Systematic construction of density functionals based on matrix product state computations

This content has been downloaded from IOPscience. Please scroll down to see the full text.
2016 New J. Phys. 18 083039
(http://iopscience.iop.org/1367-2630/18/8/083039)
View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 131.169.116.119
This content was downloaded on 25/08/2016 at 08:59

Please note that terms and conditions apply.

You may also be interested in:

Exact maps in density functional theory for lattice models
Tanja Dimitrov, Heiko Appel, Johanna I Fuks et al.

The Hubbard dimer: a density functional case study of a many-body problem
D J Carrascal, J Ferrer, J C Smith et al.

Existence, uniqueness, and construction of the density-potential mapping in time-dependent density-functional theory
Michael Ruggenthaler, Markus Penz and Robert van Leeuwen

Transfer matrices and excitations with matrix product states
V Zauner, D Draxler, L Vanderstraeten et al.

Unifying projected entangled pair state contractions
Michael Lubasch, J Ignacio Cirac and Mari-Carmen Bañuls

The entanglement of few-particle systems when using the local-density approximation
J P Coe and I D'Amico

Variational cluster approximation to the thermodynamics of quantum spin systems
S Filor and T Pruschke
Systematic construction of density functionals based on matrix product state computations

Michael Lubasch\textsuperscript{1,5}, Johanna I Fuks\textsuperscript{3}, Heiko Appel\textsuperscript{4,6}, Angel Rubio\textsuperscript{3,4}, J Ignacio Cirac\textsuperscript{1} and Mari-Carmen Bañuls\textsuperscript{3}

\textsuperscript{1} Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Straße 1, D-85748 Garching, Germany
\textsuperscript{2} Department of Physics and Astronomy, Hunter College and the Graduate Center of the City University of New York, 695 Park Avenue, New York, NY 10065, USA
\textsuperscript{3} Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin-Dahlem, Germany
\textsuperscript{4} Max-Planck-Institut für Struktur und Dynamik der Materie, Luruper Chaussee 149, D-22761 Hamburg, Germany
\textsuperscript{5} Present address: Department of Physics, University of Oxford, Parks Road, Oxford OX1 3PU, UK.

E-mail: michael.lubasch@mpq.mpg.de

Keywords: matrix product states, density functional theory, exchange-correlation energy, local density approximation

Abstract

We propose a systematic procedure for the approximation of density functionals in density functional theory that consists of two parts. First, for the efficient approximation of a general density functional, we introduce an efficient ansatz whose non-locality can be increased systematically. Second, we present a fitting strategy that is based on systematically increasing a reasonably chosen set of training densities. We investigate our procedure in the context of strongly correlated fermions on a one-dimensional lattice in which we compute accurate training densities with the help of matrix product states. Focusing on the exchange-correlation energy, we demonstrate how an efficient approximation can be found that includes and systematically improves beyond the local density approximation. Importantly, this systematic improvement is shown for target densities that are quite different from the training densities.

1. Introduction

The formulation of quantum mechanics in terms of density functionals instead of wave functions, following the ground-breaking works of Hohenberg, Kohn, and Sham \cite{Hohenberg:1964, Kohn:1965}, made numerical simulations of quantum mechanical systems ranging from the microscopic to the macroscopic world feasible \cite{Perdew:1981, Levy:1974, Perdew:1986}. The enormous success of DFT is thus deeply rooted in the Kohn–Sham theorem \cite{Kohn:1965}, which states that the ground state energy \( E \) of a system can be decomposed into a kinetic, an interaction and a potential part. By means of a fictitious non-interacting system, namely the KS system, the non-interacting part of the kinetic energy, \( T^0 \), can be obtained efficiently, which represents a large contribution to the full interacting kinetic energy \( T \). Further, part of the interaction energy is accounted for by the Hartree energy \( E_{\text{H}} \). The potential part \( V \) can be exactly computed efficiently for any ground state density \( n \). Finally, the remaining part of the total ground state energy defines the xc density functional, \( E_{\text{xc}} \), which is computed from the Kohn–Sham equations.

