



Guaranteed Convergence of the Kohn-Sham Equations

Lucas O. Wagner,^{1,2} E. M. Stoudenmire,¹ Kieron Burke,^{1,2} and Steven R. White¹

¹*Department of Physics and Astronomy, University of California, Irvine, California 92697, USA*

²*Department of Chemistry, University of California, Irvine, California 92697, USA*

(Received 13 May 2013; revised manuscript received 31 July 2013; published 28 August 2013)

A sufficiently damped iteration of the Kohn-Sham (KS) equations with the exact functional is proven to always converge to the true ground-state density, regardless of the initial density or the strength of electron correlation, for finite Coulomb systems. We numerically implement the exact functional for one-dimensional continuum systems and demonstrate convergence of the damped KS algorithm. More strongly correlated systems converge more slowly.

DOI: [10.1103/PhysRevLett.111.093003](https://doi.org/10.1103/PhysRevLett.111.093003)

PACS numbers: 31.15.E-, 05.10.Cc, 71.15.Mb

Kohn-Sham density functional theory (KS-DFT) [1] is a widely applied electronic structure method. Standard approximate functionals yield accurate ground-state energies and electron densities for many systems of interest [2], but often fail when electrons are strongly correlated. Ground-state properties can be qualitatively incorrect [3], and convergence can be very slow [4,5]. To remedy this, several popular schemes augment Kohn-Sham theory, such as LDA + U [6]. Others seek to improve approximate functionals [7] within the original formulation. But what if the exact functional does not exist for strongly correlated systems? Even if it does, what if the method fails to converge? Either plight would render KS-DFT useless for strongly correlated systems, and render fruitless the vast efforts currently underway to treat, e.g., oxide materials [8], with KS-DFT.

The Kohn-Sham (KS) approach employs a fictitious system of noninteracting electrons, defined to have the same density as the interacting system of interest. The potential characterizing this KS system is unique if it exists [9]. Because the KS potential is a functional of the density, in practice one must search for the density and KS potential together using an iterative, self-consistent scheme [10]. The converged density is in principle the ground-state density of the original, interacting system, whose ground-state energy is a functional of this density.

Motivated by concerns of convergence and existence, we have been performing KS calculations with the exact functional for one-dimensional (1D) continuum systems [11,12]. Even when correlations are strong, we never find a density whose KS potential does not exist, consistent with the results of Ref. [13]. Nor do we find any system where the KS scheme does not converge, although convergence can slow by orders of magnitude as correlation is increased, just as in approximate calculations [4,5].

Exact statements about the unknown density functional inform the construction of all successful density functional theory (DFT) approximations [14–17]. More importantly, they distinguish between what a KS-DFT calculation can *possibly* do, and what it cannot. The most notorious example is the demonstration that the KS band gap of a semiconductor

does not equal the true charge gap, even when the exact functional is used [11,18]. Our key result is an analytic proof that a simple algorithm guarantees convergence of the KS equations for all systems, weakly or strongly correlated, independent of the starting point. Thus multiple stationary points and failures to converge are artifacts of approximate functionals. Studies of convergence are well known in applied mathematics, but almost all concern simple approximations, such as LDA [19], Hartree-Fock [20], etc., and not those in current use in many calculations.

The basic idea lies in a single step of the KS scheme, which proceeds from an input density to produce an output density. For a strongly correlated system as in Fig. 1(a), the output density can differ strongly from the input density, and be further from the true ground-state density. Nevertheless, by proving that the initial slope is always

