

Chapter 4

The Strong-Interaction Limit of Density Functional Theory



Gero Friesecke, Augusto Gerolin, and Paola Gori-Giorgi

Abstract This is a comprehensive review of the strong-interaction limit of density functional theory. It covers the derivation of the limiting strictly correlated electrons (SCE) functional from exact Hohenberg–Kohn DFT, basic aspects of SCE physics such as the nonlocal dependence of the SCE potential on the density, equivalent formulations and the mathematical interpretation as optimal transport with Coulomb cost, rigorous results (including exactly soluble cases), approximations, numerical methods, integration into Kohn–Sham DFT (KS SCE), and applications to molecular systems, an example being that KS SCE, unlike the local density approximation or generalized gradient approximations, dissociates H_2 correctly. We have made an effort to make this review accessible to a broad audience of physicists, chemists, and mathematicians.

4.1 Introduction

The strong-interaction limit of DFT is the inhomogeneous low-density limit associated with the uniform coordinate scaling

$$\rho_\gamma(\mathbf{r}) = \gamma^3 \rho(\gamma\mathbf{r})$$

G. Friesecke (✉)

Department of Mathematics, Technische Universität München, Garching bei München, Germany
e-mail: gf@ma.tum.de

A. Gerolin (✉)

Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, Canada

Department of Mathematics and Statistics, University of Ottawa, Ottawa, Canada

P. Gori-Giorgi (✉)

Department of Theoretical Chemistry, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

AI4Science Microsoft Research, Amsterdam, The Netherlands

of the single-particle density at fixed particle number, with $\gamma \rightarrow 0$. In this limit, the Levy–Lieb functional, which gives the minimum kinetic and interaction energy subject to the given density, has the leading order asymptotics

$$F_{\text{LL}}[\rho_\gamma] \sim \gamma V_{\text{ee}}^{\text{SCE}}[\rho],$$

and the corresponding optimal wavefunction Ψ_γ has the asymptotics

$$\sum_{s_1, \dots, s_N \in \mathbb{Z}_2} |\Psi_\gamma(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N)|^2 \sim \gamma^{3N} \rho_N(\gamma \mathbf{r}_1, \dots, \gamma \mathbf{r}_N),$$

where ρ_N solves the variational principle of having minimal Coulomb energy subject to the given density ρ and $V_{\text{ee}}^{\text{SCE}}[\rho]$ denotes the resulting minimal energy.

This appears to be the only case in which one can obtain insight into how to extract information about the interaction energy directly from the density. As it turns out, in this limit none of the ingredients from the traditional “Jacob’s ladder” of DFT approximations (local density, local density gradients, Kohn–Sham kinetic energy density, Hartree–Fock exchange, virtual orbitals) play any role. Instead, maps based on integrals not derivatives of the density appear. These maps are mathematically related to the field of optimal transport, and physically describe strictly correlated electrons (SCE). The SCE functional $V_{\text{ee}}^{\text{SCE}}$ appearing above is the limiting Hartree–exchange–correlation functional.

While the strong-interaction limit is, of course, not reached in nature, it points the way towards the real physics happening in molecular systems containing strong correlations, without having to leave the realm of Kohn–Sham DFT. Two important examples whose physics is missed by Kohn–Sham DFT with semilocal or hybrid exchange–correlation functionals but captured correctly by integrating the SCE functional into Kohn–Sham DFT (KS SCE) are weakly charged nanosystems, see Fig. 4.11, and H_2 near the dissociation limit, see Fig. 4.14.

This chapter provides a self-contained introduction to this limit and its fascinating physics and mathematics which has been unearthed in the past two decades, and reviews the current state of the art.

4.2 The Many-Electron Schrödinger Equation and Universal Density Functional

In this section we quickly introduce the time-independent electronic Schrödinger equation and the exact reformulation of the ground state problem via a universal density functional.

4.2.1 The Many-Electron Schrödinger Equation

We consider a quantum mechanical system of N non-relativistic electrons (of mass m_e and charge $-e$), moving around classical nuclei with positions $\mathbf{R}_1, \dots, \mathbf{R}_M \in \mathbb{R}^d$ and charges Z_1e, \dots, Z_Me (Born–Oppenheimer approximation). Our main interest is in the physical space \mathbb{R}^3 , but we consider the general space dimension $d \geq 1$ since it will be instructive to illustrate key properties of the strong interaction limit with lower dimensional examples. The electrons are described by a wave function $\Psi : (\mathbb{R}^d \times \mathbb{Z}_2)^N \rightarrow \mathbb{C}$ of N positions $\mathbf{r}_i \in \mathbb{R}^d$ and spin coordinates $s_i \in \{\uparrow, \downarrow\} = \mathbb{Z}_2$.

The Pauli exclusion principle states that the electronic wave function must be antisymmetric with respect to permutations of the electron coordinates,

$$\Psi(\mathbf{r}_{\sigma(1)}, s_{\sigma(1)}, \dots, \mathbf{r}_{\sigma(N)}, s_{\sigma(N)}) = \text{sign}(\sigma) \Psi(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N), \quad \sigma \in \mathfrak{S}_N, \quad (4.1)$$

where \mathfrak{S}_N denotes the group of permutations of the indices $1, \dots, N$. The set of square-integrable N -electron wave functions, $\{\Psi \in L^2((\mathbb{R}^d \times \mathbb{Z}_2)^N; \mathbb{C}) : (4.1)\}$, will be denoted $\bigwedge_{i=1}^N L^2(\mathbb{R}^d \times \mathbb{Z}_2; \mathbb{C})$. The one-body density of an electronic wave function $\Psi \in \bigwedge_{i=1}^N L^2(\mathbb{R}^d \times \mathbb{Z}_2; \mathbb{C})$ is defined by

$$\rho_\Psi(\mathbf{r}_j) = N \sum_{s_1, \dots, s_N \in \mathfrak{S}_N} \int_{\mathbb{R}^{d(N-1)}} |\Psi(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N)|^2 \prod_{i \neq j} d\mathbf{r}_i, \\ \forall j \in \{1, \dots, N\}.$$

The energy $E[\Psi, v]$ of a fermionic state Ψ with external potential $v : \mathbb{R}^d \rightarrow \mathbb{R}$ is given, in atomic units $\hbar = m_e = e = 1$, by

$$E[\Psi, v] = T[\Psi] + V_{\text{ec}}[\Psi] + V_{\text{nc}}[\Psi, v], \quad (4.2)$$

where $T[\Psi]$ is the *kinetic energy*,

$$T[\Psi] = \frac{1}{2} \sum_{s_1 \in \mathbb{Z}_2} \int_{\mathbb{R}^d} \cdots \sum_{s_N \in \mathbb{Z}_2} \int_{\mathbb{R}^d} \sum_{i=1}^N |\nabla_{\mathbf{r}_i} \Psi(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N)|^2 d\mathbf{r}_1 \cdots d\mathbf{r}_N;$$

$V_{\text{ec}}[\Psi]$ is the electron–electron interaction energy

$$V_{\text{ec}}[\Psi] = \\ \sum_{s_1 \in \mathbb{Z}_2} \int_{\mathbb{R}^d} \cdots \sum_{s_N \in \mathbb{Z}_2} \int_{\mathbb{R}^d} \sum_{1 \leq i < j < N} w(\mathbf{r}_i - \mathbf{r}_j) |\Psi(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N)|^2 d\mathbf{r}_1 \cdots d\mathbf{r}_N,$$

and $V_{\text{nc}}[\Psi, v]$ is the electron-nuclei interaction energy,

$$V_{\text{nc}}[\Psi, v] = \sum_{s_1 \in \mathbb{Z}_2} \int_{\mathbb{R}^d} \cdots \sum_{s_N \in \mathbb{Z}_2} \int_{\mathbb{R}^d} \sum_{i=1}^N v(\mathbf{r}_i) |\Psi(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N)|^2 d\mathbf{r}_1 \dots d\mathbf{r}_N,$$

where $w : \mathbb{R}^d \rightarrow \mathbb{R}$ is an interaction potential satisfying $w(\mathbf{r}) = w(-\mathbf{r})$, so that the total interaction potential

$$V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{1 \leq i < j \leq N} w(\mathbf{r}_i - \mathbf{r}_j) \quad (4.3)$$

is symmetric.¹ Typically,

$$w(\mathbf{r}) = |\mathbf{r}|^{-1} \quad (4.4)$$

is the Coulomb electron repulsion and v is the Coulomb potential generated by M nuclei which are at positions \mathbf{R}_ν with charges Z_ν ,

$$v(\mathbf{r}) = - \sum_{\nu=1}^M \frac{Z_\nu}{|\mathbf{r} - \mathbf{R}_\nu|}. \quad (4.5)$$

If additional fields are present, the external potential v contains extra terms.

The central quantity of interest is the *ground state energy* of the system. By the *Rayleigh–Ritz variational principle*, it is given by

$$E_0[v] = \inf\{E[\Psi, v] : \Psi \in \mathcal{W}^N\}, \quad (4.6)$$

where the infimum is taken over the class \mathcal{W}^N of wavefunctions which are antisymmetric and have finite kinetic energy,

$$\mathcal{W}^N = \left\{ \Psi \in \bigwedge_{i=1}^N H^1(\mathbb{R}^d \times \mathbb{Z}_2; \mathbb{C}) : \sum_{s_1, \dots, s_N \in \mathbb{Z}_2} \int_{\mathbb{R}^{dN}} |\nabla \Psi|^2 d\mathbf{r}_1 \dots d\mathbf{r}_N < +\infty, \|\Psi\| = 1 \right\}. \quad (4.7)$$

Here H^1 is the usual Sobolev space of square-integrable functions with square-integrable gradient, and $\|\Psi\|$ denotes the L^2 norm of Ψ . The ground state energy (4.6) is well defined whenever the potentials v and w are sufficiently

¹ We follow the usual convention to use the same letter V_{ee} both for the total interaction potential, a function on \mathbb{R}^{dN} , and the associated quadratic form, a functional on the wavefunction space \mathcal{W}^N .

regular so that the functional E is well defined on $H^1((\mathbb{R}^d \times \mathbb{Z}_2)^N)$. A simple sufficient condition in dimension $d = 3$ which encompasses (4.5), (4.4) is $v, w \in L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$.

Whether or not the infimum in (4.6) is actually a minimum, that is, a minimizing Ψ exists, is much more subtle. For neutral or positively charged molecules in dimension $d = 3$ ((4.4), (4.5) with $Z = \sum_{i=1}^M Z_i > N - 1$) the answer is yes, as was proved by Zhislin [145] via a careful spectral analysis of the underlying Hamiltonian operator. For an alternative proof based on variational methods see Friesecke [48].

4.2.2 Universal Density Functional

In quantum mechanics, the absolute value squared $|\Psi(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N)|^2$ of a wave function $\Psi \in \mathcal{W}^N$ corresponds to an N -point probability distribution: it gives the probability density of finding the electrons at positions $\mathbf{r}_i \in \mathbb{R}^d$ with spins $s_i \in \mathbb{Z}_2$, $i \in \{1, \dots, N\}$.

By integrating the N -point probability distribution over the spins, we obtain the N -point position density,

$$\pi_N^\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) := \sum_{s_1, \dots, s_N \in \mathbb{Z}_2} |\Psi(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N)|^2, \quad \Psi \in \mathcal{W}^N. \quad (4.8)$$

The single particle density $\rho_\Psi(\mathbf{r}_j)$ is then obtained by integrating out all but one electron position $\mathbf{r}_j \in \mathbb{R}^d$,

$$\rho_\Psi(\mathbf{r}_j) := N \int_{\mathbb{R}^{d(N-1)}} \pi_N^\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_j, \dots, \mathbf{r}_N) \prod_{i \neq j} d\mathbf{r}_i, \quad \forall j \in \{1, \dots, N\}. \quad (4.9)$$

We denote by $\Psi \mapsto \rho$ the relation between Ψ and ρ given by Eqs. (4.8) and (4.9). This means that the wave function Ψ has single-electron density ρ .

Following the work of Hohenberg and Kohn [78], Levy [91] and Lieb [97] showed that the electronic ground state problem (4.6) can be recast as a minimization over single-electron densities ρ instead of many-electron wavefunctions Ψ :

$$E_0[v_{\text{ne}}] = \inf_{\rho \in \mathcal{D}^N} \left\{ F_{\text{LL}}[\rho] + N \int_{\mathbb{R}^d} v_{\text{ne}}(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \right\}, \quad (4.10)$$

with

$$F_{\text{LL}}[\rho] = \min \left\{ T[\Psi] + V_{\text{ce}}[\Psi] : \Psi \in \mathcal{W}^N, \Psi \mapsto \rho \right\}, \quad (4.11)$$

where F_{LL} is the Levy–Lieb functional. The above direct definition of F_{LL} by a constrained search replaced an earlier, indirect existence proof of a universal functional satisfying (4.10) [78]. The space \mathcal{D}^N is defined as the set of densities ρ coming from a wave function $\Psi \in \mathcal{W}^N$ (i.e. $\Psi \mapsto \rho$), i.e., the N -representable one-particle densities. It can be fully characterized [97] and is given by

$$\mathcal{D}^N = \left\{ \rho \in L^1(\mathbb{R}^d) : \rho \geq 0, \sqrt{\rho} \in H^1(\mathbb{R}^d), \int_{\mathbb{R}^d} \rho = N \right\}. \quad (4.12)$$

Also, it is known that the minimum in (4.11) is attained. For more details about these matters see Chap. 3 by Lewin et al.

4.3 The Strictly Correlated Electrons (SCE) Functional

4.3.1 Constrained-Search Definition

From the early days of DFT it has been clear that a useful approximation to the kinetic energy contribution in (4.11) is given by the functional

$$T_{S,LL}[\rho] = \min_{\Psi \in \mathcal{W}^N, \Psi \mapsto \rho} \langle \Psi | T | \Psi \rangle \quad (4.13)$$

and by its further approximation $T_S[\rho]$ obtained by Kohn and Sham [85] via restricting the above search to Slater determinants built from orthonormal spin orbitals,

$$T_S[\rho] = \min \left\{ \sum_{i=1}^N \sum_{s \in \mathbb{Z}_2} \int_{\mathbb{R}^d} \frac{1}{2} |\nabla \varphi_i(\mathbf{r}, s)|^2 : \varphi_i \in H^1(\mathbb{R}^d \times \mathbb{Z}_2; \mathbb{C}) \forall i, \right. \\ \left. \langle \varphi_i | \varphi_j \rangle = \delta_{ij} \forall i, j, \sum_{i=1}^N \sum_{s \in \mathbb{Z}_2} |\varphi_i(\mathbf{r}, s)|^2 = \rho(\mathbf{r}) \forall \mathbf{r} \right\}. \quad (4.14)$$

The natural analogue of $T_{s,LL}$ for the interaction energy contribution in (4.11) is the SCE functional

$$V_{ee}^{\text{SCE}}[\rho] = \inf_{\Psi \in \mathcal{W}^N, \Psi \mapsto \rho} \langle \Psi | V_{ee} | \Psi \rangle, \quad (4.15)$$

which was introduced by Seidl [122]. The acronym SCE stands for *strictly correlated electrons*, and will be explained shortly. As detailed in the next section, the functional (4.15) is a rigorous leading-order asymptotic limit of $F_{LL}[\rho]$ in the low-

density regime, where interaction dominates, just as the kinetic functional (4.13) is a leading-order asymptotic limit at high density, where the kinetic energy dominates.

What is more, there also exists a natural analogue to T_S for interaction, which approximates, in the case of N electrons in the physical space \mathbb{R}^3 , the high-dimensional minimization over wavefunctions on $3N$ dimensional space in (4.15) by a minimization over just N maps on \mathbb{R}^3 ; see Sect. 4.3.4.

4.3.2 Derivation as a Low-Density or Strong-Interaction Limit of the Levy–Lieb Functional

For any given N -particle density ρ on \mathbb{R}^d , consider its dilation obtained by uniform coordinate scaling

$$\rho^\gamma(\mathbf{r}) = \gamma^d \rho(\gamma\mathbf{r}),$$

where $\gamma > 0$ is a scaling factor. Note that this scaling preserves the total particle number,

$$\int_{\mathbb{R}^d} \rho^\gamma(\mathbf{r}) = \int_{\mathbb{R}^d} \rho(\mathbf{r}) = N.$$

We are interested in the small- γ regime, which corresponds to a low-density limit.

If Ψ is a wavefunction with density ρ , then the scaled wavefunction

$$\Psi^\gamma(\mathbf{r}_1, s_1, \dots, \mathbf{r}_N, s_N) = \gamma^{\frac{dN}{2}} \Psi(\gamma\mathbf{r}_1, s_1, \dots, \gamma\mathbf{r}_N, s_N)$$

has density ρ^γ . But as first noticed by Levy and Perdew [92], scaling does not commute with constrained search. Instead, by an elementary change of variables,

$$T[\Psi^\gamma] = \gamma^2 T[\Psi], \quad V_{\text{ee}}[\Psi^\gamma] = \gamma V_{\text{ee}}[\Psi],$$

and therefore

$$\begin{aligned} F_{\text{LL}}[\rho^\gamma] &= \min_{\Psi^\gamma \in \mathcal{W}_N, \Psi^\gamma \mapsto \rho^\gamma} \langle \Psi^\gamma | T + V_{\text{ee}} | \Psi^\gamma \rangle \\ &= \gamma \min_{\Psi \in \mathcal{W}_N, \Psi \mapsto \rho} \langle \Psi | \gamma T + V_{\text{ee}} | \Psi \rangle \\ &= \gamma^2 F^{1/\gamma}[\rho], \end{aligned} \tag{4.16}$$

where

$$F^\lambda[\rho] = \min_{\Psi \in \mathcal{W}_N, \Psi \mapsto \rho} \left(T[\Psi] + \lambda V_{\text{ee}}[\Psi] \right) \quad (4.17)$$

is a Levy–Lieb functional with coupling constant λ . This suggests, assuming that the minimization in the second line of (4.16) commutes with taking the limit $\gamma \rightarrow 0$,

$$F_{\text{LL}}[\rho^\gamma] \underset{\gamma \rightarrow 0}{\sim} \gamma V_{\text{ee}}^{\text{SCE}}[\rho] \quad (4.18)$$

or equivalently, by starting from the Levy–Lieb functional with coupling constant, Eq. (4.17), as done in [122, 127]

$$\lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} F^\lambda[\rho] = V_{\text{ee}}^{\text{SCE}}[\rho]. \quad (4.19)$$

Mathematically, as pointed out in [127] it is not obvious whether the minimization in the second line of (4.16) commutes with passing to the limit $\gamma \rightarrow 0$ since the optimal wavefunction depends on γ . Nevertheless the above leading-order asymptotics can be rigorously justified; see Theorem 4.3 in the next section.

Repeating the calculation in (4.16) without the kinetic energy and replacing “min” by “inf” shows that

$$V_{\text{ee}}^{\text{SCE}}[\rho^\gamma] = \gamma V_{\text{ee}}^{\text{SCE}}[\rho], \quad (4.20)$$

whence the asymptotic result (4.18) can also be re-written as

$$F_{\text{LL}}[\rho^\gamma] \underset{\gamma \rightarrow 0}{\sim} V_{\text{ee}}^{\text{SCE}}[\rho^\gamma]. \quad (4.21)$$

Off the low-density limit, we remark that $V_{\text{ee}}^{\text{SCE}}$ still provides a rigorous lower bound for the Levy–Lieb functional,

$$F_{\text{LL}}[\rho] \geq V_{\text{ee}}^{\text{SCE}}[\rho] \quad \forall \rho \in \mathcal{D}^N. \quad (4.22)$$

This is a trivial consequence of the constrained-search Definitions (4.11) and (4.15) and the nonnegativity of the kinetic energy functional T . For typical atomic densities on \mathbb{R}^3 , this lower bound is a significant improvement over the Lieb–Oxford bound with best known constant.

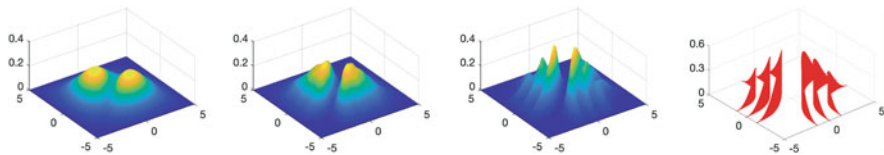


Fig. 4.1 Numerically computed ground state wave functions for $F^\lambda[\rho]$ in (4.17) for $N = 4$ and one-body density $\rho(r) = \frac{1}{2L}(1 + \cos(\frac{\pi}{L}r))$, $r \in [-L, L]$ ($L = 5$) for different values of λ : $\lambda = 0.1, 1, 10, \infty$. Shown: pair density $\sum_{s_1, s_2, s_3, s_4} \int dr_3 dr_4 |\Psi^\lambda(r_1, s_1, r_2, s_2, r_3, s_3, r_4, s_4)|^2$. Picture from [21], see also [109] for a numerical approximation of (4.17) with $N = 2$. The pair density on the left is governed by exchange effects, whereas the one on the right is governed purely by Coulombic correlations

4.3.3 Enlarging the Constrained Search to Probability Measures

The variational principle underlying the definition of $V_{ee}^{SCE}[\rho]$ in (4.15),

$$\begin{aligned} \text{Minimize } \langle \Psi | V_{ee} | \Psi \rangle &= \int_{\mathbb{R}^{dN}} V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N) \pi_N^\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) d\mathbf{r}_1 \dots d\mathbf{r}_N \\ &\text{over } \{ \Psi \in \mathcal{W}^N : \Psi \mapsto \rho \}, \end{aligned} \tag{4.23}$$

with N -point density π_N^Ψ as in (4.8), typically has no minimizer. That is, no minimizing wavefunction $\Psi \in \mathcal{W}^N$ exists and the infimum in (4.15) is not attained.² Physically, this reflects the phenomenon that if $\Psi_\lambda[\rho]$ is a sequence of square-integrable functions depending on a parameter $\lambda > 0$ such that $\langle \Psi_\lambda | V_{ee} | \Psi_\lambda \rangle$ approaches the infimum in (4.15) as λ tends to infinity—prototypical is the Ψ_λ that minimizes $\langle \Psi | T + \lambda V_{ee} | \Psi \rangle$ subject to $\Psi \mapsto \rho$ —then $|\Psi_\lambda|^2$ integrates to 1 but is typically concentrating on a lower dimensional subset, as depicted in Fig. 4.1.

This basic shortcoming of (4.23)—that wavefunctions which are closer and closer to being optimal in the constrained search (4.15) do not converge to any proper wavefunction—can be overcome as follows [14, 32]. First, interpret the variational principle (4.23) as a variational principle for the N -point density as suggested by the second expression in (4.23); second, enlarge the space of admissible N -point densities to the space $\mathcal{P}(\mathbb{R}^{dN})$ of *probability measures* on \mathbb{R}^{dN} with density ρ . Then the constrained search becomes well-posed, that is, optimizers exist. See Theorem 4.1 below. This enlargement allows N -point densities to concentrate on lower dimensional subsets as in Fig. 4.1. The condition that a probability measure $\Pi \in \mathcal{P}(\mathbb{R}^{dN})$ has density ρ now means that Π has marginals equal to the density divided by the particle number, $\frac{\rho}{N}$:

²This is not cured by dropping the requirement in (4.7) that Ψ must have square-integrable gradient and requiring mere square-integrability, i.e. replacing $\bigwedge_{i=1}^N H^1(\mathbb{R}^d \times \mathbb{Z}_2; \mathbb{C})$ by $\bigwedge_{i=1}^N L^2(\mathbb{R}^d \times \mathbb{Z}_2; \mathbb{C})$.

$$\int_{(\mathbb{R}^d)^{j-1} \times A_j \times (\mathbb{R}^d)^{N-j}} d\Pi = \int_{A_j} \frac{\rho}{N} \quad \text{for all } j \in \{1, \dots, N\} \text{ and all open sets } A_j \text{ in } \mathbb{R}^d. \quad (4.24)$$

We denote the relation given by Eq. (4.24) by $\Pi \mapsto \rho$. This yields the variational principle

$$\text{Minimize } \int_{\mathbb{R}^{dN}} V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad \text{over } \{\Pi \in \mathcal{P}(\mathbb{R}^{dN}) : \Pi \mapsto \rho\} \quad (4.25)$$

and the following enlarged-constrained-search definition of the SCE functional

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \min_{\Pi \in \mathcal{P}(\mathbb{R}^{dN}), \Pi \mapsto \rho} \int_{\mathbb{R}^{dN}} V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N). \quad (4.26)$$

This alternative definition of $V_{\text{ee}}^{\text{SCE}}[\rho]$ and the underlying enlarged variational problem (4.25) were introduced by Buttazzo et al. [14] and Cotar et al. [32], along with the insight that minimizers now exist (see Theorem 4.1 (1) below) and (4.26) is mathematically an *optimal transport problem* and can be usefully analyzed with methods from optimal transport theory (see Sect. 4.3.6). We call (4.25) the SIL variational principle, the acronym SIL standing for strong-interaction limit.

The notation in (4.26) (“min” instead of “inf”; using the same notation for the ensuing density functional even though a priori the right-hand side of (4.26) could be lower than that in (4.15) since the minimization is over a larger set) is justified because of:

Theorem 4.1 *Let ρ be any N -particle density in the class \mathcal{D}^N (see (4.12)), and let $w(\mathbf{r}) = |\mathbf{r}|^{-1}$ be the Coulomb interaction.*

- (1) *The minimum in (4.26) is attained; that is, there exists a minimizing probability measure Π .*
- (2) *[10, 32, 33] The minimum value in (4.26) is equal to the infimum in (4.15).*

Statement (1) is a special case of general existence theorems in optimal transport theory. For a textbook account see [50]. Proofs of such results rely on Prokhorov’s theorem from probability theory as well as on approximation and lower semi-continuity results for functionals of the form $\Pi \mapsto \int V_{\text{ee}} d\Pi$.

Statement (2), although plausible, is mathematically much more subtle. It rests on the nontrivial result that arbitrary symmetric probability measures $\Pi \in \mathcal{P}(\mathbb{R}^{dN})$ with marginal ρ can be approximated by N -point densities of quantum wavefunctions $\Psi \in \mathcal{W}^N$ with the *same* marginal. Note that such wavefunctions must be antisymmetric and must have a square-integrable gradient; but applying standard smoothing techniques from mathematics—such as mollification—to a given probability measure with marginal ρ does not preserve the marginal, nor does it yield the N -point density of an antisymmetric function. This result, and the

ensuing statement (2), was first proved for $N = 2$ [32], and later extended to $N = 3$ [10] and general N [33] (see also [95] for a similar extension to general N allowing mixed states).

Remark 4.2 (Symmetrization) The minimum value in (4.26) is unchanged, and still attained, when the minimization over arbitrary probability measures with marginal ρ/N , $\{\Pi \in \mathcal{P}(\mathbb{R}^{dN}) : \Pi \mapsto \rho\}$, is restricted to *symmetric* probability measures with marginal ρ/N , where a probability measure $\Pi \in \mathcal{P}(\mathbb{R}^{dN})$ is said to be symmetric if

$$\int_{A_1 \times \dots \times A_N} d\Pi = \int_{A_{\sigma(1)} \times \dots \times A_{\sigma(N)}} d\Pi \text{ for all open sets } A_1, \dots, A_N \text{ in } \mathbb{R}^d$$

and all permutations σ .

This is because whenever Π is a probability measure in $\mathcal{P}(\mathbb{R}^{dN})$ with marginals ρ , Eq. (4.24), then so is its symmetrization $S_N \Pi$ defined by

$$(S_N \Pi)(A_1 \times \dots \times A_N) = \frac{1}{N!} \sum_{\sigma} \Pi(A_{\sigma(1)} \times \dots \times A_{\sigma(N)}), \quad (4.27)$$

the sum being over all permutations of $\{1, \dots, N\}$; and the integral on the r.h.s. of (4.26) for Π agrees with that for $S_N \Pi$, thanks to the permutation symmetry of V_{ee} .

Next we rigorously justify the asymptotic relations (4.18), (4.19), (4.21) and complement them with an asymptotic result on the associated constrained-search wavefunctions.

Theorem 4.3 ([33]) *For any N -electron density ρ in the class \mathcal{D}^N (see (4.12)), and with $w(\mathbf{r}) = |\mathbf{r}|^{-1}$ being the Coulomb interaction, the asymptotic results (4.18), (4.19), (4.21) hold. Moreover if $\Psi_{\lambda}[\rho]$ is any minimizer in the constrained-search definition of $F^{\lambda}[\rho]$ (see (4.17)), then every limit point³ Π of the sequence of N -point densities $\pi^{\Psi_{\lambda}[\rho]}$ is a minimizer in the enlarged-search Definition (4.26) of $V_{\text{ee}}^{\text{SCE}}[\rho]$.*

³ By a limit point Π of a sequence Π_{λ} of probability measures we mean a limit point in the sense of narrow convergence, that is, convergence of the integrals $\int f d\Pi_{\lambda}$ to $\int f d\Pi$ for any bounded continuous function f .

Proof of (4.18), (4.19), and (4.21) The proof, taken from [33], is easy, so we include it. We show (4.19), the other statements being equivalent. Fix ρ . First, pick any minimizer $\Psi^\lambda[\rho]$ in the constrained-search definition of $F^\lambda[\rho]$, then

$$\frac{1}{\lambda} F^\lambda[\rho] = \frac{1}{\lambda} \left(T[\Psi^\lambda[\rho]] + \lambda V_{\text{ee}}[\Psi^\lambda[\rho]] \right) \geq V_{\text{ee}}[\Psi^\lambda[\rho]] \geq V_{\text{ee}}^{\text{SCE}}[\rho], \quad (4.28)$$

that is, the SCE functional is a lower bound of the left-hand side. To show that it is also an asymptotic upper bound for large λ , we fix any positive number ϵ and pick a wavefunction $\tilde{\Psi}[\rho]$ in \mathcal{W}^N such that $V_{\text{ee}}[\tilde{\Psi}[\rho]] \leq V_{\text{ee}}^{\text{SCE}}[\rho] + \epsilon$. It follows that

$$\frac{1}{\lambda} F^\lambda[\rho] \leq \frac{1}{\lambda} \left(T[\tilde{\Psi}[\rho]] + \lambda V_{\text{ee}}[\tilde{\Psi}[\rho]] \right).$$

Since $\tilde{\Psi}$ belongs to \mathcal{W}^N , its kinetic energy $T[\tilde{\Psi}]$ is finite, and so

$$\limsup_{\lambda \rightarrow \infty} \frac{1}{\lambda} F^\lambda[\rho] \leq V_{\text{ee}}[\tilde{\Psi}[\rho]] \leq V_{\text{ee}}^{\text{SCE}}[\rho] + \epsilon.$$

Since $\epsilon > 0$ was arbitrary,

$$\limsup_{\lambda \rightarrow \infty} \frac{1}{\lambda} F^\lambda[\rho] \leq V_{\text{ee}}^{\text{SCE}}[\rho]. \quad (4.29)$$

Combining (4.28) and (4.29) yields (4.19). \square

The above simple argument only shows that the asymptotic error in (4.19) is $o(1/\lambda)$, but does not give its order, which turns out to be $O(1/\lambda^{1/2})$, see Sect. 4.3.5.

