



HAL
open science

Multicomponent density functional theory with density fitting

Daniel Mejía-Rodríguez, Aurélien de la Lande

► **To cite this version:**

Daniel Mejía-Rodríguez, Aurélien de la Lande. Multicomponent density functional theory with density fitting. *The Journal of Chemical Physics*, 2019, 150 (17), pp.174115. 10.1063/1.5078596. hal-02358606

HAL Id: hal-02358606

<https://hal.science/hal-02358606>

Submitted on 12 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MULTICOMPONENT DENSITY FUNCTIONAL THEORY WITH DENSITY FITTING

Daniel Mejía-Rodríguez^{1,2}, Aurélien de la Lande^{1*}

1: Laboratoire de Chimie Physique, Université Paris Sud/CNRS, Université Paris Saclay. 15 avenue Jean Perrin 91405, Orsay, France.

2: Present address: Quantum Theory Project, Department of Physics, University of Florida, Gainesville, Florida 32611, USA

*To whom correspondence should be addressed: aurelien.de-la-lande@u-psud.fr

ABSTRACT

Multicomponent Density Functional Theory (MDFT) is a promising methodology to incorporate nuclear quantum effects, such as zero-point energy or tunneling, or to simulate other types of particles such as muons or positrons using particle densities as basic quantities. As for standard electronic DFT, a still ongoing challenge is to achieve the most efficient implementations. We introduce a multicomponent DFT implementation within the framework of auxiliary density functional theory, focusing on molecular systems comprised of electrons and quantum protons. We introduce a dual variational procedure to determine auxiliary electron and proton densities which leads to a succession of approximate energy expressions. Electronic and protonic fitted densities are employed i) in electron-electron, proton-proton and electron-proton classical Coulomb interactions, ii) in electron exchange-correlation, proton-proton exchange, and electron-proton correlation potentials. If needed, exact exchange among electrons or among protons are computed by the variational fitting of the corresponding Fock potential. The implementation is carried out in deMon2k. We test various electron proton correlation functionals on proton affinities. We find that auxiliary densities can be safely used in electron-electron, proton-proton, and electron-proton classical Coulomb interactions as well as in electron-proton correlation, albeit with some precautions related to the choice of the electronic auxiliary basis set that must be flexible enough. Computational tests reported in the last section indicate that introduction of density fitting in MDFT is clearly advantageous in terms of computational effort with good scaling properties with respect to the number of electron and protons treated at the DFT level.

I. INTRODUCTION

The Born-Oppenheimer (BO) approximation, by which one decouples the motion of light electrons from that of heavy nuclei, lies at the heart of our understanding of molecular structure or of chemical reactivity. The BO approximation (BOA) allows one to dramatically simplify the electronic-nuclear problem by building electronic Hamiltonians in which nuclear positions enter as parameters. The electronic energy obtained from electronic structure theory calculations then gives access to the potential energy surface which is obviously of prime importance for the understanding of reactivity or conformational changes of molecules. However, many interesting physical chemical problems require an account of the coupling with nuclei. This is, for example, the case of electron transfers¹ or coupled electron-proton transfers². Tunneling corrections to proton transfer reaction barriers taking place in various media (solids, solutions, enzymes...) is another class of phenomena where non-BO corrections are needed^{3,4}.

Various methodologies going beyond the BOA emerged in the literature after the pioneering work of Thomas in the late 1960s^{5,6}. Thomas computed the energy of ammonia, water, and hydrogen fluoride, expressing the proton wave functions as Slater determinants⁷. This original idea found flourishing developments in the following decades with the DEMO (Dynamical Extended Molecular Orbital)^{8,9}, the NOMO (Nuclear Orbital Molecular Orbital)^{10,11}, the MCMO (Multicomponent Molecular Orbital)¹², the NEO (Nuclear-Electronic-Orbital)^{13,14,15,16} or the APMO (All Particle Molecular Orbital) frameworks. The intrinsic appeal of these methods is that once an adequate *ansatz* for the total wave function is defined, one can reuse, *mutatis mutandis*, the algorithms developed for the electronic problem. As a good illustration of this strategy we shall mention, non-exhaustively, the coupling with Hartree-Fock (NOMO-HF¹⁷ or NEO-HF¹³), Multi-Configuration Self-Consistent Field (NEO-MCSCF)¹³, Configuration-Interaction (NOMO-CI¹⁸ or NEO-CI¹⁹, DFT (NOMO-DFT²⁰ or NEO-DFT²¹), Time-Dependent DFT (NEO-TDDFT)²² frameworks.

We are interested in multicomponent DFT (MDFT) which relies on a generalization of the Hohenberg and Kohn theorems to assemblies of quantum particles of different masses, charges or spins^{21,23}. MDFT is ideally suited to include nuclear quantum effects for a selected subset of nuclei while retaining most of the (heavy) nuclei in a classical description for which the BOA applies. MDFT holds great promise to include nuclear quantum effects (zero-point energy, tunneling...) within large molecular systems at appealing computational cost/accuracy ratio. The objective of this work is to test the reliability of MDFT in the context of density fitting^{24,25} to reduce computational cost. Indeed, in purely electronic DFT a great deal of effort has been spent to devise efficient algorithms to build and solve the Kohn Sham equations. Among the important developments made on that line, density fitting (DF) or the Resolution-of-the-Identity²⁶ (RI) approaches have been shown to lead to considerable speed-up in the

calculation of Coulomb contributions by avoiding the evaluation of four-center Electronic Repulsions Integrals (ERIS). The scaling of DFT implementations based on DF or RI is also more advantageous and they are now available in many academic or commercial packages. With respect to MDFT, Moncada et al. investigated a few years ago the compatibility of DF for multicomponent systems²⁷. In this work, the authors employed the electronic auxiliary density to accelerate the calculation of classical electron-electron Coulomb interaction and the electronic exchange-correlation terms. They did not use auxiliary densities for particles other than electrons. In the present work, we assess whether fitted densities for protons can also be introduced to further improve the efficiency of the implementations without jeopardizing the accuracy of MDFT results. This article is also the occasion to report our first implementation of MDFT in deMon2k.

In Section 2 we present the working equations of MDFT in deMon2k. A few electron-proton correlation functionals have been implemented and are reviewed on this occasion. Various levels of approximation corresponding to different degrees of involvement of the fitted proton densities are introduced in the energy expressions. Optimized nuclear auxiliary basis sets are provided. Details on a second order Self-Consistent-Field procedure implemented in deMon2k are also provided. In Section 3 we validate our implementation on calculations of proton affinities of amines or carboxylates, and on the calculation of hydrogen bond energies between organic molecules. We finally provide some performance tests showing the applicability of our implementation for a microsolvated coumarine molecule.

2. THEORETICAL FRAMEWORK

2.1 Kohn-Sham Multicomponent DFT

Our implementation of multicomponent DFT in deMon2k relies on the Kohn-Sham framework combined with the LC-GTO (Linear-Combination of Gaussian-Type-Orbitals) formalism. Our approach is similar to that found in the literature under various acronyms (*e.g.* NEO-DFT, APMO-DFT or NOMO-DFT). The implementation we have carried out is general and allows the treatment of many types of quantum particles (protons, deuterons, tritons, muons, positrons...). For the sake of simplicity though, we will restrict our attention to ensembles of electrons, quantum protons and classical nuclei of which there will be N_e , N_p and M , respectively. We invoke the BO separation between the quantum particles and the classical nuclei, while non-BO effects between quantum protons and electrons are to be captured by density functional approximations. In Hartree atomic units the full Hamiltonian for such a composite system reads:

$$\begin{aligned}
H = & - \sum_{i=1}^{N_e} \frac{1}{2} \nabla_i^2 - \sum_{a=1}^{N_p} \frac{1}{2m_p} \nabla_a^2 - \sum_{i=1}^{N_e} \sum_{A=1}^M \frac{Z_A}{|\mathbf{r}_i^e - \mathbf{R}_A|} + \sum_{a=1}^{N_p} \sum_{A=1}^M \frac{Z_A}{|\mathbf{r}_a^p - \mathbf{R}_A|} \\
& + \sum_{i=1}^{N_e} \sum_{j=1}^{N_e} \frac{1}{2} \frac{1}{|\mathbf{r}_i^e - \mathbf{r}_j^e|} + \sum_{a=1}^{N_p} \sum_{b=1}^{N_p} \frac{1}{2} \frac{1}{|\mathbf{r}_a^p - \mathbf{r}_b^p|} - \sum_{i=1}^{N_e} \sum_{a=1}^{N_p} \frac{1}{|\mathbf{r}_i^e - \mathbf{r}_a^p|}
\end{aligned} \tag{1}$$

In Eq. 1 m_p is the mass of the proton, Z_A the charge of nucleus A, \mathbf{r}^e , \mathbf{r}^p and \mathbf{R}_A are the positions of electrons (indices i, j), of quantum protons (indices a, b) and of classical nuclei, respectively (index A). The first two terms on the r.h.s. of eq. 1 are the kinetic energy operators for the electrons and the quantum protons, respectively. The third and fourth terms are the Coulomb operators between the electrons and the nuclei, and between the quantum protons and the nuclei, respectively. The fifth and sixth terms refer to the repulsion between electrons and between protons. Finally, the last term is the Coulomb attraction between electrons and quantum protons. Other terms can be added to this Hamiltonian, for instance, the electric field created by sets of external charges as customarily done in hybrid QM/MM (Quantum Mechanical/Molecular Mechanical) approaches^{28, 29}. The total energy of such a composite system can be written as³⁰

$$E = \int \rho^e(\mathbf{r}^e) v^e(\mathbf{r}^e) d\mathbf{r}^e + \int \rho^p(\mathbf{r}^p) v^p(\mathbf{r}^p) d\mathbf{r}^p + F[\rho^e, \rho^p] \tag{2}$$

where ρ^e and ρ^p are the electron and proton densities respectively and F is the universal multicomponent density functional. $v^e(\mathbf{r}^e)$ and $v^p(\mathbf{r}^p)$ are the external potential felt by the electrons and by the quantum protons, respectively. In the absence of external electromagnetic fields, they are defined as:

$$v^e(\mathbf{r}^e) = - \sum_A^M \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}^e|} \tag{3}$$

$$v^p(\mathbf{r}^p) = \sum_A^M \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}^p|} \tag{4}$$

In this setup, since $v^e(\mathbf{r}) = -v^p(\mathbf{r})$ one could use the simple v symbol for both electrons and protons. However, we keep distinct the more general v^e and v^p notations.

