

**Guaranteed Convergence of a Regularized Kohn-Sham Iteration in Finite Dimensions**

Markus Penz\*

*Max Planck Institute for the Structure and Dynamics of Matter, 22761 Hamburg, Germany*

Andre Laestadius

*Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, University of Oslo, 0315 Oslo, Norway*

Erik I. Tellgren

*Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, University of Oslo, 0315 Oslo, Norway*

Michael Ruggenthaler

*Max Planck Institute for the Structure and Dynamics of Matter, 22761 Hamburg, Germany*

(Received 22 March 2019; published 19 July 2019)

The exact Kohn-Sham iteration of generalized density-functional theory in finite dimensions with a Moreau-Yosida regularized universal Lieb functional and an adaptive damping step is shown to converge to the correct ground-state density.

DOI: [10.1103/PhysRevLett.123.037401](https://doi.org/10.1103/PhysRevLett.123.037401)

The Kohn-Sham (KS) scheme [1] of ground-state density-functional theory (DFT) is the cornerstone of electronic structure calculations in quantum chemistry and solid-state physics [2]. It maps a complicated system of interacting electrons onto an auxiliary, noninteracting KS system. This yields a set of coupled one-particle equations that need to be solved self-consistently. Since a direct solution is unfeasible, practical approaches are variations of self-consistent field methods taking the form of fixed-point iterations or energy minimization algorithms [3–8]. To date, no method has been rigorously shown to converge to the correct ground-state density. Convergence results for approximate schemes are available for auxiliary assumptions [9], and reliably achieving convergence in systems with small band gaps or for transition metals remains a hard practical challenge [10]. Approximation techniques face the problem of an exponential growth of local minima with increasing number of particles [11]. Such local minima appear as “false” solutions in the energy landscape and distract from the global, absolute minimum [12]. Hence, a method with mathematically guaranteed convergence to the correct minimizer is of central importance and has been listed as one of the twelve outstanding problems in DFT [13].

An early insight is that iterations commonly fail unless oscillations between trial states are damped [3]. Work by Cancès and Le Bris [14,15] led to the optimal damping algorithm based on energy minimization by line search along the descent direction. Wagner *et al.* [16,17] presented a similar scheme and claimed to have proven convergence in the setting of exact DFT, while only the strict descent of energies was secured. In such efforts, functional differentiability is almost always tacitly assumed or wrongly

claimed, prominently in Ref. [18], Eq. (2.105), while the underlying universal functionals are known to be non-differentiable [19]. This means the usual presentations of DFT already *assume* some form of regularization of the functionals. Other special forms of DFT like with internal magnetic fields [20] or finite temperatures [21,22] automatically include regularization effects.

This issue was addressed in Laestadius *et al.* [23], where a similar iterative scheme was proposed that proved a weak type of convergence after Moreau-Yosida (MY) regularization to ensure differentiability of the universal Lieb functional [24]. Weak-type convergence here means that the energy converges to either the correct energy or an upper bound. MY regularization has been introduced to DFT by Kvaal *et al.* [25].

A rich study of possible strategies for self-consistent field iteration was recently put forward by Lammert [26]. Yet in all those works the question of a limit density and corresponding KS potential was left open. On the other hand, the result in Laestadius *et al.* [23] is applicable to not only standard DFT, but to all DFT flavors that fit into the given framework of reflexive Banach spaces. It has already been successfully applied to paramagnetic current DFT (CDFT) [27]. This general approach is also pursued in this Letter.

In what follows we give a fully rigorous proof of convergence for the KS scheme in a finite-dimensional state space. Because of the techniques involved, the new iteration scheme was baptized the Moreau-Yosida Kohn-Sham optimal damping algorithm (MYKSODA) by Laestadius *et al.* [27]. The employed damping critically depends on MY regularization that bounds the curvature of the universal Lieb functional from above.