4.3.4 The SCE Ansatz

The SIL variational principle (4.25) still requires minimization over a high-dimensional space of N -point probability measures.

Seidl [122] (see also [127]) proposed the following low-dimensional ansatz: we restrict minimization over N -point probability measures to minimization over singular probability measures of the special form

$$d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{\rho(\mathbf{r}_1)}{N} \prod_{n=2}^N \delta(\mathbf{r}_n - f_{n-1}(\mathbf{r}_1)) d\mathbf{r}_1 \dots d\mathbf{r}_N \quad (4.30)$$

where, for any $\mathbf{r}_1 \in \mathbb{R}^d$, $\delta(\mathbf{r}_n - f_{n-1}(\mathbf{r}_1))$ denotes the delta function of \mathbf{r}_n (alias Dirac measure) centered at $f_{n-1}(\mathbf{r}_1)$, and f_1, \dots, f_{N-1} are maps from \mathbb{R}^d to \mathbb{R}^d . The singular densities (4.30) are concentrated on the d -dimensional set

$$\Omega_0 = \{(\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{dN} : \mathbf{r}_2 = f_1(\mathbf{r}_1), \dots, \mathbf{r}_N = f_{N-1}(\mathbf{r}_1)\}. \quad (4.31)$$

From a physical point of view, such a density describes a state in which the position of one of the electrons, say \mathbf{r}_1 , can be freely chosen according to the density ρ , but this then uniquely fixes the position of all the other electrons through the functions f_2, \dots, f_N , that is, $\mathbf{r}_2 = f_1(\mathbf{r}_1)$ etc. Thus states of form (4.30) are called *strictly correlated states*, or SCE states for short. The f_i are called *co-motion functions* or *transport maps*.

The marginal constraint that Π must have marginals ρ , Eq. (4.24), turns into the following constraint on the maps f_n : the f_n must transport the density ρ to itself,

$$f_{i\#}\rho = \rho \quad \forall i \in \{2, \dots, N\} \quad (4.32)$$

where, for any measurable map $f : \mathbb{R}^p \rightarrow \mathbb{R}^q$ and any measure μ on \mathbb{R}^p , the push-forward $f_{\#}\mu$ is the measure on \mathbb{R}^q defined by

$$(f_{\#}\mu)(B) = \mu(f^{-1}(B)) \text{ for all open sets } B \text{ in } \mathbb{R}^q. \quad (4.33)$$

More explicitly, if $p = q$, μ is absolutely continuous with density ρ , f is a diffeomorphism, and the density of the push-forward $f_{\#}\mu$ is denoted by $f_{\#}\rho$, we have

$$(f_{\#}\rho)(\mathbf{r}') = |\det Df^{-1}(\mathbf{r}')| \rho(f^{-1}(\mathbf{r}')).$$

By substituting this formula for the push-forward into (4.32) and changing variables $f^{-1}(\mathbf{r}') = \mathbf{r}$, the constraint (4.32) turns—provided the f_n are diffeomorphisms—into the following nonlinear first-order partial differential equation:

$$\rho(f_i(\mathbf{r})) = \frac{\rho(\mathbf{r})}{|\det Df_i(\mathbf{r})|} \quad \forall i \in \{2, \dots, N\}.$$

Plugging the ansatz (4.30) into the SIL variational principle (4.25) and integrating out the variables $\mathbf{r}_2, \dots, \mathbf{r}_N$ yields the SCE variational principle

$$\begin{aligned} \text{Minimize } \int_{\mathbb{R}^d} V_{\text{ee}}(\mathbf{r}_1, f_1(\mathbf{r}_1), \dots, f_{N-1}(\mathbf{r}_1)) \frac{\rho(\mathbf{r}_1)}{N} d\mathbf{r}_1 \\ \text{over maps } f_1, \dots, f_{N-1} \in \mathcal{T}_\rho, \end{aligned} \quad (4.34)$$

with the minimization being over maps in the admissible class

$$\mathcal{T}_\rho = \{f : \mathbb{R}^d \rightarrow \mathbb{R}^d : f \text{ measurable, } f_{\#}\rho = \rho\}. \quad (4.35)$$

Thanks to Theorem 4.4 (1) below, this yields a third construction of the SCE functional,

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \inf_{f_1, \dots, f_{N-1} \in \mathcal{T}_\rho} \int_{\mathbb{R}^d} V_{\text{ee}}(\mathbf{r}, f_1(\mathbf{r}), \dots, f_{N-1}(\mathbf{r})) \frac{\rho(\mathbf{r})}{N} \, d\mathbf{r}. \quad (4.36)$$

In the Coulomb case, (4.4), and denoting $f_0(\mathbf{r}) = \mathbf{r}$, we thus have

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \inf_{f_1, \dots, f_{N-1} \in \mathcal{T}_\rho} \sum_{0 \leq i < j \leq N-1} \int_{\mathbb{R}^d} \frac{1}{|f_i(\mathbf{r}) - f_j(\mathbf{r})|} \frac{\rho(\mathbf{r})}{N} \, d\mathbf{r}. \quad (4.37)$$

Physically, this means that one needs to minimize the mutual Coulomb repulsion of the co-motion functions. This construction of the SCE functional was introduced by Seidl [122]. A priori it is not clear, but was conjectured by Seidl, that it is equivalent to the original construction (4.15). This is now rigorously known (see Corollary 4.6 below).

The construction (4.36) should be considered the analogue for interaction of the classical Kohn–Sham kinetic energy functional T_S . Just as T_S is determined by N low-dimensional functions (the Kohn–Sham spin orbitals $\varphi_1, \dots, \varphi_N : \mathbb{R}^3 \times \mathbb{Z}_2 \rightarrow \mathbb{C}$), $V_{\text{ee}}^{\text{SCE}}$ is determined by $N - 1$ low-dimensional maps (the co-motion functions or transport maps $f_1, \dots, f_{N-1} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$), which can be easily stored on a computer. Moreover—like the Kohn–Sham orbitals—the co-motion functions are obtained by just minimizing a 3-dimensional integral.

The reader is warned, however, that the behavior of the SCE variational principle and its relationship to the SIL variational principle is subtle, and open questions remain. In particular, it is not known—except in special cases—whether minimizers in (4.37) exist. The following results have been rigorously proved.

Theorem 4.4 *Let $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ be any N -particle density in the class \mathcal{D}^N (see (4.12)), and let $w(\mathbf{r}) = |\mathbf{r}|^{-1}$ be the Coulomb interaction.*

- (1) *The infimum in (4.36) is equal to the minimum in (4.26).*
- (2) *For two electrons ($N = 2$), and in arbitrary space dimension d , the infimum in (4.36) is attained; that is, there exists a minimizing map f_1 . Moreover f_1 is unique, and the induced probability measure (4.30) is the unique minimizer of the SIL variational principle (4.25).*
- (3) *In one space dimension ($d = 1$), and for arbitrary N , the infimum in (4.36) is attained; that is, there exist minimizing maps f_1, \dots, f_{N-1} . Moreover the symmetrization (see Remark 4.2) of the associated probability measure (4.30) is the unique symmetric minimizer of the SIL variational principle (4.25).*

Statement (1) is a consequence, pointed out in [28], of a general theorem of Ambrosio [5] and Pratelli [116] in optimal transport theory. For $N = 2$ or $d = 1$, use of the Ambrosio–Pratelli theorem can be avoided since the assertion follows from (2) respectively (3).

The existence of optimal maps in (2) and (3) is subtle and depends on special Coulombic features. For non-Coulombic counterexamples see Remark 4.7 below. In the Coulomb case, it is an open question whether the infimum in (4.36) is attained for general (physically reasonable) densities ρ when $d > 1$ and $N \geq 3$.

Statement (2) completely justifies Seidl’s SCE ansatz for $N = 2$: the SCE problem

$$\text{Minimize } \int_{\mathbb{R}^d} \frac{1}{|\mathbf{r} - f_1(\mathbf{r})|} \text{ over maps } f_1 \in \mathcal{T}_\rho$$

has a unique minimizer and the associated SCE state

$$d\Pi(\mathbf{r}_1, \mathbf{r}_2) = \frac{\rho(\mathbf{r}_1)}{2} \delta(\mathbf{r}_2 - f_1(\mathbf{r}_1)) d\mathbf{r}_1 d\mathbf{r}_2 \quad (4.38)$$

is the unique minimizer of the SIL problem

$$\text{Minimize } \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} d\Pi(\mathbf{r}_1, \mathbf{r}_2) \text{ over } \Pi \mapsto \rho.$$

This was proved in [32], by modifying the analysis by Gangbo and McCann [55] of optimal transport with costs $w(\mathbf{r}, \mathbf{r}')$ which are convex or concave in the displacement $\mathbf{z} = \mathbf{r} - \mathbf{r}'$. Note that the Coulomb cost is neither: near any displacement $\mathbf{z}_0 \neq 0$, it is convex in radial direction and concave in all perpendicular directions. A simpler proof using Kantorovich duality (see Sect. 4.3.7) was suggested in [14], and made rigorous in [42]. The SCE map is given by

$$f_1(\mathbf{r}) = \mathbf{r} + \frac{\nabla u(\mathbf{r})}{|\nabla u(\mathbf{r})|^{3/2}}, \quad (4.39)$$

for some function $u : \mathbb{R}^d \rightarrow \mathbb{R}$ (Kantorovich potential). The notion of Kantorovich potential will be explained in Sect. 4.3.7. Equation (4.39) follows by solving Eq. (4.59) for f_1 .

Statement (3), together with an explicit construction of the optimal maps given in Sect. 4.3.11, was suggested in the original paper by Seidl [122] on grounds of physical arguments, and was rigorously proved in [27] with the help of cyclical monotonicity methods from optimal transport theory. See Sect. 4.3.11 for more information.

The uniqueness statements in (2) and (3) are somewhat surprising: the optimal N -point densities arising from Levy–Lieb constrained search in the strongly interacting

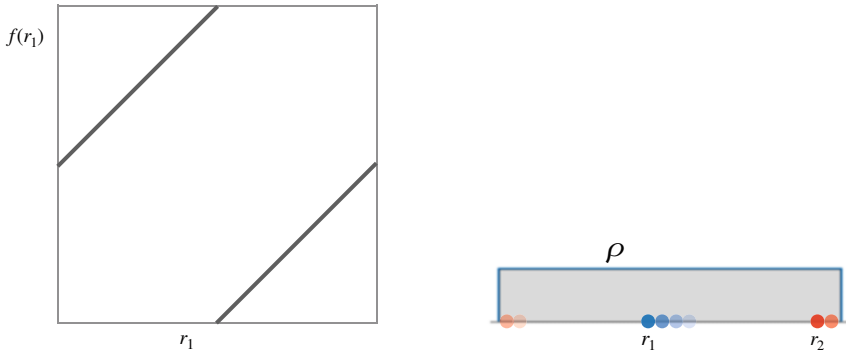


Fig. 4.2 SCE state of a two-electron system with homogeneous density in a one-dimensional domain. By Theorems 4.3 and 4.4, this state is an asymptotically exact approximation to the true quantum ground state at low density. Left: optimal co-motion function or transport map f . Right: position of the two electrons in the one-dimensional domain. The position of the second electron, r_2 , is determined by that of the first electron, r_1 , through the equation $r_2 = f(r_1)$, with the optimal f keeping the electrons at a constant distance, of half the domain size. The position of the first electron varies over the whole domain according to the density ρ (see Eq. (4.38)). As the first electron (depicted in blue) passes through the mid-point, the position of the second electron (depicted in red) jumps from the right end to the left end, causing a discontinuity of f

limit are always *unique* when either $N = 2$ or $d = 1$! No analogue holds off the strongly interacting limit.

Example 4.5 Consider a two-electron system with uniform density in a one-dimensional interval $[0, L]$. The unique minimizer $f_1 = f$ of the SCE variational principle (4.34) can be shown (see Sect. 4.3.11) to be

$$f(r_1) = \begin{cases} r_1 + \frac{L}{2} & \text{if } r_1 \leq \frac{L}{2} \\ r_1 - \frac{L}{2} & \text{if } r_1 > \frac{L}{2}. \end{cases} \tag{4.40}$$

See Fig. 4.2.

By combining Theorems 4.4 and 4.1 we obtain:

Corollary 4.6 *Let $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ be any N -particle density in the class \mathcal{D}^N (see (4.12)), and let $w(\mathbf{r}) = |\mathbf{r}|^{-1}$ be the Coulomb interaction. Then the infimum in (4.36) is equal to that in (4.15).*

Proof This follows from the fact that both quantities are equal to the minimum value of the SIL variational principle (4.25), by Theorem 4.1 (2) respectively Theorem 4.4 (1). □

We remark that no proof is known which bypasses the SIL variational principle, even though the corollary was conjectured before the latter was introduced.

We close this introductory section on the SCE ansatz with some remarks.

Remark 4.7 (Nonattainment) For simple non-Coulombic counterexamples to attainment of the infimum in (4.36) for $N = 3$ even in one space dimension, see [49, 59]. For instance, one can take the uniform density in the interval $[0, 3]$ and the interaction potential $w(r) = r^4/4 - r^3/3$ [49]. Earlier more intricate counterexamples can be found in [107]. Such a nonattainment has the undesirable consequence that numerically computed optimal maps will necessarily exhibit wilder and wilder oscillations as the mesh is refined or the basis set approaches completeness, and fail to converge in any pointwise sense to actual optimal maps.

Remark 4.8 (Existence of Non-SCE Minimizers) For a Coulombic example for $N = 3$ in three space dimensions showing that the SIL variational principle can possess minimizers which are not of SCE form, see [113]. This example exhibits nonuniqueness and it is not known whether it also admits minimizers which are of SCE form.

Remark 4.9 (Alternative Formulations of the SCE Ansatz) Denoting $f_0(\mathbf{r}) = \mathbf{r}$, one can write the SCE ansatz (4.30) in the following form in which all coordinates $\mathbf{r}_1, \dots, \mathbf{r}_N$ appear on an equal footing:

$$d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \int \frac{\rho(\mathbf{r})}{N} \prod_{n=1}^N \delta(\mathbf{r}_n - f_{n-1}(\mathbf{r})) d\mathbf{r}. \quad (4.41)$$

Also, one can work with the symmetrized form of this ansatz,

$$d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{N!} \sum_{\sigma} \int \frac{\rho(\mathbf{r})}{N} \prod_{n=1}^N \delta(\mathbf{r}_n - f_{\sigma(n-1)}(\mathbf{r})) d\mathbf{r} \quad (4.42)$$

(where σ runs over the permutations of the SCE map indices $0, \dots, N - 1$); the symmetrization doesn't change the energy $\int V_{\text{ee}} d\Pi$, and the symmetrized form (4.42) minimizes the SIL problem (4.26) if and only if the unsymmetrized form (4.41) does, as was explained in Remark 4.2.

Remark 4.10 (Nonsmoothness of Optimal Maps) The reader might wonder why, in SCE theory, no differentiability and not even continuity is imposed on the competing maps (the maps in the admissible class (4.35) are merely required to be measurable). This is because optimal maps, when they exist, are typically discontinuous. This important effect can be understood intuitively from simple examples as in Fig. 4.2. As the first electron passes through the midpoint of the domain, the position of the second electron jumps from the right end of the domain to the left end, yielding the discontinuous map depicted in the Figure. For a radial density in three dimensions ($d = 3$), an analogous discontinuity occurs in that spheres near zero are mapped to spheres near infinity [32]. For general densities and general N , the presence of discontinuities across unknown surfaces makes Eq. (4.36) very challenging for numerical computations.

4.3.5 The Next Leading Term

So far we have treated the limit of the Levy–Lieb functional at infinite coupling strength λ (or, equivalently, at extreme low density). One could ask how this limit is approached, or, in other words, what is the next leading term in Eqs. (4.18)–(4.19).

The strategy employed in [68] to compute this next leading term relies on the assumption that the minimizer in (4.25) is of the SCE or Monge type, see the detailed discussion in the previous Sect. 4.3.4. Under this assumption, as shown in Sect. 4.3.8, the classical potential energy

$$E_{\text{pot}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_{i=1}^N v_{\text{SCE}}(\mathbf{r}_i), \quad (4.43)$$

with $v_{\text{SCE}}(\mathbf{r})$ defined by Eqs. (4.59) and (4.64), attains its minimum on the manifold Ω_0 parametrized by the co-motion functions,

$$\Omega_0 = \{(\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{dN} : \mathbf{r}_1 = \mathbf{r}, \mathbf{r}_2 = \mathbf{f}_2(\mathbf{r}), \dots, \mathbf{r}_N = \mathbf{f}_N(\mathbf{r})\}. \quad (4.44)$$

When λ in Eq. (4.17) is very large but finite, we can expect that the support of the minimizer in Eq. (4.17) be strongly localized around Ω_0 , as illustrated by Fig. 4.1 in Sect. 4.3.3. We can then expand E_{pot} around its minimum through second order. The corresponding Hessian matrix $\mathbb{H}(\mathbf{r})$ evaluated on Ω_0 , for any fixed \mathbf{r} , will have d zero eigenvalues (along the manifold Ω_0) and $dN - d$ positive eigenvalues. By using curvilinear coordinates along the manifold Ω_0 and orthogonal to it, the sought next leading term is determined by adding the kinetic energy to the second-order expansion of E_{pot} , which corresponds to the Hamiltonian of zero-point oscillations in the space orthogonal to Ω_0 [68]. The final result is that Eqs. (4.18)–(4.19) are extended to [66, 68]

$$F_{\text{LL}}[\rho^\gamma] \underset{\gamma \rightarrow 0}{\sim} \gamma V_{\text{ee}}^{\text{SCE}}[\rho] + \gamma^{3/2} F^{\text{ZPE}}[\rho] \quad (4.45)$$

$$F_\lambda \underset{\lambda \rightarrow \infty}{\sim} \lambda V_{\text{ee}}^{\text{SCE}}[\rho] + \sqrt{\lambda} F^{\text{ZPE}}[\rho], \quad (4.46)$$

where

$$F^{\text{ZPE}}[\rho] = \frac{1}{2} \int_{\mathbb{R}^d} \frac{\rho(\mathbf{r})}{N} \text{Tr} \left(\sqrt{\mathbb{H}(\mathbf{r})} \right). \quad (4.47)$$

In [71] this term has been computed explicitly for $N = 2$ electrons in 1d and it has been compared with accurate numerical calculations for the Levy–Lieb functional at very large λ , finding excellent agreement.

The intuition that the next term of the Levy–Lieb functional at infinite coupling strength λ should be given by zero-point oscillations around the manifold

parametrized by the co-motion functions appeared for the first time in Seidl's seminal work [122]. He also carried out explicit calculations in 3D for the spherically-symmetric case with $N = 2$ electrons, using the co-motion function introduced in Sect. 4.3.12. This idea was extended to the general many-electron case in [68], where it was also found that Seidl's original calculation had a wrong factor 2. Very recently, a rigorous proof of Eqs. (4.45)–(4.47) for the many-electron 1d case has been provided by Colombo et al. [29].

4.3.5.1 The Fermionic Statistics

Equations (4.45)–(4.47) are the first-order correction due to kinetic energy in the large- λ (or $\hbar \rightarrow 0$) limit of the Levy–Lieb functional. This correction is still independent of the particle statistics. A natural question to ask is then at which order will the fermionic antisymmetry enter.

In Refs. [67, 68] it has been conjectured that the particle statistics enters in the $\lambda \rightarrow \infty$ limit at orders $\sim e^{-\sqrt{\lambda}}$. The physical intuition behind this idea is simply that the effect on the energy of antisymmetrization vanishes as the overlap between gaussians centered at each set of strictly-correlated positions (each \mathbf{r} value in Ω_0). The scaling $\sqrt{\lambda}$ of such gaussians comes from the zero-point hamiltonian. This conjecture has been confirmed numerically [71] for the case of $N = 2$ electrons in 1D, again by comparison with accurate numerical calculations of the exact Levy–Lieb functional at large λ .

4.3.6 The Strongly Interacting Limit of DFT from the Point of View of Optimal Transport

We now introduce a fruitful interpretation of the strongly interacting limit of DFT as “optimal transport with Coulomb cost”.

Optimal transport theory (see [50, 117, 119, 136] for textbook accounts) is concerned with the following two problems, introduced in special cases in fundamental work by Kantorovich [80] respectively Monge [108]:

- (a) *Kantorovich optimal transport problem:* For given probability measures μ_1, \dots, μ_N defined on closed subsets X_1, \dots, X_N of \mathbb{R}^d , find a joint probability measure Π on the product space $X = X_1 \times \dots \times X_N \subseteq \mathbb{R}^{Nd}$ which minimizes a cost functional

$$C[\Pi] = \int_X c(\mathbf{r}_1, \dots, \mathbf{r}_N) d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

subject to the marginal constraints

$$\int_{X_1 \times \dots \times X_{i-1} \times A_i \times X_{i+1} \times \dots \times X_N} d\Pi = \int_{A_i} d\mu_i \text{ for all measurable sets } A_i \subseteq X_i$$

and all $i \in \{1, \dots, N\}$.

Here $c : X_1 \times \dots \times X_N \rightarrow \mathbb{R} \cup \{+\infty\}$ is some given cost function, and the validity of the above constraint is denoted $\Pi \mapsto \mu_1, \dots, \mu_N$.

- (b) *Monge optimal transport problem:* For given probability measures μ_1, \dots, μ_N defined on measurable subsets X_1, \dots, X_N of \mathbb{R}^d of positive volume which possess integrable densities p_1, \dots, p_N (i.e. $p_i \in L^1(X_i)$), and a cost function c as above, find measurable maps f_1, \dots, f_{N-1} with $f_i : X_1 \rightarrow X_{i+1}$ which minimize

$$I[f_1, \dots, f_{N-1}] = \int_{X_1} c(\mathbf{r}_1, f_1(\mathbf{r}_1), \dots, f_{N-1}(\mathbf{r}_N)) d\mu_1$$

subject to the marginal constraints

$$f_{i\#} p_1 = p_{i+1} \quad \text{for } i \in \{1, \dots, N-1\}.$$

This corresponds to making the ansatz

$$d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) = d\mu_1(\mathbf{r}_1) \delta(\mathbf{r}_2 - f_1(\mathbf{r}_1)) \cdots \delta(\mathbf{r}_N - f_{N-1}(\mathbf{r}_1)) d\mathbf{r}_2 \cdots d\mathbf{r}_N \tag{4.48}$$

or equivalently—using the notion of push-forward introduced in (4.33)—

$$\Pi = (\text{id}, f_1, \dots, f_N)\# \mu_1 \tag{4.49}$$

in the Kantorovich problem, where id denotes the identity map $\text{id}(\mathbf{r}_1) = \mathbf{r}_1$.

Example 4.11 (N Equal Marginals, Coulomb Cost) If we take

$$X_1 = \dots = X_N = \mathbb{R}^d, \mu_1 = \dots = \mu_N = \frac{\rho}{N}, c(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{1 \leq i < j \leq N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

the Kantorovich optimal transport problem is precisely the SIL variational problem, (4.25), and the Monge optimal transport problem is precisely the SCE variational problem, (4.34).

Thus the strongly interacting limit of DFT can be viewed as *optimal transport with Coulomb cost*. This viewpoint, introduced by Buttazzo et al. [14] and Cotar et al. [32], opened the door to much of the current understanding of the strong-interaction limit of DFT.

Example 4.12 (Two Unequal Marginals, Positive Power Cost) The prototype problem of classical optimal transport theory going back to [80, 108] is to instead take

$$N = 2, \quad X_1 = X_2 = \mathbb{R}^d, \quad c(\mathbf{r}_1, \mathbf{r}_2) = |\mathbf{r}_1 - \mathbf{r}_2|^p, \quad p \geq 1.$$

That is, one considers:

- only two marginals;
- unequal instead of equal marginals;
- a positive instead of a negative power of the Euclidean distance as cost.

Denoting $\mu_1 = \mu, \mu_2 = \nu, f_1 = T, \mathbf{r}_1 = x, \mathbf{r}_2 = y$, the Kantorovich problem then becomes

$$\begin{aligned} \text{Minimize } C[\Pi] &= \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\Pi \\ (x, y) \text{ over } \Pi &\in \mathcal{P}(\mathbb{R}^{2d}) \text{ subject to } \Pi \mapsto \mu, \nu \end{aligned} \quad (4.50)$$

and the Monge problem becomes

$$\begin{aligned} \text{Minimize } I[T] &= \int_{\mathbb{R}^d} |x - T(x)|^p d\mu(x) \text{ over measurable maps } T : \mathbb{R}^d \rightarrow \mathbb{R}^d \\ &\text{subject to } T_{\#}\mu = \nu. \end{aligned} \quad (4.51)$$

The analogue of the SCE functional is the optimal cost as a functional of the two prescribed marginals,

$$C_{\text{opt}}[\mu, \nu] = \min\{C[\Pi] : \Pi \mapsto \mu, \nu\} = \inf\{I[T] : T_{\#}\mu = \nu\}.$$

Its p -th root, $W_p(\mu, \nu) = (C_{\text{opt}}[\mu, \nu])^{1/p}$, is the celebrated p -Wasserstein distance, which is a metric on the space of probability measures.

Thus the SCE functional can be thought of as the *Coulomb analogue of the Wasserstein distance*.

We remark that the motivation of Monge and Kantorovich for considering Example 4.12 came from civil engineering, respectively economics, and explains the name *optimal transport*: Monge thought of moving a given pile of sand on a construction site into a given hole in a way that minimizes the overall distance of transport, with $T(x)$ describing the target position of sand originally located at x and with pile and hole modelled, respectively, by μ and ν . Kantorovich thought of transporting some economic good, say steel, from producers (steel mines) to consumers (factories), at minimal transportation cost; $\Pi(x, y)$ then describes the

density of goods transported from location x to location y , and is called a *transport plan*. In the latter context it is natural *not* to make the Monge ansatz

$$d\Pi(x, y) = d\mu(x)\delta(y - T(x))dy$$

but instead allow one producer located at x to supply several consumers located at different positions y , i.e. consider the general problem (4.50).

The general question for which costs and marginals the Monge and Kantorovich problems are equivalent, i.e. the Kantorovich problem admits minimizers of Monge form, is not well understood. A sufficient condition [119] for $N = 2$ (and, say, compact convex sets X_1 and X_2 and continuously differentiable costs c) is that the marginal measure μ_1 is absolutely continuous and c satisfies the so-called twist condition that the map $\mathbf{r}_2 \mapsto \nabla_{\mathbf{r}_1}c(\mathbf{r}_1, \mathbf{r}_2)$ be injective for every r_1 . For $N > 2$, generalized twist conditions have been studied by Pass [111, 114, 115]; unfortunately these are not satisfied for the Coulomb cost.

4.3.7 Dual Construction of the SCE Functional

We now introduce a fourth—dual—construction of the SCE functional.

A cornerstone principle of optimal transport theory, *Kantorovich duality*, says that the minimum of a given Kantorovich optimal transport problem (see Sect. 4.3.6) equals the supremum of an associated explicit dual problem. The general form of the dual is recalled in Appendix 4.6. For the SIL problem (4.25), the dual problem is the following (see Appendix 4.6 for a quick derivation from general OT theory): maximize the functional

$$J[u] = \sum_{i=1}^N \int_{\mathbb{R}^d} u(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} \quad (4.52)$$

over potentials $u : \mathbb{R}^d \rightarrow \mathbb{R}$ which must satisfy the pointwise constraint

$$\sum_{i=1}^N u(\mathbf{r}_i) \leq V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad \forall (\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{dN}. \quad (4.53)$$

Maximization is over the admissible class

$$\mathcal{A} = \{u : \mathbb{R}^d \rightarrow \mathbb{R} \mid u \text{ bounded and measurable, } u \text{ satisfies (4.53)}\}. \quad (4.54)$$

This yields the following alternative definition of the SCE functional:

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \sup_{u \in \mathcal{A}} \int_{\mathbb{R}^d} u(\mathbf{r}) \rho(\mathbf{r}) \, \text{d}\mathbf{r}. \tag{4.55}$$

This construction is due to Buttazzo, DePascale, and Gori-Giorgi [14]. Note that the optimization here is not over N -point densities, but over (suitable) external potentials u . Optimizers are called *Kantorovich potentials*. Heuristically, they can be thought of as Lagrange multipliers associated with the marginal constraints in the original problem (4.25). This is explained in our discussion of optimality conditions in Sect. 4.3.8.

It can be rigorously shown that the new construction yields, again, the SCE functional, and that optimal potentials exist:

Theorem 4.13 *Let $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ be any N -particle density in the class \mathcal{D}^N (see (4.12)), and let $w(\mathbf{r}) = |\mathbf{r}|^{-1}$ be the Coulomb interaction. Then:*

- (1) [14] *The supremum in (4.55) is equal to the minimum in (4.26).*
- (2) [14, 38] *The supremum in (4.55) is attained; that is, there exists a maximizing potential u in the class (4.54).*
- (3) [13, 42] *If, in addition, $\rho > 0$ everywhere, there exists a maximizing potential which is in addition Lipschitz continuous.*

Statement (1) follows directly from the general Kantorovich duality theorem of OT theory; see Appendix 4.6. The question of existence and regularity of optimal potentials is more delicate. Note that the Coulomb potential V_{ee} which upper-bounds $u(\mathbf{r}_1) + \dots + u(\mathbf{r}_N)$ tends to plus infinity as the distance $\mathbf{r}_i - \mathbf{r}_j$ between any two position coordinates goes to zero; so one might a priori think that u 's are favourable which also tend to plus infinity at certain places. But statement (2) in the above theorem says that this does not happen; the existence proof of bounded optimal potentials is due to [14] for $N = 2$ and to [38] for general N .

