Multireference perturbation CI methods for solvated systems described within the polarizable continuum model

ARTICLE *in* THE JOURNAL OF CHEMICAL PHYSICS · OCTOBER 1999

Impact Factor: 2.95 · DOI: 10.1063/1.480048

3 AUTHORS, INCLUDING:

Benedetta Mennucci
Università di Pisa
249 PUBLICATIONS **24,272** CITATIONS
SEE PROFILE

Chiara Cappelli
Scuola Normale Superiore di Pisa
117 PUBLICATIONS **1,866** CITATIONS
SEE PROFILE

Multireference perturbation configuration interaction methods for solvated systems described within the polarizable continuum model

Benedetta Mennucci[a)]
Dipartimento di Chimica e Chimica Industriale, Università di Pisa, via Risorgimento 35, 56126 Pisa, Italy

Alessandro Toniolo
Dipartimento di Chimica Fisica ed Elettrochimica, Università di Milano, via Golgi 19, 20133 Milano, Italy

Chiara Cappelli
Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

(Received 3 June 1999; accepted 3 August 1999)

We present a method to include solvent effects described within the polarizable continuum model into the CIPSI multireference perturbation algorithm. In the methodology we have formulated and implemented, solvent interactions are explicitly included in the configuration interaction scheme and in the following perturbative corrections, through proper operators. The nonlinear character induced by such operators leads to an iterative procedure in which solute and solvent can mutually equilibrate. Applications to the electronic excitation spectrum of formamide are considered. In these cases, effects due to an incomplete electrostatic response of the solvent (nonequilibrium model) as well as repulsion interactions between solute and solvent have been included in the quantum mechanical description. © *1999 American Institute of Physics.* [S0021-9606(99)30540-7]

I. INTRODUCTION

The evaluation of correlation energy still represents a challenge for computational quantum chemistry. Different approaches, variational and perturbative, are applied routinely when a single determinant or configuration is a reasonable description for the state of the system. The ground state of several molecules satisfies this condition; in this case methods such as second-order (MP2) and fourth order (MP4) Moller–Plesset perturbation or single-double configuration interaction (SDCI) have proved to be good approximations. The same methods fail when applied to the study of states where several determinants are necessary for a first approximate description as in bond breaking, transition metal compounds and, specifically, excited states. These systems are properly treated by multireference perturbation CI methods, such as the configuration interaction by perturbation with multiconfiguration zeroth-order wave function selected by iterative process (CIPSI) method first developed in 1973 by Huron, Malrieu, and Rancurel[1] and further extended by the Toulouse group[2,3] and by the Pisa-Ferrara group.[4,5]

The CIPSI method has been shown to be a very accurate tool to study many different chemical and photochemical systems *in vacuo*; until now, however, no extensions have been presented to more complex phenomena involving solvent effects. The extension of methods accounting for the external environment to more and more accurate quantum mechanical descriptions is a very important field in the present theoretical chemistry research. Many efforts have been done in recent years in particular as concerns effective Hamiltonian (EH) methods exploiting a continuum description of the solvent.[6,7] A significant contribution in this direction has been given by the large, and still active, development achieved by the polarizable continuum model (PCM) originally formulated in Pisa.[8] In recent years this solvation method has been deeply modified[9–11] so as to assume a structure easily generalizable to different quantum mechanical levels of theory; examples for standard CI methods, multiconfiguration self-consistent-field (MCSCF), density functional (DFT), and MP2 techniques have been already presented.[12] Here we shall report and discuss the strategy we have formulated to link the CIPSI algorithm with PCM (in its last version known as integral equation formalism, IEF[11]) so as to have a computational tool in which both correlation and solvent effects are properly merged into a variation-perturbation method.

To make this article self-contained and to document the working equations of the new implementation, in Secs. II and III we summarize the basic aspects of both CIPSI and PCM. The new strategy formulated to merge them is presented in Sec. IV, while computational details and numerical applications to the study of the excitation spectrum of formamide both *in vacuo* and in solution are reported in Sec. V.