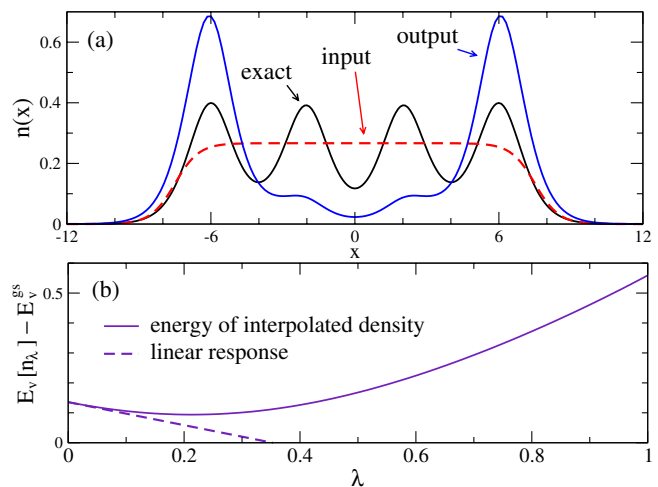


FIG. 1 (color online). (a) The input and output densities for a single step of the Kohn-Sham scheme, as well as the exact density, of a one-dimensional, strongly correlated four atom, four electron system. (b) The energy of the system which interpolates between the input and output densities $E_v[n_\lambda]$, measured from the ground-state energy E_v^{gs} . Also shown is the linear-response approximation with slope given by Eq. (12).

negative as in Fig. 1(b), we show there is always a linear combination of the input and output densities that lowers the energy. By sufficiently damping each KS step, the energy is always reduced each iteration, yielding the ground-state density and energy to within a given tolerance in a finite number of iterations.

The KS algorithm is designed to minimize the energy as a functional of the electron density $n(\mathbf{r})$. For an N -electron system with a reasonable [21] external potential $v(\mathbf{r})$, the energy functional is [1]

$$E_v[n] = T_S[n] + \int d^3r n(\mathbf{r})v(\mathbf{r}) + E_{\text{HXC}}[n], \quad (1)$$

where $T_S[n]$ is the kinetic energy of noninteracting (NI) electrons having density $n(\mathbf{r})$, and $E_{\text{HXC}}[n]$ is the Hartree-exchange-correlation (HXC) energy [22,23]. The KS equations are, in atomic units,

$$-\frac{1}{2}\nabla^2\phi_j(\mathbf{r}) + (v(\mathbf{r}) + v_{\text{HXC}}[n](\mathbf{r}))\phi_j(\mathbf{r}) = \epsilon_j\phi_j(\mathbf{r}), \quad (2)$$

where $v_{\text{HXC}}[n](\mathbf{r}) = \delta E_{\text{HXC}}[n]/\delta n(\mathbf{r})$ is the HXC potential, $\phi_j(\mathbf{r})$ are the electron orbitals, and ϵ_j are their eigenvalues. (In this work, we consider spin-unpolarized systems for simplicity.) An output density $n'(\mathbf{r})$ is found by doubly occupying the lowest-energy orbitals

$$n'(\mathbf{r}) = 2 \sum_{j=1}^{\infty} f_j |\phi_j(\mathbf{r})|^2, \quad (3)$$

where $0 \leq f_j \leq 1$ and $\sum_j f_j = N/2$. Fractional occupation is only allowed for the highest occupied orbitals if they are degenerate, where f_j is chosen to minimize the difference between $n(\mathbf{r})$ and $n'(\mathbf{r})$ [24].

Consider convergence of the following simple algorithm. Given an input density $n(\mathbf{r})$, solve the KS equations to obtain the output density $n'(\mathbf{r})$. Define

$$\eta \equiv \frac{1}{N^2} \int d^3r (n'(\mathbf{r}) - n(\mathbf{r}))^2. \quad (4)$$

Choose some small $\delta > 0$, and if $\eta < \delta$, then the calculation has converged. Otherwise, the next input is

$$n_\lambda(\mathbf{r}) = (1 - \lambda)n(\mathbf{r}) + \lambda n'(\mathbf{r}), \quad (5)$$

for some $\lambda \in (0, 1]$, and repeat. An ensemble- v -representable $n(\mathbf{r})$ is the ground-state density (or an ensemble mixture of degenerate ground-state densities) for some local potential $v[n](\mathbf{r})$ [26,27]. For NI electrons, this potential is $v_S[n](\mathbf{r})$. We call $n(\mathbf{r})$ physical when both potentials exist, and we require all $n_\lambda(\mathbf{r})$ to be physical. We refer to a single iteration of Eqs. (2)–(5) as one step of the KS algorithm. Taking full steps with $\lambda = 1$ does not usually lead to a fixed point. But taking damped steps with $\lambda < 1$ ensures the algorithm converges, as we now prove.

