

FIRST-ORDER SEMIDEFINITE PROGRAMMING FOR THE TWO-ELECTRON TREATMENT OF MANY-ELECTRON ATOMS AND MOLECULES

DAVID A. MAZZIOTTI¹

Abstract. The ground-state energy and properties of any many-electron atom or molecule may be rigorously computed by variationally computing the two-electron reduced density matrix rather than the many-electron wavefunction. While early attempts fifty years ago to compute the ground-state 2-RDM directly were stymied because the 2-RDM must be constrained to represent an N -electron wavefunction, recent advances in theory and optimization have made direct computation of the 2-RDM possible. The constraints in the variational calculation of the 2-RDM require a special optimization known as a semidefinite programming. Development of first-order semidefinite programming for the 2-RDM method has reduced the computational costs of the calculation by orders of magnitude [Mazziotti, *Phys. Rev. Lett.* **93** (2004) 213001]. The variational 2-RDM approach is effective at capturing multi-reference correlation effects that are especially important at non-equilibrium molecular geometries. Recent work on 2-RDM methods will be reviewed and illustrated with particular emphasis on the importance of advances in large-scale semidefinite programming.

Mathematics Subject Classification. 90C22, 81Q05, 52A40.

Received September 10, 2005.

1. INTRODUCTION

Because electrons are indistinguishable with only pairwise interactions, the energy of any atom or molecule may be expressed as a linear functional of the *two-electron reduced density matrix* (2-RDM) [32, 33]. This functional tantalizingly suggests that a variational calculation of the wavefunction might be replaced with a variational determination of the 2-RDM. The first attempts to use the 2-RDM fifty years ago, however, yielded energies that were unphysically low. The literature [13, 57] realized that the 2-RDM must be constrained to represent an N -electron wavefunction. In 1963 the mathematician John Coleman called these constraints *N -representability conditions* [8]. Coleman [7, 8] and mathematicians Claude Garrod and Jerome Percus [19] derived a set of N -representability conditions that required restricting three two-particle density matrices to be positive semidefinite. A matrix is *positive semidefinite* if and only if it has nonnegative eigenvalues. While these conditions were initially implemented in the mid-1970s for four-electron systems like beryllium and the helium dimer, uncertainties in the strength of the conditions and the feasibility of the optimization caused interest in the development of 2-RDM methods to wane.

Keywords and phrases. Semidefinite programming, electron correlation, reduced density matrices, N -representability conditions.

¹ Department of Chemistry and the James Franck Institute, The University of Chicago, Chicago, IL 60637, USA.
damazz@uchicago.edu

The variational calculation of the 2-RDM constrained by necessary N -representability conditions yields a *lower bound* on the ground-state energy in a given finite basis set. The strict lower bound occurs because the 2-RDM is optimized over a set that contains all correlated N -electron wavefunctions. Within wave mechanics the challenge is to introduce sufficient variational flexibility into the wavefunction, but in reduced-density-matrix mechanics the challenge is to limit the 2-RDM, which has the flexibility to model all correlated wavefunctions, to represent only realistic N -electron wavefunctions. A practical consequence is that the 2-RDM method has the potential to produce ground-state energies with useful accuracy even if the wavefunction is challenging to parameterize as in transition-state structures or other stretched geometries of a potential energy surface [22, 40, 42, 43, 50].

In the mid-1990s the 2-RDM method re-emerged through the non-variational calculation of the 2-RDM by solving a contracted Schrödinger equation [1, 2, 9–11, 31, 34–39, 51, 52, 60]. This research led to a better understanding of the conditions of Coleman [8] and Garrod and Percus [19] as well as a systematic extension of these original conditions to form a hierarchy of conditions, which have been called p -positivity conditions [16, 47]. These conditions enforce large classes of the generalized uncertainty relations [47]. While research on the 2-RDM was receiving renewed attention, significant strides were made in developing optimization algorithms for minimizing a function with matrices of the variables constrained to be positive semidefinite. Research was particularly driven by applications in control theory, combinatorial optimization, and finance. This family of optimization problems is called *semidefinite programming* [58]. Built upon these two advances, the variational calculation of the ground-state 2-RDM has been achieved for a variety of atoms and molecules without the many-electron wavefunction [16, 20–28, 30, 40–47, 49, 50, 61]. The author has recently developed a first-order semidefinite-programming algorithm for the 2-RDM calculation that, by reducing the computational scaling by orders of magnitude, enables the treatment of larger molecules and basis sets [42, 43]. In this article we will review and discuss the development of the variational 2-RDM methods with emphasis on the role of semidefinite programming.

2. VARIATIONAL PRINCIPLE AND N -REPRESENTABILITY

In quantum mechanics the ground-state energy of any many-electron atom or molecule with Hamiltonian \hat{H} may be calculated in principle from the expectation value of the Hamiltonian with respect to the ground-state wavefunction Ψ

$$E = \langle \Psi | \hat{H} | \Psi \rangle. \quad (1)$$

Because the electrons interact pairwise, the many-electron Hamiltonians for all atoms and molecules may be written as

$$\hat{H} = \sum_{i,j,k,l} {}^2K_{k,l}^{i,j} \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k, \quad (2)$$

where \hat{a}^\dagger and \hat{a} are the second-quantized creation and annihilation operators, the indices refer to the spin-orbital basis set, and the symbol 2K denotes the two-electron reduced Hamiltonian matrix, containing the one- and two-electron integrals as well as the number N of electrons [34]. Substitution of the second-quantized Hamiltonian into equation (1), which emphasizes the two-particle nature of the electron interaction, generates an energy expression for many-electron systems

$$E = \sum_{i,j,k,l} {}^2K_{k,l}^{i,j} {}^2D_{k,l}^{i,j} \quad (3)$$

in terms of the two-electron reduced density matrix (2-RDM) [32, 33] which is defined in second-quantization as

$${}^2D_{k,l}^{i,j} = \langle \Psi | \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k | \Psi \rangle. \quad (4)$$

From a knowledge of the 2-RDM we can compute any two-electron properties of an atom or molecule.