II. EVALUATION OF CORRELATION ENERGY: THE CIPSI ALGORITHM

In this section we recall the formulation of the CIPSI algorithm and the recent extensions and improvements due to the groups in Pisa and in Ferrara.[13–15] Full CI (FCI) calculations involve too large numbers of configurations even for molecular systems and basis sets of modest size. The necessity of truncating the configurational basis has led to

a)Author to whom correspondence should be addressed; electronic mail: bene@dcci.unipi.it

devise different methods partitioning the CI space in classes, each corresponding to a different level of approximation in the treatment. The CIPSI method is based on a dynamic partition of the FCI space into two subspaces called $\mathcal{S}$ and $\mathcal{Q}$ (two class calculation). The $\mathcal{S}$ space contains the most important determinants describing the electronic states under investigation. The Hamiltonian projected onto the $\mathcal{S}$ space is diagonalized and the eigenstates are taken as the zero-order description of the true wave functions:

$$\hat{P}\hat{H}\hat{P}|\Psi_k^{(0)}\rangle = E_k^{(MR)}|\Psi_k^{(0)}\rangle, \tag{1}$$

$$\hat{P} = \sum_{N\in\mathcal{S}} |N\rangle\langle N|; |\Psi_k^{(0)}\rangle = \sum_{N\in\mathcal{S}} C_N^k |N\rangle. \tag{2}$$

The contributions of the determinants $|I\rangle \in \mathcal{Q}$ are evaluated by a second-order Rayleigh–Schrödinger treatment: thus, only those $|I\rangle$ that are single or double excitations from some $|N\rangle \in \mathcal{S}$ contribute. Given a partition of the CI Hamiltonian:

$$\hat{H} = \hat{H}_0 + \hat{V}, \tag{3}$$

the second-order energy correction to the kth electronic state is:

$$E_k^{(2)} = -\sum_{I\in\mathcal{P}} \frac{\langle I|\hat{V}|\Psi_k^{(0)}\rangle^2}{E_I - E_k^{(0)}} \tag{4}$$

or

$$E_k^{(2)} = -\sum_{M,N\in\mathcal{S}} C_M^k C_N^k \sum_{I\in\mathcal{P}} \frac{\langle M|\hat{V}|I\rangle\langle I|\hat{V}|N\rangle}{E_I - E_k^{(0)}}. \tag{5}$$

Here, $\mathcal{P}$ is the subspace of $\mathcal{Q}$ collecting all single and double excitations from $\mathcal{S}$. Equations (4) and (5) are totally equivalent at this level of the treatment; they correspond to the traditional and the diagrammatic implementation of the algorithm, respectively.

Two different partitions of the Hamiltonian $\hat{H}$ are implemented in our program; both satisfy the condition that the basis $\{|\Psi_k^{(0)}\rangle \in \mathcal{S}, |I\rangle \in \mathcal{P}\}$ forms a set of eigenstates of $\hat{H}_0$:

$$\hat{H}_0|\Psi_k^{(0)}\rangle = E_k^{(0)}|\Psi_k^{(0)}\rangle$$

$$\hat{H}_0|I\rangle = E_I|I\rangle.$$

In the Møller–Plesset baricentric (MPB) partition the eigenvalues of H_0 are values of the diagonal part of the Fock operator:

$$\begin{aligned} E_k^{(0)} &= \langle\Psi_k^{(0)}|\hat{F}_{\text{diag}}|\Psi_k^{(0)}\rangle = \sum_{N\in\mathcal{S}} |C_N^k|^2 E_N \\ E_I &= \langle I|\hat{F}_{\text{diag}}|I\rangle. \end{aligned} \tag{6}$$

E_N and E_I are sums of one-electron matrix elements F_{ii} over the occupied spin orbitals of the determinants $|N\rangle$ and $|I\rangle$. The F_{ii} coincide with the ''orbital energies'' in the case of canonical SCF orbitals.