DFT in its most widely used form, namely Kohn–Sham (KS) DFT,\textsuperscript{2} requires the xc density functional in order to be able to compute ground state energies and densities. By virtue of the Hohenberg–Kohn theorem \cite{Kohn:1965}, all ground state observables are functionals of the ground state density \( n \), and so \( E_{\text{xc}} = E_{\text{xc}}[n] \). The ground state energy \( E = E[n] \) of a system can be decomposed into a kinetic, an interaction and a potential part. By means of a fictitious non-interacting system, namely the KS system, the non-interacting part of the kinetic energy, \( T^0 = T^0[n] \), can be obtained efficiently, which represents a large contribution to the full interacting kinetic energy \( T \). Further, part of the interaction energy is accounted for by the Hartree energy \( E_{\text{H}}[n] \). The potential part \( V[n] \) can be exactly computed efficiently for any ground state density \( n \). Finally, the remaining part of the total ground state energy defines the xc density functional, \( E_{\text{xc}} = E[n] - T^0[n] - E_{\text{H}}[n] - V[n] \). DFT is in principle exact, but in practice determining the precise form of the xc density functional is QMA-hard \cite{Verstraete:2009}.

Therefore, KS DFT can only make use of approximations of \( E_{\text{xc}} \). The enormous success of DFT is thus deeply connected to the successful construction of good approximations for the xc energy density functional.

In the history of DFT and quest for a universally applicable approximate \( E_{\text{xc}} \) \cite{Perdew:1986}, mainly two different paths have been followed: one is the non-empirical approach pioneered by Perdew \cite{Perdew:1981} and the other is the semi-empirical approach initiated by Becke \cite{Becke:1988}. The non-empirical approach makes use of exact conditions, that a
physical system must fulfill, to find approximations for the xc density functional. Within this approach a 'Jacob’s
ladder’ of functionals was built where each functional on a higher rung of the ladder is supposed to improve
upon the ones on the lower rungs [16, 17]. On the lowest rung of the ‘Jacob’s ladder’ resides the local density
approximation (LDA), which was already introduced by Kohn and Sham in [2]. The higher rungs are supposed to
systematically improve upon the LDA, which, in practice, does not always happen [17]. Additionally, at the
moment, the more precise functionals on the higher rungs are so much more difficult to compute that further
improvements of DFT following this non-empirical approach seem very hard to achieve. In the semi-empirical
approach, an ansatz for the functional form of $E_{xc}$ is fitted using experimental data, accurate theoretical
reference data, or other constraints. However, often relatively small training sets are used in these fits and then
the resulting functionals can be biased towards their training [14].

Alternatively, we might obtain further improvements of $E_{xc}$ away from but using concepts of both the semi-
empirical and the non-empirical approach, e.g. by using a large set of accurate training densities and
corresponding values of $E_{xc}$, and by fitting an efficient ansatz to these data that includes some exact conditions.
Obviously, a difficulty of this alternative scheme is that it requires a possibly large number of accurate solutions
for the quantum many-body problem. However, nowadays, tensor network states provide precise results for
quantum many-body systems, e.g. [18, 19], in particular with respect to ground state properties. We remark that
tensor network methods are currently limited to low-dimensional, i.e. one- and some two-dimensional,
quantum lattice problems while DFT usually handles three-dimensional continuous quantum systems. Since
DFT can be applied to a wide range of realistic quantum systems, it is a useful algorithm for a large community
and thus worth improving.

In this article, we want to analyze the feasibility of constructing an approximate xc density functional of a
specific form, when large training sets of ground state densities and corresponding values of $E_{xc}$ are available.
The specific form for the ansatz of our approximation is inspired by the semi-empirical approach [17]; it includes
the LDA [2, 20] and allows a systematic improvement beyond it. For this feasibility study, we focus on discrete
lattice problems and the one-dimensional case, and we use matrix product states (MPS) for the computation of
accurate ground state energies and densities [21, 22]. The specific discrete lattice problem considered here can be
derived from discretization of continuous space, i.e. the usual scenario of DFT. Then our ansatz can be seen as
the discretized version of a continuous function. Although we could approach the continuum solution by
successively decreasing the discretization, taking the continuum limit is beyond the scope of this work.

The structure of this article is as follows. In section 2 we introduce the considered Hamiltonian and
observables. The corresponding exact LDA is presented in section 3. We then propose, fit, and assess our ansatz
in section 4. Finally, in section 5 we conclude this work and give an outlook.

