Pure-state noninteracting v-representability of electron densities from Kohn-Sham calculations with finite basis sets

De Silva, Piotr; Wesolowski, Tomasz Adam

How to cite

DE SILVA, Piotr, WESOLOWSKI, Tomasz Adam. Pure-state noninteracting v-representability of electron densities from Kohn-Sham calculations with finite basis sets. In: Physical review, A, Atomic, molecular, and optical physics, 2012, vol. 85, n° 3, p. 032518/1–9. doi: 10.1103/PhysRevA.85.032518

This publication URL: https://archive-ouverte.unige.ch/unige:21865
Publication DOI: 10.1103/PhysRevA.85.032518

Pure-state noninteracting v-representability of electron densities from Kohn-Sham calculations with finite basis sets

Piotr de Silva[1,2] and Tomasz A. Wesolowski[1]

[1]*Université de Genève, Département de Chimie Physique 30, quai Ernest-Ansermet, CH-1211 Genève 4, Switzerland*

[2]*K. Gumiński Department of Theoretical Chemistry, Faculty of Chemistry, Jagiellonian University, R. Ingardena 3, 30-060 Kraków, Poland*

(Received 18 January 2012; published 22 March 2012)

Within the linear combination of atomic orbitals (LCAO) approximation, one can distinguish two different Kohn-Sham potentials. One is the potential available numerically in calculations, and the other is the exact potential corresponding to the LCAO density. The latter is usually not available, but can be obtained from the total density by a numerical inversion procedure or, as is done here, analytically using only one LCAO Kohn-Sham orbital. In the complete basis-set limit, the lowest-lying Kohn-Sham orbital suffices to perform the analytical inversion, and the two potentials differ by no more than a constant. The relation between these two potentials is investigated here for diatomic molecules and several atomic basis sets of increasing size and quality. The differences between the two potentials are usually qualitative (wrong behavior at nuclear cusps and far from the molecule even if Slater-type orbitals are used) and δ-like features at nodal planes of the lowest-lying LCAO Kohn-Sham orbital. Such nodes occur frequently in LCAO calculations and are not physical. Whereas the behavior of the potential can be systematically improved locally by the increase of the basis sets, the occurrence of nodes is not correlated with the size of the basis set. The presence of nodes in the lowest-lying LCAO orbital can be used to monitor whether the effective potential in LCAO Kohn-Sham equations can be interpreted as the potential needed for pure-state noninteracting v-representability of the LCAO density. Squares of such node-containing lowest-lying LCAO Kohn-Sham orbitals are nontrivial examples of two-electron densities which are not pure-state noninteracting v-representable.

DOI: 10.1103/PhysRevA.85.032518 PACS number(s): 31.15.E−

I. INTRODUCTION

Any pure-state noninteracting v-representable electron density ρ is uniquely associated with the exact Kohn-Sham potential $v_s(\mathbf{r})$. Up to a constant, the exact Kohn-Sham potential is, therefore, a functional of the density:

$$v_s(\mathbf{r}) = v_s[\rho](\mathbf{r}). \tag{1}$$

Reversibly, in a system of $N = \int \rho(\mathbf{r})d\mathbf{r}$ noninteracting electrons in the external potential $v_s(\mathbf{r})$, the ground-state electron density is just $\rho(\mathbf{r})$. In numerical methods to solve Kohn-Sham equations, two major approximations are commonly applied: approximated expressions for the exchange-correlation part of the Kohn-Sham potential and the finite basis-set expansion for the Kohn-Sham orbitals. Other factors affecting the relationship between the potential and density, such as numerical integration, fitting the electrostatic potential, and terminating the self-consistent procedure at some arbitrary level, are of a more technical nature. The present work concerns the second of these two approximations, and all subsequent considerations are equally applicable for either the exact or approximated exchange-correlation potential case. For molecules, the linear combination of atomic orbitals (LCAO) approximation is commonly applied. Such calculations employ the effective Kohn-Sham potential $[v^{\mathrm{KS(LCAO)}}(\mathbf{r})]$ calculated from a given expression involving external, Hartree, and exchange-correlation potentials:

$$v^{\mathrm{KS(LCAO)}}(\mathbf{r}) = \sum_{i=1}^{N_{\mathrm{nuc}}} -\frac{Z_i}{|\mathbf{r}-\mathbf{R}_i|} + \int \frac{\rho^{\mathrm{tot}}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + \tilde{v}_{\mathrm{xc}}[\rho^{\mathrm{tot}}](\mathbf{r}). \tag{2}$$

The LCAO ground-state density $\rho_{\mathrm{tot}}^{\mathrm{LCAO}}(\mathbf{r})$ is reached at the end of the iterative procedure. The exact Kohn-Sham potential $(v_s[\rho_{\mathrm{tot}}^{\mathrm{LCAO}}](\mathbf{r}))$ is a potential for which the same density would be the ground-state solution obtained from an explicit differential equation. Although $v^{\mathrm{KS(LCAO)}}(\mathbf{r})$ is self-consistent with $\rho_{\mathrm{tot}}^{\mathrm{LCAO}}(\mathbf{r})$, it is not necessarily equal to $v_s[\rho_{\mathrm{tot}}^{\mathrm{LCAO}}](\mathbf{r})$:

$$v^{\mathrm{KS(LCAO)}}(\mathbf{r}) \neq v_s\left[\rho_{\mathrm{tot}}^{\mathrm{LCAO}}\right](\mathbf{r}). \tag{3}$$

It is frequently tacitly assumed that the two potentials are equal or that, with a sufficiently large basis set, the differences can be neglected. The principal objective of the present work is to verify such assumptions. We should emphasize that, for a particular approximation for the exchange-correlation functional, the completeness of the basis set used in the LCAO approximation can be easily monitored in practice by analyzing the convergence of the calculated observables of the principal interest. The total energy is especially useful because it is bound from below. There exists a rich literature on this subject (see Ref. [1], for instance). The effect of incompleteness of the basis set on the relation between the potentials $v^{\mathrm{KS(LCAO)}}(\mathbf{r})$ and $v_s[\rho_{\mathrm{tot}}^{\mathrm{LCAO}}](\mathbf{r})$ was not a subject of such extensive analysis. This relation is, however, instrumental in the interpretation of the quantities obtained in the numerical procedures to solve the Kohn-Sham equations. For instance, the use of a finite basis sets leads to a violation of the unique correspondence between the ground-state density and the effective potential, as discussed in detail by Staroverow, Scuseria, and Davidson [2]. Another manifestation of the lack of uniqueness are numerical problems at inverting Kohn-Sham equations, i.e., obtaining the Kohn-Sham potential for a given density $\rho(\mathbf{r})$ [3–5]. As a result, the exact exchange-correlation potential for a given density cannot be approached with arbitrary accuracy. In

the present work, we analyze the relationship between the potentials $v^{\mathrm{KS(LCAO)}}(\mathbf{r})$ and $v_s[\rho_{\mathrm{tot}}^{\mathrm{LCAO}}](\mathbf{r})$ obtained for a few diatomic molecular systems to address the following issues: (i) the existence of the potential $v_s(\mathbf{r})$ for a density $\rho_{\mathrm{tot}}^{\mathrm{LCAO}}(\mathbf{r})$ obtained within the LCAO approximation at the end of the self-consistent procedure, and (ii) the numerical differences between $v^{\mathrm{KS(LCAO)}}(\mathbf{r})$ and $v_s[\rho_{\mathrm{tot}}^{\mathrm{LCAO}}](\mathbf{r})$.

