
Group Functions Approach Based on the Combination of Strictly Local Geminals and Molecular Orbitals

A. M. TOKMACHEV,¹ A. L. TCHOUGRÉEFF²

¹*Institute of Inorganic Chemistry, RWTH Aachen, Prof.-Pirlet-Str. 1, 52056 Aachen, Germany*

²*Karpov Institute of Physical Chemistry, Vorontsovo pole 10, 105064 Moscow, Russia*

Received 20 June 2004; accepted 20 June 2004

Published online 11 October 2005 in Wiley InterScience (www.interscience.wiley.com).

DOI 10.1002/qua.20795

ABSTRACT: The group functions technique is a natural way to introduce local description into quantum chemistry. It can also be a basis for construction of numerically effective computational schemes having almost linear growth of computational costs with that of the size of the system. Previously, we constructed a family of computationally efficient semiempirical methods based on the variationally determined strictly local geminals (SLGs). It was implemented with four popular parameterization schemes (MINDO/3, MNDO, AM1, and PM3). Because of construction details, its applicability was restricted only to compounds with well-defined two-electron two-center chemical bonds and lone pairs. We generalize the previous treatment to make the electronic structure calculations possible for a wider class of compounds without loss of computational efficiency. The proposed scheme (SLG/SCF) is based on the general group function approach combining different descriptions for different electron groups: essentially local two-electron ones are described by geminals, while those with other numbers of electrons are described in the one-electron approximation. We implement the RHF, UHF, and ROHF approaches for the groups with delocalized electrons. This approach is tested for a series of radicals and molecules with extended π -electron systems. It is shown that the SLG/SCF-based methods describe the experimental data not worse than the corresponding SCF procedures and provide a good starting point for calculations of polyatomic molecular systems. © 2005 Wiley Periodicals, Inc. *Int J Quantum Chem* 106: 571–587, 2006

Key words: geminals; local states; hybrid schemes; MNDO; π -systems

Correspondence to: A. M. Tokmachev; e-mail: andrei.tokmachev@ac.rwth-aachen.de

Contract grant sponsor: RFBR.

Contract numbers: 02-03-32087; 04-03-32146; 04-03-32206.

Introduction

The progress in modern quantum chemistry is essentially achieved due to the possibility of making reliable predictions on the electronic structure and properties of large molecular systems. The standard ab initio and density functional theory (DFT) techniques require rather extensive computational resources that also grow rapidly, while the size of the system to be studied increases (up to N^7 , where N is the number of one-electron functions in the basis set). It restricts their applicability to systems of moderate size. Even in the case of semiempirical methods, the growth of computational costs is as large as N^3 . When the system at hand is truly large, but a detailed study of its potential energy surface (PES) is necessary, specially designed computational schemes should be used, e.g., for biologically relevant systems.

Two different general strategies have been proposed in the literature for solving the problem of rapid growth of computational costs for polyatomic molecules. The first one is based on development of quantum chemical methods with almost linear dependence of computational costs on the size of the system (so-called $O(N)$ -methods). The acceptable scalability properties of quantum chemical methods can be achieved in different ways. They are thoroughly reviewed in [1, 2]. It should only be stated that most $O(N)$ methods use the local representation of the wave function. The localization of electronic degrees of freedom either is based on the exponential decay of the one-electron density matrix elements in the real space [3] or exploits the principle of nearsightedness for static observables formulated by Kohn [4]. The direct use of localized one-electron states for achieving linear scalability of required computational resources [5, 6] seems promising. Regrettably, only a small number of schemes of that type can compete in quality with standard methods of ab initio quantum chemistry [7].

The second way to treat large molecular systems is to combine descriptions with different computational costs and, therefore, different capabilities. (These so-called hybrid quantum mechanical/molecular mechanical [QM/MM] methods were introduced in practice by Warshel and Levitt [8]; the credit for the corresponding idea, however, should be paid to Lennard-Jones [9].) QM/MM methods tentatively use the well-known observation that the chemical transformations are local and touch only

relatively small part of the molecule, usually only several chemical bonds (reaction center), while the rest remains relatively inert. The reaction center is usually described by a more or less precise quantum chemical method while the environment is described by "force fields" (molecular mechanics [MM]). A considerable number of schemes are used within this hybrid methodology [10, 11], which currently becomes a practical tool. The essential feature of hybrid schemes is the necessity to correctly describe the frontier and the interaction between the quantum and classical subsystems. The problem of intersubsystem junction is not completely solved which is discussed in Ref. [12]. In fact, most QM/MM methods construct the junction in an ad hoc manner, neglecting some physical interactions and compensating their absence by adding artificial ones [12].

In a series of papers [13–15] we proposed a general strategy for the construction of junction in hybrid QM/MM schemes. It is based on distribution of the basis set of local one-electron functions among subsystems and sequential construction of the effective electronic Hamiltonian for the quantum subsystem. The whole construction is essentially based on the underlying wave function for the inert subsystem, which has to reproduce the structure of the MM approximation, for example, to operate with local fragments of its electronic structure, such as chemical bonds and lone pairs. We have chosen the trial wave function in the form of the antisymmetrized product of strictly local geminals (SLG) as one underlying the MM description. The possibility to derive an MM-like description of molecular PES from the SLG trial wave function was proved in Refs. [16, 17]. It allowed us to obtain explicit formulae for renormalization of the Hamiltonian parameters for the quantum subsystem in the presence of the classical one described by the MM [12, 14] and the PES for complex systems [15]. In contrast, the explicit expressions for the forces and torques acting on the atoms in the MM subsystem and arising due to variations of the one- and two-electron densities of the QM boundary bond are also derived [12].

In view of practical value of the SLG wave function, we developed semiempirical quantum chemical schemes based on it [18, 19]. These schemes use the MINDO/3 [20], MNDO [21], AM1 [22], and PM3 [23] semiempirical parameterizations for the molecular Hamiltonian. The essential characteristics of the SLG approximation important for motivation of the approach developed in the present

work are the following. First, the SLG approach [18, 19] uses variationally determined local orbitals. Most of the geminal-based approaches proposed in the literature use somehow predefined one-electron states, either empirically constructed or extracted from preceding Hartree–Fock calculations by that or another localization procedure [24, 25]. The problem of optimizing the one-electron states for the geminal-based wave functions was rarely addressed in the literature [26, 27]. At the same time, the correct form of local one-electron states is important for the overall quality of the wave function. Also, these states can be readily used for predicting properties of molecules and for molecular design [28, 29].

Another characteristic of the SLG approach is its computational efficiency. Dividing the molecule into parts and calculating the electronic structure for each part separately is a direct way to obtain linear dependence of computational costs on the size of the system. It was shown [18] that the SLG-MINDO/3 method linearly scales with the system size. In the case of the SLG-MINDO, SLG-AM1, and SLG-PM3 methods [19] the size-quadratic contribution to the computational costs has significant weight due to numerous calculations of two-center molecular integrals in the basis of hybrid orbitals throughout the optimization of one-electron states. In the case of the NDDO parameterizations, the computational costs can be significantly reduced by using multipole–multipole interaction schemes for the Coulomb interactions between nonbonded atoms [30]. The comparison of the SLG and the self-consistent field (SCF) procedures with analogous parameterizations has unequivocally demonstrated that the quality of numerical results on the heats of formation, molecular geometries, and vertical ionization potentials obtained using two essentially different trial wave functions is similar with some advantage of the SLG-based methods. Also, the SLG trial wave function, unlike the SCF one, possesses correct asymptotic behavior under cleavage of chemical bonds, and its essentially local character allows interpreting chemical data in a transparent fashion.

All these features—linear scaling, which incidentally leads to correct asymptotic behavior, variational determination of one-electron states, and a direct way to classical force fields and QM/MM interface, accurate description of experimental data—make the SLG-based semiempirical quantum chemistry an important alternative to the well-known and widely accepted SCF-based semiem-

pirical quantum chemical procedures. The serious drawback of the SLG methods is their nonuniversality. The SLG trial wave function assumes the molecule to have well-defined connectivity matrix, i.e., its structure is uniquely depicted by combination of two-electron chemical bonds and lone pairs neither with significant electron delocalization between different geminals nor with resonance of thus defined different structures. It is clear that many important classes of molecular systems do not fit to this picture, for example, the systems with odd number of electrons (radicals), delocalized π -systems, coordination compounds, etc.

