

On extending Kohn-Sham density functionals to systems with fractional number of electrons

Cite as: J. Chem. Phys. **146**, 214109 (2017); <https://doi.org/10.1063/1.4982951>

Submitted: 07 December 2016 . Accepted: 22 April 2017 . Published Online: 05 June 2017

Chen Li , Jianfeng Lu , and Weitao Yang

COLLECTIONS

Paper published as part of the special topic on [JCP Editors' Choice 2017](#)



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Communication: DFT treatment of strong correlation in 3d transition-metal diatomics](#)

The Journal of Chemical Physics **146**, 211105 (2017); <https://doi.org/10.1063/1.4985084>

[Grand canonical electronic density-functional theory: Algorithms and applications to electrochemistry](#)

The Journal of Chemical Physics **146**, 114104 (2017); <https://doi.org/10.1063/1.4978411>

[Perspective: Kohn-Sham density functional theory descending a staircase](#)

The Journal of Chemical Physics **145**, 130901 (2016); <https://doi.org/10.1063/1.4963168>

Lock-in Amplifiers
up to 600 MHz



Watch



On extending Kohn-Sham density functionals to systems with fractional number of electrons

Chen Li,¹ Jianfeng Lu,^{1,2,a)} and Weitao Yang^{3,4,b)}

¹Department of Chemistry, Duke University, Durham, North Carolina 27708, USA

²Departments of Mathematics and Physics, Duke University, Durham, North Carolina 27708, USA

³Department of Chemistry and Department of Physics, Duke University, Durham, North Carolina 27708, USA

⁴Key Laboratory of Theoretical Chemistry of Environment, School of Chemistry and Environment, South China Normal University, Guangzhou 510006, China

(Received 7 December 2016; accepted 22 April 2017; published online 5 June 2017)

We analyze four ways of formulating the Kohn-Sham (KS) density functionals with a fractional number of electrons, through extending the constrained search space from the Kohn-Sham and the generalized Kohn-Sham (GKS) non-interacting v -representable density domain for integer systems to four different sets of densities for fractional systems. In particular, these density sets are (I) ensemble interacting N -representable densities, (II) ensemble non-interacting N -representable densities, (III) non-interacting densities by the Janak construction, and (IV) non-interacting densities whose composing orbitals satisfy the Aufbau occupation principle. By proving the equivalence of the underlying first order reduced density matrices associated with these densities, we show that sets (I), (II), and (III) are equivalent, and all reduce to the Janak construction. Moreover, for functionals with the ensemble v -representable assumption at the minimizer, (III) reduces to (IV) and thus justifies the previous use of the Aufbau protocol within the (G)KS framework in the study of the ground state of fractional electron systems, as defined in the grand canonical ensemble at zero temperature. By further analyzing the Aufbau solution for different density functional approximations (DFAs) in the (G)KS scheme, we rigorously prove that there can be one and only one fractional occupation for the Hartree Fock functional, while there can be multiple fractional occupations for general DFAs in the presence of degeneracy. This has been confirmed by numerical calculations using the local density approximation as a representative of general DFAs. This work thus clarifies important issues on density functional theory calculations for fractional electron systems. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4982951>]

I. INTRODUCTION

With fifty years of development and successful applications to various fields, the Kohn-Sham density functional theory (KS-DFT)¹ has become the most widely used theoretical method for electronic structure calculations.

Despite its popularity in applications, the most famous approximate functionals within the Kohn-Sham or generalized Kohn-Sham (GKS) (called Hartree-Fock Kohn-Sham previously)² framework have been known to suffer from intrinsic errors,³ many of which have been shown to be related to the violation of exact conditions on fractional charges and fractional spins.⁴⁻¹³ The consequences of the fractional charge error include wrong band gap predictions, inaccurate reaction barriers, inaccurate molecular dissociation limit, and wrong charge transfers.^{4,5}

In the study of fractional system energies as a function of the electron number N , or E vs N curve, as was first carried out in Refs. 4 and 5, the ground state energy of a fractional-charge system is obtained by minimizing the functional of the non-interacting density in a KS calculation or first order reduced density matrix (1-RDM) in a GKS calculation that

follows the Aufbau protocol, i.e., the composing orbitals are fractionally occupied at the Fermi level (or frontier level), fully occupied below, and empty above. Note that this prescribes the structure of the non-interacting density or 1-RDM at the solution; it assumes that by performing minimization over such non-interacting densities or 1-RDMs for the exact functional, one can obtain the exact density and energy for the corresponding fractional system. While it is physically motivated and leads to all the useful insights, it has not been proved that this Aufbau implementation is consistent with the exact density functional formulation for fractional electron systems established by Perdew, Parr, Levy and Balduz (PPLB) in the 1980s,¹⁴ where the functional of the electron density has been defined through a constrained search over many-body grand canonical ensemble density matrices at zero temperature that deliver the density. In particular, it is not clear how the PPLB ground state is connected to the Aufbau solution.

In this paper, we will prove that for functionals with the ensemble v -representable assumption at the minimizer, the Aufbau construction is equivalent to searching over the densities that come from an ensemble of determinantal (i.e., KS) wavefunctions (denoted as Ens-KS) and is also equivalent to searching over the PPLB density domain. Therefore we justify that (under the same assumption) for the exact functional formulated in the (G)KS manner, the Aufbau protocol faithfully

a) jianfeng@math.duke.edu

b) weitao.yang@duke.edu

conducts the Ens-KS or PPLB formulation and delivers the right ground state solution of the fractional electron system. As a corollary, for a density functional approximation (DFA) in the (G)KS scheme, calculations using the Aufbau protocol, as previously performed, deliver the Ens-KS or PPLB solution of that DFA for a fractional number of electrons.

The rest of the paper is structured as follows. In Section II we will review the density functional theory development and introduce four search spaces (sets) of extending the (generalized) Kohn-Sham density domain to allow for fractional electron systems. These sets are the ensemble interacting N -representable set (PPLB), the ensemble non-interacting N -representable set (Ens-KS), the Janak set (non-interacting densities by the Janak construction), and the Aufbau set (non-interacting densities whose composing orbitals satisfy the Aufbau occupation principle). In Sections III and IV, we will prove that the PPLB, Ens-KS, and Janak sets are all equivalent through showing the equivalence of the underlying sets of 1-RDMs. Moreover, under the assumption that the minimizing 1-RDM within the Janak set is ensemble v -representable, the Janak construction reduces to the Aufbau construction. As a result, the Aufbau construction can deliver the PPLB solution. In Section V we will further analyze the structure of the Aufbau solution for different DFAs and finally close with some concluding remarks. In Secs. II–V, we will use the symbol Ω to describe the set of densities, while we use D for the set of 1-RDMs and \mathbf{D} for the set of many-body density matrices. Moreover, we will denote k for an integer number of electrons, to be distinguished from general N (could be integer as well as fractional) electron systems.

II. EXACT DENSITY FUNCTIONAL IN PURSUIT AND APPROXIMATIONS IN PRACTICE

A. Integer k -electron systems

Kohn-Sham density functional theory has been established by the existence theorem of Hohenberg and Kohn (HK)¹⁵ and the variational formalism by Kohn and Sham¹ since 1960s. In particular, given an external potential $v(\mathbf{r})$, Kohn-Sham (KS) assumes that the minimizing density that is interacting v -representable is simultaneously non-interacting v -representable, so that it can come from a single Slater determinant, corresponding to the ground state of a fictitious non-interacting system with some local potential $v_s(\mathbf{r})$ (different from the external potential $v(\mathbf{r})$ of the interacting system). Therefore its variational search space (domain), Ω_{KS} , is the intersection of interacting v -representable densities and non-interacting v -representable densities with local potential.¹⁶ In practice, however, given an N electron system ($N = k$), the variational procedure is carried out over the non-interacting N -representable (determinantal) densities, Ω_S^k , with the assumption that the minimizer is in Ω_{KS} ,

$$E_{\text{KS}}(k) = \min_{\rho \in \Omega_{\text{KS}}} E_{\text{KS}}[\rho] = \min_{\rho \in \Omega_S^k} E_{\text{KS}}[\rho]. \quad (1)$$

Here Ω_S^k is structured as follows:

$$\Omega_S^k = \left\{ \rho \mid \rho \text{ comes from } \gamma \in D_S^k \right\}, \quad (2)$$

where D_S^k is the set of determinantal 1-RDMs,

$$D_S^k = \left\{ \gamma = \sum_{i=1}^k |\psi_i\rangle\langle\psi_i| \mid \langle\psi_i|\psi_j\rangle = \delta_{ij} \right\}, \quad (3)$$

with ψ_i 's being spin-orbitals. The density functional $E_{\text{KS}}[\rho]$, either the exact one or a DFA, is written in terms of the Kohn-Sham decomposition,

$$E_{\text{KS}}[\rho] = T_s[\rho] + V_{\text{ex}}[\rho] + J[\rho] + E_{\text{xc}}[\rho]. \quad (4)$$

Here, T_s is the non-interacting kinetic energy, which reads

$$\begin{aligned} T_s[\rho] &= \min_{D_S^k \ni \gamma \rightarrow \rho} \text{tr} \left(-\frac{1}{2} \nabla^2 \gamma \right) \\ &= \min_{D_S^k \ni \gamma \rightarrow \rho} \int -\frac{1}{2} \nabla_{\mathbf{r}}^2 \gamma(\mathbf{r}'\sigma, \mathbf{r}\sigma) \Big|_{\mathbf{r}'=\mathbf{r}} d\mathbf{r} d\sigma, \end{aligned} \quad (5)$$

with \mathbf{r} and σ being the spatial and spin variables, respectively. V_{ex} is the external potential energy,

$$V_{\text{ex}}[\rho] = \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r}. \quad (6)$$

J is the classical Coulomb (Hartree) energy, given by

$$J[\rho] = \frac{1}{2} \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \quad (7)$$

Finally, E_{xc} is the exchange correlation energy, whose exact form is defined by

$$E_{\text{xc}}^{\text{exact}}[\rho] = E^{\text{exact}}[\rho] - T_s[\rho] - V_{\text{ex}}[\rho] - J[\rho]. \quad (8)$$

Here $E^{\text{exact}}[\rho]$ can be given by the HK definition if the density is v -representable or by the constrained search definition (as will be given below) for general N -representable densities. In practical calculations, E_{xc} has been approximated by DFAs. In particular, if E_{xc} is approximated by an explicit functional of $\rho(\mathbf{r})$, it corresponds to the conventional KS functional; if E_{xc} is approximated by an implicit functional of $\rho(\mathbf{r})$, it corresponds to the optimized effective potential (OEP) functional.