The numerical analyses are performed for the H_2, N_2, and LiH molecules and a series of atomic basis sets of the Slater form. This type of function in the basis set was chosen due to its correct qualitative behavior at nuclear cusps and in the tail regions. The present work concerns the LCAO approximation and either the exact or approximate exchange-correlation component of the Kohn-Sham potential, as all the considerations concern electron densities without the need to decompose the Kohn-Sham potential into its canonical components: external, electron-repulsion, and exchange-correlation potential. This work is organized as follows. We start with an overview of the known formal results concerning the exact potential $v_s[\rho_{\mathrm{tot}}](\mathbf{r})$, which is associated with a given ground-state electron density $\rho_{\mathrm{tot}}(\mathbf{r})$ and the most tightly bound Kohn-Sham orbital for this potential $[\phi_1(\mathbf{r})]$. Subsequently, examples are provided showing that the LCAO approximation leads frequently to a violation of the exact property of $\phi_1(\mathbf{r})$ (its nodelessness). The main body of the numerical results concerns the numerical differences between the potentials $v^{\mathrm{KS(LCAO)}}(\mathbf{r})$ and $v_s[\rho_{\mathrm{tot}}^{\mathrm{LCAO}}](\mathbf{r})$ for a representative sample of atomic basis sets.

II. THEORY

A. Nodelessness of the ground-state solution

A consequence of the variational principle is that the ground state of any Hamiltonian is nodeless (for a rigorous proof, see Ref. [6]; a more sketchy argument can be found in Ref. [7]). In principle, this statement is valid for any many-body system, however symmetry requirements may exclude the nodeless solution as being unphysical. This is certainly the case for more than two electrons, as the spatial part of a fermionic wave function cannot be totally symmetric, therefore it has to have nodes. Nevertheless, for one electron or two spin-compensated electrons, the spatial part of the wave function can be symmetric. The ground state is, therefore, nodeless. We will refer to this statement as *Observation I*. In particular, the lowest-lying molecular orbital (MO) of any system of noninteracting electrons, such as the Kohn-Sham system, is nodeless.

B. Reconstruction of the KS potential from the density of the lowest-lying orbital

The first Hohenberg-Kohn theorem ensures that there is a unique mapping between the density of a system and the external potential. In the most common approach to density functional theory (DFT), the density is constructed from the set of Kohn-Sham orbitals. Assume that the system comprises $2N$ electrons in an external potential $v(\mathbf{r})$. Kohn-Sham orbitals $\{\phi_i\}$ are the solutions to the following Kohn-Sham equation (atomic units are used throughout this paper):

$$\left[-\tfrac{1}{2}\nabla^2 + v_s(\mathbf{r})\right]\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}), \quad i = 1,2,\ldots,N. \tag{4}$$

Then one can construct the total electron density as the sum of orbital densities:

$$\rho_{\mathrm{tot}}(\mathbf{r}) = \sum_{i=1}^{N} 2|\phi_i(\mathbf{r})|^2. \tag{5}$$

Often the object of interest is solving the inverted problem, that is, finding the external potential $v(\mathbf{r})$, being given the total density $\rho(\mathbf{r})$. Several procedures have been developed to perform this inversion in practice [3–5]. For one electron or two spin-compensated electrons, the problem can be solved analytically by inverting the Kohn-Sham equation [Eq. (4)] for $\phi(\mathbf{r}) = \sqrt{\frac{1}{2}\rho(\mathbf{r})}$, provided that $\rho(\mathbf{r})$ is noninteracting v-representable. The potential $v_s(\mathbf{r})$ thus reads

$$\begin{aligned} v_s^{\mathrm{inv}}(\mathbf{r}) &= \frac{1}{2}\frac{\nabla^2\sqrt{\rho(\mathbf{r})}}{\sqrt{\rho(\mathbf{r})}} + \mathrm{const} \\ &= \frac{1}{4}\frac{\nabla^2\rho(\mathbf{r})}{\rho(\mathbf{r})} - \frac{1}{8}\frac{|\nabla\rho(\mathbf{r})|^2}{\rho^2(\mathbf{r})} + \mathrm{const}. \end{aligned} \tag{6}$$

The choice of the constant is arbitrary and usually it would be set in such a way that $\lim_{r\to\infty} v_s(\mathbf{r}) = 0$. This requirement sets const $= \varepsilon$, where ε is the corresponding eigenvalue.

In the above formula, we recognize the functional derivative of the von Weizsäcker kinetic energy functional, thus

$$v_s^{\mathrm{inv}}[\rho(\mathbf{r})] = -\frac{\delta T_W[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}. \tag{7}$$

If $\phi(\mathbf{r})$ changes sign, the densities $\rho(\mathbf{r}) = |\phi(\mathbf{r})|^2$ or $\rho(\mathbf{r}) = 2|\phi(\mathbf{r})|^2$ are not v-representable. This follows directly from Eq. (6). At any point where $\phi(\mathbf{r})$ has a node, the $\frac{\nabla^2\sqrt{\rho(\mathbf{r})}}{\sqrt{\rho(\mathbf{r})}}$ term is discontinuous because $\sqrt{\rho(\mathbf{r})} = |\phi(\mathbf{r})|$ [or $\sqrt{\rho(\mathbf{r})} = \sqrt{2}|\phi(\mathbf{r})|$].

For larger Kohn-Sham systems, more than one orbital is available but still each of them contains all the information about the system. This fact was noticed by Levy and Ayers in the context of evaluating the kinetic energy from a single Kohn-Sham orbital [8]. For any orbital $\phi_i(\mathbf{r})$, taking $\rho_i(\mathbf{r}) = |\phi_i(\mathbf{r})|^2$ and inserting it into Eq. (6) would give the external potential for the whole system. However, if $\phi_i(\mathbf{r})$ possesses nodal surfaces, the densities $|\phi_i(\mathbf{r})|^2$ and $2|\phi_i(\mathbf{r})|^2$ are not v-representable.