We will now present the mathematical framework. For a much more detailed discussion of generalized KS schemes in Banach spaces that can also be infinite dimensional, we refer to Laestadius *et al.* [23]. The spaces for densities and potentials are chosen to be the Hilbert space  $X = X^* = \ell^2(M)$ ,  $M \in \mathbb{N}$ , which corresponds to a finite one-particle basis, a lattice system with  $M$  sites, or many other possible settings. The reason for this dual choice of spaces is how densities and potentials couple in the energy expression. What is denoted  $\int_{\mathbb{R}^3} v \rho dx$  in standard DFT is a finite sum in the given setting and will further be written  $\langle v, \rho \rangle$  with  $\rho \in X$ ,  $v \in X^*$ . For the internal energy of the *full system*, a universal functional  $\tilde{F}$ , like the one defined by constrained search [24,28] over all  $N$ -particle density matrices  $\Gamma$  that yield a given density  $\rho \in X$ , is introduced:

$$\tilde{F}(\rho) = \inf_{\Gamma \rightarrow \rho} \{\text{Tr}[(H_{\text{kin}} + H_{\text{int}})\Gamma]\}. \quad (1)$$

Here  $H_{\text{kin}}$  stands for the kinetic energy and  $H_{\text{int}}$  for interactions. Consequently, the functional  $\tilde{F}$  is defined on a set  $\tilde{X} \subset X$  of *physical densities* that come from an  $N$ -particle density matrix (ensemble  $N$ -representability). This set  $\tilde{X}$  will be assumed bounded in  $X$ . Since all physical densities are normalized in the  $\ell^1$  norm and all norms are equivalent in finite dimensions, this follows naturally. It also holds for CDFT on a finite lattice, since the current density is bounded by the hopping parameter [Ref. [29], Eq. (25)], and for one-body reduced density matrix functional theory in finite basis sets, since the off-diagonal elements of the reduced density matrix are bounded by the diagonal ones that give the usual density [Ref. [22], Eq. (3.49)].

On the other hand, elements in  $X$  will in general not constitute physical densities. In standard DFT this means that an arbitrary  $x \in X$  does not have to be normalized or even positive. Such an  $x \in X$  will thus be called a *quasi-density*. We reserve the notation  $\rho$  for physical densities.

The total energy is the infimum of  $\tilde{F}(\rho)$  plus the potential energy coming from a given external potential  $v \in X^*$ , taken over all physical densities:

$$E(v) = \inf_{\rho \in \tilde{X}} \{\tilde{F}(\rho) + \langle v, \rho \rangle\}. \quad (2)$$

It is linked to a functional  $F$  on  $X$  by the Legendre-Fenchel transformation (convex conjugate). Then  $F$  can be transformed back to the same  $E$  as

$$F(x) = \sup_{v \in X^*} \{E(v) - \langle v, x \rangle\}, \quad (3)$$

$$E(v) = \inf_{x \in \tilde{X}} \{F(x) + \langle v, x \rangle\}. \quad (4)$$

The functional  $F$  is by construction convex and lower semicontinuous (Ref. [24], Theorem 3.6) and has  $F(x) = +\infty$  whenever  $x$  is not in the domain  $\tilde{X}$  of  $\tilde{F}$ . Minimizers of

Eq. (2) are the ground-state densities, which establishes a link to the Schrödinger equation. They stay the same if one switches from  $\tilde{F}$  to  $F$ , and thus minimizers of Eq. (4) are always in  $\tilde{X}$ . Finding such minimizers  $\rho$  of Eq. (4) is equivalent to determining the superdifferential of  $E$ , i.e., the set of functionals in  $X^{**} = X$  that yield a graph completely above  $E$ , written  $\rho \in \bar{\partial}E(v) \subset \tilde{X}$ .

The MY regularization of the functional  $F$  on  $X$  is defined as

$$F_\varepsilon(x) = \inf_{y \in \tilde{X}} \left\{ F(y) + \frac{1}{2\varepsilon} \|x - y\|^2 \right\}. \quad (5)$$

The visual understanding of this is the following. As the vertex of the regularization parabola  $(1/2\varepsilon)\|x\|^2$  moves along the graph of  $F$ , the regularized  $F_\varepsilon$  is given by the traced out lower envelope (the ‘‘Moreau envelope,’’ Ref. [30], Definition 1.22). It also means the regularization puts an upper bound of  $\varepsilon^{-1}$  on the (positive) curvature of  $F_\varepsilon$ . This will be an important ingredient in the convergence proof: A bound on the curvature means the convex functional  $F_\varepsilon$  cannot change from falling to rising too quickly, yielding a secure bound on the possible step length for descent.