In the Epstein–Nesbet (EN) partition, $\hat{H}_0$ is the diagonal part of the matrix representation of $\hat{H}$ in the basis $\{|\Psi_k^{(0)}\rangle \in \mathcal{S}, |I\rangle \in \mathcal{P}\}$, in fact we assume

$$\begin{aligned} E_k^{(0)} &= \langle\Psi_k^{(0)}|\hat{H}_0|\Psi_k^{(0)}\rangle = E_k^{(MR)} \\ E_I &= \langle I|\hat{H}_0|I\rangle. \end{aligned} \tag{7}$$

In this case $E_k^{(0)}$ coincides with the multireference energy $E_k^{(MR)}$.

In both cases the total energies are the sum of the multireference energy and the second-order correction, $E_k^{(MR)} + E_k^{(2)}$. For a discussion about advantages and problems of the two partitions see, for example, Refs. 16 and 17. Generally it is very useful to compare the results of both partitions.

A typical applications of the CIPSI method consists of few (usually 3–4) steps, with progressively enlarged $\mathcal{S}$ spaces: all the most important determinants for the perturbed wave function will be included in the $\mathcal{S}$ space of the next step. In this way the determinants belonging to $\mathcal{S}$ are individually selected by means of an iterative procedure. At every step, the appropriate choice of the $\mathcal{S}$ space is crucial, because it determines the accuracy of the perturbation corrections. The basic requirement is that the quality of $\mathcal{S}$ should be uniform throughout the set of calculations under consideration.

Recalling that the first-order correction to the wave function is:

$$|\Psi_k^{(1)}\rangle = \sum_{I\in\mathcal{P}} C_I^k |I\rangle,$$

with coefficients:

$$C_I^k = \frac{\langle I|\hat{V}|\Psi_k^{(0)}\rangle}{E_I - E_k^{(0)}},$$

we define, as a measure of the quality of the perturbation CI calculation, the square norm of the first-order correction to the wave function:

$$\sigma_k = \sum_{I\in\mathcal{P}} (C_I^k)^2 = \sum_{I\in\mathcal{P}} \frac{\langle I|\hat{V}|\Psi_k^{(0)}\rangle^2}{(E_I - E_k^{(0)})^2}. \tag{8}$$

If large interactions $\langle I|\hat{V}|\Psi_k^{(0)}\rangle$ and/or small denominators $E_I - E_k^{(0)}$ occur, the perturbative results are not accurate, and σ_k is correspondingly large. When the $\mathcal{S}$ space is extended, the most important contributions in Eq. (8) are eliminated; therefore σ_m usually decreases, although new terms involving determinants $|I\rangle \in \mathcal{P}$ are added. Recently, an algorithm to select the $\mathcal{S}$ space so as to obtain a preset value of σ for all states and geometries has been developed. Such procedure is called ''*aimed* selection.'' In order to set up an *aimed* selection, we must be able to predict the σ value we would obtain at the step $n+1$ of a CIPSI calculation, from the results of the step n, at least with some approximations (see for details Ref. 14).

Each step involves the following points (see the flow chart in Fig. 1):

(i) Once given a configurational space $\mathcal{S}_n$ the lowest desired eigensolutions are computed by the Davidson diagonalization procedure. The eigenstates are called zero-order states.

(ii) The first order perturbative contribution to the zero-order states, which arises from single and double excitations