Hunter [9] has shown that a wave function with nodes can be replaced by its absolute value by adding δ barriers to the potential. As a consequence, the density of any orbital can be obtained as the ground-state density in some generalized potential function. If $\rho_1(\mathbf{r}) = |\phi_1(\mathbf{r})|^2$ is chosen, then Eq. (6) takes the following form:

$$\begin{aligned} v_s^{\mathrm{inv}}(\mathbf{r}) &= \frac{1}{2}\frac{\nabla^2\sqrt{\rho_1(\mathbf{r})}}{\sqrt{\rho_1(\mathbf{r})}} + \mathrm{const} \\ &= \frac{1}{2}\frac{\nabla^2|\phi_1(\mathbf{r})|}{|\phi_1(\mathbf{r})|} + \mathrm{const}. \end{aligned} \tag{8}$$

Following *Observation I*, this potential will be smooth in all points except for the nuclei positions, where the density has a cusp and is therefore nondifferentiable. This holds also for the Kohn-Sham method with any approximated multiplicative exchange-correlation potential which is continuous

everywhere,

$$v_{\mathrm{KS}}(\mathbf{r}) = \sum_{i=1}^{N_{\mathrm{nuc}}} -\frac{Z_i}{|\mathbf{r}-\mathbf{R}_i|} + \int \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \tilde{v}_{\mathrm{xc}}(\mathbf{r}). \quad (9)$$

As the ground-state solution of Eq. (4) with approximate potential is nodeless as well, the inverted potential [Eq. (8)] is equal to that of Eq. (9),

$$v_s^{\mathrm{inv}}[\rho_1](\mathbf{r}) = v_{\mathrm{KS}}(\mathbf{r}). \quad (10)$$

C. Finite basis-set approach to solving the Kohn-Sham equation

The usual strategy to solving the Kohn-Sham equation is to introduce a finite basis set $\{\chi_i\}$ and expand molecular orbitals in it,

$$\phi_i = \sum_j^{N_{\mathrm{basis}}} c_{ji}\chi_j. \quad (11)$$

Then Eq. (4), which is a differential equation, is replaced by an algebraic generalized eigenvalue problem, which has the familiar form of Roothan-Hall equations [10,11],

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon}, \quad (12)$$

where $F_{\mu\nu} = \langle\chi_\mu| -1/2\nabla^2 + v_{\mathrm{KS}}|\chi_\nu\rangle$ and $S_{\mu\nu} = \langle\chi_\mu|\chi_\nu\rangle$.

This approach proved to be extremely successful in electronic-structure theory, and many properties can be calculated accurately with sufficiently large basis sets. However, replacing Eq. (4) with Eq. (12) has important theoretical implications, resulting from the fact that the external potential and the electron density are never fully consistent.

Observation II. First of all, the first Hohenberg-Kohn theorem is not valid for finite basis-set models [12]. The consequence is that a given density can be a ground-state density of infinitely many potentials. As noted by Staroverov *et al*. [2], this is not an issue in most practical applications, but it becomes a problem in the optimized effective potential (OEP) method, making the OEP potential nonunique.

Observation III. As a consequence of *Observation I*, the exact ground-state solution $\phi_1^{\mathrm{exact}}(\mathbf{r})$ of the equation

$$\left[-\tfrac{1}{2}\nabla^2 + v_{\mathrm{KS}}^{\mathrm{LCAO}}(\mathbf{r})\right]\phi_1^{\mathrm{exact}}(\mathbf{r}) = \varepsilon_1\phi_1^{\mathrm{exact}}(\mathbf{r}), \quad (13)$$

where $v_{\mathrm{KS}}^{\mathrm{LCAO}}(\mathbf{r})$ is the Kohn-Sham potential obtained at the end of the self-consistent procedure within the LCAO approximation, does not possess nodal surfaces. This holds for any approximated potential $\tilde{v}_{\mathrm{xc}}[\rho_{\mathrm{tot}}](\mathbf{r})$ and follows from the fact that the potential given in Eq. (9) is continuous everywhere except at the nuclei. Again, this statement does not have to be true for the first orbital obtained from Eq. (12). The presence of such artificial nodes would mean that the LCAO ground-state orbital density is not v-representable, although it was obtained from a smooth Kohn-Sham potential [Eq. (9)].

For H_2, N_2, and LiH diatomics, Kohn-Sham calculations were performed in order to check whether lowest-lying orbitals exhibit nodal surfaces. The calculations were carried out with a Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional and a selection of Slater-type orbital (STO) and Gaussian-type orbital (GTO) basis sets, using ADF [13] and GAMESS [14] packages, respectively. In order to minimize errors due to technical limitations of numerical procedures, precision parameters were set far beyond the defaults for both programs. The self-consistent-field (SCF) convergence criterion was set to 10^{-10}, extensive integration grids were used, and cutoff values for linear scaling were practically switched off. The results are collected in Table I. It can be clearly seen that the presence of nodes is a rule in many electron systems. Using any other local-density approximation (LDA) or generalized-gradient approximation (GGA) to the exchange-correlation potential leads to the same observation. Contrary to what one may expect, improvement of a basis set does not lead to any systematic changes in the appearance of nodal surfaces, e.g., they do not move away from the molecule. Their shape, symmetry, and location are intrinsic properties of the specific basis set.

TABLE I. The occurrence of nodal surfaces for the lowest-lying Kohn-Sham orbital in selected STO and GTO basis sets.

Basis set	H_2	N_2	LiH
	Slater-type Orbitals		
SZ	No	No	No
DZ	No	Yes	Yes
TZP	Yes	Yes	Yes
pVQZ	Yes	Yes	Yes
QZ4P	No	Yes	Yes
	Gaussian-type Orbitals		
STO-6G	No	Yes	Yes
cc-pVDZ	No	Yes	Yes
cc-pVTZ	No	Yes	Yes
cc-pVQZ	No	Yes	Yes
aug-cc-pVQZ	No	Yes	Yes

Although the precision criteria used in calculations were very stringent, still one may suspect that artificial nodes would vanish in the absolute self-consistency limit. To clarify this issue, consider a simple model of a hydrogen atom in a double-ζ basis of Slater-type orbitals. Then the singly occupied orbital takes the following form:

$$\phi(r) = c_1\exp[-\alpha_1 r] + c_2\exp[-\alpha_2 r], \quad (14)$$

where $\alpha_2 > \alpha_1 > 0$ are the fixed basis set's exponential factors and c_i are the MO's coefficients obtained from Eq. (12). If the orbital possesses a nodal surface, it appears at the distance

$$r_{\mathrm{node}} = \frac{1}{\alpha_2 - \alpha_1}\ln\left[-\frac{c_2}{c_1}\right]. \quad (15)$$

From this equation, the condition follows immediately for the presence of a node: $c_2/c_1 < -1$, that is, MO's coefficients have to be of the opposite sign and $|c_2| > |c_1|$. Solving Eq. (12) with $v_{\mathrm{KS}} = -1/r$ reveals that this condition is fulfilled whenever both exponential factors are smaller than 1, which is the exact value for the $1s$ orbital [15]. From this example, it is evident that the existence of node-containing ground-state solutions is inherent to the LCAO approximation, and its occurrence depends on the basis set used.