The Hohenberg-Kohn universal functional in MDFT is given by:

$$F[\rho^e, \rho^p] = \min_{\Psi \rightarrow \rho^e, \rho^p} \langle \Psi | \hat{T} + \hat{W} | \Psi \rangle \tag{5}$$

where

$$\hat{T} = \hat{T}_e + \hat{T}_p = - \sum_{i=1}^{N_e} \frac{1}{2} \nabla_i^2 - \sum_{a=1}^{N_p} \frac{1}{2m_p} \nabla_a^2 \quad (6)$$

and

$$\hat{W} = \hat{W}_{ee} + \hat{W}_{pp} + \hat{W}_{ep} = \sum_{i=1}^{N_e} \sum_{j=1}^{N_e} \frac{1}{2} \frac{1}{|\mathbf{r}_i^e - \mathbf{r}_j^e|} + \sum_{a=1}^{N_p} \sum_{b=1}^{N_p} \frac{1}{2} \frac{1}{|\mathbf{r}_a^p - \mathbf{r}_b^p|} - \sum_{i=1}^{N_e} \sum_{a=1}^{N_p} \frac{1}{|\mathbf{r}_i^e - \mathbf{r}_a^p|} \quad (7)$$

Importantly, it must be emphasized that in absence of any other external potential the MDFT approach developed in this article would not be appropriate. In such case the densities would become constant due to the translational invariance of the resulting Hamiltonian. We therefore assume here that $v^e(\mathbf{r}^e) \neq 0$, $v^p(\mathbf{r}^p) \neq 0$, and not constant. The electron-proton attraction is included in the definition of the universal functional F and, therefore, not present in $v^e(\mathbf{r}^e)$ and $v^p(\mathbf{r}^p)$. The ground state energy is obtained by minimization of this energy functional under the constraints that the total number of quantum particles are preserved, *i.e.* $N_e = \int \rho^e(\mathbf{r}^e) d\mathbf{r}^e$ and $N_p = \int \rho^p(\mathbf{r}^p) d\mathbf{r}^p$. Applying this procedure within the Kohn-Sham scheme, by defining a (fully) non-interacting auxiliary system, leads to a set of coupled electron-proton Kohn-Sham equations

$$\left(-\frac{1}{2} \nabla^2 + v_{KS}^e \right) \psi_i^e = \varepsilon_i^e \psi_i^e \quad (8)$$

$$\left(-\frac{1}{2m_p} \nabla^2 + v_{KS}^p \right) \psi_i^p = \varepsilon_i^p \psi_i^p \quad (9)$$

ψ_i^e and ψ_i^p are the Kohn-Sham MOs for the N_e electrons and N_p protons, respectively. In deMon2k these are expanded over atom-centered basis functions (LC-GTO framework) as:

$$\psi_i^e(\mathbf{r}^e) = \sum_{\mu} c_{\mu i}^e \mu^e(\mathbf{r}^e) \quad (10)$$

$$\psi_i^p(\mathbf{r}^p) = \sum_{\mu} c_{\mu i}^p \mu^p(\mathbf{r}^p) \quad (11)$$

The sums run over the number of electron- and proton-atomic basis functions, respectively. We use Greek letters to denote both basis functions and to index them. Although sharing similar mathematical structures, the electron and proton basis sets are different, owing to the fact that proton densities are much more localized than electron densities. Gaussian-type nuclear basis sets have been optimized by various groups^{10, 13} and can be used in deMon2k as well *via* the BASIS external library file of the program. Electron and proton densities are expressed in matrix notation as

$$\rho^{e/p}(\mathbf{r}^{e/p}) = \sum_{\mu,\nu} P_{\mu\nu}^{e/p} \mu^{e/p}(\mathbf{r}^{e/p}) \nu^{e/p}(\mathbf{r}^{e/p}) \quad (12)$$

where $P_{\mu\nu}^{e/p}$ is a matrix element of the density matrix for either the electrons or quantum protons.

These matrices are given by

$$P_{\mu\nu}^e = 2 \sum_i^{N_e/2} c_{\mu i}^e c_{\nu i}^e \quad (13)$$

$$P_{\mu\nu}^p = \sum_i^{N_p} c_{\mu i}^p c_{\nu i}^p \quad (14)$$

where we assume electronic closed-shell densities and high-spin open-shell proton densities; however, our code is operative for open-shell electronic structures too.

We express the total ground state energy of the composite system as

$$\begin{aligned} E^0 = & \sum_{\mu,\nu} P_{\mu\nu}^e H_{\mu\nu}^{e-core} + \frac{1}{2} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu}^e P_{\sigma\tau}^e \langle \mu^e \nu^e \| \sigma^e \tau^e \rangle + E_{xc}^e[\rho^e] \\ & + \sum_{\mu,\nu} P_{\mu\nu}^p H_{\mu\nu}^{p-core} + \frac{1}{2} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu}^p P_{\sigma\tau}^p \langle \mu^p \nu^p \| \sigma^p \tau^p \rangle + E_x^p[\rho^p] \\ & - \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu}^p P_{\sigma\tau}^e \langle \mu^p \nu^p \| \sigma^e \tau^e \rangle + E_c^{ep}[\rho^e, \rho^p] + E_{CN} \end{aligned} \quad (15)$$

or, introducing a more compact notation

$$\begin{aligned} E^0 = & \sum_{\mu,\nu} P_{\mu\nu}^e H_{\mu\nu}^{e-core} + \frac{1}{2} \langle \rho^e \| \rho^e \rangle + E_{xc}^e[\rho^e] + \sum_{\mu,\nu} P_{\mu\nu}^p H_{\mu\nu}^{p-core} + \frac{1}{2} \langle \rho^p \| \rho^p \rangle \\ & + E_x^p[\rho^p] - \langle \rho^e \| \rho^p \rangle + E_c^{ep}[\rho^e, \rho^p] + E_{CN} \end{aligned} \quad (16)$$

In these equations, the notation of the two-particle Coulomb integrals is the one used by the deMon2k community where the double-bar stands for the Coulomb operator $(1/|\mathbf{r}_1^{e/p} - \mathbf{r}_2^{e/p}|)$:

$$\langle \mu\nu \| \sigma\tau \rangle \equiv \int \int \frac{\mu(\mathbf{r}_1)\nu(\mathbf{r}_1)\sigma(\mathbf{r}_2)\tau(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \quad (17)$$

E_{CN} is the energy associated with the classical Coulomb repulsion between pairs of nuclei. $H_{\mu\nu}^{e/p-core}$ are the matrix elements of core Hamiltonians for electrons and protons which collect contributions from the kinetic energy operator and the interaction with the external field produced by the classical

nuclei. E_{xc}^e , E_x^p and E_c^{ep} are, respectively, the electron exchange-correlation energy, the proton exchange energy and the electron-proton correlation energies. Note that we neglect proton-proton correlation as a consequence of the highly local character of proton densities. For E_{xc}^e , we use standard energy functionals developed in the past for purely electronic DFT calculations, without reparameterization³¹. Standard LDA (Local Density Approximation), GGA (Generalized Gradient Approximation), Meta-GGA or hybrids (*i.e.* including a fraction of exact exchange), already available in deMon2k, can therefore be used in MDFT calculations. The E_x^p energy functional is taken to be exact exchange³². This is necessary to exactly remove self-interaction from the classical Coulomb proton-proton repulsion energy ($\langle \rho^p || \rho^p \rangle$). Exact exchange contributions entering E_{xc}^e (in the case of hybrid functionals) and E_x^p are obtained *via* the variational fitting of the Fock potentials as proposed in Reference 33. Correlation between electrons and protons is critical for reliable MDFT results due to the strong attractive nature of their interaction. The development of E_c^{ep} functionals is still in its infancy and the current subject of intense research^{30, 31, 34, 35, 36}. We will come back to the different functionals implemented in deMon2k. The Kohn-Sham matrix elements are obtained by differentiating the total energy with respect to the density matrix elements both for protons and electrons. $K_{\mu\nu}^e \equiv \frac{\partial E^0}{\partial P_{\mu\nu}^e}$ and $K_{\mu\nu}^p \equiv \frac{\partial E^0}{\partial P_{\mu\nu}^p}$

$$K_{\mu\nu}^e = H_{\mu\nu}^{e-core} + \sum_{\sigma,\tau} P_{\sigma\tau}^e \langle \mu^e \nu^e || \sigma^e \tau^e \rangle + \frac{\partial E_{xc}^e}{\partial P_{\mu\nu}^e} - \sum_{\sigma,\tau} P_{\sigma\tau}^p \langle \mu^e \nu^e || \sigma^p \tau^p \rangle + \frac{\partial E_c^{ep}}{\partial P_{\mu\nu}^e} \quad (18)$$

$$K_{\mu\nu}^p = H_{\mu\nu}^{p-core} + \sum_{\sigma,\tau} P_{\sigma\tau}^p \langle \mu^p \nu^p || \sigma^p \tau^p \rangle + \frac{\partial E_x^p}{\partial P_{\mu\nu}^p} - \sum_{\sigma,\tau} P_{\sigma\tau}^e \langle \mu^p \nu^p || \sigma^e \tau^e \rangle + \frac{\partial E_c^{ep}}{\partial P_{\mu\nu}^p} \quad (19)$$

2.2 MDFT with fitted densities

The occurrence of four energy terms involving quadruple loops over atomic orbitals in Eqs. 12-13 makes the calculation of the KS potential cumbersome. This difficulty can be overcome with density fitting techniques^{24, 25}. As mentioned in the Introduction, Moncada *et al.* tested a few years ago the accuracy of MDFT using electronic auxiliary density functions²⁷. We note that electron-proton correlation was neglected in this study ($E_c^{ep} = 0$). We pursue here the road opened by Moncada *et al.* by introducing both auxiliary electron and proton densities. A series of successive approximations can be made in the expression for the total energy. They correspond to different degrees of involvement of auxiliary electron or proton densities in the energy expressions. Table 1 summarizes these alternatives. The notation (X)^a means that the energy term X entering the MDFT energy is calculated with a fitted density for particles *a*. The different levels of approximations will now be detailed together with the expressions for the energies and Kohn-Sham potentials.

Table 1: Introduction of electron and proton fitted densities in MDFT calculations. A cross indicates that the corresponding energy term and related potential are calculated with the help of auxiliary densities.