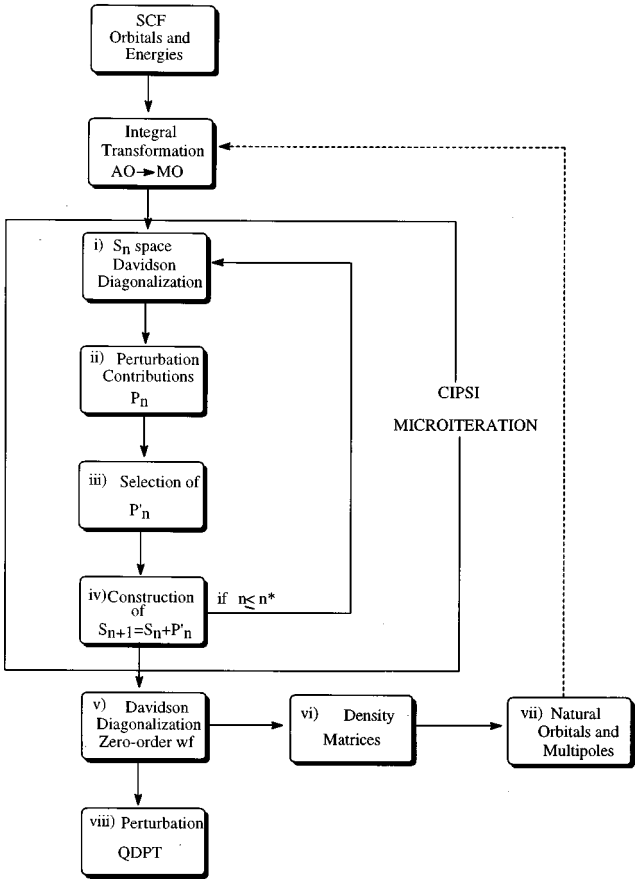

FIG. 1. Flow chart of CIPSI algorithm.

from all determinants belonging to the $\mathcal{S}_n$ space is computed. The space obtained with the single and double excitations from $\mathcal{S}_n$ is called $\mathcal{P}_n$.

(iii) A subspace $\mathcal{P}'_n$ of $\mathcal{P}_n$ is selected so that the next step with $\mathcal{S}_{n+1}=\mathcal{S}_n+\mathcal{P}'_n$ is supposed to give a norm of the first-order correction to the wave function equal to a target value σ^{n+1}. Notice that σ^{n+1} is the same for all desired electronic states.

(iv) The determinantal space $\mathcal{P}'_n$ is added to $\mathcal{S}_n$ giving the new variational space $\mathcal{S}_{n+1}$.

This procedure (from now on indicated as microiteration) is repeated n^* (usually 3 or 4) times until the dimension of the $\mathcal{S}$ space is large enough (10^4-10^5 determinants).

At this point final results are obtained through the following sequence:

(v) The CI matrix of the last $\mathcal{S}_{n*}$ space obtained through (i–v) is diagonalized in order to get the zero-order states.

(vi) The static and transition one-electron density matrices are calculated from the zero-order states and, if desired, corrected at the first order.[15] For example, the zero-order transition density matrix between the states k and j is written as:

$$^{kj}\rho(\mathbf{q},\mathbf{q}')=\sum_{rs}{}^{kj}\rho_{rs}\varphi_r(\mathbf{q})\varphi_s(\mathbf{q}'),$$

where in second-quantization formalism:

$$^{kj}\rho_{rs}=\sum_{M,N\in\mathcal{S}}(C_M^k)^*C_N^j\langle M|a_s^\dagger a_r|N\rangle,$$

and the $\{\varphi\}$ functions are the molecular orbitals.

(vii) Natural orbitals and properties such that static and transition multipoles are computed from the density matrices.

(viii) The second-order perturbation energy is computed with the diagrammatic code taking into account many more determinants (up to $\sim 10^{11}-10^{13}$).

As in the diagrammatic formulation the perturbation contribution of a single determinant is spread over different diagrams and never handled in full, we use the original formulation to select the determinants; only in the last step, the most time consuming, the diagrammatic method is employed. Both the diagrammatic and the original version of CIPSI allow a quasidegenerate perturbation theory calculation to be performed,[5] still adopting the CI eigenvectors in the $\mathcal{S}$ space as zero-order wave functions to take into account the interaction between electronic states at perturbative level.

The scheme above is the standard, simplest application of the CIPSI algorithm, however, many further modifications and/or additions are possible; in the following we shall use an improved technique which exploits natural orbitals of point (vii) as new molecular orbitals to start a further microiteration scheme (the dashed line in the flow chart). Another even more radical modification is given by the new strategy we have specifically formulated to include solvent effects described within the IEF-PCM framework.

III. PCM-IEF: THE THEORY

A. The basic formulation

In the PCM method the solvent S is represented by a homogeneous continuum medium which is polarized by the solute M placed in a cavity built in the bulk of the dielectric. The solute–solvent interactions are described in terms of a solvent reaction potential. The basic hypothesis of this kind of model is that one can always define a free energy functional, $\mathcal{G}(\Psi)$, depending on the solute electronic wave function Ψ. This energy functional can be expressed in the following form:

$$\mathcal{G}(\Psi)=\langle\Psi|\hat{H}^0|\Psi\rangle+\langle\Psi|V_p^R+\tfrac{1}{2}\hat{V}_{wf}^R|\Psi\rangle. \tag{9}$$

In Eq. (9) the Born–Oppenheimer (BO) approximation is employed.