The 2-RDM must be constrained by additional rules (or conditions) to derive from an N -electron wavefunction. Coleman described these necessary and sufficient rules as N -representability conditions [8]. Unsuccessful back-of-the-envelope 2-RDM calculations of Coleman [8] and Tredgold [57] employed four basic requirements for a density matrix of indistinguishable fermions [8]: the matrix should be (i) normalized to conserve particle number, (ii) Hermitian, (iii) antisymmetric under particle exchange, and (iv) positive semidefinite to keep probabilities nonnegative. Although these conditions are sufficient to guarantee that the 2-RDM is a density matrix, they are not sufficient for the matrix to be representable by an N -electron density matrix, or N -representable. What further conditions should be imposed on a 2-RDM to restrict it to be N -representable? Recent progress has revealed the importance of a class of N -representability conditions, called positivity conditions [16, 47], which was initially discussed by Coleman [8] and Garrod and Percus [19]. Erdahl and Jin [16] and Mazziotti and Erdahl [47] generalized these conditions to a hierarchy of N -representability conditions, and demonstrated that each level of the hierarchy corresponds to enforcing the generalized uncertainty relations for a class of operators.

Let us define a general set of operators $\{\hat{C}_i\}$ whose application to the ground-state wavefunction generates a complete set of N -particle basis functions $\{|\Phi_i\rangle\}$

$$|\Phi_i\rangle = \hat{C}_i|\Psi\rangle. \quad (5)$$

The metric (or overlap) matrix M for these basis functions

$$M_j^i = \langle\Phi_i|\Phi_j\rangle = \langle\Psi|\hat{C}_i^\dagger \hat{C}_j|\Psi\rangle \quad (6)$$

must be positive semidefinite, which we denote as $M \geq 0$. If the set of operators $\{\hat{C}_i\}$ is restricted so that the elements of the metric matrix can be evaluated only with a knowledge of the p -RDM, then the condition $M \geq 0$ is a p -positivity condition which restricts the N -representability of the p -RDM [47].

Consider the 1-positivity conditions whose metric matrices may be evaluated from a knowledge of the 1-RDM only. If for each \hat{C}_j a second-quantized operator \hat{a}_j which kills a particle in the j th orbital is chosen, we obtain that the 1-particle reduced density matrix (1-RDM), whose elements are

$${}^1D_j^i = \langle\Psi|\hat{a}_i^\dagger \hat{a}_j|\Psi\rangle, \quad (7)$$

must be positive semidefinite. Similarly, if for each \hat{C}_i an operator \hat{a}_i^\dagger is chosen which creates a particle in the i th orbital, we obtain that the 1-hole reduced density matrix (1-HRDM) whose elements

$${}^1Q_i^j = \langle\Psi|\hat{a}_j \hat{a}_i^\dagger|\Psi\rangle, \quad (8)$$

must be positive semidefinite. Physically, the positivity constraint ${}^1D \geq 0$ ensures that the probability for finding one electron in any orbital remains nonnegative while the complementary constraint ${}^1Q \geq 0$ ensures that the probability for not finding one electron (or finding a hole which is the absence of a particle) remains nonnegative.

The expectation value of the *fermionic* anticommutation relation

$$\hat{a}_j \hat{a}_i^\dagger + \hat{a}_i^\dagger \hat{a}_j = \delta_j^i \quad (9)$$

yields a linear mapping between the matrix elements of 1D and 1Q

$${}^1Q_i^j + {}^1D_j^i = \delta_j^i. \quad (10)$$

From this relation we observe that the eigenfunctions (or orbitals) that diagonalize the 1-particle RDM also diagonalize the 1-hole RDM because the Kronecker delta function is invariant under unitary transformations.

Therefore, the i th eigenvalue (or occupation number) of the 1-hole RDM is expressible in terms of the i th occupation number of the 1-particle RDM n_i as $1 - n_i$. The nonnegativity of the occupation numbers, $n_i \geq 0$ and $1 - n_i \geq 0$, from 1-positivity conditions implies the Pauli exclusion principle that the occupation numbers of the spin orbitals must lie between zero and one, that is $0 \leq n_i \leq 1$. Coleman [8] first proved in 1963 that the four basic conditions for a fermion density matrix plus the 1-positivity conditions are not only necessary but also sufficient constraints for the ensemble N -representability of the 1-RDM; in other words, every one-electron density matrix that obeys the 1-positivity conditions derives from at least one N -electron ensemble density matrix.