Quantum Analogue We remark that the dual construction of the SCE functional in Eq. (4.55) admits a quantum analogue. In [97], Lieb proposed an extension of the Levy–Lieb functional (4.11) to mixed states, i.e. $F_L : \mathcal{D}_N \rightarrow \mathbb{R}$,

$$F_L[\rho] = \min \left\{ \text{Tr} \left(-\frac{1}{2} \sum_{j=1}^N \Delta_{\mathbf{r}_j} + V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) \right) \Gamma : \Gamma = \Gamma^* \geq 0, \text{Tr}(\Gamma) = 1, \Gamma \mapsto \rho \right\}, \tag{4.56}$$

where Γ is an operator acting on the fermionic Hilbert space and, similarly to (4.9), $\Gamma \mapsto \rho$ denotes the relation $\rho(\mathbf{r}) = N \int_{\mathbb{R}^{dN}} \Gamma(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N; \mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N) \, \text{d}\mathbf{r}_2 \dots \text{d}\mathbf{r}_N$. In [91], M. Levy introduced a similar functional requiring in addition that $\Gamma = |\psi\rangle\langle\psi|$ be a rank-one operator. An advantage of the Lieb functional F_L is that

it is convex. Ignoring issues of rigor, (4.56) admits a dual formulation

$$F_L[\rho] = \sup \left\{ \int_{\mathbb{R}^3} u(\mathbf{r})\rho(\mathbf{r}) \, d\mathbf{r} : \sum_{i=1}^N u(\mathbf{r}_i) \leq -\frac{1}{2} \sum_{j=1}^N \Delta \mathbf{r}_j + V_{\text{ee}} \right\}, \quad (4.57)$$

with the above inequality understood in the sense of self-adjoint operators. For a rigorous discussion of Eq. (4.57) see Chap. 3 by Lewin et al. This equation is the quantum analogue (for mixed states) of the dual construction of the SCE functional. Note that because the right-hand side of the constraint on u now contains an additional positive term, the value of the supremum will be higher than in (4.55), as it should be.

4.3.8 Optimality Conditions

With the help of Kantorovich duality one obtains very interesting necessary conditions for solutions to the SIL variational principle (4.25). In particular, for optimizers of SCE (alias Monge) form one can express the gradient of the Kantorovich potential u in terms of the co-motion functions (alias transport maps).

We follow the rigorous presentation for general OT problems in [50], but specialize throughout to the SIL problem. For the benefit of less mathematically minded readers, we also include a heuristic derivation at the end of this section.

Theorem 4.14 (Optimality Conditions [50]) *Let $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ be any N -particle density in the class \mathcal{D}^N (see (4.12)). Let $V_{\text{ee}} : \mathbb{R}^{dN} \rightarrow \mathbb{R} \cup \{+\infty\}$ be any interaction potential which is symmetric, bounded from below, lower semi-continuous, and has the property that the minimum in (4.26) is finite. Suppose Π is a solution to the SIL problem (4.26), and u is a solution to the dual problem, i.e. a maximizer of J in the class \mathcal{A} .*

(1) Π is zero outside the set

$$\mathcal{M} = \{(\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{dN} : V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_{i=1}^N u(\mathbf{r}_i) = \min\}.$$

(2) Π is an unconstrained minimizer (i.e., a minimizer on $\mathcal{P}(\mathbb{R}^{dN})$) of the modified functional

$$\mathcal{L}[\Pi] = \int_{\mathbb{R}^{dN}} \left(V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_{i=1}^N u(\mathbf{r}_i) \right) d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N).$$

(3) At any point $(\mathbf{r}_1, \dots, \mathbf{r}_N)$ in \mathcal{M} where the function in (1) is differentiable with respect to \mathbf{r}_1 ,

$$\nabla u(\mathbf{r}_1) = \nabla_{\mathbf{r}_1} V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N). \quad (4.58)$$

In particular, if Π is of SCE form, (4.30), and $V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is the Coulomb interaction $\sum_{1 \leq i < j \leq N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$,

$$\nabla u(\mathbf{r}) = - \sum_{i=1}^{N-1} \frac{\mathbf{r} - f_i(\mathbf{r})}{|\mathbf{r} - f_i(\mathbf{r})|^3}$$

at any point \mathbf{r} where u is differentiable and $\rho(\mathbf{r}) > 0$. (4.59)

The physical and mathematical meaning of these results is as follows.

(1) says that the classical potential energy

$$E_{\text{pot}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_i u(\mathbf{r}_i)$$

is minimal on the manifold of configurations which occur with nonzero probability under the optimal plan Π . In particular, when Π is of the SCE or Monge type, the classical potential energy is minimal on the manifold (4.31) parametrized by the co-motion functions. Besides being interesting in its own right, this underlies the derivation of the next leading term of the Levy–Lieb functional outlined in Sect. 4.3.5.

(3) says that the Kantorovich potential u is an *effective one-body potential emulating the many-body system*, in the following sense: its gradient at the point r is precisely the classical repulsive force exerted on an electron at \mathbf{r} by the other electrons at positions $f_i(\mathbf{r})$. Equation (4.59) is called the *force equation*.

(2) can be viewed as an infinite-dimensional Lagrange multiplier rule, with any Kantorovich potential (i.e. any optimizer of the dual variational principle (4.55)) playing the role of a Lagrange multiplier associated with the constraint $\Pi \mapsto \rho$.

We remark that results of the above form have a long history in OT theory; for the two-marginal problem with interaction potential $|\mathbf{r}_1 - \mathbf{r}_2|$ respectively $|\mathbf{r}_1 - \mathbf{r}_2|^2$, (1) goes back to Kantorovich himself [80], while the differential version (3) and its usefulness were first realized by Knott and Smith [83].

Proof The following proof, taken from [50], is simple and illuminating, so we include it. By Kantorovich duality (in the form of Theorem 4.13 (1)),

$$0 = \int_{\mathbb{R}^{dN}} V_{ee} d\Pi - \sum_{i=1}^N \int_{\mathbb{R}^d} u(\mathbf{r}_i) \frac{\rho(\mathbf{r}_i)}{N} d\mathbf{r}_i.$$

Since Π has equal marginals $\frac{\rho}{N}, \int_{\mathbb{R}^d} u(\mathbf{r}_i) \frac{\rho(\mathbf{r}_i)}{N} d\mathbf{r}_i = \int_{\mathbb{R}^{dN}} u(\mathbf{r}_i) d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N)$, and so

$$0 = \int_{\mathbb{R}^{dN}} \left(V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_i u(\mathbf{r}_i) \right) d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N).$$

But since u satisfies the constraint (4.149) at every point in \mathbb{R}^{dN} , the integrand is nonnegative. So the minimum value of the integrand must be zero and attained, and Π must vanish wherever the integrand is positive. This establishes (1) and (2). The elementary calculus fact that the gradient of a differentiable function vanishes at minimum points now yields (4.58). Finally, (4.59) follows since the point $(\mathbf{r}_1, f_1(\mathbf{r}_1), \dots, f_{N-1}(\mathbf{r}_1))$ belongs to \mathcal{M} whenever the density ρ is positive at \mathbf{r}_1 . \square

We complete this section with a more heuristic derivation of the optimality conditions.

Heuristic derivation of Theorem 4.14 Let us re-write the SIL variational principle (4.25) in the form

$$\text{Minimize } C[\Pi] = \int_{\mathbb{R}^{dN}} V_{\text{ee}} d\Pi \text{ subject to the constraints} \tag{4.60}$$

$$G^{(\mathbf{r}_1)}[\Pi] = \rho(\mathbf{r}_1) \forall \mathbf{r}_1 \in \mathbb{R}^d,$$

where $G^{(\mathbf{r}_1)}[\Pi]$ is the functional which assigns to an N -point probability measure Π the value of its single-particle density at the point \mathbf{r}_1 , and where the minimization is over symmetric probability measures (see Remark 4.2). Let us now postulate the existence of a family of Lagrange multipliers $(\lambda(\mathbf{r}_1))_{\mathbf{r}_1 \in \mathbb{R}^d}$, one for each $G^{(\mathbf{r}_1)}$, such that minimizers of C subject to the constraints $G^{(\mathbf{r}_1)}[\Pi] = \rho(\mathbf{r}_1)$ are unconstrained minimizers of the Lagrangian

$$\mathcal{L}[\Pi] = C[\Pi] - \int_{\mathbb{R}^d} \lambda(\mathbf{r}_1) G^{(\mathbf{r}_1)}[\Pi] d\mathbf{r}_1.$$

But since $G^{(\mathbf{r}_1)}[\Pi]$ is the one-body density of Π , and Π is symmetric,

$$\begin{aligned} \mathcal{L}[\Pi] &= \int_{\mathbb{R}^{dN}} \left[V_{\text{ee}}[\mathbf{r}_1, \dots, \mathbf{r}_N] - \lambda(\mathbf{r}_1) N \right] d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) \\ &= \int_{\mathbb{R}^{dN}} \left[V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_{i=1}^N \lambda(\mathbf{r}_i) \right] d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N), \end{aligned} \tag{4.61}$$

so the Lagrangian coincides with the functional in (2) with $\lambda = u$. It is clear that minimizers of the Lagrangian must be concentrated on the set of pointwise

minimizers of the integrand, yielding (1). Statement (3) now follows as in the rigorous proof.

The above argumentation obtains the Kantorovich potential u quickly but non-rigorously as a Lagrange multiplier. In fact, with such a heuristic construction of u , statements (1) and (3) were already derived in [127] before the discovery of the *SCE theory/optimal transport* connection.

But readers are put on notice that there is no such thing as a general and rigorous Lagrange multiplier rule which would guarantee the existence of Lagrange multipliers for infinite-dimensional non-smooth problems like Levy–Lieb constrained search or its strongly interacting limit (4.25). In DFT (in its original form with both kinetic energy and electron repulsion present), the existence problem for Lagrange multipliers—i.e., the existence of one-body potentials which, when added to the Hamiltonian $T + V_{ee}$, reduce a constrained search to an unconstrained search—is known as the *v-representability problem*. This is a longstanding open problem, see e.g. [76, 84, 90, 97, 133]. For variants of the problem at positive temperature respectively quantum lattices see [19, 20]; *v*-representability for a regularization of the exact Levy–Lieb functional is discussed in Chap. 5 by Kvaal.

4.3.9 Solution of the Purely-Interacting *v*-Representability Problem

We now show that in the strongly interacting limit the *v*-representability problem, alias the problem of existence of Lagrange multipliers for density functionals defined by constrained search, can be completely solved. As we will see, this fact follows by combining known results. We assume in this section that $w(\mathbf{r}) = |\mathbf{r}|^{-1}$ is the Coulomb interaction.

Recall that a density $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ is called

- *N*-representable if it comes from a wave function $\Psi \in \mathcal{W}^N$ (i.e. $\Psi \mapsto \rho$)
- *v*-representable if it comes from a minimizer of $\langle \Psi | T + V_{ee} + \sum_i v(\mathbf{r}_i) | \Psi \rangle$ on \mathcal{W}^N for some potential $v : \mathbb{R}^d \rightarrow \mathbb{R}$
- non-interacting *v*-representable if it comes from a minimizer of $\langle \Psi | T + \sum_i v(\mathbf{r}_i) | \Psi \rangle$ on \mathcal{W}^N for some potential $v : \mathbb{R}^d \rightarrow \mathbb{R}$
- purely-interacting *v*-representable if it comes from a minimizer of $\int_{\mathbb{R}^{dN}} (V_{ee} + \sum_i v(\mathbf{r}_i)) d\Pi$ on $\mathcal{P}(\mathbb{R}^{dN})$ for some potential $v : \mathbb{R}^d \rightarrow \mathbb{R}$.

Theorem 4.15 (*N*-Representability Implies Purely-Interacting *v*-Representability) Any *N*-representable ρ , i.e. any ρ belonging to the class \mathcal{D}^N (see (4.12)), is purely-interacting *v*-representable by some bounded measurable potential $v : \mathbb{R}^d \rightarrow \mathbb{R}$. Explicitly, the following choice will do:

$$v = -u \tag{4.62}$$

where u is any bounded Kantorovich potential for ρ (see Theorem 4.13 for existence of the latter).

This result is quite remarkable, given that—to our knowledge—not much is known on the rigorous level off the strongly interacting limit.

Proof of Theorem 4.15 By Theorem 4.13 (2), there exists a bounded maximizer u of the dual functional, i.e. an associated Kantorovich potential. Let $v = -u$. By Theorem 4.1 (1), there exists a minimizer $\Pi[\rho]$ of the SIL variational problem (4.25). By Theorem 4.14 (2), this $\Pi[\rho]$ is a minimizer of $\int_{\mathbb{R}^{dN}} (V_{\text{ee}} + \sum_i v(\mathbf{r}_i)) d\Pi$ on $\mathcal{P}(\mathbb{R}^{dN})$. Since by construction Π has density ρ , it follows that v represents ρ . \square

If in addition $\rho > 0$ everywhere, the above proof together with Theorem 4.13 (3) shows that ρ is even purely-interacting v -representable by some Lipschitz continuous potential.

4.3.10 Functional Derivative and SCE Potential

It is not difficult to deduce from Theorem 4.15 that when the density ρ is sufficiently nice (say, continuous and everywhere positive) and the Kantorovich potential $u[\rho]$ (i.e. the maximizer of the dual problem (4.55)) is unique, the SCE functional is functionally differentiable at ρ with functional derivative

$$\frac{\delta V_{\text{ee}}^{\text{SCE}}[\rho]}{\delta \rho}[\rho] = u[\rho] + \text{const} \quad (4.63)$$

where const is an arbitrary additive constant. Here for any functional F on \mathcal{D}^N the functional derivative $\frac{\delta F}{\delta \rho}[\rho]$ at some density ρ (if it exists) is defined by the requirement

$$\left. \frac{d}{dt} F[\rho + t \eta] \right|_{t=0} = \int_{\mathbb{R}^d} \frac{\delta F}{\delta \rho}[\rho](\mathbf{r}) \eta(\mathbf{r}) d\mathbf{r}$$

for all smooth mass-preserving localized perturbations $\eta : \mathbb{R}^d \rightarrow \mathbb{R}$ (mathematically: $\eta \in C_0^\infty(\mathbb{R}^d)$, $\int \eta = 0$), and is unique up to an additive constant. For an informal derivation of Eq. (4.63) see e.g. [22], and for a rigorous proof under suitable assumptions see [41].

As for any Hartree-exchange-correlation functional, the Hartree-exchange-correlation potential associated to the SCE functional is the functional derivative

with additive constant chosen so that the potential vanishes at infinity, in our case

$$v_{\text{SCE}}[\rho](\mathbf{r}) = u[\rho](\mathbf{r}) + C[\rho], \quad C[\rho] \text{ a constant that ensures}$$

$$\lim_{|\mathbf{r}| \rightarrow \infty} v_{\text{SCE}}[\rho](\mathbf{r}) = 0. \quad (4.64)$$

This functional derivative is called the SCE potential.

To summarize: *the SCE potential for the strongly correlated limit of DFT agrees up to a shift with the Kantorovich potential from optimal transport theory.*

Assume now that the density is everywhere positive, that the ground state of (4.25) is an SCE state, and that

$$\text{the values } f_1(\mathbf{r}), \dots, f_{N-1}(\mathbf{r}) \text{ stay in a bounded region as } |\mathbf{r}| \rightarrow \infty. \quad (4.65)$$

It then follows from (4.59) that the SCE potential has the correct asymptotic behavior

$$v_{\text{SCE}}[\rho](\mathbf{r}) \sim_{|\mathbf{r}| \rightarrow \infty} \frac{N-1}{|\mathbf{r}|}. \quad (4.66)$$

By contrast, Hartree-exchange-correlation potentials for all semilocal functionals (LDA, GGAs) are well known to have the wrong asymptotics on physical (i.e. exponentially decaying) densities,

$$v_{\text{Hxc}}^{\text{semiloc}}[\rho](\mathbf{r}) \sim_{|\mathbf{r}| \rightarrow \infty} \frac{N}{|\mathbf{r}|}. \quad (4.67)$$

Open Problem Rigorously justify (4.65), and hence (4.66), for general densities ρ . Note that for $N = 2$ and radial densities, or any N and arbitrary densities in one dimension, assumption (4.65) follows from the explicit formulae for the f_i in [32] respectively [27] (Fig. 4.3).⁴

Example 4.16 Let $N = 2$, and let $\rho(r) = 2/\pi(1+r^2)$ be the one-dimensional Lorentzian density, normalized so that $\int \rho = 2$. The co-motion function $f_1 = f$ can be computed explicitly and is given by $f(r) = -1/r$, see Example 4.18 in Sect. 4.3.11. The SCE potential must satisfy the differential equation (4.59) which in our case reads

$$v'_{\text{SCE}}(r) = \frac{\text{sgn}(r)}{[r - f(r)]^2} = \text{sgn}(r) \frac{r^2}{(r^2 + 1)^2}.$$

The boundary condition $v_{\text{SCE}}(r) \rightarrow 0$ for $r \rightarrow \infty$ (Eq. (4.64)) yields the solution

$$v_{\text{SCE}}(r) = \frac{\text{sgn}(r)}{2} \left[\arctan r - \frac{r}{r^2 + 1} \right] - \frac{\pi}{4}.$$

⁴ Note added in proof: this problem has recently been solved in <https://arxiv.org/abs/2210.07830>.

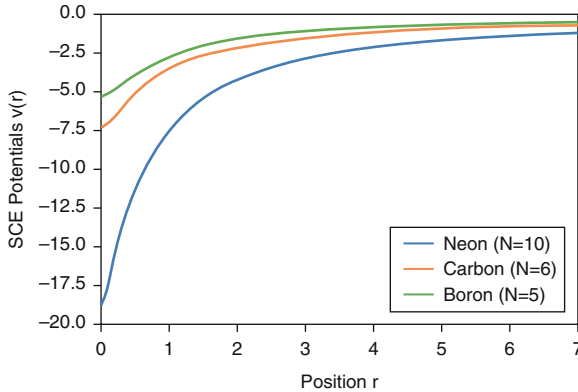


Fig. 4.3 SCE potentials $v_{SCE}(r)$ corresponding to the (radially symmetric) densities of Neon ($N = 10$), Carbon ($N = 6$) and Boron ($N = 5$), Fig. 9 in [127]. Data obtained by the following steps: (i) compute the density $\rho(r)$ by an accurate full CI or quantum Monte Carlo computation; (ii) compute the SGS maps corresponding to ρ as described in Section 4.3.12 below; (iii) obtain the corresponding SCE potentials v_{SCE} via Eqs. (4.59) and (4.64)

4.3.11 Strictly Correlated Electrons in One Dimension

In one dimension the strong interaction limit (Eq. (4.25)) can be solved exactly. The minimizing probability measure is given by an SCE state (4.30) with explicit comotion functions alias transport maps. The minimizer was found by Seidl himself in the original paper [122], on grounds of physical intuition. A proof of its optimality was found much later by Colombo, De Pascale and Di Marino [27].

Seidl's Construction For a given integrable density $\rho : \mathbb{R} \rightarrow \mathbb{R}$ with $\rho \geq 0$ and $\int \rho = N$, begin by choosing $f_1 : \mathbb{R} \rightarrow \mathbb{R}$ so that the amount of density between r and $f_1(r)$ is 1. Now choose f_2 so that the amount of density between $f_1(r)$ and $f_2(r)$ is again 1, and so on, i.e., denoting $f_0(r) = r$,

$$\int_{f_i(r)}^{f_{i+1}(r)} \rho(r') \, dr' = 1 \quad (4.68)$$

for all $i = 0, \dots, N - 1$. For Eq. (4.68) to possess a solution $f_{i+1}(r)$ in $\mathbb{R} \cup \{+\infty\}$ we must have $\int_{f_i(r)}^{\infty} \rho \geq 1$; otherwise one needs to integrate first up to $+\infty$ and then onwards from $-\infty$ so as to obtain a total value of 1,

$$\int_{f_i(r)}^{\infty} \rho(r') \, dr' + \int_{-\infty}^{f_{i+1}(r)} \rho(r') \, dr' = 1. \quad (4.69)$$

Physically this means that, given that the first electron is at some position $x_1 = r$, all the other electrons at $x_2 = f_1(r), \dots, x_N = f_{N-1}(r)$ are separated by an equal

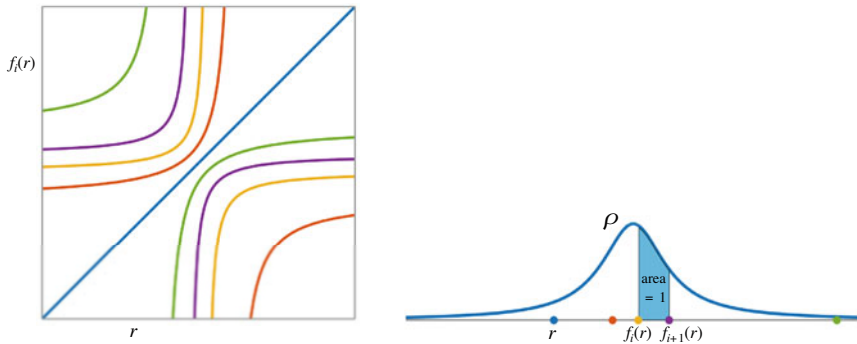


Fig. 4.4 SCE state of N electrons in one dimension. *Right*: Electron positions. Given that the first electron (depicted in blue) is at r , the positions of the other electrons are completely determined by the requirement that neighboring electrons are separated by an amount of density of 1 (blue area). The co-motion functions or transport maps f_i ($i = 1, \dots, N - 1$) of SCE theory are defined as the positions of the other electrons as a function of the first position r . The latter is distributed according to the given single-particle density ρ . *Left*: Graphs of the maps f_i , with $f_0(r) = r$ also shown. The figure corresponds to the Lorentzian density (Example 4.18) and $N = 5$

amount of density between nearest neighbors. See Fig. 4.4, right panel. As always for SCE states, the first electron position is distributed according to the given density ρ .

The above construction can be expressed concisely in terms of the cumulative distribution function

$$G_\rho(r) = \int_{-\infty}^r \frac{\rho(r')}{N} dr' \tag{4.70}$$

and its generalized inverse

$$G_\rho^{-1}(y) = \inf\{r \in \mathbb{R} : G_\rho(r) > y\}. \tag{4.71}$$

(When ρ is continuous and everywhere positive, G_ρ^{-1} is just the usual inverse function; the above definition has the virtue that it works for any nonnegative integrable ρ with integral N .) Equations (4.68) and (4.69) now take the form $G_\rho(f_{i+1}(r)) - G_\rho(f_i(r)) = 1/N$ respectively $1 - G_\rho(f_i(r)) + G_\rho(f_{i+1}(r)) = 1/N$, so by solving for f_{i+1} in terms of f_i and using $f_0(r) = r$

$$f_i(r) = \begin{cases} G_\rho^{-1}(G_\rho(r) + \frac{i}{N}) & \text{if } G_\rho(r) \leq \frac{N-i}{N} \\ G_\rho^{-1}(G_\rho(r) + \frac{i}{N} - 1) & \text{otherwise,} \end{cases} \tag{4.72}$$

for $i \in \{1, \dots, N - 1\}$.

Optimality This construction is indeed optimal:

Theorem 4.17 *Let $w(r) = |r|^{-1}$. For any nonnegative integrable density $\rho : \mathbb{R} \rightarrow \mathbb{R}$ with $\int \rho = N$, the SCE state (4.30) with f_1, \dots, f_{N-1} given by the Seidl construction (4.72) is a minimizer of the SIL problem*

$$\text{Minimize } \int_{\mathbb{R}^N} V_{\text{ee}}(r_1, \dots, r_N) \, d\Pi(r_1, \dots, r_N) \text{ over } \{\Pi \in \mathcal{P}(\mathbb{R}^N) : \Pi \mapsto \rho\}.$$

Moreover when ρ is everywhere positive, this minimizer is unique for $N = 2$, and its symmetrization (see Remark 4.2) is the unique symmetric minimizer for arbitrary N .

This theorem is due to [32] for $N = 2$ and to [27] for arbitrary N . Despite the intuitive nature of the optimizer, the proof is not elementary. It is based on a careful analysis of the structure of V_{ee} -cyclically monotone sets in \mathbb{R}^N , and strongly relies on both optimal transport theory and the ordering properties of the real line. Note that uniqueness cannot hold for $N \geq 3$ unless symmetry is required, as re-labelling the f_i then yields another solution. This is purely a mathematical, not a physical effect since solutions to the SIL problem arising as low-density limits of N -point densities of quantum wavefunctions (as described by Theorem 4.3) are always symmetric, corresponding to the symmetrization of the state (4.30) and (4.72).

Group Law Formula (4.72) implies an interesting group law for the co-motion functions, already noticed in [122]: the i th function is the i -fold composition of the first function with itself,

$$f_i = \underbrace{f_1 \circ \dots \circ f_1}_{i \text{ times}},$$

and the N -fold composition of the first function gives the identity $f_0(r) = r$.

Explicit Examples The following examples further illustrate the nonlinear governing Eqs. (4.68)–(4.69), and may serve as useful benchmarks for numerical simulations in the strongly interacting limit (or close to it).

Example 4.5, ctd Consider a two-electron system with ρ being the uniform density in a one-dimensional interval $[0, L]$. In this case we have $G_\rho(r) = r/L$, and formula (4.72) readily yields the co-motion function (4.40). For its graph, see Fig. 4.2. Mathematically this map switches the right and left half of the interval; note that its composition with itself indeed gives the identity, as it must by the group law.

Example 4.18 ([68, 71]) Let $\rho(r) = N/\pi(1 + r^2)$ be the Lorentzian density, normalized so that $\int \rho = N$. In this case $G_\rho(r) = \frac{1}{\pi} \arctan r + \frac{1}{2}$ and so Eq. (4.68)

for f_1 in the region $G_\rho(r) \leq \frac{N-1}{N}$ is, recalling the notation $f_0(r) = r$,

$$\arctan f_1(r) = \arctan r + \frac{\pi}{N}. \tag{4.73}$$

When $N = 2$ it follows that

$$f_1(r) = -\frac{1}{r} \tag{4.74}$$

(note that then the derivatives of both sides of (4.73) agree, as do their values at $r = 0$). From now on let us assume $N \geq 3$. In this case we can use the addition formula for the tangent, $\tan(x+y) = (\tan x + \tan y)/(1 - \tan x \tan y)$ for $x, y, x+y \notin \pi/2 + \mathbb{Z}$, and obtain

$$f_1(r) = \frac{r + t_1}{1 - t_1 r}, \quad t_1 = \tan \frac{\pi}{N}. \tag{4.75}$$

In the region $G_\rho(r) > \frac{N-1}{N}$, or equivalently $\arctan r > \frac{\pi}{2} - \frac{\pi}{N}$, or equivalently (because $\tan(\frac{\pi}{2} - x) = 1/\tan x$) $r > 1/t_1$, Eq. (4.69) for f_1 is

$$\arctan f_1(r) - \left(-\frac{\pi}{2}\right) = \arctan r - \frac{\pi}{2} + \frac{\pi}{N},$$

that is to say $\arctan f_1(r) = \arctan r + \frac{\pi}{N} - \pi$. Using the addition formula for the tangent and $\tan x = \tan(x - \pi)$ we again find that f_1 is given by (4.75), so this formula describes f_1 on the whole real line. It remains to compute its i -fold composition f_i . Here we give a different derivation as compared to [68, 71]. Note that mathematically f_1 is a Moebius map, i.e. a map of the form $M_a(r) = (r + a)/(1 - ar)$. Using the (elementary to check) composition formula $M_a \circ M_b = M_{\frac{a+b}{1-ab}}$ and the addition formula for the tangent we find

$$f_i(r) = \frac{r + t_i}{1 - t_i r}, \quad t_i = \tan \frac{i\pi}{N} \quad (i \in \{1, \dots, N - 1\}). \tag{4.76}$$

Moreover, setting $i = N$ in the above formula we recover the abstract fact that the N -fold composition of f_1 must be the identity. Hence the co-motion functions for the Lorenzian density form a discrete subgroup of the Moebius group. For the graph of these functions when $N = 5$ see Fig. 4.4.

4.3.12 Radially Symmetric Densities

When the one-body density ρ is radially symmetric, Seidl, Gori-Giorgi and Savin [127] conjectured an explicit minimizing probability measure in (4.25) of a radial-

symmetry-preserving SCE form which is related to the explicit SCE state of one-dimensional systems.⁵ Let us describe their conjecture in detail.

The starting point is the following reduction to a 1d problem with effective interaction.

Lemma 4.19 (Reduction to a 1d Problem, [9, 113]) *Let $\rho : \mathbb{R}^d \rightarrow \mathbb{R}$ be an integrable density with $\rho \geq 0$ and $\int \rho = N$ which is radially symmetric, that is, $\rho(\mathbf{r}) = \rho_0(|\mathbf{r}|)$ for some function ρ_0 , and let*

$$\rho_{\text{rad}}(r) = \omega_d r^{d-1} \rho_0(r),$$

where ω_d is the area of the unit sphere in \mathbb{R}^d (for $d = 3$, $\omega_d = 4\pi$). Then the SCE functional defined by (4.26) reduces to

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \min_{\eta \in \mathcal{P}([0, \infty)^N), \eta \mapsto \rho_{\text{rad}}} \int_{[0, \infty)^N} V_{\text{ee}}^{\text{rad}}(r_1, \dots, r_N) \, d\eta(r_1, \dots, r_N), \tag{4.77}$$

where $V_{\text{ee}}^{\text{rad}}$ is the reduced Coulomb cost

$$V_{\text{ee}}^{\text{rad}}(r_1, \dots, r_N) = \min \left\{ \sum_{1 \leq i < j \leq N} \frac{1}{|\mathbf{r}_j - \mathbf{r}_i|} : |\mathbf{r}_i| = r_i \, \forall i = 1, \dots, N \right\}. \tag{4.78}$$

Moreover, $\Pi \in \mathcal{P}(\mathbb{R}^{dN})$ is a minimizer for the full SIL variational principle (4.25) in d dimensions if and only if its radial projection Π_{rad} , defined by

$$\int_{A_1 \times \dots \times A_N} d\Pi^{\text{rad}}(r_1, \dots, r_N) = \int_{\{|\mathbf{r}_1| \in A_1\} \times \dots \times \{|\mathbf{r}_N| \in A_N\}} d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

for all intervals A_1, \dots, A_N , is a minimizer for the right-hand side of (4.77) and $V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = V_{\text{ee}}^{\text{rad}}(|\mathbf{r}_1|, \dots, |\mathbf{r}_N|)$ Π -a.e.

