Fast noniterative orbital localization for large molecules

Francesco Aquilante[a] and Thomas Bondo Pedersen[b]
Department of Theoretical Chemistry, Chemical Center, University of Lund, P.O. Box 124, S-221 00 Lund, Sweden

Alfredo Sánchez de Merás
Department of Physical Chemistry, Institute of Molecular Science, University of Valencia, E-46100 Burjassot, Valencia, Spain

Henrik Koch
Department of Chemistry, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

(Received 9 June 2006; accepted 12 September 2006; published online 1 November 2006)

We use Cholesky decomposition of the density matrix in atomic orbital basis to define a new set of occupied molecular orbital coefficients. Analysis of the resulting orbitals ("Cholesky molecular orbitals") demonstrates their localized character inherited from the sparsity of the density matrix. Comparison with the results of traditional iterative localization schemes shows minor differences with respect to a number of suitable measures of locality, particularly the scaling with system size of orbital pair domains used in local correlation methods. The Cholesky procedure for generating orthonormal localized orbitals is noniterative and may be made linear scaling. Although our present implementation scales cubically, the algorithm is significantly faster than any of the conventional localization schemes. In addition, since this approach does not require starting orbitals, it will be useful in local correlation treatments on top of diagonalization-free Hartree-Fock optimization algorithms. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2360264]

I. INTRODUCTION

Localized occupied molecular orbitals (MOs) are of central importance in quantum chemistry for two reasons. First, they provide the link between Hartree-Fock (HF) theory and the concept of chemical bonds formed between two atoms[1,2] and second, they are indispensable for exploiting the short-range nature of electron correlation to achieve reductions in the computational effort.[3–13] The conventional formulation of HF theory, i.e., through diagonalization of the Fock matrix, leads to strongly delocalized canonical MOs. As HF theory is invariant under rotations among the occupied and among the virtual orbitals,[14] localized MOs may be obtained by a suitable unitary transformation of the canonical ones.

Several schemes have been developed for choosing such a unitary transformation of the occupied orbitals. While the scheme usually attributed to Boys[1,2,15] minimizes the spatial extent of the orbitals by maximizing the distances between orbital centroids, the Edmiston-Ruedenberg[15] (ER) procedure aims at maximizing the self-repulsion energy of the orbitals, thus minimizing the exchange energy between them. Preserving the separation between σ and π orbitals of double bonds, the latter gained some popularity over Boys localization which lacks this separation, producing instead the so-called banana (τ) orbitals which are linear combinations of σ and π orbitals. Formally, the ER optimization shows $\mathcal{N}^5$ computational scaling, where $\mathcal{N}$ is a suitable measure of the size of the system, whereas the Boys procedure scales as $\mathcal{N}^3$. For most applications, the lower computational scaling renders Boys localization preferable to ER localization. It must be stressed, however, that linear scaling is achievable in the ER scheme, provided that the initial orbitals are sufficiently local.[16] The most widely used localization scheme today was introduced by Pipek and Mezey[17] in 1989. The Pipek-Mezey (PM) localization procedure seeks to minimize the number of atomic centers over which each MO extends by maximizing the sum of squares of gross atomic Mulliken population of the MOs. The PM scheme presents a computational scaling of $\mathcal{N}^4$ (using a straightforward Jacobi sweep algorithm, see below); for a fixed number of occupied orbitals, however, it scales cubically with the basis set size. Moreover, like the ER procedure, the PM localization preserves σ-π separation. Besides the Boys, ER, and PM schemes, other sets of localized orbitals have been suggested (see, e.g., Refs. 18 and 19) with minor success and similar orbital localization schemes are commonly used in solid state physics.[20,21]

In practice, the Boys, ER, and PM procedures are formulated as an optimization problem in which a localization functional, $\zeta(\phi)$, is maximized with respect to rotations among the occupied orbitals. The orbital localization thus becomes an iterative procedure. The localization functionals for the three schemes may be written as

$$\zeta_{\mathrm{Boys}}(\phi) = \sum_i \langle \phi_i | \mathbf{r} | \phi_i \rangle^2, \tag{1}$$

$$\zeta_{\mathrm{ER}}(\phi) = \sum_i (\phi_i \phi_i | \phi_i \phi_i), \tag{2}$$

[a]Electronic mail: francesco.aquilante@teokem.lu.se
[b]Electronic mail: thomas.pedersen@teokem.lu.se

 © 2006 American Institute of Physics

Downloaded 29 Jan 2010 to 147.156.182.23. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

$$\zeta_{\mathrm{PM}}(\phi) = \sum_i \sum_A \langle \phi_i | P_A | \phi_i \rangle^2, \tag{3}$$

where the vector $\mathbf{r}$ is the electronic position operator, $(\phi_i \phi_j | \phi_k \phi_l)$ denotes a two-electron repulsion integral in the Mulliken notation, and P_A is the projection operator onto the space of atomic orbitals (AOs) centered on atom A. Indices i, j, k, and l denote occupied orbitals. The generally adopted optimization algorithm consists in consecutive pairwise rotations,[15,17] so-called Jacobi sweeps,[16] and may be very slowly convergent. Only recently a more modern approach to the optimization problem has been suggested. Subotnik *et al.*[16] presented a direct maximization based on an exponential parametrization of the unitary orbital rotation matrix and demonstrated that direct inversion of the iterative subspace (DIIS)-type algorithms can be applied to significantly speed up convergence.