Observation IV. At large distances from the nuclei, the asymptotic behavior of the density is determined by the least

negative eigenvalue corresponding to the occupied orbitals [16],

$$\rho(r) \sim \exp[-2(-2\varepsilon_{\max})^{1/2} r]. \tag{16}$$

The ability to reproduce this asymptotic behavior depends on the type of used basis set. The most common basis sets within the LCAO method are GTOs, which by construction never yield correct asymptotics. On the other hand, STOs have the correct exponential form, therefore they have the possibility to decay properly. Assume that the Kohn-Sham orbitals are linear combinations of STOs,

$$\phi_i^{\mathrm{LCAO}}(\mathbf{r}) = \sum_j c_{ji}(\mathbf{r}) \exp[-\alpha_j r], \tag{17}$$

where $c_{ji}(\mathbf{r})$ is a polynomial and $\alpha_j > 0$. Then, the asymptotic behavior of the density does not depend on the size of the basis set but is determined by the slowest decaying component in the expansion, thus

$$\rho(r) \sim \exp[-2\alpha_{\min} r]. \tag{18}$$

As a consequence, if the function with $\alpha_j = \sqrt{-2\varepsilon_{\max}}$ is not present in the basis set, adding other functions does not improve the asymptotics.

The above-mentioned consequences of the LCAO approximation in the Kohn-Sham method result in

$$v_s^{\mathrm{inv}}\left[\rho_1^{\mathrm{LCAO}}\right](\mathbf{r}) \neq v_{\mathrm{KS}}(\mathbf{r}) = \sum_{i=1}^{N_{\mathrm{nuc}}} -\frac{Z_i}{|\mathbf{r}-\mathbf{R}_i|} + \int \frac{\rho^{\mathrm{tot}}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + \tilde{v}_{\mathrm{xc}}[\rho^{\mathrm{tot}}](\mathbf{r}). \tag{19}$$

Note that the inequality above does not stem from the use of an approximate exchange-correlation functional, but from the fact that the basis set is incomplete. Following *Observation II*, the inverted potential is not the unique one. From *Observation III*, we can expect the presence of artificial delta barriers in the inverted potential because the orbital $\phi_1^{\mathrm{LCAO}}(\mathbf{r})$ can have nodes. Finally, taking into consideration *Observation IV*, it is not possible to obtain an accurate potential far from the nuclei. Having those limitations in mind, the results of applying Eq. (8) to obtain the Kohn-Sham potential from the density of the lowest LCAO molecular orbital are presented in the next section.

III. RESULTS

The analysis was performed for three diatomics: H_2, N_2, and LiH. The considered finite basis sets comprise STOs of increasing size. They are labeled following the convention of the Amsterdam Density Functional (ADF) package [13] as SZ, DZ, TZP, pVQZ, and QZ4P. Although GTOs are most commonly used in simulating chemical molecules, they are less useful for the present analysis. The use of GTOs leads to qualitatively wrong asymptotics of the density at nuclei and far from the system. Such flaws would make the analysis of this work difficult because they are reflected in the inverted potential [see Eqs. (8) and (16), in the case in which GTOs are used in Eq. (17)]. As the choice of an approximate exchange-correlation potential is not crucial for our analyses, only one functional has been selected, namely PBE [17].

The inverted potential $v_s^{\mathrm{inv}}(\mathbf{r})$ was obtained by means of Eq. (8) using the most tightly bound Kohn-Sham orbital and choosing the *constant* to be 0. To analyze the effect of a basis set, the difference $\Delta v(\mathbf{r})$ between the inverted potential and the Kohn-Sham potential used to obtain the converged density [Eq. (9)] is calculated:

$$\Delta v(\mathbf{r}) = v_s^{\mathrm{inv}}(\mathbf{r}) - v_{\mathrm{KS}}(\mathbf{r}). \tag{20}$$

In the complete basis-set limit (CBSL), this difference is just a constant determined by the eigenvalue of $\phi_1(\mathbf{r})$:

$$\Delta v_{\mathrm{CBSL}}(\mathbf{r}) = -\varepsilon_1. \tag{21}$$

Figures 1–3 show $\Delta v(\mathbf{r})$ calculated with various STO basis sets for H_2, N_2, and LiH. In each case, the difference $\Delta v(\mathbf{r})$ is not constant. The shape of $\Delta v(\mathbf{r})$ depends also on the basis set for each considered molecule. It could be expected that increasing the size of the basis sets to approach the complete basis-set limit would lead to fewer variations in $\Delta v(\mathbf{r})$. This assumption is analyzed below. For the H_2 molecule and the SZ basis set (Fig. 1), the potential difference $\Delta v(\mathbf{r})$ is far from constant. At each nucleus there is a broad deep well. Between the hydrogens, $\Delta v(\mathbf{r})$ goes above the target value $-\varepsilon_1$, exhibiting a peak in the very middle of the bond. This peak results from a drop of $v_{\mathrm{KS}}(\mathbf{r})$ in the center of the molecule and seems to be an artifact of the underlying PBE functional. Such a feature does not occur for other semilocal functionals such as BLYP [18,19]. Going to DZ improves the shape in the outer region, causing the shoulders of the graph to lean toward the target value. For TZP, the shoulders become more flat and the tops of the wells get narrower. For pVQZ, the $\Delta v(\mathbf{r})$ gets closer to the $-\varepsilon_1$ line also between the nuclei. Finally, $\Delta v(\mathbf{r})$ is almost constant outside the molecule for QZ4P. Compared to other basis sets, high peaks arise instead of deep wells in the positions of the nuclei. Between the hydrogen nuclei, this function is still not flat but goes close to $-\varepsilon_1$ in the center of the bond.

The case of N_2 is similar. For SZ, $\Delta v(\mathbf{r})$ is not constant anywhere, becoming more and more flat with the increase of the basis set's quality. Depending on the basis set, peaks or wells occur in the positions of the nuclei. Their bases also get narrower in larger basis sets and in general are much more confined than for H_2. However, the range where $\Delta v(\mathbf{r})$ remains almost flat is smaller, reaching approximately ± 1.3 Å in the best case.