In [122, 127], the following interesting explicit state was conjectured to be optimal for the reduced problem in (4.77):

$$d\eta(r_1, \dots, r_N) = \rho_{\text{rad}}(r_1) \prod_{n=2}^N \delta(r_n - S^{(n)}(r_1)), \tag{4.79}$$

⁵ The original conjecture concerned the physical case $d = 3$. Subsequently, two-dimensional models have also been considered in the literature [125, 127].

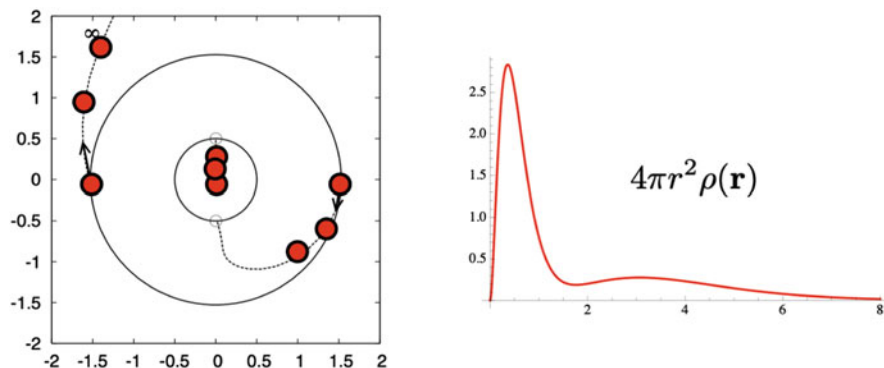


Fig. 4.5 SGS state (left panel) for the Lithium atom density (right panel). The extreme angular and radial correlation exhibited by this state is illustrated here while sending one of the electrons (the leftmost) to infinity

where $S^{(n)}$ denotes the n -fold composition $S \circ \dots \circ S$ and $S : [0, \infty) \rightarrow [0, \infty)$ is defined as follows. Let $0 = a_0 < a_1 < \dots < a_{N-1} < a_N = \infty$ be such that the intervals $A_n = [a_{n-1}, a_n)$ between successive a_n 's carry equal mass, that is, $\int_{A_n} \rho^{\text{rad}} = 1$ for all n , and let $S|_{A_n}$ be the unique function such that

$$S|_{A_n} \text{ decreasing, } S \text{ transports } \rho^{\text{rad}}|_{A_n} \text{ to } \rho^{\text{rad}}|_{A_{n+1}} \quad (4.80)$$

(with the convention $A_{N+1} = A_1$). In terms of the original SIL problem (4.25), this ansatz corresponds to the SCE ansatz (4.30) with maps satisfying the additional property

$$|f_n(\mathbf{r})| = S^{(n)}(|\mathbf{r}|)$$

for the above explicit S and with suitably chosen angles so that $\Pi \mapsto \rho$ and $V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = V_{\text{ee}}^{\text{rad}}(|\mathbf{r}_1|, \dots, |\mathbf{r}_N|)$ Π -a.e. We call S the *SGS map*, and the probability measure η given by (4.79), (4.80) the *SGS state*. See Figs. 4.5 and 4.6.

The SGS state has been rigorously proved to be optimal in some specific cases.

Example 4.20 (SCE for Radially Symmetric Densities, [32, Theorem 4.10]) Let $N = 2$, and let ρ be a radially symmetric density on \mathbb{R}^d such that $\rho(\mathbf{r}) > 0$ for all \mathbf{r} . Then the optimal co-motion function f is given by

$$f(\mathbf{r}) = s(|\mathbf{r}|) \frac{\mathbf{z}}{|\mathbf{z}|}$$

for some function $s : [0, +\infty) \rightarrow \mathbb{R}$ such that $s \leq 0$, s is increasing, $\lim_{r \rightarrow +\infty} s(z) = 0$, and $\lim_{r \rightarrow 0^+} s(r) = -\infty$.

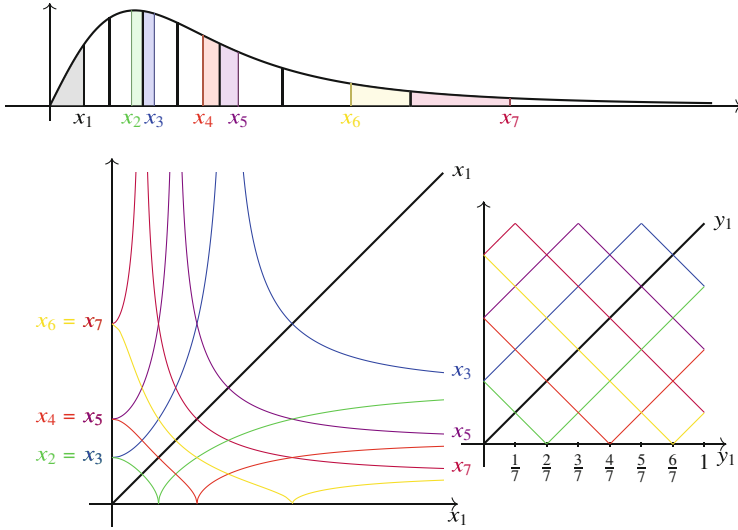


Fig. 4.6 SGS state when $N = 7$. Top: a radial measure $\rho_{\text{rad}}(r)$. Bottom left: the maps $S, S^{(2)}, \dots, S^{(6)} : [0, \infty) \rightarrow [0, \infty)$, plotted with colors green, blue, red, violet, yellow, and brown. Bottom right: the same graphs under a change of variables $y_1 = \int_0^{x_1} \frac{\rho_{\text{rad}}}{N}$ which transports ρ_{rad} to the uniform density on the interval $[0, 1]$. Picture from [124]

The function s in the above example corresponds to minus the SGS map, i.e. $s = -S$ in the $N = 2$ case.

Optimality of the SGS state has also been proved for some special class of densities ρ when $N = d = 3$ [30, 57, 125] and $N = 3$ and $d = 2$ [11].

Recently, counterexamples of radially symmetric probability densities were found for which the SGS state is not optimal. The simplest one is a uniform density on a thin annulus:

Example 4.21 ([30], see also [11, 57, 124, 125] for related examples) Let $N = 3$. For sufficiently small $\varepsilon > 0$, and the density

$$\rho_\varepsilon^{\text{rad}} = c_\varepsilon 1_{[1, 1+\varepsilon]} \tag{4.81}$$

(with the constant c_ε chosen such that $\int \rho^{\text{rad}} = 3$), the SGS state η_ε defined by (4.79)–(4.80) is not optimal for the variational problem (4.77).

This example illustrates that guessing the optimal SCE states can be a tricky business even for 1d problems, and makes it all the more remarkable that optimality of Seidl’s guess for the 1d Coulomb problem is a rigorous theorem (Theorem 4.17). The proof of nonoptimality relies on a Taylor expansion of the reduced interaction $V_{\text{ee}}^{\text{rad}}$ (defined in Eq.(4.78)) at the point $(1, 1, 1)$ and on cyclical monotonicity methods from optimal transport theory.

While this counterexample disproves optimality of the SGS state in general, the density (4.81) is quite different from typical atomic densities and the following remains an interesting mathematical problem.

Open Problem Find sufficient conditions on radial densities ρ such that the SGS state is optimal for (4.77).

4.3.13 An Example with Irregular Co-motion Functions for Repulsive Harmonic Interactions

One of the more challenging properties of co-motion functions is that they are typically discontinuous. Here we give an extreme example with modified electron-electron interaction which is discontinuous *everywhere*, due to Di Marino et al. [42].

Example 4.22 Let d be arbitrary, $V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N) = -\sum_{1 \leq i < j \leq N} |\mathbf{r}_i - \mathbf{r}_j|^2$ (repulsive harmonic interaction), and $N = 3$. Let $\rho = 3 \cdot 1_{[0,1]^d}$ (uniform density on a cube in \mathbb{R}^d). Then there exists a nowhere continuous map $T : [0, 1]^d \rightarrow [0, 1]^d$ which transports ρ to itself such that

$$d\Pi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \frac{\rho(\mathbf{r}_1)}{N} \delta(\mathbf{r}_2 - T(\mathbf{r}_1)) \delta(\mathbf{r}_3 - T(T(\mathbf{r}_1))) \tag{4.82}$$

is an optimal probability measure for the SIL problem (4.25).

The map T is an explicit fractal map. For $d = 1$ it is depicted in Fig. 4.7 and constructed as the unique fixed point of the iteration

$$f \mapsto g(x) = \begin{cases} \frac{1}{3}f(3x) + \frac{1}{3} & \text{for } 0 \leq x < \frac{1}{3} \\ \frac{1}{3}f(3x - 1) + \frac{2}{3} & \text{for } \frac{1}{3} \leq x < \frac{2}{3} \\ \frac{1}{3}f(3x - 2) & \text{for } \frac{2}{3} \leq x < 1, \end{cases} \tag{4.83}$$

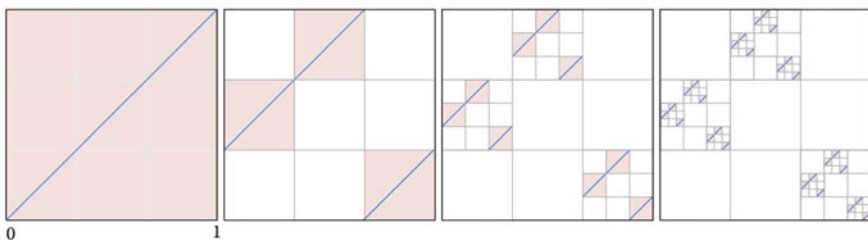


Fig. 4.7 Construction of the optimal map T in Example 4.22. The picture shows the graph (in blue) of the first few iterations of Eq. (4.83); each graph consists of three scaled copies of the previous one. The exact map is reached in the limit of infinitely many iterations

starting with $f(x) = x$. To see what the iteration is doing, divide $[0, 1]^2$ into a 3×3 grid of squares and put scaled copies of the graph of the original function into the two squares directly above the diagonal and the bottom right square. Optimality of the resulting fractal SCE state (4.82) is easy to see from the following special property of the repulsive harmonic cost which was first observed by Pass [112]: thanks to the identity $V_{ee} = |\mathbf{r}_1 + \dots + \mathbf{r}_N|^2 - N \sum_{i=1}^N |\mathbf{r}_i|^2$ and the fact that the integral of the second term against a probability measure Π only depends on its marginal, the minimizers of the SIL problem are precisely the probability measures supported on the surface $\mathbf{r}_1 + \dots + \mathbf{r}_N = 0$.

The above example and construction works for arbitrary N , see [42].

Open Problem Do such extreme examples also occur for the Coulomb interaction? Note that the repulsive harmonic interaction arises by locally Taylor-expanding the Coulomb interaction in angular direction.

4.3.14 *Minimizers of the Discretized SIL Variational Principle Are Quasi-Monge States*

We now come back to the important issue that the SIL variational principle (4.25) still requires minimization over a high-dimensional space of N -point probability measures, whereas the low-dimensional SCE ansatz (4.30) can fail to yield an optimizer of (4.25). One can ask whether some modified low-dimensional ansatz is enough to solve (4.25) exactly. In other words, can one achieve Seidl's original goal of solving the strongly interacting limit of DFT with a low-dimensional ansatz that can be easily stored on a computer?

For the discretization of (4.25) on a grid, Friesecke and Vögler [53] found a modified ansatz which achieves this, for arbitrary space dimensions, densities and interaction potentials:

$$\begin{aligned} d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) &= S_N \int_{\mathbb{R}^d} \alpha(\mathbf{r}) \prod_{n=1}^N \delta(\mathbf{r}_n - f_{n-1}(\mathbf{r})) \, d\mathbf{r} & (4.84) \\ &= \frac{1}{N!} \sum_{\sigma} \int_{\mathbb{R}^d} \alpha(\mathbf{r}) \prod_{n=1}^N \delta(\mathbf{r}_n - f_{\sigma(n-1)}(\mathbf{r})) \, d\mathbf{r}, \end{aligned}$$

where σ runs over all permutations of the indices $0, \dots, n-1$, α is some (free to choose) probability density on the single-particle space \mathbb{R}^d , and the f_n are maps from \mathbb{R}^d to \mathbb{R}^d . States of this form are called *quasi-Monge states* or *quasi-SCE states*. With the specific choice $\alpha = \frac{\rho}{N}$, the quasi-Monge ansatz (4.84) reduces precisely to the SCE (alias Monge) ansatz in its symmetric form (4.42). The novelty is the additional freedom of choosing the auxiliary density α . For the quasi-Monge ansatz, the marginal constraint $\Pi \mapsto \rho$ takes, instead of the conditions $f_{n\mp}\rho = \rho$

($n = 0, \dots, N - 1$) (Eq. (4.32)), the form of a single condition,

$$\frac{1}{N} \sum_{n=0}^{N-1} f_n \# \alpha = \frac{\rho}{N}. \tag{4.85}$$

That is, the *average* push-forward of the auxiliary density α under the quasi-SCE maps must be the (suitably normalized) physical density.

Plugging the ansatz (4.84) into the SIL variational principle (4.25) and integrating out the variables $\mathbf{r}_2, \dots, \mathbf{r}_N$ yields the *quasi-Monge* or *quasi-SCE* variational principle

$$\begin{aligned} \text{Minimize } \int_{\mathbb{R}^d} V_{\text{ee}}(f_0(\mathbf{r}), \dots, f_{N-1}(\mathbf{r})) \alpha(\mathbf{r}) \, d\mathbf{r} \text{ over probability densities } \alpha \\ \text{and maps } f_0, \dots, f_{N-1}, \end{aligned} \tag{4.86}$$

with the minimization being subject to the constraint (4.85).

Theorem 4.23 (Justification of the Quasi-Monge Ansatz, [53]) *Let ρ be any discrete N -particle density on \mathbb{R}^d , that is to say $\rho(\mathbf{r}) = \sum_{i=1}^{\ell} \rho_i \delta(\mathbf{r} - \mathbf{a}_i)$ for some distinct discretization points $\mathbf{a}_i \in \mathbb{R}^d$ and some $\rho_i \geq 0$ with $\sum_i \rho_i = N$, and let $V_{\text{ee}} : \mathbb{R}^{dN} \rightarrow \mathbb{R} \cup \{+\infty\}$ be any interaction potential which is symmetric in the electron coordinates (e.g., the Coulomb interaction $V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i < j} 1/|\mathbf{r}_i - \mathbf{r}_j|$). Then the SIL problem (4.25) possesses a minimizer which is a quasi-Monge state (4.84). Equivalently, it possesses a minimizer of the form (4.88), i.e., a superposition of at most ℓ symmetrized Dirac measures.*

This result rigorously reduces the number of unknowns from exponential to linear with respect to the number of electrons; more precisely, from ℓ^N (the dimension of the space of N -point probability measures supported on $\{\mathbf{a}_1, \dots, \mathbf{a}_\ell\}^N$) to $\ell \cdot (N + 1)$ (ℓ unknowns for each of the N quasi-Monge maps, and another ℓ unknowns for the auxiliary density α).

The above result fails if the class of quasi-Monge states is narrowed to Monge (alias SCE) states, see [49]. For continuous ρ 's, it is an open question whether the SIL problem always (or at least in the Coulomb case) admits minimizers of quasi-Monge form.

Proof of Theorem 4.23 (Following [53].) Let us explain the intuition and reasoning behind the quasi-Monge ansatz and Theorem 4.23, which comes from convex geometry. Before passing to a geometric viewpoint, we note that by the symmetry of V_{ee} the minimization in (4.25) can be restricted to symmetric probability measures

(see Remark 4.2); moreover any symmetric probability measure $\Pi \in \mathcal{P}(\mathbb{R}^{dN})$ with $\Pi \mapsto \rho$ must be of the form

$$d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{i_1, \dots, i_N=1}^{\ell} \gamma_{i_1 \dots i_N} \delta(\mathbf{r}_1 - \mathbf{a}_{i_1}) \cdots \delta(\mathbf{r}_N - \mathbf{a}_{i_N}) d\mathbf{r}_1 \dots d\mathbf{r}_N$$

for some symmetric tensor $(\gamma_{i_1 \dots i_N}) \in \mathbb{R}^{\ell \times \dots \times \ell}$ with nonnegative entries which sum to 1. Now geometrically, for fixed discretization points $\mathbf{a}_1, \dots, \mathbf{a}_\ell$ the set of these probability measures is a finite-dimensional convex polytope; let us denote it by $\mathcal{P}_{\text{sym}}(\{\mathbf{a}_1, \dots, \mathbf{a}_\ell\}^N)$. The subset satisfying the marginal constraint $\Pi \mapsto \rho$, i.e.

$$\sum_{i_2, \dots, i_N=1}^{\ell} \gamma_{i_1 i_2 \dots i_N} = \rho_{i_1} \quad \forall i_1 \in \{1, \dots, \ell\}, \tag{4.87}$$

is also a convex polytope called the *Kantorovich polytope*; let us denote it by $\mathcal{P}_\rho(\{\mathbf{a}_1, \dots, \mathbf{a}_\ell\}^N)$. While general probability measures in these sets possess a huge number of coefficients which increases combinatorially with the number N of particles, the key point is that the *extreme points*⁶ of these sets are very sparse, with only a small number of nonzero coefficients. The extreme points of $\mathcal{P}_\rho(\{\mathbf{a}_1, \dots, \mathbf{a}_\ell\}^N)$ are easily seen to be symmetrized products of delta functions, $S_N \delta(\mathbf{r}_1 - \mathbf{a}_{i_1}) \cdots \delta(\mathbf{r}_N - \mathbf{a}_{i_N})$, where S_N is the symmetrization operator. Now consider a subset of a convex polytope satisfying *one* linear constraint, geometrically: the intersection of the polytope with a hyperplane. It is geometrically expected (and not difficult to prove) that all extreme points of this new set are convex combinations of just two extreme points of the original polytope. Analogously, by a well-known result in convex geometry the intersection of a convex polytope with k hyperplanes has extreme points given by convex combinations of just $k+1$ of the original extreme points. Since the marginal condition (4.87) imposes $\ell - 1$ constraints (note that one of the ℓ constraints is redundant due to the sum of the $\gamma_{i_1 \dots i_N}$ being 1), the extreme points of the Kantorovich polytope are convex combinations of just ℓ symmetrized delta functions, i.e., probability measures of the form

$$\sum_{v=1}^{\ell} \alpha_v S_N \delta(\mathbf{r}_1 - \mathbf{a}_{i_1^{(v)}}) \cdots \delta(\mathbf{r}_N - \mathbf{a}_{i_N^{(v)}}) \tag{4.88}$$

for some nonnegative coefficients α_v . Defining the maps f_n by $f_{n-1}(\mathbf{a}_v) = \mathbf{a}_{i_n^{(v)}}$ yields that all extreme points are quasi-Monge states (4.84). Theorem 4.23 now follows from the general principle that the minimum of a linear functional (such as $\int V_{\text{ec}} d\Pi$) over a convex polytope is always attained at some extreme point. \square

⁶ These are the points that cannot be written as convex combinations of any other points in the set.

The quasi-Monge (or quasi-SCE) ansatz and Theorem 4.23 underlie the numerical method described in Sect. 4.4.5.3.

4.3.15 Entropic Regularization of the SCE Functional

We have seen in Fig. 4.1 and Sect. 4.3.8 that in the strongly interacting limit, the N -body density concentrates on the lower-dimensional manifold on which the classical effective potential energy $V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_i v_{SCE}(\mathbf{r}_i)$ is minimal. A regularization of the SCE functional which has nice mathematical properties and smears out the N -body density is the following:

$$V_{ee}^\tau[\rho] = \inf_{\pi \in \mathcal{P}(\mathbb{R}^{dN}) \cap L^1(\mathbb{R}^{dN}), \pi \mapsto \rho} \mathcal{V}_{ee}[\pi] + \tau S[\pi]. \tag{4.89}$$

Here $\tau > 0$ is a small parameter, \mathcal{V}_{ee} is the usual electron interaction energy, and S is (minus) the Shannon–Von Neumann entropy,

$$S[\pi] = \int_{\mathbb{R}^{dN}} \pi(\mathbf{r}_1, \dots, \mathbf{r}_N) (\log \pi(\mathbf{r}_1, \dots, \mathbf{r}_N) - 1) d\mathbf{r}_1 \dots d\mathbf{r}_N. \tag{4.90}$$

As shown in Lemma 4.27 below, the negative part of the entropy density has finite integral under very mild conditions on ρ (e.g., finite first moment suffices), and so definitions (4.89)–(4.90) make rigorous sense. The existence of a minimizer in (4.89) can be obtained assuming that $\rho \log \rho \in L^1(\mathbb{R}^d)$ [60]. Physically, the right-hand side in (4.89) can be viewed as the free energy of N classical particles with interaction potential V_{ee} and density ρ at inverse temperature τ . But our goal here is not to model a physical system at finite temperature, but instead to approximate the SCE functional.

Figure 4.8 illustrates the effect of the entropy term in a two-electron example: the larger the regularization parameter τ , the more the minimizers π are spread out around the support of the SCE state. (Recall that by Theorem 4.4, when $N = 2$ the SIL variational principle is uniquely minimized by an SCE state.)



Fig. 4.8 Numerically computed optimal density $\pi^\tau(r_1, r_2)$ in (4.89) for two electrons in one dimension and the density $\rho(r) = c_0/\cosh r$, $r \in [-10, 10]$, for different values of the regularization parameter τ . Here c_0 is a normalization constant, and we have used the effective Coulomb interaction [7] $V_{ee}(r_1, r_2) = 1.07 e^{-|r_1 - r_2|/2.39}$

The corresponding regularization for the Wasserstein distance squared instead of the SCE functional (Example 4.12 instead of Example 4.11) in fact goes back to Erwin Schrödinger in 1931 [121], and had a completely different motivation: Schrödinger was looking for models for the “most likely” evolution law between two probability distributions of particle positions which have been empirically observed at different times, perhaps hoping to re-discover his—then still controversial—quantum mechanics in a novel way.

An equivalent entropic problem has been considered by Chayes et al. in [19, 20]. In their setting, the integration in the entropy functional S is not against the Lebesgue measure but against the product measure $\prod_{i=1}^N \rho(\mathbf{r}_i)/N$, which constitutes a natural model in classical statistical mechanics. The role of the reference measure will be explained below and can be understood via Eq. (4.93).

In optimal transport, entropic regularization became a popular basis for computational methods following an influential paper by Cuturi [34] in machine learning and Galichon and Salanié in economics [54]; see Sect. 4.4 for a computational algorithm. Regularization by entropies other than the Shannon–Von Neumann one is considered in [41, 99].

4.3.15.1 Basic Properties

Let us now informally discuss the basic properties of (4.89).

Unique Minimizer Assuming that $\rho \log \rho \in L^1(\mathbb{R}^d)$, a minimizer π^τ in (4.89) exists [60]. This is expected from the convexity of the functional $\mathcal{V}_{ee} + \tau S$. Since the functional is strictly convex on the domain where it is finite, minimizers must be unique.

Euler–Lagrange Equation; Form of Minimizer Assume a Lagrange multiplier rule as in (4.61). That is, assume the existence of Lagrange multipliers $(\lambda(\mathbf{r}_i))_{\mathbf{r}_i \in \mathbb{R}^d}$ such that π^τ is the unconstrained minimizer of the Lagrangian

$$\mathcal{L}[\pi] = \int_{\mathbb{R}^{dN}} \left(V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_{i=1}^N \lambda(\mathbf{r}_i) \right) \pi(\mathbf{r}_1, \dots, \mathbf{r}_N) d\mathbf{r}_1 \dots d\mathbf{r}_N + \tau S[\pi].$$

Thus the function $\lambda(\mathbf{r})$ has the usual physical interpretation of DFT as minus the potential that enforces the density constraint, and will in the following be denoted $u^\tau(\mathbf{r})$. It follows that

$$0 = \left. \frac{d}{dt} \right|_{t=0} \mathcal{L}[\pi + t\eta]$$

for all variations η with $\int \eta = 0$ and $\pi \pm \eta \geq 0$. That is to say, $0 = \int [(V_{ee} - \sum_i u^\tau(\mathbf{r}_i)) + \tau \log \pi] \eta$ and therefore

$$V_{ee} - \sum_i u^\tau(\mathbf{r}_i) + \tau \log \pi = \text{const.}$$

By solving for π and adjusting u^τ by an additive constant, it follows that

$$\pi^\tau(\mathbf{r}_1, \dots, \mathbf{r}_N) = \prod_{i=1}^N a^\tau(\mathbf{r}_i) e^{-\frac{V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N)}{\tau}} \quad \text{with } a^\tau(\mathbf{r}_i) = e^{\frac{1}{\tau} u^\tau(\mathbf{r}_i)}. \quad (4.91)$$

The function a^τ can—independently of its construction above with the help of Lagrange multipliers—be interpreted as an entropic weight function which makes the probability density π^τ satisfy the constraint $\pi^\tau \mapsto \rho$. Note that by this constraint and Eq. (4.91), a^τ must satisfy the following governing equation in which Lagrange multipliers no longer appear:

$$a^\tau(\mathbf{r}_j) \int_{\mathbb{R}^{d(N-1)}} \prod_{i \neq j} a^\tau(\mathbf{r}_i) e^{-\frac{V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N)}{\tau}} \prod_{i \neq j} d\mathbf{r}_i = \frac{\rho(\mathbf{r}_j)}{N} \quad \forall j \in \{1, \dots, N\}. \quad (4.92)$$

The above equations constitute the so-called (multi-marginal) Schrödinger system. When the density ρ is Gaussian and the interaction potential w is taken to be the repulsive or attractive harmonic interaction, the entropically regularized problem can be solved exactly, see [58], for the one-dimensional case, and [39, 79, 103] for the general case. In [17], Carlier and Laborde showed the existence of a solution of the system (4.92) via an inverse function theorem argument by assuming that the one-body density ρ belongs to $L^\infty(\mathbb{R}^d)$.

Relative Entropy Formulation The functional $\mathcal{V}_{ee} + \tau S$ agrees up to an additive constant with the Kullback–Leibler divergence (or minus the relative entropy)⁷ between π and a kernel function \mathcal{K} of the electronic interaction V_{ee} [89]:

$$\mathcal{V}_{ee}[\pi] + \tau S[\pi] = \tau \text{KL}(\pi|\mathcal{K}) - \tau \quad \text{with } \mathcal{K} = e^{-V_{ee}/\tau}.$$

Thus the optimizer π^τ is the density with marginal ρ which has minimal relative entropy with respect to the kernel \mathcal{K} .

⁷ The KL divergence between two nonnegative densities with possibly unequal mass is formally defined as $\text{KL}(f|g) = \int f \log \frac{f}{g}$.

The Role of the Reference Measure In the literature, the entropy functionals which are typically studied replace integration against the Lebesgue measure in (4.90) by integration against the product of the marginals $\mu^{\otimes N} = \otimes_{i=1}^N \mu$ (or any other finite reference measure), where $\rho/N = \mu$. As shown in Lemma 1.5 in [40] (see also [60] for the Coulomb case), both problems are equivalent since the following identity holds

$$V_{\text{ce}}^\tau[\rho] = \inf_{\pi \mapsto \rho} \left\{ \mathcal{V}_{\text{ce}}[\pi] + \tau \int_{\mathbb{R}^{dN}} \frac{d\pi}{d\mu} \left(\log \frac{d\pi}{d\mu} - 1 \right) d\mu \right\} + \tau \int_{\mathbb{R}^d} \rho \log \frac{\rho}{N} d\mathbf{r}. \quad (4.93)$$

Therefore, whenever at least one side of the equality above is finite, the original variational problem from the definition of $V_{\text{ce}}^\tau[\rho]$ (Eq. (4.89)) and the variational problem defined on the right-hand side of (4.93) have the same minimizers.

Dual Formulation As for the exact (unregularized) strong-interaction limit of DFT, there is a dual variational principle for the Lagrange multiplier and an associated dual construction of $V_{\text{ce}}^\tau[\rho]$. We have

$$V_{\text{ce}}^\tau[\rho] = \sup_u J[u], \quad (4.94)$$

where

$$J[u] = \int_{\mathbb{R}^d} u(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} - \tau \int_{\mathbb{R}^{dN}} e^{-\frac{1}{\tau} [V_{\text{ce}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_i u(\mathbf{r}_i)]} d\mathbf{r}_1 \dots d\mathbf{r}_N \quad (4.95)$$

and the supremum in (4.94) is over a suitable class of potentials. The second term in (4.95) can be viewed as a soft version of the inequality constraint $E_{\text{pot}}(\mathbf{r}_1, \dots, \mathbf{r}_N) = V_{\text{ce}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - u(\mathbf{r}_1) + \dots + u(\mathbf{r}_N) \geq 0$ in the unregularized theory (see (4.149)), as it penalizes deviations from this inequality. Indeed, via the Laplace principle we have that, whenever the second term in (4.94) is finite,

$$\begin{aligned} \lim_{\tau \rightarrow 0^+} -\tau \log \left(\int_{\mathbb{R}^{dN}} e^{-\frac{1}{\tau} [V_{\text{ce}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_i u(\mathbf{r}_i)]} d\mathbf{r}_1 \dots d\mathbf{r}_N \right) \\ = \inf_{\mathbf{r}_1, \dots, \mathbf{r}_N \in \mathbb{R}^d} \{E_{\text{pot}}(\mathbf{r}_1, \dots, \mathbf{r}_N)\}. \end{aligned}$$

In the discrete setting, this is precisely the LogSumExp formula. The existence of an optimizer u^τ for the dual problem and the representation formulae (4.91), (4.96) with this u^τ were proved in [40, 41] under the assumption that $\rho \log \rho \in L^1(\mathbb{R}^d)$ and V_{ce} is measurable and bounded.