2. Model

In the following, we consider two species of fermions with long-ranged soft-Coulomb interaction on a finite
one-dimensional lattice of length $L$ with hard-wall boundary conditions, as represented by the Hamiltonian:

$$
\hat{H} = \hat{T} + \hat{W} + \hat{V}
$$

(1)

with

$$
\hat{T} = -i \sum_{l=1}^{L-1} \sum_{\sigma = \uparrow, \downarrow} \left( c_{l,\sigma}^\dagger c_{l+1,\sigma} + c_{l+1,\sigma}^\dagger c_{l,\sigma} \right),
$$

(2a)

$$
\hat{W} = U \sum_{l=1}^{L} \left( \hat{n}_{l,\uparrow} \hat{n}_{l,\downarrow} + \sum_{m=l+1}^{L} \frac{\hat{n}_m \hat{n}_m}{\sqrt{(m-l)^2 + 1}} \right),
$$

(2b)

$$
\hat{V} = \sum_{l=1}^{L} (\nu \hat{n}_l - \mu) \hat{n}_l.
$$

(2c)

Here, $c_{l,\sigma}^\dagger$ creates and $c_{l,\sigma}$ annihilates a fermion of species $\sigma = \uparrow, \downarrow$ on lattice site $l$, $\hat{n}_{l,\sigma} = c_{l,\sigma}^\dagger c_{l,\sigma}$ is the
respective occupation number operator and $\hat{n}_l = \hat{n}_{l,\uparrow} + \hat{n}_{l,\downarrow}$. The total particle number is denoted by
$N \equiv \langle \sum_{l=1}^{L} \hat{n}_l \rangle$. We obtain ground states with different total particle number by choosing different values for the
chemical potential $\mu$, which plays the role of a Lagrange multiplier fixing $N$. Such a Hamiltonian can also
describe the discretized continuous problem with lattice spacing $\Delta$ when in (2b) $t$ is replaced by $1/(2 \Delta^2)$ and in
(2b) the denominator $\sqrt{(m-l)^2 + 1}$ is replaced by $\sqrt{(m-l)^2 + \Delta^2}$ and $1$. The solution for different
discretizations can be very precisely computed with MPS [23, 24] and so our approach should yield highly
accurate training densities. If we would like to obtain the solution for continuous space, we would have to run
our computations repeatedly with decreasing lattice spacing $\Delta$ and extrapolate our results to $\Delta = 0$. Here we
see (1) as the Hamiltonian of the problem, and not a discrete version of a more fundamental one, thus we set \( t = 1/2 \) and \( U = 1 \) and fix the number of lattice sites to \( L = 21 \) from now on.

On a finite lattice, densities \( n_i = (\hat{n}_i) = (\hat{n}_i^+ + \hat{n}_i^-) \) can be written into a vector \( \mathbf{n} = (n_1, n_2, \ldots, n_L)^T \)—where \( T \) denotes the transpose—such that every density functional \( F \) can be written as a function of such density vectors \( F = F(\mathbf{n}) \). In the following, we will consider the universal Hohenberg-Kohn functional \( F_{HK}(\mathbf{n}) \), the Hartree–energy \( E_1(\mathbf{n}) \), and the non-interacting kinetic energy \( T^0(\mathbf{n}) \), e.g. [20]. For the above Hamiltonian (1) these functionals read:

\[
F_{HK}(\mathbf{n}) := E(\mathbf{n}) - \sum_{l=1}^{L} (\psi_{l}^{\text{ext}} - \mu)n_{l},
\]

\[
E_{1}(\mathbf{n}) := U \sum_{l=1}^{L} \left( \frac{n_{l}^{2}}{4} + \sum_{m=1}^{L} \frac{n_{m}n_{m+l}}{\sqrt{(m-l)^{2}+1}} \right),
\]

\[
T^{n}(\mathbf{n}) := E^{n}(\mathbf{n}) - \sum_{l=1}^{L} (\psi_{l}^{n} - \mu^{n})n_{l},
\]

where \( E(\mathbf{n}) \) denotes the ground state energy of an interacting density \( \mathbf{n} \), i.e. corresponding to (1) with \( W \), and \( E^{n}(\mathbf{n}) \) denotes the ground state energy of a non-interacting density \( \mathbf{n} \), i.e. corresponding to (1) without \( W \). Knowing the values of these functionals (3a), (3b), and (3c) for a particular density \( \mathbf{n} \) allows to calculate the xc energy \( E_{xc} \) for that density:

\[
E_{xc}(\mathbf{n}) = F_{HK}(\mathbf{n}) - E_{1}(\mathbf{n}) - T^{n}(\mathbf{n}).
\]

