



Variational properties of auxiliary density functionals

Daniel Mejía-Rodríguez¹ · S. B. Trickey¹

Received: 28 October 2020 / Accepted: 5 March 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

The evolution of variational Coulomb fitting from a purely practical scheme to reduce computational burden to a formal variant of Hohenberg–Kohn–Sham density functional theory (auxiliary density functional theory, ADFT) is discussed. After a summary of the historical evolution, an analysis of its connection with the Hohenberg–Kohn theorem is given, some implications for the Euler equation and for time-dependent DFT are given and some implications for the deMon2k code delineated.

Keywords deMon2k · Auxiliary density functional theory · Density-fitting · Density functional theory · Resolution-of-the-identity · Variational properties

1 Motivation and historical context

1.1 Motivation

Auxiliary density functional theory (ADFT) [1], in particular its structure and underpinnings, is the focus of this paper. Its implementation provides a computational cost scaling advantage for the deMon2k code [2] that is of particular value for studying very large molecules and clusters. An example in which we participated is a very recent study of a spin-crossover complex. The target molecule is known in chemical shorthand as Mn(taa) [3]. It has 53 atoms and 224 electrons, rather low symmetry, and a high-spin state nearly degenerate with its low-spin ground state. The accurate determination of such small energy difference requires basis sets of triple- or quadruple-zeta quality. Thus, Mn(taa) exemplifies the kind of system for which deMon2k is a very useful exploratory tool.

ADFT also is interesting because it is a formal structure that has emerged from a series of developments of pragmatic computational approximations. Hence, some aspects of ADFT seem rather little analyzed. We begin, therefore, with a summary of its antecedents and origins. They are valuable for understanding ADFT more clearly and formally

than heretofore. We remark that the paper deliberately is written in somewhat pedantic style in the hope that, among other benefits, it will be helpful to students who use and improve deMon2k.

1.2 Historical context

At the dawn of molecular and solid electronic structure computation and codes, the obvious computational bottleneck was Hartree–Fock exchange. To circumvent that, Slater [4] introduced what subsequently became recognized as the local density approximation for the exchange (X) functional in Hohenberg–Kohn–Sham density functional theory (DFT) [5, 6]. Similarly, the $X\alpha$ parametrized form [7] later came to be recognized as a semi-empirical approximation to the KS exchange-correlation (XC) functional.

Introduction of the KS potential (or its Slater antecedent) and DFT more generally immediately connected the many-electron problem to the one-electron, local potential problem that already was familiar then (e.g., central-field atom, empirical band-structure Hamiltonians, etc.). In the context of nuclear-site-centered basis function expansions of one-electron orbitals and the resulting secular equation [8, 9], DFT and its precursors also exposed two computational bottlenecks. One is shared with Hartree–Fock, the N_e^4 scaling of the classical Coulomb repulsion (Hartree energy, E_H) matrix elements. (N_e is the number of electrons.) The other is the nonlinear dependence of the XC energy and potential upon the electron number density

✉ Daniel Mejía-Rodríguez
dmejiaarodriguez@ufl.edu

¹ Department of Physics, Center for Molecular Magnetic Quantum Materials, Quantum Theory Project, University of Florida, Gainesville, FL 32611, USA

$$n(\mathbf{r}) := \sum_i f_i |\phi_i(\mathbf{r})|^2, \quad (1)$$

Here the ϕ_i and f_i are the KS orbitals and occupation numbers, respectively. Slater local exchange depends functionally on $n^{1/3}$. As is well-known by now, more accurate approximations have more complicated nonlinear n -dependence.

These two problems were addressed pragmatically quite early via fitted auxiliary densities. For the $V_H := \delta E_H[n]/\delta n$ matrix element problem, Baerends, Ellis, and Roos (BER) [10] used ordinary least-squares fitting to give an auxiliary density in the form of a finite linear expansion in a basis $\{\alpha\}$,

$$\tilde{n}^{\text{BER}}(\mathbf{r}) := \sum_i x_i^{\text{BER}} \alpha_i(\mathbf{r}) \approx n(\mathbf{r}). \quad (2)$$

Replacement of $E_H[n]$ by $E_H[\tilde{n}^{\text{BER}}]$ and evaluation of the XC energy $E_{\text{xc}}[n]$ by quadrature on a grid then gave computational scaling $\propto N_c^2$. But the total energies were rather poor because of the non-variational determination of the coefficients x_i^{BER} . Amelioration of that problem by use of large charge-fitting basis sets introduced numerical stability problems.

In the BER treatment, both the orbital and auxiliary basis sets were of Slater-type orbitals (STOs). As the computational advantages of Gaussian basis functions [11] drove their adoption in preference to STOs, Sambe and Felton [12, 13] introduced an exchange fit for $n^{1/3}(\mathbf{r})$ with a Gaussian-type basis separate from both the Gaussian-type orbital (GTO) basis and the Gaussian-type charge-fitting basis. The advantage conferred was that all of the matrix elements in the secular equation could be evaluated analytically with three-center integrals at most. Fitting and associated numerical integral weights were required only for auxiliary representations of the bland, nodeless density and its cube root. Again, however, the non-variational procedure necessitated very large basis sets and introduced numerical stability difficulties. (Independently, a small basis (STO-3G) variant was done with X-density fitting only [14]. Perhaps because of the resulting N_c^4 scaling, it had essentially no impact on subsequent development.)

The $E_H[n]$ variational stability problem was solved with the introduction of variational Coulomb fitting (also known as robust fitting) by Dunlap, Connolly, and Sabin (DCS hereafter) [15]. The idea had been presented earlier by Whitten [16] but gone unnoticed. See Ref. [17] for a review of subsequent literature and of the relationship between variational Coulomb fitting and superficially similar but non-variational algorithms known generically as “RI methods.” Essentially, DCS showed that minimization of the density-error Coulomb repulsion

$$\Delta E_H[n] = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \frac{\Delta n(\mathbf{r}_1) \Delta n(\mathbf{r}_2)}{r_{12}} \quad (3)$$

$$:= [\Delta n | \Delta n] \quad (4)$$

with

$$\Delta n(\mathbf{r}) := n(\mathbf{r}) - \tilde{n}(\mathbf{r}) \quad (5)$$

provides a variational (lower bound) determination of the coefficients in the *auxiliary* density defined as

$$\tilde{n}(\mathbf{r}) = \sum_i x_i \alpha_i(\mathbf{r}). \quad (6)$$

Thus, if everything else in the secular equation were to be handled so as to preserve the HK variational property, variational Coulomb fitting would present a min-max problem for coding [18].