Lemma.—Consider two finite [28] systems of N electrons, with ground-state densities $n(\mathbf{r})$, $n'(\mathbf{r})$, and potentials $v[n](\mathbf{r}) \neq v[n'](\mathbf{r})$, by which we mean the potentials differ by more than a constant. Then [9]

$$\int d^3r (v[n'](\mathbf{r}) - v[n](\mathbf{r}))(n'(\mathbf{r}) - n(\mathbf{r})) < 0. \quad (6)$$

Proof.—Following Ref. [9], we apply the variational principle. Since $n(\mathbf{r})$ is the ground-state density of the potential $v[n](\mathbf{r})$, we have $E_{v[n]}[n] < E_{v[n]}[n']$, or

$$\int d^3r v[n](\mathbf{r})(n(\mathbf{r}) - n'(\mathbf{r})) < F[n'] - F[n], \quad (7)$$

where $F[n] \equiv T_S[n] + E_{\text{HXC}}[n]$. It is also true that $E_{v[n']}[n'] < E_{v[n']}[n]$, so we may switch primes with unprimes in Eq. (7). Adding the resulting equation to the original yields Eq. (6). ■

Note that the lemma is true for any interaction between electrons, including none.

Theorem.—Given an arbitrary physical density $n(\mathbf{r})$ as input into the KS algorithm,

$$E'_v[n] \equiv \left. \frac{dE_v[n_\lambda]}{d\lambda} \right|_{\lambda=0} \leq 0, \quad (8)$$

where $n_\lambda(\mathbf{r})$ is defined as in Eq. (5). If equality holds, then $n(\mathbf{r})$ is a stationary point of $E_v[n]$.

Proof.—Consider ΔE_v resulting from $\lambda \Delta n(\mathbf{r}) \equiv \lambda(n'(\mathbf{r}) - n(\mathbf{r})) = n_\lambda(\mathbf{r}) - n(\mathbf{r})$. Then

$$E'_v[n] = \int d^3r \frac{\delta E_v[n]}{\delta n(\mathbf{r})} \Delta n(\mathbf{r}). \quad (9)$$

For a physical density, the functional derivative is [27]

$$\frac{\delta E_v[n]}{\delta n(\mathbf{r})} = -v_S[n](\mathbf{r}) + v(\mathbf{r}) + v_{\text{HXC}}[n](\mathbf{r}). \quad (10)$$

Since $v(\mathbf{r}) + v_{\text{HXC}}[n](\mathbf{r})$ defines $v_S[n'](\mathbf{r})$ [$n'(\mathbf{r})$ is the output density of Eq. (2)], we have

$$\frac{\delta E_v[n]}{\delta n(\mathbf{r})} = v_S[n'](\mathbf{r}) - v_S[n](\mathbf{r}). \quad (11)$$

Combining Eqs. (11) and (9) gives

$$E'_v[n] = \int d^3r (v_S[n'](\mathbf{r}) - v_S[n](\mathbf{r}))(n'(\mathbf{r}) - n(\mathbf{r})). \quad (12)$$

Two cases arise: if $v_S[n'](\mathbf{r}) \neq v_S[n](\mathbf{r})$, use the lemma applied to NI systems: then $E'_v[n]$ must be less than zero. Otherwise, $v_S[n'](\mathbf{r}) = v_S[n](\mathbf{r})$, so both $E'_v[n]$ and the rhs of Eq. (11) are zero, and $n(\mathbf{r})$ is a stationary point of $E_v[n]$. ■

We illustrate the theorem in Fig. 1(b), where we plot $E_v[n_\lambda]$ and its linear-response approximation for the input density of Fig. 1(a).