	J_{ee}	J_{pp}	J_{ep}	E_{xc}^e	E_x^p	E_c^{ep}
E^0	-	-	-	-	-	-
E^1	X	-	-	-	-	-
E^2	X	-	-	X	-	-
E^3	X	X	-	X	X	-
E^4	X	X	X	X	X	-
E^5	X	X	X	X	X	X

Starting from E given by Eq. 9 or 10, a first step is to introduce electronic fitted densities in the calculation of J_{ee} alone or in J_{ee} and E_{xc} , in the spirit of Ref. 27, leading to the energy expression E^1 and E^2 , respectively. To write down these energies we need first to recall the working equations of DF. For electrons, the auxiliary density function $\tilde{\rho}^e$ is expressed as a linear combination of auxiliary basis functions \bar{k}^e :

$$\tilde{\rho}^e(\mathbf{r}^e) = \sum_{\bar{k}} x_{\bar{k}}^e \bar{k}^e(\mathbf{r}^e) \quad (20)$$

where the $x_{\bar{k}}^e$ are the so-called Coulomb fitting coefficients. In deMon2k, the \bar{k} are primitive Hermite GTOs³⁷. The $x_{\bar{k}}^e$ are obtained upon minimization of the self-interaction-energy $\varepsilon^2 = \frac{1}{2} \langle \rho^e - \tilde{\rho}^e | \rho^e - \tilde{\rho}^e \rangle$.²⁵ This procedure leads to a set of homogeneous equations:

$$\mathbf{X}^e = \mathbf{G}^{-1} \mathbf{J}^e \quad (21)$$

where \mathbf{J}^e is the so-called Coulomb vector,

$$\mathbf{J}^e = \begin{pmatrix} \sum_{\mu,\nu} P_{\mu\nu}^e \langle \mu^e \nu^e | \bar{1}^e \rangle \\ \sum_{\mu,\nu} P_{\mu\nu}^e \langle \mu^e \nu^e | \bar{2}^e \rangle \\ \vdots \\ \sum_{\mu,\nu} P_{\mu\nu}^e \langle \mu^e \nu^e | \bar{m}^e \rangle \end{pmatrix} \quad (22)$$

and,

$$\mathbf{G} = \begin{pmatrix} \langle \bar{1}^e | \bar{1}^e \rangle & \dots & \langle \bar{1}^e | \bar{m}^e \rangle \\ \vdots & \ddots & \vdots \\ \langle \bar{m}^e | \bar{1}^e \rangle & \dots & \langle \bar{m}^e | \bar{m}^e \rangle \end{pmatrix} \quad (23)$$

is the matrix of Coulomb repulsion integrals between auxiliary functions. Because ε^2 is positive semidefinite, the inequality $\frac{1}{2}\langle\rho^e\|\rho^e\rangle\geq\langle\rho^e\|\tilde{\rho}^e\rangle-\frac{1}{2}\langle\tilde{\rho}^e\|\tilde{\rho}^e\rangle$ always holds and the classical Coulomb repulsion among electrons appearing in Eq. 9 can be substituted by the r.h.s of the inequality, thereby eliminating the need of four center integrals.

$$E^1 = \sum_{\mu,\nu} P_{\mu\nu}^e H_{\mu\nu}^{e-core} + \sum_{\mu,\nu} P_{\mu\nu}^p H_{\mu\nu}^{p-core} + \langle\rho^e\|\tilde{\rho}^e\rangle - \frac{1}{2}\langle\tilde{\rho}^e\|\tilde{\rho}^e\rangle + \frac{1}{2}\langle\rho^p\|\rho^p\rangle - \langle\rho^e\|\rho^p\rangle \quad (24)$$

$$+ E_{xc}^e[\rho^e] + E_x^p[\rho^p] + E_c^{ep}[\rho^e, \rho^p]$$

This is a variational procedure that removes the quartic scaling factor over the number of electronic atomic orbitals. Density fitting has been proved to be a very robust approximation in electronic DFT calculations³⁸. In deMon2k, significant supplementary computational saving is obtained by introducing fitted densities in the evaluation of the XC contribution³⁹, *i.e.* replacing $E_{xc}^e[\rho^e]$ by $E_{xc}^e[\tilde{\rho}^e]$. This scheme is usually referred as Auxiliary DFT in the framework of electronic DFT^{38, 39}.

$$E^2 = \sum_{\mu,\nu} P_{\mu\nu}^e H_{\mu\nu}^{e-core} + \sum_{\mu,\nu} P_{\mu\nu}^p H_{\mu\nu}^{p-core} + \langle\rho^e\|\tilde{\rho}^e\rangle - \frac{1}{2}\langle\tilde{\rho}^e\|\tilde{\rho}^e\rangle + \frac{1}{2}\langle\rho^p\|\rho^p\rangle \quad (25)$$

$$- \langle\rho^e\|\rho^p\rangle + E_{xc}^e[\tilde{\rho}^e] + E_x^p[\rho^p] + E_c^{ep}[\rho^e, \rho^p]$$

The Kohn-Sham matrix elements read (in case of E^2 MDFT energy expression):

$$K_{\mu\nu}^e = H_{\mu\nu}^{e-core} + \sum_k \langle\mu^e\nu^e\|\bar{k}^e\rangle(x_k^e + z_k^e) - \sum_{\sigma,\tau} P_{\sigma\tau}^p \langle\mu^e\nu^e\|\sigma^p\tau^p\rangle + \frac{\partial E_c^{ep}[\rho^e, \rho^p]}{\partial P_{\mu\nu}^e} \quad (26)$$

$$K_{\mu\nu}^p = H_{\mu\nu}^{p-core} + \frac{1}{2} \sum_{\sigma,\tau} P_{\sigma\tau}^p \langle\mu^p\nu^p\|\sigma^p\tau^p\rangle - \sum_{\sigma,\tau} P_{\mu\nu}^e \langle\mu^p\nu^p\|\sigma^e\tau^e\rangle + \frac{\partial E_x^p[\rho^p]}{\partial P_{\mu\nu}^p}$$

$$+ \frac{\partial E_c^{ep}[\rho^e, \rho^p]}{\partial P_{\mu\nu}^p}$$

where³⁹

$$z_k^e = \sum_l G_{kl}^{-1} \langle\bar{l}^e|v_{xc}\rangle \quad (28)$$

Analogously, we now further introduce an auxiliary proton density. The repulsion energy term $\frac{1}{2}\langle\rho^p\|\rho^p\rangle$ of Eq. 19 can be substituted by $\langle\rho^p\|\tilde{\rho}^p\rangle - \frac{1}{2}\langle\tilde{\rho}^p\|\tilde{\rho}^p\rangle$, leading to the energy expression E^3 .

$$E^3 = \sum_{\mu,\nu} P_{\mu\nu}^e H_{\mu\nu}^{e-core} + \sum_{\mu,\nu} P_{\mu\nu}^p H_{\mu\nu}^{p-core} + \langle \rho^e \| \tilde{\rho}^e \rangle - \frac{1}{2} \langle \tilde{\rho}^e \| \tilde{\rho}^e \rangle + \langle \rho^p \| \tilde{\rho}^p \rangle - \frac{1}{2} \langle \tilde{\rho}^p \| \tilde{\rho}^p \rangle \quad (29)$$

$$- \langle \rho^e \| \rho^p \rangle + E_{xc}^e[\tilde{\rho}^e] + E_x^p[\tilde{\rho}^p] + E_c^{ep}[\rho^e, \rho^p]$$

We also introduce a variational density fitting procedure to evaluate exact exchange among protons ($E_x^p[\tilde{\rho}^p]$) like the one followed for exact exchange among electrons³³. The KS matrix elements for electrons remain the same as those derived from of E^2 MDFT energy expression but the expression for the KS matrix elements for protons now simplify to:

$$K_{\mu\nu}^p = H_{\mu\nu}^{p-core} + \sum_k \langle \mu^p \nu^p \| \bar{k}^p \rangle x_k^p - \sum_{\sigma,\tau} P_{\sigma\tau}^e \langle \mu^p \nu^p \| \sigma^e \tau^e \rangle + \frac{\partial E_x^p[\tilde{\rho}^p]}{\partial P_{\mu\nu}^p} + \frac{\partial E_c^{ep}[\rho^e, \rho^p]}{\partial P_{\mu\nu}^e} \quad (30)$$

At this stage, four-center electron-proton attraction integrals (EPAIs) remain present in the energy expression, hence in the Kohn-Sham potential of both protons and electrons. To remove these cumbersome integrals we seek a simplified way to calculate the attraction energy between electron and nuclei using again fitted densities. In this article we consider the possibility of replacing $\langle \rho^e \| \rho^p \rangle$ either by $\langle \tilde{\rho}^e \| \rho^p \rangle$, by $\langle \rho^e \| \tilde{\rho}^p \rangle$ or by $\langle \tilde{\rho}^e \| \tilde{\rho}^p \rangle$. Since in typical applications of MDFT the number of electrons, and of electronic basis functions, largely exceeds that of quantum nuclei it seems more advantageous to replace $\langle \rho^e \| \rho^p \rangle$ by $\langle \tilde{\rho}^e \| \rho^p \rangle$. This leads to the E^4 energy expression:

$$E^4 = \sum_{\mu,\nu} P_{\mu\nu}^e H_{\mu\nu}^{e-core} + \sum_{\mu,\nu} P_{\mu\nu}^p H_{\mu\nu}^{p-core} + \langle \rho^e \| \tilde{\rho}^e \rangle - \frac{1}{2} \langle \tilde{\rho}^e \| \tilde{\rho}^e \rangle + \langle \rho^p \| \tilde{\rho}^p \rangle \quad (31)$$

$$- \frac{1}{2} \langle \tilde{\rho}^p \| \tilde{\rho}^p \rangle - \langle \tilde{\rho}^e \| \rho^p \rangle + E_{xc}^e[\tilde{\rho}^e] + E_x^p[\tilde{\rho}^p] + E_c^{ep}[\rho^e, \rho^p]$$

with

$$\langle \tilde{\rho}^e \| \rho^p \rangle = \sum_{\mu,\nu} \sum_k P_{\mu\nu}^p x_k^e \langle \mu^p \nu^p \| \bar{k}^e \rangle \quad (32)$$

The matrix elements entering the proton KS potential are given by:

$$\frac{\partial \langle \tilde{\rho}^e \| \rho^p \rangle}{\partial P_{\mu\nu}^p} = \sum_k x_k^e \langle \mu^p \nu^p \| \bar{k}^e \rangle \quad (33)$$

Following the development proposed in Ref. 39 for expressing XC contributions in the context of Auxiliary DFT, we derive the electron KS potential from the density fitting coefficients *via* repeated application of the chain rule:

$$\frac{\partial \langle \tilde{\rho}^e \| \rho^p \rangle}{\partial P_{\mu\nu}^e} = \int \frac{\delta \langle \tilde{\rho}^e \| \rho^p \rangle}{\delta \tilde{\rho}^e(\mathbf{r}^e)} \frac{\partial \tilde{\rho}^e(\mathbf{r}^e)}{\partial P_{\mu\nu}^e} d\mathbf{r}^e \quad (34)$$

The derivative of the electronic fitted density with respect to the density matrix elements is obtained from Eq. 14:

$$\frac{\partial \tilde{\rho}^e(\mathbf{r}^e)}{\partial P_{\mu\nu}^e} = \sum_k \frac{\partial x_k^e}{\partial P_{\mu\nu}^e} \bar{k}^e(\mathbf{r}^e) \quad (35)$$