In the generalized Kohn-Sham scheme, $E_{\text{xc}}^{\text{DFA}}$ is extended to explicit functionals of non-interacting 1-RDM, while the total energy $E^{\text{DFA}} = T_s + V_{\text{ex}} + J + E_{\text{xc}}^{\text{DFA}}$ can still be thought of as an implicit functional of $\rho(\mathbf{r})$ (see the argument below), and the variational density domain is extended to Ω_{GKS}^k , the intersection of (A) interacting v -representable densities and (B) non-interacting non-local v -representable densities. Note that (B) is a subset of Ω_S^k by definition. Moreover, we can show that they are equivalent. This is because given any density ρ in Ω_S^k which comes from $\det(\psi_1, \psi_2, \dots, \psi_k)$, one can define an effective Hamiltonian $h_s = \sum_{i=1}^{\infty} \epsilon_k |\psi_k\rangle\langle\psi_k|$, with $\epsilon_1 < \epsilon_2 < \dots$. It follows that one can formally define non-local $v(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | h_s | \mathbf{r}' \rangle - \langle \mathbf{r} | t | \mathbf{r}' \rangle$, where t is the one-electron kinetic energy operator, so that the non-interacting system with this potential yields ground state density ρ , suggesting that ρ is also in (B), so that set (B) is the same as Ω_S^k . Similar arguments apply to 1-RDMs, where one can show that the set of non-interacting non-local v -representable 1-RDMs is the same as D_S^k .

To see that E^{DFA} as an explicit functional of non-interacting γ can be linked to an implicit functional of ρ , we can introduce the following constrained search:

$$E^{\text{DFA}}[\rho] = \min_{\gamma \rightarrow \rho \in \Omega_{\text{GKS}}^k} E^{\text{DFA}}[\gamma]. \quad (9)$$

Let the domain of γ be $D_{\text{GKS}}^k = \{\gamma | \rho_\gamma \in \Omega_{\text{GKS}}^k\}$, where $\rho_\gamma(\mathbf{r}) = \sum_{\sigma} \gamma(\mathbf{r}\sigma, \mathbf{r}\sigma)$. It follows that

$$\begin{aligned} E_{\text{GKS}}^{\text{DFA}}(k) &= \min_{\rho \in \Omega_{\text{GKS}}^k} E^{\text{DFA}}[\rho] = \min_{\rho \in \Omega_{\text{GKS}}^k} \min_{\gamma \rightarrow \rho} E^{\text{DFA}}[\gamma] \\ &= \min_{\gamma \in D_{\text{GKS}}^k} E^{\text{DFA}}[\gamma]. \end{aligned} \quad (10)$$

Therefore, minimizing the functional as an implicit functional of ρ delivers the same result as minimizing as an explicit functional of non-interacting γ , while the latter way is implemented in practice. Moreover, as in KS, we enlarge the variational search space from Ω_{GKS}^k (the intersection of v -representable densities and Ω_S^k) or D_{GKS}^k (the intersection of v -representable 1-RDMs and D_S^k) to Ω_S^k or D_S^k , with the assumption that the minimizer within Ω_S^k or D_S^k is also interacting v -representable, i.e.,

$$E_{\text{GKS}}^{\text{DFA}}(k) = \min_{\rho \in \Omega_{\text{GKS}}^k} E^{\text{DFA}}[\rho] = \min_{\rho \in \Omega_S^k} E^{\text{DFA}}[\rho], \quad (11)$$

which is equivalent as

$$E_{\text{GKS}}^{\text{DFA}}(k) = \min_{\gamma \in D_{\text{GKS}}^k} E^{\text{DFA}}[\gamma] = \min_{\gamma \in D_S^k} E^{\text{DFA}}[\gamma]. \quad (12)$$

Here the second equalities in Eqs. (11) and (12) are due to the above assumption. As a remark, when E_{xc}^{DFA} is an explicit functional of ρ , it is still valid to perform the minimization in the GKS scheme, which can be reduced to KS. In the case of E_{xc}^{DFA} being an implicit functional of ρ , optimization within KS is equivalent to OEP, which requires the minimizing density to be local v -representable. OEP calculations lead to slightly higher energies compared with the GKS calculations because of the smaller variational space due to the local potential restriction. In the following, our discussions on DFAs will focus on the minimization in the GKS scheme.

In the GKS scheme for an integer number of electrons, the density minimization has been transformed into the minimization over the orbitals composing a Slater determinant, and the minimizer corresponds to a GKS non-interacting reference system with possible non-local potential. Although nicely formulated as a computational procedure, the GKS domain has been shown to be non-convex (initially shown for KS but also applicable to GKS), an undesirable feature that is problematic especially in the case of degenerate ground state densities.¹⁶ To have a better formulation of the functional on a nicely characterized convex set, in the late 1970s, Levy and subsequently Lieb extended the Kohn-Sham domain to the pure-state N -representable densities, Ω_p^k , and defined the exact functional through a constrained search formulation.^{16,17} In particular, the Levy-Lieb functional reads

$$E_{\text{exact}}^{\text{pure}}[\rho] = \min_{\Psi \rightarrow \rho} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r}. \quad (13)$$

Here \hat{T} and \hat{V}_{ee} are the many-body kinetic and electron-electron repulsion energy operators, respectively. The Levy functional expression is not convex with respect to ρ , while Lieb modified the Levy-Lieb constrained search and defined a convex functional through Legendre transform.¹⁶ Subsequently, Valone extended it to canonical ensemble N -representable densities with an integer number of electrons.¹⁸

B. Fractional N -electron systems

The Levy-Lieb domain is valid only for integer electron systems. Yet it can be further extended to allow densities that integrate to a noninteger number of electrons and therefore extending the DFT formulation to fractional electron systems, for example, open systems that can share electrons with other systems. This is in contrast to wavefunction theory within the Hilbert space that can only be defined for integer systems. The extension of DFT to a fractional number of electrons was made later in the 1980s by the work of Perdew, Parr, Levy, and Balduz (PPLB),¹⁴ where the constrained search was extended to grand canonical ensemble density matrices rather than pure state wavefunctions, and the search space of the electron density is the ensemble N -representable densities, Ω_p^N . In particular, for any N , either integer or fractional, the total electronic energy functional can be defined as

$$E_{\text{exact}}^{\text{ens}}[\rho] = \min_{\Gamma \rightarrow \rho} \text{tr}((\hat{T} + \hat{V}_{ee})\Gamma) + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r}, \quad (14)$$

where Γ is the many-body ensemble N -representable density matrix, i.e., a statistical mixture of pure states of different electron numbers, which delivers the average electron number N (this will be defined more explicitly later in Eq. (28)). The total energy of the system is then given by

$$E^{\text{exact}}(N) = \min_{\rho \in \Omega_p^N} E_{\text{exact}}^{\text{ens}}[\rho]. \quad (15)$$

With density functional approximations, we have

$$E^{\text{DFA}}(N) = \min_{\rho \in \Omega_p^N} E^{\text{DFA}}[\rho]. \quad (16)$$

Now if Eq. (11) is the GKS realization of the constrained search formula of Eq. (13) for practical calculations for integer systems, it is straightforward to formulate the GKS realization of Eq. (14) for fractional systems through extending the search space Ω_S^k to Ω_S^N ,

$$E_{\text{GKS}}^{\text{DFA}}(N) = \min_{\rho \in \Omega_{\text{GKS}}^N} E^{\text{DFA}}[\rho] = \min_{\rho \in \Omega_S^N} E^{\text{DFA}}[\rho]. \quad (17)$$

Here Ω_{GKS}^N is the extension of Ω_{GKS}^k , which is not particularly interesting to us and its details are omitted. A more interesting set is Ω_S^N , the set of ensemble non-interacting N -representable densities, i.e., densities coming from an ensemble of determinants of different electron numbers. Again, the ensemble v -representability at the minimizer of Ω_S^N is assumed in the second equality of Eq. (17).

The search space of Ω_S^N has clear physics, however, it is not convenient to be directly used for practical calculations because the set is too implicit in terms of how the densities are structured. Instead, what is often implemented^{4,5} is the Aufbau construction, given by

$$E_{\text{Aufbau}}^{\text{DFA}}(N) = \min_{\rho \in \Omega_A^N} E^{\text{DFA}}[\rho], \quad (18)$$

where

$$\Omega_A^N = \{\rho | \rho \text{ comes from } \gamma \in D_A^N\}, \quad (19)$$

with

$$D_A^N = \left\{ \gamma = \sum_{i=1}^{\infty} n_i |\psi_i\rangle\langle\psi_i| \mid \langle\psi_i|\psi_j\rangle = \delta_{ij} \right\}, \quad (20)$$

and the occupation numbers n_i follow the Aufbau protocol, i.e., the orbitals are fractionally occupied at the Fermi level ϵ_f , fully occupied below, and empty above,

$$n_i = \begin{cases} 1, & \epsilon_i < \epsilon_f, \\ n_i^f, & \epsilon_i = \epsilon_f, \\ 0, & \epsilon_i > \epsilon_f. \end{cases} \quad (21)$$

Here ϵ_i 's are GKS orbital energies, i.e., the eigenvalues of the GKS effective Hamiltonian, and $0 \leq n_i^f \leq 1$. As a remark, ρ in the Aufbau domain can be thought of as coming from an ensemble of determinants. In particular, in the case of only one fractional occupation, Eq. (21) implies two determinants, $\det(\psi_1, \psi_2, \dots, \psi_{[N]})$ and $\det(\psi_1, \psi_2, \dots, \psi_{[N]+1})$, with the linear combination coefficients being $[N] + 1 - N$ and $N - [N]$, respectively, with $[N]$ being the largest integer no greater than N . In the case of multiple fractional occupations, one can construct the corresponding determinants similarly.

