. From this perspective, a systematically improvable basis set^{60,84,85} for both OBS and PBS has the potential to significantly improve the results of iKS.

Guide Potentials. The WY original method⁶⁹ uses the Fermi–Amaldi (FA) potentials^{56,69}

$$v_{FA}(\mathbf{r}) = \frac{N-1}{N} \int d\mathbf{r}' \frac{n_{in}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (20)$$

as a guide potential $v_0(\mathbf{r})$ in eq 11, following Zhao et al.⁵⁶ Because the FA potential can be understood as a Hartree potential $v_H(\mathbf{r}) = \int d\mathbf{r}' \frac{n_{in}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$ with a part from exchange–correlation that partially prevents self-interaction, we also test $v_0(\mathbf{r}) = v_H(\mathbf{r})$ below. Because PBS usually has poor behavior in the asymptotic region, $\sum_t b_t \phi_t(\mathbf{r})$ often decays to 0 too quickly (see Figure 3). In other words, $v_0(\mathbf{r})$ should have the asymptotic behavior of $v_H(\mathbf{r}) + v_{xc}(\mathbf{r})$ because $\sum_t b_t \phi_t(\mathbf{r})$ has negligible contributions far from the nuclei. In the case of LDA densities, $v_0(\mathbf{r}) = v_H(\mathbf{r})$ fits this requirement better. However, the XC potentials obtained using $v_0(\mathbf{r}) = v_{FA}(\mathbf{r})$ usually have an asymptotic behavior that runs much closer to the exact one and leads to better satisfaction of the ionization potential theorem,^{86,87} $\epsilon_{HOMO} = -I$. This is true even when the LDA input density is used.^{50,57}

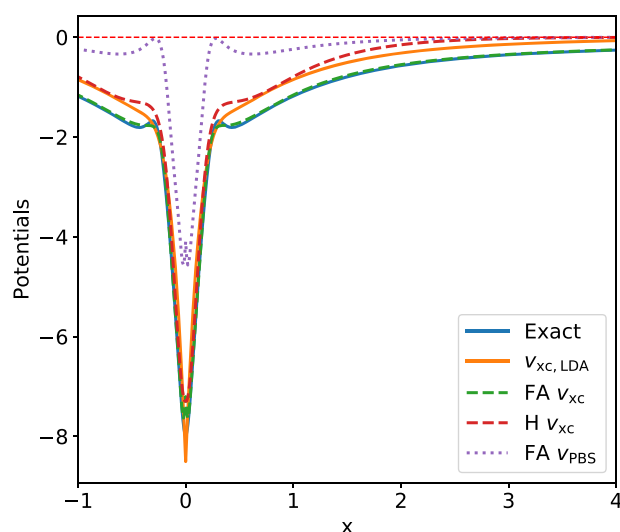


Figure 3. Neon atom XC potential and v_{PBS} (eq 11) for different (Fermi–Amaldi and Hartree) guide potentials (v_0) calculated with WY in C5/C5. H stands for Hartree, and FA stands for Fermi–Amaldi. Note that for a Hartree guide potential, $v_{\text{XC}} = v_{\text{PBS}}$.

Regularization. The effects of regularization are usually significant. The response function in eq 6 and the Hessian matrix in eq 14c are supposed to sum over all unoccupied KS orbitals. The real response functions are invertible, but the Hessian matrices approximated with a finite number of KS orbitals can be very singular, especially when the PBS differs significantly or is larger than the OBS,³⁷ as shown in Figure 4.

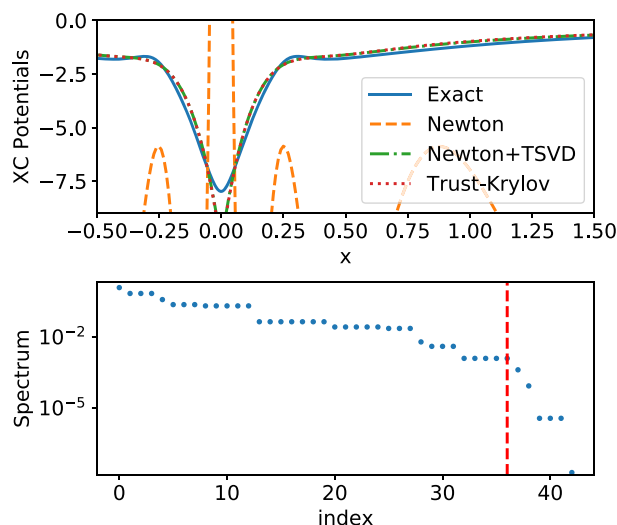


Figure 4. Neon atom XC potential using Newton+TSVD. The top panel shows different results from different optimization methods. The bottom graph is the spectrum of singular values of the Hessian matrix before the convergence. The vertical red dashed line is the TSVD cutoff. CD/CT basis sets are used.