From Eqs. 15-16:

$$x_k^e = \sum_l (G_{kl}^e)^{-1} J_l^e = \sum_l (G_{kl}^e)^{-1} \sum_{\mu,\nu} P_{\mu\nu}^e \langle \mu^e \nu^e \| \bar{l}^e \rangle \quad (36)$$

therefore:

$$\frac{\partial \tilde{\rho}^e(\mathbf{r}^e)}{\partial P_{\mu\nu}^e} = \sum_k \left(\sum_l \langle \mu^e \nu^e \| \bar{l}^e \rangle (G_{lk}^e)^{-1} \right) \bar{k}^e(\mathbf{r}^e) \quad (37)$$

Inserting this equation into Eq. 25, one obtains

$$\frac{\partial \langle \tilde{\rho}^e \| \rho^p \rangle}{\partial P_{\mu\nu}^e} = \sum_l \langle \mu^e \nu^e \| \bar{l}^e \rangle \sum_k (G_{lk}^e)^{-1} \langle \bar{k}^e \| \rho^p \rangle \quad (38)$$

Finally, in E^5 the electron-proton correlation potential is further calculated with the help of the electron auxiliary density.

$$\begin{aligned} E^{(JXC)^{ep}} = & \sum_{\mu,\nu} P_{\mu\nu}^e H_{\mu\nu}^{e-core} + \sum_{\mu,\nu} P_{\mu\nu}^p H_{\mu\nu}^{p-core} + \langle \rho^e \| \tilde{\rho}^e \rangle - \frac{1}{2} \langle \tilde{\rho}^e \| \tilde{\rho}^e \rangle + \langle \rho^p \| \tilde{\rho}^p \rangle \\ & - \frac{1}{2} \langle \tilde{\rho}^p \| \tilde{\rho}^p \rangle - \langle \tilde{\rho}^e \| \rho^p \rangle + E_{xc}^e[\tilde{\rho}^e] + E_x^p[\tilde{\rho}^p] + E_c^{ep}[\tilde{\rho}^e, \rho^p] \end{aligned} \quad (39)$$

The matrix elements deriving from E_c^{ep} are obtained as for the classical Coulomb interaction. In the proton Kohn-Sham equations the matrix elements are obtained as $\partial E_c^{ep}[\tilde{\rho}^e, \rho^p] / \partial P_{\mu\nu}^e$ assuming fixed (auxiliary) electron density. As for Eq. 28, the matrix elements for the electron-proton correlation potential felt by the electrons are obtained with the chain rule

$$\frac{\partial E_c^{ep}[\tilde{\rho}^e, \rho^p]}{\partial P_{\mu\nu}^e} = \sum_l \langle \mu^e \nu^e \| \bar{l}^e \rangle \sum_k (G_{lk}^e)^{-1} \langle \bar{k}^e | v_{ce}^{ep} \rangle \quad (40)$$

where

$$v_{ce}^{ep} \equiv \frac{\delta E_c[\tilde{\rho}^e, \rho^p]}{\delta \tilde{\rho}^e} \quad (41)$$

2.3 Electron proton correlation functionals

A key ingredient of MDFT calculation is the inclusion of electron-proton correlation (E_c^{ep})^{21, 34}. The ability to devise reliable EPC functionals is still an active research field. EPC should correct for the mean-field approximation made by considering electron-proton attraction with a classical Coulomb attraction between the densities of the two types of quantum species. EPC is more difficult to address than electron-electron correlation because of the attractive nature of the interaction between electrons and protons⁴⁰. A consequence of the lack of EPC is the over-localization of proton densities, hence a strong overestimation of vibrational frequencies involving quantum particles. Some groups developed fully non local EPC functionals based on explicitly correlated electron nuclear wave functions^{31, 41, 42}. Another family of EPC functionals relies on adaptations of the Colle-Salvetti model of electron-electron correlation^{43, 44}. Such local functionals have been proposed by Nakai and coll.^{20, 45}, Tachikawa and coll.⁴⁶ and by Hammes-Schiffer and co-workers^{35, 36}. The former authors applied the EPC functional as a correction on top of the SCF procedure while in the latter authors introduced EPC self consistently. As already indicated this a currently active research field and not all functionals provide accurate, or even sometimes physically correct, results. However, they have been implemented, self consistently, in deMon2k for methodological purposes and the choice of a particular EPC functional remains the responsibility of the user. The EPC17-2 correlation energy reads³⁶:

$$E^{epc17-2} = - \int \frac{\rho^e(\mathbf{r})\rho^p(\mathbf{r})}{a - b(\rho^e(\mathbf{r}))^{1/2}(\rho^p(\mathbf{r}))^{1/2} + c\rho^e(\mathbf{r})\rho^p(\mathbf{r})} d\mathbf{r} \quad (42)$$

with a , b and c being three parameters set to 2.25, 2.4 and 6.6 by the authors of the original reference.

The Non-Singular-Form of the Colle-Salvetti functional developed by Inamura et al. reads²⁰:

$$E^{NSF07} = - \int \rho^e(\mathbf{r})\rho^p(\mathbf{r}) \frac{\pi Z^2 [(24 + (-9 + 2\sqrt{2})\pi)Z + 4(-4 + \pi)\sqrt{\pi}\beta]}{2\pi\beta^4 \exp\left[\frac{4Z(Z - \sqrt{\pi}\beta)}{\pi\beta^2}\right]} d\mathbf{r} \quad (43)$$

$$\beta(\mathbf{r}) = q \cdot (\rho^e(\mathbf{r}))^{1/3} \quad (44)$$

Z is the nuclear charge, i.e. 1.0 in the present context and β defines the correlation length. q is a parameter that was optimized to the value of 4.971 by the authors of reference²⁰. Recently Udagawa, Tsuneda and Tachikawa proposed an alternative formulation in which the EPC energy is given by⁴⁶:

$$E^{UTT14} = - \int \rho^e(\mathbf{r})\rho^p(\mathbf{r})(F_0(\beta) + F_1(\beta) + F_2(\beta)) d\mathbf{r} \quad (45)$$

$$F_0(\beta) = \frac{-3\pi}{\beta^3} \quad (46)$$

$$F_1(\beta) = \frac{32\beta^3 + (8\sqrt{2} - 32\sqrt{\pi})Z\beta^2 + 24Z^2\beta + (\sqrt{2\pi} - 8\sqrt{\pi})Z^3}{64\beta^5} \Phi(\beta)$$

$$F_2(\beta) = \frac{384\beta^6 - 192\sqrt{2\pi}Z\beta^5 + 384Z^2\beta^4 - 196\sqrt{2\pi}Z^3\beta^3 + 112Z^4\beta^2 - 15\sqrt{2\pi}Z^5\beta + 8Z^6}{1536\beta^8} \Phi^2(\beta)$$

The explicit form of $\Phi(\beta)$ was not specified in ⁴⁶. We derive it here following the Colle-Salvetti approach by requesting that the Jastrow factor $\phi(r, R) = \exp(-\beta^2 r^2)(1 - \Phi(R) \left(1 - r + \frac{1}{2}r^2 - \frac{1}{6}r^3\right))$ retains the normalization condition, i.e. by solving $4\pi \int (\phi^2(r, 0) - 2\phi(r, 0))r^2 dr = 0$. This leads to:

$$\Phi(\beta) = \frac{16(48\sqrt{2\pi} - 192\sqrt{\pi})\beta^6 + 4608\beta^5 + 16(18\sqrt{2\pi} - 144\sqrt{\pi})\beta^4 + 1792\beta^3 + 16\sqrt{2}\beta^3\sqrt{f}}{35\sqrt{2\pi} - 384\beta + 420\sqrt{2\pi}\beta^2 - 2048\beta^3 + 1152\sqrt{2\pi}\beta^4 - 3072\beta^5 + 768\sqrt{2\pi}\beta^6} \quad (47)$$

$$\begin{aligned} f &= 18432\pi\beta^6 \\ &+ 18432\beta^5\sqrt{\pi}(\sqrt{2} - 5) \\ &+ 1728\beta^4(2\sqrt{2}\pi + 15\pi + 24) \\ &+ 192\beta^3\sqrt{\pi}(71\sqrt{2} - 456) \\ &+ 72\beta^2(34\sqrt{2\pi} + 131\pi + 448) \\ &+ 2592\beta\sqrt{\pi}(\sqrt{2} - 8) \\ &+ 7(15\pi(4\sqrt{2} - 1) + 896) \end{aligned} \quad (48)$$

The UTT14 functional uses a different form of the $\beta(\mathbf{r})$ function, which is given in equation 10 of Ref. 46 and uses information from the exchange hole to improve the description of the correlation hole. EPC functional energies and potentials are calculated numerically on a standard fixed-type grid used for electronic DFT. This type of grid was found to be surprisingly accurate to integrate the EPC potentials based on the Colle-Salvetti model.

2.4 Nuclear auxiliary basis sets

The formalism described in the previous section supposes to define basis sets to expand the electron and proton auxiliary density functions. For electrons, deMon2k uses Hermine Gaussian auxiliary functions that are grouped by function sets sharing the same coefficients³⁷. Auxiliary basis sets are generated by an automatic procedure implemented in deMon2k that depends on the exponents of the atomic orbital basis set. The GEN-An auxiliary function sets contain groups of auxiliary functions with *s* and *spd* angular momenta. The index *n* determines the number of auxiliary function sets, *i.e.* the number of these sets increase with increasing *n*. A star indicates the addition of functions of higher angular momenta, namely, *f* and *g* functions. For protons, we have optimized various even-tempered

(ET) auxiliary basis sets with exponents ranging from twice the minimum to twice the maximum exponent in the original nuclear basis sets. We have called this auxiliary function sets ET-8*, ET-10 and ET-10* (Table 2). As in the electronic case, the proton auxiliary function sets are grouped by function sharing the same coefficients. It is important to note that since customary nuclear basis sets use high angular momentum even for the tightest functions, the nuclear auxiliary function basis sets developed are grouped in sets of *spd*- and *spdfg*-type functions, *ie* no *s*-type sets are included. The LMAX denotes the maximum angular momentum for a given set.

Table 2: Exponents for the even-tempered ET-8, ET-8* and ET-10* proton auxiliary basis sets.