Rather than defining locality through a functional, we observe that the sparsity of the one-electron density matrix in AO basis defines regions of interaction between the AOs and thus, indirectly, the locality of the MOs. Hence, using that the density matrix is positive semidefinite with rank equal to the number of occupied orbitals, we propose to define local MOs by Cholesky decomposition of the one-electron density matrix in AO basis. This has several advantages. First, Cholesky decomposition is a numerically stable and fast algorithm that can be made linear scaling[22] for matrices with linear scaling number of nonzero elements. Second, being a noniterative procedure, complicated optimization techniques are not needed. Third, as no initial orbitals need be given, the procedure is particularly well suited for determining local MOs directly from density matrix-based HF theory.[22–26] In this paper we discuss these Cholesky MOs in more detail for the case of a closed-shell restricted HF (RHF) wave function.

II. THEORY

Within the linear combination of atomic orbitals approach, the MOs are expanded in a basis of nonorthogonal AOs $\{\chi_\mu\}$ as

$$\phi_p = \sum_\mu \chi_\mu C_{\mu p}. \tag{4}$$

The orthonormality condition reads

$$\sum_{\mu\nu} C_{\mu p} S_{\mu\nu} C_{\nu q} = \delta_{pq}, \tag{5}$$

where $S_{\mu\nu} = \langle \chi_\mu | \chi_\nu \rangle$ is the AO overlap matrix and the orbitals are assumed real. Here and in the following greek subscripts are used to refer to AOs, whereas p and q are used to denote general MOs. For occupied and virtual MOs we use subscripts i,j and a,b, respectively. For simplicity, we restrict our discussion to closed-shell HF wave functions such that all occupied MOs are doubly occupied.

A. Occupied orbitals

For a closed-shell HF wave function, the one-electron density matrix in AO basis can be written in terms of the canonical MO coefficients as

$$D_{\mu\nu} = 2 \sum_i C_{\mu i} C_{\nu i}. \tag{6}$$

For insulators, the density matrix is local in the sense that $\rho(\mathbf{r},\mathbf{r}') = \Sigma_{\mu\nu} D_{\mu\nu} \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r}')$ decays exponentially with the distance $|\mathbf{r}-\mathbf{r}'|$; see, e.g., Ref. 27. The density matrix thus provides a measure of the interaction between different regions of space in the molecule and for this reason it is intimately related to the concept of orbital localization.[28,29] In addition, locality in the density matrix is crucial for achieving linear scaling in various electronic structure theories.[30–35]

Exploiting that the density matrix is positive semidefinite, we propose to compute a new set of occupied MOs that inherit locality from the density matrix by Cholesky decomposition (see Ref. 36 for a simple algorithm), i.e., we compute a set of Cholesky MOs, $\tilde{\phi}_i = \Sigma_\mu \chi_\mu X_{\mu i}$, such that

$$\sum_i X_{\mu i} X_{\nu i} = \tfrac{1}{2} D_{\mu\nu}. \tag{7}$$

We stress that the number of Cholesky MOs is exactly equal to the rank of the density matrix, i.e., the number of occupied orbitals. This is ensured by using full column pivoting, which is the standard algorithm for incomplete Cholesky decomposition. Pivoting is mandatory for numerical stability and, provided that there are no degeneracies among the density matrix diagonal elements at any iteration of the decomposition, ensures that the decomposition is unique. Degeneracies, however, certainly will occur whenever the molecular system is symmetric. In order to make the decomposition unique for such cases, one would have to incorporate some physically motivated algorithm for selecting among the degenerate density matrix diagonal elements. No attempt has been made at devicing such a selection algorithm and all results presented in this paper are thus determined by the ordering of atomic orbitals inherited from the HF program.

As demonstrated in the appendix, the Cholesky MOs constitute an orthonormal set (with metric $\mathbf{S}$). If needed, the orthogonal transformation $\mathbf{U}$ from canonical to Cholesky MOs,

$$X_{\mu i} = \sum_j C_{\mu j} U_{ji}, \tag{8}$$

can be computed according to

$$U_{ij} = \sum_{\mu\nu} C_{\mu i} S_{\mu\nu} X_{\nu j}. \tag{9}$$

The computational cost of a Cholesky decomposition of the density matrix formally is $N_o^2 N$, where N_o is the number of occupied MOs and N the number of AOs, and thus scales cubically with the size of the system. Using sparse matrix techniques, however, the scaling can be brought down to linear in exactly the same way as has been done for the Cholesky decomposition of the overlap matrix.[22] Our implementation uses dense matrix techniques (level 1 BLAS) and therefore scales cubically. Nevertheless, owing to the small prefactor and to being noniterative, the Cholesky localization is substantially faster than the conventional functional optimizations.