Let us now consider the metric matrices that can be evaluated with the 2-RDM only. If these matrices are constrained to be positive semidefinite, we obtain the 2-positivity conditions [19, 40, 47]. The basis functions in equation (5) that produce these 2-RDM metric matrices may be naturally grouped into three orthogonal vector spaces according to the number of creation operators in \hat{C} : (i) $\hat{C} \in \{\hat{a}_i \hat{a}_j\}$, (ii) $\hat{C} \in \{\hat{a}_i^\dagger \hat{a}_j^\dagger\}$, and (iii) $\hat{C} \in \{\hat{a}_i^\dagger \hat{a}_j\}$. The elements of the three metric matrices M from these groups may be expressed as:

$${}^2D_{i,j}^{p,q} = \langle \Psi | \hat{a}_q^\dagger \hat{a}_p^\dagger \hat{a}_i \hat{a}_j | \Psi \rangle, \quad (11)$$

$${}^2Q_{i,j}^{p,q} = \langle \Psi | \hat{a}_q \hat{a}_p \hat{a}_i^\dagger \hat{a}_j^\dagger | \Psi \rangle, \quad (12)$$

$${}^2G_{i,j}^{p,q} = \langle \Psi | \hat{a}_q^\dagger \hat{a}_p \hat{a}_i^\dagger \hat{a}_j | \Psi \rangle. \quad (13)$$

The condition ${}^2D \geq 0$ prevents the probability distribution for finding two particles in two orbitals from being anywhere negative while the condition ${}^2Q \geq 0$ enforces this nonnegativity for two holes. Analogously, the constraint ${}^2G \geq 0$ restricts the probability for finding a particle and a hole in any two spin orbitals from being negative.

By rearranging the creation and annihilation operators according to the anticommutation relations, we generate *linear* relations between pairs of the 2-positive metric matrices, 2D , 2G , and 2Q [40]:

$${}^2Q_{i,j}^{p,q}/2 = (\delta_i^p - 2 {}^1D_i^p) \wedge \delta_j^q + {}^2D_{i,j}^{p,q}/2 \quad (14)$$

and

$${}^2G_{i,j}^{p,q} = \delta_p^i {}^1D_j^q - {}^2D_{p,j}^{i,q}, \quad (15)$$

where the 1-RDM arises from the integration (or contraction in a matrix representation) of the 2-RDM

$${}^1D_i^p = \frac{1}{N-1} \sum_j {}^2D_{i,j}^{p,j} \quad (16)$$

and the symbol \wedge denotes the antisymmetric tensor product known as the Grassmann wedge product [34]. It may be demonstrated by contraction that the 2-positivity conditions contain the 1-positivity conditions. Unlike the 1-positivity conditions for the 1-RDM, the 2-positivity conditions are necessary but not sufficient for restricting the 2-RDM to be N -representable. The 2-positivity conditions have proven fairly effective constraints for ground-state calculations of many-electron atoms and molecules at both equilibrium and non-equilibrium geometries [22, 40, 42, 43, 50].

What are the N -representability conditions not included by the 2-positivity conditions? Erdahl and Jin [16] and Mazziotti and Erdahl [47] have shown that the 2-positivity conditions may be generalized to p -positivity conditions that arise from the metric matrices that can be expressed from only a knowledge of the p -RDM. At the level of 3-positivity there are four distinct metric matrices for fermions that correspond to restricting the probability distributions to be nonnegative for three particles ${}^3D \geq 0$, two particles and one hole ${}^3E \geq 0$, one particle and two holes ${}^3F \geq 0$, and three holes ${}^3Q \geq 0$ [40, 47]. Calculations on spin systems and atoms and molecules [45] show that the accuracy of the energy and properties converges rapidly with the level of p -positivity with 3-positivity yielding excellent results at all interaction strengths even where traditional wavefunction

methods fail [26, 47, 56]. Two different types of particle 3-positivity conditions have been proposed: (i) the lifting conditions [26, 40], and (ii) the T_1 and T_2 conditions [14, 44, 61]. The latter conditions correspond to the following restrictions: $T_1 = {}^3D + {}^3Q \geq 0$ and $T_2 = {}^3E + {}^3F \geq 0$. Originally proposed by Erdahl [14], the T_1 and T_2 conditions have been applied by Zhao *et al.* [61] to computing the ground states of atoms and molecules at equilibrium geometries with similar or better accuracy than many standard wavefunction methods. The author has applied a spin- and symmetry-adapted version of the T_2 condition to molecules with similar accuracy at both equilibrium and stretched geometries in non-minimal basis sets [44].

The p -positivity conditions guarantee the nonnegativity of important probability distributions involving different mixtures of particles and holes. Their strength, however, may be further appreciated from their enforcement of important classes of the *generalized uncertainty relations* [47]. Mazziotti and Erdahl have shown that the $2p$ -positivity conditions enforce the generalized uncertainty relations for all pairs of Hermitian p -body operators \hat{C}_1 and \hat{C}_2 , that is

$$\sigma(\hat{C}_1)\sigma(\hat{C}_2) \geq \frac{1}{2}|\langle\Psi|[\hat{C}_1, \hat{C}_2]|\Psi\rangle|, \quad (17)$$

where

$$\sigma(\hat{C}) = \langle\Psi|(\hat{C} - \langle\Psi|\hat{C}|\Psi\rangle)^2|\Psi\rangle^{1/2} \quad (18)$$

and the expectation values may be evaluated with the $2p$ -RDMs. Therefore, the 2-positivity conditions imply the uncertainty relations for all pairs of one-body operators.