ET-8 (LMAX = 2)	ET-8* (LMAX = 4)	ET-10* (LMAX = 4)
37.647578	37.647578	128.000000000
27.505323	27.505323	90.509667990
20.095388	20.095388	64.000000000
14.681689	14.681689	45.254834000
10.726442	10.726442	32.000000000
7.836738	7.836738	22.617417000
5.725520	5.725520	16.000000000
4.183064	4.183064	11.313708500
		8.000000000
		5.656854249

2.5 Resolution of the coupled MDFT Kohn-Sham equations

Experience from other groups has shown that convergence of the coupled MDFT Kohn-Sham equations is hard to achieve, taking sometimes thousands of cycles⁴⁷, a convergence difficulty presumably originating from the non-inclusion of electron proton correlation self-consistently. We decided to tackle this problem through three different schemes. The first one is to perform one minimization step for the electrons followed by one minimization step for the protons (minimization-after-minimization). The second one is to perform one minimization step for the electrons followed by the full convergence of the proton MDFT Kohn-Sham equations while leaving the electronic density frozen (convergence-after-minimization). The last one is to fully converge the electronic Kohn-Sham equations with the protonic density frozen followed by the full convergence of the proton Kohn-Sham equations while keeping the electronic density frozen (convergence-after-convergence). In practice, the minimization-after-minimization approach is not stable enough for the majority of cases, while the convergence-after-minimization and convergence-after-convergence perform rather similar.

The electronic SCF can be converged with Pulay's direct-inversion of iterative subspace (DIIS)^{48, 49} with the fitting coefficients as drivers⁵⁰, or the approximate second-order SCF (SOSCF) method originally proposed by Fischer and Almlöf and extended by others^{51, 52, 53}. If requested, the electronic SOSCF method starts when the electronic orbital gradient drops below 10^{-2} Ha for HF-type calculations or 10^{-3} Ha for DFT-type ones. Before the activation of the SOSCF approach, DIIS can be used to accelerate. Only the SOSCF scheme is available to accelerate the convergence of the nuclear SCF and it is started when the nuclear orbital gradient drops below 5×10^{-3} Ha. Electronic SCF convergence is achieved when the MinMax electronic SCF error achieves 10^{-6} Ha independent of the SCF accelerator chosen, while nuclear SCF convergence is achieved when the infinite norm of the nuclear orbital gradient is below 10^{-6} Ha.

III TESTS AND PERFORMANCE

III.1 Computational details

All calculations have been carried out with a recent developer version of deMon2k (4.3.4) in which we implemented the formalism described in the previous section. The electronic Def2-TZVPP atomic basis set⁵⁴ has been used on all atoms. For protons, unless otherwise stated, we used the even-tempered HET-5S5P3D nuclear basis set⁵⁵ to expand the proton MOs, combined with the ET-8* auxiliary nuclear basis set. The electron XC, as well as the electron-proton correlation energies and potentials, have been integrated numerically on pruned fixed grids comprising 99 radial shells and 590 Lebedev angular points. The B3LYP^{44, 56, 57} electron XC and EPC17-2^{35, 36} EPC functionals have been used. We also tested EPC functionals developed by Inamura et al^{20, 45} (NSF07, eq. 43-44) or by Tachikawa and coll. (UTT14, eq. 45-48).⁴⁶ It is important to note that all three EPC functionals were used as part of self-consistent calculations. SCF convergence criteria were set to 10^{-8} Ha, 10^{-6} Ha and 10^{-6} Ha for the total MDFT energy, proton energy and for the density fitting energies. For each molecule, we first ran a purely electronic DFT before switching on the MDFT using the electron KS MOs as the guess for MDFT SCF. The second order SCF algorithm described in the previous section was used for both electron and proton SCF. This computational set-up permits the SCF procedure to always converge smoothly, although sometimes slowly.

III.2 Proton affinities

We consider in this subsection the set of proton binding affinities recently devised by the Hammes-Schiffer group in the context of the validation of the EPC17-2 electron-proton correlation functional³⁶. The PA23 set contains six amines (NH_3 , CH_3NH_2 , $\text{CH}_3\text{CH}_2\text{NH}_2$, $\text{CH}_3\text{CH}_2\text{CH}_2\text{NH}_2$, $(\text{CH}_3)_2\text{NH}$ and $(\text{CH}_3)_3\text{N}$),

three aromatic molecules ($C_6H_5NH_2$, $C_6H_5O^-$ and $C_6H_5COO^-$), three inorganic compounds (CN^- , HS^- and NO_2^-) and finally eleven carboxylates ($HCOO^-$, CH_3COO^- , $CH_3CH_2COO^-$, $CH_3CH_2CH_2COO^-$, $CH_3CH_2CH_2CH_2COO^-$, CH_3COCOO^- , CH_2FCOO^- , CHF_2COO^- , CF_3COO^- , $ClCH_2COO^-$, $ClCH_2CH_2COO^-$). Experimental PA are available for these molecules. To be consistent with Ref. 36 we have used the EPC17-2 functional too. For molecules with more than one proton, only the most acidic is treated at the quantum level by MDFT. We optimized each molecule of the PA23 set at the B3LYP/Def2-TZVPP/GEN-A2* level of theory (Cartesian geometries can be found in SI). The proton affinities are defined as $PA(A) = -\Delta E + RT$ where ΔE is the variation of total energy of the molecule of interest upon protonation encompassing electronic, translational, rotational and vibrational contributions. If we neglect the change in rotational energy, approximate the change of translational energy to be $(-3/2)RT$ and assume the vibrational states of the molecule but that involving the added proton are constant, then³⁶ $PA(A) \approx E(A) - E(AH^+) + (5/2)RT$, where $E(A)$ is calculated with purely electronic DFT and $E(AH^+)$ with MDFT. We report Root Mean Square Errors (RMSE) and Mean Signed Error (MSE) and maximum absolute error (MAX) taking experimental values as references.

Table 3 gathers the results obtained from deMon2k. For this series of calculations, we have chosen the very large GEN-A4* electronic auxiliary basis. Let us first consider the E^1 energies in which auxiliary electronic density enters in the electron-electron classical repulsion only. As in Ref. 36 MDFT with the EPC17-2 functional gives good values with a global RMSE of 2.2 kcal/mol. This value is slightly larger than in Ref. 36 (RMSE 1.4 kcal/mol). Note that different (larger) electronic and nuclear basis sets were used in Ref. 36, which explains the different RMSE values.

Moving to E^2 in which auxiliary electronic densities are also used to integrate electronic XC contributions (see Table 1), we don't observe any decrease in accuracy as the RMSE and MUE are identical to the E^1 energy based scheme. This conclusion is reminiscent of the work of Moncada et al.²⁷ and of the experience gained in purely electronic DFT calculations³⁹. In passing, we confirm that neglecting electron-proton correlation leads to very unreliable results with RMSE values around 15 kcal/mol (E^2 column, values in brackets).

We now move to E^4 MDFT approximation in which we introduce auxiliary proton densities to evaluate electron-proton attraction (see Table 1). Note that we don't consider in this sub-section the E^3 MDFT energy approximation because only one proton is treated at the DFT level, therefore $E^3 = E^2$. The E^4 values are very close to E^2 . This shows that introducing DF for protons using the ET-8* auxiliary nuclear basis has no impact on the accuracy of the computed proton affinities.

We now consider the most optimized E^5 in which auxiliary densities are involved in all the Coulomb interaction terms, exchange terms and in the electron-electron or electron-proton correlation (see

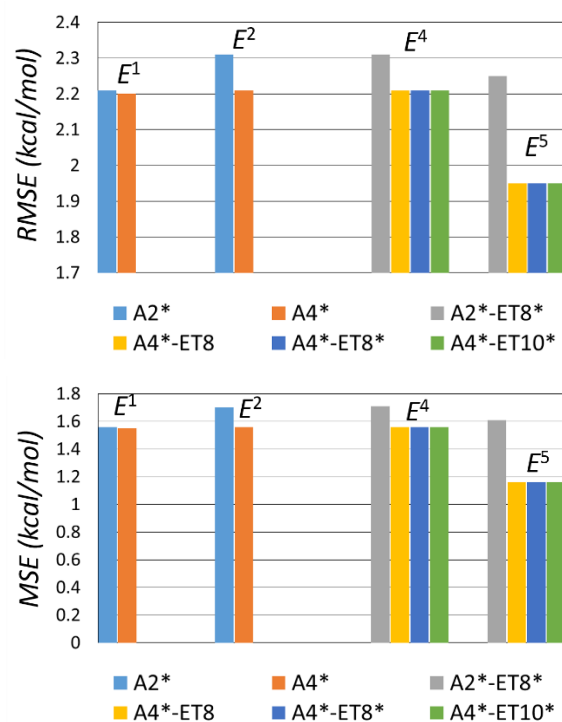
Table 1). Interestingly E^5 values are getting closer to experiments, suggesting subtle cancellation of errors due to the introduction of the electronic auxiliary density in the proton-electron correlation term. These are encouraging results that justify *a posteriori* the objective of this article to combine MDFT to density fitting techniques.

Table 3: Proton affinities (kcal/mol), computed with MDFT as $PA(A) = E(A) - E(AH^+) + (5/2)RT$. The numbers in parentheses corresponds to MDFT calculations with no electron-proton correlation.

	EXP	MDFT			
		E^1	E^2	E^4	E^5
		GEN-A4*	GEN-A4*	GEN-A4*	GEN-A4*
NH3	204.1	205.7	205.7 (188.2)	205.3	205.3
CH ₃ NH ₂	214.9	219.1	219.1 (202.2)	218.7	218.7
CH ₃ CH ₂ NH ₂	217.9	219.8	219.8 (202.9)	219.4	219.4
CH ₃ (CH ₂) ₂ NH ₂	219.3	221.2	221.2 (204.2)	220.8	220.8
(CH ₃) ₂ NH	222.1	223.0	223.0 (206.1)	222.7	222.7
(CH ₃) ₃ N	226.9	227.0	227.0 (210.1)	226.7	226.7
RMSE	-	2.17	2.16 (15.4)	2.16	1.86
MSE	-	1.90	1.91 (15.0)	1.91	1.52
C ₆ H ₅ NH ₂	211.0	211.3	211.4 (194.4)	211.4	211.0
C ₆ H ₅ O ⁻	351.4	350.9	350.9 (333.7)	350.9	350.5
C ₆ H ₅ COO ⁻	340.1	342.8	342.9 (325.6)	342.9	342.4
RMSE	-	1.59	1.60 (16.4)	1.60	1.42
MSE	-	0.84	0.84 (16.3)	0.84	0.45
CN ⁻	353.1	351.8	351.8 (334.5)	351.8	351.4
HS ⁻	353.1	355.2	355.3 (338.0)	355.3	354.8
NO ₂ ⁻	340.1	343.9	343.9 (326.6)	343.9	343.5
RMSE	-	2.62	2.64 (15.9)	2.64	2.40
MSE	-	1.56	1.59 (15.7)	1.59	1.17
HCOO ⁻	345.2	346.7	346.7 (329.5)	346.7	346.3
CH ₃ COO ⁻	348.5	350.6	350.6 (333.4)	350.6	350.2
CH ₃ CH ₂ COO ⁻	347.5	349.9	349.9 (332.6)	349.9	349.5
CH ₃ (CH ₂) ₂ COO ⁻	346.6	349.4	349.4 (332.1)	349.4	349.0
CH ₃ (CH ₂) ₃ COO ⁻	346.1	349.1	349.1 (331.9)	349.1	348.7
CH ₃ COCOO ⁻	333.5	334.6	334.6 (317.4)	334.6	334.2
CH ₂ FCOO ⁻	339.2	339.8	339.8 (322.6)	339.8	339.4
CHF ₂ COO ⁻	330.2	332.8	332.8 (315.6)	332.8	332.4
CF ₃ COO ⁻	322.6	326.0	326.0 (308.8)	326.0	325.6
ClCH ₂ COO ⁻	336.2	336.6	336.6 (319.3)	336.6	336.2