Downloaded 29 Jan 2010 to 147.156.182.23. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

Finally, for planar molecules the density matrix can be regarded as a sum of independent matrices derived from orbital components of different symmetry with respect to the plane of the molecule, namely, the σ or π canonical orbitals. Consequently, the σ-π separation is preserved by the Cholesky decomposition algorithm without imposing any *ad hoc* constraint.

B. Virtual orbitals

It is straightforward to extend the above scheme to virtual MOs by Cholesky decomposing the density-type matrix

$$D^v_{\mu\nu} = \sum_a C_{\mu a} C_{\nu a}, \tag{10}$$

yielding an orthonormal set of virtual Cholesky MOs, as shown in the Appendix. This approach requires knowledge of the virtual MOs and may therefore be of limited interest in practice.

Linearly dependent projected AOs (PAOs) spanning the virtual space are obtained from the AOs by projecting out components of the occupied space as[3,5,8]

$$\tilde{\chi}_\mu = \sum_\nu \chi_\nu \left(\mathbf{1} - \tfrac{1}{2}\mathbf{DS}\right)_{\nu\mu} \equiv \sum_\nu \chi_\nu Q_{\nu\mu}. \tag{11}$$

The linear dependence may be eliminated by diagonalizing the overlap matrix $\tilde{\mathbf{S}} = \mathbf{Q}^T\mathbf{S}\mathbf{Q}$ and subsequently removing orbitals corresponding to zero eigenvalues. In an exact formulation, this procedure is not only computationally expensive but may also lead to loss of locality. For these reasons, elimination of linear dependence is usually performed for each of the small subsets of PAOs that are accessible by excitation from a given occupied orbital (orbital domains).[8]

An alternative method is based on the density-type matrix

$$\tilde{D}_{\mu\nu} = \sum_\rho Q_{\mu\rho} Q_{\nu\rho}, \tag{12}$$

which is positive semidefinite by construction and its rank is exactly equal to the number of virtual orbitals, N_v. Thus, by Cholesky decomposition

$$\sum_a R_{\mu a} R_{\nu a} = \tilde{D}_{\mu\nu}, \tag{13}$$

we obtain N_v vectors $R_{\mu a}(a=1,2,\ldots,N_v)$ that constitute a linearly independent nonorthogonal set spanning the virtual space. A set of orthonormal Cholesky PAOs may be computed according to

$$Y_{\mu a} = \sum_b R_{\mu b} \bar{S}^{-1/2}_{ba}, \tag{14}$$

where $\bar{S}^{-1/2}_{ba} = (\bar{\mathbf{S}}^{-1/2})_{ba}$ and

$$\bar{S}_{ab} = \sum_{\mu\nu} R_{\mu a} S_{\mu\nu} R_{\nu b} \tag{15}$$

is the overlap matrix of the nonorthogonal Cholesky PAOs. According to the Carlson-Keller theorem,[37] the orthogonal orbitals thus constructed most closely resemble the initial nonorthogonal set. In this sense, they are expected to retain most of the localized character of $R_{\mu a}$, especially for compact atomic orbital sets.

As the $\mathbf{Q}$ matrix is sparse, it is reasonable to expect that the nonorthonormal Cholesky PAOs are about as local as the original PAOs. As demonstrated in the supplementary material,[38] however, the Cholesky PAOs are significantly less sparse than the linearly dependent PAOs and we shall not pursue virtual orbitals any further here.

TABLE I. Value of the various localization functionals for a given orbital set using the RHF/cc-pVDZ optimized wave function for linear $(Gly)_{20}$. The value of the Boys functional is given relative to the value obtained using canonical orbitals.

	ζ		
Orbital set	Boys	PM	ER
Canonical	0	93.2	349.7
Boys	211 763	221.1	498.7
PM	211 741	223.6	495.3
ER	211 761	221.2	500.3
Cholesky	211 183	190.8	450.3

III. SAMPLE CALCULATIONS

In order to investigate the properties of the Cholesky MOs we compare to the Boys, ER, and PM localized orbitals. While we use the (slowly convergent) method of Jacobi sweeps[15,17] for Boys and PM optimization, the ER orbitals are localized using the η steps of Subotnik *et al.*[16] without DIIS-type convergence enhancement. In all calculations we have optimized the HF wave function using Cholesky decomposed two-electron integrals.[39–42] All calculations are performed with a development version of the MOLCAS quantum chemistry software.[43,44]

The localization method presented here is expected to be effective for large molecules only. The one-electron density matrix of a small molecule is not sparse and the Cholesky MOs can therefore not be expected to be local. For alkanes of small size we have observed poor performance of the Cholesky localization. If applied to the benzene molecule, on the other hand, the Cholesky decomposition of the density matrix is able to produce orbitals that qualitatively resemble those obtained from the PM localization. This seems to be ascribable to the preservation of σ-π separation.