In fact, as a more general formulation, one can introduce another way of fractional formulation of functionals (let us call it the Janak construction),

$$E_{\text{Janak}}^{\text{DFA}}(N) = \min_{\rho \in \Omega_J^N} E^{\text{DFA}}[\rho], \quad (22)$$

where Ω_J^N is the set of densities associated with canonical 1-RDMs by the Janak construction, given by

$$\Omega_J^N = \{\rho | \rho \text{ comes from } \gamma \in D_J^N\}, \quad (23)$$

with

$$D_J^N = \left\{ \gamma = \sum_{i=1}^{\infty} n_i |\psi_i\rangle\langle\psi_i| \mid \langle\psi_i|\psi_j\rangle = \delta_{ij} \right\}, \quad (24)$$

and the occupation numbers satisfy

$$0 \leq n_i \leq 1, \quad \sum_{i=1}^{\infty} n_i = N. \quad (25)$$

The above four ways (Eqs. (16)–(18) and (22)) of fractional formulations of GKS differ only in the constrained search space, namely, Ω_P^N , Ω_S^N , Ω_A^N , and Ω_J^N . It is not immediately clear how the four spaces are related and thus how the four formulations for fractional systems are connected. To the best of our knowledge, this has never been explored in the previous literature and is the main contribution of the present paper. Here we note that although some of the conclusions in the present paper can be applied to density functional theory in the constrained search formulation without invoking GKS, we will focus on extending the generalized Kohn-Sham DFT using DFAs because the latter provides us with a framework for practical calculations and is most relevant to applications.

In Secs. III and IV, we will show that the sets Ω_P^N , Ω_S^N , and Ω_J^N are completely equivalent; therefore, the corresponding fractional formulation is the same as the Janak construction. Moreover, the Janak construction reduces to the Aufbau construction for general differentiable density functionals in the GKS scheme with the ensemble ν -representable assumption at the minimizer.

III. EQUIVALENCE PROOFS

To prove the equivalence of the sets of densities, it suffices to prove the equivalence of the underlying 1-RDM sets that

deliver the density sets. In Subsection III A, we will introduce the underlying 1-RDM sets for Ω_P^N and Ω_S^N , denoted by D_P^N and D_S^N (see Eqs. (30) and (31)), respectively, and then prove that $D_P^N = D_S^N = D_J^N$, and D_J^N collapses into D_A^N once we perform the energy minimization with respect to the non-interacting 1-RDM, so that the corresponding density set Ω_J^N reduces to Ω_A^N .

Here we note that although in the practical calculations the spin-DFT has been implemented, i.e., treating $E^{\text{DFA}} = E^{\text{DFA}}[\gamma^\alpha(\mathbf{r}, \mathbf{r}'), \gamma^\beta(\mathbf{r}, \mathbf{r}')]$, here for the simplicity of discussion and notations, we first focus on the simple case of $E^{\text{DFA}} = E^{\text{DFA}}[\gamma(\mathbf{x}, \mathbf{x}')]$ ($\mathbf{x} \equiv \mathbf{r}\sigma$ is the combined spatial and spin variable) to draw a conclusion and then prove in Section IV that this conclusion is also applicable to spin-DFT. Moreover, by replacing \mathbf{x} by \mathbf{r} , the conclusion in this section automatically holds for the spatial 1-RDM of a specific spin in the absence of the other spin.

A. Set equivalence of D_P^N , D_S^N , and D_J^N

We start with the definition of D_P^N , the set of ensemble N -representable 1-RDMs. In fact, for any integer k , the set of many-body ensemble k -representable density matrices has been known and defined by the convex combinations of pure state k -representable density matrices,¹⁹

$$\mathbf{D}^k = \left\{ \sum_{i=1}^{\infty} m_i \Gamma(\Psi_i^{(k)}) \mid \Psi_i^{(k)} \in \mathcal{H}^{(k)}, \right. \\ \left. 0 \leq m_i \leq 1, \sum_{i=1}^{\infty} m_i = 1 \right\}. \quad (26)$$

Here $\mathcal{H}^{(k)}$ is the set of antisymmetric wavefunctions of k electrons, and $\Gamma(\Psi_i^{(k)})(\mathbf{x}^k, (\mathbf{x}^k)') = \Psi_i^{(k)}(\mathbf{x}^k) [\Psi_i^{(k)}((\mathbf{x}^k)')]^*$ is the many-body density matrix corresponding to the pure state $\Psi_i^{(k)}(\mathbf{x}^k) \equiv \Psi_i^{(k)}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k)$. For $k=0$, $\mathcal{H}^{(0)}$ denotes the Hilbert space for the vacuum state. For each $\Gamma_k \in \mathbf{D}^k$, the corresponding 1-RDM is given by

$$\gamma_k(\mathbf{x}, \mathbf{x}') = k \int \Gamma_k(\mathbf{x}, \mathbf{x}_2, \dots, \mathbf{x}_k; \mathbf{x}', \mathbf{x}_2, \dots, \mathbf{x}_k) \\ \times d\mathbf{x}_2 d\mathbf{x}_3 \cdots d\mathbf{x}_k. \quad (27)$$

We then denote the set of all possible γ_k 's, i.e., the ensemble k -representable 1-RDMs as D_P^k .

For a fractional number N , now we define the set of ensemble N -representable density matrices within the Fock space as the direct sum of ensemble k -representable density matrices,

$$\mathbf{D}^N = \left\{ \bigoplus_{k=0}^{\infty} n_k \Gamma_k \mid \Gamma_k \in \mathbf{D}^k \right\}, \quad (28)$$

where the combination coefficients satisfy

$$0 \leq n_k \leq 1, \quad \sum_{k=0}^{\infty} n_k = 1, \quad \sum_{k=0}^{\infty} k n_k = N. \quad (29)$$

Moreover, we define the set of ensemble N -representable 1-RDMs as the convex combinations of the corresponding

integer 1-RDMs,

$$D_P^N = \left\{ \gamma = \sum_{k=0}^{\infty} n_k \gamma_k \mid \gamma_k \in D_P^k (k = 0, 1, \dots), \right. \\ \left. 0 \leq n_k \leq 1, \sum_{k=0}^{\infty} n_k = 1, \sum_{k=0}^{\infty} k n_k = N \right\}. \quad (30)$$

Similarly, we can follow the above procedure and define ensemble non-interacting N -representable 1-RDMs, i.e., an ensemble of 1-RDMs that come from Slater determinants,

$$D_S^N = \left\{ \gamma = \sum_{k=0}^{\infty} n_k \gamma_k \mid \gamma_k \in D_S^k (k = 0, 1, \dots), \right. \\ \left. 0 \leq n_k \leq 1, \sum_{k=0}^{\infty} n_k = 1, \sum_{k=0}^{\infty} k n_k = N \right\}. \quad (31)$$

Here D_S^k is the set of determinantal 1-RDMs of k electrons, as defined in Eq. (3). By definition, $D_S^N \subset D_P^N$. Now we want to show that $D_S^N = D_P^N = D_J^N$. Rewriting D_J^N as

$$D_J^N = \{ \gamma \mid \gamma = \gamma^\dagger, 0 \leq \gamma \leq 1, \text{tr}(\gamma) = N \}, \quad (32)$$

we will prove $D_S^N = D_P^N = D_J^N$ by showing that $D_J^N \subset D_S^N \subset D_P^N \subset D_J^N$. It suffices to show that $D_P^N \subset D_J^N$ and $D_J^N \subset D_S^N$.

(I) $D_P^N \subset D_J^N$. For any $\gamma \in D_P^N$, there exists $\Gamma_k \in \mathcal{D}^k (k = 0, 1, \dots)$ and $\Gamma_k \rightarrow \gamma_k(\mathbf{r}, \mathbf{r}')$, such that

$$\gamma = \sum_{k=0}^{\infty} n_k \gamma_k, \quad (33)$$

where $\sum n_k = 1$ and $\sum k n_k = N$. It follows that $\text{tr}(\gamma) = \sum n_k \text{tr}(\gamma_k) = N$.

Since it has been proved by Coleman that $D_P^k = D_J^k$,¹⁹ we have $\gamma_k = \gamma_k^\dagger$ and $0 \leq \gamma_k \leq 1$. It is obvious that $\sum n_k \gamma_k = \sum n_k \gamma_k^\dagger$, and $0 \leq \sum n_k \gamma_k \leq 1$, where the last inequality is due to the fact that convex combinations of positive semi-definite operators are still positive semi-definite. Thus $\gamma = \gamma^\dagger$ and $0 \leq \gamma \leq 1$. Therefore, $\gamma \in D_J^N$, and hence $D_P^N \subset D_J^N$.

(II) $D_J^N \subset D_S^N$. Given arbitrary $\gamma \in D_J^N$, let

$$\gamma(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^{\infty} \theta_i \phi_i(\mathbf{x}) \phi_i(\mathbf{x}'), \quad (34)$$

where $1 \geq \theta_1 \geq \theta_2 \geq \dots \geq 0$, and $\sum_{i=1}^{\infty} \theta_i = N$. We construct the following γ_k 's,

$$\gamma_k(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^k \phi_i(\mathbf{x}) \phi_i(\mathbf{x}'), \quad (35)$$

and construct a sequence $\{n_k\}$ as

$$n_0 = 1 - \theta_1, \quad (36)$$

$$n_k = \theta_k - \theta_{k+1}, \quad (k \geq 1). \quad (37)$$

It is easy to see that $0 \leq n_k \leq 1$, $\sum_{k=0}^{\infty} n_k = 1$. Moreover,

$$\sum_{k=0}^{\infty} k n_k = \sum_{k=1}^{\infty} k(\theta_k - \theta_{k+1}) = \sum_{k=1}^{\infty} k \theta_k - \sum_{k=2}^{\infty} (k-1) \theta_k \\ = \sum_{k=1}^{\infty} \theta_k = N. \quad (38)$$

Therefore, γ can be written as $\gamma = \sum_{k=0}^{\infty} n_k \gamma_k \in D_S^N$. Thus $D_J^N \subset D_S^N$. Consequently, (I) and (II) imply that $D_S^N = D_P^N = D_J^N$; thus, they all reduce to the Janak construction. Note our proof of $D_P^N = D_J^N$ is an extension of Coleman's theorem from integer to fractional N . Moreover, the fact that $D_P^N = D_S^N$ suggests that the non-interacting N -representable 1-RDMs defined in this paper exhaust the set of interacting N -representable 1-RDMs. Similar argument applies for the corresponding densities. Furthermore, we remark that in the case of integer, $D_S^k = D_P^k = D_J^k$ as a corollary of our proof. This suggests that any 1-RDM coming from an ensemble of k -representable wavefunctions can be represented as the density corresponding to an ensemble of determinantal wavefunctions of different electron numbers, with the average electron number being k . In fact, we can prove a stronger claim that any 1-RDM that comes from an ensemble of k -representable wavefunctions can be written as a convex combination of determinantal 1-RDMs of k electrons. This is because $D_P^k = D_J^k$ is a convex set, and all the extreme points in this set are determinantal 1-RDMs. Therefore, any 1-RDM in the set is a linear combination of determinantal 1-RDMs, hence our claim is true.