LiH is an asymmetric system, for which the lowest-lying Kohn-Sham orbital is mostly localized on lithium and resembles the $1s$ orbital of the isolated lithium atom. As in previously considered symmetric diatomics, $\Delta v(\mathbf{r})$ varies significantly, exhibiting broad peaks at nuclei for SZ. Already for DZ, $\Delta v(\mathbf{r})$ is flat around a deep and narrow well at the position of the lithium nucleus. At the hydrogen nucleus, there is a peak with a broad basis between the nuclei and $\Delta v(\mathbf{r})$ goes below $-\varepsilon_1$ outside the bond. For TZP, $\Delta v(\mathbf{r})$ becomes more flat in the bonding region. This trend is retained for pVQZ and QZ4P basis sets. For QZ4P, the peak on the hydrogen atom becomes a well. The behavior of $\Delta v(\mathbf{r})$ outside the bonding region, in the immediate vicinity of the hydrogen, is rather erratic and mostly dictated by the presence of nodes in the most tightly bound orbital (see below).

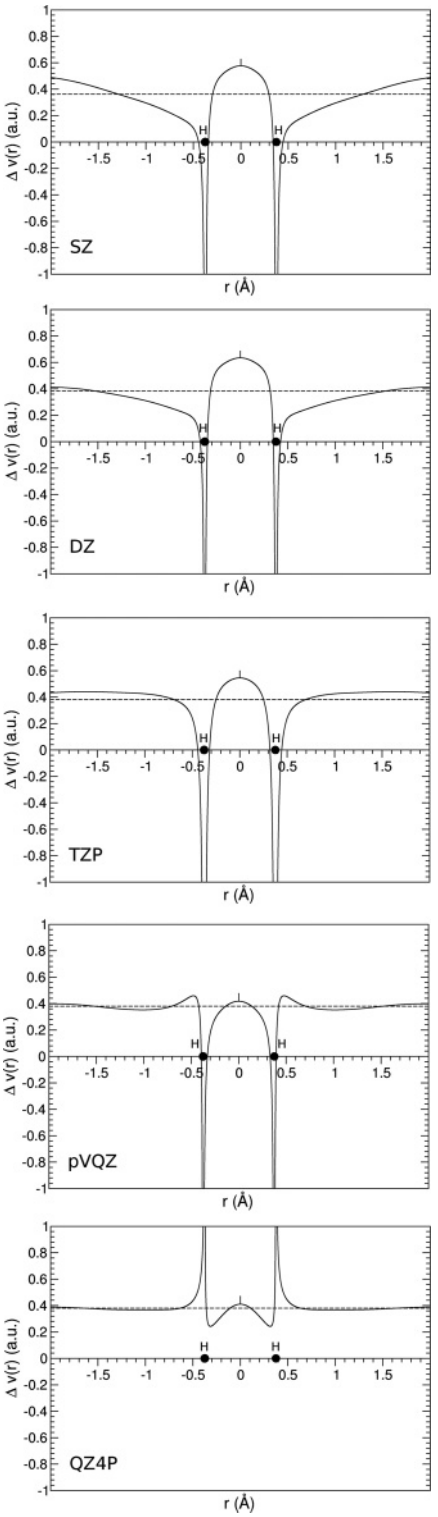

FIG. 1. $\Delta v(\mathbf{r})$ along the bonding axis calculated with various STO basis sets for the H_2 molecule. The broken line indicates the value of $-\varepsilon_1$.

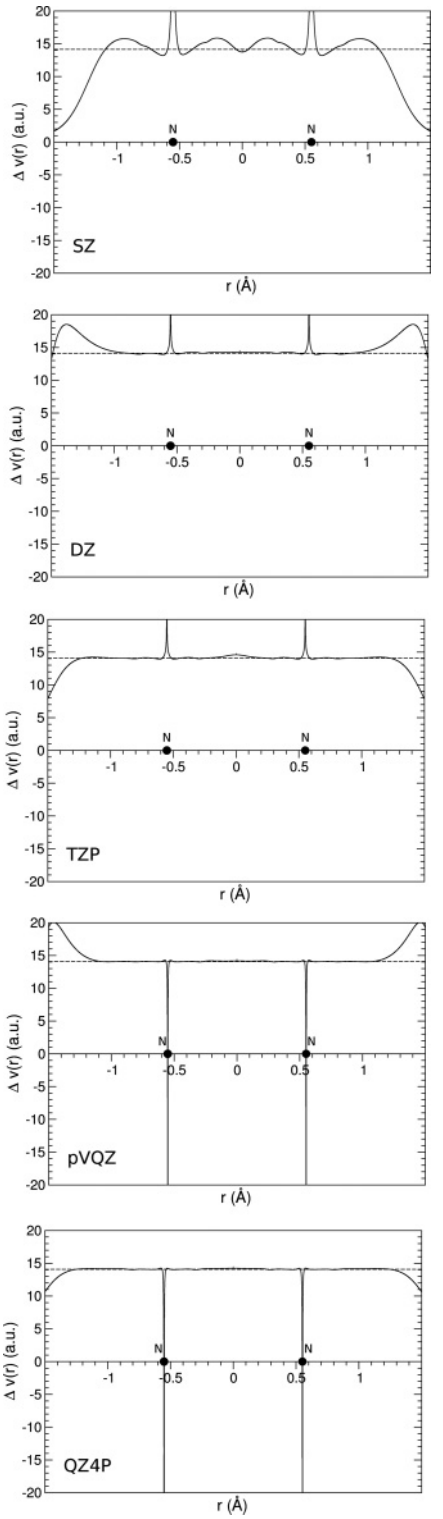

FIG. 2. $\Delta v(\mathbf{r})$ along the bonding axis calculated with various STO basis sets for the N_2 molecule. The broken line indicates the value of $-\varepsilon_1$.