Cl(CH ₂) ₂ COO ⁻	340.8	338.1	338.1 (320.9)	338.1	337.7
RMSE	-	2.27	2.27 (15.8)	2.27	2.01
MSE	-	1.55	1.56	1.56	1.15
<hr/>					
Global					
RMSE		2.17	2.17 (15.4)	2.17	1.92
MSE		1.55	1.56 (15.6)	1.56	1.16
MAX		4.2	4.2(20.0)	4.4	3.8

Figure 1 provides a graphical representation of the dependence of RMSE, MSE and MAXimum error with the size of the auxiliary basis sets. Full tables of values can be found in Supplementary Material (SM, Table S1). For the electrons we compare results with GEN-A2* and GEN-A4*. For nuclei we test the ET8, ET8* and ET10* (Table 2) basis sets. Remember that with E^1 and E^2 schemes no auxiliary proton densities are used. With E^2 , moving from GEN-A2* to GEN-A4* induces a small decrease of RMSE and maximum error of by 0.1 and 0.2 kcal/mol, respectively. With E^5 , this same change of electronic auxiliary basis results in an improvement is of 0.3 kcal/mol on the RMSE. Figure 1 also shows that the nuclear auxiliary basis is well converged even with the smaller ET8 set.



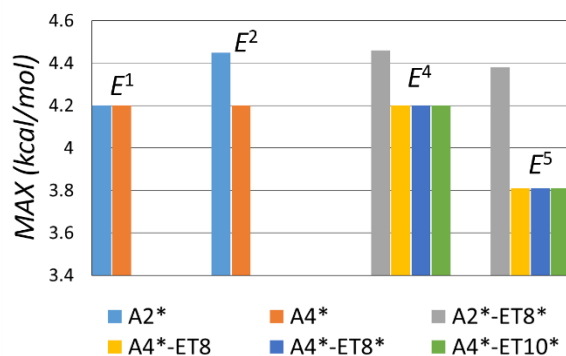


Figure 1: Convergence of proton affinities with electronic ($A2^*=GEN-A2^*$, $A4^*=GEN-A4^*$) and nuclear ($ET8$, $ET8^*$ and $ET10^*$) auxiliary basis sets.

To conclude this section, we compare different electron-proton correlation functionals to calculate proton affinities. We report in Table 4 the RMSE, MSE and MAX (complete results can be found in SM Table S2). The NSF07 is the least satisfactory functional, with a RMSE of 5.36 kcal/mol, well above the RMSE obtained with EPC17-2. On the other hand, UTT14 exhibits the best performance, with RMSE, MSE and maximum error of 1.85, 1.0 and 2 kcal/mol. These data indicate that accurate EPC functionals start to become available for MDFT calculations with protons.

Table 4: Influence of electron-proton correlation functional on the proton affinities (kcal/mol) of reference set. The E^1 energy expression is considered with $GEN-A2^*/ET8^*$ auxiliary sets.

	EPC17-2	NSF07	UTT14
RMSE	2.20	5.26	1.85
MSE	1.55	5.1	1.0
MAX	4.2	7.9	2.0

III.3 Hydrogen bond energies

In the second series of tests, we consider hydrogen-bonded dimers of molecules. The geometries are extracted from the S66 dataset⁵⁸. The dimers involve water, methanol, small peptide models, methylamine, uracyl, pyridine or acetamide. In these calculations, all the protons are treated as quantum particles. Results are gathered in Table 5. Compared to Table 3, we add the supplementary Column E^3 which corresponds to MDFT energies for Coulomb repulsion among protons evaluated with fitted proton densities. We take as reference MDFT results with $E^1/GEN-A4^*$. We find that application of DF to proton densities lead excellent results with RMSE below 0.05 kcal/mol w.r.t. the reference values (putting aside the $GEN-A2$ series). The maximum deviation is also small (0.15 kcal/mol).

Table 5: Hydrogen bond energy computed by MDFT/DF with different degrees of approximations. All values are given in kcal/mol. Reference geometries extracted from the S66 data set⁵⁸. RMSE, MUE and MAX are computed with respect to the $E^{(J)}$ ^g GEN-A4* series.

Partner 1	Partner 2	E^1	E^1	E^1	E^2	E^3	E^4	E^5
		GEN-A4*	GEN-A2	GEN-A2*	GEN-A2*	GEN-A2*	GEN-A2*	GEN-A2*
Water	Water	-5.92	-5.90	-5.92	-5.89	-5.89	-5.91	-5.89
Water	MeOH	-6.06	-6.07	-6.05	-6.05	-6.05	-6.08	-6.05
Water	MeNH2	-7.88	-7.79	-7.87	-7.88	-7.88	-7.91	-7.87
Water	Peptide	-8.18	-8.12	-8.17	-8.18	-8.19	-8.22	-8.20
MeOH	MeOH	-6.06	-6.03	-6.06	-6.05	-6.05	-6.06	-6.03
MeOH	MeNH2	-7.73	-7.57	-7.72	-7.75	-7.75	-7.80	-7.76
MeOH	Peptide	-8.20	-8.03	-8.19	-8.18	-8.18	-8.23	-8.19
MeOH	Water	-5.85	-5.76	-5.84	-5.81	-5.81	-5.84	-5.84
MeNH2	MeOH	-2.31	-2.32	-2.30	-2.33	-2.33	-2.32	-2.30
MeNH2	MeNH2	-3.13	-3.11	-3.12	-3.11	-3.11	-3.11	-3.06
MeNH2	Peptide	-3.48	-3.38	-3.48	-3.48	-3.48	-3.49	-3.45
MeNH2	Water	-7.84	-7.78	-7.83	-7.86	-7.86	-7.89	-7.85
Peptide	MeOH	-5.26	-5.21	-5.25	-5.24	-5.24	-5.23	-5.17
Peptide	MeNH2	-6.53	-6.35	-6.52	-6.56	-6.56	-6.56	-6.48
Peptide	Peptide	-6.73	-6.60	-6.72	-6.70	-6.70	-6.68	-6.64
Peptide	Water	-5.29	-5.25	-5.29	-5.25	-5.25	-5.22	-5.24
Uracil	Uracil	-17.21	-16.17	-17.19	-17.07	-17.07	-17.19	-17.19
Water	Pyridine	-7.36	-7.31	-7.35	-7.35	-7.35	-7.38	-7.36
MeOH	Pyridine	-7.40	-7.17	-7.39	-7.40	-7.40	-7.45	-7.41
AcOH	AcOH	-20.92	-20.31	-20.90	-20.87	-20.87	-21.07	-21.07
AcNH2	AcNH2	-16.70	-15.94	-16.68	-16.60	-16.60	-16.68	-16.63
AcOH	Uracil	-20.46	-19.67	-20.45	-20.37	-20.37	-20.55	-20.55
RMSE		-	0.36	0.01	0.05	0.05	0.05	0.05
MUE		-	0.21	0.01	0.03	0.03	0.04	0.04
MAX		-	1.04	0.02	0.14	0.14	0.15	0.15

We provide in Figure 4 a graphical representation of the difference of absolute energies of the dimers between the MDFT E^1 scheme, with GEN-A4* electronic auxiliary basis set, and the various DF based approximations. Table S3 in SM gathers the computed values. E^1 obtained with GEN-A2* slightly deviate from the reference with an RMSE of 0.07 kcal/mol. Upon introduction of proton fitted densities the RMSE and MUE increase to 1.39 and 1.29 kcal/mol, respectively, for E^2 and E^3 with GEN-A2* but drops to 0.06 kcal/mol with GEN-A4*. For very accurate absolute energies with E^2 and E^3 compared

to E^1 it is therefore recommended to use GEN-A4* which guarantees better energies compared to less extended electronic auxiliary basis set. As seen previously the three proton auxiliary basis sets gives same RMSE, attesting the completeness of the basis set. Moving to E^4 MDFT energies, where now electron-proton classical attraction is calculated from electronic density fitting, we find a RMSE of 0.36 and 3.3 kcal/mol when using GEN-A2* and GEN-A4*, respectively. A similar trend, although with smaller error is obtained with E^5 . It is therefore the introduction of electron-proton Coulomb attraction with electronic fitted density that deteriorate the accuracy of absolute energy in E^4 and E^5 schemes. This might be due to the non-variational substitution operated in Eq. 31 that is the source of this trend. Research is on-going in our laboratory to improve the E^4 and E^5 energy expressions. Note however that we discuss here absolute energies. Relative energies are expected to benefit from systematic error cancelations, as was seen for the proton affinities.

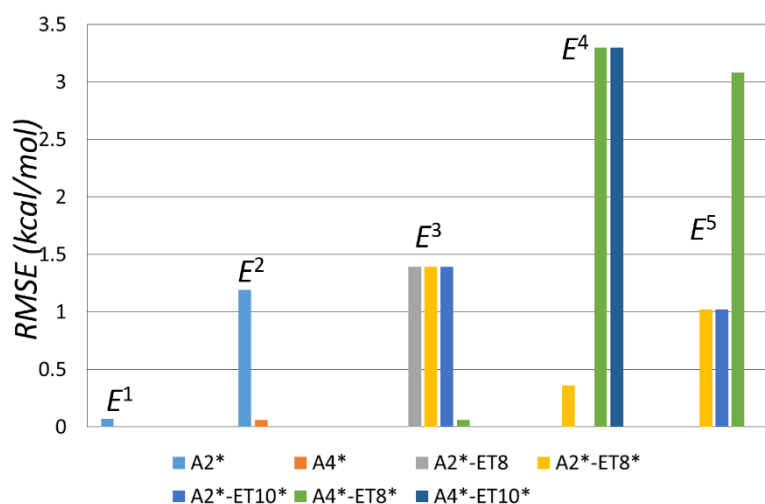


Figure 2: Dependence of RMSE with respect to E^1 /GEN-A4* for absolute energies (in kcal/mol) on electronic (A2*=GEN-A2*, A4*=GEN-A4*) and nuclear (ET8, ET8* and ET10*) auxiliary basis sets

III.3 Computational performance

In this last section, we assess the performance of MDFT calculations in deMon2k using density fitting. Toward this end, we consider a microsolvated coumarine molecule (Figure 3). The system encompasses a total of 125 atoms, among which 76 are protons, and 436 electrons. We consider seven situations for which we successively increase the number of protons treated by DFT, namely 6, 10, 16, 22, 28, 44 and 68. We have used the 3s3p nuclear basis set optimized by Nakai and co-workers¹⁰ in combination to the herein developed ET-8* auxiliary basis set, which should yield close to worst-case scenario in terms of computational performance. The EPC-17³⁶ and PBE⁵⁹ functionals are used for the electron-proton correlation and for the electron exchange-correlation respectively.