Regarding the physical interpretation of the minimizing 1-RDM for a fractional system, now we have the freedom to interpret it in three perspectives. However, for the parallel comparison with the fictitious non-interacting 1-RDM for integer systems in the Kohn-Sham theory, here for fractional systems we adopt the Ens-KS interpretation and view the 1-RDM as coming from a fictitious ensemble of non-interacting systems, and only the energy and electron density have physical interpretations.

B. Equivalence of Janak and Aufbau constructions

Consider a general DFA in the GKS scheme. Here we assume that the functional is differentiable and the effective Hamiltonian is hermitian. By definition, $D_A^N \subset D_J^N$, thus the minimizer of $E_{\text{Aufbau}}^{\text{DFA}}(N)$, call it γ_A , has to be in D_J^N . To prove that the minimizers within Janak and Aufbau constructions are identical, it suffices to show that the minimizer of $E_{\text{Janak}}^{\text{DFA}}(N)$, call it γ_J , is in D_A^N . We will prove this by contradiction. Let $h_s = \frac{\delta E^{\text{DFA}}}{\delta \gamma} \Big|_{\gamma=\gamma_J} = \sum_i \epsilon_i |\psi_i\rangle \langle \psi_i|$. Here we assume $\epsilon_1 \leq \epsilon_2 \leq \dots$. Suppose there exists $\epsilon_i < \epsilon_j$ such that $0 \leq \langle \psi_i | \gamma_J | \psi_i \rangle < 1$ and $0 < \langle \psi_j | \gamma_J | \psi_j \rangle \leq 1$. Then for $\eta > 0$ and sufficiently small, let

$$\gamma'_J = \gamma_J + \eta |\psi_i\rangle \langle \psi_i| - \eta |\psi_j\rangle \langle \psi_j|. \quad (39)$$

It is obvious that γ'_J is a legitimate 1-RDM in D_J^N . Moreover, the energy change up to the first order of η is given by

$$E^{\text{DFA}}[\gamma'_J] - E^{\text{DFA}}[\gamma_J] = \text{tr} \left(\frac{\delta E^{\text{DFA}}}{\delta \gamma} \Big|_{\gamma=\gamma_J} (\gamma'_J - \gamma_J) \right) \\ = \eta \text{tr} (h_s (|\psi_i\rangle \langle \psi_i| - |\psi_j\rangle \langle \psi_j|)) \\ = \eta (\epsilon_i - \epsilon_j) < 0, \quad (40)$$

which leads to a contradiction. Therefore, there exists a Fermi energy level ϵ_f , below which all the orbitals are completely occupied ($\langle \psi_i | \gamma_J | \psi_i \rangle = 1$) and above which all the orbitals are empty ($\langle \psi_j | \gamma_J | \psi_j \rangle = 0$). Orbitals at the Fermi level can be

empty, fractionally occupied, or fully occupied. Thus,

$$\gamma_J(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^L \phi_i(\mathbf{x})\phi_i(\mathbf{x}') + \sum_{i=L+1}^{L+M} n_i \phi_i(\mathbf{x})\phi_i(\mathbf{x}'), \quad (41)$$

where $L \leq [N]$ and $0 \leq n_i \leq 1$ ($L+1 \leq i \leq L+M$) for some M (allowed to be infinity); ϕ_i 's are orthonormal. Moreover, $\{\phi_i\}_{i=1}^L$ span the space below the Fermi level; $\{\phi_i\}_{i=L+1}^{L+M}$ are within the space spanned by orbitals at the Fermi level, so that each of them is an eigen-orbital of h_s with eigenvalue ϵ_f . Therefore, γ and h_s commute, so that the eigen-orbitals of h_s , $\{\psi_i\}$, are also eigen-orbitals of γ . Thus, we can replace ϕ_i by ψ_i in Eq. (41) and have

$$\gamma_J(\mathbf{x}, \mathbf{x}') = \sum_{i=1}^L \psi_i(\mathbf{x})\psi_i(\mathbf{x}') + \sum_{i=L+1}^{L+M} n_i \psi_i(\mathbf{x})\psi_i(\mathbf{x}'), \quad (42)$$

where ψ_i 's satisfy the (G)KS equation,

$$h_s \psi_i(\mathbf{x}) = \epsilon_i \psi_i(\mathbf{x}). \quad (43)$$

This suggests that $\gamma_J \in D_A^N$ and the Janak construction is equivalent to the Aufbau construction. Here we note that our proof is in a similar spirit as Gilbert's proof²⁰ on the structure of minimizing 1-RDMs in the case of density matrix functionals, or Janak's²¹ and generalized Janak's theorem⁶ in terms of energy derivative with respect to occupation numbers for density functionals in the case of integer electrons, yet our proof here extends to a fractional number of electrons.

IV. EQUIVALENCE PROOF FOR COLINEAR SPIN-DFT

In colinear spin-DFT, the functional becomes $E^{\text{DFA}}[\gamma_\alpha, \gamma_\beta]$, and the ground state is given by

$$E_{\text{GKS}}^{\text{DFA}}(N) = \min_{\gamma \in \bar{D}_J^N} E^{\text{DFA}}[\gamma_\alpha(\mathbf{r}, \mathbf{r}'), \gamma_\beta(\mathbf{r}, \mathbf{r}')], \quad (44)$$

where the spin 1-RDM $\gamma_\sigma(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r}\sigma | \gamma | \mathbf{r}'\sigma \rangle = \gamma(\mathbf{r}\sigma, \mathbf{r}'\sigma)$ for $\sigma = \alpha, \beta$, and

$$\bar{D}_J^N = \left\{ \gamma(\mathbf{x}, \mathbf{x}') | \gamma = \gamma^\dagger, 0 \leq \gamma \leq 1, \right. \\ \left. \gamma \rightarrow (\gamma_\alpha, \gamma_\beta), \gamma_\alpha \in D_J^{N_\alpha}, \gamma_\beta \in D_J^{N_\beta} \right\}. \quad (45)$$

Here $N_\sigma = \text{tr}(\gamma_\sigma)$ and the representability of $\gamma_\sigma(\mathbf{r}, \mathbf{r}')$ should be understood as referring to the spatial 1-RDM, i.e., replacing \mathbf{x} by \mathbf{r} .

The variational set of \bar{D}_J^N is explicit in terms of γ but implicit in terms of each γ_σ . To define an explicit minimization over γ_σ , one can introduce the following two-step minimization:

$$E_{\text{GKS}}^{\text{DFA}}(N_\alpha, N_\beta) = \min_{\gamma_\sigma \in D_J^{N_\sigma}} E^{\text{DFA}}[\gamma_\alpha(\mathbf{r}, \mathbf{r}'), \gamma_\beta(\mathbf{r}, \mathbf{r}')], \quad (46)$$

$$E_{\text{GKS}}^{\text{DFA}}(N) = \min_{N_\alpha + N_\beta = N} E_{\text{GKS}}^{\text{DFA}}(N_\alpha, N_\beta). \quad (47)$$

To justify that this two-step minimization is equivalent to the one step minimization in Eq. (44), it suffices to show that $D_J^N = \bar{D}_J^N$, i.e., for $\gamma(\mathbf{x}, \mathbf{x}') \in D_J^N$, the sufficient and necessary condition is that its spin components $\gamma_\sigma(\mathbf{r}, \mathbf{r}') \in D_J^{N_\sigma}$ ($\sigma = \alpha, \beta$). By the set equivalence of $D_P^N = D_S^N = D_J^N$, this translates into the following claim: if $\gamma(\mathbf{x}, \mathbf{x}')$ is ensemble N -representable (or ensemble non-interacting N -representable), then its spin

components $\gamma_\sigma(\mathbf{r}, \mathbf{r}')$ must be ensemble N_σ -representable (or ensemble non-interacting N_σ -representable). Once again, we prove the set equality by showing $\bar{D}_J^N \subset D_J^N$ and $D_J^N \subset \bar{D}_J^N$.

(I) $\bar{D}_J^N \subset D_J^N$. $\forall \gamma \in \bar{D}_J^N$, by definition, $\gamma = \gamma^\dagger, 0 \leq \gamma \leq 1$. Moreover, $\text{tr}(\gamma) = \text{tr}(\gamma_\alpha) + \text{tr}(\gamma_\beta) = N_\alpha + N_\beta = N$. Thus, $\gamma \in D_J^N \Rightarrow \bar{D}_J^N \subset D_J^N$.

(II) $D_J^N \subset \bar{D}_J^N$. $\forall \gamma \in D_J^N$, we only need to show that $\gamma_\sigma \in D_J^{N_\sigma}$ or $0 \leq \gamma_\sigma \leq 1$. In fact, $\forall \phi(\mathbf{r})$,

$$\langle \phi | \gamma_\sigma | \phi \rangle = \langle \phi | (\langle \sigma | \gamma | \sigma \rangle) | \phi \rangle = \langle \phi \sigma | \gamma | \phi \sigma \rangle = \langle \psi | \gamma | \psi \rangle, \quad (48)$$

where $\psi(\mathbf{x}') = \psi(\mathbf{r}'\sigma) = \phi(\mathbf{r}')\sigma(\sigma')$. Since $0 \leq \gamma \leq 1$, we have $0 \leq \langle \psi | \gamma | \psi \rangle \leq 1$ and thus $0 \leq \langle \phi | \gamma_\sigma | \phi \rangle \leq 1 \Rightarrow 0 \leq \gamma_\sigma \leq 1$. This suggests that $\gamma \in D_J^N \Rightarrow D_J^N \subset \bar{D}_J^N$. Thus $D_J^N = \bar{D}_J^N$.

For DFAs in spin-DFT, we can now borrow the results in Sec. III and conclude that the PPLB and Ens-KS formulations reduce to the Janak construction of Eqs. (46) and (47) and that the Janak construction reduces to the Aufbau construction.

As a remark, the Aufbau construction in spin-DFT is given by

$$n_i^\sigma = \begin{cases} 1, & \epsilon_i^\sigma < \epsilon_f, \\ n_{i\sigma}^f, & \epsilon_i^\sigma = \epsilon_f, \\ 0, & \epsilon_i^\sigma > \epsilon_f, \end{cases} \quad (49)$$

where $0 < n_{i\sigma}^f < 1$, and the Fermi level is given by the highest occupied orbital energy of both spins.

Another remark is on the exact conditions which is fulfilled by the exact functional in the case of fractional as well as integer systems.

