Tensor product expansions for correlation in quantum many-body systems

Gábor Csányi
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

T. A. Arias
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853

(Received 2 November 1999)

We explore a class of computationally feasible approximations of the two-body density matrix as a finite sum of tensor products of single-particle operators. Physical symmetries then uniquely determine the two-body matrix in terms of the one-body matrix. Representing dynamical correlation alone as a single tensor product results in a theory that predicts near zero dynamical correlation in the homogeneous electron gas at moderate to high densities. But, representing both dynamical and statistical correlation effects together as a tensor product leads to the recently proposed ’’natural orbital functional.’’ We find that this latter theory has some asymptotic properties consistent with established many-body theory but is no more accurate than Hartree-Fock in describing the homogeneous electron gas for the range of densities typically found in the valence regions of solids.

The fundamental difficulties associated with ab initio solutions of the quantum-mechanical many-body problem stem directly from the large dimensionality of the wave function. Rather than dealing directly with the many-body wave function, large scale electronic structure calculations determine the ground-state energy by minimizing an energy functional over the much more manageable space of single-particle orbitals. Hohenberg-Kohn-Sham theory\textsuperscript{1,2} places this approach on a firm theoretical footing by establishing that exact electronic ground-state energies can be determined in this manner by minimizing a universal functional of such orbitals.

Although the requisite universal energy functional is not known exactly, the local density approximation\textsuperscript{2} and various improvements thereof\textsuperscript{3–7} are sufficient to resolve bond energies to a little better than one-tenth of an electron Volt. With this accuracy, these functionals make possible highly predictive first-principles studies in diverse areas such as the study of surfaces, point defects, plastic deformation, and chemical reactions. (For a review see Ref. 8.) Although sufficient for many studies, this error is too large to allow accurate prediction of the rates of microscopic processes at room temperature, thus limiting the ultimate predictive power of the ab initio density-functional approach. Whereas, further improvement of energy functionals for single-particle theories is an active area of research, much less work has been done to construct energy functionals of the one-body density matrix.\textsuperscript{3,9} Such functionals are promising because the density matrix contains more explicit information than do the Kohn-Sham orbitals, and an accurate energy functional thus is likely to be simpler in form. The exact kinetic-energy functional, for instance, is known for the density matrix, but not the Kohn-Sham orbitals.

To date, the only extensively used density-matrix energy functional is the Hartree-Fock approximation. One may view Hartree-Fock theory as approximating the two-body matrix as a sum of Hartree and exchange terms, each of which is a tensor product of the one-body density matrix. Hartree-Fock then uses the known, exact energy functional of the two-body density matrix to produce an energy functional of the one-body density matrix. In this paper, we build upon this perspective and consider tensor product approximations to the two-body density matrix, which not only serve as generators of energy functionals of the one-body density matrix, but also provide estimates of the two-body matrix in terms of the one-body matrix, and thereby shed light on the nature of correlation in quantum many-body systems.

Notation. For a system of $N$ electrons, the exact energy functional of the two-body density matrix $\gamma(x_1x_2, x'_1x'_2)$

\[ E = \int dx_1 \left[ -\frac{1}{2} \nabla^2 r_1 + U(r_1) \right] n(x_1, x'_1) \bigg|_{x'_1 = x_1} + \int dx_1 dx_2 \frac{\gamma(x_1x_2, x_1x'_2)}{|r_1 - r_2|}. \]  

(1)

Here and throughout, we work in atomic units, and the $x_i$ are compound coordinates representing both position $\mathbf{r}_i$ and spin $s_i$, so that integration over $x_i$ represents integration over space and summation over spin channels. The external potential and one-body density matrix are $U(\mathbf{r})$ and $n(x_1, x'_1) = \langle \Psi | \hat{\psi}^\dagger(x'_1) \hat{\psi}(x_1) | \Psi \rangle$, respectively. Finally, the one-body density matrix comes directly from $\gamma$ through the sum rule

\[ \left( \frac{N-1}{2} \right) n(x_1x'_1) = \int dx_2 \gamma(x_1x_2, x'_1x_2). \]  