In this paper we focus on larger molecules for which the density matrix is expected to possess a certain degree of sparsity. In order to compare the Cholesky orbitals to the other sets of orbitals, we use several measures of locality. An obvious choice is the evaluation of the functionals of Eqs. (1)–(3). In Table I, we report the values of the three localization functionals computed for each of the localized orbital sets and for the canonical orbitals of the extended linear 20-unit glycine peptide $(Gly)_{20}$. Although the deviation from the optimal value is in all cases larger for the Cholesky MOs than for the other localized orbitals, it is still considerably smaller than for the canonical ones. Thus, the Cholesky MOs may be employed as initial orbitals for any of the standard localizations, usually reducing the number of iterations

Downloaded 29 Jan 2010 to 147.156.182.23. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

TABLE II. Classification of pair domains generated by different sets of occupied orbitals from the RHF wave function for linear $(Gly)_{10}$.

		Number of pairs			
Basis set	Pair classification	Canonical	PM	Boys	Cholesky
STO-3G	Strong	8846	697	737	1137
	Weak	1522	2826	2791	2881
	Distant	328	1298	1297	1267
	Very distant	1394	7269	7265	6805
	Max. domain size[a]	57	3	7	7
	Max. pair domain size[a]	71	6	13	14
cc-pVDZ	Strong	8669	655	621	1263
	Weak	1649	2839	2893	2784
	Distant	320	1300	1294	1237
	Very distant	1452	7296	7292	6806
	Max. domain size[a]	49	4	3	9
	Max. pair domain size[a]	68	8	6	18

[a]Measured as the number of atoms in the largest domain.

needed for convergence. The requirement of a set of localized MOs as initial orbitals is an issue of particular relevance for the maximization algorithm suggested by Subotnik *et al.*,[16] both for a faster convergence and for achieving the subquadratic (and potentially linear) scaling of the ER method. A confirmation that the Cholesky orbitals are comparable in terms of locality with the standard localized orbitals can be obtained by inspecting the sparsity of the MO coefficient matrices; see the supplementary material, Ref. 38

An important ingredient in modern local correlation methods[3–13] is the use of orbital domains to restrict the number of excitations from a given localized occupied orbital into the virtual space spanned by (approximately) atom-centered PAOs. An orbital domain is defined by the PAOs arising from AOs centered on the atoms needed to span the localized MO with a certain accuracy according to the completeness criterium of Boughton and Pulay.[4] In short, atoms are added to the domain $[i]$ of the occupied orbital ϕ_i in the order of decreasing Mulliken charge until the function

$$f(\mathbf{a}) = \min \int (\phi_i - \phi_i')^2 d\tau \tag{16}$$

is smaller than a prescribed threshold, typically 0.02 which is used in all calculations presented in this paper. In Eq. (16), $\phi_i' = \Sigma_{\mu\in[i]} \chi_\mu a_{\mu i}$. Once the orbital domains have been defined, single excitations are obtained by exciting electrons from a local occupied orbital ϕ_i into the PAOs of Eq. (11) centered on the atoms of the corresponding domain $[i]$. For double excitations, excitations are made from both ϕ_i and ϕ_j into their pair domain, which is defined as the union of the individual domains $[i]$ and $[j]$. Moreover, based on the short range nature of the pair correlation energy, an exact correlation treatment is carried out only for pair domains composed of orbital domains spatially close (strong pairs). Different levels of approximation are then used for pair domains whose components are more and more distant from each other (weak pairs, distant pairs), and the most distant pairs are completely neglected (very distant pairs). It is therefore of interest to compare the Cholesky MOs with the standard localized orbitals in terms of the classification (strong, weak, distant, very distant) of the pair domains.

In Table II the pair domains are classified according to the criteria used for the linear scaling local coupled cluster singles and doubles (LCCSD) method by Schütz and Werner.[12] Let R be the smallest distance in bohr between the atoms of domain $[i]$ and the atoms of domain $[j]$. The pair domain $[ij] \equiv [i] \cup [j]$ is then classified as a strong pair if $R=0$, i.e., if the domains share at least one atom. If $0<R\leqslant 10$, $[ij]$ is a weak pair. If $10<R\leqslant 15$, $[ij]$ is a distant pair. Finally, if $R>15$, $[ij]$ is a very distant pair. The canonical orbitals are highly delocalized, some extending over 80% of the molecule, and thus give rise to a very large number of strong pairs. In contrast, PM and Boys orbitals produce similar results, with a very limited number of strong pairs. The Cholesky orbitals yield almost twice as many strong pairs as the PM and Boys orbitals, still significantly smaller than the canonical orbitals. With the STO-3G basis set, the largest orbital domain for both the Cholesky and the Boys orbitals comprises seven atoms, less than 10% of the spatial extent of the molecule. However, while the Boys orbitals produce only one orbital domain of this size, there are nine such domains with the Cholesky orbitals. With the cc-pVDZ basis set, the largest domain comprises nine atoms in the case of Cholesky MOs, three times more than in the Boys case. A more detailed analysis of the size of the domains is reported in the histograms of Figs. 1 and 2. From the histograms we conclude that there are significantly fewer Cholesky domains comprising two atoms than observed for the Boys and PM orbitals. In other words, the Cholesky procedure does not yield the common chemical bond between two atoms. Instead, most of the Cholesky MOs extend over three to four atoms with the STO-3G basis set and over two to five atoms with the cc-pVDZ basis set.