(i) PPLB condition for fractional charge. Perdew, Parr, Levy, and Balduz have shown that the fractional system energy by the exact functional should be a linear interpolation between the neighboring integer energies,¹⁴ i.e.,

$$E^{\text{exact}}(N) = (1 - \delta)E^{\text{exact}}(K) + \delta E^{\text{exact}}(K + 1), \quad (50)$$

where $K = [N]$, $\delta = N - K$. Here it is assumed that the integer energies are convex in the sense that $2E^{\text{exact}}(K) < E^{\text{exact}}(K - 1) + E^{\text{exact}}(K + 1)$.

(ii) Constancy condition for fractional spin. For integer N , in the case of degenerate ground states, Yang, Zhang, and Ayers have shown that the exact functional evaluated at any linear combination of the degenerate ground state densities gives the same ground state energy in 2000.²² In particular, when the degeneracy occurs through spin channels, this has been known as the fractional spin condition.²³ In the language of the present paper, this corresponds to the case where Eq. (47) has multiple minimizers, i.e., for all pairs (N_α, N_β) satisfying $N_\alpha + N_\beta = N$ and $N_\alpha \in [K_\alpha, K_{\alpha+1}]$ for some K_α ,

$$E^{\text{exact}}(N_\alpha, N_\beta) = E^{\text{exact}}(N). \quad (51)$$

(iii) Flat plane condition. Combining conditions (i) and (ii), Mori-Sánchez, Cohen, and Yang proved the general

conditions combining fractional charges and degenerate states and developed the flat plane condition specifically for the fractional charge and fractional spins in 2009.²⁴ For spin-DFT, in particular, this corresponds to

the general case where we can express $E^{\text{exact}}(N_\alpha, N_\beta)$ in terms of a flat plane energy associated with the integer energies at the edges of the plane (denote $K_\sigma = [N_\sigma]$ and $\delta_\sigma = N_\sigma - K_\sigma$),

$$E^{\text{exact}}(N_\alpha, N_\beta) = \begin{cases} (1 - \delta_\alpha - \delta_\beta)E(K_\alpha, K_\beta) + \delta_\beta E(K_\alpha, K_\beta + 1) + \delta_\alpha E(K_\alpha + 1, K_\beta), & 0 \leq \delta_\alpha + \delta_\beta < 1, \\ (1 - \delta_\alpha)E(K_\alpha, K_\beta + 1) + (1 - \delta_\beta)E(K_\alpha + 1, K_\beta) + (\delta_\alpha + \delta_\beta - 1)E(K_\alpha + 1, K_\beta + 1), & 1 \leq \delta_\alpha + \delta_\beta < 2. \end{cases} \quad (52)$$

Here we have abbreviated $E^{\text{exact}}(K_\alpha, K_\beta)$ as $E(K_\alpha, K_\beta)$.

Eq. (52) is valid with the assumption that

$$E(K_\alpha + 1, K_\beta) + E(K_\alpha, K_\beta + 1) < E(K_\alpha, K_\beta) + E(K_\alpha + 1, K_\beta + 1). \quad (53)$$

When this assumption is violated, there has been another type of flat plane condition constructed recently.²⁵ Note that both types of flat plane condition, Eq. (52) and the one in Ref. 25, are specific consequences of the general results in Ref. 24.

V. FURTHER ANALYSIS ON AUFBAU OCCUPATION

The Aufbau construction allows multiple fractional occupancies at the Fermi level; however, it cannot deliver further information about how many fractional occupations should occur at the Fermi level. In fact, this is dependent on the choice of DFAs. In this section, we will rigorously prove that for the Hartree Fock (HF) functional, there can be one and only one fractional occupation through one of the spin channels at the Fermi level. For the local density approximation (LDA) functional as a representative of the semi-local functionals, there can be multiple fractional occupations at the Fermi level in the presence of degeneracy, and whether this occurs or not is system dependent.

A. Aufbau principle for HF functional

In the following, we prove that only one of the occupations in γ_J is fractional for HF in the absence of spin. In fact, it has been proved that for integer systems, all the occupation numbers of the HF minimizer have to be integer.^{26,27} Here our claim can be viewed as an extension to fractional N .

Suppose γ_J consists of at least two fractionally occupied states, i.e., there exist $i \neq j$ such that $0 < n_i, n_j < 1$. We will use a similar technique as in Ref. 26 to arrive at a contradiction. Let $\eta = \min\{n_i, n_j, 1 - n_i, 1 - n_j\}$, and define

$$\gamma_J^1(\mathbf{r}, \mathbf{r}') = \gamma_J(\mathbf{r}, \mathbf{r}') + \eta \psi_i(\mathbf{r})\psi_i(\mathbf{r}') - \eta \psi_j(\mathbf{r})\psi_j(\mathbf{r}'), \quad (54)$$

$$\gamma_J^2(\mathbf{r}, \mathbf{r}') = \gamma_J(\mathbf{r}, \mathbf{r}') + \eta \psi_j(\mathbf{r})\psi_j(\mathbf{r}') - \eta \psi_i(\mathbf{r})\psi_i(\mathbf{r}'). \quad (55)$$

It is obvious that $\gamma_J^1, \gamma_J^2 \in D_J^N$. The HF functional, as a functional of 1-RDM is defined by

$$E[\gamma(\mathbf{r}, \mathbf{r}')] = \int -\frac{1}{2} \nabla_r^2 \gamma(\mathbf{r}, \mathbf{r}') \Big|_{\mathbf{r}=\mathbf{r}'} d\mathbf{r} + \int \gamma(\mathbf{r}, \mathbf{r}) v(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint \frac{\gamma(\mathbf{r}, \mathbf{r})\gamma(\mathbf{r}', \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' - \frac{1}{2} \iint \frac{\gamma(\mathbf{r}, \mathbf{r}')\gamma(\mathbf{r}', \mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \quad (56)$$

We now calculate $E[\gamma_J^1] + E[\gamma_J^2] - 2E[\gamma_J]$. Since $\gamma_J^1 + \gamma_J^2 - 2\gamma_J = 0$, the linear terms on the 1-RDMs vanish, and we only have to collect the quadratic terms as follows:

$$\begin{aligned} & E[\gamma_J^1] + E[\gamma_J^2] - 2E[\gamma_J] \\ &= \eta^2 \iint \frac{[|\psi_i(\mathbf{r})|^2 - |\psi_j(\mathbf{r})|^2][|\psi_i(\mathbf{r}')|^2 - |\psi_j(\mathbf{r}')|^2]}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \\ &\quad - \eta^2 \iint \frac{|\psi_i(\mathbf{r})\psi_i(\mathbf{r}') - \psi_j(\mathbf{r})\psi_j(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \\ &= -\eta^2 \iint \frac{|\psi_i(\mathbf{r})\psi_j(\mathbf{r}') - \psi_i(\mathbf{r}')\psi_j(\mathbf{r})|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' < 0. \end{aligned} \quad (57)$$

Thus either $E[\gamma_J^1] < E[\gamma_J]$ or $E[\gamma_J^2] < E[\gamma_J]$, suggesting that γ_J is not the minimizer, which is a contradiction. Therefore, there is one and only one occupation number being fractional, and the minimizer can be written as

$$\gamma(\mathbf{r}, \mathbf{r}') = \sum_{k=1}^K \psi_k(\mathbf{r})\psi_k(\mathbf{r}') + \delta \psi_{K+1}(\mathbf{r})\psi_{K+1}(\mathbf{r}'), \quad (58)$$

where $K = [N]$ and $\delta = N - K$.

In the case of unrestricted HF where both spin channels are present, the above conclusion can be easily extended. In particular, since the quadratic terms of each γ_J^σ are similar to the quadratic ones of γ in Eq. (56), there can be one fractional occupation in each spin at most (proof omitted). Furthermore, we can prove that the scenario of fractional occupation in both spins cannot occur. In other words, there can be one and only one fractional occupation through one of the spin channels in the minimizer of unrestricted HF. The proof is similar as above, see the Appendix for details.

Here we remark that the claim of single fractional occupation for HF is out of reach of the proof in Section III B or the generalized Janak's theorem. In other words, the first order energy derivative analysis dictates the occupation rule for the orbitals below the HOMO (they should be fully occupied), whereas leaves the ambiguity of the degeneracy at the HOMO level. It is the second order derivative analysis as done in the first part of the proof that addresses the HOMO occupation problem.

To have a better understanding of the general picture of the $E_v(N)$ curve, in Figure 1, we use the example of a carbon atom to illustrate the HF energy as a function of the total number of electrons N . Here the calculations are performed using our in-house built QM4D program,²⁸ and the basis set used is 6-311++G(3df,3pd). It has been shown that the integer energies by HF is monotone non-increasing, i.e.,

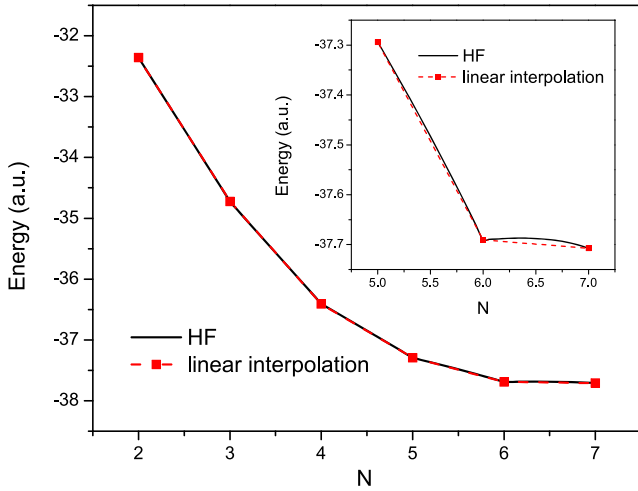


FIG. 1. HF energy as a function of the fractional number of electrons of carbon atom. The black solid curve shows the behavior of the HF energy, while the red dashed line is the piecewise interpolations between neighboring integer energies. The black curve is no lower than the red line, although the deviation is small and they seem to overlap each other. The inset shows the enlarged plot for $5 \leq N \leq 7$, where the relative positions of the two lines become clearer, and the concavity of the HF functional at $[K, K + 1]$ is pronounced.

$E_v^{\text{HF}}(K + 1) \leq E_v^{\text{HF}}(K)$ and eventually reaches a constant as $K \rightarrow \infty$ due to ionization.²⁹ Moreover, it is often assumed (yet still an open conjecture) that $E_v^{\text{HF}}(N)$ is convex at integers, although recently we have proved that HF is concave restricted to $[K, K + 1]$ using the Aufbau construction,³⁰ which actually motivated the present study of equivalence of fractional formulations. As can be seen, all these features have been verified by Figure 1.