In their paper and their code, DCS retained the Sambe–Felton XC fitting approach. The XC auxiliary density expansion used a separate Gaussian basis with least-squares determination of coefficients against the true density, i.e., Eq. (1). The same was true for the extension to periodic systems by Mintmire, Sabin, and Trickey [19]. (Periodically bounded systems also require fitting to neutral densities. That is not an issue here.) In both codes, the computational cost issue with the Sambe–Felton XC auxiliary density is that it requires construction of $n(\mathbf{r})$ at each SCF cycle, itself a task quadratic in the orbital basis. In essence the Hartree energy computational cost had been reduced sufficiently that the XC energy cost was the new bottleneck.

So the next innovation and simplification was Boettger’s “fit-to-fit.” Introduced almost invisibly (mentioned in only one dependent clause) in Ref. [20] as part of the evolution of the periodic system code FILMS, on a practical level the innovation was simple. One does the Sambe–Felton least-squares XC fitting to the auxiliary variational-Coulomb-fitted density, \tilde{n} , instead of to n . Since \tilde{n} is linear in auxiliary basis functions, using it as the reference density is much faster than the original Sambe–Felton XC fit. Boettger’s reasoning was straightforward physics: since \tilde{n} is good enough for calculating a large contribution to the total energy, $E_H[n]$, and can be refined systematically, \tilde{n} surely ought to be good enough for calculating a contribution that is smaller in magnitude, E_{xc} . Fit-to-fit was carried over into GTOFF, the successor 2D and 3D periodic code to FILMS [21].

Fit-to-fit was incorporated in only one other code to our knowledge, ParaGauss [22]. That aside, conceptually fit-to-fit is important because it comes close to converting the original KS problem, determination of $n(\mathbf{r})$, into the problem of determining a new density functional, namely

$E[\tilde{n}[n]]$, with the original self-consistent KS equation used to solve for \tilde{n} and evaluated on it. Only the non-variational second fit in fit-to-fit prevents that mapping from being the case.

It appears that the first *direct* use of an auxiliary density in the context of the XC contribution was by Fan and Ziegler [24]. A critical point is that they used \tilde{n}^{BER} (in terms of STOs), not the variational \tilde{n} . They separated E_{xc} into what they called a local and (inaccurately) a nonlocal part. In fact those were a local density approximation (LDA) part and a generalized gradient approximation (GGA) correction, respectively. They focused on the resulting decomposition of the XC potential into $\delta E_{xc}/\delta n = v_{xc} \approx v_x^L + v_c^L + v_x^{\text{GGA}} + v_c^{\text{GGA}}$ in modern notation and used the valence auxiliary density, in our notation \tilde{n}_v^{BER} ($\tilde{\rho}_v$ in theirs), to evaluate the GGA part. They state that “It was found computationally expedient to evaluate $v_x^{\text{GGA}} + v_c^{\text{GGA}}$ from \tilde{n}_v^{BER} . Tests revealed that the error introduced by using \tilde{n} rather than n in evaluating $v_x^{\text{GGA}} + v_c^{\text{GGA}}$ is less than 0.001 a.u. ...”. (Again, we have updated the notation.) That is pure substitution, not fit-to-fit nor ADFT. Nothing was said about how the “local” part, $v_x^{\text{LDA}} + v_c^{\text{LDA}}$, was done nor how E_{tot} was calculated. Nor, regrettably, did they give any details about what was “computationally expedient.”

Somewhat similar direct use of a non-variational auxiliary density to evaluate E_{xc} approximately (as a “Harris functional”) appears in Ref. [25]. Perhaps because both that paper and Fan and Ziegler used STO basis functions, there is little discernible impact of them on the mainstream GTO-basis-method literature.

Independently of fit-to-fit and somewhat later, the developers of the deMon2k code came to the idea of direct use of \tilde{n} from variational Coulomb fitting in the evaluation of both E_{xc} and $v_{xc} = \delta E_{xc}/\delta n$ as well as in evaluation of $E_{\text{H}}[n]$ [1, 26]. Again, this was a re-discovery of an idea, in this case apparently first put forth by Laikov [27] who, ironically enough, did not mention fit-to-fit. Laikov implemented his scheme in the PRIRODA suite, which has seen much use in the Russian Federation [28]. Similarly, upon learning (later) of fit-to-fit, the lead deMon2k developers viewed it as confirmation of the merit of the direct use of \tilde{n} they were pursuing. Direct use of \tilde{n} takes the physical motivation of fit-to-fit and eliminates the overhead and uncontrolled approximation associated with least-squares fitting for the XC density. The result appears to be a theoretical construct for finding $\tilde{n}[n]$. It has been dubbed ADFT [1] and arguments given for its validity. A quite similar argument and implementation, called the model density approach (MDA), appeared very shortly thereafter [29], although the original implementation in ParaGauss appears to date as far back as 1997 [30, 31].

The emergence of ADFT as a theoretical construct within DFT is rather remarkable. As just sketched, use of auxiliary densities in KS DFT calculations began and grew as

thoroughly pragmatic steps for computational efficiency. What has emerged is a formalism that in a sense is reverse-engineered from those computational approximations. It is a somewhat subtle, comparatively little-studied, and (we think) inadequately justified formal structure. We address those issues by revisiting the line of reasoning presented in Refs. [1, 27, 29] and connecting that reasoning directly to the underlying DFT theorems.

The main realization of AFDT is in the deMon2k code. (MDA is implemented in ParaGauss but that code is not widely distributed.) Therefore, in the next section we outline relevant aspects of ADFT. After that we refine the original ADFT reasoning to identify relationships with the underlying HKS quantities, and then give a formal analysis connecting ADFT to the constrained-search form of the HK theorems. We make connection with time-dependent DFT (tdDFT) and conclude with a few observations and remarks about technical implications.