Obviously, local correlation methods based on orbital domains and pair domains will scale linearly only if the number of strong, weak, and distant pairs scales linearly with

Downloaded 29 Jan 2010 to 147.156.182.23. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

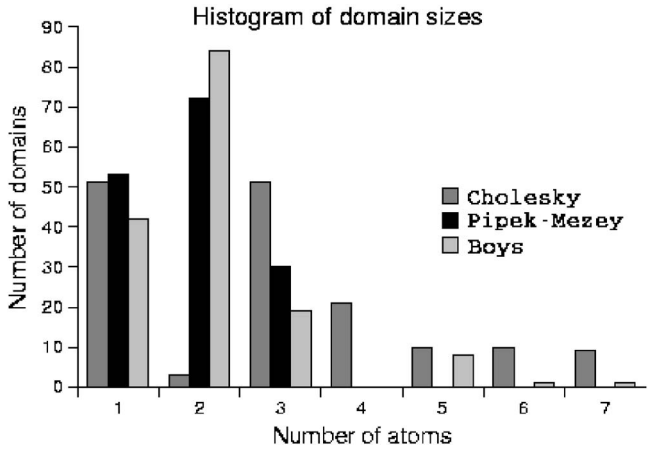

FIG. 1. Linear $(Gly)_{10}$ RHF/STO-3G wave function. Size of the orbital domains of each set of localized orbitals.

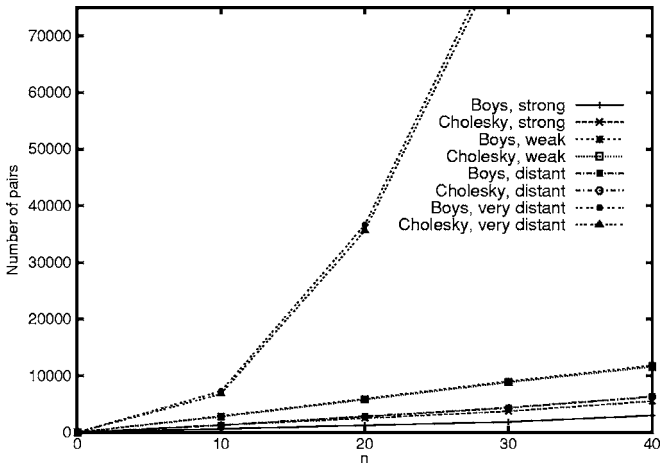

FIG. 3. Linear glycine chains $(Gly)_n$ RHF/cc-pVDZ wave function. Classification of pair domains.

the size of the system. Figures 3 and 4 display the number of pairs as a function of glycine and alkane chain lengths, respectively, for the Boys and Cholesky localizations. The results obtained with PM orbitals are similar to those obtained with Boys orbitals and are therefore not shown in the figures. Evidently, the number of Cholesky strong, weak, and distant pairs does scale linearly. For the glycines, the number of Cholesky strong pairs is approximately twice the number of Boys strong pairs, independent of peptide chain length. For the alkanes, there are about three times as many Cholesky strong pairs as Boys strong pairs. Therefore, combining local correlation methods and Cholesky MOs will lead to linear scaling, although the prefactor will be larger than that for Boys (or PM) orbitals. However, since the Cholesky orbitals are generated directly from the density matrix, diagonalization-free Hartree-Fock techniques can be combined with local correlation methods to give a complete linear scaling formulation. Table III shows timings of different localization procedures for the linear glycines and alkanes. The PM localization is considerably more expensive than the Boys scheme when treating molecules with a large number of atoms and, due to its noniterative character, the Cholesky procedure is an order of magnitude faster than Boys localization. The Boys and Cholesky procedures show the same scaling with system size, but we stress that the latter may straightforwardly be formulated in a linear scaling fashion.

IV. CONCLUDING REMARKS

Using that the electronic density matrix in AO basis is positive semidefinite, we investigate in this paper the orbitals resulting from its Cholesky decomposition. The resulting Cholesky MOs inherit locality from the sparsity of the density matrix and are therefore suitable for large molecules only. The main advantages of this procedure compared to the conventional localization functional optimizations are the following:

(1) Cholesky decomposition is fast (and may be made linear scaling) and numerically stable,
(2) it is noniterative, and
(3) initial orbitals are not needed.