The regularized  $F_\varepsilon$  is then differentiable and even has a continuous gradient  $\nabla F_\varepsilon$  (Fréchet differentiability) (Ref. [23], Theorem 9), something that will also become important in the convergence proof. We define the associated energy functional

$$E_\varepsilon(v) = \inf_{x \in \tilde{X}} \{F_\varepsilon(x) + \langle v, x \rangle\}. \quad (6)$$

The functional in Eq. (6) is *not* the MY regularization of  $E$  but the Legendre-Fenchel transformation of  $F_\varepsilon$ . If  $z \in X$  is a minimizer in Eq. (6), called the ground-state quasidensity, then the gradient of  $F_\varepsilon + v$  at  $z$  must be zero:

$$\nabla F_\varepsilon(z) + v = 0. \quad (7)$$

Since the regularized functional is differentiable everywhere, the usual problem of  $v$ -representability is avoided. For  $E_\varepsilon$ , which is still not differentiable, we can resort to the superdifferential. Since any such element  $z \in \bar{\partial}E_\varepsilon(v)$  automatically solves Eq. (7), it is the ground-state quasidensity of the regularized problem with potential  $v$ . Two important properties of  $E_\varepsilon$  are (Ref. [23], Theorem 10 and Corollary 11)

$$E(v) = E_\varepsilon(v) + \frac{\varepsilon}{2} \|v\|^2, \quad (8)$$

$$\bar{\partial}E(v) = \bar{\partial}E_\varepsilon(v) + \varepsilon v \subset \tilde{X}, \quad (9)$$

which relate the regularized problem back to the unregularized one. Since  $E$  is already concave, the subtraction of a

parabola in Eq. (8) makes  $E_\varepsilon$  strongly concave. Note that the  $\varepsilon$  in Eq. (9) takes a role comparable to that of *permittivity*, linking potentials to densities.

To set up a KS scheme we define a *reference system* that is noninteracting by

$$\tilde{F}^0(\rho) = \inf_{\Gamma \rightarrow \rho} \{\text{Tr}(H_{\text{kin}}\Gamma)\} \quad (10)$$

on the same  $\tilde{X} \subset X$  and define  $E^0$ ,  $F_\varepsilon^0$ , and  $E_\varepsilon^0$  analogously. The analogue of Eq. (7) for  $F_\varepsilon^0$  at the same quasidensity  $z \in X$  is  $\nabla F_\varepsilon^0(z) + v_{\text{KS}}^\varepsilon = 0$  and defines the KS potential  $v_{\text{KS}}^\varepsilon$ . Simply equating this equation and Eq. (7) gives

$$\nabla F_\varepsilon(z) + v = \nabla F_\varepsilon^0(z) + v_{\text{KS}}^\varepsilon, \quad (11)$$

where the ground-state quasidensity  $z$  and the auxiliary  $v_{\text{KS}}^\varepsilon$  for the reference system are still unknown and neither  $F_\varepsilon$  nor  $F_\varepsilon^0$  have a simple, explicit expression. The trick is to determine  $z$  and  $v_{\text{KS}}^\varepsilon$  in an iterative algorithm by replacing them with sequences  $x_i \rightarrow z$ ,  $v_i \rightarrow v_{\text{KS}}^\varepsilon$ . The indicated convergence is our major concern in the following proof. We get an update rule for the potential sequence  $(v_i)_i$  directly from Eq. (11),

$$v_{i+1} = v + \nabla F_\varepsilon(x_i) - \nabla F_\varepsilon^0(x_i), \quad (12)$$

and determine the next quasidensity by solving for the ground state of the regularized reference system with  $v_{i+1}$ . This iteration has the stopping condition  $v_{i+1} = -\nabla F_\varepsilon^0(x_i)$ , which means  $v = -\nabla F_\varepsilon(x_i)$  by Eq. (12). Then  $x_i$  is already the sought-after ground-state quasidensity  $z$  and thus also  $v_{i+1} = v_{\text{KS}}^\varepsilon$  is the respective KS potential that yields the same quasidensity for the reference system.