However, given an arbitrary density vector \( \mathbf{n} \), only \( E_{1}(\mathbf{n}) \) is trivial to compute: \( F_{HK}(\mathbf{n}) \) requires the knowledge of the external potential as a function of the density, \( \psi_{l}^{\text{ext}} = \psi_{l}^{\text{ext}}(\mathbf{n}) \), and \( T^{n}(\mathbf{n}) \) requires the knowledge of the effective non-interacting KS potential \( \psi_{l}^{n} = \psi_{l}^{n}(\mathbf{n}) \). This process of calculating the external potential \( \psi_{l}(\mathbf{n}) \), in which the ground state has the given density \( \mathbf{n} \), is called inversion and can be performed efficiently only in the non-interacting case or for two fermions [25]. In general, there exists no efficient inversion procedure for the interacting case and we use a slight modification of the iteration proposed in [26, 27]: aiming at the target density \( \mathbf{n}_{0}^{\text{ext}} \), we iterate \( \psi_{l}(i+1) = \psi_{l}(i) + \gamma(i)(n_{l}(i) - \psi_{l}^{\text{ext}}) \) until \( \| \mathbf{n}(i) - \mathbf{n}_{0}^{\text{ext}} \| \approx 0 \) where \( \| \mathbf{d} \| \) denotes Euclidean norm—is below a desired precision threshold. Here, \( \mathbf{n}(i) \) is the ground state density in the external potential \( \psi(\mathbf{i}) \) at the iteration step \( i \), and \( \gamma(i) > 0 \) is adjusted during the iterations to speed up the convergence. Since interacting inversion necessitates several ground state computations to attain an approximate solution, it is not efficient. Even more sophisticated iteration schemes cannot circumvent that some densities require incredibly many iterations, i.e. ground state computations, until convergence [28]. Therefore, in general, interacting inversion represents a computationally demanding task.

3. Exact LDA

As the exact form of \( E_{xc} \) from (4) is not known, in practice, approximations are used. One of the simplest and most successful approximations is the LDA [2, 20].

The exact LDA \( E_{xc}^{\text{LDA}} \) is defined via the homogeneous electron gas, i.e. via exactly homogeneous densities \( \mathbf{n} = (n_1, n_2, \ldots, n_L)^T = (n, n, \ldots, n)^T \) in the thermodynamic limit \( L \to \infty \) [2, 20]:

\[
e_{xc}^{\text{LDA}}(\mathbf{n}) := \lim_{L \to \infty} E_{xc}(n, n, \ldots, n)/L.
\]

This quantity is then used to approximate the xc energy of a finite system by

\[
E_{xc}(\mathbf{n}) \approx E_{xc}^{\text{LDA}}(\mathbf{n}) := \sum_{l=1}^{L} e_{xc}^{\text{LDA}}(n_{l}).
\]

Because \( E_{xc}^{\text{LDA}}(\mathbf{n}) \) is the exact xc energy for exactly homogeneous densities in the thermodynamic limit, it represents a good approximation for relatively homogeneous densities on large lattices \( L \gg 1 \).

Our feasibility study here assumes a relatively small lattice of size \( L = 21 \) with hard-wall boundary conditions, such that finite size and boundary effects play a role. We therefore derive our own LDA for this system and do not make use of the existing results in [29]. Figure 1 shows our \( e_{xc}(\mathbf{n}) \) obtained from numerically exactly homogeneous densities computed by means of non-interacting (for \( T^{n}(\mathbf{n}) \)) and interacting (for \( F_{HK}(\mathbf{n}) \)) inversions on \( L = 21 \) lattice sites for all possible total particle numbers \( N = 0, 1, \ldots, 42 \). To allow for an efficient evaluation of \( E_{xc}^{\text{LDA}} \), we parametrize the function \( e_{xc}(\mathbf{n}) \) using a finite number of parameters. A simple way to achieve this is to assume a polynomial form \( p_{d}(\mathbf{n}) \) of certain degree \( d \) and to fit our results using different values of \( d \). In the fit of each polynomial \( p_{d}(\mathbf{n}) \), we impose the physically reasonable constraint \( p_{d}(0) = 0 = p_{d}(2) \), which trivially holds for the exact \( E_{xc} \), as can be seen in (4): Obviously \( e_{xc}(0) = E_{xc}(0, 0, \ldots, 0)/L = 0 \) because
every term in (4) vanishes independently for zero total particle number, and $e_{xc}(2) = E_{xc}(2, 2, \ldots, 2)/L = 0$ because $F_{IK}(2, 2, \ldots, 2) = E_{II}(2, 2, \ldots, 2)$ and $T^{2}(2, 2, \ldots, 2) = 0$ due to impossible tunneling. Apparently, our exact LDA $e_{xc}$ is well approximated by polynomials of low degree since, on the scale of figure 1, the $d = 8$ fit seems indistinguishable from the $d = 4$ fit.