(2)

Computational considerations. Although Eq. (1) is exact, the two-body density matrix $\gamma$ is a function of four variables, and direct computation with such functions is infeasible. Visible techniques, however, exist for dealing directly with two-variable functions such as the one-body density matrix.\textsuperscript{10–15} Accordingly, we expand $\gamma$ in terms of two-variable functions,
\[ \hat{\gamma} = \sum_i \hat{g}_i \otimes \hat{h}_i, \]  
(3)

which, by separation of variables, is always possible. Here, \( \hat{g}_i \) and \( \hat{h}_i \) are one-body operators (functions of two variables), we view the two-body density matrix as a two-body operator (a function of four variables), and the tensor product denotes one of the three possible choices for separating four variables,

\[ \hat{g} \otimes \hat{h} = g(x_1, x'_1) h(x_2, x'_2), \]

\[ \hat{g} \otimes \hat{h} = g(x_1, x_2) h(x'_2, x'_1), \]

\[ \hat{g} \otimes \hat{h} = g(x_1, x'_2) h(x_2, x'_1). \]

**Quantum statistical considerations.** We wish to maximize the physical content of a truncation of the expansion (3) to a limited number of terms. Physically, \( n(x_1, x'_1) \) is the quantum amplitude for the insertion of a hole at position \( x_1 \) and its instantaneous removal from \( x'_1 \), whereas \( \gamma(x_1 x_2 x'_1 x'_2) \) is the quantum amplitude for the insertion of a pair of holes at \( x_1, x_2 \) and their removal from \( x'_1, x'_2 \). Approximating the latter events as independent gives \( \gamma = \gamma_H + \gamma_{xc} \), the familiar Hartree approximation, where the factor of 2 maintains the normalization implicit in Eq. (2). We may then refine this mean-field behavior and define \( \gamma = \gamma_H + \gamma_{xc} \), where \( \gamma_{xc} \) represents exchange and correlation effects, which we then expand, without loss of generality, according to Eq. (3).

The most significant drawback of the Hartree approximation is that \( \gamma_H \) is not properly antisymmetric with respect to particle exchange \( (x'_1 \leftrightarrow x_2) \), so that, potentially, many terms will be required in \( \gamma_{xc} \) to restore this symmetry. Alternately, we may explicitly ensure antisymmetry and take \( \gamma = \gamma_{HF} + \gamma_{xc} \), where \( \gamma_{HF} \) is precisely the familiar Hartree-Fock approximation. Again, we may expand the unknown \( \gamma_{xc} \) as in Eq. (3).

Finally, we note that a fundamental difficulty exists when using a finite number of terms of Type II. Because of the instantaneous nature of the quantum event that \( \gamma \) represents, we expect for normal systems that \( \gamma \rightarrow 0 \) as the insertion \( x'_1 \) and removal \( x_1 \) locations are placed at ever further distances from one another. In an extended system with the expansion for \( \gamma \) containing a finite sum of terms of Type II, taking this limit while keeping the insertion points near one another and the removal points near one another will violate this asymptotic condition unless each Type-II term is identically zero. Accordingly, we focus below on approximating \( \gamma_{xc} \) and \( \gamma_{HF} \), with terms of Type I and III.

**Symmetry.** Physical symmetries of \( \gamma \) significantly restrict the allowable forms for the tensor products appearing in expansions of the form (3) for \( \gamma_{xc} \) and \( \gamma_{HF} \). We first consider Hermiticity \( \gamma(x_1 x_2, x'_1 x'_2) = \gamma^\dagger(x_1 x'_2, x_2 x'_1) \) and particle permutation symmetry \( \gamma(x_1 x_2, x'_1 x'_2) = \gamma(x_1 x'_1, x_2 x'_2) \) for each of the three types of tensor product, and defer discussion of Fermionic antisymmetry until after the sum rule immediately below. Because both \( \gamma_H \) and \( \gamma_{HF} \) respect these symmetries, so must tensor products representing \( \gamma_{xc} \) and \( \gamma_{HF} \). These symmetries imply relations with arbitrary constants of proportionality between the one-body operators of such products, which may be absorbed into the one-body operators. Table I summarizes the final result of such considerations for all three tensor products. We note that in all cases symmetry restricts the tensor product to involve only a single one-body operator, which remains to be determined.