The most important ingredient of *practical* KS calculations enters by giving suitable approximations for the expression  $\nabla F_\varepsilon - \nabla F_\varepsilon^0$  (Hartree-exchange-correlation potential including the correlated kinetic energy). For the purpose of showing convergence it is not crucial that this object comes from the exact functional or that it is the result of an approximation, as long as  $F_\varepsilon$ ,  $F_\varepsilon^0$  have the stated properties.

The MYKSODA algorithm is then the following. In step (a) get the new potential by Eq. (12) above. In step (b) solve the (simpler) ground-state problem for the reference system by choosing the next quasidensity from  $\bar{\partial}E_\varepsilon^0(v_{i+1})$ . From Eq. (9) it follows that the set of quasidensities  $\bar{\partial}E_\varepsilon^0(v_{i+1})$  can be determined from the set of ground-state densities of the reference system, which means solving the noninteracting Schrödinger equation. Finally, to ensure a strictly descending energy and to show convergence of  $(x_i)_i$ ,  $(v_i)_i$ , we include a damping step (c) with an adaptively chosen step length.

*MYKSODA iteration scheme.*—Assume  $\tilde{X}$  bounded and  $E^0$  finite everywhere. For  $v \in X^*$  fixed, set  $v_1 = v$  and select  $x_1 \in \bar{\partial}E_\varepsilon^0(v)$ . Iterate  $i = 1, 2, \dots$  according to

(a) set  $v_{i+1} = v + \nabla F_\varepsilon(x_i) - \nabla F_\varepsilon^0(x_i)$  and stop if  $v_{i+1} = -\nabla F_\varepsilon^0(x_i) = v_{\text{KS}}^\varepsilon$ ,

(b) select  $x'_{i+1} \in \bar{\partial}E_\varepsilon^0(v_{i+1})$  and get the step direction  $y_i = (x'_{i+1} - x_i) / \|x'_{i+1} - x_i\|$ ,

(c) choose the step length  $\tau_i = -\varepsilon \langle \nabla F_\varepsilon(x_i) + v, y_i \rangle > 0$  and set  $x_{i+1} = x_i + \tau_i y_i$ .

We prove below that this algorithm guarantees convergence to the correct KS potential,  $v_i \rightarrow v_{\text{KS}}^\varepsilon$ , and to the ground-state quasidensity,  $x_i \rightarrow z$ , of both the full system with  $v$  and the reference system with  $v_{\text{KS}}^\varepsilon$ . The corresponding energy is then determined by  $E_\varepsilon(v) = F_\varepsilon(z) + \langle v, z \rangle > -\infty$ . These are still solutions of the regularized problem, but with Eqs. (8) and (9) a transformation back to the unregularized setting is easily achieved. This, unlike the usually assumed unregularized KS scheme, gives *different* ground-state densities for the noninteracting and the interacting system, while circumventing all problems of differentiability and thus of  $v$ -representability. The assumption that  $E^0$  is finite everywhere is trivially fulfilled in a finite-dimensional setting because  $E^0$  is a finite sum. It is still kept here to connect more closely to standard DFT and CDFT, where  $E^0$  is finite even in the infinite-dimensional setting; see Lieb [24] [Theorem 3.1(iii)] and Laestadius *et al.* [27] (Lemma 20), respectively.

*Convergence proof.*—We refer to the first part of the proof of Theorem 12 in Laestadius *et al.* [23] to show that the superdifferential  $\bar{\partial}E_\varepsilon^0(v_{i+1})$  is everywhere nonempty because of  $E^0$  finite, guaranteeing that the (regularized) ground-state problem in (b) always has at least one solution. The directional derivative of  $F_\varepsilon + v$  at  $x_i$  in direction  $x'_{i+1} - x_i$  can be rewritten by (a)

$$\langle \nabla F_\varepsilon(x_i) + v, x'_{i+1} - x_i \rangle = \langle v_{i+1} + \nabla F_\varepsilon^0(x_i), x'_{i+1} - x_i \rangle. \quad (13)$$