Truncated singular value decomposition (TSVD) regularization is often necessary. Wu and Yang used TSVD³⁸ without providing a systematic way to determine the truncation threshold.⁶⁹ Bulat et al. introduced an additional λ -regularizer

$$-\lambda \|\nabla v_{\text{PBS}}(\mathbf{r})\|^2 \rightarrow -2\lambda \mathbf{b}^T \mathbf{T} \mathbf{b} \quad (21)$$

where $\mathbf{b}$ is the coefficient vector for $v_{\text{PBS}}(\mathbf{r})$ and $\mathbf{T}$ is the kinetic contribution to the Fock matrix.³⁸ The L-curve method was introduced to search for the regularization parameter λ .^{38,39} Both methods are widely used as standard regularization for numerical optimization. They both contribute to the stability of the inversions irrespective of the basis-set employed and help overcome the overfitting problems especially around the nuclei. To demonstrate this, we introduce here two simple but effective tricks: (1) Cliff-plotting to determine the optimum parameters for TSVD and (2) T_s -focusing to optimize λ for λ -regularization.

(1) The truncating parameters for TSVD depend on the systems and basis sets. The most straightforward method to determine these parameters is to plot the spectrum (the diagonal vector of the $\mathbf{S}$ matrix in SVD; see Figure 4). In many cases, the spectrum is of a “cliff” shape. Cutting at the edge of the cliff usually works. The main idea is to eliminate the subspace whose variations do not change the Fock matrices in eq 19. This simple method can be tricky when the OBS is much smaller than the PBS, where the spectrum shows several cliffs or disconnections and one has to decide on which one to cut. On the other hand, when the OBS and PBS are very close or the same, TSVD is often unnecessary. Moreover, optimizers like trust-Krylov⁸⁹ and L-BFGS⁹⁰ yield a result similar to that of the Newton method with TSVD (Figure 4).

(2) Even with TSVD, there can still be overfitting, especially around the nuclei (see Figure 4). We find that λ -regularization is more reliable in this region. In addition to the L-curve method introduced by Bulat et al.,³⁸ we test here an alternative method to search for λ . We calculate $T_s(\lambda)$ as a function of λ (see Figure 5). The point of $T_s(\lambda)$ with the largest λ that is close enough to

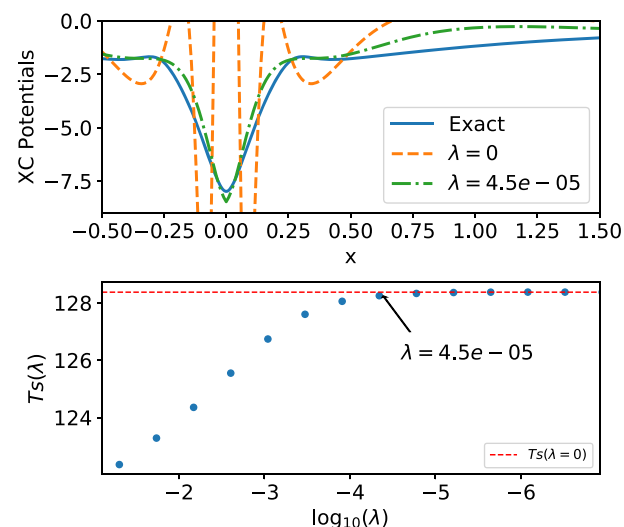


Figure 5. Neon atom XC potentials from the Wu–Yang method with λ -regularization. Newton is used on CD/C5. TSVD is used when $\lambda = 0$.

$T_s(\lambda = 0)$ is chosen. The main idea is to adopt the simplest potential (corresponding to the largest λ) which does not essentially change the optimized result. Obtaining $T_s(\lambda)$ curves such as those of Figures 5 and 6 requires multiple inversion calculations, but the efficiency of constrained optimization methods on finite PBS makes it acceptable in practice.

Motivated by the very similar overfitting features exhibited by the WY and PDE-CO methods, as evidenced in Figure 2, we now add λ -regularization to the Lagrangian of the PDE-CO method (eq 9). Note the sign for the regularization term needs

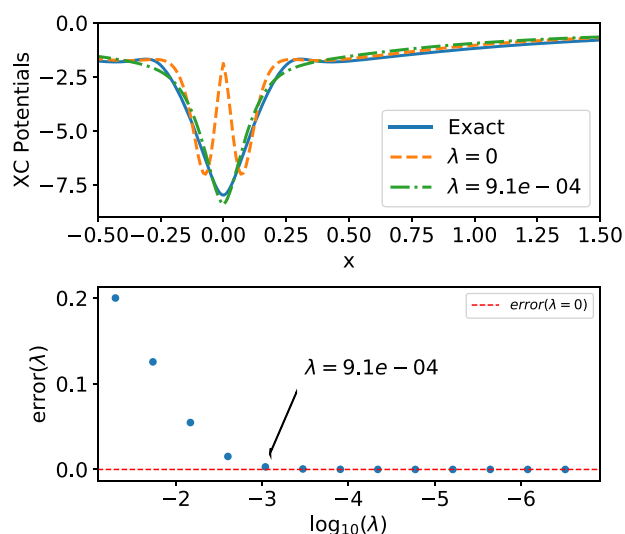


Figure 6. Neon atom XC potentials from the PDE-constrained optimization with λ -regularization. $\text{error}(\lambda)$ is the density error function defined in eq 7. L-BFGS is used on CD/CS.

to be changed when compared to eq 21 because eq 9 is minimized while eq 1 is concave for a given PBS. We follow a similar strategy to search for λ as we have done for WY. Rather than $T_s(\lambda)$, the error (eq 7) as a function of λ is utilized, and the largest λ for a small enough error is selected (Figure 6). In general it is impractical to search for a perfect λ for a specific OBS/PBS. The simple method introduced above usually effectively eliminates the oscillations.