Functional Derivative As in exact SCE theory, the functional derivative of the energy functional is formally given by the optimal potential in the dual problem, that is to say

$$\frac{\delta V_{\text{ee}}^\tau[\rho]}{\delta \rho} = u^\tau + \text{const}, \quad (4.96)$$

where u^τ is the maximizer of (4.95) (assuming such a maximizer exists and is unique). As in SCE theory, a natural choice of the additive constant is to require $\lim_{|\mathbf{r}| \rightarrow \infty} (u^\tau(\mathbf{r}) + \text{const}) = 0$. The ensuing potential $v^\tau = u^\tau + \text{const}$ is then an approximation to the SCE potential.

4.3.15.2 Relation with the Levy–Lieb Functional

Just like the SCE functional itself, its entropic regularization is a rigorous lower bound of the exact functional, provided the regularization parameter τ is chosen suitably. More precisely:

Theorem 4.24 ([125]) *Let Ψ be any N -electron wavefunction in the space \mathcal{W}^N (see (4.7)), or alternatively any bosonic wavefunction in $H^1(\mathbb{R}^{dN})$, and suppose $\Psi \mapsto \rho$. Let V_{ee} be the Coulomb interaction. Then the scaled Levy–Lieb functional defined in Eq. (4.17) satisfies*

$$\frac{F^\lambda[\rho]}{\lambda} \geq V_{\text{ee}}^\tau[\rho] \quad \text{with } \tau = \frac{\pi}{2\lambda}. \quad (4.97)$$

In particular, the original Levy–Lieb functional (4.11) satisfies

$$F_{\text{LL}}[\rho] \geq V_{\text{ee}}^{\pi/2}[\rho]. \quad (4.98)$$

This result is a consequence of the logarithmic Sobolev inequality (LSI). We include a proof, following Seidl et al. [125]. We begin by recalling a standard version of the LSI.

Theorem 4.25 (LSI, Corollary 7.3 in [70]) *Let $\nu \in \mathcal{P}(\mathbb{R}^n)$ such that $\nu(\mathbf{r}) = e^{-V(\mathbf{r})}$ with $D^2V \geq \kappa \text{Id}$. Then, for every locally integrable function $f \geq 0$ on \mathbb{R}^n such that $f\nu \in \mathcal{P}(\mathbb{R}^n)$ we have that $\int_{\mathbb{R}^n} f \log f \, d\nu \leq \frac{2}{\kappa} \int |\nabla \sqrt{f}|^2 \, d\nu$.*

This implies the following LSI for the Lebesgue measure:

Corollary 4.26 (LSI for the Lebesgue Measure, [125]) *Let $f \geq 0$ be a function such that $\sqrt{f} \in H^1(\mathbb{R}^n)$ and $f \in \mathcal{P}(\mathbb{R}^n)$. Then $\int_{\mathbb{R}^n} f \log f \, d\mathbf{r} \leq \frac{1}{\pi} \int_{\mathbb{R}^n} |\nabla \sqrt{f}|^2 \, d\mathbf{r}$.*

Proof of Corollary 4.26

1. In the LSI in Theorem 4.25, the requirement on f that $\int f \, d\nu = 1$ can be relaxed to $0 < \int f \, d\nu \leq 1$. This follows by applying the LSI to f/α , $\alpha = \int f \, d\nu$, and noting that the extra term $-(1/\alpha) \int f \log \alpha \, d\nu$ on the left-hand side is ≥ 0 .
2. Take $\nu_{\mathbf{r}_2}(\mathbf{r}_1) = e^{-\pi|\mathbf{r}_1-\mathbf{r}_2|^2}$, then ν satisfies the assumption of the LSI with $\kappa = 2\pi$, and moreover $\int f \, d\nu_{\mathbf{r}_2} \leq 1$. Hence by the LSI,

$$\int f \log f \, d\nu_{\mathbf{r}_2} \leq \frac{1}{\pi} \int |\nabla \sqrt{f}|^2 \, d\nu_{\mathbf{r}_2}.$$

3. Integrate over \mathbf{r}_2 and use that $\int e^{-\pi|\mathbf{r}_1-\mathbf{r}_2|^2} \, d\mathbf{r}_2 = 1$. This yields the assertion. □

Proof of Theorem 4.24 Let $\Psi \in \mathcal{W}^N$, $\Psi \mapsto \rho$, and let Π be its N -point position density (4.8). By a version of the Hoffmann-Ostenhof inequality [77],⁸ $\sqrt{\Pi} \in H^1(\mathbb{R}^{dN})$ and

$$T[\Psi] \geq \frac{1}{2} \int_{\mathbb{R}^{dN}} |\nabla \sqrt{\Pi}|^2 \, d\mathbf{r}_1 \dots d\mathbf{r}_N.$$

This together with the LSI for the Lebesgue measure (Corollary 4.26) applied to π gives

$$\begin{aligned} \frac{1}{\lambda} T[\Psi] + V_{\text{ee}}[\Psi] &\geq \frac{\pi}{2\lambda} \int_{\mathbb{R}^{dN}} \Pi \log \Pi \, d\mathbf{r}_1 \dots d\mathbf{r}_N + \int_{\mathbb{R}^{dN}} V_{\text{ee}} \Pi \, d\mathbf{r}_1 \dots d\mathbf{r}_N \\ &= \frac{\pi}{2\lambda} (S(\Pi) + 1) + V_{\text{ee}}[\Pi]. \end{aligned}$$

Taking the infimum over $\Psi \in \mathcal{W}^N$ yields $F^\lambda[\rho]/\lambda \geq V_{\text{ee}}^\tau[\rho] + \tau$, with τ as in the theorem. □

Although Theorem 4.24 provides a lower bound for the Levy–Lieb functional (4.11), in practice this bound can be rather loose [58].

⁸ Strictly speaking, this inequality and related ones are proved in [77] under the tacit assumption that $\sqrt{\Pi}$ (or related reduced quantities) belong to H^1 and can be differentiated by the chain rule. For further discussion of this point see Chap. 5 by Kvaal in this volume.

4.3.15.3 Well Definedness of Entropy and Convergence to the SCE Functional

We now show that the entropy is well defined under very mild conditions on ρ (e.g., finite first moment suffices), and that the entropically regularized functional V_{ee}^τ converges to the SCE functional when the regularization parameter tends to zero.

Note that a priori both the positive and the negative part of the integral (4.90) could be divergent; Lemma 4.27 excludes this for the negative part, and so the integral always has a well defined value in $\mathbb{R} \cup \{+\infty\}$.

Lemma 4.27 (Well-Definedness of Entropy and of the Regularized SCE Functional) *Let $\rho \in L^1(\mathbb{R}^d)$, $\rho \geq 0$, $\int \rho = N$, and assume ρ has finite first moment, that is to say $\int |\mathbf{r}|\rho(\mathbf{r})d\mathbf{r} < \infty$. Let $\pi \in \mathcal{P}(\mathbb{R}^{dN}) \cap L^1(\mathbb{R}^{dN})$ with $\pi \mapsto \rho$. Then the negative part $(\pi \log \pi)_-$ has finite integral; more precisely, for some constant $A_\rho > 0$ which depends only on ρ but not on π*

$$\int (\pi \log \pi)_- \geq -A_\rho > -\infty,$$

where $f_-(\mathbf{r}) = \min\{f(\mathbf{r}), 0\}$ denotes the negative part of a function f . Hence S as defined by (4.90) is well defined as a functional

$$S : \{\pi \in \mathcal{P}(\mathbb{R}^{dN}) \cap L^1(\mathbb{R}^{dN}) : \pi \mapsto \rho\} \rightarrow \mathbb{R} \cup \{+\infty\},$$

and $V_{ee}^\tau[\rho]$ as defined by (4.89) is well defined as an element of $\mathbb{R} \cup \{+\infty\}$.

The assumption that ρ has finite first moment cannot be omitted. For instance, for $N = 2$ and $d = 1$ the N -body density

$$\pi(r_1, r_2) = c_0 \prod_{i=1}^2 \frac{1}{r_i (\log r_i)^2} \text{ on } [2, \infty)^2,$$

continued by zero to \mathbb{R}^2 and with c_0 chosen such that $\int \pi = 1$, belongs to $L^1(\mathbb{R}^2)$ but satisfies $\int (\pi \log \pi)_- = -\infty$, as the interested reader can check using that $\int_2^\infty \frac{1}{z|\log z|^\alpha} dz = \infty$ for $\alpha = 1$ but $< \infty$ for $\alpha > 1$. In particular, in such a case the equivalence described in (4.93) does not necessarily hold.

The lemma implies that for any interaction potential V_{ee} on \mathbb{R}^{dN} which is symmetric and bounded from below, such as the Coulomb interaction, $V_{ee}^\tau[\rho]$ is well defined as an element of $\mathbb{R} \cup \{+\infty\}$.

Proof of Lemma 4.27 $\pi \log \pi$ is ≤ 0 precisely in the region $\Omega = \{\mathbf{r} \in \mathbb{R}^{dN} : \pi(\mathbf{r}) \leq 1\}$. Split Ω into $\Omega_< = \{\mathbf{r} \in \Omega : 0 \leq \pi(\mathbf{r}) < e^{-(|\mathbf{r}_1|+\dots+|\mathbf{r}_N|)}\}$ and $\Omega_> = \{\mathbf{r} \in \Omega : \pi(\mathbf{r}) \geq e^{-(|\mathbf{r}_1|+\dots+|\mathbf{r}_N|)}\}$. Since $g(z) = z \log z$ satisfies $|g(z)| \leq C\sqrt{z}$ in

$[0, 1]$ for some constant C ,

$$\begin{aligned} \int_{\mathbb{R}^{dN}} |(\pi \log \pi)_-| &= \int_{\Omega_{<}} |(\pi \log \pi)_-| + \int_{\Omega_{>}} |(\pi \log \pi)_-| \\ &\leq C \int_{\Omega_{<}} e^{-(|\mathbf{r}_1|/2+\dots+|\mathbf{r}_N|/2)} \mathbf{d}\mathbf{r}_1 \dots \mathbf{d}\mathbf{r}_N \\ &\quad + \int_{\Omega_{>}} \pi(\mathbf{r}_1, \dots, \mathbf{r}_N) (|\mathbf{r}_1| + \dots + |\mathbf{r}_N|) \mathbf{d}\mathbf{r}_1 \dots \mathbf{d}\mathbf{r}_N \\ &\leq C \left(\int_{\mathbb{R}^d} e^{-|\mathbf{r}_1|/2} \mathbf{d}\mathbf{r}_1 \right)^N + \int_{\mathbb{R}^d} \rho(\mathbf{r}_1) |\mathbf{r}_1| \mathbf{d}\mathbf{r}_1 =: A_\rho. \end{aligned} \tag{4.99}$$

By the assumption that ρ has finite first moment, the right-hand side is finite, completing the proof of the lemma. \square

Finally, we prove that—as intuitively expected—the entropically regularized functional V_{ee}^τ converges to the exact SCE functional when the regularization parameter tends to zero. The corresponding Γ -convergence result was obtained in [60].

Theorem 4.28 *Let ρ be any N -electron density which belongs to the class \mathcal{D}^N (see (4.12)) and has finite first moment, and let V_{ee} be the Coulomb interaction. Then*

$$\lim_{\tau \rightarrow 0} V_{ee}^\tau[\rho] = V_{ee}^{\text{SCE}}[\rho]. \tag{4.100}$$

Proof We combine the upper bound on $V_{ee}^\tau[\rho]$ from Theorem 4.24, the asymptotic result on $F^\lambda[\rho]/\lambda$ in Eq.(4.19) (see Theorem 4.3), and the lower bound from Lemma 4.27. By inequality (4.99) we have for any $\pi \mapsto \rho$

$$\tau S[\pi] \geq \tau \int_{\mathbb{R}^{dN}} (\pi \log \pi)_- - \tau \geq -\tau (A_\rho + 1)$$

and hence, by adding $\mathcal{V}_{ee}[\pi]$ to both sides and taking the infimum over π

$$V_{ee}^\tau[\rho] \geq V_{ee}^{\text{SCE}}[\rho] - \tau (A_\rho + 1).$$

Obviously this lower bound converges to $V_{ee}^{\text{SCE}}[\rho]$ as $\tau \rightarrow 0$. On the other hand, by Theorem 4.24 we have $V_{ee}^\tau[\rho] \leq F^\lambda[\rho]/\lambda$ and by Theorem 4.3 this upper bound also converges to $V_{ee}^{\text{SCE}}[\rho]$; hence so must $V_{ee}^\tau[\rho]$. \square

4.4 Numerical Methods and Approximations

The SCE functional cannot at the moment be accurately and efficiently computed for general three-dimensional densities and large N . But accurate numerical methods are available for small N or special situations, novel methods aimed at large N are under development, and less accurate approximations can already be computed for large N . We review these methods and approximations in this section, and their use within Kohn–Sham DFT in Sect. 4.5.

4.4.1 Numerical Methods Based on Co-motion Functions

Numerical implementations using co-motion functions were confined to the following cases:

- the exact maps are known: general N in one dimension (see Sect. 4.3.11);
- an explicit ansatz, able to get very close to the true minimum, exists: spherically symmetric (radial) case (see Sect. 4.3.12).

In addition, co-motion functions can be extracted from optimal plans in the case

- $N = 2$, for which the existence of the map is proven and there are 1-1 correspondences between map, optimal plan, and Kantorovich potential (see Eqs. (4.38) and (4.39)).

We review here and in the following section the implementation for these three classes of problems. Their use in combination with Kohn–Sham DFT is then discussed in Sect. 4.5.

4.4.1.1 One-Dimensional N -Electron Systems

The SCE functional has been implemented for one-dimensional (1D) many-electron systems using the exact co-motion functions (maps) of Seidl [122], which we reported and illustrated in Sect. 4.3.11. These applications typically aim at modeling physical systems in which electrons are confined in elongated traps (quantum wires): the interaction used is thus 3D Coulomb renormalized for small interparticle distances. The idea is that at long range the electrons feel the $1/|x|$ interaction, but at short range they can avoid each other due to the finite thickness of the wire, which is mimicked by removing the divergence at $x = 0$. For example, a widely used effective quasi-1D interaction is obtained by integrating the 3D Coulomb interaction over normalized gaussians in two of the three spatial directions [64], modeling

harmonic confinement within a wire of thickness b ,

$$v_{\text{ee}}^{\text{wire}}(x) = \frac{1}{4\pi b^2} \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} dz \frac{e^{-\frac{1}{4b^2}(x^2+b^2)}}{\sqrt{x^2+y^2+z^2}} = \frac{\sqrt{\pi}}{2b} \exp\left(\frac{x^2}{4b^2}\right) \operatorname{erfc}\left(\frac{|x|}{2b}\right). \quad (4.101)$$

This interaction is finite at $x = 0$, where it has a cusp, behaves as $1/|x|$ for large x and it is convex for $x \geq 0$. Other popular quasi-1D interactions are the soft Coulomb and the regularized Coulomb,

$$v_{\text{ee}}^{\text{soft}}(x) = \frac{1}{\sqrt{x^2+a^2}}, \quad (4.102)$$

$$v_{\text{ee}}^{\text{reg}}(x) = \frac{1}{|x|+a}. \quad (4.103)$$

Notice, however, that the 1D maps of Seidl [123] are exact only for interactions (costs) that are convex for $x \geq 0$ [27]. This means that when using $v_{\text{ee}}^{\text{soft}}(x)$, which is concave for $x \in [0, a/\sqrt{2}]$, the Seidl maps are not guaranteed to yield the true minimizer, as illustrated, for example, in Fig. 2 of Ref. [71].

Numerical realizations of the 1D Seidl maps are reported in Refs. [71, 72, 100–102, 104]. The implementation of the maps directly follows from Sect. 4.3.11: given a density $\rho(x)$ on a grid, the cumulant function $F_\rho(x)$ is evaluated on the same grid, and its inverse $F_\rho^{-1}(x)$ is simply obtained by swapping the columns. The grid can be restored by using a spline interpolation for $F_\rho^{-1}(x)$, and the maps are readily obtained. Numerical issues can appear in regions where the density is close to zero, with $F_\rho^{-1}(x)$ raising extremely steeply. An alternative method to obtain the 1D maps without the need to construct $F_\rho^{-1}(x)$ is discussed in Ref. [72].

4.4.1.2 Spherically Symmetric Densities

For spherically symmetric densities the radial SGS maps (4.79)–(4.80) conjectured in [122, 127] have been implemented in Refs. [63, 127] for the 3D case using numerical densities for atoms from He to Ne, and in Ref. [105] for the 2D case, where the SCE functional has been combined self-consistently with Kohn–Sham DFT to describe electrons confined in a parabolic potential at low density.

The construction of the radial maps is implemented as in the 1D case. However, the computational complexity is now higher due to the evaluation of the reduced radial cost of Eq. (4.77), which requires an angular minimization for given radial distances. For the two-dimensional case treated in Ref. [105], where the number of relative angles to minimize was equal to $N-1$, the procedure has been the following. For an initial non-degenerate radial configuration and given initial starting angles, the quasi-Newton Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm was used

to find the closest local minimum. Then the radial position of the “first” electron was changed in small discrete steps, the radial positions of the remaining electrons were computed using the SGS maps, and the angles were optimized using the BFGS algorithm, with starting angles taken from the previous step. This procedure rests on the assumption that the optimal angles change continuously with the radial configuration. The starting angles for the initial radial configuration can be chosen by using simulated annealing as a global optimization strategy. It should be stressed that the angular minimization does not need to be performed for the whole set N_{grid} of radial grid points. In fact, the N radial distances are periodic, as each circular shell $r \in [a_i, a_{i+1}]$ (with $a_i = F_\rho^{-1}(i)$, $i \in \mathbb{N}$), corresponds to the same physical situation,[127] simply describing a permutation of the set of distances occurring in the first shell $r \in [0, a_1]$. Thus, by keeping track of the minimizing angles, and by readapting the grid in every circular shell, it is possible to do the angular minimization only N_{grid}/N times rather than N_{grid} times.

4.4.2 Methods Based on Linear Programming

Direct discretization of the SIL variational principle (4.25) yields a linear program, which is numerically tractable when $N = 2$.

4.4.2.1 The $N = 2$ Case

For two-electron systems in 3D with general density, Chen et al. [22] have implemented a method to directly solve the SIL variational principle via linear programming and extract the co-motion function and the SCE potential from the SIL solution. They used this approach to compute the co-motion function and the KS-SCE binding curve of the H_2 molecule (see Figs. 4.9 and 4.14).

One truncates \mathbb{R}^3 to a bounded domain, discretizes it into ℓ finite regions e_1, \dots, e_ℓ , and represents each element by a point \mathbf{a}_ℓ located at its barycenter. The single-particle density becomes a vector in \mathbb{R}^ℓ with components $\rho_\ell = \int_{e_\ell} \rho(\mathbf{r}) \, d\mathbf{r}$. The two-particle density π is represented by a matrix $\gamma = (\gamma_{ij}) \in \mathbb{R}^{\ell \times \ell}$ with $\gamma_{ij} = \int_{e_i} \int_{e_j} d\pi(\mathbf{r}_1, \mathbf{r}_2)$, and the interaction $V_{\text{ee}}(\mathbf{r}_1, \mathbf{r}_2)$ becomes a matrix $(c_{ij}) \in \mathbb{R}^{\ell \times \ell}$ with $c_{ij} = \frac{1}{|\mathbf{a}_i - \mathbf{a}_j|}$. The SIL problem (4.25) then becomes

$$\min_{\gamma \in \mathbb{R}^{\ell \times \ell}} \sum_{1 \leq i, j \leq \ell} c_{ij} \gamma_{ij} \quad (4.104)$$

$$\text{s/to } \sum_{j=1}^{\ell} \gamma_{ij} = \frac{\rho_i}{2}, \quad i = 1, \dots, \ell, \quad \sum_{i=1}^{\ell} \gamma_{ij} = \frac{\rho_j}{2}, \quad j \in \{1, \dots, \ell\},$$

$$\gamma_{ij} \geq 0.$$

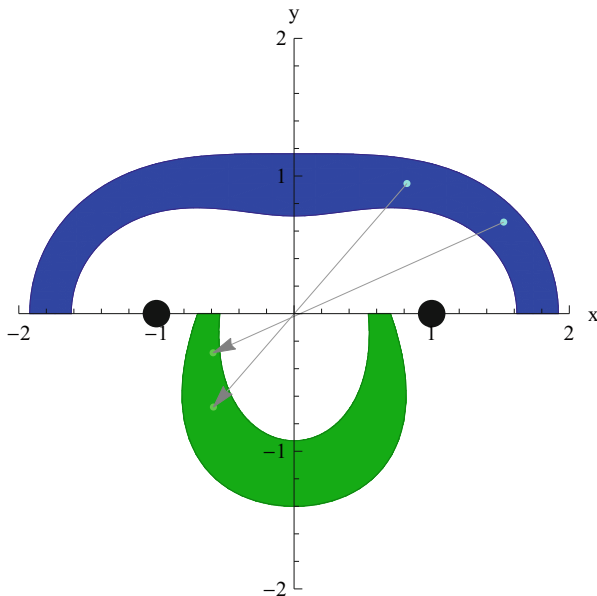


Fig. 4.9 Co-motion function for the H_2 molecule [22]. The blue region—corresponding to the points in a half plane adjacent to the molecular axis with density between 0.04 and 0.08—is mapped to the green region in the opposite half plane. The black dots indicate the positions of the nuclei

This is a standard linear programming problem of the form $\min_x f^T x$ subject to $Ax = b$, $x_k \geq 0$, where x is the vector containing the entries of γ . The solution can be obtained with a standard linear programming software (in [22], the authors used MOSEK). For a uniform discretization of the density, the number of degrees of freedom in the linear program would still be huge; instead an adaptive mesh was used in which all elements contain roughly the same amount of density, that is to say the mesh is much finer in the high-density region near the nuclei. (For automated generation of such a mesh, the finite element package PHK was used. See Chap. 10 by Dai and Zhou for more information about this package.) The solution to (4.104) entails an approximation to the co-motion function f at the barycenters $\{\mathbf{a}_i\}_{i=1}^\ell$, namely the barycenter of the image of \mathbf{a}_i under the transport plan γ :

$$f^{(\ell)}(\mathbf{a}_i) = \sum_{j=1}^{\ell} \frac{\gamma_{ij}}{\rho_i/2} \mathbf{a}_j \quad i \in \{1, \dots, \ell\}, \quad (4.105)$$

where γ_{ij} can be regarded as the mass transported from \mathbf{a}_i to \mathbf{a}_j and the normalization factor $\rho_i/2$ guarantees that the barycentric weights sum to 1. Since, for $N = 2$, the optimal N -point density π for the continuum problem is unique and of SCE form (see (4.41) and Theorem 4.4), it follows that if the discretization is sufficiently fine,

i.e. ℓ is large enough, $f^{(\ell)}$ is a good approximation to f . The resulting co-motion function for the H_2 molecule is depicted in Fig. 4.9.

4.4.2.2 The $N > 2$ Case and the Curse of Dimension

Since the above method uses a real-space discretization of the SIL variational principle whose unknown is the N -particle density on \mathbb{R}^{3N} , it is limited in practice to $N = 2$, to keep the number of computational degrees of freedom manageable. Indeed, for general N the N -particle density π must be represented by an order- N tensor $(\gamma_{i_1 \dots i_N}) \in \mathbb{R}^{\ell \times \dots \times \ell}$ with entries $\gamma_{i_1 \dots i_N} = \int_{e_{i_1} \times \dots \times e_{i_N}} d\pi(\mathbf{r}_1, \dots, \mathbf{r}_N)$. Since π can be assumed to be symmetric (see Remark 4.2), γ can be assumed to be symmetric under permutation of indices and Eq. (4.104) becomes

$$\begin{aligned} \min_{(\gamma_{i_1 \dots i_N}) \in \mathbb{R}^{\ell \times \dots \times \ell} \text{ symmetric}} & \sum_{1 \leq i_1, \dots, i_N \leq \ell} V_{ee}(a_{i_1}, \dots, a_{i_N}) \gamma_{i_1 \dots i_N} & (4.106) \\ \text{s/to} & \sum_{i_2, \dots, i_N=1}^{\ell} \gamma_{i_1 i_2 \dots i_N} = \frac{\rho_{i_1}}{N}, \quad i_1 \in \{1, \dots, \ell\}, \\ & \gamma_{i_1 \dots i_N} \geq 0. \end{aligned}$$

This is still a linear program, but in ℓ^N (or, using symmetry, $\binom{N+\ell-1}{\ell+1}$) variables. For rigorous hardness results on problems of the form (4.106) with pairwise interaction see [4, 52].

4.4.3 Methods Based on the Dual Formulation

Mendl and Lin [104] have implemented a method for solving the dual formulation of the SCE functional, Eqs. (4.55) and (4.53), and applied it to the Beryllium atom, a four-electron quantum wire in 1D, and a model trimer in 3D. In the 3D case, they parametrized the (unknown) Kantorovich potential by a pseudocharge,

$$v(\mathbf{r}) = \int \frac{m(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}',$$

with m given by a small number of Gaussians and satisfying $\int m = N - 1$ to account for the asymptotic behavior $v(\mathbf{r}) \sim (N - 1)/|\mathbf{r}|$ for large $|\mathbf{r}|$ (see (4.66)). They showed that the constrained maximization in (4.55), (4.53) is equivalent to a

nested pair of unconstrained optimizations,

$$V_{\text{ee}}^{\text{SCE}}[\rho] = \sup_v \left(\int v(\mathbf{r})\rho(\mathbf{r})\mathbf{d}\mathbf{r} + g[v] \right) \quad \text{with} \quad (4.107)$$

$$g[v] = \min_{(\mathbf{r}_1, \dots, \mathbf{r}_N) \in \mathbb{R}^{dN}} \left(V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N) - \sum_{i=1}^N v(\mathbf{r}_i) \right). \quad (4.108)$$

The inner optimization for given v was implemented by a quasi-Newton method and the outer optimization via a gradient-free simplex algorithm. For the Beryllium atom, using just two Gaussians for m resulted in a relative error of the SCE energy of only 1.6% compared to the SCE energy obtained via the SGS co-motion functions for radially symmetric densities [127] as described in Sect. 4.4.1.2. Also, the obtained SCE potential was in good agreement with the one based on the radial co-motion functions.

As the authors point out, this approach is in practice limited to small systems, because the inner optimization is high-dimensional, nonlinear, and highly degenerate for the optimal v (recall that the set of minimizers is typically d -dimensional), and the outer optimization (4.107) is nonlinear and nonsmooth, and hence unsuitable for numerical optimization over a large number of degrees of freedom.

4.4.4 Multi-Marginal Sinkhorn Algorithm

In optimal transport, a standard computational method [35] is to pass to the entropic regularization (in our case, problem (4.89)) and solve the latter via the Sinkhorn algorithm. This is a simple and robust algorithm which goes back to Sinkhorn in the context of estimating Markov transition matrices [131]; it was introduced into two-marginal optimal transport in [34] and generalized to several marginals in [8]. The multi-marginal Sinkhorn algorithm with Coulomb cost was implemented by Benamou et al. [9] (see also [110]) to compute the SCE energy and potential for the He and Li atoms.

The multi-marginal Sinkhorn algorithm goes as follows; we state it here in the continuous setting. One starts from the exact form (4.91) of the optimizer. One now allows the N entropic weight functions $a_j(\mathbf{r}_j) = a^\tau(\mathbf{r}_j)$ in this form to be different (so as to be able to update them one by one). One updates them iteratively so as to enforce the j -th marginal constraint, (4.92) for j :

$$a_j(\mathbf{r}_j) \int_{\mathbb{R}^{d(N-1)}} \prod_{i \neq j} a_i(\mathbf{r}_i) e^{-V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N)/\tau} \prod_{i \neq j} \mathbf{d}\mathbf{r}_i \stackrel{!}{=} \rho(\mathbf{r}_j)/N. \quad (4.109)$$

Solving for a_j yields an explicit formula for a_j in terms of the other a_i . Thus a single updating cycle consists of the N steps

$$a_j^{\text{new}}(\mathbf{r}_j) = \frac{\rho(\mathbf{r}_j)/N}{\left[\int_{\mathbb{R}^{d(N-1)}} \prod_{i < j} a_i^{\text{new}}(\mathbf{r}_i) \prod_{i > j} a_i^{\text{old}}(\mathbf{r}_i) \times e^{-V_{\text{ee}}(\mathbf{r}_1, \dots, \mathbf{r}_N)/\tau} \prod_{i \neq j} d\mathbf{r}_i \right]}, \quad j = 1, \dots, N. \quad (4.110)$$

One then repeats the cycle until convergence.

Convergence of the Sinkhorn algorithm is rigorously guaranteed under mild conditions on the interaction potential and the density (e.g., bounded potentials and $\rho \log \rho \in L^1$ are sufficient); see [131] for the discretized $N = 2$ case, [118] for the general $N = 2$ case, and [40] for $N \geq 2$. The (linear) rate of convergence for the Sinkhorn algorithm was obtained in [23, 46] in the $N = 2$ case, and in [15] for the multi-marginal Sinkhorn algorithm. For a two-electron example in dimension one computed with the Sinkhorn algorithm see Fig. 4.8.

In [9], Benamou et al. demonstrated that for the He atom (and the choice $\tau = 0.02$) the algorithm yields an accurate approximation to the SCE energy and the SCE potential compared to the (in this case rigorously justified) SGS map-based solution; the relative error of the potential in the L^∞ norm was only 0.4%. Moreover, for the Li atom the numerical Sinkhorn solution exhibited very good qualitative agreement with the SGS solution.

Some regularization is essential for the Sinkhorn approach. As τ approaches zero—so that the entropic regularization $V_{\text{ee}}^\tau[\rho]$ from (4.89) approaches the exact SCE functional (4.26)—the convergence speed of the algorithm also goes to zero (see e.g. [35, 50]), and numerical instabilities can appear associated with the extremely small order $e^{-1/\tau}$ of the integrand (see e.g. [8]).