While the second point implies that the convergence problems occasionally encountered with the conventional localizations are completely avoided, the last point allows for computation of localized orbitals without knowledge of the canonical ones, thus making Cholesky localization useful for

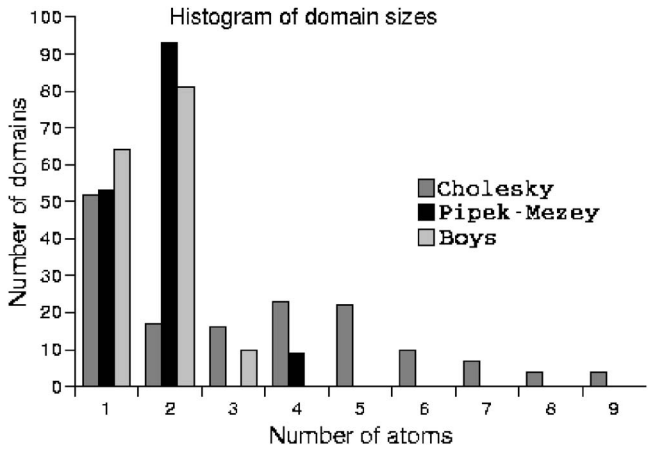

FIG. 2. Linear $(Gly)_{10}$ RHF/cc-pVDZ wave function. Size of the orbital domains of each set of localized orbitals.

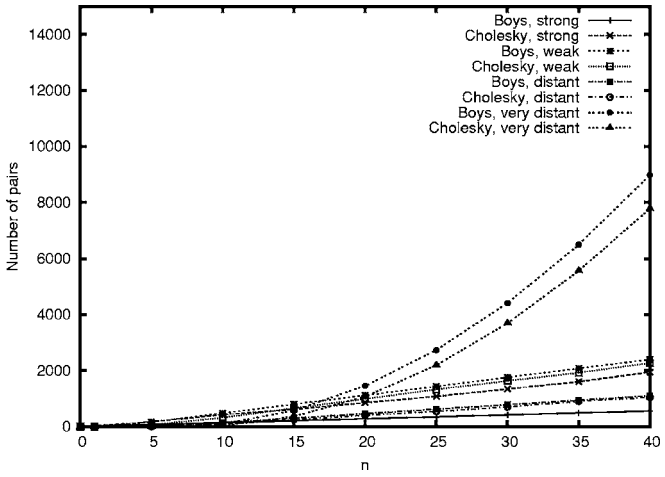

FIG. 4. Alkane chains C_nH_{2n+2} RHF/cc-pVDZ wave function. Classification of pair domains.

Downloaded 29 Jan 2010 to 147.156.182.23. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

TABLE III. Timings (in seconds, AMD Opteron 2.4 GHz processor) for the various localization methods based on the RHF/cc-pVDZ optimized wave function. For Cholesky localization the timing includes the calculation of the density matrix from canonical orbitals.

	Localization		
Molecule[a]	PM[b]	Boys[b]	Cholesky
$(Gly)_{10}$	66.6	1.9	0.1
$(Gly)_{20}$	1264.7	17.1	1.0
$(Gly)_{30}$	5433.3	81.2	3.5
$(Gly)_{40}$		295.9	7.7
C_5H_{12}	0.02	0.03	0.00
$C_{10}H_{22}$	0.19	0.05	0.01
$C_{15}H_{32}$	1.11	0.17	0.01
$C_{20}H_{42}$	5.07	0.28	0.03
$C_{25}H_{42}$	13.07	0.58	0.06
$C_{30}H_{52}$	37.24	0.90	0.10
$C_{35}H_{72}$	70.14	1.37	0.16
$C_{40}H_{82}$	184.40	2.38	0.24

[a]Linear geometry used for the glycines.
[b]Cholesky MOs used as initial orbitals.

correlation methods formulated on top of diagonalization-free Hartree-Fock theory where the density matrix is optimized directly. The main disadvantage is that the Cholesky MOs are less local than the localized orbitals obtained by the conventional Boys or PM procedures, as judged by a number of locality measures. However, we have demonstrated that linear scaling may still be achieved using Cholesky MOs in conjunction with local correlation methods, albeit with a larger prefactor. The major culprit seems to be the inability of the Cholesky localization to reproduce two-center MOs (i.e., the common chemical bond).

We have observed that using the Cholesky MOs as initial orbitals for the Boys, ER, and PM schemes usually reduces the number of iterations needed for convergence. Thus, combining Cholesky MOs with the DIIS-type convergence enhancement developed by Subotnik *et al.*[16] might become a powerful tool for conventional orbital localization in large molecules. In addition, the subquadratic (and potentially linear) scaling of the ER procedure[16] should be ensured by using Cholesky MOs as initial orbitals.