**Sum rule.** Remarkably, when a single term satisfying Hermiticity and particle permutation is used to represent \( \gamma_{xc} \) or \( \gamma_{HF} \), we can "invert" the sum rule (2) and determine the two-body density matrix \( \gamma \) in terms of the one-body matrix \( n \).

Inserting the resulting form for \( \gamma \) into Eq. (1) then generates an energy functional of the one-body matrix, which implicitly satisfies the sum rule, an important property for obtaining good ground-state energies.

When terms of Type I represent \( \gamma_{xc} \) or \( \gamma_{HF} \), the sum rule becomes \( \hat{g} (\text{Tr} \hat{g} ) = \hat{\delta} \), where \( \hat{\delta} = \hat{n} \), where \( \hat{n} = \hat{n}_1 \) (1 = \hat{n}_2, respectively). The solution of this equation is \( \hat{g} = \pm \sqrt{\delta} / \sqrt{\text{Tr} \hat{\delta}} \). In either case, the extensivity of \( \text{Tr} \hat{\delta} \) makes the resulting solution irrelevant in the thermodynamic limit \( N \rightarrow \infty \). For products of Type III, the sum rule combined with appropriate symmetries gives \( \hat{g}^2 = \hat{\delta} \), so that \( \hat{g} = \sqrt{\hat{\delta}} \) with \( \hat{\delta} \) defined as above.

**Fermionic antisymmetry.** Fermionic antisymmetry, \( \gamma(x_1 x_2, x'_1 x'_2) = -\gamma(x_1 x'_2, x_2 x'_1) \), is a stronger condition than the particle permutation symmetry condition considered above and imposes more complex constraints. Two-body density matrices \( \gamma \) that satisfy this condition may always be written with Type-I and -III products appearing in corresponding pairs of the form \( \hat{g}_1 \hat{g}_1 - \hat{g}_2 \hat{g}_2 \) and with Type-II products (were we to consider such) whose one-body terms are separately antisymmetric.

The two viable functionals remaining after the above considerations involve a Type-III representation for either \( \gamma_{xc} \) or \( \gamma_{HF} \). Unfortunately, these forms do not consist of symmetric pairs of Type-I and -III products. To satisfy antisymmetry, one could take \( \gamma_{xc} = (\hat{n}_1 \hat{n}_2 - \hat{n}_2 \hat{n}_1) / 2 \), so that \( \gamma = (\hat{n}_1 \hat{n}_2 - \hat{n}_2 \hat{n}_1) / 2 \), which is simply the Hartree-Fock approximation. To go beyond this, antisymmetry requires representing \( \gamma \) as a pair of terms so that \( \gamma = (\hat{n}_1 \hat{n}_2 - \hat{n}_2 \hat{n}_1 + \hat{g}_1 \hat{g}_1 - \hat{g}_2 \hat{g}_2) / 2 \).

The sum rule for this latter extension on Hartree-Fock is \( \hat{g}^2 - \hat{g} (\text{Tr} \hat{g} ) = \hat{n}(1 - \hat{n}) \), which has no solution for \( \hat{g} \) in extended systems unless \( \text{Tr} \hat{g} \) vanishes in the thermodynamic limit. Under this condition, we have the solution \( \hat{g} = \pm \sqrt{\hat{n}(1 - \hat{n})} \), where the signs of the eigenvalues in the square root must be chosen to ensure \( \text{Tr} \hat{g} = 0 \). For the paramagnetic phase we consider below, the natural choice is to take the upper sign for the spin-up block of the density matrix and the lower sign for the spin-down block. The structure of the energy functional (1), however, is such that for this choice the resulting energy functional of \( \hat{n} \) is equal to