Null-Space Correction. When using TSVD, it is possible that there is not a single clear edge and cutting at one edge does not lead to a smooth result. Also, cutting a Hessian matrix in this way “wastes” the large PBS selected. Instead of λ -regularization, a correction to the Hessian matrix can be defined based on an Unsöld approximation⁹¹ (this follows the derivation given by Gidopoulos and Lathiotakis³⁷ for OEP):

$$\tilde{H}_{mn} = \sum_i \langle \psi_i | \phi_m \phi_n | \psi_i \rangle - \sum_i \sum_{p \in \{i, a_0\}} \langle \psi_i | \phi_m | \psi_p \rangle \langle \psi_p | \phi_n | \phi_i \rangle \quad (22)$$

where the subscript i denotes the occupied KS orbitals and a_0 denotes the finite set of unoccupied KS orbitals. The correction potential on PBS can be defined and derived as

$$\tilde{\mathbf{b}} = \mathbf{V}^\mu (\mathbf{A}^{-1} \mathbf{U}^{\mu T} \tilde{\mathbf{g}} - (\mathbf{U}^{\mu T} \tilde{\mathbf{H}} \mathbf{V}^\mu)^{-1} \mathbf{U}^{\mu T} \tilde{\mathbf{H}} \mathbf{b}_0) \quad (23)$$

A detailed derivation and the definition for each term can be found in the Supporting Information. This correction fixes unphysical oscillations well, as shown in Figure 7.

Optimization Methods. The choice of optimization methods can also be important. Even for the WY method, which is concave, the optimized results can be very different with similar T_s and density error (eq 24) as we mentioned above. Here, we compare the performance of some of the most widely used methods that are available from the Scipy.optimize library:⁹² quasi-Newton (BFGS⁸³ and L-BFGS-B⁹⁰) and trust-region methods (exact⁹³ and Krylov⁸⁹). We also implement Newton with TSVD regularization manually. As Wu and Yang pointed out,⁶⁹ the BFGS does not require TSVD and can often give a better convergence regarding $W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]$, though more iterations are required.

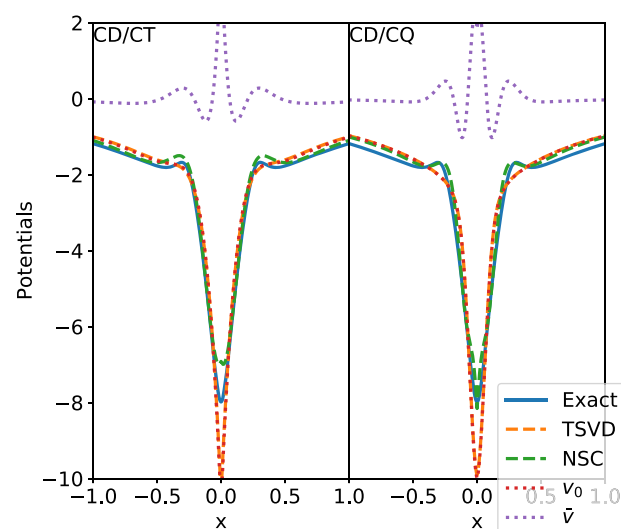


Figure 7. Neon atom potentials in CD/CT (left) and CD/CQ (right). TSVD and null-space corrected (NSC) XC potentials are compared. ν_0 is the original result, and $\tilde{\nu}$ is the correction term (i.e., $\text{NSC} = \nu_0 + \tilde{\nu}$).

We consider four different criteria including $W[\Psi_{\text{det}}[\nu_{\text{KS}}], \nu_{\text{KS}}]$, $T_s[\Psi_{\text{det}}]$, the gradient (eq 13a) norm $\|\text{grad}\|$, and density error

$$N_{\text{error}} = \int d\mathbf{r} |n(\mathbf{r}) - n_{\text{in}}(\mathbf{r})| \quad (24)$$

The gradient norm $\|\text{grad}\|$ can be understood as metrics of optimization on a given PBS. The density error is more general and unlimited by the PBS. The gradient norm optimization is a necessary but insufficient condition for density error optimization because of the limitation of finite PBS (Table 3). The gradient norm is the main criterion used by optimizers for convergence.

Table 4 and Figure 8 illustrate the performance of these different methods on the Ne atom. Trust-exact can still converge to a very overfitted potential when the PBS is not balanced. BFGS does not overcome the ill-condition of the accurate Hessian matrix in extreme cases. L-BFGS-B, originally developed to save the memory for the Hessian matrices and maintaining a lower-rank approximated Hessian matrix, has a better performance than BFGS. Trust region methods and Newton’s method are more efficient than BFGS and L-BFGS-B as one can expect. In all the systems we tested, trust-Krylov is the most robust method. Also, the results suggest severe ill-conditioning problems. Note that PDE-CO has similar levels of accuracy as WY, which is about 1 order of magnitude worse than those achieved by PDE-CO implemented on a finite-element basis.⁶⁰ This confirms again the limitation of finite basis sets and the advantage of using systematically improvable basis.^{60,84,85}

Inverse Crimes. An “inverse crime” occurs when the methods used to calculate the input data are the same as those used to perform the inversion. For example, an inverse crime is committed when one calculates the PBE $\nu_{\text{xc}}(\mathbf{r})$ by inverting a PBE density that has been obtained by solving the forward PBE problem on, e.g., the same grid.⁵⁹ Inverse crimes should be dealt with to avoid trivial results.⁹⁵ This is not a prevalent problem in KS inversions because the input densities are often taken from wave function-based calculations. Nevertheless, we compare the results from a PBE density inversion and a CCSD density inversion as an illustration in Figure 9 for the FC₂Cl molecule.