Realizing that  $x'_{i+1} \in \bar{\partial}E_\varepsilon^0(v_{i+1})$  from (b) and  $x_i \in \bar{\partial}E_\varepsilon^0(\nabla F_\varepsilon^0(x_i))$  from invertibility (Ref. [23], Lemma 4), we rewrite the right-hand side of Eq. (13) with the help of Eq. (9), substituting

$$\tilde{x}'_{i+1} = x'_{i+1} + \varepsilon v_{i+1} \in \bar{\partial}E^0(v_{i+1}), \quad (14)$$

$$\tilde{x}_i = x_i + \varepsilon \nabla F_\varepsilon^0(x_i) \in \bar{\partial}E^0(\nabla F_\varepsilon^0(x_i)), \quad (15)$$

which gives

$$\langle v_{i+1} + \nabla F_\varepsilon^0(x_i), \tilde{x}'_{i+1} - \tilde{x}_i \rangle - \varepsilon \|v_{i+1} + \nabla F_\varepsilon^0(x_i)\|^2. \quad (16)$$

Now, since  $\tilde{x}'_{i+1}$ ,  $\tilde{x}_i$  are selected from the superdifferential of  $E^0$  for the respective potentials  $v_{i+1}$ ,  $\nabla F_\varepsilon^0(x_i)$ , the inner product is always smaller or equal to zero (Ref. [23], Lemma 5). This property is called *monotonicity of  $\bar{\partial}E^0$*  and

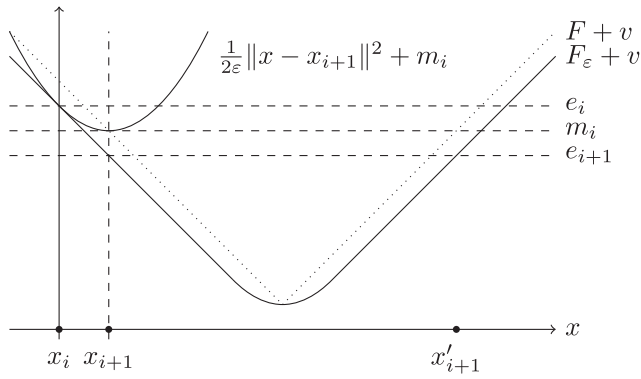


FIG. 1. Illustration of one iteration step.

directly follows from concavity of  $E^0$ . What follows is *strong* monotonicity of  $\bar{\partial}E_\varepsilon^0$ , i.e.,

$$\begin{aligned} \langle \nabla F_\varepsilon(x_i) + v, x'_{i+1} - x_i \rangle &= \langle v_{i+1} + \nabla F_\varepsilon^0(x_i), x'_{i+1} - x_i \rangle \\ &\leq -\varepsilon \|v_{i+1} + \nabla F_\varepsilon^0(x_i)\|^2 \\ &= -\varepsilon \|\nabla F_\varepsilon(x_i) + v\|^2. \end{aligned} \quad (17)$$

The last line follows from (a) and is strictly smaller than zero if not  $\|\nabla F_\varepsilon(x_i) + v\| = 0$ , which would mean that we have already converged to the ground-state quasidensity. We thus infer that, unless converged, we always have a negative directional derivative of  $F_\varepsilon + v$  at  $x_i$  in the step direction  $y_i$  which is parallel to  $x'_{i+1} - x_i$ , i.e.,  $\langle \nabla F_\varepsilon(x_i) + v, y_i \rangle < 0$ . Such a negative directional derivative means the left leg of the regularization parabola is aligned tangentially to the (differentiable) energy functional  $F_\varepsilon + v$ , like depicted in Fig. 1. The next quasidensity  $x_{i+1} = x_i + \tau_i y_i$  is then chosen at the vertex of this regularization parabola. This corresponds to a choice of step length  $\tau_i$  where the directional derivatives at  $x_i$  in direction  $y_i$  of the regularization parabola  $(1/2\varepsilon)\|\cdot - x_{i+1}\|^2$  and of  $F_\varepsilon + v$  are equal:

$$\begin{aligned} \langle \nabla F_\varepsilon(x_i) + v, y_i \rangle &= -\frac{1}{\varepsilon} \langle x_{i+1} - x_i, y_i \rangle \\ &= -\frac{1}{\varepsilon} \|x_{i+1} - x_i\| = -\frac{\tau_i}{\varepsilon}. \end{aligned} \quad (18)$$