2 deMon2k code

deMon2k [2] is a KS DFT quantum-chemistry code, that is, a code for isolated molecules or clusters. Its design objectives include high computational efficiency and low memory requirements for use on large systems. It uses the linear combination of atomic orbitals (LCAO) ansatz [8]. Thus, each KS molecular orbital (MO) is expressed as a linear combination of atom-centered (nuclear-site centered) basis functions

$$\phi_i(\mathbf{r}) = \sum_{\mu} C_{\mu i} \chi_{\mu}(\mathbf{r}). \quad (7)$$

The basis functions, $\chi(\mathbf{r})$, commonly are called atomic orbitals (AOs). The MO coefficients, $C_{\mu i}$, are to be determined variationally. The exact form of the AOs is unimportant for this discussion; we note only that deMon2k uses contracted GTOs. Expressed in terms of the AOs, the KS electron number density, Eq. (1), is

$$n(\mathbf{r}) = \sum_{\mu, \nu} P_{\mu \nu} \chi_{\mu}(\mathbf{r}) \chi_{\nu}(\mathbf{r}) \quad (8)$$

with

$$P_{\mu \nu} := \sum_i f_i C_{\mu i} C_{\nu i} \quad (9)$$

called the AO density matrix. It follows, as already noted, that computation of the Hartree energy

$$E_{\text{H}}[n] = \frac{1}{2} [n | n] \quad (10)$$

scales formally as $\mathcal{O}(N_{\text{basis}}^4)$, where N_{basis} is the number of basis functions $\propto N_e$.

To avoid that, deMon2k uses the auxiliary density, $\tilde{n}(\mathbf{r})$, Eq. (6) from variational Coulomb fitting. The precise form of the auxiliary basis functions $\alpha_j(\mathbf{r})$ in deMon2k is irrelevant here (but see remarks in Sect. 6). They are uncontracted Hermite–Gaussian-type orbitals.

The effect is to approximate $E_H[n]$ with some functional $\tilde{E}_H[\tilde{n}, n]$ of $\tilde{n}(\mathbf{r})$ and, possibly, of $n(\mathbf{r})$ explicitly as well. Note the tilde on \tilde{E}_H . It signifies an important distinction. For the ensuing discussion, a convenient formulation of variational Coulomb fitting as done in deMon2k is to determine \tilde{E}_H through linear order in $\Delta n(\mathbf{r})$. One has

$$E_H[n] = E_H[\tilde{n} + \Delta n] = E_H[\tilde{n}] + [\tilde{n} | \Delta n] + E_H[\Delta n] \equiv \tilde{E}_H[\tilde{n}, n] + E_H[\Delta n]. \quad (11)$$

The error in approximating $E_H[n]$ by $\tilde{E}_H[\tilde{n}, n]$ is $E_H[\Delta n] \geq 0$, so the DCS variational minimization of that error to optimize $\tilde{n}(\mathbf{r})$ is

$$\delta(E_H[n] - \tilde{E}_H[\tilde{n}, n]) = \int d\mathbf{r} \frac{\delta \tilde{E}_H[\tilde{n}, n]}{\delta \tilde{n}(\mathbf{r})} \delta \tilde{n}(\mathbf{r}) = 0 \quad (12)$$

Given a fixed auxiliary basis $\{\alpha(\mathbf{r})\}$, the only variational freedom is in the fitting coefficients $\{x_i\}$. Thus, the extremum gives

$$\sum_j [\alpha_i | \alpha_j] x_j = [n | \alpha_i] \quad (13)$$

$$[\tilde{n} | \alpha_i] = [n | \alpha_i].$$

In notation currently used in conjunction with deMon2k, this is

$$\mathbf{G} \mathbf{x} = \mathbf{J}[n] \quad (14)$$

with $G_{ij} := [\alpha_i | \alpha_j]$ and $J_i[n] := [n | \alpha_i]$.

It is important to note that deMon2k uses an *unconstrained* density fit. There is no constraint corresponding to $\int d\mathbf{r} \tilde{n}(\mathbf{r}) = N_e$. To the best of the authors' knowledge, the constraint was dropped since version 2.1 of the code around 2005 (note that Refs. [1] and [38] mentioned the constraint.) We return to that point below, and note only that *a posteriori* normalization checks typically show a density normalization error of one part in a thousand or less. Here, the absence of that constraint becomes important because Eqs. (13) and (14) mean that

$$2E_H[\tilde{n}] = [n | \tilde{n}]. \quad (15)$$

It follows easily that

$$\tilde{E}_H[\tilde{n}] = E_H[\tilde{n}] \quad (16)$$

exactly in this unconstrained fit. If a norm-preserving constraint with Lagrange multiplier λ were to be used, the corresponding identity would be

$$\tilde{E}_H[\tilde{n}] = E_H[\tilde{n}] + \lambda N_e. \quad (17)$$

Observe that Eq. (13) still holds. It is straightforward to demonstrate that

$$\lambda = \{[\tilde{n} | n] - [\tilde{n} | \tilde{n}]\} / N_e \quad (18)$$

thereby confirming that λ vanishes for a perfect fit (against the Coulomb metric) and is small for large systems with realistic fitting precision.

The computational outcome of this variational Coulomb fitting is the reduction in the formal scaling in deMon2k to $\mathcal{O}(N_{\text{basis}}^2 N_{\text{auxis}})$, with “basis” and “auxis” being deMon2k key words that refer to the AO and auxiliary basis sets, respectively. For most applications, satisfactory auxiliary basis sets have cardinality $N_{\text{auxis}} \approx 4N_{\text{basis}}$. The resulting formal cubic scaling can be reduced further by means of integral screening techniques (based on the Cauchy–Schwarz inequality) and asymptotic expansions [35, 36]. Though valuable, those deMon2k aspects are not relevant here.