B. Aufbau principle for LDA

Here we assume the fractional occupations occur only in one spin channel, which is a reasonable ansatz due to the static correlation or the fractional spin error,²³ similar as in the HF case although this cannot be strictly proved for LDA. Moreover, we prescribe that the α electron number is no less than the β electron number. As a remark, for the exact functional, which is fractional spin error free, the scenario of fractional occupation in both spin channels could arise.^{22,24}

Once again we suppress the spin index for the ease of notations. By Janak's theorem or similar arguments as in Section III B, the minimizer of LDA with fractional N electrons is given by (note that the basic variable reduces to the electron density ρ for LDA)

$$\rho(\mathbf{r}) = \sum_{i=1}^L |\psi_i(\mathbf{r})|^2 + \sum_{k=L+1}^{L+M} n_k |\psi_k(\mathbf{r})|^2, \quad (59)$$

where $L \leq [N]$ and $\epsilon_1 \leq \epsilon_2 \leq \dots \leq \epsilon_L < \epsilon_f$. Here $\epsilon_f = \epsilon_{L+1} = \epsilon_{L+2} = \dots = \epsilon_{L+M}$ is the degenerate HOMO energy.

Now suppose $M \geq 2$. Then we perform the electron moving trick as in Section III B and let

$$\begin{aligned} \rho'_{ij}(\mathbf{r}) &= \rho(\mathbf{r}) + \eta (|\psi_{L+i}(\mathbf{r})|^2 - |\psi_{L+j}(\mathbf{r})|^2) \\ &= \rho(\mathbf{r}) + \eta g_{ij}(\mathbf{r}), \end{aligned} \quad (60)$$

where $1 \leq i < j \leq M$, $\eta \leq \min\{n_{L+i}, 1 - n_{L+i}\}$, and $g_{ij}(\mathbf{r}) \equiv |\psi_{L+i}(\mathbf{r})|^2 - |\psi_{L+j}(\mathbf{r})|^2$. We calculate $E[\rho'_{ij}] - E[\rho]$ up to

second order of η . Since the orbital energies of ψ_{L+i} and ψ_{L+j} are degenerate, the first order term of η vanishes and it follows that

$$\begin{aligned} E[\rho'_{ij}] - E[\rho] &= \frac{1}{2} \eta^2 \iint \frac{\delta^2 E}{\delta \rho(\mathbf{r}) \delta \rho(\mathbf{r}')} \frac{\partial \rho(\mathbf{r})}{\partial \eta} \frac{\partial \rho(\mathbf{r}')}{\partial \eta} d\mathbf{r} d\mathbf{r}' + \mathcal{O}(\eta^3) \\ &= \frac{1}{2} \eta^2 \left\{ \iint \frac{g_{ij}(\mathbf{r}) g_{ij}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \int f_{xc}(\mathbf{r}) |g_{ij}(\mathbf{r})|^2 d\mathbf{r} \right\} \\ &\quad + \mathcal{O}(\eta^3), \end{aligned} \quad (61)$$

where $f_{xc}(\mathbf{r}) = \frac{\delta^2 E_{xc}}{\delta \rho^2(\mathbf{r})}$ is the exchange correlation kernel. A necessary condition for Eq. (59) to be the minimizer is that for all pairs (i, j) , and any η small enough,

$$E[\rho'_{ij}] - E[\rho] \geq 0, \quad (62)$$

which is equivalent to

$$\kappa_{ij}^{\vec{n}} = \iint \frac{g_{ij}(\mathbf{r}) g_{ij}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \int f_{xc}(\mathbf{r}) |g_{ij}(\mathbf{r})|^2 d\mathbf{r} \geq 0, \quad (63)$$

for all (i, j) pairs. Here $\vec{n} = (1, 1, \dots, 1, n_{L+1}, n_{L+2}, \dots, n_{L+M})$ refers to the occupation configuration.

Thus, a sufficient condition for Eq. (59) not to be the minimizer is that

$$\min_{i,j} \kappa_{ij}^{\vec{n}} < 0. \quad (64)$$

And taking one step further, a sufficient condition for the equivalence of Janak and Aufbau constructions is that

$$\max_{\vec{n}} \min_{i,j} \kappa_{ij}^{\vec{n}} < 0. \quad (65)$$

For XLDA (LDA with exchange only functional) in particular, $E_{xc}[\rho] = -C_x \int [\rho(\mathbf{r})]^{4/3} d\mathbf{r}$, where $C_x = \frac{3}{2} (\frac{3}{4\pi})^{1/3}$ and $f_{xc}(\mathbf{r}) = -\frac{4}{9} C_x [\rho(\mathbf{r})]^{-2/3}$. Eq. (63) reads

$$\kappa_{ij}^{\vec{n}} = \iint \frac{g_{ij}(\mathbf{r}) g_{ij}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' - \frac{4}{9} C_x \int [\rho(\mathbf{r})]^{-2/3} |g_{ij}(\mathbf{r})|^2 d\mathbf{r}. \quad (66)$$

The first term of Eq. (66) is positive, while the second term is negative, suggesting a competition between the two terms in determining the sign of $\kappa_{ij}^{\vec{n}}$. In principle, one cannot make a conclusion on its sign or the optimal occupation configuration but resort to the numerical experiment, as will be presented in the following.

In Tables I and II, we select several fractionally charged atoms with spatial degeneracies and evaluate the LDA energies with different occupation configurations. Note that for systems without ground state degeneracy, only one fractional occupation can occur by the Janak's theorem. Therefore, we only have to test on systems with symmetric electronic structures in search for counterexamples. All the calculations below are done using our in-house built QM4D program,²⁸ for which the self-consistent fractional KS-DFT programs have been implemented and one can specify non-Aufbau occupation numbers during the SCF iterations. Here to reduce the degrees of freedom, we restrict the fractional electron to evenly distribute

TABLE I. Comparison of relative LDA energies with different occupation configurations for fractionally charged atoms whose neutral valence configuration is ns^2 or ns^2np^1 or ns^2nd^1 . Here we set the fractional charge $\delta_1 = 0.3$ for anions and $\delta_2 = 0.7$ for cations. For each configuration, the fractional electron (or hole) is restricted to evenly distribute over the fractional occupations, i.e., $n_{L+i} = \frac{\delta}{M}$, for $1 \leq i \leq M$. We set the zero energy to be the minimizer with only one fractional occupation and tabulate the relative energy of other configurations. If not specially noted, the fractionally occupied orbitals within a minimizer are degenerate. The basis used for the calculations are aug-cc-pVTZ for Be, B, Mg, Al, and Ga; cc-pVTZ for Ca, Sc, and Zn; TZP for Sr; ATZP for In; and ADZP for Ba and Tl.

	$\Delta E \times 10^4$ (a.u.)					$\Delta E \times 10^4$ (a.u.)			
	$M = 1$	$M = 2$	$M = 3$	$M = 5$		$M = 1$	$M = 2$	$M = 3$	$M = 5$
Be $^{\delta_1-}$	0	3.14	4.33	59.7 ^a	B $^{\delta_2+}$	0	5.02	6.93	358 ^a
Mg $^{\delta_1-}$	0	3.14	4.37	45.6 ^a	Al $^{\delta_2+}$	0	3.13	4.37	223 ^a
Ca $^{\delta_1-}$	0	0.89	1.68	2.60	Sc $^{\delta_2+}$	0	15.4	16.2 ^b	29.8 ^b
Zn $^{\delta_1-}$	0	3.08	4.29	159 ^a	Ga $^{\delta_2+}$	0	2.33	3.28	226 ^a
Sr $^{\delta_1-}$	0	2.23	3.09	38.1 ^a	In $^{\delta_2+}$	0	0.60	0.91	241 ^a
Ba $^{\delta_1-}$	0	2.67	3.74	268 ^a	Tl $^{\delta_2+}$	0	1.31	1.89	201 ^a

^aFrontier orbital degeneracy breaks up.

^bThe fractionally occupied orbitals are nearly degenerate.

among M fractionally occupied orbitals of the same spin. In Table I, we select atoms whose neutral valence configuration is singlet ns^2 or doublet ns^2np^1 or ns^2nd^1 , while in Table II we focus on atoms whose neutral valence configuration is triplet ns^2np^2 or quartet ns^2np^3 and whose neutral α valence configuration is ns^1np^2 or ns^1np^3 .

As can be seen, M cannot be too large; otherwise, the dictated orbital degeneracy exceeds the possible spatial degeneracy of the effective KS potential, which leads to the break-up of the fractionally occupied orbital energies. Such configuration violates the frontier degeneracy rule in the Aufbau principle and thus cannot be the ground state. Moreover, for the several allowed choices of M , Table I shows that the energy becomes less favorable when M increases, and the lowest energy configuration is given by the one with only one fractional occupation. This, however, is not true in Table II, where we have found several examples where the optimum energy is achieved by configurations with multiple fractional occupancies, such as the cationic Si and Ge, which confirms that the Aufbau constructions do not necessarily lead to single fractional occupation.

In fact, such phenomena not only exist in fractionally charged systems; they are observed in integer systems as well. In Table III, we tabulate the relative energies of different electron configurations for some spatially symmetric atoms and molecular cations. Here for molecular cations, the geometries are taken from the G2-97 set.³¹ Based on the numbers shown, for atoms the situation is somewhat similar to Table II. For example, the ground state configuration of a B atom by LDA has no fractional occupancy, which forms a distinct contrast to Si and Ge, whose ground state is given by multiple fractional occupancies. This can be ascribed to the coexistence and competition of static correlation error and delocalization error⁷ of LDA, where in the B atom case the former dominates, leading to a positive error of fractional occupancies relative to the integer occupancies; while in the Si or Ge cases, the latter wins and results in a negative error. In the absence of both errors, the two configurations should be degenerate in the exact functional. On the other hand, for molecular cations with spatial degeneracy, such as CH₄⁺, we have not found an example where the ground state is given by multiple occupations.

TABLE II. Comparison of relative LDA energies with different occupation configurations for fractionally charged atoms whose neutral valence configuration for α electron is ns^1np^2 or ns^1np^3 . For each configuration, the fractional electron (or hole) is restricted to evenly distribute over M orbitals (all in the α spin channel) that have been specified in the table. We set the zero energy to be the minimizer with only one fractional occupation and tabulate the relative energy of other configurations. If not specially noted, the fractionally occupied orbitals within a minimizer are degenerate. Here the three np orbitals are linear combinations of np_x , np_y , and np_z , here denoted as np_1 , np_2 , and np_3 , and we denote $(p^\delta)^3 \equiv p_1^\delta p_2^\delta p_3^\delta$. The basis used for the calculations are aug-cc-pVTZ for all the atoms.