Table 3. Two Different Norm Comparison at convergence

basis set	CD/STO-3G		CD/CD		CD/CQ	
	$\ \text{grad}\ $	N_{error}	$\ \text{grad}\ $	N_{error}	$\ \text{grad}\ $	N_{error}
Be	7.5×10^{-6}	1.8×10^{-2}	4.8×10^{-5}	1.6×10^{-2}	8.8×10^{-7}	4.2×10^{-6}
Ne	2.7×10^{-6}	3.1×10^{-2}	1.5×10^{-5}	7.8×10^{-4}	5.9×10^{-5}	1.5×10^{-3}
Ar	5.5×10^{-8}	9.2×10^{-2}	2.1×10^{-6}	6.8×10^{-2}	5.2×10^{-5}	4.7×10^{-3}

Table 4. Optimizer Performance for the Ne Atom^a

method	$W[\Psi_{\text{det}} \nu_{\text{KS}}]$	$T_s[\Psi_{\text{det}}]$	$\ \text{grad}\ $ ^b	N_{error}	ϵ_{homo} ^c	N_f ^d
Newton ^e +TSVD	128.609986	128.597	2.4×10^{-2}	4.9×10^{-3}	−0.6280	42
BFGS	128.609810	128.615	1.6×10^{-3}	8.1×10^{-3f}	−0.7452 ^g	12
L-BFGS-B	128.609800	128.614	4.0×10^{-3}	7.5×10^{-3h}	−0.7439 ⁱ	10
trust-exact	128.609948	128.593	1.5×10^{-3}	5.2×10^{-3}	−0.7010	3
trust-Krylov	128.609926	128.609	6.6×10^{-5}	3.0×10^{-3}	−0.6988	4
Newton-CG	128.609823	128.605	2.2×10^{-2}	4.2×10^{-3}	−0.7408	5

^aBasis Set: CT/C5. ^b L^2 . ^cThe experimental I is 0.7925.⁹⁴ ^dThe number of function evaluations. ^eStrong Wolfe line search is used. PDE-CO: ^f 3.5×10^{-3} ; ^g−0.7460; ^h 3.4×10^{-3} ; ⁱ−0.7377.

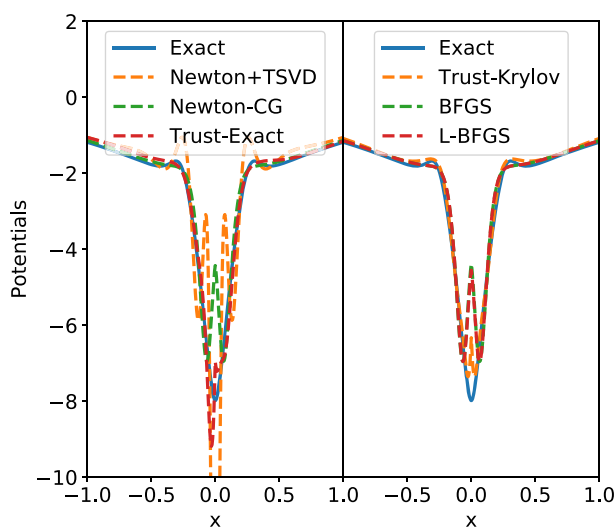


Figure 8. Neon atom XC potentials for common optimization methods in CT/C5.

One can see that the PBE density inversion is much more stable than the CCSD inversion even when no regularization is used. The CCSD inverted potential has unphysical features around the nucleus. However, it clearly attempts to reproduce more details of the reference calculation (mRKS), because CCSD densities are generally more accurate than those of PBE. The inversion from an approximate self-consistent KS density is easier than from a wave function density.

A Final Piece of Advice. Generally (but not strictly), there is a trade-off between accuracy and efficiency in iKS problems. When implemented on finite PBS, with very desirable efficiency, the results are more sensitive to all the factors discussed here. Given our experience, some of which is shown in the [Supporting Information I: Example Calculations](#), we recommend using the largest basis sets you can afford. Trust region methods, especially trust-Krylov with λ -regularization, are usually more reliable (the search for λ could start in a region $[10^{-4}, 10^{-6}]$). If a careful tuning is preferred, the Newton method with the null-space correction discussed in [Supporting Information II: Null-Space Correction](#) can be considered. Our code with all of the methods compared and discussed in this Perspective will be open-sourced very soon, but a word of caution for the user is in order: pure iKS