Unlike the optimization of a trial wavefunction that produces an upper bound on the ground-state energy by the Rayleigh-Ritz variational principle, the variational 2-RDM method with necessary N -representability constraints like the positivity conditions yields a lower bound on the ground-state energy. For this reason the variational 2-RDM method has also been called the lower bound method in the literature [16]. The lower bound on the energy, however, is only a signature for a more fundamental difference between the wavefunction and 2-RDM approaches to many-electron correlation. All wavefunction methods, whether variational or not, depend upon a parametrization of the N -electron wavefunction. In the variational case this parametrization leads to the energy upper bound of the Rayleigh-Ritz variational principle while in non-variational methods it generates accuracy through a certain order of perturbation theory. In the variational 2-RDM method, however, the set of 2-RDMs contains all N -representable 2-RDMs, and hence, all wavefunctions, whether single- or multi-referenced, are correctly parametrized. In fact, because the set of 2-RDMs defined by necessary positivity conditions is larger than the set of N -representable 2-RDMs, the 2-RDM is slightly “over-parameterized”, which leads to the lower bounds to the energies of N -electron atoms and molecules. Nevertheless, because the positivity conditions do not depend on a specific reference configuration like the Slater determinant for a single-reference wavefunction technique, the variational 2-RDM method can treat both single- and multi-reference correlation effects with consistent accuracy. As we will show in the applications, this is especially useful for describing the shape of potential energy surfaces.

3. SEMIDEFINITE PROGRAMMING

Variational calculation of the energy with respect to the 2-RDM constrained by 2-positivity conditions requires minimizing the energy in equation (3) while restricting the 2D , 2Q , and 2G to be not only positive semidefinite but also interrelated by the linear mappings in equations (14–16). This is a special optimization problem known as a *semidefinite program*. The solution of a semidefinite program is known as *semidefinite programming* [54, 58, 59].

A semidefinite program may be written in two complementary formulations which are known as the *primal* and *dual programs*. For convenience we define the map M which transforms any vector $|x\rangle$ of length n^2 into an $n \times n$ matrix $M(x)$ by creating each column of the matrix sequentially with the elements of the vector.

The *primal* formulation of the semidefinite program may be expressed in general notation as

$$\begin{aligned} & \text{minimize} && \langle c|x\rangle \\ & \text{such that} && A|x\rangle = |b\rangle \\ & && M(x) \geq 0, \end{aligned} \tag{19}$$

where the vector $|c\rangle$ defines the system, the vector $|x\rangle$ denotes the primal solution, the $m \times n$ matrix A and the m -dimensional vector $|b\rangle$ enforce m linear constraints upon the solution $|x\rangle$, and the matrix $M(x)$ is restricted to be positive semi-definite. Similarly, the *dual* formulation of the semidefinite program may be expressed generally as

$$\begin{aligned} & \text{maximize} && \langle b|y\rangle \\ & \text{such that} && |z\rangle = |c\rangle - A^T|y\rangle \\ & && M(z) \geq 0, \end{aligned} \tag{20}$$

where the vector $|y\rangle$ of length m is the dual solution, A^T is the $n \times m$ transpose of the matrix A , and the $n \times n$ matrix $M(z)$ is constrained to be positive semidefinite.

In the variational 2-RDM method with 2-positivity the solution $|x\rangle$ of the primal program is a vector of the three metric matrices from the 2-RDM, the 2D , the 2Q , and the 2G matrices

$$M(x) = \begin{pmatrix} {}^2D & 0 & 0 \\ 0 & {}^2Q & 0 \\ 0 & 0 & {}^2G \end{pmatrix}, \tag{21}$$

the vector $|c\rangle$ holds specific information about the quantum system in the form of the two-particle reduced Hamiltonian

$$M(c) = \begin{pmatrix} {}^2K & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{22}$$

and the matrix A and the vector $|b\rangle$ contain the linear mappings among the 2D , 2Q , and 2G matrices in equations (14) and (15), the contraction (Eq. (16)) and trace conditions ($\text{Tr}({}^1D) = N$), as well as any spin constraints. The constraint $M(x) \geq 0$ in equation (19) restricts the 2D , 2Q , and 2G matrices to be positive semidefinite.

Feasible $|x\rangle$ and $|y\rangle$ give upper and lower bounds on the optimal value of the objective function, which in the 2-RDM problem is the ground-state energy in a finite basis set. The primal and dual solutions, $|x\rangle$ and $|y\rangle$, are *feasible* if they satisfy the primal and dual constraints in equations (19) and (20) respectively. The difference between the feasible primal and the dual objective values, called the *duality gap* μ , which equals the inner product of the vectors $|x\rangle$ and $|z\rangle$,

$$\mu = \langle x|z\rangle \geq 0, \tag{23}$$

vanishes if and only if the solution is a global extremum. This optimization theorem for semidefinite programming was first proved by Erdahl in 1979 in the context of 2-RDM theory [15]. For the variational 2-RDM method the duality gap furnishes us with a mathematical guarantee that we have determined the optimal energy within the convex set defined by the positivity conditions. With necessary N -representability conditions the optimal energy is a *lower-bound* to the energy from full configuration interaction in the selected basis set.

In the mid-1990s a powerful family of algorithms, known as *primal-dual interior-point algorithms*, was developed for solving semidefinite programs [58]. The phrase *interior-point* means that the method keeps the trial primal and dual solutions on the *interior* of the feasible set throughout the solution process. In these algorithms a good initial guess for the 2-RDM is a scalar multiple of the two-particle identity matrix. Advantages of the interior-point methods are: (i) rapid quadratic convergence from the identity matrix to the optimal 2-RDM

for a set of positivity conditions; and (ii) a rigorous criterion in the duality gap for convergence to the global minimum. These benefits, however, are accompanied by large memory requirements and a significant number of floating-point operations per iteration, specifically $O(nm^3 + n^2m^2)$ where n is the number of variables and m is the number of constraints. With m and n proportional to the number of elements in the 2-RDM ($\approx r^4$), the method scales approximately as r^{16} where r is the rank of the one-particle basis set [40, 41]. The variational 2-RDM method has been explored for minimal basis sets with the primal-dual interior-point algorithm, but the computational scaling significantly limits both the number of active electrons and the size of the basis set [30, 40, 41, 49, 50, 61].