3 ADFT formulation

Use of $E_H[\tilde{n}]$ to ease the $E_H[n]$ computational burden brings us, as already discussed, to the numerical integration of the XC potential. That task scales formally as $\mathcal{O}(N_{\text{basis}}^2 N_{\text{grid}})$, where N_{grid} is the number of numerical integration grid points. Reduction in that burden in ADFT is by substitution of $E_{\text{xc}}[\tilde{n}]$ for $E_{\text{xc}}[n]$ [1]. The analysis given in that paper (and in Refs. [27, 29] as well) focused upon functional derivatives of the substitute, \tilde{n} -dependent expression that are related properly to the functional derivatives $\delta/\delta n$ in the KS equation. But no analysis of the relationship between functionals themselves, that is, the original n -dependent and substituted \tilde{n} -dependent ones, was presented. We therefore address that issue.

Begin by reiterating the obvious. The auxiliary density is a functional of the KS electron density. In the auxiliary basis representation that is explicit:

$$\begin{aligned} \tilde{n}[n; \mathbf{r}] &= \sum_i x_i[n] \alpha_i(\mathbf{r}) \\ &= \sum_{i,j} (\mathbf{G}^{-1})_{ij} J_j[n] \alpha_i(\mathbf{r}). \end{aligned} \quad (19)$$

The second line follows from Eq. (14). The matrix \mathbf{G}^{-1} is guaranteed to exist since the Coulomb operator $1/r_{12}$ induces a metric (in other words, it defines a positive definite kernel).

Before proceeding a remark is needed. In robust Coulomb fitting there is a crucial but usually unstated assumption about the auxiliary basis $\{\alpha\}$, namely that it is capable of representing the target density n in a meaningful way. The mathematical difficulty arises from the constrained search

formulation of DFT. For any given finite basis, there might be “perverse” densities for which the minimization, Eq. (12) either leaves unacceptably large errors $E_H[\Delta n]$ or, worse for the argument below, for which $\delta\tilde{n}[n]/\delta n$ does not exist or has difficult properties. Known conditions [32] on what constitutes a proper density include

$$\int d\mathbf{r} n(\mathbf{r}) = N_e \quad (20)$$

$$n(\mathbf{r}) \geq 0 \text{ almost everywhere} \quad (21)$$

$$\int d\mathbf{r} |\nabla n^{1/2}(\mathbf{r})|^2 < \infty \quad (22)$$

$$\int d\mathbf{r} |\nabla n^{1/3}(\mathbf{r})|^3 < \infty. \quad (23)$$

Typical practice is to choose the auxiliary basis $\{\alpha\}$ so that it resembles closely the orbital basis product in Eq. (8) and gives \tilde{n} that obeys at least (21) and (22) and satisfies (20) at least to acceptable numerical precision. In what follows we assume that is enough to provide well-defined $\delta\tilde{n}/\delta n$. See discussion below also.

In consequence of the fact that \tilde{n} is a functional of n , both $E_H[\tilde{n}]$ and $E_{xc}[\tilde{n}]$ also are implicit functionals of $n(\mathbf{r})$. They are *not*, in general, identical with the original ones. Instead, they are new auxiliary density functionals that are intended to be made as close as possible to the originals in the context of use of \tilde{n} . We can say that, in a sense, ADFT is a black-box that transforms a given density functional into an auxiliary density functional. The issue raised is the extent to which general properties of that auxiliary functional can be established.

For the classical Coulomb repulsion, we define the auxH functional as

$$E_{\text{auxH}}[n] := \tilde{E}_H[n, \tilde{n}[n]]. \quad (24)$$

As just discussed, it is identical to $E_H[\tilde{n}[n]]$ when unconstrained variational Coulomb fitting is used. Recall Eq. (16). In ADFT, the auxXC functional is defined as

$$E_{\text{auxXC}}[n] := E_{xc}[\tilde{n}[n]] \quad (25)$$

Since both auxH and auxXC are implicit functionals of $n(\mathbf{r})$, the functional derivative chain rule gives their respective potentials in the Euler equation (KS equation) for the original $n(\mathbf{r})$:

$$v_{\text{auxH}}(\mathbf{r}) := \int d\mathbf{r}_1 \frac{\delta E_H[\tilde{n}]}{\delta \tilde{n}(\mathbf{r}_1)} \frac{\delta \tilde{n}[n; \mathbf{r}_1]}{\delta n(\mathbf{r})} \quad (26)$$

$$\equiv \int d\mathbf{r}_1 \tilde{v}_H[\tilde{n}(\mathbf{r}_1)] \frac{\delta \tilde{n}[n; \mathbf{r}_1]}{\delta n(\mathbf{r})} \quad (27)$$

and

$$v_{\text{auxXC}}(\mathbf{r}) := \int d\mathbf{r}_1 \frac{\delta E_{xc}[\tilde{n}]}{\delta \tilde{n}(\mathbf{r}_1)} \frac{\delta \tilde{n}[n; \mathbf{r}_1]}{\delta n(\mathbf{r})} \quad (28)$$

$$\equiv \int d\mathbf{r}_1 \tilde{v}_{XC}[\tilde{n}] \frac{\delta \tilde{n}[n; \mathbf{r}_1]}{\delta n(\mathbf{r})}. \quad (29)$$

Note the tilde atop $\tilde{v}_H[\tilde{n}]$ and $\tilde{v}_{XC}[\tilde{n}]$. That notation is to make explicit the distinction with the standard KS potentials. Equation (26) has $\delta E_H[\tilde{n}]/\delta \tilde{n}$. That does *not* denote a simple change of functional variable. Rather, the functional derivative is taken *only* with respect to auxiliary densities representable by the prescribed auxiliary basis set $\{\alpha\}$. The same is true for Eq. (28). The same notation will appear for higher functional derivatives, see below. The functional derivative of the auxiliary density with respect to the KS density follows from Eq. (19) as

$$\begin{aligned} \frac{\delta \tilde{n}[n; \mathbf{r}_1]}{\delta n(\mathbf{r})} &= \sum_{ij} \alpha_i(\mathbf{r}_1) (\mathbf{G}^{-1})_{ij} \frac{\delta J_j[n]}{\delta n(\mathbf{r})} \\ &= \sum_{ij} \alpha_i(\mathbf{r}_1) (\mathbf{G}^{-1})_{ij} \int d\mathbf{r}_3 \frac{\alpha_j(\mathbf{r}_3)}{|\mathbf{r} - \mathbf{r}_3|}. \end{aligned} \quad (30)$$