The idea of regularization underlying the algorithm fits well into our DFT context as the optimal N -point density is smeared out anyway off the strongly interacting limit. However, a significant limitation from the point of view of DFT is the high-dimensionality of the integral in (4.109), (4.110). For a discretization of the one-body density by ℓ gridpoint values, the cost of a single integral evaluation for fixed \mathbf{r}_j is $O(\ell^{N-1})$, limiting the method to small N .

4.4.5 Towards Large N

Very recently, some promising methods have been proposed which should, at least in principle, be suitable for tackling the case of large N . These have been demonstrated to show good performance on one-dimensional test examples where the Seidl solution from Sect. 4.3.11 is available for comparison. At the time of writing, it has yet to be demonstrated that any of these methods is capable of accurately computing the SCE energy for large N in three dimensions.

4.4.5.1 Semidefinite Convex Relaxation

The starting point of this method, introduced by Khoo and Ying [82], is the fact that the SIL problem 4.25 can, due to the fact that V_{ee} is a two-body potential (4.3), be reformulated as a minimization over N -representable 2-point probability measures:

$$V_{ee}^{SCE}[\rho] = \min_{\substack{\Gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \\ \Gamma \text{ } N\text{-representable}, \Gamma \mapsto \rho}} \binom{N}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} w_{ee}(\mathbf{r} - \mathbf{r}') d\Gamma(\mathbf{r}, \mathbf{r}'). \quad (4.111)$$

Here a two-point probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ is called N -representable if it is the 2-marginal of a symmetric N -point probability measure on \mathbb{R}^{dN} . This two-body formulation of the SCE functional was introduced in [51], and is a direct adaptation of the well-known two-body reduced density matrix formulation [26] of the Rayleigh–Ritz variational principle (4.6) to the strongly correlated limit of DFT.

After discretization as described in Sect. 4.4.2.1, the two-point marginal becomes a matrix $\Gamma = (\Gamma_{ij}) \in \mathbb{R}^{\ell \times \ell}$, and N -representability means that Γ is obtained from some symmetric tensor $(\gamma_{i_1 \dots i_\ell}) \in \mathbb{R}^{\ell \times \dots \times \ell}$ with nonnegative entries which sum to one by $\Gamma_{i_1 i_2} = \sum_{i_3, \dots, i_N} \gamma_{i_1 i_2 i_3 \dots i_N}$.

The extreme points of the set of discrete N -representable 2-marginals have been determined explicitly [53] (see [16, 82] for generalizations to 3-marginals respectively general k -marginals).

Theorem 4.29 ([53]) *The set of extreme points of the set \mathcal{R}_2 of discrete N -representable 2-marginals is*

$$\mathcal{R}_2^{\text{ext}} = \left\{ \frac{N}{N-1} \lambda \lambda^T - \frac{1}{N-1} \text{diag}(\lambda) : \lambda \in \mathbb{R}^\ell, \lambda_i \geq 0 \forall i, \mathbf{1}^T \lambda = 1, \lambda_i \in \left\{0, \frac{1}{N}, \frac{2}{N}, \dots\right\} \right\}.$$

In particular, \mathcal{R}_2 is the convex hull of $\mathcal{R}_2^{\text{ext}}$.

Here $\mathbf{1}$ denotes the vector in \mathbb{R}^ℓ with all components equal to 1. The discretized problem is then

$$\begin{aligned} \min_{\Gamma} \sum_{1 \leq i, j \leq \ell} c_{ij} \Gamma_{ij} & \quad (4.112) \\ \text{s/to } \Gamma \in \mathcal{R}_2, \Gamma \mathbf{1} = \frac{\rho}{N}. & \end{aligned}$$

Khoo and Ying [82] introduced the following convex relaxation of this problem in which \mathcal{R}_2 is replaced by a slightly larger but simpler set:

$$\begin{aligned} \min_{\Gamma} \quad & \sum_{1 \leq i, j \leq \ell} c_{ij} \Gamma_{ij} & (4.113) \\ \text{s/to } \Gamma \in \widetilde{\mathcal{R}}_2 = \quad & \left\{ \frac{N}{N-1} \Lambda - \frac{1}{N-1} \text{diag}(\Lambda \mathbf{1}) : \Lambda_{ij} \geq 0 \forall i, j, \Lambda \geq 0, \mathbf{1}^T \Lambda \mathbf{1} = 1 \right\}, \\ & \Gamma \mathbf{1} = \frac{\rho}{N}. \end{aligned}$$

Here $\Lambda \geq 0$ means matrix positivity of Λ .

It is clear that $\widetilde{\mathcal{R}}_2 \supset \mathcal{R}_2$, since $\widetilde{\mathcal{R}}_2$ is convex and—by inspection—contains the set of extreme points of \mathcal{R}_2 given in Theorem 4.29. A theoretical argument in support of the approximation (4.113) is:

Theorem 4.30 ([82]) *The extreme points of the true set \mathcal{R}_2 of discrete N -representable 2-marginals are still extreme points of $\widetilde{\mathcal{R}}_2$.*

Intuitively this means that, at least near the extreme points of the exact set \mathcal{R}_2 of N -representable 2-marginals, the relaxation is very tight.

Viewed as a minimization over Λ , (4.112) is a semidefinite program (SDP), i.e. a problem of minimizing a linear cost subject to finitely many linear equalities or inequalities and a matrix positivity constraint. It has been implemented in [82] using a uniform grid and the large-scale SDP solver SDPNAL+. For 1D problems with $N = 8$, up to $\ell = 1600$ gridpoints, and different one-body densities, the solutions reported in [82] are in excellent qualitative agreement with the pair density of the exact Seidl solution. The relative energy error compared to the unapproximated discrete problem (4.111) is estimated to be of the order of 10^{-2} to 10^{-4} , depending on the choice of one-body density. Also, (4.113) is solved for 6 electrons in 2D with a Gaussian density on a 10×10 grid.

Khoo and Ying [82] also give a dual formulation of the SDP (4.113) which yields an approximation to the Kantorovich potential. For 1D test problems with 8 electrons and 200 gridpoints, a relative accuracy of 10^{-2} to 10^{-3} in the L^2 norm is reported compared to the exact potential obtained from the Seidl solution and Eq. (4.59).

4.4.5.2 Langevin Dynamics with Moment Constraints

This approach was proposed by Alfonsi et al. [1, 2]. The idea is to only discretize the density constraint, but not the N -point density, and then use a stochastic particle method to simulate the many-electron density. One performs a Galerkin (or “moment”) discretization of the marginal constraint (4.24) by requiring only a

fixed number M of integral constraints, of the form

$$\int_{\mathbb{R}^{Nd}} \varphi_m(\mathbf{r}_i) d\gamma(\mathbf{r}_1, \dots, \mathbf{r}_N) = \int_{\mathbb{R}^d} \varphi_m d\mu \quad \forall i = 1, \dots, N, \quad \forall m = 1, \dots, M, \tag{4.114}$$

where $\mu = \rho/N$ is the prescribed single-particle density and $\varphi_1, \dots, \varphi_M$ are suitable single-particle basis functions on \mathbb{R}^d . Moreover since the marginal constraint has been relaxed, one introduces a mild additional constraint on the class of admissible N -electron densities γ to prevent mass from escaping to infinity,

$$\int_{\mathbb{R}^{dN}} \sum_{i=1}^N \theta(|\mathbf{r}_i|) d\gamma(\mathbf{r}_1, \dots, \mathbf{r}_N) \leq A \tag{4.115}$$

for some nonnegative increasing function $\theta : [0, \infty) \rightarrow [0, \infty)$ with $\theta(r) \rightarrow \infty (r \rightarrow \infty)$ and some constant $A > 0$. The SIL problem (4.25) is now approximated by:

$$\text{Minimize } \int_{\mathbb{R}^{Nd}} V_{ee} d\gamma \text{ over } \gamma \in \mathcal{P}(\mathbb{R}^{Nd}) \text{ subject to (4.114) and (4.115).} \tag{4.116}$$

Under suitable assumptions on the basis functions, and for A chosen sufficiently large, the minimum value of (4.116) can be shown to converge to the SCE energy $V_{ee}^{\text{SCE}}[\rho]$ as the number M of basis functions tends to infinity [2]. The key property of (4.116) opening the door to numerical methods is the following.

Theorem 4.31 ([2]) *Assume $\mu \in \mathcal{P}(\mathbb{R}^d)$, and suppose that the basis functions $\varphi_1, \dots, \varphi_M : \mathbb{R}^d \rightarrow \mathbb{R}$ are continuous, belong to $L^1(d\mu)$, and satisfy the growth bound $|\varphi_m(\mathbf{r})| \leq \text{const} (1 + \theta(|\mathbf{r}|))^s$ for some $s \in (0, 1)$. Assume that $V_{ee} : \mathbb{R}^{dN} \rightarrow \mathbb{R} \cup \{+\infty\}$ is nonnegative and $\int V_{ee} d\gamma$ is finite for some γ satisfying (4.114), and that A is sufficiently large. Then there exists a minimizer of (4.116) of the form $d\gamma(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{v=1}^K \alpha_v S_N \delta(\mathbf{r}_1 - \mathbf{a}_1^{(v)}) \dots \delta(\mathbf{r}_N - \mathbf{a}_N^{(v)})$ for some $K \leq M + 2$, some coefficients $\alpha_v \geq 0$, and some $\mathbf{a}_i^{(v)} \in \mathbb{R}^d$.*

Thus a sparse ansatz for the many-electron density consisting of $K \leq M + 2$ symmetrized Dirac measures (where M is the number of constraints discretizing the marginal condition) is sufficient. This result generalizes Theorem 4.23 from discrete problems to semi-discrete problems with continuous state space and discretized marginal constraint.

In order to numerically solve (4.116), in [1] a stochastic particle method in continuous state space has been implemented. More precisely, the authors use constrained overdamped Langevin dynamics in the potential V_{ee} , which is a natural stochastic evolution equation for minimizing V_{ee} , applied to weighted sums of K symmetrized Dirac measures moving on the constraint manifold (4.114). For 5 electrons in a one-dimensional interval and the regularized Coulomb interaction (4.103)

with $a = 0.1$, up to $M = 40$ basis functions taken to be Legendre polynomials, and superpositions of up to $K = 10,000$ symmetrized Dirac measures, the method achieves good agreement with the Seidl solution described in Sect. 4.3.11. The implementation uses an iterative method to maintain the constraints (which are nonlinear in the particle positions), as well as judicious choices of the time steps, temperature profile, and numbers of symmetrized Diracs to balance accuracy and computational efficiency.

An attractive feature of this method besides its feasibility for large numbers of electrons is the fact that space is not discretized. In [1] simulations are reported for 100 electrons in three dimensions subject to 52 marginal constraints, again using superpositions of 10,000 symmetrized Dirac measures. At the time of writing, it remains an interesting open question to assess, in such situations, the accuracy of the model (4.116) and its numerical solutions.

4.4.5.3 Genetic Column Generation

This method was proposed recently by Friesecke et al. [52]. It directly solves the discretized SIL problem (4.106), by combining the sparse but exact quasi-SCE or quasi-Monge ansatz (see Theorem 4.23), the method of column generation from discrete optimization, and basic ideas from machine learning.

The idea is to alternate between solving the SIL problem on a small but otherwise unconstrained subset of the many-electron configuration space, and updating the subset based on the (primal and dual) SIL solution. Recall that after discretization, the many-electron density becomes a density γ on X^N , where $X = \{\mathbf{a}_1, \dots, \mathbf{a}_\ell\}$ is a set of discretization points (e.g., a grid) for the single-electron configuration space \mathbb{R}^d . One now starts from the quasi-SCE or quasi-Monge ansatz in the form (4.88), which suffices to solve the discrete SIL problem (4.106) exactly (see Theorem 4.23), but—for computational reasons—allows a slightly larger number of delta functions:

$$\gamma(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{v=1}^{\ell'} \alpha_v S_N \delta(\mathbf{r}_1 - \mathbf{r}_1^{(v)}) \dots \delta(\mathbf{r}_N - \mathbf{r}_N^{(v)}), \quad \ell \leq \ell' \leq \beta\ell. \quad (4.117)$$

Here the $\mathbf{r}^{(v)} = (\mathbf{r}_1^{(v)}, \dots, \mathbf{r}_N^{(v)})$ are arbitrary N -point configurations in X^N and $\beta > 1$ is a hyperparameter (taken to be 5 in [52]) which limits the number of N -point configurations to $O(\ell)$ instead of the naively required $O(\ell^N)$. To achieve a unique correspondence between symmetrized Diracs and N -point configurations one restricts the $\mathbf{r}^{(v)}$ to the sector $X_{\text{sym}}^N = \{(\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_N}) \in X^N : i_1 \leq \dots \leq i_N\}$, making the expansion coefficients α_v in (4.117) unique.

The ansatz (4.117) involves two sets of degrees of freedom, the subset $\Omega = \{\mathbf{r}^{(1)}, \dots, \mathbf{r}^{(\ell')}\}$ of the many-electron configuration space and the coefficient vector $(\alpha_1, \dots, \alpha_{\ell'})$, which are updated alternately. For fixed Ω , the coefficient vector is governed by the SIL problem (4.106) restricted to the ansatz (4.117), which reads,

using that $S_N \delta(\mathbf{r}_1 - \mathbf{r}_1^{(v)}) \dots \delta(\mathbf{r}_N - \mathbf{r}_N^{(v)})$ has single-particle density $\rho^{(v)}(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i^{(v)})$,

$$\min_{\alpha \in \mathbb{R}^{\ell'}} \sum_{\nu=1}^{\ell'} \alpha_{\nu} V_{\text{ee}}(\mathbf{r}^{(\nu)}) \quad \text{s/to} \quad \sum_{\nu=1}^{\ell'} \alpha_{\nu} \rho^{(\nu)}(\mathbf{a}_i) = \frac{\rho_i}{N}, \quad i \in \{1, \dots, \ell\}, \quad \alpha_{\nu} \geq 0. \quad (4.118)$$

This is just a small linear program with an $\ell \times O(\ell)$ constraint matrix. Updating the set Ω is done in a simple but subtle manner, as standard methods would incur the curse of dimension (see below). One also uses the dual problem

$$\max_{u: X \rightarrow \mathbb{R}} \sum_{i=1}^{\ell} u(\mathbf{a}_i) \rho_i \quad \text{s/to} \quad u(\mathbf{r}_1^{(v)}) + \dots + u(\mathbf{r}_N^{(v)}) \leq V_{\text{ee}}(\mathbf{r}^{(v)}) \quad \forall v \in \{1, \dots, \ell'\}, \quad (4.119)$$

whose solution u is an approximation to the Kantorovich potential.

An updating cycle in the genetic column generation (GenCol) method goes as follows:

1. Given a set $\Omega \subset X_{\text{sym}}^N$ of N -particle configurations, update the primal solution α and the dual solution u by solving (4.118) and (4.119).
2. Given the updates α^{new} and u^{new} , update Ω by the following genetic learning method:

- pick a random “parent” configuration $\mathbf{r}^{(v)} \in \Omega$ satisfying $\alpha_{\nu}^{\text{new}} > 0$;
 - create a random “child” $\mathbf{r}^* \in X_{\text{sym}}^N$ by moving one electron position to a nearest neighbor;
 - repeat these steps until $u^{\text{new}}(\mathbf{r}_1^*) + \dots + u^{\text{new}}(\mathbf{r}_N^*) > V_{\text{ee}}(\mathbf{r}^*)$ and set $\Omega^{\text{new}} = \Omega \cup \{\mathbf{r}^*\}$.
- (4.120)

Steps 1. and 2. are iterated until convergence, with the oldest configurations which do not contribute to the current optimal plan (i.e. satisfy $\alpha_{\nu}^{\text{new}} = 0$) being deleted from Ω whenever its size ℓ' exceeds the maximum allowed size $\beta\ell$.

The simple but powerful genetic learning aspect of the search rule in (4.120) is that only “successful” N -electron configurations in Ω (i.e. ones that contribute to the current optimal plan (4.117) with a nonzero coefficient α_{ν}) are allowed to bear offspring. Numerical observations and theoretical considerations show that this is essential for overcoming the curse of dimension. An unbiased random search of new configurations, or the updating step in the classical column generation method

of solving the so-called pricing problem,⁹ would merely turn the curse of dimension with respect to the size of the state space into a curse of dimension with respect to the number of search steps.

The rationale behind the acceptance criterion in (4.120) is that any new configuration \mathbf{r}^* satisfying it represents a constraint of the full dual problem (Eq. (4.119) with the $\mathbf{r}^{(v)}$ being replaced by all configurations in X_{sym}^N) which the current dual solution u^{new} violates. Adding this configuration to the set Ω “cuts off” u^{new} from the optimization domain of the dual problem, yielding a new dual solution and an energy decrease. For a rigorous justification see [52].

Figure 4.10, taken from [52], shows the solution of the SIL problem (4.25) computed by the GenCol algorithm for 10 electrons in a 1D interval discretized by 100 gridpoints. In this example, the grid spacing is normalized to 1, the density is taken to be $\rho(x) = \text{const}(0.2 + \sin^2(\frac{x}{\ell+1}))$, and the interaction is the soft Coulomb potential (4.102) with $a = 0.1$. With the initial set of many-electron configurations chosen randomly, the algorithm always found the exact Seidl solution (see Sect. 4.3.11) of the discretized problem to machine precision using less than 7000 iterations and less than 5 samples per iteration. This means that only a tiny fraction of the configuration space was accessed. The energy decreased steadily at an exponential rate.

Tests reported in [52] on larger 1D systems with up to $N = 30$ electrons on 120 grid points (corresponding to a space of N -point densities of dimension $\ell^N \approx 2.4 \times 10^{62}$) show only a slow polynomial growth in N of the number of iterations required to find the exact solution to machine precision, with the average number of samples needed per iteration to satisfy the acceptance criterion remaining approximately constant.

Apart from its simplicity and efficiency in high dimensions, attractive features of the genetic column generation method are that after discretization no further approximations are made (and the discrete SIL problem is solved accurately), and that the method also provides the Kantorovich potential for use within Kohn–Sham DFT.

Tests for accurately discretized three-dimensional densities are not yet available at the time of writing.

4.4.6 Approximations

As explained, there are presently no efficient algorithms to solve the SCE problem in an exact or very accurate way for the general three-dimensional case. In the usual spirit of DFT, several approximations for the functional $V_{\text{ee}}^{\text{SCE}}[\rho]$ have

⁹ which consists in our case in finding a configuration \mathbf{r}^* which maximizes the difference $u^{\text{new}}(\mathbf{r}_1^*) + \dots + u^{\text{new}}(\mathbf{r}_N^*) - V_{\text{ee}}(\mathbf{r}^*)$.

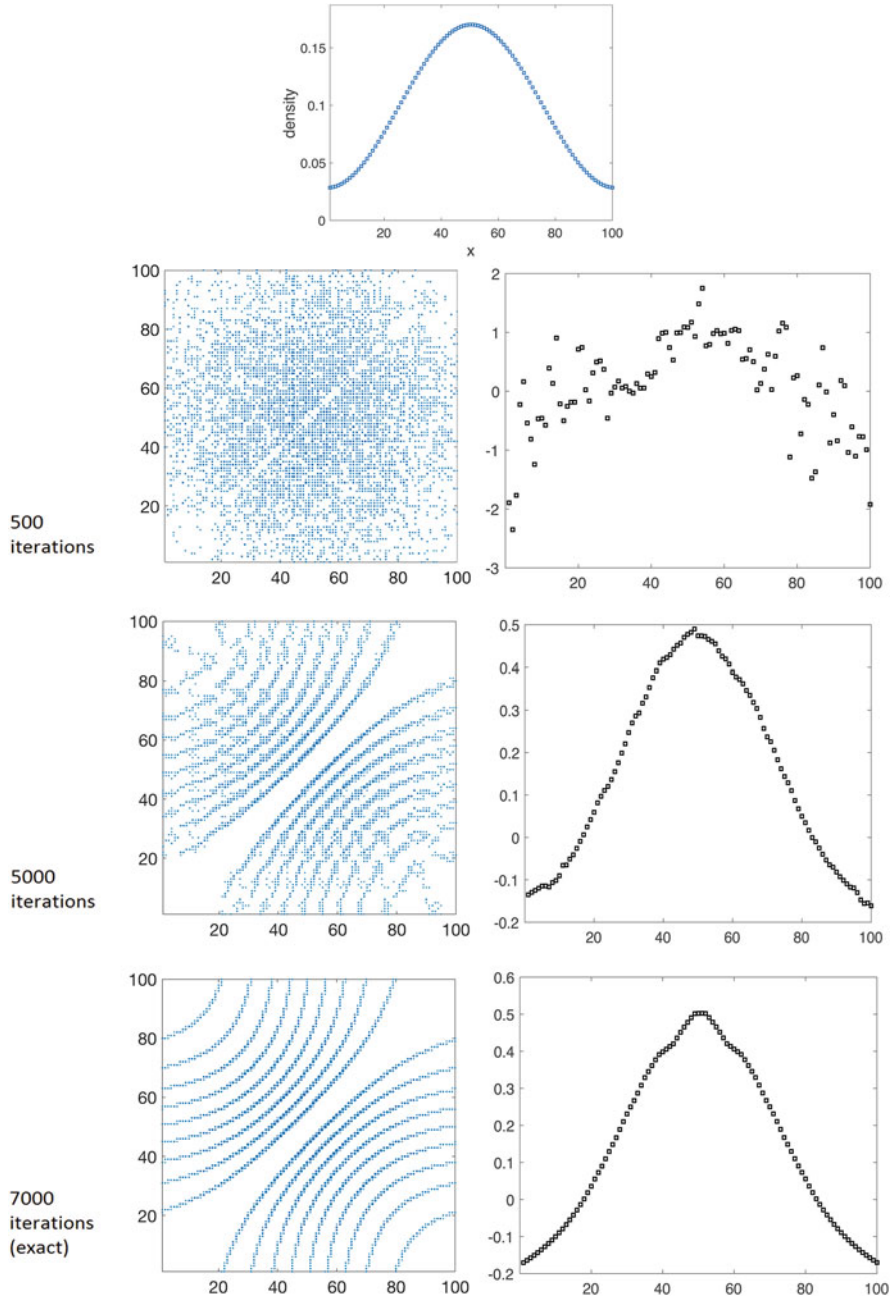


Fig. 4.10 Solution to the SIL problem (4.25) for 10 electrons in 1D with the GenCol algorithm [52]. *Top*: prescribed density. *Left*: Evolution under GenCol of the N -electron density from a random initial state, visualized via its two-point marginal (pair density). Gridpoints with nonzero values (i.e., “successful” configurations) are shown in blue, with larger markers indicating higher values. *Right*: Evolution of the Kantorovich potential. The final N -point density recovers Seidl’s SCE state for the discretized problem with machine precision

been proposed and used in combination with Kohn–Sham DFT. We review the approximations in this section, and their use within Kohn–Sham DFT in Sect. 4.5.

4.4.6.1 Gradient Expansion: Point-Charge-Plus-Continuum Model (PC)

The first gradient expansion approximation (GEA) for the indirect energy functional $W_\infty[\rho] = V_{\text{ee}}^{\text{SCE}}[\rho] - U[\rho]$ has been proposed by Seidl et al. [128], and it is called the point-charge-plus continuum (PC) model,

$$W_\infty^{\text{PC}}[\rho] = \int d\mathbf{r} \left[A \rho(\mathbf{r})^{4/3} + B \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})^{4/3}} \right], \quad (4.121)$$

where $A = -\frac{9}{10}(\frac{4\pi}{3})^{1/3}$ and $B = \frac{3}{350}(\frac{3}{4\pi})^{1/3}$. The model is built from the physical interpretation of $W_\infty[\rho]$ as the electrostatic energy of a system of perfectly correlated electrons with density ρ inside a classical background with the same charge density ρ of opposite sign [128]. Notice that the electrons are not allowed to relax in this fictitious external potential, as they are kept in the SCE state with the prescribed density. Only when the density is uniform is the energy of the SCE state the same as the one we would obtain by letting the electrons relax in the positive background external potential [96]. The idea of the PC model is that when the density is slowly varying the energy should be well approximated by surrounding each electron by a PC cell (given by the combined effect of the background and the remaining electrons) that neutralizes its charge and it is such that the electron plus its cell have zero dipole moment [128].

The PC approximation works rather well: for example, for the atomic densities from He to Ne, the values $W_\infty^{\text{PC}}[\rho]$ agree within 1% with the values obtained by using the radial co-motion functions (maps) described in Sect. 4.3.12, as shown in Table I of Ref. [127]. This is quite remarkable as, usually, gradient expansions for the exchange-correlation functionals fail in providing accurate quantitative results.

4.4.6.2 Generalized Gradient Approximations: The Modified PC Model

Although quantitatively accurate for the SIL energy, the main drawback of the PC model is that its functional derivative,

$$v_{\text{xc}}^{\text{PC}}(\mathbf{r}) = \frac{\delta W_\infty^{\text{PC}}[\rho]}{\delta \rho(\mathbf{r})}, \quad (4.122)$$

diverges to $-\infty$ in the tail of atomic and molecular densities [45], making a self-consistent Kohn–Sham calculation impossible. Moreover, the PC model fails for quasi-2D and quasi-1D systems [31].

To overcome these problems, Constantin [31] has proposed a generalized gradient approximation (GGA) for $W_\infty[\rho]$, called the modified PC model (mPC), which reads

$$W_\infty^{\text{mPC}}[\rho] = A \int d\mathbf{r} \rho(\mathbf{r})^{4/3} \frac{1 + a s(\mathbf{r})^2}{1 + (a + 0.14) s(\mathbf{r})^2},$$

$$s(\mathbf{r}) = \frac{|\nabla \rho(\mathbf{r})|}{2(3\pi)^{1/3} \rho(\mathbf{r})^{4/3}}, \quad (4.123)$$

where A has the same value as in the original PC model, and $a = 2$. This approximation is less accurate for the SIL of atomic densities with respect to the original PC model (with errors around 9–10%), but has the advantage of a well-behaved functional derivative, and of achieving a physical description of the crossover from three to two dimensions.

4.4.6.3 Approximations with Some Non-Locality: The Non-local Radius (NLR) and the Shell Model

The PC and mPC are semilocal approximations, while, as we have seen, the exact SIL physics has an extreme non-local dependence on the density. Approximations that retain some (albeit limited) non-locality are the non-local radius (NLR) [140] and the shell models [6]. Both approximations use as key ingredient the spherically averaged density $\tilde{\rho}(\mathbf{r}, u)$ around a given position \mathbf{r} , obtained by integrating out the angular dependence of \mathbf{u} ,

$$\tilde{\rho}(\mathbf{r}, u) = \int \rho(\mathbf{r} + \mathbf{u}) \frac{d\hat{\mathbf{u}}}{4\pi}, \quad (4.124)$$

and, in analogy with the SCE structure for spherical densities conjectured in Ref. [127] and illustrated in Sect. 4.3.12, its cumulant

$$N_e(\mathbf{r}, u) = \int_0^u 4\pi x^2 \tilde{\rho}(\mathbf{r}, x) dx. \quad (4.125)$$

In the NLR model [140] the functional $W_\infty[\rho]$ is approximated as

$$W_\infty^{\text{NLR}}[\rho] = - \int d\mathbf{r} \rho(\mathbf{r}) \int_0^{R(\mathbf{r})} 2\pi \tilde{\rho}(\mathbf{r}, u) u du, \quad (4.126)$$

where the radius $R(\mathbf{r})$ is defined by the condition that the underlying exchange-correlation hole be normalized:

$$N_e(\mathbf{r}, R(\mathbf{r})) = 1. \quad (4.127)$$

This simple approximation is less accurate than the PC and mPC models for the case of the uniform electron gas, giving a too high energy. For non-uniform densities, the NLR has the advantage, with respect to the PC and mPC models, of being exact for one-electron systems. For atomic densities, NLR makes errors, with respect to the SCE results of Ref. [127], of the order of 8–9% [140].

The shell model [6] substantially improves the NLR approximation, by making it exact for a uniform density, and reducing its error with respect to the SCE results for atomic densities by almost a factor of 10. While the NLR model approximates the exchange-correlation hole with a sphere depleting one electron from the spherically averaged density, the shell model adds a single positive oscillation, and reads

$$W_{\infty}^{\text{shell}}[\rho] = \int d\mathbf{r} \rho(\mathbf{r}) 2\pi \left(- \int_0^{u_s(\mathbf{r})} \tilde{\rho}(\mathbf{r}, u) u du + \int_{u_s(\mathbf{r})}^{u_c(\mathbf{r})} \tilde{\rho}(\mathbf{r}, u) u du \right), \quad (4.128)$$

where for all \mathbf{r} we have $u_s = 0.849488 u_c$, which is the condition needed to make the model exact for a uniform density. The value of $u_c(\mathbf{r})$ is then obtained again by the normalization condition,

$$2 N_e(\mathbf{r}, 0.849488 u_c(\mathbf{r})) - N_e(\mathbf{r}, u_c(\mathbf{r})) = 1. \quad (4.129)$$

4.5 Kohn–Sham Combined with the Strong-Interaction Limit

4.5.1 Kohn–Sham with the SCE Functional (KS SCE)

The Kohn–Sham scheme with the SCE functional (KS SCE) was first proposed and implemented in [100], and corresponds to a crude, but well-defined approximation for the HK functional,

$$F_{\text{KS SCE}}[\rho] = T_s[\rho] + V_{\text{ee}}^{\text{SCE}}[\rho], \quad (4.130)$$

in which we replace the minimum of the sum of kinetic energy and electron-electron repulsion at fixed density, with the sum of the two minima. As such, the KS SCE will always provide a lower bound for the HK functional. When implemented self-consistently, the KS SCE scheme yields the usual KS equations with the Hartree-exchange-correlation potential given by the SCE or Kantorovich potential (written below for simplicity for a closed-shell system),

$$-\frac{1}{2} \nabla^2 \varphi_i(\mathbf{r}) + (v_{\text{SCE}}(\mathbf{r}, [\rho]) + v_{\text{ne}}(\mathbf{r})) \varphi_i(\mathbf{r}) = \epsilon_i \varphi_i(\mathbf{r}),$$

$$\rho(\mathbf{r}) = 2 \sum_{i=1}^{N/2} |\varphi_i(\mathbf{r})|^2, \quad (4.131)$$

where the SCE potential is equal to

$$v_{\text{SCE}}(\mathbf{r}, [\rho]) = u(\mathbf{r}, [\rho]) + C[\rho], \quad (4.132)$$

with $u(\mathbf{r}, [\rho])$ the maximizer in Eq. (4.55), and the constant $C[\rho]$ a shift that ensures $\lim_{\mathbf{r} \rightarrow \infty} v_{\text{SCE}}(\mathbf{r}) = 0$, see Eq. (4.64). This shift is the same, in the $\lambda \rightarrow \infty$ limit of the density-fixed adiabatic connection, as the one introduced by Levy and Zahariev [94, 139]. If we want to compute the ground-state density and the ground state energy only, one could better work with u instead of v_{SCE} , as with the former the energy becomes simply [22, 94, 139] the sum of the occupied orbital energies, $E_0 = 2 \sum_{i=1}^{N/2} \epsilon_i$. The shift is needed if we want to estimate the ionisation potential $I = E_0^{N-1} - E_0^N$ from the highest occupied molecular orbital energy (HOMO), as $I = -\epsilon_{N/2}$ holds only when the exchange-correlation potential goes to zero far from the barycentre of nuclear charge [3, 93]. For further discussion of this point see Chap. 1 by Toulouse in this volume.