The author has recently developed a large-scale semidefinite programming algorithm for solving the semidefinite program in the variational 2-RDM method [42, 43]. The optimization challenge in the 2-RDM method is to constrain the metric matrices to be positive semidefinite while the ground-state energy is minimized. The algorithm constrains the solution matrix M to be positive semidefinite by a matrix factorization

$$M = RR^*, \quad (24)$$

where for the 2-positivity conditions M is given in equation (21). Such a matrix factorization was previously considered in the context of 2-RDM theory by Mihailović and Rosina [48], Harriman [29], and the author [34], and it was recently employed for solving large-scale semidefinite programs in combinatorial optimization [5]. The applications in Mazziotti [42, 43] and Burer and Choi [4] are the first to apply the matrix factorization to semidefinite programs with multiple diagonal blocks in the solution matrix M . The linear constraints, including the trace, the contraction, and the interrelations between the metric matrices, become quadratic in the new independent variables R . Therefore, the factorization in equation (24) converts the semidefinite program into a *nonlinear program* where the energy must be minimized with respect to R while nonlinear constraint equalities are enforced. Burer and Choi [4] examine semidefinite programs which permit a low-rank factorization of the solution matrix M . The semidefinite program for electronic structure in references [42, 43] has a solution matrix M whose rank is approximately half of its dimension. Even without a low rank for M , however, the matrix factorization in an electronic problem has a much lower cost than the matrix operations in interior-point methods.

We solve the nonlinear formulation of the semidefinite program by the augmented Lagrange multiplier method for constrained nonlinear optimization [3, 5, 12, 17, 42, 43, 55]. Consider the augmented Lagrangian function

$$L(R) = E(R) - \sum_i \lambda_i c_i(R) + \frac{1}{\mu} \sum_i c_i(R)^2, \quad (25)$$

where R is the matrix factor for the solution matrix M , $E(R)$ is the ground-state energy as a function of R , $\{c_i(R)\}$ is the set of equality constraints, $\{\lambda_i\}$ is the set of Lagrange multipliers, and μ is the penalty parameter. For an appropriate set of multipliers $\{\lambda_i\}$ the minimum of the Lagrangian function with respect to R corresponds to the minimum of the energy $E(R)$ subject to the nonlinear constraints $c_i(R)$. The positive third term in the augmented Lagrangian function, known as the quadratic penalty function, tends to zero as the constraints are satisfied.

The augmented Lagrange multiplier algorithm finds the energy minimum of the constrained problem with an iterative, three-step procedure:

- (1) For a given set of Lagrange multipliers $\{\lambda_i^{(n)}\}$ and penalty parameter $\mu^{(n)}$, minimize the Lagrangian function $L(R)$ to obtain an improved estimate R_{n+1} of the factorized 2-RDM at the energy minimum.
- (2) If the maximum absolute error in the constraints $\max\{c_i(R_{n+1})\}$ is below a chosen threshold (*i.e.* $0.25 \max\{c_i(R_n)\}$), then the Lagrange multipliers are updated by a first-order correction

$$\lambda_i^{(n+1)} = \lambda_i^{(n)} - c_i(R_{n+1})/\mu^{(n)},$$

while the penalty parameter remains the same

$$\mu^{(n+1)} = \mu^{(n)}.$$

- (3) If step 2 is not executed, then the penalty parameter is decreased to enforce better the constraints

$$\mu^{(n+1)} = 0.1\mu^{(n)},$$

while the Lagrange multipliers remain the same

$$\lambda^{(n+1)} = \lambda^{(n)}.$$

Steps 1–3 are repeated until the maximum absolute error in the constraints falls below a target threshold. Before the first iteration the Lagrange multipliers may be initialized to zero, and the penalty parameter set to 0.1. The constraints are not fully enforced until convergence, and the energy in the primal program approaches the optimal value from below. Even though the matrix factorization of the primal matrix produces a nonconvex optimization problem in equation (25), Burer and Monteiro [6] have shown that the Lagrange multiplier algorithm converges to the optimal (global) solution of the semidefinite program.

The cost of the algorithm is dominated by r^6 floating-point operations [42], mainly from the matrix multiplication of the block-diagonal R matrix with itself, where r is the rank of the one-particle basis set. Storage of the factorized 2-RDM, several copies of its gradient, and the Lagrange multipliers scales as r^4 . In comparison with the primal-dual interior-point approach, which scales as r^{16} and r^8 in floating-point operations and memory storage, the first-order nonlinear algorithm for the variational 2-RDM method [42, 43] provides a significant improvement in computational efficiency.

4. APPLICATIONS OF THE 2-RDM

The variational 2-RDM method has been applied to a variety of atoms and molecules at both equilibrium and stretched geometries. We illustrate the method briefly here by reviewing a calculation of the potential energy curve of the hydrogen fluoride molecule [44]. The 2-RDM method with 2-positivity plus the T_2 condition produces a semidefinite program for hydrogen fluoride (after applying spin and symmetry adaptation [23]) with 336 664 variables, 168 432 constraints, 650 583 non-zero elements in the constraint matrix A , and 38 blocks in the solution matrix M . Semidefinite programs with several million variables and constraints have been solved with the method [24, 27, 28, 44, 45]. Comparisons of the first-order algorithm with second-order primal-dual interior-point algorithms have been reported in the literature [23, 42].