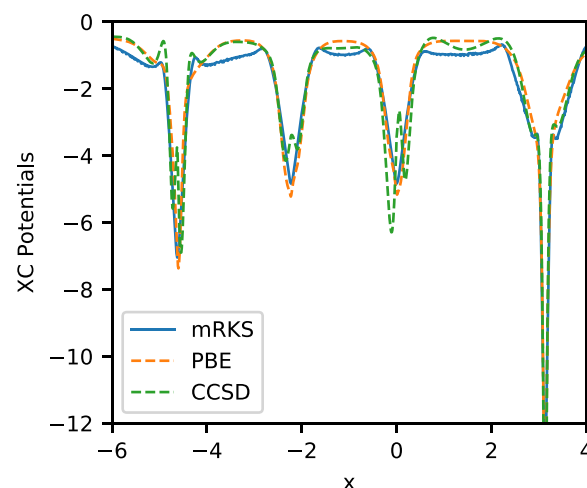


Figure 9. FC₂Cl molecule XC potentials in CD/CQ. The three potentials are mRKS XC potential inverted from a CISD⁹⁶ density, Wu–Yang XC potential inverted from the PBE density, and Wu–Yang XC potential inverted from the CCSD density. The Wu–Yang method with the BFGS optimizer is used for optimization.

problems, especially in finite basis sets and achieved by constrained optimization, remain challenging. New ideas are needed.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpclett.1c00752>.

Supporting Information I: Example Calculations (PDF)

Supporting Information II: Null-Space Correction (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Yuming Shi — Department of Physics and Astronomy, Purdue University, West Lafayette, Indiana 47907, United States; Email: shi449@purdue.edu

Adam Wasserman — Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, United States; Department of Physics and Astronomy and Purdue Quantum Science and Engineering Institute, Purdue University, West

Lafayette, Indiana 47907, United States; orcid.org/0000-0002-8037-4453; Email: awasser@purdue.edu

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jpclett.1c00752>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Grant No. CHE-1900301. The authors thank Cyrus Umrigar for providing the QMC data and Qin Wu and Egor Ospadov for helpful discussions.