Within the framework of valence bond (VB) theories, the problem of description of resonance between different local wave functions is typically solved by configuration interaction for several VB (and, possibly, SCF) structures (see, e.g., Refs. [31, 32]). The SLG approach can be extended to treat more general situation, when the electronic groups with essential delocalization are allowed. It can be made in the framework of the general group function theory [33] treating different electron groups by different procedures. Some of these groups can be described by geminals and some of them are described by single Slater determinants build upon molecular orbitals. This possibility brings us to the realm of the hybrid methodology, but of its particular kind, when different parts are treated using different methods of quantum chemistry [34, 35]. A similar form of the wave function is used in the general setting of the GVB-PP method [36], but an important point here is how to determine the carrier spaces for molecular orbitals and geminals. Even in the case of semiempirical implementation rapid GVB schemes with general one-electron states are based on very cumbersome working equations [37]. A general method based on the strongly orthogonal groups is implemented in the VB-2000 program suite [38]. At the same time, the particular *ab initio* implementation allows calculations of small molecules only. The limitations in the applicability of the geminal-type wave function have led Rassolov [27] to construct a so-called singlet-type strongly orthogonal geminals (SSG) wave function. The *ab initio* SSG method [27] is based on the antisymmetrized product of geminals and orbitals both delocalized over the whole molecule. The SSG method is positioned as one applicable to all molecular systems. In practice, however, it was tested only on diatomics [27]. Our extension of the geminal approach is essentially based on the strict locality of one-electron states and the ultimate need to

conserve the above-mentioned attractive properties of the SLG-based methods. In the next section, we consider the general structure of a new hybrid method and working formulae for the total energy and for the effective energy operators for the groups. We then test the proposed procedure, using an MNDO-based parameterization of the electronic Hamiltonian with special attention paid to the π -electron systems. The analysis of the calculated electronic structure and geometry of the test systems gives some insight into the relative importance of electron delocalization and correlation in typical organic molecules and allows us to draw conclusions about the general applicability of the scheme proposed, which are given in the final section. We also discuss the relation between numerical experiments on π -conjugated systems performed within the SLG/SCF approach and the traditional π -electron theory.

Method

Construction of the method is based on the electron group functions formalism. The entire molecular system is divided into groups. These groups represent chemical bonds, lone pairs, local unpaired valences, or delocalized π -electron systems, or they can be constructed in another way based on chemical intuition. The definition of any group includes the number of electrons in it and the set of one-electron basis functions spanning its carrier space. The settings for the group also include its spin state and a method (approximate form for the trial wave function) to be used for electronic structure calculation. The choice of the carrier space is the question of primary importance. The intuitive concepts of chemical bonding are mainly based on the hybrid orbitals (HOs), which are constructed from the atomic orbitals (AOs) located on one atom. The concept of hybridization proved very fruitful [28]. However, it imposes strong restrictions on the molecular wave function. The usual representation of chemical bonds and lone pairs by localized molecular orbitals (LMOs) relies on the transformation of canonical MOs on the basis of that or another localization criterion [39–41]. The orbitals thus obtained are not strictly local and have nonzero amplitudes on many atoms.

Although the use of the HO basis sets is an important constraint on the flexibility of wave function, it allows construction of linearly scaling procedures and is useful for separation of electronic

variables in a complex system [12, 42]. The construction of HOs reduces to orthogonal rotation of the AOs centered on some atom A :

$$t_{p\sigma}^+ = \sum_{i \in A} h_{pi}^A a_{i\sigma}^+ \quad (1)$$

where $|t_{p\sigma}^+\rangle = t_{p\sigma}^+|0\rangle$ is the HO, and $SO(4)$ matrix h^A determines the transformation from the AO to the HO basis set on the atom A bearing the valence sp -basis. The orthogonality of matrices h^A ensures the orthogonality of the carrier spaces assigned to the groups and, therefore, the strong orthogonality of electronic wave functions for the latter. The transformations h^A (and, therefore, HOs) are determined on the ground of variational principle for the energy. The transformation is meaningful only in the case of nonhydrogen atom with orbitals belonging to at least two different electron groups (frontier atom). The optimal transformation matrix can be determined by minimizing the total energy with respect to six angular variables per heavy (nonhydrogen) atom providing parameterization of the $SO(4)$ manifold by a product of six two-dimensional Jacobi rotations.

Each HO should be uniquely assigned to some electron group. If the group is to be described by an SCF-based method the molecular spin-orbitals are constructed as linear combinations of the hybrid spin-orbitals:

$$b_{i\sigma}^+ = \sum_{p \in \{\text{SCF}\}} c_{ip} t_{p\sigma}^+ \quad (2)$$

In the case of a geminal representing a two-center two-electron bond, the HOs at the “right” and “left” ends of the bond can be denoted as $|r_m\rangle$ and $|l_m\rangle$.

Each geminal is then a superposition of three singlet two-electron configurations:

$$g_m^+ = u_m r_{m\alpha}^+ r_{m\beta}^+ + v_m l_{m\alpha}^+ l_{m\beta}^+ + w_m (r_{m\alpha}^+ l_{m\beta}^+ + l_{m\alpha}^+ r_{m\beta}^+), \quad (3)$$

which are two ionic configurations (both electrons are on the same end of the chemical bond) and the covalent (Heitler–London type) one, respectively. The normalization condition for the geminal amplitudes u_m , v_m , and w_m reads:

$$\langle 0 | g_m g_m^+ | 0 \rangle = u_m^2 + v_m^2 + 2w_m^2 = 1. \quad (4)$$

In the case of a lone electron pair, only one configuration survives with the amplitude equal to unity:

$$g_m^+ = r_{m\alpha}^+ r_{m\beta}^+ \quad (5)$$

The wave function for the entire molecule is represented by the antisymmetrized product of geminals and molecular orbitals:

$$|\Psi\rangle = \left(\prod_{i\sigma} b_{i\sigma}^+ \right) \left(\prod_m g_m^+ \right) |0\rangle, \quad (6)$$

which we abbreviate as SLG/SCF. This wave function is written as if only one SCF group were present. At the same time, the presence of several SCF-treated groups leads only to grouping of summations over HOs in the definition Eq. (2) with respect to assignment of the HOs to the corresponding SCF-treated group. Therefore, the wave function Eq. (6) is quite general. It is convenient to introduce the density matrices for separate groups (one-electron density matrix is block-diagonal and nonvanishing intergroup elements of the two-electron density matrix are always products of elements of one-electron density matrix). In the case of the SCF groups, the spin-dependent density matrices are

$$P_{pq}^\sigma = \sum_{i\sigma} c_{i\sigma p} c_{i\sigma q} \quad (7)$$

In the case of (singlet) geminals, one- and two-electron density matrices are spin independent:

$$\begin{aligned} P_m^{tt'} &= \langle 0 | g_m t_{m\sigma}^+ t'_{m\sigma} g_m^+ | 0 \rangle, \quad P_m^{rr} = u_m^2 + w_m^2, \\ P_m^{ll} &= v_m^2 + w_m^2, \quad P_m^{rl} = P_m^{lr} = (u_m + v_m) w_m, \\ \Gamma_m^{tt'} &= \langle 0 | g_m t_{m\beta}^+ t'_{m\alpha}^+ t_{m\alpha} t_{m\beta} g_m^+ | 0 \rangle, \\ \Gamma_m^{rr} &= u_m^2, \quad \Gamma_m^{ll} = v_m^2, \quad \Gamma_m^{rl} = \Gamma_m^{lr} = w_m^2. \end{aligned} \quad (8)$$

We consider the electronic Hamiltonian in the NDDO approximation and transform it to the basis of HOs. This transformation generally preserves the overall structure of the Hamiltonian. It remains the sum of one- and two-center contributions:

$$H = \sum_A H_A + \sum_{A<B} H_{AB}, \quad (9)$$

where the one-center contributions have the form

$$\begin{aligned} H_A &= \sum_{tt' \in A} \left(U_{tt'}^A + \sum_{B \neq A} V_{tt',B}^A \right) \sum_{\sigma} t_{\sigma}^+ t'_{\sigma} \\ &+ \frac{1}{2} \sum_{tt't'' \in A} (tt'|t''t''')^A \sum_{\sigma\tau} t_{\sigma}^+ t_{\tau}'' t_{\tau}''' t_{\sigma}, \end{aligned} \quad (10)$$

while the two-center contributions have the form

$$\begin{aligned} H_{AB} &= - \sum_{\substack{t \in A \\ t' \in B}} \beta_{tt'}^{AB} \sum_{\sigma} (t_{\sigma}^+ t'_{\sigma} + h.c.) \\ &+ \sum_{\substack{tt' \in A \\ t''t''' \in B}} (tt'|t''t''')^{AB} \sum_{\sigma\tau} t_{\sigma}^+ t_{\tau}'' t_{\tau}''' t_{\sigma}. \end{aligned} \quad (11)$$

The expressions for molecular integrals in the basis of HOs through the transformation matrices h^A and molecular integrals in the basis of AOs can be found in Refs. [18, 19].