For selected bond distances along the potential energy curve of HF in a valence double-zeta basis set the ground-state energies from the variational 2-RDM method with 2-positivity conditions (DQG) as well as 2-positivity and T_2 conditions (DQGT2) are compared in Figure 1 with the energies from several wavefunction methods including Hartree-Fock (HF), second-order many-body perturbation theory (MP2), coupled-cluster singles-doubles (CCSD), coupled-cluster singles-doubles with a perturbative triples correction (CCSD(T)) and full configuration interaction (FCI). The core orbital of fluorine is frozen. The 2-RDM curve with DQG conditions reveals the presence of a systematic shift of the curve from the lower-bound nature of the 2-RDM method which, however, does not affect the *shape* of the potential energy curve. The shape of the curve is most important for chemistry because it determines spectroscopic constants and reaction kinetics. In calculations on nitrogen and carbon dimers as well as bond stretching in organic molecules we have easily adjusted the 2-RDM curves by shifting them by a constant to agree with CCSD(T) at the equilibrium molecular geometry [21, 22]. Addition of the T_2 condition to the calculation improves the absolute errors in the energies. In the region around its equilibrium geometry the energies from the 2-RDM method with 2-positivity plus T_2 (DQGT2) conditions are as accurate as the energies from coupled cluster with a perturbative triples correction (CCSD(T)). For all approximate methods considered the absolute magnitude of the error in the energy increases with the bond distance, but it remains the smallest in the 2-RDM method with the DQGT2 constraints.

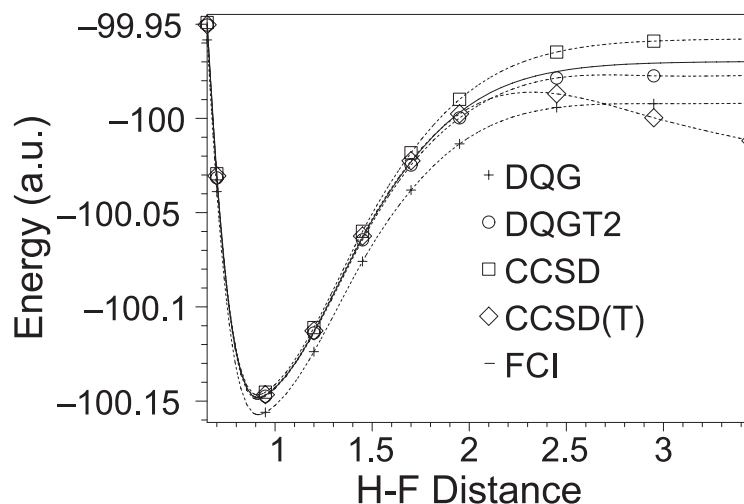


FIGURE 1. Potential energy curves of HF are generated in a valence double-zeta basis set from the two variational 2-RDM methods, one with the 2-positivity conditions (DQG) and one with both the 2-positivity and the T_2 conditions (DQGT2). Although the absolute magnitude of the error in the energy increases with bond distance for all approximate methods considered, it remains the smallest in the 2-RDM method with DQGT2 constraints.

5. A LOOK AHEAD

Since early work in the 1950s recognized the promise and challenges of computing the energies and properties of atoms and molecules without the many-electron wavefunction, quantum chemistry has experienced many important advances towards the accurate treatment of electron correlation including the development of density functional theory, coupled cluster theory, Monte Carlo methods, and multi-reference perturbation theory. The recent progress in the 2-RDM methods contributes both a new perspective and tool for describing energies and properties of atoms and molecules in which correlation effects are accounted. The variational 2-RDM, which yields a systematic lower bound on the ground-state energy in a finite basis set, is especially applicable to treat molecular systems with difficult-to-parameterize wavefunctions that arise in chemistry at transition states and other nonequilibrium geometries as well as in describing many open-shell molecules. The 2-RDM method also has applications to the study of correlation phenomena in Bose condensation [20] and solid-state physics.

A critical part of realizing a “quantum chemistry without wavefunctions” through the variational 2-RDM method is the development of robust algorithms for large-scale semidefinite programming. As discussed in the article a large-scale algorithm, developed by the author, reduces the computational scaling of the 2-RDM method by orders of magnitude in both floating-point operations and memory. A key feature of the algorithm is the conversion of the semidefinite program into a nonlinear constrained optimization, which is then solved by the method of augmented Lagrange multipliers. Zhao *et al.* [61] has examined improving the performance of the primal-dual interior-point methods by redefining the statement of the semidefinite program, and M. Nayakkankuppam [53] has explored the use of a first-order spectral bundle algorithm. Further advances in large-scale semidefinite programming will have an important impact on the variational 2-RDM method, and similarly, the problems in electronic structure offer fertile ground for testing and benchmarking new large-scale algorithms. Advances will also have a broad impact on many other scientific problems in areas like control theory, combinatorial optimization, quantum information, and finance. Although still in its early stages, the 2-RDM method for computing energies and properties without the many-electron wavefunction offers a new approach to computing the electronic structure of atomic and molecular systems.

Acknowledgements. The author expresses his appreciation to Dudley Herschbach, Herschel Rabitz, John Coleman, and Alexander Mazziotti for their support and encouragement. The author thanks the NSF, the Henry-Camille Dreyfus Foundation, and the Alfred P. Sloan Foundation for their support.