REFERENCES

- (1) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864.
- (2) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133.
- (3) Burke, K. Perspective on density functional theory. *J. Chem. Phys.* **2012**, *136*, 150901.
- (4) Verma, P.; Truhlar, D. G. Status and Challenges of Density Functional Theory. *Trends in Chemistry* **2020**, *2*, 302–318.
- (5) Kurth, S.; Perdew, J. P. Role of the exchange-correlation energy: Nature's glue. *Int. J. Quantum Chem.* **2000**, *77*, 814–818.
- (6) Nam, S.; McCarty, R. J.; Park, H.; Sim, E. KS-pies: Kohn-Sham inversion toolkit. *J. Chem. Phys.* **2021**, *154*, 124122.
- (7) Unsleber, J. P.; Dresselhaus, T.; Klahr, K.; Schnieders, D.; Böckers, M.; Barton, D.; Neugebauer, J. Serenity: A subsystem quantum chemistry program. *J. Comput. Chem.* **2018**, *39*, 788–798.
- (8) Runge, E.; Gross, E. K. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997.
- (9) Helbig, N.; Tokatly, I.; Rubio, A. Exact Kohn-Sham potential of strongly correlated finite systems. *J. Chem. Phys.* **2009**, *131*, 224105.
- (10) Hodgson, M. J.; Kraissler, E.; Schild, A.; Gross, E. K. How Interatomic Steps in the Exact Kohn-Sham Potential Relate to Derivative Discontinuities of the Energy. *J. Phys. Chem. Lett.* **2017**, *8*, 5974–5980.
- (11) Thiele, M.; Kummel, S. Photoabsorption spectra from adiabatically exact time-dependent density-functional theory in real time. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4631–4639.
- (12) Fuks, J. I.; Nielsen, S. E. B.; Ruggenthaler, M.; Maitra, N. T. Time-Dependent density functional theory beyond Kohn-Sham Slater determinants. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20976–20985.
- (13) Maitra, N. T.; Zhang, F.; Cave, R. J.; Burke, K. Double excitations within time-dependent density functional theory linear response. *J. Chem. Phys.* **2004**, *120*, 5932–5937.
- (14) Maitra, N. T. Perspective: Fundamental aspects of time-dependent density functional theory. *J. Chem. Phys.* **2016**, *144*, 220901.
- (15) Ramsden, J.; Godby, R. Exact Density-Functional Potentials for Time-Dependent Quasiparticles. *Phys. Rev. Lett.* **2012**, *109*, 036402.
- (16) Mancini, L.; Ramsden, J.; Hodgson, M.; Godby, R. Adiabatic and local approximations for the Kohn-Sham potential in time-dependent Hubbard chains. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *89*, 195114.
- (17) Nam, S.; Song, S.; Sim, E.; Burke, K. Measuring Density-Driven Errors Using Kohn-Sham Inversion. *J. Chem. Theory Comput.* **2020**, *16*, 5014–5023.
- (18) Naito, T.; Ohashi, D.; Liang, H. Improvement of functionals in density functional theory by the inverse Kohn-Sham method and density functional perturbation theory. *J. Phys. B: At., Mol. Opt. Phys.* **2019**, *52*, 245003.
- (19) Accorto, G.; Brandolini, P.; Marino, F.; Porro, A.; Scalesi, A.; Colò, G.; Roca-Maza, X.; Vigezzi, E. First step in the nuclear inverse Kohn-Sham problem: From densities to potentials. *Phys. Rev. C: Nucl. Phys.* **2020**, *101*, 024315.
- (20) Accorto, G.; Naito, T.; Liang, H.; Niksic, T.; Vretenar, D.; et al. Nuclear energy density functionals from empirical ground-state densities. *Phys. Rev. C: Nucl. Phys.* **2021**, *103*, 044304.
- (21) Brockherde, F.; Vogt, L.; Li, L.; Tuckerman, M. E.; Burke, K.; Müller, K.-R. Bypassing the Kohn-Sham equations with machine learning. *Nat. Commun.* **2017**, *8*, 872.
- (22) Nagai, R.; Akashi, R.; Sasaki, S.; Tsuneyuki, S. Neural-network Kohn-Sham exchange-correlation potential and its out-of-training transferability. *J. Chem. Phys.* **2018**, *148*, 241737.
- (23) Kalita, B.; Li, L.; McCarty, R. J.; Burke, K. Learning to Approximate Density Functionals. *Acc. Chem. Res.* **2021**, *54*, 818.
- (24) Manzhos, S. Machine learning for the solution of the Schrödinger equation. *Machine Learning: Science and Technology* **2020**, *1*, 013002.
- (25) Jacob, C. R.; Neugebauer, J. Subsystem density-functional theory. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **2014**, *4*, 325–362.
- (26) Wesolowski, T. A.; Shedge, S.; Zhou, X. W. Frozen-Density Embedding Strategy for Multilevel Simulations of Electronic Structure. *Chem. Rev.* **2015**, *115*, 5891–5928.
- (27) Goodpaster, J. D.; Ananth, N.; Manby, F. R.; Miller, T. F., III Exact nonadditive kinetic potentials for embedded density functional theory. *J. Chem. Phys.* **2010**, *133*, 084103.
- (28) Nafziger, J.; Wu, Q.; Wasserman, A. Molecular binding energies from partition density functional theory. *J. Chem. Phys.* **2011**, *135*, 234101.
- (29) Huang, C.; Pavone, M.; Carter, E. A. Quantum mechanical embedding theory based on a unique embedding potential. *J. Chem. Phys.* **2011**, *134*, 154110.
- (30) Kummel, S.; Kronik, L. Orbital-dependent density functionals: Theory and applications. *Rev. Mod. Phys.* **2008**, *80*, 3–60.
- (31) Levy, M. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v -representability problem. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76*, 6062–6065.
- (32) Perdew, J. P.; Ruzsinszky, A.; Tao, J.; Staroverov, V. N.; Scuseria, G. E.; Csonka, G. I. Prescription for the design and selection of density functional approximations: More constraint satisfaction with fewer fits. *J. Chem. Phys.* **2005**, *123*, 062201.
- (33) Jin, Y.; Yang, Y.; Zhang, D.; Peng, D.; Yang, W. Excitation energies from particle-particle random phase approximation with accurate optimized effective potentials. *J. Chem. Phys.* **2017**, *147*, 134105.
- (34) Jin, Y.; Zhang, D.; Chen, Z. H.; Su, N. Q.; Yang, W. T. Generalized Optimized Effective Potential for Orbital Functionals and Self-Consistent Calculation of Random Phase Approximations. *J. Phys. Chem. Lett.* **2017**, *8*, 4746–4751.
- (35) Flick, J.; Schafer, C.; Ruggenthaler, M.; Appel, H.; Rubio, A. Ab Initio Optimized Effective Potentials for Real Molecules in Optical Cavities: Photon Contributions to the Molecular Ground State. *ACS Photonics* **2018**, *5*, 992–1005.
- (36) Yang, W.; Wu, Q. Direct Method for Optimized Effective Potentials in density-functional Theory. *Phys. Rev. Lett.* **2002**, *89*, 143002.
- (37) Gidopoulos, N. I.; Lathiotakis, N. N. Nonanalyticity of the optimized effective potential with finite basis sets. *Phys. Rev. A: At., Mol., Opt. Phys.* **2012**, *85*, 052508.
- (38) Bulat, F. A.; Heaton-Burgess, T.; Cohen, A. J.; Yang, W. Optimized effective potentials from electron densities in finite basis sets. *J. Chem. Phys.* **2007**, *127*, 174101.
- (39) Heaton-Burgess, T.; Yang, W. Optimized effective potentials from arbitrary basis sets. *J. Chem. Phys.* **2008**, *129*, 194102.
- (40) Aryasetiawan, F.; Stott, M. Effective potentials in density-functional theory. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *38*, 2974.
- (41) Görling, A. Kohn-Sham potentials and wave functions from electron densities. *Phys. Rev. A: At., Mol., Opt. Phys.* **1992**, *46*, 3753.
- (42) Zhao, Q.; Parr, R. G. Quantities $T_s[n]$ and $T_c[n]$ in density-functional theory. *Phys. Rev. A: At., Mol., Opt. Phys.* **1992**, *46*, 2337.