The total energy of the entire molecule can be found as is usually done, by averaging the electronic Hamiltonian over the electronic wave function and adding the core-core repulsion:

$$\begin{aligned} E^{\text{total}} &= E^{\text{el}} + E^{\text{core-core}}, \\ E^{\text{core-core}} &= \sum_{A<B} Z^A Z^B [(ss|ss)^{AB} + D^{AB}], \\ E^{\text{el}} &= \sum_A (E_A^{(1)} + E_A^{(2)}) + \sum_{A<B} (E_{AB}^{(1)} + E_{AB}^{(2)}), \end{aligned} \quad (12)$$

where D^{AB} is specific contribution characteristic for semiempirical methods employed. The electronic energy in the present approximation is a sum of contributions of four types: one- and two-center and one- and two-electron ones. The first type of contributions to the energy arises from attraction of electrons to the cores. For an arbitrary atom A it can be written as

$$\begin{aligned} E_A^{(1)} &= \sum_{pq \in \{A, \text{SCF}\}} \left(U_{pq}^A + \sum_{B \neq A} V_{pq,B}^A \right) \sum_{\sigma} P_{pq}^{\sigma} \\ &+ 2 \sum_{tm \in \{A, \text{SLG}\}} \left(U_{tm}^A + \sum_{B \neq A} V_{tm,B}^A \right) P_m^{tt}. \end{aligned} \quad (13)$$

The second type of contribution to the energy corresponds to one-center electron-electron repulsion and can be written as

$$\begin{aligned}
 E_A^{(2)} = & \sum_{pp'qq' \in \{A, \text{SCF}\}} \left\{ (pp'|qq')^A P_{pp'}^\alpha P_{qq'}^\beta + \frac{1}{2} [(pp'|qq')^A \right. \\
 & - (pq'|qp')^A] \sum_{\sigma} P_{pp'}^\sigma P_{qq'}^\sigma \left. \right\} + \sum_{\substack{pq \in \{A, \text{SCF}\} \\ t_m \in \{A, \text{SLG}\}}} [2(pq|t_m t_m)^A \\
 & - (pt_m|t_m q)^A] P_m^{tt} \sum_{\sigma} P_{pq}^\sigma + \sum_{t_m t_n \in \{A, \text{SLG}\}} [4(t_m t_m|t_n' t_n')^A \\
 & - 2(t_m t_n|t_n' t_m')^A] P_m^{tt} P_n^{t't'} + \sum_{t_m \in \{A, \text{SLG}\}} (t_m t_m|t_m t_m)^A \Gamma_m^{tt}. \quad (14)
 \end{aligned}$$

The resonance (one-electron transfer) contributions to the energy are nonzero only for intragroup electron transfers:

$$E_{AB}^{(1)} = -2 \sum_{\substack{p \in \{A, \text{SCF}\} \\ q \in \{B, \text{SCF}\}}} \beta_{pq}^{AB} \sum_{\sigma} P_{pq}^\sigma - 4 \sum_{\substack{t_m \in \{A, \text{SLG}\} \\ \tilde{t}_m \in \{B, \text{SLG}\}}} \beta_{t_m \tilde{t}_m}^{AB} P_m^{\tilde{t}t}. \quad (15)$$

The last type of contribution to the energy arises from the two-center electron–electron repulsion. It takes the following form for any pair of atoms A and B :

$$\begin{aligned}
 E_{AB}^{(2)} = & \sum_{\substack{pp' \in \{A, \text{SCF}\} \\ qq' \in \{B, \text{SCF}\}}} (pp'|qq')^{AB} \left\{ \left(\sum_{\sigma} P_{pp'}^\sigma \right) \left(\sum_{\tau} P_{qq'}^\tau \right) \right. \\
 & - \sum_{\sigma} P_{pq'}^\sigma P_{qp'}^\sigma \left. \right\} + 2 \sum_{\substack{t_m \in \{A, \text{SLG}\} \\ pq \in \{B, \text{SCF}\}}} (t_m t_m|pq)^{AB} P_m^{tt} \sum_{\sigma} P_{pq}^\sigma \\
 & + 2 \sum_{\substack{pq \in \{A, \text{SCF}\} \\ t_m \in \{B, \text{SLG}\}}} (pq|t_m t_m)^{AB} P_m^{tt} \sum_{\sigma} P_{pq}^\sigma \\
 & - 2 \sum_{\substack{t_m \in \{A, \text{SLG}\} \\ \tilde{t}_m \in \{B, \text{SLG}\}}} \sum_{\substack{p \in \{A, \text{SCF}\} \\ q \in \{B, \text{SCF}\}}} (t_m p|q \tilde{t}_m)^{AB} P_m^{\tilde{t}t} \sum_{\sigma} P_{pq}^\sigma \\
 & + 2 \sum_{\substack{t_m \in \{A, \text{SLG}\} \\ t_n' \in \{B, \text{SLG}\}}} (t_m t_m|t_n' t_n')^{AB} [2(1 - \delta_{mn}) P_m^{tt} P_n^{t't'} + \delta_{mn} \Gamma_m^{tt'}] \\
 & - 4 \sum_{\substack{t_m \in \{A, \text{SLG}\} \\ \tilde{t}_m \in \{B, \text{SLG}\}}} \sum_{\substack{t_n' \in \{A, \text{SLG}\} \\ \tilde{t}_n' \in \{B, \text{SLG}\}}} (t_m t_n'| \tilde{t}_n' \tilde{t}_m)^{AB} P_m^{\tilde{t}t} P_n^{t't'}. \quad (16)
 \end{aligned}$$

The energy of the molecule is minimized according to variational principle with respect to the electronic structure parameters. We implemented an iteration scheme that alternates optimizations of HOs and determinations of density matrices for electron groups. The optimal HOs are obtained by gradient minimization of the energy with respect to sextuples of parameters determining transformation matrices h^A . The geminal amplitudes u_m , v_m , and w_m are obtained by diagonalizing 3×3 effective

Hamiltonians for each geminal representing a chemical bond. For determination of molecular orbitals and one-electron density matrices in the SCF-treated groups, we implemented three procedures that are well known in the literature [33]: restricted Hartree–Fock (RHF), restricted open-shell Hartree–Fock (ROHF), and unrestricted Hartree–Fock (UHF). The construction of the effective Fock operators for these methods takes into account the presence of other electron groups. They modify one-electron matrix elements of the effective operator. The modified one-electron operator has the following matrix elements \tilde{h}_{pq} for the orbitals $p \in A$ and $q \in B$:

$$\begin{aligned}
 \tilde{h}_{pq} = & \delta_{AB} \left\{ U_{pq}^A + \sum_{C \neq A} V_{pq,C}^A + \sum_{t_m \in \{A, \text{SLG}\}} [2(pq|t_m t_m)^A \right. \\
 & - (pt_m|t_m q)^A] P_m^{tt} + 2 \sum_{t_m \in \{C \neq A, \text{SLG}\}} (pq|t_m t_m)^{AC} P_m^{tt} \left. \right\} \\
 & - (1 - \delta_{AB}) \left\{ \beta_{pq}^{AB} + \sum_{\substack{t_m \in \{A, \text{SLG}\} \\ \tilde{t}_m \in \{B, \text{SLG}\}}} (t_m p|q \tilde{t}_m)^{AB} P_m^{\tilde{t}t} \right\}. \quad (17)
 \end{aligned}$$

The averaging of two-electron operators acting within any of the SCF-treated groups and producing the Fock operator is not affected by the presence of other groups. It is clearly seen that the number of elementary steps in the optimization procedure is proportional to the size of the system if all the groups are chosen to be small. Therefore, the important property of linear scalability of computational resources provided by the SLG-based methods is preserved in this more general approach. Based on the above expressions for the energy and corresponding analytic gradients, we also implemented the procedure for determination of the optimal molecular geometries.

Results and Discussion

The SLG/SCF scheme is implemented for four semiempirical Hamiltonians: MINDO/3, MNDO, AM1, and PM3. In this section, we discuss only one of them, MNDO, since the qualitative conclusions are similar for all the mentioned parameterizations. Our main purpose is to study the effect of trial wave function upon calculated molecular properties. Physically, the difference between the SCF and SLG(/SCF) wave functions can be understood as a

combination of two factors: in the case of the SLG/SCF scheme the delocalization between electron groups is totally suppressed, while in the SCF scheme it is taken into account; on the other hand the SLG/SCF scheme takes into account the intrabond electron correlation, while the SCF scheme lacks it.

A type of scheme in which the electronic structure of molecule is represented as a Slater determinant, but each molecular orbital is a linear combination of only two HOs is known in the literature, corresponding to the wave function of strictly localized molecular orbitals (SLMO) [43]. This wave function is obviously less precise than that of the SCF method due to restrictions on the form of MOs, especially when the structure of localized molecular orbitals is not optimized. At the same time, it can be used as a starting point for constructing perturbational treatments, analogous to that implemented in the PCILO scheme [44]. Within the framework of the SLG/SCF approach, one of the variants of the SLMO scheme can be obtained by setting the SCF method as a means of calculation of all electron groups. The important difference of this SLMO scheme from those proposed earlier is the means of determining one-electron states forming the localized MOs, i.e., their constrained variational determination. It allows defining the delocalization energy as the difference between energies obtained by the SCF and SLMO schemes.