So far, the shape of $v_s^{\text{inv}}(\mathbf{r})$ in the bonding region, i.e., where the density of the lowest-lying orbital is non-negligible, was shown. In these regions, the potentials $v_s^{\text{inv}}(\mathbf{r})$ and $v_{\text{KS}}(\mathbf{r})$ resemble each other better upon improvement of a basis set. The positions of the nuclei require special care. At the nuclei, either peaks or wells occur. This means that none of the basis sets used is able to describe correctly the density at nuclear cusps. However, these wells or peaks become more confined with the enlargement of a basis set, making the incorrectly described volume of space smaller. The subsequent

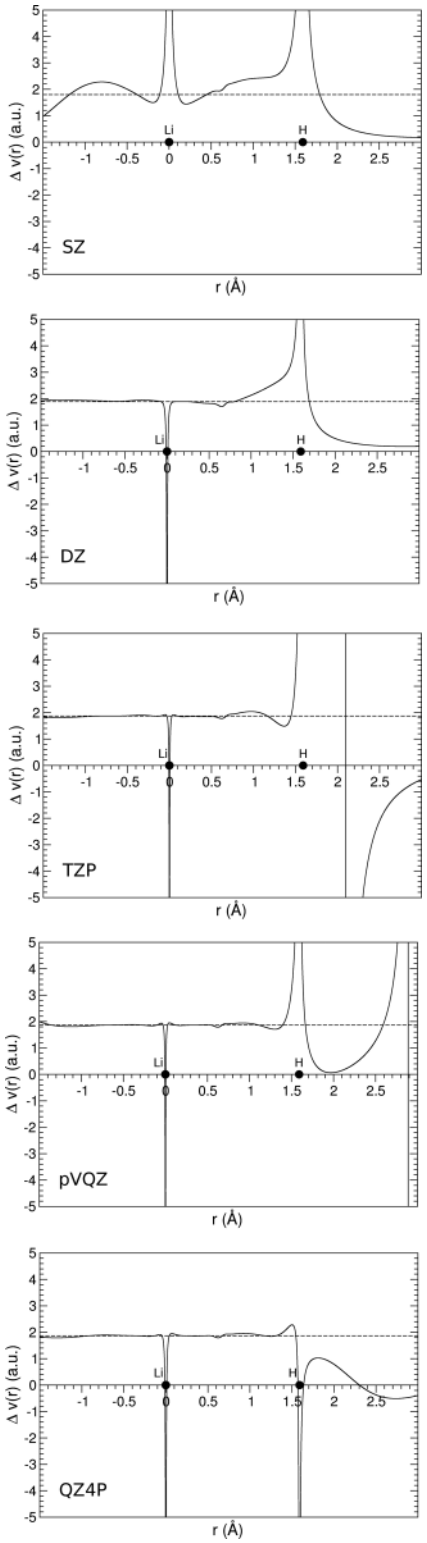

FIG. 3. $\Delta v(\mathbf{r})$ along the bonding axis calculated with various STO basis sets for the LiH molecule. The broken line indicates the value of $-\varepsilon_1$.

part concerns the inverted potential beyond the ranges shown in Figs. 1–3. First of all, the orbital density of the lowest-lying Kohn-Sham orbital in LCAO calculations approaches zero usually faster than the total density. This makes the right-hand side of Eq. (8) more vulnerable to basis-set deficiencies even

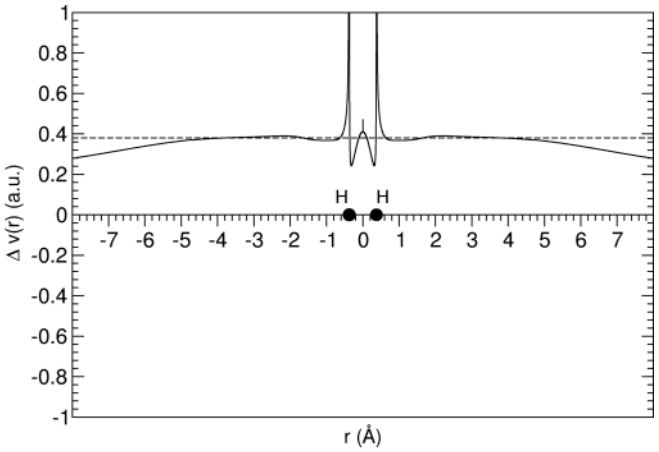

FIG. 4. $\Delta v(\mathbf{r})$ along the bonding axis calculated with the QZ4P basis set for the H_2 molecule. The broken line indicates the value of $-\varepsilon_1$.

in regions where the total density is still non-negligible. Moreover, as discussed in the Introduction (see Table I), the lowest-lying Kohn-Sham orbitals possess frequently artificial nodes in these regions. As a result, the right-hand side of Eq. (8) is not well-defined at the nodal surfaces. In the inverted potential, the nodes are reflected by Dirac δ-like barriers, which numerically show up as discontinuities with $\Delta v(\mathbf{r})$ going to $\pm\infty$ on both sides of the singularity. Figures 4–6 show $\Delta v(\mathbf{r})$ obtained using the basis set QZ4P (the highest quality among the considered basis sets) in a wider range. For H_2, $\Delta v(\mathbf{r})$ does not possess any discontinuities and $\Delta v(\mathbf{r})$ is close to $-\varepsilon_1$ (Fig. 4). Starting from about ± 4 Å, the graph falls gently below the reference line. In contrast, for N_2 (Fig. 5), $\Delta v(\mathbf{r})$ drops abruptly at ± 1.3 Å. In addition to this unphysical behavior, the first artificial nodes are present at approximately ± 2 Å and they result in discontinuities. Further on, $\Delta v(\mathbf{r})$ goes much below the $-\varepsilon_1$ value. The behavior for LiH is similar (Fig. 6). The presence of artificial nodes and the lack of symmetry of the system makes the plot highly erratic. Since the $\phi_1(\mathbf{r})$ orbital is mostly the core orbital of the lithium atom, $\Delta v(\mathbf{r})$ is close to

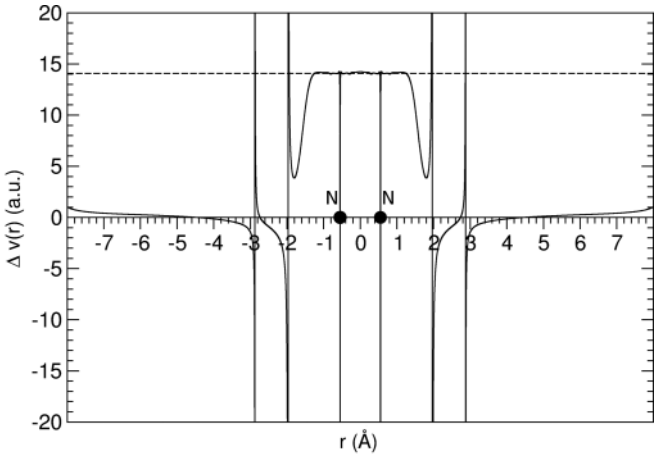

FIG. 5. $\Delta v(\mathbf{r})$ along the bonding axis calculated with the QZ4P basis set for the N_2 molecule. The broken line indicates the value of $-\varepsilon_1$.