- (43) Zhao, Q.; Parr, R. G. Constrained-search method to determine electronic wave functions from electronic densities. *J. Chem. Phys.* **1993**, *98*, 543–548.
- (44) Wang, Y.; Parr, R. G. Construction of exact Kohn-Sham orbitals from a given electron density. *Phys. Rev. A: At., Mol., Opt. Phys.* **1993**, *47*, R1591.
- (45) Peirs, K.; Van Neck, D.; Waroquier, M. Algorithm to derive exact exchange-correlation potentials from correlated densities in atoms. *Phys. Rev. A: At., Mol., Opt. Phys.* **2003**, *67*, 012505.
- (46) Kadantsev, E. S.; Stott, M. Variational method for inverting the Kohn-Sham procedure. *Phys. Rev. A: At., Mol., Opt. Phys.* **2004**, *69*, 012502.
- (47) Wagner, L. O.; Baker, T. E.; Stoudenmire, E.; Burke, K.; White, S. R. Kohn-Sham calculations with the exact functional. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *90*, 045109.
- (48) Ou, Q.; Carter, E. A. Potential Functional Embedding Theory with an Improved Kohn-Sham Inversion Algorithm. *J. Chem. Theory Comput.* **2018**, *14*, 5680–5689.
- (49) Kumar, A.; Harbola, M. K. Using random numbers to obtain Kohn-Sham potential for a given density. *arXiv* **2020**, 2006.00324.
- (50) Callow, T. J.; Lathiotakis, N. N.; Gidopoulos, N. I. Density-inversion method for the Kohn-Sham potential: Role of the screening density. *J. Chem. Phys.* **2020**, *152*, 164114.
- (51) Kumar, A.; Singh, R.; Harbola, M. K. Universal nature of different methods of obtaining the exact Kohn-Sham exchange-correlation potential for a given density. *J. Phys. B: At., Mol. Opt. Phys.* **2019**, *52*, 075007.
- (52) Kumar, A.; Harbola, M. K. A general penalty method for density-to-potential inversion. *Int. J. Quantum Chem.* **2020**, *120*, e26400.
- (53) Ryabinkin, I. G.; Kohut, S. V.; Staroverov, V. N. Reduction of electronic wave functions to Kohn-Sham effective potentials. *Phys. Rev. Lett.* **2015**, *115*, 083001.
- (54) Ospadov, E.; Ryabinkin, I. G.; Staroverov, V. N. Improved method for generating exchange-correlation potentials from electronic wave functions. *J. Chem. Phys.* **2017**, *146*, 084103.
- (55) Kumar, A.; Singh, R.; Harbola, M. K. Accurate effective potential for density amplitude and the corresponding Kohn-Sham exchange-correlation potential calculated from approximate wavefunctions. *J. Phys. B: At., Mol. Opt. Phys.* **2020**, *53*, 165002.
- (56) Zhao, Q.; Morrison, R. C.; Parr, R. G. From electron densities to Kohn-Sham kinetic energies, orbital energies, exchange-correlation potentials, and exchange-correlation energies. *Phys. Rev. A: At., Mol., Opt. Phys.* **1994**, *50*, 2138.
- (57) Wu, Q.; Ayers, P. W.; Yang, W. Density-functional theory calculations with correct long-range potentials. *J. Chem. Phys.* **2003**, *119*, 2978–2990.
- (58) Nafziger, J.; Jiang, K.; Wasserman, A. Accurate reference data for the nonadditive, noninteracting kinetic energy in covalent bonds. *J. Chem. Theory Comput.* **2017**, *13*, 577–586.
- (59) Jensen, D. S.; Wasserman, A. Numerical methods for the inverse problem of density functional theory. *Int. J. Quantum Chem.* **2018**, *118*, e25425.
- (60) Kanungo, B.; Zimmerman, P. M.; Gavini, V. Exact exchange-correlation potentials from ground-state electron densities. *Nat. Commun.* **2019**, *10*, 4497.
- (61) Hadamard, J. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin* **1902**, 49–52.
- (62) Chayes, J.; Chayes, L.; Ruskai, M. B. Density functional approach to quantum lattice systems. *J. Stat. Phys.* **1985**, *38*, 497–518.
- (63) Ullrich, C.; Kohn, W. Degeneracy in density functional theory: Topology in the v and n spaces. *Phys. Rev. Lett.* **2002**, *89*, 156401.
- (64) Lammert, P. E. Coarse-grained V representability. *J. Chem. Phys.* **2006**, *125*, 074114.
- (65) Kohn, W. v -Representability and density functional theory. *Phys. Rev. Lett.* **1983**, *51*, 1596.
- (66) Li, L.; Hoyer, S.; Pederson, R.; Sun, R.; Cubuk, E. D.; Riley, P.; Burke, K.; et al. Kohn-Sham equations as regularizer: Building prior knowledge into machine-learned physics. *Phys. Rev. Lett.* **2021**, *126*, 036401.
- (67) Nagai, R.; Akashi, R.; Sugino, O. Completing density functional theory by machine learning hidden messages from molecules. *npj Computational Materials* **2020**, *6*, 43.
- (68) Gaiduk, A. P.; Ryabinkin, I. G.; Staroverov, V. N. Removal of basis-set artifacts in Kohn-Sham potentials recovered from electron densities. *J. Chem. Theory Comput.* **2013**, *9*, 3959–3964.
- (69) Wu, Q.; Yang, W. A direct optimization method for calculating density functionals and exchange-correlation potentials from electron densities. *J. Chem. Phys.* **2003**, *118*, 2498–2509.