Corollary 1.—The KS algorithm described above is guaranteed to converge to a stationary point of the functional, if (1) only physical densities are encountered, (2) the energy functional is convex, and (3) appropriate values for λ are used, e.g., from the algorithm of Ref. [29], because it is effectively a gradient-descent algorithm [30].

Corollary 2.—When using the exact functional, the KS algorithm using appropriate λ 's converges to the exact

ground-state density, as long as the first input density is a physical density. This is because we can choose each subsequent input density as a physical density [31], and the exact ensemble functional [22,33] is convex. The only stationary point of the exact functional, when considering physical densities, is the ground-state density [34].

Numerical implementation.—To find the KS energy functional exactly when there is no degeneracy, we must find the many-electron wave function $\Psi[n]$ that minimizes $\langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle$ (the kinetic and electron-electron repulsion energies) with density $n(\mathbf{r})$ [22,35]. To perform this very demanding [36] interacting inversion, start with a guess for the potential $\tilde{v}(\mathbf{r})$. Then solve the many-body system for the ground-state wave function $\tilde{\Psi}$ and density $\tilde{n}(\mathbf{r})$. Using a quasi-Newton method [37], modify $\tilde{v}(\mathbf{r})$ and repeat, minimizing the difference between $\tilde{n}(\mathbf{r})$ and the target density $n(\mathbf{r})$. Once converged, the procedure is repeated for NI electrons. The HXC energy is then

$$E_{\text{HXC}}[n] = \langle \Psi[n] | \hat{T} + \hat{V}_{ee} | \Psi[n] \rangle - T_S[n], \quad (13)$$

and the HXC potential is

$$v_{\text{HXC}}[n](\mathbf{r}) = v_S[n](\mathbf{r}) - v[n](\mathbf{r}). \quad (14)$$

We implement these functionals for 1D continuum systems [11,12], obtaining highly accurate many-body solutions with the density matrix renormalization group [38,39]. These are the first such inversions for systems with more than two electrons [40,41]. Because, in one dimension, degeneracy (beyond spin) does not occur, we find pure states $\Psi[n]$. More generally, one should invert using an ensemble $\Gamma[n]$ and take a trace in Eq. (13) [22,33].

To illustrate convergence of the damped KS algorithm using the exact functional, we plot the output densities and KS potentials for a four-electron, four-atom system in Fig. 2. We choose the interatomic spacing $R = 3$ to be

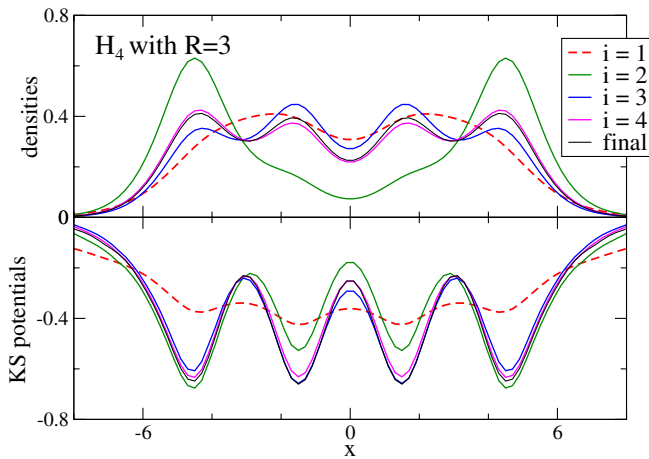


FIG. 2 (color online). KS procedure for a moderately correlated four-electron system (four hydrogen atoms with $R = 3$), showing the first few iterations. Using a fixed $\lambda = 0.30$, we converge to $\eta < 10^{-6}$ using Eq. (4) within 13 iterations.

roughly twice the equilibrium spacing of H_2 (when the interaction between nuclei is the same as that between electrons), making this a moderately correlated system. Taking $\lambda = 0.30$, the algorithm converges to the exact density (computed separately using DMRG) to $\eta < 10^{-6}$ using Eq. (4), within 13 steps.