The LDA resides on the lowest rung of ‘Jacob’s ladder’ [17] and the most successful approximations of $E_{xc}$ beyond the LDA were built on top of it [6–12]. Analogously, we will use the LDA computed above as our reference, and we will try to improve upon it with a more general ansatz for the functional.

4. Our ansatz

Our approach for the construction of an improved xc energy approximation consists of two parts. Firstly, it requires an efficient variational density functional ansatz, denoted by $G$, to approximate $E_{xc}$. Secondly, a set of $M$ external potentials has to be specified, which will be called training scenario, such that the corresponding $M$ exact ground state densities $n'$, called training densities, and exact values $E_{xc}(n')$ are used to determine $G$ by minimizing a cost function

$$d(G) = \sum_{i=1}^{M} [E_{xc}(n') - G(n')]^2$$

over the variational parameters of $G$.

We are interested in an ansatz that, firstly, includes the LDA and, secondly, allows for a systematic improvement over it by including non-local terms. In this spirit, we propose a two-site ansatz of the following form:

$$G^X(n) = \sum_{k=0}^{X} G^{k}(n)$$

with

$$G^{k=0}(n) = \sum_{i=1}^{L} g^{0}(n_i),$$

$$G^{k>0}(n) = \sum_{i=1}^{L-k} g^{k}(n_i, n_{i+k}).$$

For $X = 0$, we have $G(n) = G^{X=0}(n) = G^{k=0}(n)$, which is completely analogous to the LDA (6). And for $X > 0$, the $k > 0$ terms allow for a more general dependence on the density with two-site functions over a range limited by $X$. In this way, increasing $X$ allows us to systematically include more non-local information and to go beyond the local LDA.

In order to have a practical functional, we want to write it in terms of a discrete set of variational parameters. Thus we need to restrict the form of the functions $g^{k}$. For simplicity we choose here a polynomial form for each term, as we did in the previous section 3 for the reference LDA. Additionally, in all following numerical experiments, we simply fix the degree of the polynomial to $d = 4$. 

Figure 1. Our exact LDA $e_{xc}(n)$ (crosses) and polynomial interpolations $p_d(n)$ of degree $d = 2$ (dotted), 4 (dashed), and 8 (solid).
When $G(n)$ is assumed to be a polynomial of the $n_l$, the variational parameters of $G$ are the polynomial coefficients. Then the desired $\text{argmin}_d d(G)$, i.e. the argument of $d$ that minimizes the cost function, results from the solution of linear equations $Ac = Ec$, where the polynomial coefficients of $G$ are vectorized in $c$, the exact values are vectorized in $E_{xc}$, and the elements in the matrix $A$ establish the correct connection to the cost function (7): $d(G) = \sum_{l=1}^{M}[E_{xc}(n_l') - G(n_l')]^2 = \sum_{l=1}^{M}[E_{xc} - (Ac)]^2$.

We want to emphasize that this approach does not need any computationally demanding interacting inversion. Because the training densities $n'$ follow from the training potentials, i.e. from $M$ different choices of $v^{\text{ext}}$ in (3a), we know $F_{\text{LDA}}(n')$. While the calculation of $E_{\text{LDA}}(n')$ is trivial, $T^{\text{\text{LDA}}}(n')$ is computed via efficient non-interacting inversion. Thus, all further ingredients for $E_{xc}(n')$ of (4) are then efficiently computable.

The first step in our approach is to consider the ansatz $G = G^d$, with $g^d(n_l) = \sum_{l=0}^{d} c^d_l n_l^d$ a polynomial in $n_l$ of degree $d$ with coefficients $c^d_l$ (and $d = 4$ in the following). The simplest possible, ‘homogeneous’, training scenario amounts to setting $v^{\text{ext}} = 0$, in which case we have one training density for each possible total particle number $N = 1, 2, \ldots, 2L$, i.e. at most $M = 42$ for $L = 21$. Figure 2 demonstrates how training our local term $g^4$ with such ground states reproduces the exact LDA. Remarkably, a very good match between our $g^4$ and the exact LDA is achieved already with $M = 30$ ground state computations. This has to be compared to the several thousands of ground state computations that were required for figure 1 due to the interacting inversion iteration.