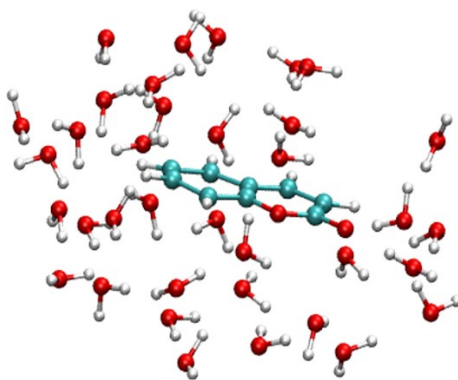
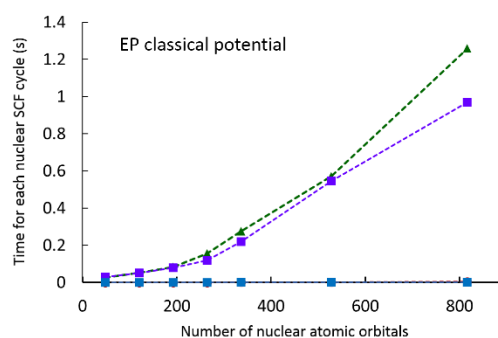
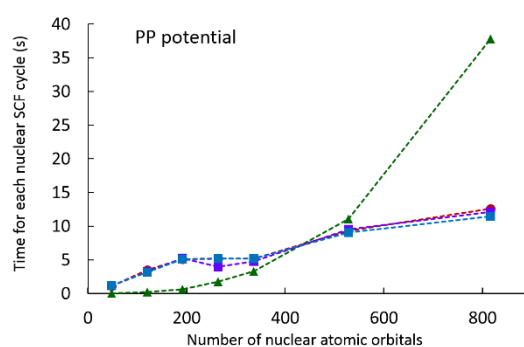


Figure 3: microsolvated coumarine molecule used for MDFT performance tests. The number of protons treated at the DFT level are progressively increased (see text for details).

Figure 4 depicts the total time to converge the SCF cycle and the time for each nuclear SCF spent i) in the proton-proton interaction, summing both contributions for classical Coulomb repulsion and exact exchange among protons (Top), ii) in the proton-electron classical Coulomb attraction (Middle), and iii) in the integration of the electron-proton correlation energy and potential (Bottom). Four MDFT energy expressions are considered, namely, E^2 , E^3 , E^4 and E^5 . The overall timing look deceptive has for less than 500 nuclear orbitals introduction of proton fitted densities is more time consuming. We also remark that when fitted densities are introduced in the electron-proton interaction SCF convergence is severely slowed down. In average two more time nuclear SCF cycles are required with E^4 and E^5 . More work will needed to improve SCF convergence, for example by introducing variational approach to treat electron-proton Coulomb interaction with density fitting. To go deeper we now analyze the computational cost of the different contributions for one SCF cycle.

We start by considering the proton-proton interactions ("PP potential", Top). The E^1 MDFT based calculation (no density fitting to proton densities) shows a rapid increase of the time spent in proton-proton interaction. For the largest system (816 proton atomic orbitals) proton-proton interactions represent 72% of the total time. The introduction of proton fitted densities in MDFT calculation based on the E^3 , E^4 or E^5 energy expressions permits to drastically alleviate computational cost with a much better scaling with the number of quantum protons. Exact exchange among protons remains the main time consuming task in proton-proton interactions. One approach that has been successfully used in the literature to solve this bottleneck is to completely avoid the calculation of proton-proton interactions³², effectively using a Hartree product for the nuclear wave function; such approach has also been implemented in deMon2k. We would also like to point out to recent improvements to the exact exchange module of deMon2k that can help to diminish the computational effort required for this task⁶⁰.

The classical Coulomb repulsion is obtained at negligible computational effort now that proton fitted densities are involved (<1% of the total SCF time). Moving to the calculation of electron-proton attraction (Figure 4, Middle) we can assess the improvement of the E^4 approximation over the E^3 since now electron-proton attraction is evaluated from electronic fitted density. First, we note that the electron-proton classical interaction is not the most time consuming part compared to calculation of the PP potential or of the EPC potential. Yet is satisfactory to see that the introduction of the fitted electronic density makes this part of MDFT calculations also negligible in terms of computational efforts. We finally move to the numerical evaluation of electron-proton correlation energy and potential (Figure 4, Bottom). It is a time consuming part of the MDFT calculation. MDFT relying on energy expressions E^2 , E^3 or E^4 exhibits very similar trends with the number of proton treated at the DFT level. In fact, it is the number of grid points that determines the computational cost for EPC. The introduction of electronic fitted density in E^5 permits a reduction of the computational cost by a factor of around 2.



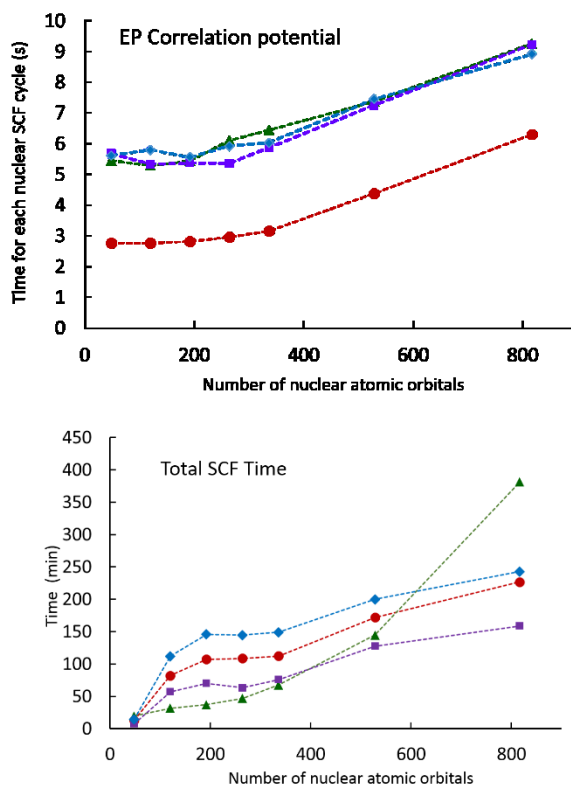


Figure 4: Average time (s) per nuclear SCF cycle spent for the evaluation of proton-proton interaction (Top, numerical integration on a grid of points) and for the evaluation of electron-proton correlation potential (Bottom). Lower panel: total SCF time in minutes. Green, violet, blue and red data correspond to MDFT calculations based on energies namely E^2 , E^3 , E^4 and E^5 respectively.

CONCLUSION

In this article, we have reported an original implementation of multicomponent DFT in deMon2k. The implementation has been parallelized using the MPI (Message Passing Interface) protocol. Resolution of the coupled electron/proton Kohn-Sham equations is carried out using an approximate second-order SCF algorithm that we implemented in the program. This algorithm greatly improves SCF convergence. We have introduced variationally fitted auxiliary densities for both electrons and protons. The ET8, ET8* and ET10* auxiliary basis sets have been optimized to express auxiliary proton densities. Five approximate energy expressions using fitted densities have been derived. Our tests indicate that all these can be used safely provided that sufficiently flexible auxiliary electronic basis sets are used (GEN-A4*). Introduction of fitted densities for electrons and protons completely remove the computational demand for classical Coulomb interactions. Exact exchange among protons and electron-proton correlation are the most time-consuming parts of MDFT as implemented in deMon2k. Further developments are underway in our laboratory to further improve computational efficiency. Computational advantages might also be obtained by switching to a standard exchange functional instead of using exact-exchange among protons.

In summary, the present implementation of MDFT in deMon2k provides a solid basis for our future developments of the MDFT methodology, which despite being still in its infancy, holds great promise for simulating several physicochemical processes involving nuclear quantum effects or "exotic" particles. The efficiency of the code (*e.g.* SCF stability, computational performance) combined with continuous progress on electron-proton correlation functionals is encouraging for applications to real-world problems. In this context, we would like to mention that, although we decided not to focus on the quality of the proton densities obtained with the EPC functionals, we have preliminary indications that EPC17-2 and UTT14 perform rather similar, and that the NSF07 functional does not improve over computations without EPC functionals. deMon2k is an open-source computer code under the license agreement and is free for academics. Interested users may contact the authors to obtain copies of the MDFT routines.

SUPPLEMENTARY MATERIAL

Supplementary material (SM) contains Tables gathering proton affinities with different combinations of electronic and proton auxiliary basis sets and with different electron proton correlation energy functional. These data were used for Figure 1 and Table 4. Optimized geometries (in Cartesian coordinates) used to build Table 3 can also be found in SM. We also provide absolute energies of hydrogen bond complexes that were used to draw Figure 2.

ACKNOWLEDGMENTS

We are grateful to Prof. Flores-Moreno for his help in the first steps of our MDFT project as well as for several illuminating discussion. This work was supported by the French *Agence Nationale de la Recherche* (Convention ANR-13-JS08-0010-01).

REFERENCES

1. Dogonadze, R. R.; Alexander, M. K.; Chernenko, A. A., THEORY OF HOMOGENEOUS AND HETEROGENEOUS ELECTRONIC PROCESSES IN LIQUIDS. *Russ. Chem. Rev.* **1965**, *34* (10), 759.
2. Hammes-Schiffer, S.; Stuchebrukhov, A. A., Theory of Coupled Electron and Proton Transfer Reactions. *Chem. Rev.* **2010**, *110* (12), 6939.
3. Alhambra, C., *et al.*, Canonical Variational Theory for Enzyme Kinetics with the Protein Mean Force and Multidimensional Quantum Mechanical Tunneling Dynamics. Theory and Application to Liver Alcohol Dehydrogenase. *J. Phys. Chem. B* **2001**, *105* (45), 11326.
4. Hwang, J.-K.; Warshel, A., How Important Are Quantum Mechanical Nuclear Motions in Enzyme Catalysis? *J. Am. Chem. Soc.* **1996**, *118* (47), 11745.
5. Thomas, I. L., Protonic Structure of Molecules. I. Ammonia Molecules. *Physical Review* **1969**, *185* (1), 90.
6. Thomas, I. L., The protonic structure of methane, ammonia, water, and hydrogen fluoride. *Chem. Phys. Lett.* **1969**, *3* (9), 705.
7. Slater, J. C., The Quantum Theory of the Equation of State. *Physical Review* **1931**, *38* (2), 237.