Consider the KS scheme applied to a simple 1D H_2 molecule with bond length R [12]. Initialize the algorithm with an asymmetric input density, a H^- density centered on the left atom. Of course, no sensible KS calculation starts with such a density, but we do this to amplify convergence issues. In Fig. 3, we quantify the convergence of the KS algorithm using η from Eq. (4) as well as energy differences from the ground state. For the equilibrium bond length ($R = 1.6$), λ may be chosen quite large (≈ 0.5), but as the atoms are stretched to $R = 3$, λ must be ≤ 0.2 . When $R = 5$, even $\lambda = 0.01$ is too large to converge the calculation (not shown). Thus, as the bond is stretched and the system develops strong static correlation [12], convergence becomes increasingly difficult. As more atoms are added to the chain (not shown), such as stretched H_4 , even a reasonable initial state converges very slowly.

Consequences for real calculations.—For approximate XC functionals, the corresponding $E_v[n]$ is not, in general, convex for every $v(\mathbf{r})$, and our corollaries do not hold. Consider H_2 in the local spin-density approximation. At and near equilibrium bond lengths, only one stationary solution exists. The approximate functional may or may not be convex. But when the bond is stretched beyond the infamous Coulson-Fischer point [42,43], an unrestricted

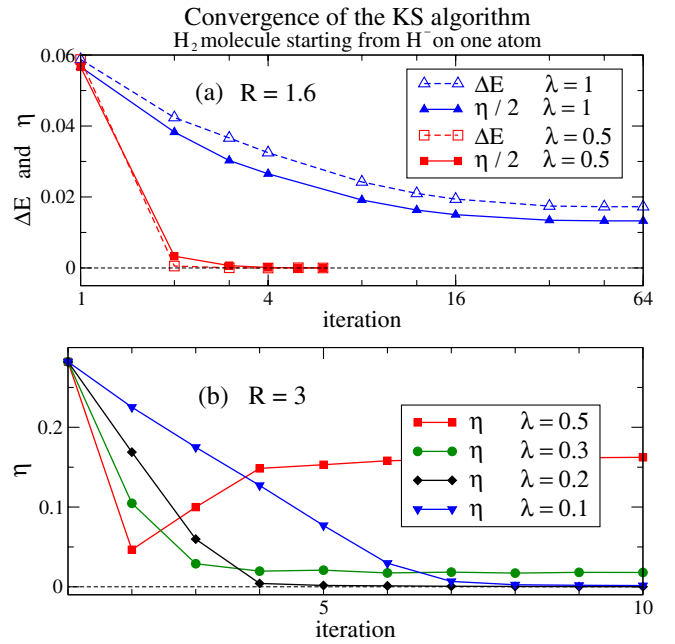


FIG. 3 (color online). Differences in the density η using Eq. (4) and the energy with $\Delta E = E_v[n'] - E_v^{\text{gs}}$, for an H_2 molecule with (a) $R = 1.6$ and (b) $R = 3$. In (b), the ΔE curves are omitted for clarity, but are like those in (a).

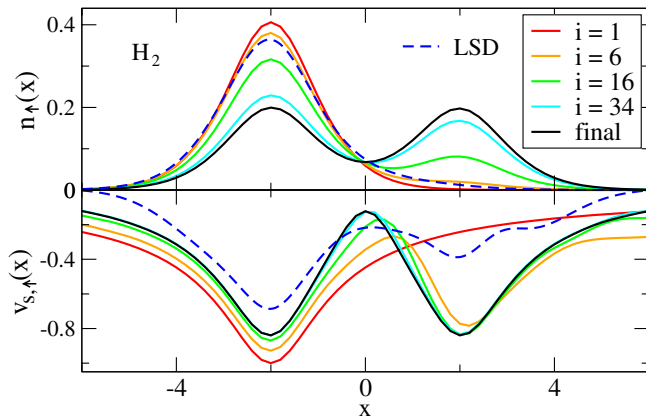


FIG. 4 (color online). Starting an exact KS calculation of stretched H_2 with a spin-polarized density still converges (with $\lambda = 0.5$) to the correct spin-singlet density. For the same initial density, the KS calculation with the local spin-density (LSD) approximation [47] converges to the broken spin-symmetry solution shown.