The correlation energy can also be readily calculated for any chemical bond: it is simply the difference between energies obtained for the geminal in the SLG and SCF approximations, with all other electronic structure parameters fixed. Thus, we can consider interplay between electron delocalization and intrabond correlation. The corresponding data are given in Table I, which presents delocalization energies for some simple test molecules and correlation energies for all chemical bonds in them. The energies of electron delocalization naturally increase with the size of the molecular system. Their additivity is only approximate, since the interaction between chemical bonds separated by one or two chemical bonds is also significant. For example, the difference between delocalization energies for ethane and methane is 18.72 kcal/mol, while the difference between delocalization energies for propane and ethane is 19.30 kcal/mol. The delocalization energy strongly depends on the number of lone electron pairs. In the case of systems with lone electron pairs, the delocalization energy is smaller

TABLE I
Energies of correlation for some chemical bonds, and energies of delocalization for some molecules as obtained for MNDO parameterization of the Hamiltonian (kcal/mol).

Molecule	Bond	E_{corr}	E_{del}	$E_{\text{del}}(PT)$
H ₂	H—H	-0.97	0.0	0.0
CH ₄	C—H	-0.92	-17.06	-22.61
C ₂ H ₆	C—C	-0.79	-35.78	-44.91
	C—H	-0.94		
C ₃ H ₈	C—C	-0.83	-55.08	-68.50
	C ¹ —H	-0.94		
	C ² —H	-0.97		
C ₂ H ₄	σ -C—C	-0.27	-34.64	-38.66
	π -C—C	-8.84		
	C—H	-0.84		
C ₂ H ₂	σ -C—C	-0.12	-25.19	-30.50
	π -C—C	-5.48		
	C—H	-0.68		
NH ₃	N—H	-1.12	-3.13	-4.54
CH ₃ NH ₂	C—N	-1.26	-32.57	-34.29
	C—H	-0.97		
	N—H	-1.12		
H ₂ O	O—H	-1.18	-1.36	-2.07
CH ₃ OH	C—O	-1.29	-34.11	-32.49
	C—H	-0.99		
	O—H	-1.20		
HF	F—H	-1.89	0.00	0.00
CH ₃ F	C—F	-1.65	-31.23	-27.58
	C—H	-0.98		

mainly because the electron transfers to these pairs are impossible.

The delocalization energies calculated as differences between the SCF and SLMO energies can be compared with the perturbation estimates of the latter. The simple one can be obtained in the second order with respect to the one-electron transfers between the geminals. Each geminal can be expressed through bonding and antibonding orbitals $|b_m\rangle$ and $|a_m\rangle$, which are orthogonal linear combinations of HOs $|r_m\rangle$ and $|l_m\rangle$:

$$g_m^+ = U_m b_{m\alpha}^+ b_{m\beta}^+ + V_m a_{m\alpha}^+ a_{m\beta}^+. \quad (18)$$

Equation (18) represents the natural expansion of the geminal. The relevant excited states are those with one electron on one of the geminals and three electrons on some other geminal, occupying bonding and/or antibonding spin-orbitals, and with all other geminals fixed. The matrix elements necessary for the perturbation estimate are calculated

using the MNDO Hamiltonian and the wave function of the SLG form. For simplicity, we neglected complicated (but incidentally relatively small) contributions to interbond one-electron transfer coming from the interatomic Coulomb interactions. The calculated delocalization energies are given in Table I. Data analysis shows that the delocalization energies obtained by different methods are quite close and have the same qualitative behavior. We can conclude that the definition of the electron delocalization energy as a difference between the SCF and SLMO energies is adequate.

The intrabond correlation energies given in Table I are calculated for equilibrium interatomic distances. In the case of large interbond distances, the effect of correlations can be essentially larger, since the SCF-based description of chemical bond leads to an incorrect dissociation limit. Nevertheless, for equilibrium geometries, the estimates for the correlation energies of σ -bonds obtained with semiempirical parameters are all close to 1 kcal/mol. Somewhat larger values are obtained for chemical bonds involving heteroatoms. The special case is provided by multiple bonds where the correlation energy for σ -bonds is very small, while that for π -bonds amounts to several kcal/mol. Another conclusion is that the correlation energies calculated for the same chemical bond in similar environments are very close: see, for example, primary C—H bonds in all alkanes, N—H bond in ammonia and methylamine, and O—H bond in water and methanol. At the same time, it can be seen that the correlation energy for the C—H bond in methyl group bonded with electronegative atom is somewhat larger than that in alkanes.

Another question that can be addressed in the present context is how the difference in the way of density matrix determination in the basis of HOs affects the composition of the latter. Each HO is characterized by its direction and the weight of the s -orbital. The latter is in a one-to-one correspondence with the quantity x in the standard sp^x -notation for the HOs. Table II represents the x values for the HOs as calculated in the framework of the SLMO and SLG schemes. The analysis shows that the difference between two ways of determination of optimal HOs is not very large. In fact, the qualitative pictures are the same. Relatively large quantitative deviations are found only for the heteroatom hybrids directed to the neighbor carbon atoms.

Up to this point, we have discussed mainly the particular variant of the SLG/SCF scheme, i.e., that

TABLE II
Indices x in the sp^x representation of hybrid orbitals as evolved from the energy minimization by the SLG-MNDO and SLMO-MNDO methods.

Molecule	Hybrid	\times (SLMO)	\times (SLG)
C ₂ H ₆	C → C	3.05	3.07
	C → H	2.98	2.98
C ₃ H ₈	C ¹ → C ²	3.11	3.07
	C ² → C ³	3.00	3.02
	C ¹ → H	2.96	2.98
	C ² → H	3.00	2.98
C ₂ H ₄	C → C	1.70	1.73
	C → H	2.18	2.15
C ₂ H ₂	C → C	0.94	0.94
	C → H	1.06	1.06
NH ₃	N → H	7.72	7.77
	N → :	0.52	0.52
CH ₃ NH ₂	C → N	3.93	4.10
	N → C	6.44	8.50
	C → H	2.76	2.73
	N → H	6.39	7.43
H ₂ O	O → :	0.68	0.52
	O → H	8.53	8.56
CH ₃ OH	O → :	1.53	1.53
	C → O	4.65	4.97
	O → C	8.90	10.31
	C → H	2.65	2.60
	O → H	7.91	8.42
HF	O → :	1.54	1.48
	F → H	5.32	4.96
CH ₃ F	C → F	5.41	5.41
	F → C	4.95	5.22
	C → H	2.55	2.55

of the SLMOs, and have studied the delocalization and correlation contributions to the energy within this framework. The important question of parameterization of the SLG/SCF method was not considered. It is not a very simple and well-defined problem because the electronic structure calculated by this method depends not only on the composition and the spatial structure of the molecule, but also on the definition of electron groups. At the same time, it looks like a reasonable solution to define different parameters for electron groups with different form of the trial electronic wave function. To achieve the best agreement of the calculated characteristics with the experimental characteristics, it is necessary to reconsider the whole set of parameters of semiempirical method, since many of them in that or another extent implicitly account for the type of trial wave function (more precisely,

TABLE III
Experimental and calculated by the SCF-MNDO and SLG/SCF-MNDO methods heats of formation for free radicals (kcal/mol).

Molecule	Expt	MNDO SCF	MNDO SLG/SCF
CH ₃	34.8	25.8	36.9
C ₂ H ₅	25	12.8	26.8
(CH ₃) ₂ CH	16.8	1.5	18.9
(CH ₃) ₃ C	4.5	-7.2	14.6
NH ₂	40.1	37.1	30.9
OH	9.5	0.5	-2.5
HCO	10.4	-0.5	10.9

to the amount of electron correlation explicitly included in it). For the time being, our purpose is not to reach the best possible quantitative agreement with experiment, but to demonstrate a general applicability of the proposed scheme. Therefore, we restrict the re-parameterization by a relatively small parameters' subset, that for the resonance (one electron hopping), since it is the most sensitive one with respect to the form of the trial wave function. This strategy previously led to a successful parameterization of the SLG method [18, 19]. Bearing this in mind, we introduce different sets of resonance parameters for the SLG and SCF electron groups.

The outlined strategy of "point" re-parameterization immediately opens access to a whole class of systems that cannot be calculated with the use of the SLG-based methods. These are free σ -radicals with a single local orbital assigned to the SCF subsystem. Their calculation does not imply any re-parameterization, since no resonance interaction appears for the singly occupied HO spanning the whole carrier space for the corresponding group. Table III represents heats of formation for a small test set of radicals calculated by the SCF-MNDO and SLG/SCF-MNDO methods, compared with the experimental data. The data in Table III show that for this class of objects, the SLG/SCF-MNDO method turns out to be preferable to the SCF-MNDO method. Only in the case of the NH₂ and OH radicals are the heats of formation calculated by the SCF-MNDO method closer to the experimental ones than those calculated by the SLG/SCF method.