Then, the auxH potential is

$$\begin{aligned} v_{\text{auxH}}(\mathbf{r}) &= \int d\mathbf{r}_1 \left[\left(\int d\mathbf{r}_2 \frac{\tilde{n}(\mathbf{r}_2)}{r_{12}} \right) \left(\sum_{ij} (\mathbf{G}^{-1})_{ij} \alpha_i(\mathbf{r}_1) \int d\mathbf{r}_3 \frac{\alpha_j(\mathbf{r}_3)}{|\mathbf{r} - \mathbf{r}_3|} \right) \right] \\ &= \sum_{ij} J_i[n] (\mathbf{G}^{-1})_{ij} \int d\mathbf{r}_3 \frac{\alpha_j(\mathbf{r}_3)}{|\mathbf{r} - \mathbf{r}_3|} \\ &= \int d\mathbf{r}_3 \frac{\tilde{n}(\mathbf{r}_3)}{|\mathbf{r} - \mathbf{r}_3|} \end{aligned} \quad (31)$$

where we used Eq. (13) twice, once each in going from the first to the second line and again from the second to the third line. The corresponding equation for the auxXC potential is

$$\begin{aligned} v_{\text{auxXC}}(\mathbf{r}) &= \int d\mathbf{r}_1 \left[\tilde{v}_{XC}[\tilde{n}(\mathbf{r}_1)] \left(\sum_{ij} \alpha_i(\mathbf{r}_1) (\mathbf{G}^{-1})_{ij} \int d\mathbf{r}_3 \frac{\alpha_j(\mathbf{r}_3)}{|\mathbf{r} - \mathbf{r}_3|} \right) \right] \\ &= \int d\mathbf{r}_3 \frac{\sum_i z_i \alpha_i(\mathbf{r}_3)}{|\mathbf{r} - \mathbf{r}_3|} = \int d\mathbf{r}_3 \frac{\tilde{n}_{xc}(\mathbf{r}_3)}{|\mathbf{r} - \mathbf{r}_3|} \end{aligned} \quad (32)$$

where

$$z_i := \sum_j (\mathbf{G}^{-1})_{ij} \int d\mathbf{r}_1 \tilde{v}_{XC}[\tilde{n}(\mathbf{r}_1)] \alpha_j(\mathbf{r}_1) \quad (33)$$

Computationally, much has been gained with this procedure. First, as in ordinary variational Coulomb fitting, we have

reduced the formal quartic scaling to cubic. Second, the part that needs a numerical integration, $E_{xc}[\tilde{n}]$ and Eq. (33), never references $n(\mathbf{r})$. Furthermore, the integral remaining in Eq. (32) can be done analytically and can be combined readily with Eq. (31) simply by summing the x and z vectors (no significant additional computational work load).

Something interesting and a bit peculiar has been introduced formally however with respect to the energy $E[n]$. We turn to that next.

4 ADFT energy ambiguity

Three distinct KS equations have emerged. Associated with each is a total energy. The issue of interest is their relationship.

The familiar, conventional KS equation of course involves no reference to an auxiliary density,

$$\left\{ -\frac{1}{2}\nabla^2 + v_H[n(\mathbf{r})] + v_{XC}[n(\mathbf{r})] + v_{ext}(\mathbf{r}) \right\} \phi_j(\mathbf{r}) = \epsilon_j \phi_j(\mathbf{r}). \quad (34)$$

It determines the target density $n(\mathbf{r})$, Eq. (1) for the given $v_{ext}(\mathbf{r})$. In molecular codes, the usual procedure for evaluating the total energy is to compute the KS kinetic energy, $T_s[n]$, explicitly but an alternative is to use the eigenvalue sum to avoid that (as is done in many solid-state codes), thus

$$E_{tot}[n] = \sum_j f_j \epsilon_j - \int d\mathbf{r} n(\mathbf{r}) \{ v_H(\mathbf{r})/2 + v_{XC}(\mathbf{r}) \} + E_{XC}[n]. \quad (35)$$

The second KS equation is the version with the “aux” potentials just introduced, Eqs. (31), $v_{auxH}(\mathbf{r})$, and (32) $v_{auxXC}(\mathbf{r})$, that handle $\tilde{n}[n]$. It is

$$\left\{ -\frac{1}{2}\nabla^2 + v_{auxH}[\tilde{n}_R(\mathbf{r})] + v_{auxXC}[\tilde{n}_R(\mathbf{r})] + v_{ext}(\mathbf{r}) \right\} \phi_{j,R}(\mathbf{r}) = \epsilon_{j,R} \phi_{j,R}(\mathbf{r}), \quad (36)$$

subject to

$$\begin{aligned} n_R(\mathbf{r}) &= \sum_j f_{j,R} |\phi_{j,R}|^2 \\ n_R(\mathbf{r}) &= \tilde{n}_R(\mathbf{r}) + \Delta n(\mathbf{r}) \quad \exists [\Delta n | \Delta n] \leq \epsilon_{tol} \\ \tilde{n}_R(\mathbf{r}) &= \sum_i^{N_{auxis}} x_i[n_R] \alpha_i(\mathbf{r}). \end{aligned} \quad (37)$$

The subscript “ R ” notation is to make explicit the existence of these last restrictions (37) upon the variational extremum from which Eq. (36) arises, namely those densities $n_R(\mathbf{r})$ that can be represented with the expansion (6) to precision criterion ϵ_{tol} imposed upon the Coulomb error from Δn . Observe that except in the limit of a complete auxiliary basis $\{\alpha\}$, the orbitals from (36) are not the same as delivered by the standard, conventional KS equation. Hence, neither is the density. Because Eq. (36) is explicitly functionally dependent only upon $\tilde{n}_R(\mathbf{r})$, at SCF convergence $n_R(\mathbf{r}) = \tilde{n}_R(\mathbf{r})$ to numerical precision. By that we mean the following. The Coulomb metric is used in practice by setting some tolerance between the fitted and orbital densities associated with Eq. (37). An exact SCF solution of that equation would reduce the difference between n_R and \tilde{n}_R to within that Coulomb metric tolerance. But the SCF solution itself has a separate tolerance. Therefore, the two densities can differ such that the Coulomb metric tolerance still is met but the Coulomb metric error is larger than would be the case if the two densities were identical. In essence $n_R = \tilde{n}_R + \Delta n_R + \Delta n_{SCF}$ such that nevertheless $[\Delta n_R + \Delta n_{SCF} | \Delta n_R + \Delta n_{SCF}] \leq \epsilon_{tol}$. Though both $n_R(\mathbf{r})$ and $\tilde{n}_R(\mathbf{r})$ are implicit functionals of the original $n(\mathbf{r})$, in the case of a fixed, finite $\{\alpha\}$ set there is no way to invert the transformation and recover $n(\mathbf{r})$.