ACKNOWLEDGMENTS

This work has been supported by the Swedish Foundation for Strategic Research, by the EU network NANOQUANT, by Japan Science and Technology Agency (JST-CREST), and by the Spanish M.C.T. Project No. CTQ2004-07768-C02-01/BQU. Computational resources were provided by the Lunarc supercomputer center in Lund, Sweden. The authors thank Dr. Valera Veryazov for help with the graphical processing of the data.

APPENDIX: ORTHONORMALITY OF THE CHOLESKY MOs

Consider the scaled density matrix

$$\mathbf{P} = \tfrac{1}{2}\mathbf{D}, \tag{A1}$$

whose decomposition defines the Cholesky MOs as

$$X_{\mu i} = \left[P_{ii} - \sum_{k=1}^{i-1} (X_{ik})^2 \right]^{-1/2} \left[P_{\mu i} - \sum_{j=1}^{i-1} X_{\mu j} X_{ij} \right], \tag{A2}$$

where we assume that the rows and columns of **P** are ordered according to the decomposition pattern (pivoting). The matrix **P** is a projection operator in the sense that **PS** projects onto the occupied space and satisfies the idempotency condition

$$\mathbf{PSP} = \mathbf{P}, \tag{A3}$$

see, e.g., Ref. 14 for details.

To demonstrate that the Cholesky MOs constitute an orthonormal set we initially assume that the first $M-1$ MOs satisfy the condition,

$$\sum_{\mu\nu} X_{\mu i} S_{\mu\nu} X_{\nu j} = \delta_{ij}, \tag{A4}$$

and proceed by showing that the Mth MO is normalized and orthogonal to the previous ones. The proof is completed by verifying that the two first MOs are orthonormal.

It is easy to show that the Mth MO is normalized,

$$\sum_{\mu\nu} X_{\mu M} S_{\mu\nu} X_{\nu M} = \left[P_{MM} - \sum_{k=1}^{M-1} X_{Mk}^2 \right]^{-1} \left[\sum_{\mu\nu} P_{\mu M} S_{\mu\nu} P_{\nu M} - 2 \sum_{i=1}^{M-1} \left(\sum_{\mu\nu} P_{\mu M} S_{\mu\nu} X_{\nu i} \right) X_{Mi} + \sum_{i,j=1}^{M-1} X_{Mi} \left(\sum_{\mu\nu} X_{\mu i} S_{\mu\nu} X_{\nu j} \right) X_{Mj} \right] = 1, \tag{A5}$$

and orthogonal to the previous MOs ($k<M$),

$$\sum_{\mu\nu} X_{\mu k} S_{\mu\nu} X_{\nu M} = \left[P_{MM} - \sum_{j=1}^{M-1} X_{Mj}^2 \right]^{-1/2} \left[\sum_{\mu\nu} X_{\mu k} S_{\mu\nu} P_{\nu M} - \sum_{i=1}^{M-1} X_{Mi} \sum_{\mu\nu} X_{\mu k} S_{\mu\nu} X_{\nu i} \right] = 0. \tag{A6}$$

Here we have used the idempotency condition, the orthonormality of the first $M-1$ Cholesky MOs, and the identity

$$\sum_{\nu\gamma} P_{\mu\gamma} S_{\gamma\nu} X_{\nu i} = X_{\mu i}. \tag{A7}$$

Similarly, the first two MOs are normalized,

$$\sum_{\mu\nu} X_{\mu 1} S_{\mu\nu} X_{\nu 1} = P_{11}^{-1} \sum_{\mu\nu} P_{\mu 1} S_{\mu\nu} P_{\nu 1} = 1, \tag{A8}$$

Downloaded 29 Jan 2010 to 147.156.182.23. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp

$$\sum_{\mu\nu} X_{\mu 2} S_{\mu\nu} X_{\nu 2} = [P_{22} - X_{21}^2]^{-1} \Bigg[\sum_{\mu\nu} P_{\mu 2} S_{\mu\nu} P_{\nu 2} - 2 \Bigg(\sum_{\mu\nu} P_{\mu 2} S_{\mu\nu} X_{\nu 1} \Bigg) X_{21} + X_{21} \Bigg(\sum_{\mu\nu} X_{\mu 1} S_{\mu\nu} X_{\nu 1} \Bigg) X_{21} \Bigg] = 1, \tag{A9}$$

and orthogonal,

$$\sum_{\mu\nu} X_{\mu 1} S_{\mu\nu} X_{\nu 2} = [P_{22} - X_{21}^2]^{-1/2} \Bigg[\sum_{\mu\nu} X_{\mu 1} S_{\mu\nu} P_{\nu 2} - \Bigg(\sum_{\mu\nu} X_{\mu 1} S_{\mu\nu} X_{\nu 1} \Bigg) X_{21} \Bigg] = 0, \tag{A10}$$

which completes the proof.

We stress that this result is valid for projection operators in general (including, obviously, the case where $\mathbf{S}=\mathbf{1}$). Note, in particular, that the virtual Cholesky MOs obtained by decomposing $\mathbf{D}^v$, Eq. (10), are orthonormal.