solution of lower energy appears [12], as in Fig. 4, so the corresponding $E_v[n]$ is not convex and convergence with our simple algorithm is not guaranteed. While the restricted solution is a saddle point, the unrestricted solution is a local minimum. Thus, only the unrestricted solution behaves locally like the solution with the exact functional, providing further rationale [43] for preferring such a solution over any restricted one. On the other hand, slowing of convergence as correlations become stronger is a real effect, and not an artifact of approximations.

We chose our simple algorithm to prove convergence, but many are more sophisticated and efficient (see, e.g., Refs. [4,5]). Mixing KS potentials instead of densities [44] can similarly be proven to converge, with the advantage that all densities encountered are NI v representable.

Finally, we expect that orbital degeneracies in three dimensions require the ensemble treatment [22,25,26,33]. Further, extending the KS approach to use fractional occupation of electron orbitals (even in the case of nondegeneracy) may speed convergence [45] and allow KS-DFT to more naturally handle strong static correlation [46].

The authors acknowledge U.S. DOE Award No. DE-SC0008696, and L. O. W. also thanks the Korean Global Research Network Grant (No. NRF-2010-220-C00017). L. O. W. thanks Stefano Pittalis, Marcos Raydan, Robert van Leeuwen, and Mel Levy for helpful discussions.

- [1] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
 [2] K. Burke, *J. Chem. Phys.* **136** 150901 (2012).
 [3] P. Mori-Sánchez, A. J. Cohen, and W. Yang, *Phys. Rev. Lett.* **102**, 066403 (2009).
 [4] A. D. Daniels and G. E. Scuseria, *Phys. Chem. Chem. Phys.* **2**, 2173 (2000).