A crucial test for the SLG/SCF method is calculation of electronic structure and properties of molecules with delocalized π -systems. In this case, the number of parameters to be modified is also very

small. We consider π -electron systems of different extent formed by orbitals of the C and N atoms. As previously, we use the MNDO parameterization but also, for comparison, we give the results obtained using a simpler MINDO/3 scheme, as it uses an essentially different scheme for parameterization of the resonance integrals. The readjusted values of the resonance parameters for the SCF π -subsystem are the following: in the case of the MNDO parameterization, $\beta_p^\pi(\text{C}) = 9.92$ eV and $\beta_p^\pi(\text{N}) = 21.97$ eV, while in the case of the MINDO/3 parameterization, $\beta_{\text{CC}}^\pi = 0.452$, $\beta_{\text{NC}}^\pi = 0.371$, and $\beta_{\text{NN}}^\pi = 0.400$. These values should be compared with the original SCF ones: $\beta_p(\text{C}) = 7.934122$ eV and $\beta_p(\text{N}) = 20.495758$ eV for the SCF-MNDO parameterization and $\beta_{\text{CC}} = 0.419907$, $\beta_{\text{NC}} = 0.410886$, and $\beta_{\text{NN}} = 0.377342$ for the SCF-MINDO/3 one.

Table IV represents heats of formation for some test molecules containing π -electron systems calculated by the SCF and SLG/SCF methods, with the MNDO and MINDO/3 parameterizations as well as the experimental values. In all cases the maximal π -system was used in the SLG/SCF calculations (e.g., four p -orbitals form a single four-electron group in butadiene and not two two-electron ones). The data presented in Table IV show that in the case of the MNDO parameterization, the level of agreement with the experiment for the SCF and SLG/SCF techniques depends on the class of the molecules. The SCF scheme better suits for molecules with multiple bonds (including those with conjugated double bonds), while the SLG/SCF scheme works better in the case of aromatic compounds. Some disbalance in the quality of description of different π -systems certifies that the parameters for the σ - and π -subsystems arising from the MNDO parameterization are not optimal. Generally, considering the whole set of the test molecules, we can draw a conclusion about similar quality of the SCF-MNDO and SLG/SCF-MNDO methods in description of the heats of formation. It should be noted that adding the delocalization energy for the π -bonds (calculated as the difference between the SCF and SLMO energies of π -system where the two-center two-electron π -bonds are described by LMOs without intergroup resonance) to the heats of formation obtained by the SLG method for compounds with conjugated double bonds can improve the agreement with the experiment: for example, in the case of *trans*-1,3-butadiene the heat of formation becomes 28.9 kcal/mol, while in the case of fulvene it becomes 52.9 kcal/mol. In the case of the MINDO/3 parameterization, the transition from

TABLE IV

Experimental and calculated by the SCF-MNDO(MINDO/3) and SLG/SCF-MNDO(MINDO/3) methods heats of formation for molecules with π -systems (kcal/mol).

Molecule	Expt	MNDO SCF	MNDO SLG/SCF	MINDO/3 SCF	MINDO/3 SLG/SCF
C ₂ H ₄	12.5	15.3	17.0	19.2	22.2
C ₂ H ₂	54.3	57.3	45.6	57.8	49.8
<i>trans</i> -1,3-Butadiene	26.0	28.9	32.2	31.9	31.9
Fulvene	47.5	53.6	58.2	67.3	60.8
Cycloheptatriene	43.9	33.6	50.4	33.6	35.9
Benzene	19.8	21.2	19.3	28.4	16.0
Toluene	12.0	13.5	15.1	22.3	15.0
Naphthalene	36.1	38.1	37.7	55.4	28.8
Allyl	40.0	35.3	42.4	43.9	46.9
Pyridine	34.6	28.7	34.3	34.4	34.7
Pyridazine	66.5	44.2	54.4	23.0	55.7
Pyrimidine	47.0	35.5	50.3	35.9	51.9
Pyrazine	46.9	38.3	51.3	41.8	53.7
Aniline	20.8	22.0	29.4	18.2	14.2
HN ₃	70.3	73.0	71.1	-13.9	50.1
CH ₃ NC	35.6	60.0	42.4	21.5	48.3

the SCF to the SLG/SCF trial wave function significantly improves the description of experimental data on the heats of formation. It is especially well seen in the example of HN₃ and relative stability of three isomers: pyridazine, pyrimidine, and pyrazine. The data in Table IV show that the MNDO-based methods are more suitable than the MINDO/3-based ones for describing the heats of formation of molecules with extended π -electron systems.

Above we performed qualitative comparison of heats of formation calculated by different methods (SCF and SLG/SCF). At the same time quantitative analysis of errors of different methods can give more insight on their particular characteristics. To perform such an analysis, we considered a set of 67 test molecules. Their heats of formation were calculated by the SCF and SLG/SCF methods with the MNDO parameterization (see Table II of Ref. [19] and Tables III and IV of the present study). For each of the molecules and calculation methods, the error in the heat of formation was calculated as the difference between calculated and experimental values. If one takes this error as a random variable a subsequent statistical treatment makes sense. In this assumption, we consider empirical distribution functions for the errors for both calculation methods. The characteristics of these distributions can be readily found. In the case of the SCF-MNDO method, the average over set of errors is approximately -3.3 kcal/mol, while in the case of the

SLG/SCF-MNDO method, this value is only 0.01 kcal/mol. These estimates indicate that the SCF-MNDO method has a noticeable systematic error in the calculated heats of formation, while the SLG/SCF-MNDO method is free of this error; its errors lie more or less symmetrically around the zero value. This preliminary conclusion can be refined by statistical analysis. For each of the methods, the normal distribution for errors in the heats of formation holds with reasonable accuracy provided four molecules (of 67) with largest deviations from the experiment are excluded from the consideration. This conclusion can be drawn from Figures 1 and 2 representing the empirical distribution functions of errors in the coordinates linearizing the normal distribution for the SCF-MNDO and SLG/SCF-MNDO methods, respectively, as well as the linear fits for both distributions. The closeness of the sets of points to the corresponding linear fits certifies the general applicability of the assumption of the normal distribution of errors. The linear fits contain the information about the parameters of the normal distribution (mean value and dispersion) describing the errors. The abscissa for the crossing of the linear fit and the x axis gives the value of the a parameter (average), while the slope of the linear fit is σ^{-1} . The numerical values of a and σ obtained are the following: $a = -3.2$ kcal/mol, $\sigma = 7.8$ kcal/mol and $a = 0.3$ kcal/mol and $\sigma = 7.0$ kcal/mol for the SCF- and SLG/SCF-based methods, respectively.

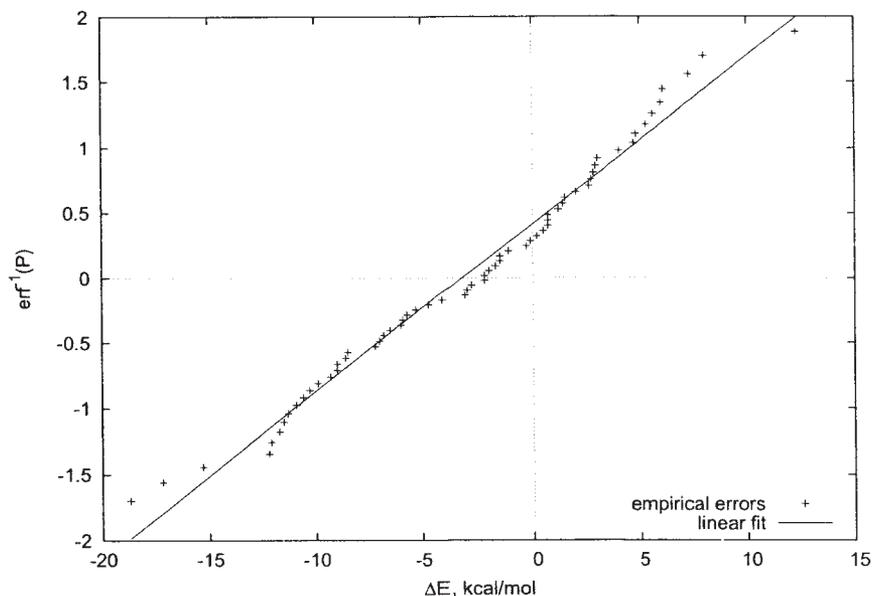


FIGURE 1. Empirical distribution of errors for heats of formation in the SCF-MNDO method.

The values of a confirm the presence of the systematic error in the SCF-MNDO method estimates of the heats of formation and its absence in the SLG/SCF-MNDO estimates.