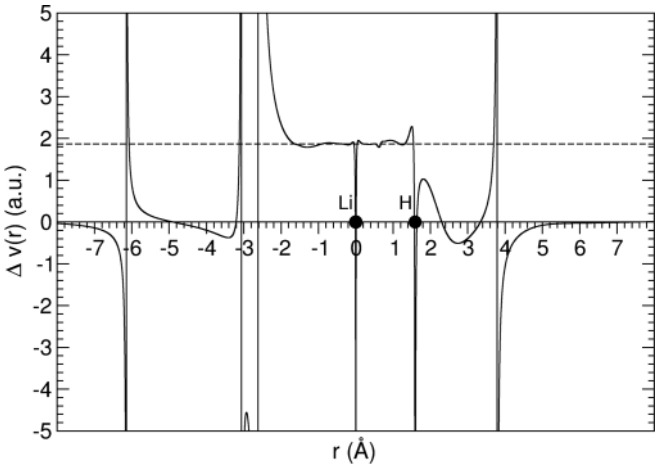

FIG. 6. $\Delta v(\mathbf{r})$ along the bonding axis calculated with the QZ4P basis set for the LiH molecule. The broken line indicates the value of $-\varepsilon_1$.

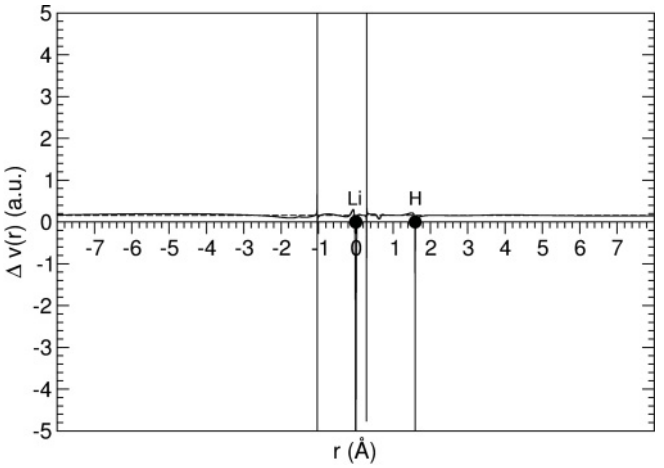

FIG. 7. Plot of $\Delta v(\mathbf{r})$ for LiH, obtained from $\phi_2(\mathbf{r})$ along bonding axis. The broken line indicates the value of $-\varepsilon_2$.

$-\varepsilon_1$ only in the vicinity of the lithium nucleus. At the position of the hydrogen, there is a deep well followed by a node at 3.8 Å approximately. The proximity of these two special points makes the graph between them irregular.

The above examples show cases of two-electron densities that are obviously not pure-state noninteracting v-representable. These are electron densities obtained as squares of the lowest-lying orbital obtained within the LCAO approximations. This is the case for N_2 and LiH and the QZ4P basis set, while the core density of H_2 is v-representable. Within the LCAO approximation, the lowest-lying orbital of the H_2 molecule can, however, possess nodes (see Table I). The square of this orbital, which is also the total electron density for this molecule, is not, therefore, pure-state noninteracting v-representable. For such densities, which are squares of a node-containing orbital, the corresponding multiplicative potential $v_s(\mathbf{r})$ does not exist within any function space, however it may be approached with generalized functions (distributions). This indicates that also the total density of a system, obtained using a finite basis set, does not have to be pure-state noninteracting v-representable.

The reason why the lowest-lying orbital was used in the inverting procedure [Eq. (8)] is that the exact orbital is nodeless (*Observation I*). The nodes occurring in practice are artifacts due to the LCAO approximation. In the case of higher orbitals, however, the nodes have a physical meaning. The corresponding orbital densities are not v-representable by definition. However, as noticed in Sec. II, they may be considered as ground-state densities in a generalized local potential, with δ barriers at the positions of nodes. In this case, the inverted potential [Eq. (6)] would be equal to the shifted Kohn-Sham potential $v_{\mathrm{KS}}(\mathbf{r})$ everywhere outside nodal surfaces. Equation (8) can be applied, therefore, anywhere outside the nodal surfaces to obtain the potential, which should be just the Kohn-Sham potential (exact) shifted by a constant. This statement does not hold when a finite basis set is employed. Figure 7 shows $\Delta v(\mathbf{r})$ derived from the $\phi_2(\mathbf{r})$ LiH orbital, obtained using the QZ4P basis set. The potential difference $\Delta v(\mathbf{r})$ follows closely the $-\varepsilon_2$ line in the whole range presented in Fig. 7. Two discontinuities, resulting from the physical nodal surface around the lithium nucleus, are present. There are no artificial nodes in this case. They may, however, appear for other basis sets, as was the case for $\phi_1(\mathbf{r})$ orbitals. Wells at the positions of nuclei and perturbations around singularities are very narrow in comparison with those obtained from $\phi_1(\mathbf{r})$ (see Fig. 6).

The data shown in Fig. 7 indicate that the inverted potential $v_s^{\mathrm{inv}}(\mathbf{r})$, derived from the second lowest-lying orbital, resembles the Kohn-Sham potential $v_{\mathrm{KS}}(\mathbf{r})$ much better than the inverted potential derived from the orbital density of the lowest-lying Kohn-Sham orbital. This results from the fact that this orbital is more diffused than $\phi_1(\mathbf{r})$, which resembles the $1s$ orbital of the lithium atom, thus it is very confined in space. The general result is that although each Kohn-Sham orbital contains complete information about the electronic structure of the whole molecular system, this is violated when a finite basis

TABLE II. Asymptotic properties of $v_s^{\mathrm{inv}}(\mathbf{r})$ far from the molecule. α_{min} is the exponent of the slowest-decaying basis function, and $\lim v_s^{\mathrm{inv}}(\mathbf{r})$ is the numerically available limit.

	α_{min}	$\frac{1}{2}\alpha_{\mathrm{min}}^2$	$\lim v_s^{\mathrm{inv}}(\mathbf{r})$
		H_2	
SZ	1.240	0.769	0.760
DZ	0.760	0.289	0.285
TZP	0.690	0.238	0.235
pVQZ	0.569	0.162	0.160
QZ4P	0.710	0.252	0.249
		N_2	
SZ	1.950	1.901	1.859
DZ	1.120	0.627	0.613
TZP	1.000	0.500	0.489
pVQZ	0.794	0.315	0.308
QZ4P	0.700	0.245	0.239
		LiH	
SZ	0.800	0.320	0.313
DZ	0.600	0.180	0.177
TZP	0.460	0.106	0.103
pVQZ	0.292	0.042	0.042
QZ4P	0.480	0.115	0.113

set is used. The discrepancy between $v_s^{\rm inv}(\mathbf{r})$ obtained from different LCAO orbitals means that they are eigenfunctions of different Hamiltonians. Only when projected onto a finite basis set can they become a solution to a common eigensystem.