- (70) Jacob, C. R. Unambiguous optimization of effective potentials in finite basis sets. *J. Chem. Phys.* **2011**, *135*, 244102.
- (71) Vogel, C. R. *Computational Methods For Inverse Problems*; SIAM, 2002.
- (72) Umrigar, C. J.; Gonze, X. Accurate exchange-correlation potentials and total-energy components for the helium isoelectronic series. *Phys. Rev. A: At., Mol., Opt. Phys.* **1994**, *50*, 3827.
- (73) Filippi, C.; Gonze, X.; Umrigar, C. Generalized gradient approximations to density functional theory: comparison with exact results. In *Recent developments and applications of modern density functional theory*; 1996; Vol. 4.
- (74) Parrish, R. M.; Burns, L. A.; Smith, D. G.; Simmonett, A. C.; DePrince, A. E., III; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; et al. Psi4 1.1: An open-source electronic structure program emphasizing automation, advanced libraries, and interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197.
- (75) Smith, D. G.; Burns, L. A.; Sirianni, D. A.; Nascimento, D. R.; Kumar, A.; James, A. M.; Schriber, J. B.; Zhang, T.; Zhang, B.; Abbott, A. S.; et al. Psi4NumPy: An interactive quantum chemistry programming environment for reference implementations and rapid development. *J. Chem. Theory Comput.* **2018**, *14*, 3504–3511.
- (76) Lieb, E. H. Density functionals for coulomb systems. *Int. J. Quantum Chem.* **1983**, *24*, 243–277.
- (77) Wu, Q. Variational nature of the frozen density energy in density-based energy decomposition analysis and its application to torsional potentials. *J. Chem. Phys.* **2014**, *140*, 244109.
- (78) Baydin, A. G.; Pearlmutter, B. A.; Radul, A. A.; Siskind, J. M. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research* **2018**, *18*, 5595–5637.
- (79) Mura, M. E.; Knowles, P. J.; Reynolds, C. A. Accurate numerical determination of Kohn-Sham potentials from electronic densities: I. Two-electron systems. *J. Chem. Phys.* **1997**, *106*, 9659–9667.
- (80) Schipper, P.; Gritsenko, O.; Baerends, E. Kohn-Sham potentials corresponding to Slater and Gaussian basis set densities. *Theor. Chem. Acc.* **1997**, *98*, 16–24.
- (81) Staroverov, V. N.; Scuseria, G. E.; Davidson, E. R. Optimized effective potentials yielding Hartree-Fock energies and densities. *J. Chem. Phys.* **2006**, *124*, 141103.
- (82) Mayer, I.; Pápai, I.; Bako, I.; Nagy, A. Conceptual problem with calculating electron densities in finite basis density functional theory. *J. Chem. Theory Comput.* **2017**, *13*, 3961–3963.
- (83) Fletcher, R. *Practical Methods of Optimization*; John Wiley & Sons, 2013.
- (84) Kanungo, B.; Gavini, V. Large-scale all-electron density functional theory calculations using an enriched finite-element basis. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *95*, 035112.
- (85) Rufus, N. D.; Kanungo, B.; Gavini, V. Fast and robust all-electron density functional theory calculations in solids using orthogonalized enriched finite elements. *arXiv* **2020**, 2010.09701.
- (86) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz, J. L., Jr. Density-functional theory for fractional particle number: derivative discontinuities of the energy. *Phys. Rev. Lett.* **1982**, *49*, 1691.
- (87) Levy, M.; Perdew, J. P.; Sahni, V. Exact differential equation for the density and ionization energy of a many-particle system. *Phys. Rev. A: At., Mol., Opt. Phys.* **1984**, *30*, 2745.
- (88) Hansen, P. C. The truncatedsvd as a method for regularization. *BIT Numerical Mathematics* **1987**, *27*, 534–553.
- (89) Gould, N. I.; Lucidi, S.; Roma, M.; Toint, P. L. Solving the trust-region subproblem using the Lanczos method. *SIAM Journal on Optimization* **1999**, *9*, 504–525.

- (90) Byrd, R. H.; Lu, P.; Nocedal, J.; Zhu, C. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* **1995**, *16*, 1190–1208.
- (91) Unsöld, A. Quantentheorie des Wasserstoffmolekülions und der Born-Landéschen Abstoßungskräfte. *Eur. Phys. J. A* **1927**, *43*, 563–574.
- (92) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272.
- (93) Conn, A. R.; Gould, N. I.; Toint, P. L. *Trust Region Methods*; SIAM, 2000.
- (94) West, R. C.; Astle, M. J. *CRC handbook of chemistry and physics*; CRC Press: Boca Raton, FL, 1987; p. D-71 1979.
- (95) Colton, D.; Kress, R. *Inverse acoustic and electromagnetic scattering theory*; Springer Nature, 2019; Vol. 93.
- (96) Handy, N. C. Multi-root configuration interaction calculations. *Chem. Phys. Lett.* **1980**, *74*, 280–283.