REFERENCES

- [1] D.R. Alcoba, F.J. Casquero, L.M. Tel, E. Perez-Romero and C. Valdemoro, Convergence enhancement in the iterative solution of the second-order contracted Schrödinger equation. *Int. J. Quantum Chem.* **102** (2005) 620–628.
- [2] M.D. Benayoun, A.Y. Lu and D.A. Mazziotti, Invariance of the cumulant expansion under 1-particle unitary transformations in reduced density matrix theory. *Chem. Phys. Lett.* **387** (2004) 485–489.
- [3] D.P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, New York (1982).
- [4] S. Burer and C. Choi, Computational enhancements in low-rank semidefinite programming. *Optim. Methods Soft.* **21** (2006) 493–512.
- [5] S. Burer and R.D.C. Monteiro, Nonlinear programming algorithm for solving semidefinite programs *via* low-rank factorization. *Math. Program. Ser. B* **95** (2003) 329–357.
- [6] S. Burer and R.D.C. Monteiro, Local minima and convergence in low-rank semidefinite programming. *Math. Program. Ser. A* **103** (2005) 427–444.
- [7] L. Cohen and C. Frishberg, Hierarchy equations for reduced density matrices, *Phys. Rev. A* **13** (1976) 927–930.
- [8] A.J. Coleman, Structure of fermion density matrices. *Rev. Mod. Phys.* **35** (1963) 668.
- [9] A.J. Coleman and V.I. Yukalov, *Reduced Density Matrices: Coulson's Challenge*. Springer-Verlag, New York (2000).
- [10] F. Colmenero and C. Valdemoro, Approximating q -order reduced density-matrices in terms of the lower-order ones. 2. Applications. *Phys. Rev. A* **47** (1993) 979–985.
- [11] F. Colmenero and C. Valdemoro, Self-consistent approximate solution of the 2nd-order contracted Schrödinger equation. *Int. J. Quantum Chem.* **51** (1994) 369–388.
- [12] A.R. Conn, I.M. Gould and P.L. Toint, *Trust-Region Methods*. SIAM: Philadelphia (2000).
- [13] C.A. Coulson, Present state of molecular structure calculations. *Rev. Mod. Phys.* **32** (1960) 170–177.
- [14] R.M. Erdahl, Representability. *Int. J. Quantum Chem.* **13** (1978) 697–718.
- [15] R.M. Erdahl, Two algorithms for the lower bound method of reduced density matrix theory. *Reports Math. Phys.* **15** (1979) 147–162.
- [16] R.M. Erdahl and B. Jin, The lower bound method for reduced density matrices. *J. Mol. Struct. (Theochem)* **527** (2000) 207–220.
- [17] R. Fletcher, *Practical Methods of Optimization*. John Wiley and Sons, New York (1987).
- [18] M. Fukuda, B.J. Braams, M. Nakata, M.L. Overton, J.K. Percus, M. Yamashita and Z. Zhao, Large-scale semidefinite programs in electronic structure calculation. *Math. Program., Ser. B* **109** (2007) 553.
- [19] C. Garrod and J. Percus, Reduction of N -particle variational problem. *J. Math. Phys.* **5** (1964) 1756–1776.
- [20] G. Gidofalvi and D.A. Mazziotti, Boson correlation energies *via* variational minimization with the two-particle reduced density matrix: Exact N -representability conditions for harmonic interactions. *Phys. Rev. A* **69** (2004) 042511.
- [21] G. Gidofalvi and D.A. Mazziotti, Application of variational reduced-density-matrix theory to organic molecules. *J. Chem. Phys.* **122** (2005) 094107.
- [22] G. Gidofalvi and D.A. Mazziotti, Application of variational reduced-density-matrix theory to the potential energy surfaces of the nitrogen and carbon dimers. *J. Chem. Phys.* **122** (2005) 194104.
- [23] G. Gidofalvi and D.A. Mazziotti, Spin- and symmetry-adapted two-electron reduced-density-matrix theory. *Phys. Rev. A* **72** (2005) 052505.
- [24] G. Gidofalvi and D.A. Mazziotti, Potential energy surface of carbon monoxide in the presence and absence of an electric field using the two-electron reduced-density-matrix method. *J. Phys. Chem. A* **110** (2006) 5481–5486.
- [25] G. Gidofalvi and D.A. Mazziotti, Computation of quantum phase transitions by reduced-density-matrix mechanics. *Phys. Rev. A* **74** (2006) 012501.
- [26] J.R. Hammond and D.A. Mazziotti, Variational two-electron reduced-density-matrix theory: Partial 3-positivity conditions for N -representability. *Phys. Rev. A* **71** (2005) 062503.
- [27] J.R. Hammond and D.A. Mazziotti, Variational reduced-density-matrix calculations on radicals: a new approach to open-shell *ab initio* quantum chemistry. *Phys. Rev. A* **73** (2006) 012509.
- [28] J.R. Hammond and D.A. Mazziotti, Variational reduced-density-matrix calculation of the one-dimensional Hubbard model. *Phys. Rev. A* **73** (2006) 062505.
- [29] J.E. Harriman, Geometry of density matrices. II. Reduced density matrices and N -representability. *Phys. Rev. A* **17** (1978) 1257–1268.
- [30] T. Juhász and D.A. Mazziotti, Perturbation theory corrections to the two-particle reduced density matrix variational method. *J. Chem. Phys.* **121** (2004) 1201–1205.
- [31] W. Kutzelnigg and D. Mukherjee, Irreducible Brillouin conditions and contracted Schrödinger equations for n -electron systems. IV. Perturbative analysis. *J. Chem. Phys.* (2004) **120** 7350–7368.