Evaluation of the corresponding total energy via the eigenvalue sum in this case and use of Eq. (25) gives

$$\begin{aligned} E_{tot,R}[n_R] &= \sum_j f_{j,R} \epsilon_{j,R} - \int d\mathbf{r} n_R(\mathbf{r}) \{ v_{auxH}[n_R(\mathbf{r})]/2 \\ &\quad + v_{auxXC}[n_R(\mathbf{r})] \} + E_{XC}[n_R]. \end{aligned} \quad (38)$$

In this case, because the variation is restricted, appeal to the constrained search formulation of DFT [33, 34] and to the properties of $\tilde{E}_H[\tilde{n}]$, Eq. (16), might seem to show that this energy would lie at or above the energy for the true density, $E_{tot,R}[n_R] \geq E_{tot}[n]$. In fact, that argument does not hold. The problem is related to the intrinsic property of variational Coulomb fitting. Not only is $E_H[n_R]$ a lower bound to $E_H[n]$, but it is unknown whether $E_{XC}[n_R]$ has a bounding relationship with $E_{XC}[n]$. For ADFT as implemented in deMon2k, because there is no normalization constraint on Δn , it is easy to demonstrate that $E_{XC}[n_R]$ can be either above or below $E_{XC}[n]$. Simply consider Slater local exchange for very small magnitude Δn . One has

$$\begin{aligned} E_x[n] &= E_x[\tilde{n} + \Delta n] \\ &= -c_x \int d\mathbf{r} (\tilde{n} + \Delta n)^{4/3} \approx -c_x \int d\mathbf{r} [\tilde{n}^{4/3} + \frac{4}{3} \tilde{n}^{1/3} \Delta n] \\ &= E_x[\tilde{n}] - \frac{4}{3} c_x \int d\mathbf{r} \tilde{n}^{1/3} \Delta n. \end{aligned} \quad (39)$$

If $\Delta n > 0$ everywhere, $E_x[\tilde{n}]$ is an upper bound to $E_x[n]$ and conversely for $\Delta n < 0$ everywhere. Because $E_H[n]$ is much larger than $|E_{XC}[n]|$ however, a common numerical outcome for deMon2k is that the ADFT total energy is slightly below the ordinary KS energy. An informal observation is that Δn usually is positive near nuclear sites, which typically makes $E_x[\tilde{n}] > E_x[n]$ the most common result. There is a small additional complication from the SCF numerical precision issue. Instead of Eq. (38) what actually would be calculated in ADFT is

$$E_{\text{tot,R}}[n_R] \approx \sum_j f_{j,R} \epsilon_{j,R} - \int d\mathbf{r} \tilde{n}_R(\mathbf{r}) \{ v_{\text{auxH}}[\tilde{n}_R(\mathbf{r})]/2 + v_{\text{auxXC}}[\tilde{n}_R(\mathbf{r})] \} + E_{\text{XC}}[\tilde{n}_R]. \quad (40)$$

Depending on the relationship of the quality of the orbital and auxiliary basis sets, there can be a disparity between the two energies. Usually it is small.

The third possible KS equation follows from straight substitution of \tilde{n} into the original HK Levy–Lieb density functional (for the same v_{ext}) and variation with respect to \tilde{n} . This amounts to standard KS development but directly on densities that are linearly expandable under variational Coulomb fitting. (The notation can be deceptive. Again, this is *not* the same as a change of variable symbols in the original KS equation.) As a result, the functional derivatives $\delta\tilde{n}/\delta n$ do not appear and immediately one has

$$\left\{ -\frac{1}{2} \nabla^2 + v_{\text{auxH}}[\tilde{n}(\mathbf{r})] + v_{\text{XC}}[\tilde{n}(\mathbf{r})] + v_{\text{ext}}(\mathbf{r}) \right\} \tilde{\varphi}_j(\mathbf{r}) = \tilde{\epsilon}_j \tilde{\varphi}_j(\mathbf{r}). \quad (41)$$

Observe that the Hartree potential is the same as in the previous Euler equation, v_{auxH} but the XC potential is not. That inconsistency arises from the fact that free variation of a functional (in this case E_{xc}) with respect to a restricted domain without any accounting of the restriction is inconsistent. The notation includes the requirement that follows from the requirement of “straight substitution,” namely that the density from the KS orbitals

$$\eta(\mathbf{r}) := \sum_i v_i |\tilde{\varphi}_i(\mathbf{r})|^2, \quad (42)$$

(v_i are occupation numbers) *must be* $\tilde{n}(\mathbf{r})$ at every step of the SCF cycle. Although this is another form of an auxiliary DFT, we are unaware of its being used, perhaps because of the internal contradiction involved. In essence, however, it is fit-to-fit. A perfect second fit to $\tilde{n}^{1/3}$ for XC would be precisely that. From the success of fit-to-fit, we surmise that the formal inconsistency is not a prohibitive practical issue if the auxiliary basis sets are adequate. However, the existence of that inconsistency makes analysis of fit-to-fit for the finite $\{\alpha\}$ case essentially impossible. ADFT stands in contrast to that.

5 Time-dependent ADFT

It is straightforward to extend the ground state ADFT formulation of Sect. 3 to time-dependent DFT (tdDFT). The requirement is to compute the second functional derivatives of both the auxH and auxXC functionals. Those kernels then can be used directly in any standard time-dependent-DFT formulation in the adiabatic approximation.