This construction yields a  $x_{i+1} \neq x_i$ , where the energy  $e_i = F_\varepsilon(x_i) + \langle v, x_i \rangle$  is always larger than the energy value  $m_i$  at the vertex; see Fig. 1. Since the regularization parabola lies fully above the energy functional  $F_\varepsilon + v$  by construction, the energy  $e_{i+1}$  at  $x_{i+1}$  must obey  $e_{i+1} \leq m_i < e_i$ . The strictly decreasing  $e_i$  is now by definition bounded below by  $E_\varepsilon(v)$  from Eq. (6) and thus converges. By determining  $e_i - m_i$  from the regularization parabola and then combining it with Eq. (18),

$$\frac{\tau_i^2}{2\varepsilon} = \frac{1}{2\varepsilon} \|x_{i+1} - x_i\|^2 = e_i - m_i \leq e_i - e_{i+1} \rightarrow 0, \quad (19)$$

we can infer convergence of  $(x_i)_i$ . Step (a) then defines an associated potential,

$$\lim_{i \rightarrow \infty} v_{i+1} = v + \nabla F_\varepsilon^1(x) - \nabla F_\varepsilon^0(x), \quad (20)$$

since the gradients are both continuous. After having proved that the densities and potentials converge, it shall be demonstrated that they converge to the expected ground-state quasidensity  $z$  and KS potential  $v_{\text{KS}}^\varepsilon$ . We come back to Eq. (17), where substituting  $x'_{i+1} - x_i = y_i \|x'_{i+1} - x_i\|$  gives

$$\|x'_{i+1} - x_i\| \langle \nabla F_\varepsilon(x_i) + v, y_i \rangle \leq -\varepsilon \|\nabla F_\varepsilon(x_i) + v\|^2, \quad (21)$$

which together with Eq. (18) results in

$$\|x'_{i+1} - x_i\| \frac{\tau_i}{\varepsilon} \geq \varepsilon \|\nabla F_\varepsilon(x_i) + v\|^2. \quad (22)$$

We already know from the convergence of densities that  $(x_i)_i$  is bounded, further

$$x'_{i+1} \in \bar{\partial}E_\varepsilon^0(v_{i+1}) = \bar{\partial}E^0(v_{i+1}) - \varepsilon v_{i+1}, \quad (23)$$

by (b) and Eq. (9). But  $\bar{\partial}E^0(v_{i+1}) \subset \tilde{X}$ , which is bounded, and  $(v_i)_i$  converges as well. Thus,  $\|x'_{i+1} - x_i\|$  is bounded, and since  $\tau_i \rightarrow 0$ , it follows  $\|\nabla F_\varepsilon(x_i) + v\| \rightarrow 0$  and  $\|\nabla F_\varepsilon^0(x_i) + v_{i+1}\| \rightarrow 0$ . This in turn means  $v = -\nabla F_\varepsilon(\lim x_i)$ , so  $\lim x_i = z$  is the ground-state quasidensity for the potential  $v$  in the full, regularized problem. Finally,  $\lim v_{i+1} = -\lim \nabla F_\varepsilon^0(x_i) = -\nabla F_\varepsilon^0(z) = v_{\text{KS}}^\varepsilon$  is the KS potential. ■

As noted above, the reference system reproduces the quasidensity  $z$  of the full system and they link back to the real densities by Eq. (9),

$$\rho = z + \varepsilon v, \quad \rho_{\text{KS}}^\varepsilon = z + \varepsilon v_{\text{KS}}^\varepsilon, \quad (24)$$

where typically  $\rho_{\text{KS}}^\varepsilon \neq \rho$ . Then  $v_{\text{KS}}^\varepsilon - v = \varepsilon^{-1}(\rho_{\text{KS}}^\varepsilon - \rho)$  is precisely the Hartree-exchange-correlation potential that depends on the regularization parameter  $\varepsilon$  here. This means every choice of  $\varepsilon$  defines a *different* reference system. A limit  $\varepsilon \rightarrow 0$  in the algorithm is unfeasible because of its relation to the step length.