We can understand that the ‘homogeneous’ training scenario leads to the exact LDA because this scenario contains relatively homogeneous training densities and because the exact LDA is constructed from exactly homogeneous densities. Now we want to consider more inhomogeneous densities. For that purpose we propose two different ways of converging $G$ with this ‘step’ training scenario. In the first way, we pick $M$ ground states randomly and study the convergence of $G$ as a function of $M$. In the second way, we fix the total particle numbers considered to $N = 1, 2, \ldots, 12$, take all possible step positions and orientations, and increase $M$ systematically together with the step height. In both schemes, convergence is quantified by comparison of the solution for $M$ with the solution for the largest considered $M_{\text{max}}$, which we fixed to 12 800 for the random and to 9612 for the systematic densities. We can then look at the quantity

$$\epsilon(M) = \langle |C_i(M) - C_i(M_{\text{max}})|/|C_i(M_{\text{max}})| \rangle,$$

where $C_i(M)$ denotes the $i$th parameter of $G$ after training with $M$ densities and $\langle \ldots \rangle = 1/P\sum_{i=1}^{\ldots}$ denotes taking the mean value over all $P$ possible values of $i$. 
Figure 4 shows our results for random densities. Interestingly, this training scenario gives rise to local terms \( g_0 \) that are very similar to the exact LDA, although many training densities of this scenario are very inhomogeneous. Furthermore, we can read off from the inset of figure 4 that convergence occurs rapidly.

To go beyond the LDA, we now include the longer-range two-site terms with \( k > 0 \) using a general polynomial ansatz:

\[
g^k(n_l, n_{l+1}) := \sum_{\delta_k \in \mathbb{Z}} c_{\delta_k} n_l^\delta n_{l+1}^k,
\]

i.e. these terms are general degree \( d \) polynomials of the density values on two lattice sites separated by distance \( k \) (and, again, we simply fix \( d = 4 \) in the following). While, as discussed above, for \( X = 0 \), our ansatz \( G^0(n) = \sum_{l=1}^{L} g^0(n_l) \) is completely analogous to the LDA of (6), for \( X > 0 \), it contains additional non-local terms, such that, by systematically increasing \( X \) in our ansatz \( G^X(n) \), we can systematically increase its non-locality beyond the local LDA-like term.
We enforce in our desired solution $G^X$ that the terms $g^k$ for increasing $k$ are obtained one after another such that each additional non-local term (i.e. corresponding to the next larger value of $k$) is a correction to the previous solution. This means that, for given $X$, we first minimize (7) only via the parameters $c^0$ which gives $G^0$. Then we minimize (7) for the remainder $E_{xc}^0(n') := E_{xc}(n') - G^0(n')$ only via the parameters $c^1$ which together with the previous solution $c^0$ gives $G^1$. Then we minimize (7) for the remainder $E_{xc}^1(n') := E_{xc}(n') - G^1(n')$ only via the parameters $c^2$ which together with the previous solutions $c^0$ and $c^1$ gives $G^2$. We continue the scheme until we have reached $c^X$ and thus $G^X$. This procedure ensures that each longer-range term is built on top of all previous shorter-ranged ones, in the same way as the functionals on higher rungs of ‘Jacob’s ladder’ are more non-local and are built on top of the more local functionals on the lower rungs [17].

In order to analyze the performance of the ansatz, we adopt now a different strategy. We will now always fit our ansatz $G^X$ with the ‘step’ training scenario of figure 3 and then we will apply it to completely different target densities. For the latter, we choose ground states of the $H_2$ dissociation problem, i.e. Hamiltonian (1) with total particle number $N = 2$ and external potential

$$\hat{V}(R) = \sum_{i=1}^{L} v_{i}^{\text{ext}}(R) \hat{n}_i + \frac{1}{\sqrt{R^2 + 1}},$$

where $R$ denotes the separation between the two $H$ atoms placed in the middle of the lattice such that we set $l_0 = 11$ for $L = 21$. Because this problem represents a realistic physical application that is significantly different from our training scenario, we consider it to be a very good benchmark for our approach.