[1] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).
[2] J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).
[3] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[4] J. W. Boughton and P. Pulay, J. Comput. Chem. **14**, 736 (1993).
[5] S. Saebo and P. Pulay, Annu. Rev. Phys. Chem. **44**, 213 (1993).
[6] E. Kapuy, F. Bogár, and C. Kozmutza, J. Mol. Struct.: THEOCHEM **297**, 365 (1993).
[7] E. Kapuy, F. Bogár, and E. Tfirst, Int. J. Quantum Chem. **52**, 127 (1994).
[8] C. Hampel and H. J. Werner, J. Chem. Phys. **104**, 6286 (1996).
[9] W. Förner, R. Knab, J. Cizek, and J. Ladik, J. Chem. Phys. **106**, 10248 (1997).
[10] J. Pipek and F. Bogár, *Topics in Current Chemistry* (Springer, New York, 1999).
[11] M. Schütz, G. Hetzer, and H. J. Werner, J. Chem. Phys. **111**, 5691 (1999).
[12] M. Schütz and H. J. Werner, J. Chem. Phys. **114**, 661 (2001).
[13] H. J. Werner, F. R. Manby, and P. J. Knowles, J. Chem. Phys. **118**, 8149 (2003).
[14] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2000).
[15] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).
[16] J. E. Subotnik, Y. Shao, W. Z. Liang, and M. Head-Gordon, J. Chem. Phys. **121**, 9220 (2004).
[17] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).
[18] V. Magnasco and A. Perico, J. Chem. Phys. **47**, 971 (1967).
[19] W. von Niessen, J. Chem. Phys. **56**, 4290 (1972).
[20] G. Berghold, C. J. Mundy, A. H. Romero, J. Hutter, and M. Parrinello, Phys. Rev. B **61**, 10040 (2000).
[21] I. Souza, N. Marzari, and D. Vanderbildt, Phys. Rev. B **65**, 35109 (2002).
[22] J. M. Millam and G. E. Scuseria, J. Chem. Phys. **106**, 5569 (1997).
[23] R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).
[24] M. Challacombe, J. Chem. Phys. **110**, 2332 (1999).
[25] T. Helgaker, H. Larsen, J. Olsen, and P. Jørgensen, Chem. Phys. Lett. **327**, 397 (2000).
[26] H. Larsen, J. Olsen, P. Jørgensen, and T. Helgaker, J. Chem. Phys. **115**, 9685 (2001).
[27] P. E. Maslen, C. Ochsenfeld, C. A. White, M. S. Lee, and M. Head-Gordon, J. Phys. Chem. A **102**, 2215 (1998).
[28] P. R. Surján, M. Kertész, A. Karpfen, and J. Koller, Phys. Rev. B **27**, 7583 (1983).
[29] P. Ordejón, D. A. Drabold, R. M. Martin, and M. P. Grumbach, Phys. Rev. B **51**, 1456 (1995).
[30] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).
[31] E. Schwegler, M. Challacombe, and M. Head-Gordon, J. Chem. Phys. **106**, 9708 (1997).
[32] R. Baer and M. Head-Gordon, Phys. Rev. Lett. **79**, 3962 (1997).
[33] C. Ochsenfeld, C. A. White, and M. Head-Gordon, J. Chem. Phys. **109**, 1663 (1998).
[34] W. Liang, H. Shao, C. Ochsenfeld, A. T. Bell, and M. Head-Gordon, Chem. Phys. Lett. **358**, 43 (2002).
[35] C. H. Yam, S. Yokojima, and G. Chen, J. Chem. Phys. **119**, 8794 (2003).
[36] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in FORTRAN*, 2nd ed. (Cambridge University Press, New York, 1992).
[37] B. C. Carlson and J. K. Keller, Phys. Rev. **105**, 102 (1957).
[38] See EPAPS Document No. E-JCPSA6-125-306640 for a graphical analysis of the sparsity of the density, virtual density, and MO coefficient matrices. This document can be reached via a direct link in the online article's HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).
[39] N. H. F. Beebe and J. Linderberg, Int. J. Quantum Chem. **7**, 683 (1977).
[40] I. Røeggen and E. Wisløff-Nielsen, Chem. Phys. Lett. **132**, 154 (1986).
[41] H. Koch, A. Sánchez de Merás, and T. B. Pedersen, J. Chem. Phys. **118**, 9481 (2003).
[42] T. B. Pedersen, A. M. J. Sánchez de Merás, and H. Koch, J. Chem. Phys. **120**, 8887 (2004).
[43] MOLCAS6 University of Lund, Sweden, 2005 (see http://www.teokem.lu.se/molcas/).
[44] G. Karlstrom, R. Lindh, P. Malmqvist *et al.*, Comput. Mater. Sci. **28**, 222 (2003).

Downloaded 29 Jan 2010 to 147.156.182.23. Redistribution subject to AIP license or copyright; see http://jcp.aip.org/jcp/copyright.jsp