A simulation of two electrons on a ring lattice [31] allows us to illustrate the above method. Compared to a previous implementation in a CDFT setting [27], the version given here uses the more conservative damping step that helped to prove convergence. To distinguish the two versions, we denote them ‘‘MYKSODA-S’’ for shorter, conservative steps, and ‘‘-L’’ for the original longer steps [23,27]. Both versions have been adapted to a pure DFT setting,  $H_{\text{kin}}$ , taking the form of a standard second-order finite difference. A radius of  $R = 1$  bohr, a uniform grid with 30 points, and the interaction energy  $H_{\text{int}} = 3\sqrt{1 + \cos(\theta_1 - \theta_2)}$  were used. As expected, larger  $\varepsilon$  leads



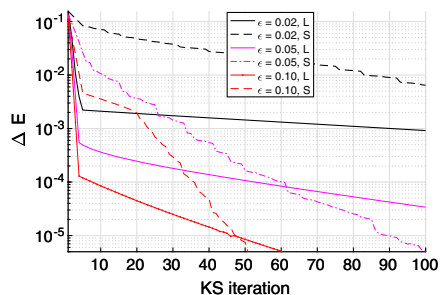


FIG. 2. Convergence of  $\Delta E_i = e_i - E_\epsilon(v)$  for a ring-lattice system with external potential  $v = \cos(2\theta) + 0.2\cos(\theta)$  and different  $\epsilon$ . The algorithm developed here is labeled “S” while the method “L” chooses the step length maximally [23,27].

to faster convergence. Also, the more conservative steps taken by MYKSODA-S often lead to slower convergence in practice. Surprisingly, however, in some cases MYKSODA-S overtakes the less conservative MYKSODA-L. An example is shown in Fig. 2. Such a crossover is possible as the two algorithms follow different paths through the space of densities and potentials. Yet, when the starting point is the same, the *first* step by MYKSODA-L always lowers the energy more than the first step by MYKSODA-S. Although it is a plausible conjecture that also MYKSODA-L, taking maximal steps, is guaranteed to converge, the present proof does not establish this. “Maximal steps” here means taking  $\tau_i$  maximally such that  $\langle \nabla F_\epsilon(x_{i+1}) + v, x'_{i+1} - x_i \rangle \leq 0$ , which yields maximal decrease in energy in the direction chosen by step (b).

In this Letter we proved convergence of the regularized KS scheme with special adaptive damping. In short, this means that KS DFT is a veritable method to calculate the correct ground-state density. This strong statement holds for all flavors of DFT that are defined on a finite-dimensional density space  $X = \ell^2(M)$  and have a linear coupling to external potentials of type  $\langle v, \rho \rangle$ . This includes CDFT, where the potential  $v$  is a combination of scalar and vector potential, and the density  $\rho$  includes the paramagnetic current density. To allow for a combination of these different entities into one Banach space setting, the respective function spaces for one-particle densities and current densities have to fulfill a condition termed “compatibility” in Laestadius *et al.* [27]. A proof of MYKSODA convergence for infinite-dimensional Banach spaces  $X$  is feasible but much more technical and will be presented elsewhere. The choice of step length in (c) is an essential part of the proof and similar choices could be of value in showing convergence of related iteration schemes and for other settings such as Hartree-Fock theory [14]. Next to this damping step the MY regularization is a vital part of the proof at hand, not only to have functional differentiability, but also for the strong monotonicity estimate needed to show convergence. How those findings can be transferred to realistic KS implementations will be the content of future

research, but it is expected that they serve as useful guidelines for better convergence results.

We express our gratitude for fruitful discussion with Simen Kvaal. M. P. is grateful for the hospitality received at the Hylleraas Centre for Quantum Molecular Sciences in Oslo and furthermore acknowledges support by the Erwin Schrödinger Fellowship J 4107-N27 of the FWF (Austrian Science Fund). E. I. T. and A. L. were supported by CoE Hylleraas Centre for Molecular Sciences Grant No. 262695, as well as the Norwegian Supercomputing Program (NOTUR) through a grant of computer time (Grant No. NN4654K). E. I. T. acknowledges support from Research Council of Norway through Grant No. 240674. A. L. acknowledges support from European Research Council under Grant Agreement No. ERC-STG-2014 639508.