	$\Delta E \times 10^4$ (a.u.)					$\Delta E \times 10^4$ (a.u.)			
	$M = 1$	$M = 2$	$M = 3$	$M = 4$		$M = 1$	$M = 2$	$M = 3$	$M = 4$
	$s^1 p_1^1 p_2^{0.8}$	$s^1 p_1^{0.9} p_2^{0.9}$	$s^1 (p^{0.6})^3$	$s^{0.7} (p^{0.7})^3$		$s^1 p_1^1 p_2^1 p_3^{0.7}$	$s^1 p_1^1 p_2^{0.85} p_3^{0.85}$	$s (p^{0.9})^3$	$s^{0.925} (p^{0.925})^3$
C $^{0.2+}$	0	-0.48	6.45	92.6 ^a	N $^{0.3+}$	0	-2.21 ^b	-3.03	309 ^a
Si $^{0.2+}$	0	-0.80	-3.50	75.2 ^a	P $^{0.3+}$	0	-2.06 ^b	-2.80	23.4 ^a
Ge $^{0.2+}$	0	-0.78 ^b	-11.3	83.7 ^a	As $^{0.3+}$	0	-2.56 ^b	-3.45	24.7 ^a

^aFrontier orbital degeneracy breaks up by rendering s orbital energy significantly lower than the three degenerate p orbitals.

^bFrontier orbital degeneracy breaks up slightly.

TABLE III. Comparison of relative LDA energies with different occupation configurations for some integer systems with doublet or triplet multiplicity (abbreviated as D or T) and with spatial symmetry. For each configuration, the valence electrons are restricted to evenly distribute over the degenerate frontier orbitals. In particular, for doublet states, the excess electron in the α spin channel or the excess hole in the β spin channel relative to the other spin is divided into M fractionally occupied orbitals. For triplet states, the occupation configuration for the two excess α electrons are $(1, \frac{1}{2}, \frac{1}{2})$ for $M = 2$, $(\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$ for $M = 3$, and $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ for $M = 4$. We set the zero energy to be the minimizer with no fractional occupation (here denoted by $M = 0$) and tabulate the relative energy of other configurations. If not specially noted, the fractionally occupied orbitals within a minimizer are degenerate. The basis used are aug-cc-pVTZ for atoms and 6-311++G(3df,3pd) for molecules.

Multiplicity		$\Delta E \times 10^4$ (a.u.)				Multiplicity		$\Delta E \times 10^4$ (a.u.)			
		$M = 0$	$M = 2$	$M = 3$	$M = 4$			$M = 0$	$M = 2$	$M = 3$	$M = 4$
B	D	0	15.3	24.2	411 ^a	CH ₄ ⁺	D	0	8.00	9.05	720 ^a
Al	D	0	7.31	12.5	261 ^a	SiH ₄ ⁺	D	0	... ^b	3.42	487 ^a
Ga	D	0	2.40	5.78	247 ^a	BF ₃ ⁺	D	0	... ^b	204 ^a	... ^b
C	T	0	4.92 ^c	2.62	1295 ^a	BF ₃ ⁻	D	0	135 ^c	225 ^c	275 ^c
Si	T	0	-3.32 ^c	-6.92	827 ^a	CH ₄ ²⁺	T	0	17.2	25.1	1493 ^a
Ge	T	0	-9.85 ^c	-15.3	775 ^a	SiH ₄ ²⁺	T	0	5.94	8.55	1035 ^a

^aFrontier orbital degeneracy breaks up by rendering one fractionally occupied orbital energy significantly different from the rest of the orbitals.

^bSelf consistent field convergence fails.

^cFrontier orbital degeneracy breaks up slightly.

To obtain some insight on when a general system should have multiple fractional occupancies as its ground state and how to assign occupations, in Table IV, we present the spectrum (here we only show the levels near the frontier energy) of a Si atom optimized with different prescribed occupation configurations. As can be seen, the overall structure of the spectrum is insensitive to the changes of the occupations, although each of the levels might change slightly. From the spectrum obtained by applying the stringent Aufbau rule (for the triplet Si case, this corresponds to the energy levels under the column $M = 0$), one can immediately know which levels can share electron occupancies in order to further decrease the total energy. The Si case has three near-degenerate p orbitals spanning 2 occupied and 1 unoccupied levels in the stringent Aufbau configuration, which can be rearranged as two electrons being shared in these orbitals, leading to the occupation $(\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$. One can then perform calculation for such configurations and compare energy with the stringent Aufbau configuration in search for the true ground state of LDA.

TABLE IV. Orbital energies and occupation numbers (for α electron) of a Si atom with different occupation configurations. Here we only present the orbitals close to the frontier level and the three np orbitals are linear combinations of np_x , np_y , and np_z , here denoted as np_1 , np_2 , and np_3 . All energies are in eV.

	$M = 0$		$M = 2$		$M = 3$		$M = 4$	
	ϵ_i	n_i	ϵ_i	n_i	ϵ_i	n_i	ϵ_i	n_i
3s	-11.308	1	-11.347	1	-11.355	1	-13.112	1
3p ₁	-4.599	1	-4.649	1	-4.638	2	-6.154	0.5
3p ₂	-4.579	1	-4.649	0.5	-4.638	2	-6.154	0.5
3p ₃	-4.578	0	-4.588	0.5	-4.638	3	-6.154	0.5
4s	-0.230	0	-0.252	0	-0.256	0	-1.285	0.5
4p ₁	0.709	0	0.707	0	0.710	0	-0.305	0
4p ₂	0.710	0	0.715	0	0.710	0	-0.305	0
4p ₃	0.765	0	0.715	0	0.710	0	-0.305	0

VI. CONCLUSION

In this paper, we have analyzed four ways of extending KS-DFT to systems with a fractional number of electrons by enlarging the search space of densities. In particular, the search space has been extended from the KS non-interacting v -representable density domain for integer systems to four different sets of densities for fractional N -electron systems, namely, the ensemble N -representable densities, the ensemble non-interacting N -representable densities, and the set of non-interacting densities by the Janak and Aufbau constructions. We have shown that the first two sets are equivalent to the Janak set by proving the set equivalence of their underlying 1-RDMs, and thus the three formulations reduce to the Janak construction. Moreover, the Janak construction reduces to the Aufbau construction for functionals with the ensemble v -representable assumption at the minimizer. Therefore, given a DFA, by solving the (G)KS equation and following the Aufbau occupation protocol, one obtains the ground state of the fractional system defined through the ensemble search. By further analyzing the structure of Aufbau solutions, we have proved that there can be one and only one fractional occupation for the HF functional, while for other DFAs, this is not true in general; there can be multiple fractional occupations at the Fermi level. By testing the LDA functional in real systems, we find that although single fractional occupation is observed in many systems in the presence of degeneracy, we can also find counterexamples, and the fact that the energy of multiple fractional occupations differs from the single occupation configuration for such systems is due to the intrinsic error of LDA in the presence of multiple fractional occupations in degenerate orbitals.

In terms of the shape of the E vs N curves of approximate functionals, we have illustrated that HF should display a piecewise concave curve interpolating the integer points, with the integer energies being monotone non-increasing, as shown previously.⁴ For general DFAs, such as semi-local density functionals, the situation becomes more complicated.

The monotone non-increase of integer energies is still true due to the same ionization argument as shown in Ref. 29, rendering non-negative electron affinity. Although there have been reports on negative electron affinities by DFAs calculated for atoms and molecules (see the supplemental information of Ref. 9, for example), these are likely due to the use of finite basis set.³² In terms of the shape of the E vs N curve of approximate DFAs restricted to $[K, K + 1]$, it has been conjectured that it is convex for LDA, for example, yet this conjecture has been recently shown to be not true in general within the Aufbau construction.³⁰

For the exact functional, the E vs N curve should be a piecewise linear interpolation between neighboring integers, under the assumption that the integer energies are convex.¹⁴ Yet the validity of this assumption remains an interesting and open problem to be addressed.

ACKNOWLEDGMENTS

Support from the National Science Foundation (CHE-13-62927 (C.L.) and DMS-14-54939 (J.L.)) is gratefully appreciated. W.Y. appreciates the support as part of the Center for the Computational Design of Functional Layered Materials, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award No. DE-SC0012575.

APPENDIX: SUPPLEMENTAL PROOF OF FRACTIONAL OCCUPATION RULE FOR HF FUNCTIONAL

Here we show that there could be one and only one fractional occupation through one of the spin channels in the minimizer of unrestricted HF. Consider the HF functional as a functional of the spin density matrix,

$$\begin{aligned} E[\gamma^\alpha(\mathbf{r}, \mathbf{r}'), \gamma^\beta(\mathbf{r}, \mathbf{r}')] &= \sum_{\sigma=\alpha,\beta} \int -\frac{1}{2} \nabla_{\mathbf{r}'}^2 \gamma^\sigma(\mathbf{r}, \mathbf{r}') \Big|_{\mathbf{r}'=\mathbf{r}} d\mathbf{r} + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \\ &+ \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \\ &- \sum_{\sigma=\alpha,\beta} \frac{1}{2} \iint \frac{\gamma^\sigma(\mathbf{r}, \mathbf{r}')\gamma^\sigma(\mathbf{r}', \mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}', \end{aligned} \quad (\text{A1})$$

where $\rho(\mathbf{r}) = \gamma^\alpha(\mathbf{r}, \mathbf{r}) + \gamma^\beta(\mathbf{r}, \mathbf{r})$. Note that the quadratic terms of γ^σ in Eq. (A1) are the same (treating $\gamma^{\sigma'}$ as constant, $\sigma' \neq \sigma$) as the quadratic dependence of the spinless HF energy functional on the spinless 1-RDM introduced in Eq. (56).

First of all, we claim that for each γ^σ , there could be at most one fractional occupation in the minimizer. Otherwise, without loss of generality, assuming there are two fractional occupations in the α spin channel, we can fix γ^β and construct two new γ^α as in Eqs. (54) and (55) to derive a contradiction (derivations are straightforward and omitted here).