Finally, we turn back to the consequences of using finite basis sets for the asymptotic behavior of the total density and their reflection in the Kohn-Sham potential. The asymptotic behavior of $v_s^{\rm inv}(\mathbf{r})$ is determined only by the slowest decaying function in the basis set, that is,

$$\lim_{r\to\infty} v_s^{\rm inv}(\mathbf{r}) = \frac{\alpha_{\rm min}^2}{2}. \tag{22}$$

To verify if Eq. (22) is fulfilled, the numerically available limits of $v_s^{\rm inv}(\mathbf{r})$ are collected in Table II and checked against the values determined by the basis set. It is evident that it is possible to reproduce numerically the asymptotic behavior dictated by the slowest-decaying basis function. This is a confirmation that the inverting procedure applied [Eq. (8)] is sound in the whole physical space, even where the density $\rho_1(\mathbf{r})$ is very small.

IV. CONCLUSIONS

The present numerical examples show that improving the atomic basis sets leads locally to better agreement between the effective potential in LCAO Kohn-Sham calculations and its exact counterpart obtained analytically from the lowest-lying LCAO orbital. This tendency holds, however, only locally in the region where the lowest-lying LCAO orbital is non-negligible. It can be expected, therefore, that expanding the size of the basis set in LCAO calculations would lead not only to a lower total energy but also to an improvement in all properties determined by the shape of the Kohn-Sham orbitals in the relevant region. Qualitative differences between the two potentials persist, however, in other regions. In particular, the potential obtained from Eq. (8) features δ-function-like singularities at artificial nodes of the lowest-lying LCAO orbital. Such nodes occur erratically if the basis set is changed. Their occurrence seems to be a rule rather than an exception in LCAO Kohn-Sham calculations, and they are usually localized in the regions where the density is very small. The node-containing lowest-lying LCAO orbital can be used to construct nontrivial one- or two-electron densities that are not pure-state noninteracting v-representable. We note that any higher-lying Kohn-Sham orbital also possesses nodes. Such nodes are, however, physical because the higher-lying orbitals must contain them in the molecular interior to be orthogonal to the lowest-lying one. The corresponding one- or two-electron densities are not pure-state noninteracting v-representable for obvious reasons in either the exact or LCAO case. They are trivial examples of nonrepresentable densities.

Once the LCAO density is fixed, not only is the exact corresponding Kohn-Sham potential uniquely defined by this density (up to a constant), but also the exact Kohn-Sham orbitals. The present study shows that neither the effective Kohn-Sham potential in LCAO calculations nor the potential obtained from Eq. (8) is the exact Kohn-Sham potential, even if the lowest-lying LCAO orbital is nodeless. In fact, using two different LCAO orbitals in Eq. (6) leads to two potentials that differ more than just by a constant and δ functions at physical nodes of the higher-lying orbitals. This indicates that the LCAO orbitals are not exact Kohn-Sham orbitals associated with the LCAO density. Concerning the lowest-lying LCAO orbital, the fact that applying it in Eq. (8) leads to an erroneous potential in the asymptotic region even if it is nodeless indicates one of the important consequences of the LCAO approximation. The lowest-lying LCAO orbital is too localized and decays too rapidly to describe properly the whole molecule. As a result, the correct asymptotics of the inverted potential can never be recovered for any LCAO density even if STOs are used. We recall here that an explicit term taking the form of the Fermi-Amaldi term [20] is used in commonly used numerical inversion methods such as the ones by Zhao, Morrisson, and Parr [3] and Wu and Yang [5]. The addition of the Fermi-Amaldi term results explicitly in the correct $-\frac{1}{r}$ long-range behavior of the exchange-correlation potential even if the LCAO is used in the inversion procedure. The reported lack of consistency here between the LCAO density and the underlying Kohn-Sham potential may explain the difficulties with applying numerical inversion techniques to obtain accurate potentials from a given density. The necessity of using large basis sets in this context was noticed by Tozer *et al*. [21].

The present work shows that even with extensive basis sets, the resulting LCAO Kohn-Sham orbitals may exhibit unphysical features such as nodes in the lowest-lying LCAO orbital. Moreover, even if the used basis set yields a nodeless lowest-lying LCAO orbital, the right-hand side of Eq. (6), which is well-defined outside the nodal surfaces also for higher-lying LCAO orbitals, depends on the choice of the LCAO orbital despite the fact that any exact Kohn-Sham orbital possesses complete information about the whole system [8]. This indicates that in the case of the LCAO approximation, the complete information about the system *is not contained in just one LCAO orbital*. Regularity or even the existence of a common potential $v(\mathbf{r})$ needed for pure-state noninteracting v-representability is not assured.

ACKNOWLEDGMENTS

This work was supported by the International PhD-studies program at the Faculty of Chemistry Jagiellonian University within the Foundation for Polish Science MPD Program co-financed by the EU European Regional Development Fund and a Grant from the Swiss National Science Foundation (200020/134791/1).

[1] F. Jensen, J. Chem. Phys. **116**, 7372 (2002).
[2] V. N. Staroverov, G. E. Scuseria, and E. R. Davidson, J. Chem. Phys. **124**, 141103 (2006).
[3] Q. Zhao, R. C. Morrison, and R. G. Parr, Phys. Rev. A **50**, 2138 (1994).
[4] R. van Leeuwen and E. J. Baerends, Phys. Rev. A **49**, 2421 (1994).
[5] Q. Wu and W. Yang, J. Chem. Phys. **118**, 2498 (2003).
[6] R. Courant and D. Hilbert, *Methods of Mathematical Physics,* Vol. I (Interscience, New York, 1962).

[7] M. Cohen, Doctoral thesis, California Institute of Technology, Pasadena, CA, 1956.
[8] M. Levy and P. W. Ayers, Phys. Rev. A **79**, 064504 (2009).
[9] G. Hunter, Int. J. Quantum Chem. **17**, 133 (1980).
[10] C. C. J. Roothaanm, Rev. Mod. Phys. **23**, 69 (1951).
[11] G. G. Hall, Proc. R. Soc. A **205**, 541 (1951).
[12] J. E. Harriman, Phys. Rev. A **27**, 632 (1983).
[13] G. te Velde, F. M. Bickelhaupt, S. J. A. van Gisbergen, C. Fonseca Guerra, E. J. Baerends, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931 (2001).
[14] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).
[15] MATHEMATICA notebook used to solve the Roothan-Hall equation is available upon request.
[16] M. M. Morrell, R. G. Parr, and M. Levy, J. Chem. Phys. **62**, 549 (1975).
[17] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
[18] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[19] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
[20] E. Fermi and E. Amaldi, Mem. R. Accad. Italia **6**, 117 (1934).
[21] D. Tozer, K. Somasundram, and N. Handy, Chem. Phys. Lett. **265**, 614 (1997).