4.5.1.1 1D Case

The self-consistent KS SCE equations have been solved for 1D systems with the interaction $v_{\text{ee}}^{\text{wire}}(x)$ of Eq. (4.101) when the external potential is harmonic, $v_{\text{ne}}(x) = \frac{1}{2}\omega^2 x^2$, [72, 100, 101], and with the soft Coulomb interaction $v_{\text{ee}}^{\text{soft}}(x)$ of Eq. (4.102) for model 1D atoms and molecules with ‘nuclei’ that attract the electrons with the same soft Coulomb potential [102]. At each KS iteration, the 1D co-motion functions [123] were computed numerically as explained in Sect. 4.4.1.1, and the potential $v_{\text{SCE}}(x, [\rho])$ was obtained by simply integrating the force equation

$$v'_{\text{SCE}}(x, [\rho]) = \sum_{i=2}^N w'(|x - f_i(x)|) \text{sgn}(x - f_i(x)), \quad (4.133)$$

with boundary condition $v_{\text{SCE}}(x \rightarrow \pm\infty, [\rho]) = 0$, and where $w(x)$ is the chosen 1D interaction (wire or soft Coulomb, see Sect. 4.4.1.1). In addition, at low density the highest occupied KS SCE eigenvalue gives a very accurate ionization energy of the system [101].

Harmonic External Potential

In Fig. 4.11 we show the self-consistent KS SCE densities for $N = 4$ electrons interacting with $v_{\text{ee}}^{\text{wire}}(x)$ of Eq. (4.101) when the external potential is harmonic, using scaled units in terms of $L = 2\omega^{-1/2}$, compared with accurate many-body results from configuration interaction (CI) and with KS within the local density approximation (LDA), provided for this interaction in Ref. [18]. We see that, as the system is driven to low density by reducing the strength of the harmonic

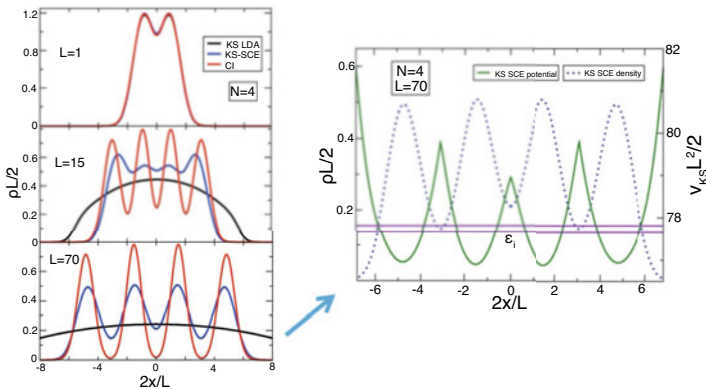


Fig. 4.11 Left: self-consistent KS SCE densities for $N = 4$ electrons interacting with $v_{\text{ec}}^{\text{wire}}(x)$ of Eq. (4.101) when the external potential is harmonic, $v_{\text{ne}}(x) = \frac{1}{2}\omega^2 x^2$, compared with very accurate configuration interaction results (CI) and with KS LDA. Right: the total KS potential at self consistency, $v_{\text{KS}} = v_{\text{ne}} + v_{\text{SCE}}$, for the most correlated case. The horizontal lines are the two eigenvalues of the occupied KS SCE orbitals. Results are in scaled units, where $L = 2\omega^{-1/2}$, and are taken from Ref. [101]

confinement (large L), the exact many-body solution undergoes a so-called “ $2k_F \rightarrow 4k_F$ ” transition, in which the number of peaks in the density is doubled. At high density, in fact, the number of peaks is dictated by the number of occupied orbitals, $N/2$ for a closed shell system. At low density, we have an incipient Wigner molecular structure, in which the electrons are well separated. Notice that with the Coulomb interaction this Wigner molecular phase exhibits different properties than the simpler case of very short-range interactions, in which the physics can be captured by making the system spin-polarized (i.e. by occupying N orbitals instead of $N/2$). This is clearly illustrated in Ref. [142].

It is well-known that the local and semilocal approximations to the XC functional, as well as exact exchange, are not able to capture this “ $2k_F \rightarrow 4k_F$ ” transition [134, 135] without introducing artificial symmetry breaking. This is also clearly shown by the KS LDA results of Fig.4.11, which become very close in this limit to the Thomas–Fermi result (minus the external potential in the classically allowed region) predicting a too delocalized density. The KS SCE self-consistent density, although not quantitatively very accurate, has the correct qualitative behavior, with two peaks at high density and four at low density, and with the correct extension. The KS SCE HOMO energy is also very close to the exact many-body ionisation potential [100, 101]. In the right panel of Fig.4.11 we show the total KS potential at self consistency, $v_{\text{KS}} = v_{\text{ne}} + v_{\text{SCE}}$, for the most correlated case. The horizontal lines are the two occupied KS SCE eigenvalues. We see that the SCE functional is able to self-consistently build barriers that create classically forbidden regions inside the harmonic trap. Classically forbidden regions for the KS

orbitals created by the Hartree-exchange-correlation potential seem to play a crucial role in describing strong correlation within KS DFT [12, 75, 144].

Model 1D Chemistry with Soft Coulomb Potential

In Ref. [102] the KS SCE method has been tested for model chemical systems in 1D, consisting of “nuclei” and electrons attracting each other with the soft-Coulomb potential (for the use of these 1D models to test DFT approximations, see also Refs. [74, 141]). While in the harmonic external potential we can drive the system to low density where the SCE becomes a very good approximation to the exact KS exchange-correlation functional, chemical systems (bound by the Coulomb external potential) are never in this regime. For this reason, KS SCE does not in general yield accurate results, with total energies that are way too low. An exception seems to be the good agreement between the eigenvalue of the highest occupied KS SCE orbital and the many-body chemical potential, as shown in Table 2 of Ref. [102].

4.5.1.2 2D Case

The circularly-symmetric 2D case of electrons interacting with the $1/r$ repulsion in the harmonic external potential has been studied with KS SCE in Ref. [105], using the SGS radial co-motion functions and the reduced radial cost of Eq. (4.78) implemented as described in Sect. 4.4.1.2. As in 1D, the aim is to model electrons strongly confined in one direction, found, for example, at the interface of semiconductor heterostructures. For this reason, the interaction remains the same as the 3D Coulomb one.

As discussed in Sect. 4.3.12, the SGS state defined by (4.79)–(4.80) is not guaranteed to yield the absolute minimum for the electron-electron interaction in a given radial density $\rho(r)$. Nonetheless, it can be proven [125] that, for a spherically-symmetric density, if we reduce the admissible class of maps \mathcal{T}_ρ in the SCE functional (4.37) to a class $\mathcal{T}_\rho^{\text{SGS}} \subset \mathcal{T}_\rho$ of maps given by the SGS ansatz defined in Eqs. (4.79)–(4.80) as an approximation for $V_{\text{ee}}^{\text{SCE}}[\rho]$, even when the SGS maps are not optimal the functional derivative of this approximate $V_{\text{ee}}^{\text{SCE}}[\rho]$ with respect to $\rho(r)$ still satisfies the force equation (written using the notation of Eq. (4.59)),

$$\nabla v_{\text{SGS}}(\mathbf{r}) = - \sum_{i=2}^N \frac{\mathbf{r} - \mathbf{f}_i^{\text{SGS}}(\mathbf{r})}{|\mathbf{r} - \mathbf{f}_i^{\text{SGS}}(\mathbf{r})|^3}, \quad (4.134)$$

which we can integrate to obtain a potential $v_{\text{SGS}}(r)$. In other words, the SGS maps provide a well-defined approximation to the exact SCE functional, with an easy to evaluate functional derivative, which, in turn, can be used in the KS equations.

In Fig. 4.12 we show the resulting KS SCE self-consistent radial density for $N = 3$ electrons for two low-density cases, compared with accurate Quantum

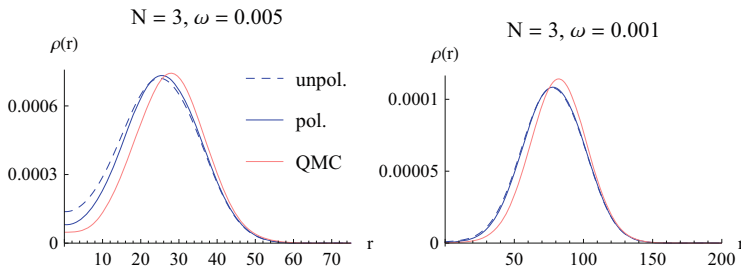


Fig. 4.12 Self-consistent radial KS SCE densities $\rho(r)$ for $N = 3$ electrons in 2D, with external potential $v_{\text{ne}}(r) = \frac{1}{2}\omega^2 r^2$ [105], compared with accurate Quantum Monte Carlo (QMC) results [61, 73]. The KS SCE densities are shown for both the unpolarized (2 orbitals, of which only the lowest is doubly occupied) and spin polarized (3 different singly occupied KS orbitals) case

Monte Carlo (QMC) results from Refs. [61, 73]. The KS SCE calculations have been done for both the unpolarized case (2 KS orbitals, of which only the lowest is doubly occupied) and the spin-polarized case (3 different singly occupied KS orbitals). We see that the KS SCE densities are very close to the QMC ones, predicting the right shell structure with one peak. Total energies are in agreement with QMC within $\sim 4 - 6\%$ [105]. Notice that at such low densities it is very hard to even obtain converged results using KS with the local-spin density (LSD) approximation. We thus see that even if the SGS maps are not optimal for these densities (see [125]), they yield very good results when used in the self-consistent KS equations at low density. However, we have to mention that QMC predicts that at such small ω 's the ground state is spin-polarized, while in KS SCE the unpolarized case always yields the lowest energy, due to the lack of any spin dependence in the SCE functional.

Figure 4.13 shows the self-consistent KS SCE total potential and density for $N = 10$ electrons (spin unpolarized) [105]. The green dashed curve is the energy of the highest occupied KS orbital. We clearly see, as in the 1D case of Fig. 4.11, that when the system is driven to low-density (small ω case), KS SCE is able to self-consistently create classically forbidden regions inside the trap.

4.5.1.3 3D Case

KS SCE has been tested on the anions of the He isoelectronic series [106] and on the dissociation curve of the H_2 molecule [22].

Anions of the He Isoelectronic Series

In this case, i.e., $N = 2$ electrons with $v_{\text{ne}}(r) = -\frac{Z}{r}$, where Z is lowered until the system can no longer bind two particles, the co-motion function and the SCE potential are simply built following the original work of Seidl [123] (see

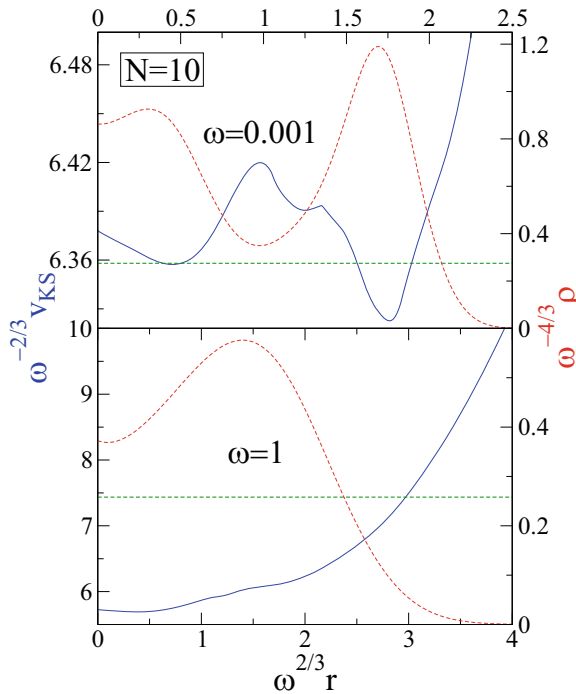


Fig. 4.13 Self-consistent KS SCE radial potential $v_{\text{KS}}(r) = v_{\text{SCE}}(r) + \frac{1}{2}\omega^2 r^2$ (blue solid line) and radial densities (red dashed line) for a strongly and weakly correlated case (top and bottom, respectively) of a 2D system composed of $N = 10$ electrons inside a circularly symmetric harmonic trap. The green dashed horizontal lines correspond to the energies of the highest occupied KS orbital. Notice the presence of classically forbidden regions inside the trap in the strongly correlated case ($\omega = 0.001$)

Example 4.20), which is a special case of the SGS maps. While very accurate wavefunction results predict [47] that one electron is lost by the system at a critical nuclear charge $Z_{\text{crit}}^{\text{exact}} \approx 0.911$, KS SCE binds two electrons down to $Z_{\text{crit}}^{\text{KS SCE}} \approx 0.7307$ [106]. This is because in the KS SCE case the two electrons can get much closer to the nucleus by perfectly avoiding each other, without raising the kinetic energy too much, which is only treated within KS.

The H₂ Molecule

The dissociation curve of the H₂ molecule has been computed within KS SCE in Ref. [22]. The result is shown in Fig. 4.14. To compute the self-consistent density and energy, an accurate adaptive three-dimensional finite element discretization was

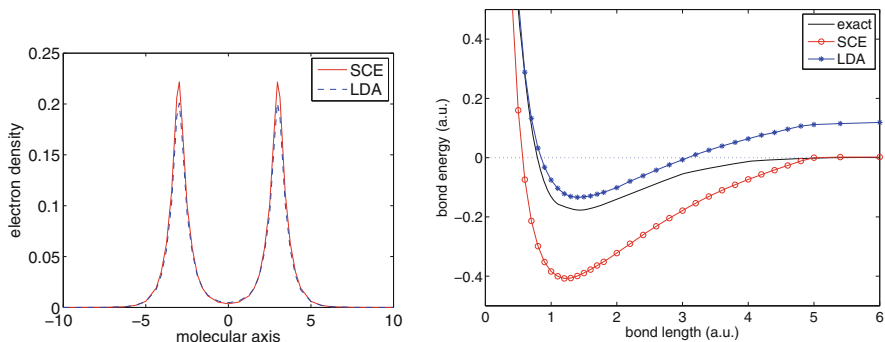


Fig. 4.14 *Right:* Dissociation curve of H_2 in KS SCE [22], that is, energy of H_2 minus twice the energy of the isolated H atom. For comparison, the KS LDA curve computed on the same mesh and the exact curve from Ref. [86] are also shown. Note that KS SCE, unlike the local density approximation, dissociates H_2 correctly. *Left:* Corresponding self-consistent KS SCE density and KS LDA density near dissociation

used and the SIL problem was solved using linear programming, as described in Sect. 4.4.2.1. The co-motion function for H_2 was then obtained from the SIL density via Eq. (4.105), and the SCE potential via the force Eqs. (4.59) and (4.64).

Not surprisingly, KS SCE predicts a binding energy that is way too low. A remarkable feature, though, is the ability of KS SCE to correctly dissociate the H_2 molecule, i.e., the molecular energy tends to twice the energy of the isolated H atom as the internuclear distance R becomes very large (see [22] for a rigorous proof). Local and semilocal approximations to the XC functionals are unable to do that, and exact exchange (or Hartree–Fock) perform even worse, unless we allow spin-symmetry breaking. Indeed, the extremely stretched H_2 molecule is often regarded as a severe test for XC functionals to check whether they are able to describe strong (or “static”) correlation [24].

Although the SCE functional yields the exact energy when $R \rightarrow \infty$, we see that at large but finite R the KS SCE curve immediately start to deviate from the exact one. We can understand this error by making the following simple analysis. With the internuclear vector \mathbf{R} directed along the x -axis, we can expand the electron–electron interaction at large R , which, without considering one-body terms and neglecting higher orders in R^{-1} yields

$$\begin{aligned} & \frac{1}{\sqrt{(x_1 - x_2 - R)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}} \\ & \sim \frac{2(x_1 - x_2)^2 - (y_1 - y_2)^2 - (z_1 - z_2)^2}{R^3}, \end{aligned} \quad (4.135)$$

where the origins of \mathbf{r}_1 and \mathbf{r}_2 are placed on their respective nuclei. The SCE functional for large R then corresponds to the minimization of this interaction at

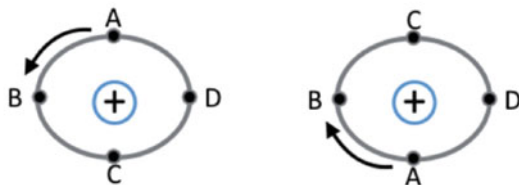


Fig. 4.15 When the distance R between the two atoms of the H_2 molecule gets very large, the optimal map describes the physics of perfectly coupled dipoles. The figure shows four pairs of electronic positions $\{\mathbf{r}, \mathbf{f}(\mathbf{r})\}$, labeled with the same letter A,B,C,D, with respect to the two positive nuclei

fixed one-body density (hence, the neglect of one-body terms that will not affect the minimizer). The SCE problem in this limit reduces then to the attractive harmonic cost [56] in the bond (x) direction and to the repulsive harmonic cost [42] in the two directions perpendicular to the bond axis. For large R , the optimal map will then approach the solution

$$f_x(\mathbf{r}) = x, \quad f_y(\mathbf{r}) = -y, \quad f_z(\mathbf{r}) = -z, \quad (4.136)$$

which corresponds to perfectly coupled dipoles (see Fig. 4.15). Such maps will give a finite (negative) expectation value for the r.h.s. of Eq. (4.135) even when the total density of the molecule is given by the sum of two spherical atomic densities, yielding an interaction energy that is too attractive, decaying as $\sim R^{-3}$ instead of the exact $\sim R^{-6}$. What is missing in the KS SCE approach is the raising in kinetic energy associated with the perfectly correlated dipoles of Fig. 4.15. A strategy to include the raise in kinetic energy in this asymptotic large- R regime is described in Ref. [88].

4.5.2 Interaction Strength Interpolation (ISI) Functionals

Another way to use the SIL in KS DFT is the interaction strength interpolation (ISI) construction, originally proposed by Seidl, Perdew and Levy (SPL) [130]. ISI is essentially the extension to non-uniform densities of Wigner's original idea [143] of approximating the energy of the uniform electron gas by interpolating between its high- and low-density asymptotics, which, by scaling, correspond to the weak- and strong-interaction limits, respectively.

The starting point is to use the Hellmann–Feynman theorem to write the exchange-correlation energy as an integral over the coupling-strength parameter λ of (4.17)¹⁰

$$E_{\text{xc}}[\rho] = \int_0^1 W_\lambda[\rho], \quad (4.137)$$

where

$$W_\lambda[\rho] = V_{\text{ee}}[\psi_\lambda[\rho]] - U[\rho], \quad (4.138)$$

with $\psi_\lambda[\rho]$ the minimizer in (4.17). The idea is then to construct approximations for the adiabatic connection integrand $W_\lambda[\rho]$ by interpolating between the $\lambda \rightarrow 0$ asymptotic expansion,

$$W_{\lambda \rightarrow 0} \sim E_x + 2\lambda E_c^{\text{GL2}} + \dots, \quad (4.139)$$

with E_x the exchange energy and E_c^{GL2} the second-order Görling-Levy perturbation theory correlation energy [69], and the large- λ limit provided by the SIL, and possibly by the conjectured next leading term of Eq. (4.47),

$$W_{\lambda \rightarrow \infty} \sim W_\infty[\rho] + \frac{F^{\text{ZPE}}[\rho]}{2\sqrt{\lambda}} + \dots. \quad (4.140)$$

For example, SPL [130] proposed the following simple form to interpolate between the two limits, without using the term with $F^{\text{ZPE}}[\rho]$:

$$W_\lambda^{\text{SPL}} = W_\infty + \frac{E_x - W_\infty}{\sqrt{1 + 2\lambda\chi}}, \quad (4.141)$$

with

$$\chi = \frac{2E_c^{\text{GL2}}}{W_\infty - E_x}. \quad (4.142)$$

The SPL XC functional then reads

$$E_{\text{xc}}^{\text{SPL}} = (E_x - W_\infty) \left[\frac{\sqrt{1 + 2\chi} - 1 - \chi}{\chi} \right] + E_x. \quad (4.143)$$

¹⁰ For further discussion see also Chap. 1 by Toulouse in this volume.

Several other interpolating functions that may or may not include $F^{\text{ZPE}}[\rho]$ have been proposed in the literature, [43, 68, 98, 128, 129] and are reported, for example, in the appendix of Ref. [87].

4.5.2.1 Global Interpolations

Interpolations such as the one of Eq. (4.141) have been implemented and tested on several chemical systems by using for the $\lambda \rightarrow \infty$ limit the PC model of Sect. 4.4.6.1 (and its extension [128] to $F^{\text{ZPE}}[\rho]$ when needed). In a practical calculation, KS orbitals with a given approximate semilocal or hybrid functional are used to compute the density ρ , the exchange energy E_x , and the second-order energy E_c^{GL2} , which are then fed into formulas such as (4.143) to obtain improved energies. The result is thus dependent on the chosen starting approximate functional used to generate the KS orbitals.

A basic problem of these *global* (in the sense that they are done on quantities that have been already integrated over all space) interpolations is the violation of size-consistency, i.e., if we take two different systems A and B that do not interact with each other, it is easy to verify from (4.143) that, in general,

$$E_{\text{xc}}^{\text{SPL}}(A + B) \neq E_{\text{xc}}^{\text{SPL}}(A) + E_{\text{xc}}^{\text{SPL}}(B), \quad (4.144)$$

an issue shared by all the other interpolation formulas proposed in the literature [25]. Notice that size-consistency of approximate electronic-structure methods is a very delicate issue when A and/or B have a degenerate ground state [65, 120]. Here we stress that even when degeneracy is not present, the fact that the input ingredients (E_x , E_c^{GL2} and W_∞) enter in a non-linear way in the ISI formulas introduces anyway a size-consistency error. However, this error can be easily corrected [137]. In fact, the reason why size-consistency is crucial in chemistry is that we are interested in interaction energies rather than total energies. All we need to do is to set the limit of a molecular dissociation curve (when A and B are infinitely far apart) at the value given by the left-hand side of Eq. (4.144) rather than the one given by the right-hand side. Notice that both sides of this equation can be evaluated at exactly the same computational cost, as all that is needed is the input ingredients of the fragments A and B [137]. With this size-consistency correction it is possible to extract meaningful interaction energies from the ISI functionals [137].

The ISI functionals have been tested on several chemical data sets and systems [44, 62, 137]. They have been found to work reasonably well for interaction energies (especially of non-covalent systems) when Hartree–Fock orbitals (rather than KS ones) are used as input. This observation has triggered the study of the strong-interaction limit in Hartree–Fock theory [37, 126], which, in turn, has led to new interpolation schemes in this framework which are able to give very accurate results for a large variety of non-covalent interaction energies, ranging from small to medium-large systems [36].

If one wants to overcome the dependence on the input orbitals, one should evaluate the energy using the ISI functionals within a fully self-consistent KS scheme. For this, their functional derivative with respect to the density is needed, which is challenging due to the presence of second-order perturbation theory. Nonetheless, first attempts in the computation of the ISI functional derivatives have been reported in Refs. [45, 132], and self-consistent calculations are likely to appear soon.

4.5.2.2 Local Interpolations

Another possibility is to build the interpolations *locally*, in each point of space, by defining an energy density $w_\lambda(\mathbf{r}; [\rho])$ for the coupling-constant integrand $W_\lambda[\rho]$ of Eq. (4.138), writing $E_{xc}[\rho]$ as

$$E_{xc}[\rho] = \int d\mathbf{r} \rho(\mathbf{r}) \int_0^1 w_\lambda(\mathbf{r}; [\rho]) d\lambda. \quad (4.145)$$

Energy densities are obviously not uniquely defined, and the only important requirement here is to use local quantities defined in the same way at weak and strong coupling. Some different choices for energy densities in the λ -interpolation context have been analyzed in [139], where it has been found that the electrostatic potential of the exchange-correlation hole¹¹ $h_{xc}^\lambda(\mathbf{r}_1, \mathbf{r}_2)$ seems to be the most suitable ,

$$w_\lambda(\mathbf{r}) = \frac{1}{2} \int \frac{h_{xc}^\lambda(\mathbf{r}, \mathbf{r}_2)}{|\mathbf{r} - \mathbf{r}_2|} d\mathbf{r}_2, \quad (4.146)$$

where $h_{xc}^\lambda(\mathbf{r}_1, \mathbf{r}_2)$ is defined in terms of the pair-density $P_2^\lambda(\mathbf{r}_1, \mathbf{r}_2)$ and the density ρ ,

$$h_{xc}^\lambda(\mathbf{r}_1, \mathbf{r}_2) = \frac{P_2^\lambda(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} - \rho(\mathbf{r}_2), \quad (4.147)$$

with P_2^λ obtained from $\Psi_\lambda[\rho]$,

$$P_2^\lambda(\mathbf{r}, \mathbf{r}') = \quad (4.148)$$

$$N(N-1) \sum_{\sigma, \sigma', \sigma_3, \dots, \sigma_N} \int |\Psi_\lambda(\mathbf{r}\sigma, \mathbf{r}'\sigma', \mathbf{r}_3, \sigma_3, \dots, \mathbf{r}_N, \sigma_N)|^2 d\mathbf{r}_3 \dots d\mathbf{r}_N.$$

¹¹ For further discussion of this point, see Chap. 1 by Toulouse in this volume.

Local interpolations within this definition have been analysed and tested in Refs. [87, 138] on small systems, with mixed results.

4.6 Appendix: Kantorovich Duality

The dual construction of the SCE functional and potential (see Theorem 4.13) relies on Kantorovich duality. In this Appendix we give a precise mathematical statement of Kantorovich duality for multi-marginal optimal transport, and show how it implies Theorem 4.13 (1).

Recall the general Kantorovich optimal transport problem introduced in Sect. 4.3.6: for given marginal measures μ_1, \dots, μ_N defined on closed subsets X_1, \dots, X_N of \mathbb{R}^d , minimize a cost functional

$$C[\gamma] = \int_{X_1 \times \dots \times X_N} c(\mathbf{r}_1, \dots, \mathbf{r}_N) d\Pi(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

over probability measures Π on the product space $X_1 \times \dots \times X_N$ subject to the marginal constraints

$$\int_{X_1 \times \dots \times X_{i-1} \times A_i \times X_{i+1} \times \dots \times X_N} d\Pi = \int_{A_i} d\mu_i \text{ for all measurable sets } A_i \subseteq X_i$$

and all $i \in \{1, \dots, N\}$.

Here $c : X_1 \times \dots \times X_N \rightarrow \mathbb{R} \cup \{+\infty\}$ is a given measurable cost function.

This problem is related to a certain dual variational problem: maximize the functional

$$J[u_1, \dots, u_N] = \sum_{i=1}^N \int_{X_i} u_i d\mu_i$$

over potentials $u_i : X_i \rightarrow \mathbb{R}$ ($i = 1, \dots, N$) which must satisfy the pointwise constraint

$$\sum_{i=1}^N u_i(\mathbf{r}_i) \leq c(\mathbf{r}_1, \dots, \mathbf{r}_N) \quad \forall (\mathbf{r}_1, \dots, \mathbf{r}_N) \in X_1 \times \dots \times X_N. \quad (4.149)$$

The following nontrivial statement, taken from the recent textbook [50], summarizes what is known in \mathbb{R}^d , and is general enough to cover the Coulomb cost.

Theorem 4.32 (Kantorovich Duality) *For given probability measures μ_1, \dots, μ_N defined on closed subsets X_1, \dots, X_N of \mathbb{R}^d , provided the cost function*

$c : X = X_1 \times \dots \times X_N \rightarrow \mathbb{R} \cup \{+\infty\}$ is bounded from below and lower semi-continuous and the optimal cost is finite,

$$\inf_{\substack{\Pi \in \mathcal{P}(X) \\ \gamma \mapsto \mu_1, \dots, \mu_N}} \int_X c \, d\Pi = \sup_{(u_1, \dots, u_N) \in \mathcal{A}(c)} \sum_{i=1}^N \int_{X_i} u_i \, d\mu_i, \tag{4.150}$$

where $\mathcal{A}(c)$ is any of the following increasingly general sets of admissible potentials:

- (1) $\mathcal{A}(c) = \{(u_1, \dots, u_N) : u_i \in C_0(X_i) \forall i, \text{ (4.149) holds } \forall \mathbf{r} \in X\}$
- (2) $\mathcal{A}(c)$ as in (1), with $C_b(X_i)$ in place of $C_0(X_i)$
- (3) $\mathcal{A}(c)$ as in (1), with $B(X_i) = \{u : X_i \rightarrow \mathbb{R} : u \text{ bounded measurable}\}$ in place of $C_0(X_i)$
- (4) $\mathcal{A}(c) = \{(u_1, \dots, u_N) : u_i \in L^1(X_i; d\mu_i) \forall i, \text{ (4.149) holds for } \mu_1 \otimes \dots \otimes \mu_N \text{-a.e. } \mathbf{r} \in X\}$.