For further evaluation of a new method, it is also necessary to compare the optimal molecular geometries as obtained by the SCF and SLG/SCF meth-

ods. In the previous studies on the SLG wave function [19], it was established that the latter allows us to improve description of the lengths of chemical bonds formed by heteroatoms and of the torsion angles as compared with the standard SCF-based semiempirical methods. Table V presents data on molecular geometries of several compounds as cal-

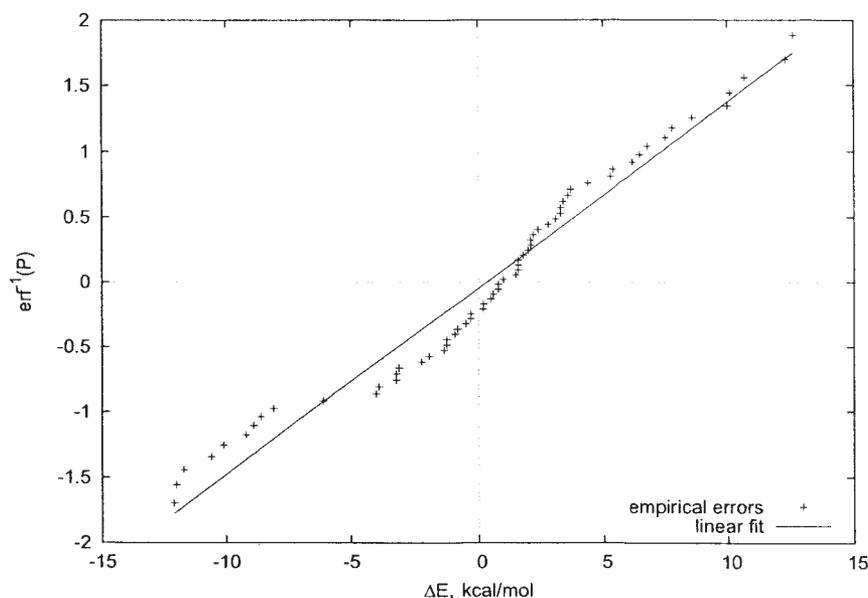


FIGURE 2. Empirical distribution of errors for heats of formation in the SLG/SCF-MNDO method.

TABLE V

Experimental and calculated by the SCF-MNDO(MINDO/3) and SLG/SCF-MNDO(MINDO/3) methods geometry parameters.

Molecule	Parameter	Expt	MNDO SCF	MNDO SLG/SCF	MINDO/3 SCF	MINDO/3 SLG/SCF
<i>trans</i> -1,3-Butadiene	C ¹ C ²	1.344	1.341	1.321	1.329	1.315
	C ² C ³	1.463	1.465	1.440	1.464	1.438
	C ¹ C ² C ³	123.3	125.7	125.0	124.7	125.3
Fulvene	C ¹ C ²	1.470	1.491	1.463	1.508	1.479
	C ² C ³	1.355	1.366	1.343	1.359	1.340
	C ³ C ⁴	1.476	1.477	1.449	1.481	1.454
	C ¹ C ⁶	1.349	1.345	1.323	1.337	1.324
Benzene	CC	1.397	1.407	1.382	1.406	1.387
	CH	1.084	1.090	1.080	1.106	1.088
Naphthalene	C ¹ C ²	1.364	1.382	1.361	1.380	1.364
	C ² C ³	1.415	1.429	1.405	1.431	1.402
	C ¹ C ⁹	1.421	1.439	1.412	1.449	1.423
	C ⁹ C ¹⁰	1.418	1.435	1.415	1.462	1.438
Pyridine	N ¹ C ²	1.338	1.353	1.355	1.335	1.355
	C ² C ³	1.394	1.411	1.379	1.406	1.375
	C ³ C ⁴	1.392	1.405	1.380	1.406	1.385
	C ² H	1.086	1.095	1.085	1.114	1.092
	C ³ H	1.082	1.089	1.081	1.105	1.089
	C ⁴ H	1.081	1.090	1.081	1.106	1.088
	C ⁶ N ¹ C ²	116.9	118.2	115.8	119.6	117.1
	N ¹ C ² C ³	123.8	122.8	123.6	122.5	122.9
	C ² C ³ C ⁴	118.5	118.8	119.5	118.2	119.4
	C ³ C ⁴ C ⁵	118.4	118.6	118.1	119.0	118.3
	HC ² C ³	120.2	121.3	120.6	121.1	120.4
	HC ³ C ²	120.1	120.3	120.1	121.1	120.4

culated by the SCF and SLG/SCF methods with the MNDO and MINDO/3 parameterizations and the corresponding experimental data. It shows that for the both parameterizations, the quality of optimal molecular geometries obtained with the different wave functions is similar. As in the case of the heats of formation, the SCF method better describes compounds with multiple bonds, while the SLG/SCF method is better suited to the aromatic compounds.

A convenient way to compare the electronic structures obtained using different methods is to compare the total charges on atoms. The corresponding data are given in Table VI. It is clearly seen that the charge distributions obtained by the SCF and SLG/SCF methods are quite different. In CH₃NC the difference is even qualitative (sign of the charge on the atom C¹). The quantitative differences can be also rather large. For example, the C—H bonds in the SLG/SCF method are significantly polarized, while in the case of the SCF approximation they are almost nonpolar.

In the present calculations of molecules with π -electron systems, the σ/π -separation is explicitly employed. It is interesting to analyze the set of parameters of the π -electron Hamiltonian compatible with our consideration. First, we consider ethylene molecule, its ground electronic state (¹A_{1g}), excited singlet state (¹B_{1u}), and excited triplet state (³B_{1u}). From the energies of the excitations ¹A_{1g} → ¹B_{1u} and ¹A_{1g} → ³B_{1u} we can estimate the difference between the one-center and two-center Coulomb repulsion $\gamma_{11} - \gamma_{12}$ and the resonance integral β , since

$$E(^1B_{1u}) - E(^3B_{1u}) = \gamma_{11} - \gamma_{12} \quad (19)$$

and

$$E(^1B_{1u}) - E(^1A_{1g}) = \frac{\gamma_{11} - \gamma_{12}}{2} + \sqrt{\left(\frac{\gamma_{11} - \gamma_{12}}{2}\right)^2 + 4\beta^2}. \quad (20)$$

TABLE VI

Atomic charges calculated by the SCF-MNDO(MINDO/3) and SLG/SCF-MNDO(MINDO/3) methods.

Molecule	Atom	MNDO SCF	MNDO SLG/SCF	MINDO/3 SCF	MINDO/3 SLG/SCF
C ₂ H ₄	C	-0.080	-0.229	-0.016	-0.148
	H	0.040	0.114	0.008	0.074
C ₂ H ₂	C	-0.155	-0.211	-0.111	-0.154
	H	0.155	0.211	0.111	0.154
Benzene	C	-0.059	-0.130	0.007	-0.056
	H	0.059	0.130	-0.007	0.056
Naphthalene	C ¹	-0.041	-0.120	-0.003	-0.081
	C ²	-0.058	-0.128	0.009	-0.048
	C ⁹	-0.038	-0.027	0.021	0.037
	H ¹	0.058	0.131	-0.010	0.056
Pyridine	H ²	0.060	0.131	-0.006	0.054
	N ¹	-0.230	-0.119	-0.165	-0.105
	C ²	0.054	-0.106	0.135	-0.008
	C ³	-0.124	-0.167	-0.065	-0.084
	C ⁴	-0.007	-0.088	0.070	-0.013
	H ²	0.084	0.164	-0.015	0.062
HN ₃	H ³	0.072	0.142	0.000	0.061
	H ⁴	0.067	0.141	-0.016	0.056
	H	0.124	0.076	0.095	0.024
	N ¹	-0.289	-0.254	-0.423	-0.409
CH ₃ NC	N ²	0.212	0.170	0.637	0.628
	N ³	-0.047	0.082	-0.309	-0.243
	H	0.001	0.103	-0.021	0.064
	C ¹	0.270	-0.002	0.083	-0.118
	N ²	-0.464	-0.432	0.097	0.066
	C ³	0.192	0.125	-0.118	-0.141

The energy of the singlet-triplet transition ${}^1A_{1g} \rightarrow {}^3B_{1u}$ is well established experimentally. It is equal to 4.36 eV [45]. The experimental value for the energy of the ${}^1A_{1g} \rightarrow {}^1B_{1u}$ transition is determined as the maximum of the band corresponding to $\pi \rightarrow \pi^*$ excitation. It is equal to 7.66 eV [46]. At the same time, it was established [47, 48] that this maximum does not correspond to the vertical transition, but to a somewhat twisted ethylene molecule. The energy of this transition accepted in the literature on the basis of detailed refinement is 8.0 eV. Taking these values for the energies of transitions as a reference, we estimate $\gamma_{11} - \gamma_{12} \approx 3.6$ eV and $\beta \approx 2.95$ eV. The values close to these were obtained in the framework of the effective valence shell Hamiltonian formalism based on the Löwdin orthogonalized orbitals [49]. We should note that the value of resonance integral coming from considerations of larger π -systems and typically accepted in theoretical considerations is significantly lower (~ 2.4 eV). Within the framework of the SLG/SCF scheme, the MNDO parameterization leads to the following values:

$\gamma_{11} - \gamma_{12} = 3.59$ eV, and $\beta = 2.16$ eV. At the same time, the "effective" resonance parameter [coefficient at the nondiagonal element of one-electron density matrix at energy expression Eq. (15)] within the SLG/SCF-MNDO scheme is not β , but $\beta_{\pi}^{\text{eff}} = \beta + \frac{1}{2}(\sigma\pi|\sigma\pi)^{\text{CC}}$ (we assume the bond order to be unity for the σ -bond), which amounts to 2.43 eV. This is precisely the value of this parameter, which has been known for decades. We see that the difference between both Coulomb integrals used in the MNDO approximation and the resonance integral (effective) are quite close to those obtained from the experiment.