<table>
<thead>
<tr>
<th>Hermiticity</th>
<th>Particle Permutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{g} \otimes \hat{h} )</td>
<td>( g = \hat{g}, \hat{h} = \hat{h} )</td>
</tr>
<tr>
<td>( \hat{g}_u \otimes \hat{h} )</td>
<td>( g = \hat{h}, \hat{g} + \hat{g}^\dagger = \pm \hat{h} )</td>
</tr>
<tr>
<td>( \hat{g}_u \otimes \hat{h}^\dagger )</td>
<td>( g = \hat{h}^\dagger, \hat{g} = \pm \hat{h} )</td>
</tr>
</tbody>
</table>

### Table I. Implications of Hermiticity and particle permutation symmetries for tensor product expansions of \( \gamma \).
that generated by representing \( \hat{\gamma} \) as a single Type -III product.

**Representation of the functionals.** The above discussion leaves the energy of two representations of the two-body matrix to explore: corrected Hartree theory, \( \hat{\gamma} = \hat{\gamma}_{\text{H}} - (\sqrt{\hat{n}} \hat{n}^\dagger \sqrt{\hat{n}})/2 \), and corrected Hartree-Fock theory, \( \hat{\gamma} = \hat{\gamma}_{\text{HF}} - [\sqrt{(1-\hat{n})} \hat{n}^\dagger \sqrt{(1-\hat{n})}]/2 \).

The corresponding energy functionals may be represented either directly in terms of the one-body density matrix for use with direct density-matrix methods, or in terms of the spectral (“natural orbital”) representation of the density matrix.

The energy functional (1) contains one-body terms that may be evaluated directly in terms of the one-body density matrix. The remaining two-body term may be computed in terms of the two-point density \( \epsilon(r_1, r_2) = \sum_{ij} \gamma(x_i x_j, x_1 x_2) \). When the above functionals represent paramagnetic states, this density has the following form in terms of the eigenvectors (natural orbitals) \( \phi_i \) and corresponding eigenvalues (occupancies) \( f_i \) of the one-body density matrix \( \hat{n} \),

\[
c(r_1, r_2) = \frac{1}{2} \left[ 4 \sum_{ij} f_i \phi_j(r_1) \phi_j^*(r_2) \phi_i^*(r_1) \phi_i(r_2) - 2 \sum_{ij} g_{ij} \phi_i(r_1) \phi_j^*(r_2) \phi_i^*(r_1) \phi_j(r_2) \right],
\]

where for corrected Hartree and Hartree-Fock theories, we take \( g_{ij} = f_i f_j \) and \( g_{ij} = f_i f_j + \sqrt{f_i (1-f_i)} f_j (1-f_j) \), respectively.

The first of the above forms corresponds to the natural orbital functional proposed by Goedecker and Umrigar as an ansatz among the many possible choices for the \( g_{ij} \), which satisfy the sum rule. Here, we see that this form is the unique Type-III correction to Hartree theory, which satisfies basic symmetries and the sum rule. This theory has been studied analytically in the case of the homogeneous electron gas in the low-density regime \( (r_\tau > 5.77) \). Here, we present results for both this theory and our new corrected Hartree-Fock theory over a range of densities that includes both this low-density regime and the regime more physically relevant in the valence region of solids, \( r_\tau < 5.77 \).

**Homogeneous electron gas.** Translational invariance in the homogeneous gas ensures that the natural orbitals are plane waves. The remaining freedom in \( \hat{n} \) lies in the occupation eigenvalues \( f_k \). For the paramagnetic phase, the energy (1) then takes the following form:

\[
E = 2 \int \frac{Vd\mathbf{k}}{(2\pi)^3} \frac{k^2}{2} f_k - \frac{1}{V} \int \frac{V^2 d\mathbf{k} d\mathbf{k}'}{(2\pi)^6} g_{kk'} \frac{4\pi}{|\mathbf{k}^\prime - \mathbf{k}|^2},
\]

where \( V \) is the volume of the system and the factor of 2 in the kinetic energy arises from the sum over spin.