Here we have used the standard notation $C_b(X_i)$ for the space of bounded continuous functions on X_i , and $C_0(X_i)$ for the space of decaying continuous functions on X_i (i.e. those u which in addition satisfy $u(\mathbf{r}_i) \rightarrow 0$ if $|\mathbf{r}_i| \rightarrow \infty$).

In the special case of two marginals defined on compact sets, cost functions c which are metrics (such as $|\mathbf{r}_1 - \mathbf{r}_2|$), and the choice (2) for the potentials, this fundamental result was discovered by Kantorovich [80]. A great many variants and modifications have subsequently appeared in the mathematics literature. Some of them replace the X_i by abstract spaces; many are worked out only for two marginals; almost all of them differ in the precise assumptions on the cost function and the class of admissible potentials. For instance, [117] (Theorems 2.1.4(b) and 2.1.1) and [81] cover bounded continuous cost functions and the class (3) for N marginals; [136] (Theorem 1.3) covers bounded-below lower semi-continuous cost functions and the class (4) for two marginals. Strictly speaking, none of the versions published prior to the discovery of the *optimal transport/SCE theory* connection applied directly to the multi-marginal Coulomb case, even though the underlying ideas essentially did. For the proof of Theorem 4.32 we refer the reader to [50].

Technical Remark From a functional analysis point of view, the natural class of admissible potentials in (4.150) is the smallest one, (1). This choice reflects the duality between potentials u_i and measures μ_i in the integral $\int u_i d\mu_i$; note that the linear hull of the space $\mathcal{P}(\mathbb{R}^d)$ of probability measures, that is, the space $\mathcal{M}(\mathbb{R}^d)$ of signed measures, is the dual of $C_0(\mathbb{R}^d)$. Enlarging this class from (1) to (2)–(4) has the virtue that the supremum of the dual problem is attained for increasingly general cost functions c .

Proof of Theorem 4.13 (1) using Theorem 4.32 Applying the Kantorovich duality theorem with X_i , μ_i , and c as in Example 4.11, and making the choice (3) for the class of admissible potentials, one obtains

$$\inf_{\substack{\Pi \in \mathcal{P}(\mathbb{R}^{Nd}) \\ \Pi \mapsto \rho}} \int_{\mathbb{R}^{Nd}} V_{ee} \, d\Pi = \\ \sup_{\substack{(u_1, \dots, u_N) : u_i \in B(\mathbb{R}^d) \, \forall i, \\ u_1(\mathbf{r}_1) + \dots + u_N(\mathbf{r}_N) \leq V_{ee}(\mathbf{r}_1, \dots, \mathbf{r}_N) \, \forall (\mathbf{r}_1, \dots, \mathbf{r}_N)}} \sum_{i=1}^N \int_{\mathbb{R}^d} u_i(\mathbf{r}_i) \frac{\rho(\mathbf{r}_i)}{N} \, d\mathbf{r}_i.$$

The left-hand side is the enlarged-search definition (4.26) of the SCE functional $V_{ee}^{\text{SCE}}[\rho]$ (which, by Theorem 4.1, is equivalent to the original definition (4.15)). The right-hand side can be simplified. For any collection of potentials (u_1, \dots, u_N) , the sum of the integrals on the right-hand side is preserved under the replacement $(u_1, \dots, u_N) \mapsto (\bar{u}, \dots, \bar{u})$, where \bar{u} denotes the average $(u_1 + \dots + u_N)/N$; moreover the constraint in (4.6) is also preserved, thanks to the symmetry of V_{ee} . Thus the right-hand side of (4.6) stays unaltered if the supremization is restricted to N equal potentials, $u_1 = \dots = u_N = u$. But in this case the right-hand side reduces to that of (4.55), establishing Theorem 4.13 (1). \square

Acknowledgments AG acknowledges partial support of his research by the Canada Research Chairs Program and Natural Sciences and Engineering Research Council of Canada. This work started when AG was at the Vrije Universiteit Amsterdam and has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. [795942].

References

1. A. Alfonsi, R. Coyaud, and V. Ehrlacher. Constrained overdamped langevin dynamics for symmetric multimarginal optimal transportation. *Mathematical Models and Methods in Applied Sciences* **32**(03), 403–455 (2022)
2. A. Alfonsi, R. Coyaud, V. Ehrlacher and D. Lombardi. Approximation of optimal transport problems with marginal moments constraints. *Math. Comp.* **90**, 689–737 (2021).
3. C.-O. Almladh and U. von Barth. Exact results for the charge and spin densities, exchange-correlation and density-functional eigenvalues. *Phys. Rev. B* **31**, 3232–3244 (1985).
4. J. Altschuler and E. Boix-Adsera. Hardness results for Multimarginal Optimal Transport problems. *Discrete Optimization* **42**, 100669 (2021).
5. L. Ambrosio. Lecture notes on optimal transport problems. In: *Mathematical Aspects of Evolving Interfaces*, pp. 1–52. vol. 1812, Springer Lecture Notes in Mathematics (2003).
6. H. Bahmann, Y. Zhou and M. Ernzerhof. The shell model for the exchange-correlation hole in the strong-correlation limit. *J. Chem. Phys.* **145**, 124104 (2016).
7. T.E. Baker, E.M. Stoudenmire, L.O. Wagner, K. Burke and S.R. White. One-dimensional mimicking of electronic structure: The case for exponentials. *Phys. Rev. B* **91**, 235141 (2015). Err. **93** 119912 (2016).

8. J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna and G. Peyré. Iterative Bregman projections for regularized transportation problems. *SIAM Journal on Scientific Computing*, **37**(2), A1111–A1138 (2015).
9. J.-D. Benamou, G. Carlier and L. Nenna. A numerical method to solve multi-marginal optimal transport problems with coulomb cost. In: *Splitting Methods in Communication, Imaging, Science, and Engineering*, edited by R. Glowinski, S. Osher and W. Yin, pp. 577–601 Springer, Berlin-Heidelberg (2016).
10. U. Bindini and L. De Pascale. Optimal transport with Coulomb cost and the semiclassical limit of density functional theory. *J. Éc. polytech. Math.* **4**, 909–934 (2017).
11. U. Bindini, L. De Pascale and A. Kausamo. On Seidl-type maps for multi-marginal optimal transport with Coulomb cost. *arXiv preprint arXiv:2011.05063* (2020).
12. M.A. Buijse, E.J. Baerends and J. G. Snijders. Analysis of correlation in terms of exact local potentials: Applications to two-electron systems. *Phys. Rev. A* **40**, 4190–4202 (1989).
13. G. Buttazzo, T. Champion and L. De Pascale. Continuity and estimates for multimarginal optimal transportation problems with singular costs. *Applied Mathematics & Optimization* **78**, 185–200 (2018).
14. G. Buttazzo, L. De Pascale and P. Gori-Giorgi. Optimal-transport formulation of electronic density-functional theory. *Phys. Rev. A* **85**, 062502 (2012).
15. G. Carlier. On the linear convergence of the multi-marginal Sinkhorn algorithm. HAL Id: hal-03176512 (2021).
16. G. Carlier, G. Friesecke and D. Vögler. Convex geometry of finite exchangeable laws and de Finetti style representation with universal correlated corrections. *Probab. Theory Relat. Fields* (2022). <https://doi.org/10.1007/s00440-022-01115-2>
17. G. Carlier and M. Laborde. *SIAM Journal on Mathematical Analysis* **52**, 709–717 (2020).
18. M. Casula, S. Sorella and G. Senatore. Ground state properties of the one-dimensional coulomb gas using the lattice regularized diffusion monte carlo method. *Phys. Rev. B* **74**, 245427 (2006).
19. J. Chayes and L. Chayes. On the validity of the inverse conjecture in classical density functional theory. *Journal of statistical physics* **36**, 471–488 (1984).
20. J. Chayes, L. Chayes and E.H. Lieb. The inverse problem in classical statistical mechanics. *Communications in Mathematical Physics* **93**, 57–121 (1984).
21. H. Chen and G. Friesecke. Pair densities in density functional theory. *Multiscale Modeling & Simulation* **13**, 1259–1289 (2015).
22. H. Chen, G. Friesecke and C.B. Mendl. Numerical methods for a Kohn–Sham density functional model based on optimal transport. *J. Chem. Theory Comput.* **10**, 4360–4368 (2014).
23. Y. Chen, T. Georgiou and M. Pavon. Entropic and displacement interpolation: a computational approach using the Hilbert metric. *SIAM Journal on Applied Mathematics* **76**, 2375–2396 (2016).
24. A. Cohen, P. Mori-Sánchez and W. Yang. Insights into current limitations of density functional theory. *Science* **321**, 792–794 (2008).
25. A.J. Cohen, P. Mori-Sánchez and W. Yang. *J. Chem. Phys.* **127**, 034101 (2007).
26. A.J. Coleman and V.I. Yukalov. *Reduced Density Matrices*. Lecture Notes in Chemistry Vol. 72, Springer (2000).
27. M. Colombo, L. De Pascale and S. Di Marino. Multimarginal optimal transport maps for one-dimensional repulsive costs. *Canad. J. Math.* **67**, 350–368 (2015).
28. M. Colombo and S. Di Marino. Equality between monge and kantorovich multimarginal problems with coulomb cost. *Annali di Matematica Pura ad Applicata*, 1–14 (2013).
29. M. Colombo, S. Di Marino and F. Stra. First order expansion in the semiclassical limit of the levy-lieb functional. *arXiv preprint arXiv:2106.06282* (2021).
30. M. Colombo and F. Stra. Counterexamples in multimarginal optimal transport with coulomb cost and spherically symmetric data. *Mathematical Models and Methods in Applied Sciences* **26**, 1025–1049 (2016).

31. L.A. Constantin. Correlation energy functionals from adiabatic connection formalism. *Phys. Rev. B* **99**, 085117 (2019).
32. C. Cotar, G. Friesecke, and C. Klüppelberg. Density functional theory and optimal transportation with coulomb cost. *Comm. Pure Appl. Math.* **66**, 548–99 (2013).
33. C. Cotar, G. Friesecke and C. Klüppelberg. Smoothing of transport plans with fixed marginals and rigorous semiclassical limit of the Hohenberg-Kohn functional. *Arch. Ration. Mech. An.* **228**, 891–922 (2018).
34. M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In: *Advances in neural information processing systems 26 (NIPS 2013)*, edited by C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani and K.Q. Weinberger, pp. 2292–2300, Curran Associates Inc., NY (2013).
35. M. Cuturi and G. Peyré. *Computational optimal transport*. Vol. 11, Now Publishers, Inc. (2019).
36. T.J. Daas, E. Fabiano, F. Della Sala, P. Gori-Giorgi and S. Vuckovic. Noncovalent interactions from models for the Møller–Plesset adiabatic connection. *The journal of physical chemistry letters* **12**, 4867–4875 (2021).
37. T.J. Daas, J. Grossi, S. Vuckovic, Z.H. Musslimani, D.P. Kooi, M. Seidl, K.J. Giesbertz and P. Gori-Giorgi. Large coupling-strength expansion of the Møller–Plesset adiabatic connection: From paradigmatic cases to variational expressions for the leading terms. *The Journal of chemical physics* **153**, 214112 (2020).
38. L. De Pascale. Optimal transport with coulomb cost. approximation and duality. *ESAIM: Math. Model. Numer. Anal.* **49**, 1643–1657 (2015).
39. E. del Barrio and J.-M. Loubes. The statistical effect of entropic regularization in optimal transportation. *arXiv:2006.05199* (2020).
40. S. Di Marino and A. Gerolin. An Optimal Transport approach for the Schrödinger bridge problem and convergence of Sinkhorn algorithm. *Journal of Scientific Computing* **85**(2), 1–28 (2020).
41. S. Di Marino and A. Gerolin. Optimal transport losses and Sinkhorn algorithm with general convex regularization. *arXiv preprint, arXiv:2007.00976* (2020).
42. S. Di Marino, A. Gerolin and L. Nenna. Optimal transport for repulsive costs. In: *Topological Optimization and Optimal Transport – In the Applied Sciences*, edited by M. Bergounioux, É. Oudet, M. Rumpf, G. Carlier, T. Champion and F. Santambrogio, chapter 9, De Gruyter (2017).
43. M. Ernzerhof. Construction of the adiabatic connection. *Chem. Phys. Lett.* **263**, 499 (1996).
44. E. Fabiano, P. Gori-Giorgi, M. Seidl, and F. Della Sala. Interaction-strength interpolation method for main-group chemistry: Benchmarking, limitations, and perspectives. *J. Chem. Theory. Comput.* **12**, 4885–4896 (2016).
45. E. Fabiano, S. Smiga, S. Giarrusso, T.J. Daas, F. Della Sala, I. Grabowski and P. Gori-Giorgi. Investigation of the exchange-correlation potentials of functionals based on the adiabatic connection interpolation. *J. Chem. Theory. Comput.* **15**, 1006–1015 (2019).
46. J. Franklin and J. Lorenz. On the scaling of multidimensional matrices. *Linear Algebra and its applications* **114**, 717–735 (1989).
47. D.E. Freund, B.D. Huxtable and J.D. Morgan. Variational calculations on the helium isoelectronic sequence. *Phys. Rev. A* **29**, 980–982 (1984).
48. G. Friesecke. The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions. *Archive for Rational Mechanics and Analysis*, **169**, 35–71 (2003).
49. A simple counterexample to the Monge ansatz in multi-marginal optimal transport, convex geometry of the set of Kantorovich plans, and the Frenkel–Kontorova model. *SIAM J. Math. Analysis* **51**, 4332–4355 (2019).
50. G. Friesecke. Lectures on optimal transport. *SIAM*, to appear (2022).
51. G. Friesecke, C.B. Mendl, B. Pass, C. Cotar and C. Klüppelberg. N-density representability and the optimal transport limit of the Hohenberg–Kohn functional. *J. Chem. Phys.* **139**, 164109 (2013).

52. G. Friesecke, A.S. Schulz and D. Vögler. Genetic column generation: Fast computation of high-dimensional multi-marginal optimal transport problems. *SIAM J. Sci. Comp.* **44**(3), A1632–A1654 (2022).
53. G. Friesecke and D. Vögler. Breaking the curse of dimension in multi-marginal Kantorovich optimal transport on finite state spaces. *SIAM J. Math. Analysis* **50**, 3996–4019 (2018).
54. A. Galichon and B. Salanié. Matching with trade-offs: Revealed preferences over competing characteristics. CEPR Discussion Paper No. DP7858 (2010).
55. W. Gangbo and R.J. McCann. The geometry of optimal transportation. *Acta Math.* **177**, 113–161 (1906).
56. W. Gangbo and A. Swiech. Optimal maps for the multidimensional Monge-Kantorovich problem. *Commun. Pure Appl. Math.* **51**, 23 (1998).
57. A. Gerolin. *Multi-marginal optimal transport and potential optimization problems for Schrödinger operators*. PhD thesis, Università degli studi di Pisa (2016).
58. A. Gerolin, J. Grossi, and P. Gori-Giorgi. Kinetic correlation functionals from the entropic regularisation of the strictly-correlated electrons problem. *Journal of Chemical Theory and Computation* **16**, 488–498 (2019).
59. A. Gerolin, A. Kausamo, and T. Rajala. Non-existence of optimal transport maps for the multi-marginal repulsive harmonic cost. *SIAM Journal on Mathematical Analysis* **51** (2019).
60. A. Gerolin, A. Kausamo, and T. Rajala. Multi-marginal Entropy-Transport with repulsive cost. *Calc. Var. PDEs* **59**, Art. 90 (2020).
61. A. Ghosal, A.D. Guclu, C.J. Umrigar, D. Ullmo and H.U. Baranger. *Nature Phys.* **2**, 336 (2006).
62. S. Giarrusso, P. Gori-Giorgi, F. Della Sala and E. Fabiano. Assessment of interaction-strength interpolation formulas for gold and silver clusters. *J. Chem. Phys.* **148**, 134106 (2018).
63. S. Giarrusso, S. Vuckovic, and P. Gori-Giorgi. Response potential in the strong-interaction limit of DFT: Analysis and comparison with the coupling-constant average. *J. Chem. Theory Comput.* **14**, 4151–4167 (2018).
64. G.F. Giuliani and G. Vignale. *Quantum Theory of the Electron Liquid*. Cambridge University Press, New York (2005).
65. P. Gori-Giorgi and A. Savin. *J. Phys.: Conf. Ser.* **117**, 012017 (2008).
66. P. Gori-Giorgi and M. Seidl. Density functional theory for strongly-interacting electrons: perspectives for physics and chemistry. *Phys. Chem. Chem. Phys.* **12**, 14405–14419 (2010).
67. P. Gori-Giorgi, M. Seidl and G. Vignale. Density-functional theory for strongly interacting electrons. *Phys. Rev. Lett.* **103**, 166402 (2009).
68. P. Gori-Giorgi, G. Vignale and M. Seidl. Electronic zero-point oscillations in the strong-interaction limit of density functional theory. *J. Chem. Theory Comput.* **5**, 743–753 (2009).
69. A. Görling and M. Levy. Correlation-energy functional and its high-density limit obtained from a coupling-constant perturbation expansion. *Phys. Rev. B* **47**, 13105 (1993).
70. N. Gozlan and C. Léonard. Transport inequalities, a survey. *Markov Processes and Related Fields* **16**, 635–736 (2010).
71. J. Grossi, D.P. Kooi, K.J.H. Giesbertz, M. Seidl, A.J. Cohen, P. Mori-Sánchez and P. Gori-Giorgi. Fermionic statistics in the strongly correlated limit of density functional theory. *J. Chem. Theory Comput.* **13**, 6089–6100 (2017).
72. J. Grossi, Z. Musslimani, M. Seidl and P. Gori-Giorgi. Kohn–Sham equations with functionals from the strictly-correlated regime: Investigation with a spectral renormalization method. *Journal of Physics: Condensed Matter*. Epub ahead of print (2020).
73. A.D. Guclu, A. Ghosal, C.J. Umrigar and H.U. Baranger. *Phys. Rev. B* **77**, 041301 (2008).
74. N. Helbig, J.I. Fuks, M. Casula, M.J. Verstraete, M. Marques, I. Tokatly and A. Rubio. Density functional theory beyond the linear regime: Validating an adiabatic local density approximation. *Physical Review A* **83**, 032503 (2011).
75. N. Helbig, I.V. Tokatly, and A. Rubio. Exact Kohn–Sham potential of strongly correlated finite systems. *J. Chem. Phys.* **131**, 224105 (2009).
76. T. Helgaker, P. Jørgensen and J. Olsen. *Density-functional Theory: A Convex Treatment*. Wiley Blackwell (2016).

77. M. Hoffmann-Ostenhof and T. Hoffmann-Ostenhof. “Schrödinger inequalities” and asymptotic behavior of the electron density of atoms and molecules. *Phys. Rev. A* **16**, 1782–1785 (1977).
78. P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev. B* **136**, 864 (1964).
79. H. Janati, B. Muzellec, G. Peyré and M. Cuturi. Entropic optimal transport between unbalanced gaussian measures has a closed form. In *Advances in Neural Information Processing Systems*, vol. 33, edited by H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan and H. Lin, pp.10468–10479, Curran Associates, Inc. (2020).
80. L.V. Kantorovich. On the transfer of masses. *Dokl. Akad. Nauk. SSSR*. **37**, 227 (1942).
81. H.G. Kellerer. Duality theorems for marginal problems. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* **67**, 399–432 (1984).
82. Y. Khoo and L. Ying. Convex relaxation approaches for strictly correlated density functional theory. *SIAM J. Sci. Comput.* **41**, B773–B795 (2019).
83. M. Knott and C.S. Smith. On the optimal mapping of distributions. *J. Optimization Theory and Appl.* **43**, 39–49 (1984).
84. W. Kohn. v -representability and density functional theory. *Physical review letters* **51**, 1596 (1983).
85. W. Kohn and L.J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133–A1138 (1965).
86. W. Kolos and C.C.J. Roothaan. Accurate electronic wave functions for the h_2 molecule. *Rev. Mod. Phys.* **32**, 219–232 (1960).
87. D.P. Kooi and P. Gori-Giorgi. Local and global interpolations along the adiabatic connection of dft: a study at different correlation regimes. *Theoretical chemistry accounts* **137**, 1–12 (2018).
88. D.P. Kooi and P. Gori-Giorgi. A variational approach to London dispersion interactions without density distortion. *The journal of physical chemistry letters* **10**, 1537–1541 (2019).
89. C. Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Cont. Dyn.-A* **34**, 1533–1574 (2014).
90. M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v -representability problem. *Proc. Natl. Acad. Sci.* **76**, 6062–6065 (1979).
91. M. Levy. Electron densities in search of hamiltonians. *Phys. Rev. A* **26**, 1200–1208 (1982).
92. M. Levy and J. Perdew. Hellman–Feynman, virial, and scaling requisites for the exact universal density functionals. Shape of the correlation potential and diamagnetic susceptibility for atoms. *Phys. Rev. A* **32**, 2010–2021 (1985).
93. M. Levy, J.P. Perdew and V. Sahni. Exact differential equation for the density and ionization energy of a many-particle system. *Phys. Rev. A* **30**, 2745–2748 (1984).
94. M. Levy and F. Zahariev. *Phys. Rev. Lett.* **113**, 113002 (2014).
95. M. Lewin. Semi-classical limit of the Levy–Lieb functional in Density Functional Theory. *C. R. Math.* **356**, 449–455 (2018).
96. M. Lewin, E.H. Lieb and R. Seiringer. Floating wigner crystal with no boundary charge fluctuations. *Physical Review B* **100**, 035127 (2019).
97. E.H. Lieb. Density functionals for Coulomb systems. *Int. J. Quantum. Chem.* **24**, 243–277 (1983).
98. Z.-F. Liu and K. Burke. Adiabatic connection in the low-density limit. *Phys. Rev. A* **79**, 064503 (2009).
99. D. Lorenz and H. Mahler. Orlicz space regularization of continuous optimal transport problems. *Appl. Math. Optim.* **85**, 14 (2022).
100. F. Malet and P. Gori-Giorgi. Strong correlation in Kohn–Sham density functional theory. *Phys. Rev. Lett.* **109**, 246402 (2012).
101. F. Malet, A. Mirtschink, J.C. Cremon, S.M. Reimann and P. Gori-Giorgi. Kohn–Sham density functional theory for quantum wires in arbitrary correlation regimes. *Phys. Rev. B* **87**, 115146 (2013).

102. F. Malet, A. Mirschink, K.J.H. Giesbertz, L.O. Wagner and P. Gori-Giorgi. Exchange-correlation functionals from the strong interaction limit of dft: applications to model chemical systems. *Phys. Chem. Chem. Phys.* **16**, 14551–14558 (2014).
103. A. Mallasto, A. Gerolin, and H.Q. Minh. Entropy-regularized 2-Wasserstein distance between gaussian measures. *Information Geometry*, 1–35 (2021).
104. C.B. Mendl and L. Lin. Kantorovich dual solution for strictly correlated electrons in atoms and molecules. *Phys. Rev. B* **87**, 125106 (2013).
105. C.B. Mendl, F. Malet and P. Gori-Giorgi. Wigner localization in quantum dots from kohnsham density functional theory without symmetry breaking. *Phys. Rev. B* **89**, 125106 (2014).
106. A. Mirschink, C.J. Umrigar, J.D. Morgan III and P. Gori-Giorgi. Energy density functionals from the strong-coupling limit applied to the anions of the He isoelectronic series. *J. Chem. Phys.* **140**, 18A532 (2014).
107. A. Moameni and B. Pass. Solutions to multi-marginal optimal transport problems concentrated on several graphs. *ESAIM: Control Optim. Calc. Var.* **23**(2), 551–567 (2017).
108. G. Monge. *Mémoire sur la théorie des déblais et des remblais*. Histoire Acad. Sciences, Paris (1781).
109. P. Mori-Sánchez and A.J. Cohen. Exact density functional obtained via the Levy constrained search. *The journal of physical chemistry letters* **9**, 4910–4914 (2018).
110. L. Nenna. *Numerical methods for multi-marginal optimal transportation*. PhD thesis, Université Paris sciences et lettres (2016).
111. B. Pass. *Structural results on optimal transportation plans*. PhD thesis, University of Toronto (2011).
112. B. Pass. On the local structure of optimal measures in the multi-marginal optimal transportation problem. *Calculus of Variations and Partial Differential Equations* **43**, 529–536 (2012).
113. B. Pass. Remarks on the semi-classical Hohenberg–Kohn functional. *Nonlinearity*, **26**(9), 2731 (2013).
114. B. Pass. Multi-marginal optimal transport and multi-agent matching problems: uniqueness and structure of solutions. *Discrete Contin. Dyn. Syst.* **34**, 1623–1639 (2014).
115. B. Pass. Multi-marginal optimal transport: theory and applications. *ESAIM: Mathematical Modelling and Numerical Analysis* **49**(6), 1771–1790 (2015).
116. A. Pratelli. On the equality between Monge’s infimum and Kantorovich’s minimum in optimal mass transportation. *Annales de l’Institut Henri Poincaré (B) Probability and Statistics* vol. **43**, 1–13 (2007).
117. S. Rachev and L. Rüschendorf. *Mass transportation problems*. Springer-Verlag, New York (1998).
118. L. Rüschendorf. Convergence of the iterative proportional fitting procedure. *The Annals of Statistics* **23**, 1160–1174 (1995).
119. F. Santambrogio. *Optimal Transport for Applied Mathematicians*. Progress in Nonlinear Differential Equations and Their Applications, Birkhäuser (2015).
120. A. Savin. Is size-consistency possible with density functional approximations? *Chem. Phys.* **356**, 91–97 (2009).
121. E. Schrödinger. *Über die umkehrung der naturgesetze*. Verlag Akademie der wissenschaften in kommission bei Walter de Gruyter u. Company (1931).
122. M. Seidl. Strong-interaction limit of density-functional theory. *Phys. Rev. A* **60**, 4387–4395 (1999).
123. M. Seidl. Adiabatic connection in density-function theory: Two electrons on the surface of a sphere. *Phys. Rev. A* **75**, 062506 (2007).
124. M. Seidl, S. Di Marino, A. Gerolin, L. Nenna, K.J. Giesbertz and P. Gori-Giorgi. The strictly-correlated electron functional for spherically symmetric systems revisited ii: Sgs conjecture. In preparation.
125. M. Seidl, S. Di Marino, A. Gerolin, L. Nenna, K.J. Giesbertz and P. Gori-Giorgi. The strictly-correlated electron functional for spherically symmetric systems revisited. *arXiv preprint, arXiv:1702.05022* (2017).

126. M. Seidl, S. Giarrusso, S. Vuckovic, E. Fabiano and P. Gori-Giorgi. Communication: Strong-interaction limit of an adiabatic connection in Hartree–Fock theory. *The Journal of Chemical Physics* **149**, 241101 (2018).
127. M. Seidl, P. Gori-Giorgi and A. Savin. Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities. *Phys. Rev. A* **75**, 042511/12 (2007).
128. M. Seidl, J.P. Perdew and S. Kurth. Density functionals for the strong-interaction limit. *Phys. Rev. A* **62**, p. 012502 (2000).
129. M. Seidl, J.P. Perdew and S. Kurth. Simulation of all-order density-functional perturbation theory, using the second order and the strong-correlation limit. *Phys. Rev. Lett.* **84**, 5070–5073 (2000).
130. M. Seidl, J.P. Perdew and M. Levy. Strictly correlated electrons in density-functional theory. *Phys. Rev. A* **59**, 51–54 (1999).
131. R. Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *The annals of mathematical statistics* **35**, 876–879 (1964).
132. S. Smiga and L.A. Constantin. Modified interaction-strength interpolation method as an important step toward self-consistent calculations. *Journal of chemical theory and computation* **16**, 4983–4992 (2020).
133. R. van Leeuwen. Density functional approach to the many-body problem: key concepts and exact functionals. *Adv. Quantum Chem.* **43**, 24–94 (2003).
134. D. Vieira. Spin-independent v -representability of Wigner crystal oscillations in one-dimensional Hubbard chains: The role of spin-charge separation. *Phys. Rev. B* **86**, 075132 (2012).
135. D. Vieira and K. Capelle. Investigation of self-interaction corrections for an exactly solvable model system: Orbital dependence and electron localization. *J. Chem. Theory Comput.* **6**, 3319–3329 (2010).
136. C. Villani. *Topics in Optimal Transportation*. Grad. Stud. Math. **58**. Amer. Math. Soc., Providence (2003).
137. S. Vuckovic, P. Gori-Giorgi, F. Della Sala and E. Fabiano. Restoring size consistency of approximate functionals constructed from the adiabatic connection. *J. Phys. Chem. Lett.* **9**, 3137–3142 (2018).
138. S. Vuckovic, T.J.P. Irons, A. Savin, A.M. Teale and P. Gori-Giorgi. Exchange–correlation functionals via local interpolation along the adiabatic connection. *J. Chem. Theory Comput.* **12**, 2598–2610 (2016).
139. S. Vuckovic, M. Levy and P. Gori-Giorgi. Augmented potential, energy densities, and virial relations in the weak-and strong-interaction limits of DFT. *J. Chem. Phys.* **147**, 214107 (2017).
140. L.O. Wagner and P. Gori-Giorgi. Electron avoidance: A nonlocal radius for strong correlation. *Phys. Rev. A* **90**, 052512 (2014).
141. L.O. Wagner, E.M. Stoudenmire, K. Burke, and S.R. White. Reference electronic structure calculations in one dimension. *Phys. Chem. Chem. Phys.* **14**, 8581 (2012).
142. J.-J. Wang, W. Li, S. Chen, G. Xianlong, M. Rontani and M. Polini. Absence of Wigner molecules in one-dimensional few-fermion systems with short-range interactions. *Physical Review B* **86**, 075110 (2012).
143. E.P. Wigner. On the interaction of electrons in metals. *Phys. Rev.* **46**, 1002 (1934).
144. Z.-J. Ying, V. Brosco, G.M. Lopez, D. Varsano, P. Gori-Giorgi and J. Lorenzana. Anomalous scaling and breakdown of conventional density functional theory methods for the description of Mott phenomena and stretched bonds. *Phys. Rev. B* **94**, 075154 (2016).
145. G.M. Zhislin. Discussion of the spectrum of Schrödinger operators for systems of many particles. *Trudy Moskovskogo matematicheskogo obschestva* **9**, 81–120 (1960).