The corresponding equations for the kernels are

$$\begin{aligned} f_{\text{auxH}}(\mathbf{r}_1, \mathbf{r}_2) &= \int d\mathbf{r}_3 \frac{\delta v_{\text{auxH}}(\mathbf{r}_1)}{\delta \tilde{n}(\mathbf{r}_3)} \frac{\delta \tilde{n}(\mathbf{r}_3)}{\delta n(\mathbf{r}_2)} \\ &= \sum_{ij} (\mathbf{G}^{-1})_{ij} \left(\int d\mathbf{r}_3 \frac{\alpha_i}{r_{13}} \right) \left(\int d\mathbf{r}_4 \frac{\alpha_j}{r_{24}} \right) \end{aligned} \quad (43)$$

$$\begin{aligned} f_{\text{auxXC}}(\mathbf{r}_1, \mathbf{r}_2) &= \int d\mathbf{r}_3 \frac{\delta v_{\text{auxXC}}(\mathbf{r}_1)}{\delta \tilde{n}(\mathbf{r}_3)} \frac{\delta \tilde{n}(\mathbf{r}_3)}{\delta n(\mathbf{r}_2)} \\ &= \sum_{i,j,k,l} (\mathbf{G}^{-1})_{ij} \left[\alpha_j \left| \frac{\delta^2 E_{\text{xc}}[\tilde{n}]}{\delta \tilde{n}(\mathbf{r}) \delta \tilde{n}(\mathbf{r}')} \right| \alpha_k \right] \\ &\quad (\mathbf{G}^{-1})_{kl} \left(\int d\mathbf{r}_3 \frac{\alpha_i}{r_{13}} \right) \left(\int d\mathbf{r}_4 \frac{\alpha_l}{r_{24}} \right) \end{aligned} \quad (44)$$

Note that $f_{\text{auxH}}(\mathbf{r}_1, \mathbf{r}_2)$ is not just the simple $1/r_{12}$ kernel. In this case, the quadratures on grids again involve only \tilde{n} and the auxiliary functions $\{\alpha\}$, as in the ground state case,

If Casida’s formulation of linear response td-DFT is used [37], the resulting linear response td-ADFT equations will be the same as previously obtained in [38] and [39]. However, these kernels also could be used in any other formulation, such as in a Sternheimer approach, or the extensions to obtain other excited state properties, e.g., Refs. [40, 41].

It is important to note that, by using the kernels just displayed, the resulting td-ADFT treatment will be internally consistent. This is different from other attempts to reduce the computational cost of td-DFT in which the perturbed density is fitted independently of the ground-state density [42, 43], or RI projectors are introduced directly into the td-DFT equations or other response equations [44–47]. It has been noted that the introduction of RI projectors for integrals of the type $[\alpha_i | f_{\text{xc}}[n] | \alpha_j]$ might lead to numerical instabilities [48]. Those instabilities arise when $f_{\text{xc}} \sim n^{-2/3}$, which is obtained in the orbital basis, diverges faster than the rate of decay of the $\alpha_i \alpha_j$ products.

6 Observations and conclusions

The foregoing analysis exposes the relationship of direct use of auxiliary densities from variational Coulomb fitting in DFT equations that is fundamental to ADFT. It also directs

attention to the restrictions imposed by the auxiliary basis set $\{\alpha\}$. Those can lead to unwelcome technical limitations. Recently we have encountered one in deMon2k.

The context is de-orbitalized meta-GGA XC functionals [49–51]. The only trait relevant here is that such functionals depend upon the density Laplacian $\nabla^2 n$. To support high-throughput exploration of large molecules, deMon2k provides a facility for automated generation of auxiliary basis sets [52]. To facilitate that, the code architecture assumes a certain structure for the auxiliary function sets. All Hermite–Gaussian-type auxiliary orbitals are primitives (i.e., uncontracted). The functions are grouped in subsets according to the Gaussian exponent. Each subset is comprised of all functions with angular momentum index from $\ell = 0$ to some maximum ℓ . The problem with this shared exponent structure seems to be that it can induce features in \tilde{n} that can lead to unphysical features in $\nabla^2 \tilde{n}$. Those in turn cause v_{auxXC} to be very different from v_{XC} for the deorbitalized meta-GGA functional known as SCAN-L. So far, the only resolution has been to use very rich auxiliary sets, e.g., GEN-A4 in deMon2k notation. Of course, this is a technical difficulty, not an intrinsic property of ADFT. But it is, we think, a useful illustration of the discussion given above about the implications of variation over a restricted representation space.

A possible relationship we have not discussed is with the local-scaling version of DFT [53, 54]. Structurally it resembles the ADFT definition, Eq. (25), of $E_{\text{auxXC}}[n]$ rather closely, though the motivating physics seems quite different. Whether results proven from local scaling DFT could be used to improve ADFT remains to be investigated.

Acknowledgements With thanks, we acknowledge many informative discussions about auxiliary density methods over the years with Andreas Köster, Gerald Geudtner, Patrizia Calaminici, Brett Dunlap, Notker Rösch, Sven Krüger, Jon Boettger, and Alberto Vela. Work supported by U.S. Dept. of Energy Energy Frontier Research Center grant DE-SC 0019330.