The minimum of this energy functional occurs when \( \delta(E - \mu N)/\delta f_k = 0 \). For both corrected Hartree and corrected Hartree-Fock, this gives to leading order as \( |k| \to \infty \),

\[
\delta(E - \mu N)/\delta f_k \sim k^4 - C \sqrt{k},
\]

where \( C \) is a constant. Hence, for large \( k \), the momentum density scales as \( f_k \sim |k|^{-8} \), in agreement with the random phase approximation and other many-body calculations.

This analysis holds for the present two functionals at all particle number densities and, in fact, for all theories of the form of Eq. (5) in which \( g_{kk'} \sim \sqrt{\mathbf{k}} \) as \( k \to \infty \) and \( f_k \to 0 \).

**Numerical Results.** To determine the total energy of the homogeneous electron gas within the corrected Hartree and Hartree-Fock approximations, we exploit spherical symmetry and reduce \( f_k \) to a single variable function \( f(k) \), which we represent on a radial mesh of variable spacing. We then minimize Eq. (5) with conjugate-gradients techniques subject to both the constraint that \( f(k) \) corresponds to a total of \( N \) electrons and the Fermi constraint that \( 0 \leq f(k) \leq 1 \). The latter constraint we implement by defining \( f(k) = (e^{-x} + 1)^{-2} \), where \( x \) is then free to range over the real line.

To evaluate the energy functional in terms of the fillings, we integrate numerically the kinetic and exchange-correlation terms in Eq. (5). A key step in producing accurate results is the kinetic energy functional and exchange-correlation terms once by parts analytically and then evaluating the result numerically. For the corrected Hartree-Fock case, we found it more efficient to approximate the Coulomb potential by a Yukawa potential, \( e^{-\kappa r}/r \), and numerically extrapolate the results to \( \kappa = 0 \). For these latter calculations, we found it more efficacious to work in the grand canonical ensemble (holding the chemical potential fixed while minimizing \( E - \mu N \)) than to impose the particle-number constraint explicitly.

The above procedures introduce four sources of numerical error: the finite spacing of the radial mesh, the finite extent of the mesh in \( k \) space, the error in the extrapolation in the case of using the Yukawa potential, and the finite number of iterations involved in our searches for the minima. Through analysis and numerical experiments, we determine these errors to be less than \( 10^{-3}, 10^{-4}, 10^{-3}, \) and \( 10^{-6} \) Hartree per particle, respectively, over the range of densities which we explored \( (0.01 < r_\tau < 30) \).

We found the minimization of the above functionals problematic for two reasons. First and foremost, the discretized representation of the one-dimensional function \( f(k) \) on the radial mesh artificially stabilizes any discontinuities in \( f(k) \), which may arise during the minimization process. This may be managed by taking care to ensure that the radial mesh provides sufficient resolution to describe any large gradients that arise in \( f(k) \). An additional factor making the calculations difficult is that both the corrected Hartree and corrected Hartree-Fock functionals suffer from a very large condition number (in excess of \( 10^6 \)).