8. Tachikawa, M.; Mori, K.; Suzuki, K.; Iguchi, K., Full variational molecular orbital method: Application to the positron-molecule complexes. *Int. J. Quantum Chem* **1998**, *70* (3), 491.
9. Ishida, M.; Tachikawa, M.; Tokiwa, H.; Mori, K.; Ishii, A., First principles calculation for hydrogen/positronium adsorption on an Si(111) surface using the dynamical extended molecular orbital method. *Surf. Sci.* **1999**, *438* (1), 47.
10. Nakai, H., Nuclear orbital plus molecular orbital theory: Simultaneous determination of nuclear and electronic wave functions without Born–Oppenheimer approximation. *Int. J. Quantum Chem* **2007**, *107* (14), 2849.
11. Tachikawa, M.; Mori, K.; Nakai, H.; Iguchi, K., An extension of ab initio molecular orbital theory to nuclear motion. *Chem. Phys. Lett.* **1998**, *290*, 437.
12. Tachikawa, M., Multi-component molecular orbital theory for electrons and nuclei including many-body effect with full configuration interaction treatment: isotope effects on hydrogen molecules. *Chem. Phys. Lett.* **2002**, *360* (5), 494.
13. Webb, S. P.; Iordanov, T.; Hammes-Schiffer, S., Multiconfigurational nuclear-electronic orbital approach: Incorporation of nuclear quantum effects in electronic structure calculations. *J. Chem. Phys.* **2002**, *117* (9), 4106.
14. Iordanov, T.; Hammes-Schiffer, S., Vibrational analysis for the nuclear–electronic orbital method. *J. Chem. Phys.* **2003**, *118* (21), 9489.
15. Pak, M. V.; Swalina, C.; Webb, S. P.; Hammes-Schiffer, S., Application of the nuclear–electronic orbital method to hydrogen transfer systems: multiple centers and multiconfigurational wavefunctions. *Chem. Phys.* **2004**, *304* (1), 227.
16. Swalina, C.; Pak, M. V.; Hammes-Schiffer, S., Alternative formulation of many-body perturbation theory for electron–proton correlation. *Chem. Phys. Lett.* **2005**, *404* (4), 394.
17. Nakai, H., Simultaneous determination of nuclear and electronic wave functions without Born–Oppenheimer approximation: Ab initio NO+MO/HF theory. *Int. J. Quantum Chem* **2002**, *86* (6), 511.
18. Nakai, H.; Sodeyama, K.; Hoshino, M., Non-Born–Oppenheimer theory for simultaneous determination of vibrational and electronic excited states: ab initio NO+MO/CIS theory. *Chem. Phys. Lett.* **2001**, *345* (1), 118.
19. Skone, J. H.; Pak, M. V.; Hammes-Schiffer, S., Nuclear-electronic orbital nonorthogonal configuration interaction approach. *J. Chem. Phys.* **2005**, *123* (13), 134108.
20. Imamura, Y.; Kiryu, H.; Nakai, H., Colle-Salvetti-type correction for electron-nucleus correlation in the nuclear orbital plus molecular orbital theory. *J. Comput. Chem.* **2007**, *29* (5), 735.
21. Chakraborty, A.; Pak, M. V.; Hammes-Schiffer, S., Development of Electron-Proton Density Functionals for Multicomponent Density Functional Theory. *Phys. Rev. Lett.* **2008**, *101* (15), 153001.
22. Yang, Y.; Culpitt, T.; Hammes-Schiffer, S., Multicomponent Time-Dependent Density Functional Theory: Proton and Electron Excitation Energies. *J. Phys. Chem. Lett.* **2018**, *9* (7), 1765.
23. Kreibich, T.; van Leeuwen, R.; Gross, E. K. U., Multicomponent density-functional theory for electrons and nuclei. *Phys. Rev. A* **2008**, *78*, 022501.
24. Mintmire, J. W.; Dunlap, B. I., Fitting the Coulomb potential variationally in linear-combination-of-atomic-orbitals density-functional calculations. *Phys. Rev. A* **1982**, *25* (1), 88.
25. Dunlap, B. I.; Rösch, N.; Trickey, S. B., Variational fitting methods for electronic structure calculations. *Mol. Phys.* **2010**, *108* (21-23), 3167.
26. Kendall, R. A.; Früchtl, H. A., The impact of the resolution of the identity approximate integral method on modern ab initio algorithm development. *Theor. Chem. Acc.* **1997**, *97* (1), 158.
27. Moncada, F.; Posada, E.; Flores-Moreno, R.; Reyes, A., Non-Born–Oppenheimer self-consistent field calculations with cubic scaling. *Chem. Phys.* **2012**, *400*, 103.
28. Warshel, A.; Levitt, M., Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103* (2), 227.
29. Salahub, D., *et al.*, QM/MM Calculations with deMon2k. *Molecules* **2015**, *20* (3), 4780.
30. Chakraborty, A.; Pak, M. V.; Hammes-Schiffer, S., Properties of the exact universal functional in multicomponent density functional theory. *J. Chem. Phys.* **2009**, *131* (12), 124115.

31. Sirjoosingh, A.; Pak, M. V.; Hammes-Schiffer, S., Multicomponent density functional theory study of the interplay between electron-electron and electron-proton correlation. *J. Chem. Phys.* **2012**, *136* (17), 174114.
32. Auer, B.; Hammes-Schiffer, S., Localized Hartree product treatment of multiple protons in the nuclear-electronic orbital framework. *J. Chem. Phys.* **2010**, *132* (8), 084110.
33. Mejía-Rodríguez, D.; Köster, A. M., Robust and efficient variational fitting of Fock exchange. *J. Chem. Phys.* **2014**, *141* (12), 124114.
34. Kreibich, T.; Gross, E. K. U., Multicomponent Density-Functional Theory for Electron and Nuclei. *Phys. Rev. Lett.* **2001**, *86* (14), 2984.
35. Yang, Y.; Brorsen, K. R.; Culpitt, T.; Pak, M. V.; Hammes-Schiffer, S., Development of a practical multicomponent density functional for electron-proton correlation to produce accurate proton densities. *J. Chem. Phys.* **2017**, *147* (11), 114113.
36. Brorsen, K. R.; Yang, Y.; Hammes-Schiffer, S., Multicomponent Density Functional Theory: Impact of Nuclear Quantum Effects on Proton Affinities and Geometries. *J. Phys. Chem. Lett.* **2017**, *8* (15), 3488.
37. Köster, A. M., Hermite Gaussian auxiliary functions for the variational fitting of the Coulomb potential in density functional methods. *J. Chem. Phys.* **2003**, *118* (22), 9943.
38. Calaminici, P., *et al.*, Auxiliary Density Functional Theory: From Molecules to Nanostructures. In *Handbook of Computational Chemistry*, Leszczynski, J., Ed. Springer Netherlands: Dordrecht, 2016; pp 1.
39. Köster, A. M.; Reveles, J. U.; del Campo, J. M., Calculation of exchange-correlation potentials with auxiliary function densities. *J. Chem. Phys.* **2004**, *121* (8), 3417.
40. Pak, M. V.; Hammes-Schiffer, S., Electron-Proton Correlation for Hydrogen Tunneling Systems. *Phys. Rev. Lett.* **2004**, *92* (10), 103002.
41. Sirjoosingh, A.; Pak, M. V.; Swalina, C.; Hammes-Schiffer, S., Reduced explicitly correlated Hartree-Fock approach within the nuclear-electronic orbital framework: Theoretical formulation. *J. Chem. Phys.* **2013**, *139* (3), 034102.
42. Sirjoosingh, A.; Pak, M. V.; Brorsen, K. R.; Hammes-Schiffer, S., Quantum treatment of protons with the reduced explicitly correlated Hartree-Fock approach. *J. Chem. Phys.* **2015**, *142*, 214107.
43. Colle, R.; Salvetti, O., Approximate calculation of the correlation energy for the closed shells. *Theoret. Chim. Acta* **1974**, *37* (4), 329.
44. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37* (2), 785.
45. Imamura, Y.; Tsukamoto, Y.; Kiryu, H.; Nakai, H., Extension of density functional theory to nuclear orbital plus molecular orbital theory: self-consistent calculations with the Colle-Salvetti Electron-Nucleus correlation functional. *Bull. Chem. Soc. Jpn.* **2009**, *82* (9), 1133.
46. Udagawa, T.; Tsuneda, T.; Tachikawa, M., Electron-nucleus correlation functional for multicomponent density-functional theory. *Phys. Rev. A* **2014**, *89* (5), 052519.
47. Roberto, F. M., *et al.*, LOWDIN: The any particle molecular orbital code. *Int. J. Quantum Chem* **2014**, *114* (1), 50.
48. Pulay, P., Convergence acceleration of iterative sequences. the case of scf iteration. *Chem. Phys. Lett.* **1980**, *73* (2), 393.
49. P., P., Improved SCF convergence acceleration. *J. Comput. Chem.* **1982**, *3* (4), 556.
50. Köster, A. M.; Campo, J. M. d.; Janetzko, F.; Zuniga-Gutierrez, B., A MinMax self-consistent-field approach for auxiliary density functional theory. *J. Chem. Phys.* **2009**, *130* (11), 114106.
51. Fischer, T. H.; Almlof, J., General methods for geometry and wave function optimization. *J. Phys. Chem.* **1992**, *96* (24), 9768.
52. Chaban, G.; Schmidt, M. W.; Gordon, M. S., Approximate second order method for orbital optimization of SCF and MCSCF wavefunctions. *Theor. Chem. Acc.* **1997**, *97* (1), 88.
53. Neese, F., Approximate second-order SCF convergence for spin unrestricted wavefunctions. *Chem. Phys. Lett.* **2000**, *325* (1), 93.

54. Weigend, F.; Ahlrichs, R., Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297.
55. Feller, D. F.; Ruedenberg, K., Systematic approach to extended even-tempered orbital bases for atomic and molecular calculations. *Theoret. Chim. Acta* **1979**, *52* (3), 231.
56. Becke, A. D., Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38* (6), 3098.
57. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648.
58. Řezáč, J.; Riley, K. E.; Hobza, P., S66: A Well-balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theor. Comput.* **2011**, *7* (8), 2427.
59. Perdew, J. P.; Burke, K.; Ernzerhof, M., Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865.
60. Delesma, F. A.; Geudtner, G.; Mejia-Rodriguez, D.; Calaminici, P.; Köster, A. M., Range-separated Hybrid Functionals with Variational Fitted Exact Exchange. *J. Chem. Theory Comput.* **2018**, *14*, 5608.