From now on, the local terms $g^k(n)$ in our two-site polynomial ansatz (8) are fixed to be the exact LDA from figure 1. And the non-local terms $g^{k>0}(n_1, n_2)$ are enforced to fulfill $g^{k>0}(0, 0) = 0 = g^{k>0}(2, 2)$ as well as $g^{k>0}(n_1, n_2) = g^{k>0}(n_2, n_1)$. These properties are physically reasonable and they reduce the number of variational parameters, which turned out to be beneficial for the convergence of our fit. In particular, this helped us to avoid the effect known as overfitting. We distinguish two different versions of our ansatz, namely constrained and unconstrained. In our constrained two-site polynomial ansatz we determine $g^{k>0}(n_1, n_2)$ under the constraint $g^{k>0}(n, n) = 0$; then $G^X(n)$ is exact for exactly homogeneous densities $n = (n, n, \ldots, n)^T$. In our unconstrained two-site polynomial ansatz we do not impose this constraint on the polynomial coefficients: the unconstrained ansatz has thus more variational parameters than the constrained ansatz.

Figure 5 shows our results after convergence with $M$ systematic densities. As we can see in (a), the constrained ansatz leads to a visible improvement over LDA close to $R = 0$, but not for larger $R$. A convergence of the non-local terms can be concluded from the inset. In (b), the unconstrained ansatz leads to an improvement over LDA for almost all values of $R$, but it produces too low energy values close to $R = 0$. The inset of (b) demonstrates that convergence of the non-local terms occurs, however, for $k > 1$ this convergence is slower than in (a). With increasing $X$, our ansatz systematically improves the LDA result at specific values of $R$: around $R = 0$ when the constrained version is used, and at larger $R$ when the unconstrained version is used. Both versions of our ansatz show a systematic improvement over LDA with increasing $X$ when the mean energy for all values of $R$ is considered. In fact, such a mean value is the correct figure of merit because the cost function (7), minimized for ‘step’ training densities by our ansatz, is also a mean value of many xc energies.

Clearly, we would like to use $G^X(n)$ to compute densities self-consistently via the KS cycle. A first step in this direction is the calculation of the xc potential $v_{xc}^X(n) := -\partial E_{xc}/\partial n_{E_{xc}}$ for the exact ground state density $n$. In the $H_2$ dissociation problem, the xc potential for larger values of $R$ is particularly interesting, since its exact form exhibits a characteristic peak that cannot be reproduced by LDA alone [30]. Figure 6 shows our results for $R = 5$. While our constrained ansatz leads to a potential that basically coincides with the one from LDA, our unconstrained ansatz leads to a small systematic improvement with increasing $X$.

### 5. Conclusions

We have analyzed the feasibility of constructing semi-empirical approximations for the xc density functional in the context of a long-range interacting many-electron system on a one-dimensional lattice. Using numerically exact ground states from MPS simulations, we proposed to fit an ansatz that includes an LDA-like part plus additional terms of increasing non-locality, by means of reasonably chosen training densities. We observed that our ansatz converges systematically within the training scenario. Additionally, when applied to completely different target densities, namely of the $H_2$ dissociation problem, our fitted ansatz improved upon the LDA systematically. This systematic improvement was demonstrated for the ground state energies of the $H_2$ problem and for a xc potential corresponding to a stretched $H_2$ molecule.
In this work, we have tested the effect of a systematic inclusion of non-local ingredients in the functional using very simple ansätze for the non-local terms, namely only two-site dependences, and for the functional form, namely only polynomials, and we considered only one training scenario. Our results show that by systematically including non-local terms, the approximation can without doubt be improved beyond LDA. Our quantitative results are nevertheless limited to the specific form of the ansatz and training set used. For instance, the fact that the dissociation curve does not significantly improve by including terms of longer range after some point seems to indicate that other non-local contributions may be more relevant. Likewise, considering functional forms for each term that go beyond a polynomial may improve the power of the ansatz. We have already run initial tests to study the effect of terms that depend on three variable densities, but we observe no clear convergence with as many as \( \approx 20000 \) densities from the ‘step’ training scenario. This clearly indicates the necessity for a different training scenario and possibly additional physical constraints that effectively reduce the number of variational parameters. The question itself of how to choose the training densities optimally is, in general, a very important one that should be further explored, as a better training scenario would always further improve our results. All in all, although such improved ansätze and training scenarios must definitely achieve better results than the ones reported here, a careful analysis is beyond the scope of this proof-of-principle work.

It would be very interesting to combine our procedure with concepts from recent works on machine learning of density functionals [31–36]. On the one hand, these works typically required less training densities than our approach. On the other hand, our work constructs systematic corrections to a standard approximation, namely the LDA, that can be applied in general, i.e. to other types of systems beyond those that were originally used for the fit. Thus, a combination of the good aspects of our procedure with the good aspects of the previous machine learning concepts could be the ultimate solution to some of the problems that both approaches currently have independently from each other.