Figure 1 compares our results for the correlation energy per particle of the electron gas with established many-body predictions. We find that at very low densities, both new theories begin to approach one another as \( r_\tau \to \infty \). For even moderately low densities \( (r_\tau > 6) \), both theories perform significantly worse than Hartree-Fock, overestimating the correlation energy by factors from 2 to 4 and higher. For the densities experienced in the valence regions of solids \( (6 > r_\tau > 1) \), corrected Hartree-Fock theory (diamonds) crosses over from overcorrelation to undercorrelation, whereas corrected Hartree theory (circles) overcorrelates by roughly the same amount by which traditional Hartree-Fock theory undercorrelates. For higher densities more representative of
atomic cores ($1 > r_s$), the corrected Hartree-Fock theory predicts near zero correlation. At these same densities, corrected Hartree theory remains overcorrelated and gives a somewhat better estimate of the correlation energy than does traditional Hartree-Fock. The asymptotic form of the corrected Hartree results in the high-density limit appears to be a straight line in our plot, which corresponds to the form of the leading order terms of the Gell-Mann-Brueckner expansion. However, the prefactors which we find by a least squares fit, $E = 0.0570 \log(r_s) - 0.0714$, are only correct to within a little better than a factor of 2; the well-known analytic values for these coefficients are 0.0311 and −0.0480, respectively. This suggests that the favorable reports for atoms for the natural orbital functional may be in part a consequence of the high densities experienced in atomic cores, the limited particle number in the low-density regions, and the self-interaction corrected nature of those calculations.

Figure 2 shows momentum distributions for $r_s = 7$ and $r_s = 1$, which are representative of what we find at low and high densities, respectively. We note that the numerical results of both theories exhibit the correct $k^{-8}$ scaling at large momenta and that we have very good agreement with the analytic results for the natural orbital functional (applicable for $r_s > 5.77$).

We observe that at low densities the momentum distributions of both corrected Hartree theory and corrected Hartree-Fock theory resemble one another quite closely and, in fact, approach one another as $r_s$ approaches infinity. This occurs because in this limit both theories predict very low occupations $f_i \ll 1$ so that the central quantities in the two theories approach one another, $f_i f_j + \sqrt{f_i (1-f_i) f_j (1-f_j)}$.

The next panel of the figure illustrates that for densities outside the range of the applicability of the analytic solution ($r_s < 5.77$), this solution predicts $f > 1$ for sufficiently low momenta. We find that for such densities, the momentum distributions of corrected Hartree and corrected Hartree-Fock theory exhibit different behaviors at low momenta. Our numerical results indicate (to better than 1 part in $10^5$) that the momentum distribution of corrected Hartree theory saturates to $f = 1$ near the Fermi level. (See inset.) Such behavior is in direct contradiction of analytic and numerical many-body results for the electron gas and is contrary to the previously reported experience with atoms for this functional. In contrast, the momentum distribution of corrected Hartree-Fock theory never saturates the Fermi occupation constraint, but rather approaches a value at $k = 0$, which remains less than unity. We find that in the high density limit the corrected Hartree-Fock momentum distribution in fact approaches the standard Hartree-Fock result, ultimately leading to zero correlation energy for this theory in the high-density limit.

In conclusion, we find that when the two-body density matrix is expanded beyond traditional Hartree and Hartree-Fock theories in terms of tensor products of one-body operators, fundamental symmetries allow the exchange-correlation hole sum rule to be inverted to give estimates for the two-body matrix, and thus correlations, in terms of the one-body matrix. By approximating the two-body density matrix in terms of the one-body density matrix, this approach also leads to energy functionals of the one-body density matrix.

The addition of the first separable term beyond Hartree theory shows some promising features when applied to the homogeneous electron gas, such as giving the correct scaling behavior of the momentum distribution in the high-momentum limit and the correct scaling behavior of the correlation energy in the high-density limit. However, the electronic states predicted by this functional are significantly overcorrelated, and attempts to improve upon it by the addition of further terms without additional constraints would only add more freedom in the minimization and lead to further overcorrelation. Hence, before adding a second term, it is important to use antisymmetry to constrain the first such term to be the traditional Fock term. This gives the second functional that we considered, corrected Hartree-Fock theory. This second theory undercorrelates the electron gas...
in the high-density limit but still overcorrelates at lower densities. This indicates that the addition of further terms will only hold the promise of producing a more accurate functional if additional, appropriate constraints are imposed.

C.S.G. would like to thank Sohrab Ismail-Beigi for useful discussions. This work was supported primarily by the MRSEC Program of the National Science Foundation under Grant No. DMR 94-00334.