Second, we show that the minimizer cannot have fractional occupations in both spin channels. Assume otherwise, let

$$\gamma^\alpha(\mathbf{r}, \mathbf{r}') = \sum_{k=1}^{N_\alpha} \psi_k^\alpha(\mathbf{r})\psi_k^\alpha(\mathbf{r}') + n_\alpha \psi_f^\alpha(\mathbf{r})\psi_f^\alpha(\mathbf{r}'), \quad (\text{A2})$$

$$\gamma^\beta(\mathbf{r}, \mathbf{r}') = \sum_{k=1}^{N_\beta} \psi_k^\beta(\mathbf{r})\psi_k^\beta(\mathbf{r}') + n_\beta \psi_f^\beta(\mathbf{r})\psi_f^\beta(\mathbf{r}') \quad (\text{A3})$$

be the minimizer, where $0 < n_\alpha, n_\beta < 1$. Let $\eta = \min\{n_\alpha, n_\beta, 1 - n_\alpha, 1 - n_\beta\}$, and define

$$\gamma_I^\alpha(\mathbf{r}, \mathbf{r}') = \gamma^\alpha(\mathbf{r}, \mathbf{r}') + \eta \psi_f^\alpha(\mathbf{r})\psi_f^\alpha(\mathbf{r}'), \quad (\text{A4})$$

$$\gamma_I^\beta(\mathbf{r}, \mathbf{r}') = \gamma^\beta(\mathbf{r}, \mathbf{r}') - \eta \psi_f^\beta(\mathbf{r})\psi_f^\beta(\mathbf{r}'), \quad (\text{A5})$$

and

$$\gamma_{II}^\alpha(\mathbf{r}, \mathbf{r}') = \gamma^\alpha(\mathbf{r}, \mathbf{r}') - \eta \psi_f^\alpha(\mathbf{r})\psi_f^\alpha(\mathbf{r}'), \quad (\text{A6})$$

$$\gamma_{II}^\beta(\mathbf{r}, \mathbf{r}') = \gamma^\beta(\mathbf{r}, \mathbf{r}') + \eta \psi_f^\beta(\mathbf{r})\psi_f^\beta(\mathbf{r}'). \quad (\text{A7})$$

We now evaluate $E[\gamma_I^\alpha, \gamma_I^\beta] + E[\gamma_{II}^\alpha, \gamma_{II}^\beta] - 2E[\gamma^\alpha, \gamma^\beta]$. Since $\gamma_I^\sigma + \gamma_{II}^\sigma - 2\gamma^\sigma = 0$, it suffices for us to evaluate the quadratic terms of the spin 1-RDMs. These terms are

$$\begin{aligned} E[\gamma_I^\alpha, \gamma_I^\beta] + E[\gamma_{II}^\alpha, \gamma_{II}^\beta] - 2E[\gamma^\alpha, \gamma^\beta] &= E_J[\gamma_I^\alpha, \gamma_I^\beta] + E_J[\gamma_{II}^\alpha, \gamma_{II}^\beta] - 2E_J[\gamma^\alpha, \gamma^\beta] \\ &- E_K[\gamma_I^\alpha, \gamma_I^\beta] - E_K[\gamma_{II}^\alpha, \gamma_{II}^\beta] + 2E_K[\gamma^\alpha, \gamma^\beta], \end{aligned} \quad (\text{A8})$$

where E_J and E_K are Hartree and exchange terms, respectively,

$$\begin{aligned} E_J[\gamma_I^\alpha, \gamma_I^\beta] + E_J[\gamma_{II}^\alpha, \gamma_{II}^\beta] - 2E_J[\gamma^\alpha, \gamma^\beta] &= \frac{1}{2} \iint \frac{\rho_I(\mathbf{r})\rho_I(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + \frac{1}{2} \iint \frac{\rho_{II}(\mathbf{r})\rho_{II}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \\ &- \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \end{aligned} \quad (\text{A9})$$

Here $\rho_I(\mathbf{r}) = \gamma_I^\alpha(\mathbf{r}, \mathbf{r}) + \gamma_I^\beta(\mathbf{r}, \mathbf{r}) = \rho(\mathbf{r}) + \eta\zeta(\mathbf{r})$, with $\zeta(\mathbf{r}) = |\psi_f^\alpha(\mathbf{r})|^2 - |\psi_f^\beta(\mathbf{r})|^2$ and $\rho_{II}(\mathbf{r}) = \rho(\mathbf{r}) - \eta\zeta(\mathbf{r})$. By straightforward calculation, Eq. (A9) reads

$$\begin{aligned} E_J[\gamma_I^\alpha, \gamma_I^\beta] + E_J[\gamma_{II}^\alpha, \gamma_{II}^\beta] - 2E_J[\gamma^\alpha, \gamma^\beta] &= \eta^2 \iint \frac{\zeta(\mathbf{r})\zeta(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \end{aligned} \quad (\text{A10})$$

On the other hand,

$$\begin{aligned} -E_K[\gamma_I^\alpha, \gamma_I^\beta] - E_K[\gamma_{II}^\alpha, \gamma_{II}^\beta] + 2E_K[\gamma^\alpha, \gamma^\beta] &= -\frac{1}{2} \sum_{\sigma=\alpha,\beta} \left\{ \iint \frac{\gamma_I^\sigma(\mathbf{r}, \mathbf{r}')\gamma_I^\sigma(\mathbf{r}', \mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \right. \\ &+ \iint \frac{\gamma_{II}^\sigma(\mathbf{r}, \mathbf{r}')\gamma_{II}^\sigma(\mathbf{r}', \mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \\ &\left. - 2 \iint \frac{\gamma^\sigma(\mathbf{r}, \mathbf{r}')\gamma^\sigma(\mathbf{r}', \mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \right\} \\ &= -\eta^2 \sum_{\sigma=\alpha,\beta} \iint \frac{|\psi_f^\sigma(\mathbf{r})|^2 |\psi_f^\sigma(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \end{aligned} \quad (\text{A11})$$

Plugging Eqs. (A10) and (A11) into Eq. (A8) and cancelling out the opposite terms, we obtain

$$\begin{aligned} E[\gamma_I^\alpha, \gamma_I^\beta] + E[\gamma_{II}^\alpha, \gamma_{II}^\beta] - 2E[\gamma^\alpha, \gamma^\beta] &= -2\eta^2 \iint \frac{|\psi_f^\alpha(\mathbf{r})|^2 |\psi_f^\beta(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' < 0, \end{aligned} \quad (\text{A12})$$

which is a contradiction. Therefore, only one of the spin channels has fractional occupation.

We remark that in the literature, it has been recognized that the energy with fractional occupations in both spin channels is often overestimated by approximate DFAs, which has been referred to as the fractional spin error.²³ Here Eq. (A12) is a quantification of the fractional spin error for HF. For other DFAs, one can derive a generic formula dependent on the specific form of E_{xc} , but the negativity of Eq. (A12) can no longer be proved.

- ¹W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- ²R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- ³A. J. Cohen, P. Mori-Sánchez, and W. Yang, *Chem. Rev.* **112**, 289 (2012).
- ⁴P. Mori-Sánchez, A. J. Cohen, and W. Yang, *J. Chem. Phys.* **125**, 201102 (2006).
- ⁵A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria, *J. Chem. Phys.* **125**, 194112 (2006).
- ⁶A. J. Cohen, P. Mori-Sánchez, and W. Yang, *Phys. Rev. B* **77**, 115123 (2008).
- ⁷P. Mori-Sánchez, A. J. Cohen, and W. Yang, *Phys. Rev. Lett.* **100**, 146401 (2008).
- ⁸A. J. Cohen, P. Mori-Sánchez, and W. Yang, *Science* **321**, 792 (2008).
- ⁹X. Zheng, A. J. Cohen, P. Mori-Sánchez, X. Hu, and W. Yang, *Phys. Rev. Lett.* **107**, 026403 (2011).
- ¹⁰J.-D. Chai and P.-T. Chen, *Phys. Rev. Lett.* **110**, 033002 (2013).
- ¹¹E. Kraisler and L. Kronik, *Phys. Rev. Lett.* **110**, 126403 (2013).
- ¹²T. Gould and J. F. Dobson, *J. Chem. Phys.* **138**, 014103 (2013).
- ¹³C. Li, X. Zheng, A. J. Cohen, P. Mori-Sánchez, and W. Yang, *Phys. Rev. Lett.* **114**, 053001 (2015).
- ¹⁴J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, *Phys. Rev. Lett.* **49**, 1691 (1982).
- ¹⁵P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964).
- ¹⁶E. H. Lieb, *Int. J. Quantum. Chem.* **24**, 243 (1983).
- ¹⁷M. Levy, *Proc. Natl. Acad. Sci. U. S. A.* **76**, 6062 (1979).
- ¹⁸S. M. Valone, *J. Chem. Phys.* **73**, 1344 (1980).
- ¹⁹A. J. Coleman, *Rev. Mod. Phys.* **35**, 668 (1963).
- ²⁰T. L. Gilbert, *Phys. Rev. B* **12**, 2111 (1975).
- ²¹J. F. Janak, *Phys. Rev. B* **18**, 7165 (1978).
- ²²W. Yang, Y. K. Zhang, and P. W. Ayers, *Phys. Rev. Lett.* **84**, 5172 (2000).
- ²³A. J. Cohen, P. Mori-Sánchez, and W. Yang, *J. Chem. Phys.* **129**, 121104 (2008).
- ²⁴P. Mori-Sánchez, A. J. Cohen, and W. Yang, *Phys. Rev. Lett.* **102**, 066403 (2009).
- ²⁵X. D. Yang, A. H. G. Patel, R. A. Miranda-Quintana, F. Heidar-Zadeh, C. E. Gonzalez-Espinoza, and P. W. Ayers, *J. Chem. Phys.* **145**, 031102 (2016).
- ²⁶E. Lieb, *Phys. Rev. Lett.* **46**(7), 457 (1981).
- ²⁷V. Bach, *Commun. Math. Phys.* **147**, 527 (1992).
- ²⁸X. Hu, H. Hu, X. Zheng, X. Zeng, P. Wu, D. Peng, Y. Jin, L. Yu, and W. Yang, QM4D, <http://www.qm4d.info>.
- ²⁹E. Lieb and B. Simon, *Commun. Math. Phys.* **53**, 185 (1977).
- ³⁰C. Li and W. Yang, *J. Chem. Phys.* **146**, 074107 (2017).
- ³¹L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, *J. Chem. Phys.* **106**, 1063 (1997).
- ³²M. J. G. Peach, A. M. Teale, T. Helgaker, and D. J. Tozer, *J. Chem. Theory Comput.* **11**, 5262 (2015).