References

- Köster AM, Reveles JU, del Campo JM (2004) *J. Chem. Phys.* 121:3417
- Köster AM, Geudtner G, Alvarez-Ibarra A, Calaminici P, Casida ME, Carmona-Espíndola J, Dominguez VD, Flores-Moreno R, Gamboa GU, Goursoot A, Heine T, Ipatov A, de la Lande A, Janetzko F, del Campo JM, Mejía-Rodríguez D, Reveles JU, Vasquez-Perez J, Vela A, Zuniga-Gutierrez B, Salahub DR (2018) deMon2k, version 6. The deMon developers, Cinvestav, Mexico City
- Mejía-Rodríguez D, Fonseca E, Chen D, Cheng H-P, Trickey SB, Hennig R Density functional approximations, computational choices, and spin-crossover energies: Mn(taa) (unpublished)
- Slater JC (1951) *Phys Rev* 81:385–390
- Hohenberg P, Kohn W (1964) *Phys Rev* 136:B864
- Kohn W, Sham LJ (1965) *Phys Rev* 140:A1133
- Slater JC (1974). In: Daudel R, Pullman B (eds) *The world of quantum chemistry*. D. Reidel, Dordrecht, pp 3–15
- Roothaan CCJ (1951) *Rev Mod Phys* 23:69
- Hall GG (1951) *Proc R Soc Lond A* 205:541
- Baerends EJ, Ellis DE, Roos P (1973) *Chem Phys* 2:41
- Boys SF (1950) *Proc R Soc A* 200:542
- Sambe BH, Felton RH (1974) *J Chem Phys* 61:3862
- Sambe BH, Felton RH *ibid* 62: 1122 (1975)
- Koller J, Zaucer M, Azman A, *Naturforsch Z* (1976) Teil A 31:1022
- Dunlap BI, Connolly JWD, Sabin JR (1979) *J Chem Phys* 71:3396–3402
- Whitten JL (1973) *J Chem Phys* 58:4496
- Dunlap BI, Rösch N, Trickey SB (2010) *Mol Phys* 108:3167–3180
- Köster AM, del Campo JM, Janetzko F, Zuniga-Gutierrez B (2009) *J Chem Phys* 130:114106
- Mintmire JW, Sabin JR, Trickey SB (1982) *Phys Rev B* 26:1743 (**and references therein**)
- Boettger JC, Trickey SB (1985) *Phys Rev B* 32:1356(R)
- Trickey SB, Alford JA, Boettger JC (2004) In: Leszczynski J (ed) *Computational materials science*, vol 15 of *theoretical and computational chemistry*. Elsevier, Amsterdam, pp 171–228
- Belling T, Gauschopf T, Krüger S, Nörtemann F, Staufer M, Mayer M, Nasluzov VA, Birkenheuer U, Rösch N (1999). In: Keil F, Mackens M, Voss H, Werther J (eds) *Scientific computing in chemical engineering II*, vol 15. Springer, Berlin, Heidelberg, pp 66–73
- Belling T, Gauschopf T, Krüger S, Nörtemann F, Staufer M, Mayer M, Nasluzov VA, Birkenheuer U, Hu A, Matveev AV, Shor AM, Fuchs-Rohr MSK, Neyman KM, Ganyushin DI, Kerdcharoen T, Woiterski A, Gordienko AB, Majumder S, Rösch N (2004) *ParaGauss* version 3.0, Technische Universität München
- Fan L, Ziegler T (1991) *J Chem Phys* 95:7401
- Delley B, Wrinn M, Lüthi HP (1994) *J Chem Phys* 100:5785
- Andreas M, Köster (1998) *Habilitation thesis*, Universität Hannover
- Laikov DN (1997) *Chem Phys Lett* 281:151–156
- Laikov DN, Ustynyuk YA (2005) *Russ Chem Bull* 54:820–826
- Birkenheuer U, Gordienko AB, Nasluzov VA, Fuchs-Rohr MK, Rösch N (2005) *Int J Quantum Chem* 102:743–761
- Birkenheuer U (1999) *Habilitation thesis*, Technische Universität München
- Belling T, Gauschopf T, Krüger S, Nörtemann F, Staufer M, Mayer M, Nasluzov VA, Birkenheuer U, Hu A, Matveev AV, Shor AM, Fuchs-Rohr MSK, Neyman KM, Ganyushin DI, Kerdcharoen T, Woiterski A, Gordienko AB, Majumder S, Rotllant MH, Ramakrishnan R, Dixit G, Nikodem A, Soini T, Roderus M, Rösch N, *ParaGauss*-GPL. <https://www.github.com/alexei-matveev/paragauss-gpl>
- Harriman JE (1985) *J Chem Phys* 83:6283–6287
- Levy M (1979) *Proc Natl Acad Sci USA* 76:6062
- Lieb EH (1983) *Int J Quantum Chem* 24:243
- Alvarez-Ibarra A, Köster AM (2013) *J Chem Phys* 139:024102
- Pedroza-Montero JN, Morales JL, Geudtner G, Alvarez-Ibarra A, Calaminici P, Köster AM (2020) *J Chem Theory Comput* 16:2965–2974
- Casida ME (1995). In: Chong DP (ed) *Recent Advances in Density Functional Methods. Part I*. World Scientific, Singapore, p 155
- Ipatov A, Fouqueau A, Perez del Valle C, Cordova F, Casida ME, Köster AM, Vela A, Jamorski CJ (2006) *J Mol Struct THEO-CHEM* 762:179–191

39. Carmona-Espíndola J, Köster AM (2013) *Can J Chem* 91:795–803
40. Furche F (2001) *J Chem Phys* 114:5982
41. Furche F, Ahlrichs R (2002) *J Chem Phys* 117:7433
42. Bauernschmitt R, Häser M, Treutler O, Ahlrichs R (1997) *Chem Phys Lett* 264:573–578
43. van Gisbergen SJA, Snijders JG, Baerends EJ (1999) *Comput Phys Commun* 118:119
44. Komornicki A, Fitzgerald G (1993) *J Chem Phys* 98:1398
45. Casida ME, Jamorski C, Casida KC, Salahub DR (1998) *J Chem Phys* 108:4439
46. Görling A, Heinze HH, Ruzankin SP, Staufer M, Rösch N (1999) *J Chem Phys* 110:2785
47. Heinze HH, Görling A, Rösch N (2000) *J Chem Phys* 113:2088
48. Bosko S (2008) Ph.d. thesis, Technische Universität München
49. Mejía-Rodríguez D, Trickey SB (2017) *Phys Rev A* 96:052512
50. Mejía-Rodríguez D, Trickey SB (2018) *Phys Rev B* 98:115161
51. Mejía-Rodríguez D, Trickey SB (2020) *Phys Rev B* 102:121109(R)
52. Calaminici P, Janetzko F, Köster AM, Mejia-Olvera R, Zuniga-Gutierrez B (2007) *J Chem Phys* 126:44108
53. López-Boada R, Ludeña EV, Karasiev VV, Pino R (1997) *J Chem Phys* 107:6722–6731
54. Ludeña EV, Karasiev VV, López-Boada R, Valderrama E, Maldonado J (1999) *J Comput Chem* 20:155

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.