\*markus.penz@mpsd.mpg.de

- [1] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [2] K. Burke, *J. Chem. Phys.* **136**, 150901 (2012).
- [3] G. Karlström, *Chem. Phys. Lett.* **67**, 348 (1979).
- [4] A. D. Rabuck and G. E. Scuseria, *J. Chem. Phys.* **110**, 695 (1999).
- [5] E. Cancès and C. Le Bris, *Math. Model. Numer. Anal.* **34**, 749 (2000).
- [6] T. Helgaker, H. Larsen, J. Olsen, and P. Jørgensen, *Chem. Phys. Lett.* **327**, 397 (2000).
- [7] K. N. Kudin and G. E. Scuseria, *J. Chem. Phys.* **116**, 8255 (2002).
- [8] S. Høst, J. Olsen, B. Jansik, L. Thøgersen, P. Jørgensen, and T. Helgaker, *J. Chem. Phys.* **129**, 124106 (2008).
- [9] X. Liu, Z. Wen, X. Wang, M. Ulbrich, and Y. Yuan, *SIAM J. Numer. Anal.* **53**, 1758 (2015).
- [10] K. N. Kudin and G. E. Scuseria, *Math. Model. Numer. Anal.* **41**, 281 (2007).
- [11] H. Fukutome, *Prog. Theor. Phys.* **45**, 1382 (1971).
- [12] F. H. Stillinger, *Phys. Rev. E* **59**, 48 (1999).
- [13] A. Ruzsinszky and J. P. Perdew, *Comput. Theor. Chem.* **963**, 2 (2011).
- [14] E. Cancès and C. Le Bris, *Int. J. Quantum Chem.* **79**, 82 (2000).
- [15] E. Cancès, *J. Chem. Phys.* **114**, 10616 (2001).
- [16] L. O. Wagner, E. M. Stoudenmire, K. Burke, and S. R. White, *Phys. Rev. Lett.* **111**, 093003 (2013).
- [17] L. O. Wagner, T. E. Baker, E. M. Stoudenmire, K. Burke, and S. R. White, *Phys. Rev. B* **90**, 045109 (2014).
- [18] E. Engel and R. M. Dreizler, *Density Functional Theory: An Advanced Course* (Springer, Berlin, 2011).
- [19] P. E. Lammert, *Int. J. Quantum Chem.* **107**, 1943 (2007).
- [20] E. I. Tellgren, *Phys. Rev. A* **97**, 012504 (2018).
- [21] J. Chayes, L. Chayes, and M. B. Ruskai, *J. Stat. Phys.* **38**, 497 (1985).
- [22] K. J. Giesbertz and M. Ruggenthaler, *Phys. Rep.* **806**, 1 (2019).
- [23] A. Laestadius, M. Penz, E. I. Tellgren, M. Ruggenthaler, S. Kvaal, and T. Helgaker, *J. Chem. Phys.* **149**, 164103 (2018).
- [24] E. H. Lieb, *Int. J. Quantum Chem.* **24**, 243 (1983).

- [25] S. Kvaal, U. Ekström, A. M. Teale, and T. Helgaker, *J. Chem. Phys.* **140**, 18A518 (2014).
- [26] P. E. Lammert, [arXiv:1807.06125](https://arxiv.org/abs/1807.06125).
- [27] A. Laestadius, M. Penz, E. I. Tellgren, M. Ruggenthaler, S. Kvaal, and T. Helgaker, *J. Chem. Theory Comput.* **15**, 4003 (2019).
- [28] M. Levy, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 6062 (1979).
- [29] M. Farzanehpour and I. V. Tokatly, *Phys. Rev. B* **86**, 125130 (2012).
- [30] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis* (Springer, Berlin, 2009), Vol. 317.
- [31] MYring, a program for Moreau–Yosida regularization of a one-dimensional quantum ring, <https://gitlab.com/et/myring>.