- [5] L. Thogersen, J. Olsen, D. Yeager, P. Jorgensen, P. Salek, and T. Helgaker, *J. Chem. Phys.* **121**, 16 (2004).
 [6] V. I. Anisimov, F. Aryasetiawan, and A. I. Lichtenstein, *J. Phys. Condens. Matter* **9**, 767 (1997).
 [7] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **118**, 8207 (2003); **124**, 219906(E) (2006).
 [8] E. Assmann, P. Blaha, R. Laskowski, K. Held, S. Okamoto, and G. Sangiovanni, *Phys. Rev. Lett.* **110**, 078701 (2013).
 [9] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
 [10] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer-Verlag, Berlin, 1990).
 [11] E. M. Stoudenmire, L. O. Wagner, S. R. White, and K. Burke, *Phys. Rev. Lett.* **109**, 056402 (2012).
 [12] L. O. Wagner, E. M. Stoudenmire, K. Burke, and S. R. White, *Phys. Chem. Chem. Phys.* **14**, 8581 (2012).
 [13] J. T. Chayes, L. Chayes, and M. Ruskai, *J. Stat. Phys.* **38**, 497 (1985).
 [14] M. Levy and J. P. Perdew, *Phys. Rev. A* **32**, 2010 (1985).
 [15] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996); **78**, 1396(E) (1997).
 [16] J. P. Perdew and S. Kurth, in *A Primer in Density Functional Theory*, edited by C. Fiolhais, F. Nogueira, and M. A. L. Marques (Springer-Verlag, Berlin, 2003), pp. 1–55.
 [17] J. P. Perdew, A. Ruzsinszky, J. Tao, V. N. Staroverov, G. E. Scuseria, and G. I. Csonka, *J. Chem. Phys.* **123**, 062201 (2005).
 [18] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, *Phys. Rev. Lett.* **49**, 1691 (1982).
 [19] A. Anantharaman and E. Cancès, *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis* **26**, 2425 (2009).
 [20] E. Cancès and C. Le Bris, *ESAIM: Math. Modell. Numer. Anal.* **34**, 749 (2000).
 [21] The precise restrictions on potentials are detailed in Ref. [22]; Coulomb potentials are allowed.
 [22] E. H. Lieb, *Int. J. Quantum Chem.* **24**, 243 (1983).
 [23] C. Fiolhais, F. Nogueira, and M. Marques, *A Primer in Density Functional Theory* (Springer, New York, 2003).
 [24] An orbital rotation among degenerate orbitals may also be required; see Ref. [25].
 [25] C. A. Ullrich and W. Kohn, *Phys. Rev. Lett.* **87**, 093001 (2001).
 [26] M. Levy, *Phys. Rev. A* **26**, 1200 (1982).
 [27] R. van Leeuwen, *Density Functional Approach to the Many-Body Problem: Key Concepts and Exact Functionals* (Academic, New York, 2003), pp. 25–94.
 [28] We restrict our attention to finite systems, since in an extended system Eq. (6) would be ill defined.
 [29] E. G. Birgin, J. Mario Martinez, and M. Raydan, *IMA J. Numer. Anal.* **23**, 539 (2003).
 [30] With respect to gradient descent minimization, we assume that functionals behave like their less dimensional counterparts, functions. A more comprehensive analysis is beyond the scope of this work.
 [31] In the language of DFT, we can find an ensemble- v -representable density which is arbitrarily close to a density $\bar{n}(\mathbf{r})$ which has reasonable properties: integrating to a finite particle number N , with a finite von Weizsäcker kinetic energy, i.e., $\sqrt{\bar{n}(\mathbf{r})} \in H^1(\mathbb{R}^3)$, and

- being nowhere negative [22,27]. This set of reasonable densities is convex [27], so that $n_\lambda(\mathbf{r})$ is reasonable if both $n(\mathbf{r})$ and $n'(\mathbf{r})$ are, which is the case. Some examples of reasonable, but non- v -representable densities are found in Ref. [32].
- [32] H. Englisch and R. Englisch, *Physica (Amsterdam)* **121A**, 253 (1983).
- [33] S.M. Valone, *J. Chem. Phys.* **73**, 4653 (1980).
- [34] J.P. Perdew and M. Levy, *Phys. Rev. B* **31**, 6264 (1985).
- [35] M. Levy, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 6062 (1979).
- [36] N. Schuch and F. Verstraete, *Nat. Phys.* **5**, 732 (2009).
- [37] C.G. Broyden, *Math. Comput.* **19**, 577 (1965).
- [38] S.R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [39] S.R. White, *Phys. Rev. B* **48**, 10345 (1993).
- [40] M. Thiele, E.K.U. Gross, and S. Kümmel, *Phys. Rev. Lett.* **100**, 153004 (2008).
- [41] J.P. Coe, K. Capelle, and I. D'Amico, *Phys. Rev. A* **79**, 032504 (2009).
- [42] C.A. Coulson and I. Fischer, *Philos. Mag. Series 7* **40**, 386 (1949).
- [43] J.P. Perdew, A. Savin, and K. Burke, *Phys. Rev. A* **51**, 4531 (1995).
- [44] G.P. Kerker, *Phys. Rev. B* **23**, 3082 (1981).
- [45] A.D. Rabuck and G.E. Scuseria, *J. Chem. Phys.* **110**, 695 (1999).
- [46] P.R.T. Schipper, O.V. Gritsenko, and E.J. Baerends, *Theor. Chem. Acc.* **99**, 329 (1998).
- [47] N. Helbig, J.I. Fuks, M. Casula, M.J. Verstraete, M.A.L. Marques, I.V. Tokatly, and A. Rubio, *Phys. Rev. A* **83**, 032503 (2011).