Further comparison of our numerical results with previous theoretical concepts and experimental data involves analysis of the famous bond order vs. bond-length relation. The values of the equilibrium C—C distances obtained in the framework of the SLG/SCF-MNDO method are somewhat smaller than the experimental ones, i.e., 1.158 Å, 1.311 Å, and 1.382 Å as compared with experiment 1.203 Å, 1.339 Å, and 1.397 Å for C₂H₂, C₂H₄, and

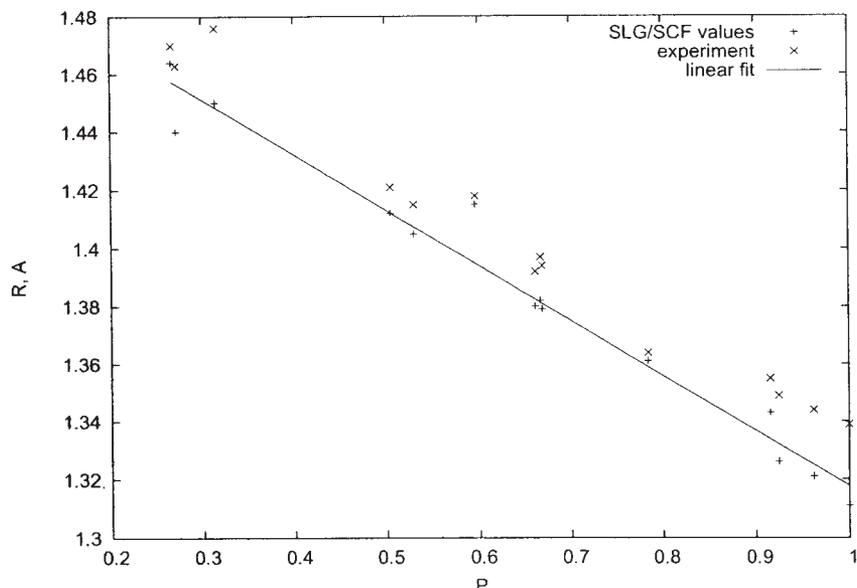


FIGURE 3. Bond order bond length (BOBL) analysis of the experimental and SLG/SCF MNDO calculated bond lengths vs. the calculated bond orders.

C_6H_6 , respectively. The error is smaller for smaller bond orders. It has been known for decades that the bond length in molecules with conjugate π -systems is related to the π -bond order by

$$R = R_0 - \omega P, \quad (21)$$

where the values $R_0 = 1.517$ and $\omega = 0.18$ (both in Å) have been determined in Ref. [50]. We considered a set of 14 π -bonds with the bond orders smaller than unity. It turned out that the linear relation $R = 1.507 - 0.19P$ satisfactorily describes the results of our calculations on bond lengths. The results are depicted in Figure 3. One can see that both R_0 and ω have the values which correspond to certain underestimation of the bond lengths in our calculations (all the + corresponding to experimental data are higher than the × corresponding to our numerical experiment): R_0 for the numerical experiment is smaller than that of Ref. [50], and ω is larger thus yielding even a stronger contraction of the bond due to π -electrons as compared with the hypothetical “ π -electron free” bond length R_0 .

In Ref. [51], the linear relationship was derived from the Hückel–Lennard–Jones theory [9], which assumes the harmonic behavior of the energy of the σ -core in the vicinity of its minimum R_0 with the elasticity constant k_σ and linear dependence of the resonance integral on the C–C distance elongation:

$\beta = \beta_0 + \beta' \delta R$ (hereinafter β_0 stands for the effective resonance parameter for the π -system introduced above and β' for its C–C distance derivative). In this case

$$\omega = -2\beta'/k_\sigma \quad (22)$$

We employed this simple possibility in the hope of obtaining independent estimates of the parameters R_0 and ω from the values of semiempirical parameters and the dependencies of the molecular integrals on the interatomic distances. To this purpose, we have taken a system of two neutral C atoms in the sp^2 -hybridization and equal densities on all hybrid orbitals as a model system. It turns out that the minimum of the σ -energy in this idealized system appears at $R_0 = 1.4196$ Å, and that the elasticity constant $k_\sigma = 68.0$ eV/Å². Taking into account our estimate for $\beta' = -3.95$ eV/Å, we get for the ω parameter the value of 0.116 Å. Clearly, these estimated values of R_0 and ω are far from those obtained from analysis of either natural or our numerical (SLG/SCF) experiments given above and the problem is to reconcile rather appealing results of the numerical fit for those parameters obtained from our model data with our independent estimates of the same parameters. The problem at the first glance is concentrated in the too large rigidity of the σ -core that comes from the SLG/SCF model.

For example, the corresponding value used in Refs. [51, 52], amounting to only $47.5 \text{ eV}/\text{\AA}^2$ (extracted from the analysis of the vibrational spectrum of benzene) would apparently lead to the required agreement at least in the value of ω . Nevertheless, the too small value of R_0 remains a problem with this approach. Another important characteristic is the dimensionless constant of electron-phonon coupling in the π -electron systems. It is defined as [51]:

$$\lambda = \frac{-2\beta'}{\beta_0} \omega = \frac{-2\beta' - 2\beta'}{\beta_0} \frac{-2\beta'}{k_\sigma}. \quad (23)$$

Its estimate obtained on the basis of parameters extracted from the spectrum of benzene [52] is ~ 0.65 . If the ω fitted to the numerical experiment is used in the above expression, this parameter is ~ 0.62 . If the independent estimate for ω is employed this value becomes too small (0.38) as compared with the previous data.

In view of the contradiction found between the results of the BOBL fits and the independent estimates of the fit parameters, we studied the BOBL relation in more details. First, we mention that in the case of the SLG/SCF NDDO approach (as in all other methods, including the Coulomb repulsion of electrons explicitly, as opposed to the repulsionless Hückel theory) the energy of the π -system cannot be represented by a linear function of the π -bond orders. In contrast, the approximation to the σ -energy provided by a system of two neutral C atoms in the sp^2 -hybridization misses the specific core-core repulsion involving neighbor atoms of the bond. At the equilibrium distance R_0 of the model diatomic system, the molecule-specific core-core repulsion acts as an additional repulsive force F , so that the harmonic approximation remains valid, but with a shifted equilibrium position. With this notion, the energy of a π -bond in the SLG/SCF approach can be represented as

$$\underbrace{-F(R - R_0) + \frac{1}{2} k_\sigma (R - R_0)^2}_{E_\sigma} - \underbrace{\left(2\beta P + \frac{1}{2} \gamma_{12} P^2 \right)}_{E_\pi}. \quad (24)$$

Optimizing this with respect to R , and applying linear approximation for the dependence of the π -energy on the interatomic distance, one obtains for the bond length

$$R = R_0 + \frac{F}{k_\sigma} - \frac{2\beta' + \frac{1}{2} \gamma'_{12} P}{k_\sigma} P, \quad (25)$$

which allows us to define effective P -dependent quantities:

$$\omega(P) = \frac{2\beta' + \frac{1}{2} \gamma'_{12} P}{k_\sigma}$$

$$\lambda(P) = \frac{2\beta' + \frac{1}{2} \gamma'_{12} P}{\beta_0 + \frac{1}{4} \gamma_{12}^0 P} \frac{2\beta' + \frac{1}{2} \gamma'_{12} P}{k_\sigma}.$$

Inserting the quantities left until the moment undefined, $\gamma_{12}^0 = 7.49 \text{ eV}$ and $\gamma'_{12} = -2.62 \text{ eV}/\text{\AA}$, we obtain that for $P = 1$ the quantity $\omega(P)$ reaches the value of 0.135 \AA and $\lambda(P)$ that of 0.59 (the values for the vanishing bond order remain as they were previously). The first of the two is significantly closer to the values extracted from the fit. If the numerical data are fit against the model Eq. (25), the result is

$$R = 1.4796 - 0.1980P + 0.0101P^2 + 0.2863F,$$

provided the core-core forces F are calculated with use of the four closest neighbors of the bond under consideration.

Further improvement of the model for the SLG/SCF-based numerical procedure can be achieved if the C-C distance dependence of the π -energy is taken into account up to second order. In this case, the second derivatives β'' and γ''_{12} of the resonance integral and the two-center Coulomb integral come into play. Their numerical values as estimated from the parameters of the SLG/SCF-MNDO method are, respectively, 5.44 and $-1.523 \text{ eV}/\text{\AA}^2$. Taking this into account, and performing as previously the bond energy minimization with respect to R , we modify Eqs. (25) and (3) by replacing k_σ by expression: $k_\sigma - 2\beta''P - \frac{1}{2} \gamma''_{12} P^2$. Our calculations show that in this modification $\omega(P)$ falls within the range of 0.116 – 0.166 for P 0 – 1 . This is already a reasonable independent estimate for the fit value of ω . This also leads to improvement of estimated interatomic distances (see Table VII). The effective softening of the harmonic σ -potential due to second order contributions of the π -bond leads to $\lambda(P) \sim 0.35$ for a wide range of bond orders (see Table VII).