- [32] P.O. Löwdin, Quantum theory of many-particle systems. 1. Physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configuration interaction. *Phys. Rev.* **97** (1955) 1474–1489.
- [33] J.E. Mayer, Electron correlation. *Phys. Rev.* **100** (1955) 1579–1586.
- [34] D.A. Mazziotti, Contracted Schrödinger equation: Determining quantum energies and two-particle density matrices without wave functions. *Phys. Rev. A* **57** (1998) 4219–4234.
- [35] D.A. Mazziotti, Approximate solution for electron correlation through the use of Schwinger probes. *Chem. Phys. Lett.* **289** (1998) 419–427.
- [36] D.A. Mazziotti, Pursuit of N -representability for the contracted Schrödinger equation through density-matrix reconstruction. *Phys. Rev. A* **60** (1999) 3618–3626.
- [37] D.A. Mazziotti, Comparison of contracted Schrödinger and coupled-cluster theories. *Phys. Rev. A* **60** (1999) 4396–4408.
- [38] D.A. Mazziotti, Correlated purification of reduced density matrices. *Phys. Rev. E* **65** (2002) 026704.
- [39] D.A. Mazziotti, A variational method for solving the contracted Schrödinger equation through a projection of the N -particle power method onto the two-particle space. *J. Chem. Phys.* **116** (2002) 1239–1249.
- [40] D.A. Mazziotti, Variational minimization of atomic and molecular ground-state energies via the two-particle reduced density matrix. *Phys. Rev. A* **65** (2002) 062511.
- [41] D.A. Mazziotti, Solution of the 1,3-contracted Schrödinger equation through positivity conditions on the 2-particle reduced density matrix. *Phys. Rev. A* **66** (2002) 062503.
- [42] D.A. Mazziotti, Realization of quantum chemistry without wavefunctions through first-order semidefinite programming. *Phys. Rev. Lett.* **93** (2004) 213001.
- [43] D.A. Mazziotti, First-order semidefinite programming for the direct determination of two-electron reduced density matrices with application to many-electron atoms and molecules. *J. Chem. Phys.* **121** (2004) 10957–10966.
- [44] D.A. Mazziotti, Variational two-electron reduced-density-matrix theory for many-electron atoms and molecules: Implementation of the spin- and symmetry-adapted T_2 condition through first-order semidefinite programming. *Phys. Rev. A* **72** (2005) 032510.
- [45] D.A. Mazziotti, Variational reduced-density-matrix method using three-particle N -representability conditions with application to many-electron molecules. *Phys. Rev. A* **74** (2006) 032501.
- [46] D.A. Mazziotti, *Reduced-Density-Matrix with Application to Many-electron Atoms and Molecules*, Advances in Chemical Physics **134**, D.A. Mazziotti Ed., John Wiley and Sons, New York (2007).
- [47] D.A. Mazziotti and R.M. Erdahl, Uncertainty relations and reduced density matrices: Mapping many-body quantum mechanics onto four particles. *Phys. Rev. A* **63** (2001) 042113.
- [48] M.V. Mihailović and M. Rosina, Excitations as ground-state variational parameters. *Nucl. Phys.* **A130** (1969) 386.
- [49] M. Nakata, H. Nakatsuji, M. Ehara, M. Fukuda, K. Nakata and K. Fujisawa, Variational calculations of fermion second-order reduced density matrices by semidefinite programming algorithm. *J. Chem. Phys.* **114** (2001) 8282–8292.
- [50] M. Nakata, M. Ehara and H. Nakatsuji, Density matrix variational theory: Application to the potential energy surfaces and strongly correlated systems. *J. Chem. Phys.* **116** (2002) 5432–5439.
- [51] H. Nakatsuji, Equation for the direct determination of the density matrix. *Phys. Rev. A* **14** (1976) 41–50.
- [52] H. Nakatsuji and K. Yasuda, Direct determination of the quantum-mechanical density matrix using the density equation. *Phys. Rev. Lett.* **76** (1996) 1039–1042.
- [53] M. Nayakkankuppam, Solving large-scale semidefinite programs in parallel. *Math. Program., Ser. B* **109** (2007) 477–504.
- [54] Y. Nesterov and A.S. Nemirovskii, *Interior Point Polynomial Method in Convex Programming: Theory and Applications*. SIAM: Philadelphia (1993).
- [55] E. Polak, *Optimization: Algorithms and Consistent Approximations*. Springer-Verlag, New York (1997).
- [56] J.H. Sebold and J.K. Percus, Model derived reduced density matrix restrictions for correlated fermions. *J. Chem. Phys.* **104** (1996) 6606–6612.
- [57] R.H. Tredgold, Density matrix and the many-body problem. *Phys. Rev.* **105** (1957) 1421–1423.
- [58] L. Vandenbergh and S. Boyd, Semidefinite programming. *SIAM Rev.* **38** (1996) 49–95.
- [59] S. Wright, *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia (1997).
- [60] K. Yasuda, and H. Nakatsuji, Direct determination of the quantum-mechanical density matrix using the density equation II. *Phys. Rev. A* **56** (1997) 2648–2657.
- [61] Z. Zhao, B.J. Braams, H. Fukuda, M.L. Overton and J.K. Percus, The reduced density matrix method for electronic structure calculations and the role of three-index representability conditions. *J. Chem. Phys.* **120** (2004) 2095–2104.