Figure 5. \( H_2 \) dissociation energy \( E(R) \) as a function of the separation \( R \): from exact LDA (dotted) and from our constrained (a) and unconstrained (b) ansatz with \( X = 1 \) (dash-double dotted), 2 (dashed–dotted), 5 (dashed), compared to the exact solution (solid). Insets: mean relative difference \( \epsilon(M) \) between the coefficients of \( g^{X}k \) for \( k = 1 \) (dash–double dotted), 2 (dashed–dotted), and 5 (dashed) after training with \( M \) systematic densities and the corresponding coefficients of \( g^{X}k \) after training with \( M_{\text{max}} = 9612 \) systematic densities—the insets show our results for \( M = 1212, 2412, \) and 4812. Here, our ansätze for \( G \) are polynomials of degree \( d = 4 \).

In this work, we have tested the effect of a systematic inclusion of non-local ingredients in the functional using very simple ansätze for the non-local terms, namely only two-site dependences, and for the functional form, namely only polynomials, and we considered only one training scenario. Our results show that by systematically including non-local terms, the approximation can without doubt be improved beyond LDA. Our quantitative results are nevertheless limited to the specific form of the ansatz and training set used. For instance, the fact that the dissociation curve does not significantly improve by including terms of longer range after some point seems to indicate that other non-local contributions may be more relevant. Likewise, considering functional forms for each term that go beyond a polynomial may improve the power of the ansatz. We have already run initial tests to study the effect of terms that depend on three variable densities, but we observe no clear convergence with as many as \( \approx 20000 \) densities from the ‘step’ training scenario. This clearly indicates the necessity for a different training scenario and possibly additional physical constraints that effectively reduce the number of variational parameters. The question itself of how to choose the training densities optimally is, in general, a very important one that should be further explored, as a better training scenario would always further improve our results. All in all, although such improved ansätze and training scenarios must definitely achieve better results than the ones reported here, a careful analysis is beyond the scope of this proof-of-principle work.

It would be very interesting to combine our procedure with concepts from recent works on machine learning of density functionals. On the one hand, these works typically required less training densities than our approach. On the other hand, our work constructs systematic corrections to a standard approximation, namely the LDA, that can be applied in general, i.e. to other types of systems beyond those that were originally used for the fit. Thus, a combination of the good aspects of our procedure with the good aspects of the previous machine learning concepts could be the ultimate solution to some of the problems that both approaches currently have independently from each other.
In this article, we focused exclusively on approximations of the density functional for the xc energy. However, in principle, any ground state observable can be written as a functional of the ground state density. Our proposed scheme allows in principle to also construct systematically density functionals for observables other than the ground state energy. This fact might now be useful for ultracold atoms in optical lattices since the densities of these systems have become experimentally accessible with remarkable resolution.

Measurements of a certain observable which are difficult to perform with the current techniques might be easier to carry out now when a density functional would be provided for that observable. Particularly interesting measurements regard two-dimensional systems with special observables, such as e.g. special order parameters. The corresponding density functionals could be constructed with the help of projected entangled pair states (PEPS). Because experiments are restricted to finite system sizes, PEPS algorithms are perfectly suited for the ground state simulation of such finite two-dimensional systems, for which they can produce accurate numerical results.

Acknowledgments

M L is very grateful to Neepa T Maitra and Garnet K-L Chan for discussions. He acknowledges funding by the EU through SIQS grant (FP7 600645) and the DFG (NIM cluster of excellence). He also thanks the Pedro Pascual Benasque Center for Science (CCBPP), where he carried out part of this work. AR acknowledges financial support from the European Research Council (ERC-2010-AdG-267374), Spanish grant (FIS2013-46159-C3-1-P), Grupos Consolidados (IT578-13).

References


Figure 6. Exchange–correlation potential $v_{xc}$ for the exact ground state density of the $H_2$ dissociation curve at $R = 5$: from exact LDA (dotted) and from our constrained (a) and unconstrained (b) ansatz with $X = 1$ (dash-double dotted), 2 (dashed–dotted), 5 (dashed), compared to the exact solution (solid). Our ansätze for $G$ here are the same as the ones in figure 5.
[12] Zhao Y and Truhlar D G 2008 The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals Theor. Chem. Accounts 120 215–41