TABLE VII

Comparison of the bond lengths calculated from the model SLG/SCF parameters according to BOBL formula.

Molecule	P	ω	λ	R_1^a	R_2^b	R_3^c	Exp.
	0	0.1161	0.3774	—	—	—	—
Butadiene ^d	0.2548	0.1264	0.3544	1.3874	1.4326	1.4417	1.463
Graphite	0.525	0.1389	0.3433	1.3467	1.3911	1.4004	1.42
Polyene	$2/\pi$	0.1446	0.3417	1.3276	1.3729	1.3826	—
Benzene	$2/3$	0.1462	0.3415	1.3222	1.3675	1.3772	1.397
Butadiene ^e	0.9670	0.1638	0.3450	1.2612	1.2949	1.3025	1.344
Ethylene	1	0.1659	0.3459	1.2537	1.2807	1.2868	1.339

$$^a R_1 = R_0 - \omega P.$$

$$^b R_2 = R_1 - F/k_\sigma.$$

$$^c R_3 = R_1 - F/(k_\sigma + \dots).$$

^d Middle bond.

^e Terminal bond.

π -bonded atoms. Despite the strong disagreement with the value of [51] (0.65), it turns out that this disagreement favors the antiferromagnetic state of infinite polyene even more over one with alternating bond orders. The latter is possible if the electron-phonon coupling is larger than the value of dimensionless parameter of antiferromagnetic coupling [53] ($\gamma_{11}/2\beta \approx 1.13$), which is definitely not going to happen within the SLG/SCF-MNDO parameterization/calculation scheme.

Conclusions

We considered the extension of the SLG approach developed previously. The approach obtained (SLG/SCF) conserves all the pleasant properties of the SLG counterpart but allows calculation of molecular systems with essentially delocalized electrons. The SCF and SLG wave functions are particular cases of the SLG/SCF wave function. We tested this wave function with the MNDO parameterization. First, we estimated separately electron correlation and delocalization contributions to the total energy. Then the heats of formation and molecular geometries were compared with those obtained in the SCF calculations and experimentally for a set of molecules with π -electron systems. The results show that even very limited re-parameterization of molecular Hamiltonian allows us to obtain the results of quality comparable with that of the SCF-based approach.

ACKNOWLEDGMENTS

This work is dedicated to Professor I. A. Misurkin on the occasion of his 70th birthday. The work was completed during the stay of A. M. T. at the RWTH, Aachen, in the frame of the Alexander von Humboldt Postdoctoral Fellowship, which is gratefully acknowledged as is the kind hospitality of Professor R. Dronskowski.

References

- Goedecker, S. *Rev Mod Phys* 1999, 71, 1085.
- Wu, S. Y.; Jayanthi, C. S. *Phys Rep* 2002, 358, 1.
- Goedecker, S. *Phys Rev B* 1998, 58, 3501.
- Kohn, W. *Phys Rev Lett* 1996, 76, 3168.
- Stewart, J. J. P. *Int J Quantum Chem* 1996, 58, 133.
- Rubio, J.; Povill, A.; Malrieu, J. P.; Reinhardt, P. *J Chem Phys* 1997, 107, 10044.
- White, S. R.; Martin, R. L. *J Chem Phys* 1999, 110, 4127.
- Warshel, A.; Levitt, M. *J Mol Biol* 1976, 103, 227.
- Lennard-Jones, J. E. *Proc R Soc A* 1937, 158, 280.
- Gao, J. In *Reviews in Computational Chemistry*; VCH: New York, 1995; p 119.
- Sherwood, P. In *Modern Methods and Algorithms of Quantum Chemistry*; Proceedings; 2nd Ed.; Grotendorst, J., Ed.; NIC Series; John von Neumann Institute for Computing: Jülich, 2000; p 257.
- Tchougréeff, A. L.; Tokmachev, A. M. In *Advanced Topics in Theoretical Chemical Physics*; Maruani, J.; Lefebvre, R.; Brändas, E., Eds.; Kluwer: Dordrecht, the Netherlands, 2003; p 27.
- Tchougréeff, A. L. *Phys Chem Chem Phys* 1999, 1, 1051.
- Tokmachev, A. M.; Tchougréeff, A. L.; Misurkin, I. A. *J Mol Struct (Theochem)* 2000, 506, 17.

15. Tokmachev, A. M.; Tchougréeff, A. L. *Int J Quantum Chem* 2001, 84, 39.
16. Tchougréeff, A. L.; Tokmachev, A. M. *Int J Quantum Chem* 2004, 96, 175.
17. Tchougréeff, A. L. *J Mol Struct (Theochem)* 2003, 630, 243.
18. Tokmachev, A. M.; Tchougréeff, A. L. *J Comput Chem* 2001, 22, 752.
19. Tokmachev, A. M.; Tchougréeff, A. L. *J Phys Chem A* 2003, 107, 358.
20. Bingham, R. C.; Dewar, M. J. S.; Lo, D. H. *J Am Chem Soc* 1975, 97, 1302.
21. Dewar, M. J. S.; Thiel, W. *J Am Chem Soc* 1977, 99, 4899.
22. Dewar, M. J. S.; Zebisch, E. G.; Healy, E.; Stewart, J. J. P. *J Am Chem Soc* 1985, 107, 3902.
23. Stewart, J. J. P. *J Comp Chem* 1989, 10, 209.
24. Surján, P. R. In *Topics in Current Chemistry*; Vol 203; Springer-Verlag: Berlin, 1999; p 63.
25. Surján, P. R.; Lázár, A. *J Mol Struct (Theochem)* 1998, 455, 175.
26. Polezzo, S.; Fantucci, P. *Mol Phys* 1978, 36, 1835.
27. Rassolov, V. A. *J Chem Phys* 2002, 117, 5978.
28. Randić, M.; Maksić, Z. B. *Chem Rev* 1972, 72, 43.
29. Hay, B. P.; Hancock, R. D. *Coord Chem Rev* 2001, 212, 61.
30. Tokmachev, A. M.; Tchougréeff, A. L. *J Phys Chem A* 2005, 109, 7613.
31. Hollauer, E.; Nascimento, M. A. C. *J Chem Phys* 1993, 99, 1207.
32. Mo, Y.; Gao, J. *J Comp Chem* 2000, 21, 1458.
33. McWeeny, R. *Methods of Molecular Quantum Mechanics*; 2nd Ed.; Academic Press: London, 1992.
34. Soudackov, A. V.; Tchougréeff, A. L.; Misurkin, I. A. *Int J Quantum Chem* 1996, 58, 161.
35. Humbel, S.; Sieber, S.; Morokuma, K. *J Chem Phys* 1996, 105, 1959.
36. Bobrowicz, F. W.; Goddard, W. A., III. In *Modern Theoretical Chemistry: Methods of Electronic Structure Theory*; Vol 3; Schaefer, H. F., III; Plenum: New York, 1977; p 79.
37. Cullen, J. M. *Int J Quantum Chem* 1995, 56, 97.
38. Li, J. B.; McWeeny, R. *Int J Quantum Chem* 2002, 89, 208.
39. Foster, J. M.; Boys, S. F. *Rev Mod Phys* 1960, 32, 300.
40. Edmiston, C.; Ruedenberg, K. *Rev Mod Phys* 1963, 35, 457.
41. Perkins, P. G.; Stewart, J. J. P. *J Chem Soc Faraday Trans* 1982, 78, 285.
42. Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J Phys Chem A* 1998, 102, 4714.
43. Smits, G. F.; Altona, C. *Theor Chim Acta* 1985, 67, 461.
44. Diner, S.; Malrieu, J. P.; Claverie, P. *Theor Chim Acta* 1969, 13, 1.
45. van Veen, E. H. *Chem Phys Lett* 1976, 41, 540.
46. Johnson, K. E.; Johnston, D. B.; Lipsky, S. *J Chem Phys* 1979, 70, 3844.
47. Petrongolo, C.; Buenker, R. J.; Peyerimhoff, S. D. *J Chem Phys* 1982, 76, 3655.
48. Lindh, R.; Roos, B. O. *Int J Quantum Chem* 1989, 35, 813.
49. Martin, C. H.; Freed, K. F. *J Chem Phys* 1994, 101, 4011.
50. Coulson, C.; Golebievski, A. *Proc Phys Soc* 1961, 78, 1310.
51. Misurkin, I. A.; Ovchinnikov, A. A. *J Struct Chem* 1965, 5, 888 [in Russian].
52. Misurkin, I. A.; Ovchinnikov, A. A. *Opt Spectr* 1964, 16, 228 [in Russian].
53. Misurkin, I. A.; Ovchinnikov, A. A. *Usp Khim* 1977, 46, 1833.