The detailed description of the various terms appearing in Eq. (9) will be given in the following. Here it is worth noting that $\hat{H}^0$ is the Hamiltonian describing the isolated molecule, and $\hat{V}_p^R+\hat{V}_{wf}^R$ represents the solvent reaction operator. In particular V_p^R is the *constant potential* not depending on the solute wave function while $\hat{V}_{wf}^R$ describes the *response function of the reaction potential* depending on Ψ.

By applying the variational principle to this functional we can derive the nonlinear Schrödinger equation specific for the system under scrutiny. If we impose that first-order variations of $\mathcal{G}$ with respect to arbitrary variations of the

solute wave function Ψ are zero, and that the latter is normalized, the nonlinear Schrödinger equation becomes:

$$\hat{H}_{\text{eff}}|\Psi\rangle=(\hat{H}^0+\hat{V}_p^R+\hat{V}_{wf}^R)|\Psi\rangle=E|\Psi\rangle, \tag{10}$$

where E is the Lagrange multiplier introduced to fulfill the normalization condition on the electronic wave function. Equation (10) defines the specific ''effective Hamiltonian,'' $\hat{H}_{\text{eff}}$, giving the name to the whole procedure.

The free energy functional $\mathcal{G}$ has a privileged role in the theory, as the solution Ψ of the Schrödinger equation yields a minimum $G(\Psi)$ of this functional, which however differs from the eigenvalue E by a quantity $-\frac{1}{2}\langle\Psi|V_{wf}^R|\Psi\rangle$. We recall that in the standard linear Hamiltonians these two quantities, the Hamiltonian eigenvalue and the variational functional, coincide. The difference between E and $\mathcal{G}$ has, however, a clear physical meaning; it represents the polarization work that the solute does to create the charge density inside the solvent.

We recall that the free energy functional defined above is limited to the electronic part; the complete expression (G) should also include the nuclear repulsion energy V_{NN} of the solute M as well as the interaction energy between solute nuclei and the solvent reaction component generated by the solute nuclear charge, $\frac{1}{2}E_{NN}$.

Once we have defined the basic aspects of the method, the attention returns to the reaction potential operator determining the solvent interactions. As we have seen, this is generally partitioned into two terms. The former, $\hat{V}_p^R$, is often assimilated to an average solvent potential, it can also play the role of an instantaneous solvent potential with respect to a dynamical behavior. The presence of such perturbative term in the molecular Hamiltonian does not lead to any difficulty, neither from the theoretical point of view, nor from the practical. Many examples are known in which an external potential is introduced in the molecular calculations. On the contrary, the treatment of the other term, the operator $\hat{V}_{wf}^R(\Psi)$, is more delicate, as it gives origin to the nonlinear character of the Schrödinger Eq. (10); this fundamental point will be considered in more details in the following.

In general, the computational strategy formulated to define the reaction potential is based on a modelization of the solvent interactions derived from the theory of intermolecular forces. Within this framework the energetic quantity G is written as a sum of contributions of different physical origin, related to dispersion, repulsion and electrostatic forces between solute and solvent molecules. In this article, we shall mainly focus on the electrostatic part of the interaction, adding repulsion effects as a further refinement. No dispersion contributions will be considered.

The electrostatic problem of a charge distribution ρ_M embedded in a cavity C surrounded by a continuum dielectric strongly depends on the macroscopic structural characteristics of the dielectric itself. For example, we can define isotropic dielectrics, characterized by a constant scalar permittivity, ϵ, anisotropic dielectrics, characterized by a constant tensorial permittivity, $\boldsymbol{\epsilon}$, and ionic solutions in which the dissolved electrolyte charges are free to move in the surrounding isotropic medium.

Within classical electrostatics, these systems are all represented by an infinite dielectric surrounding the cavity, with a given function for the permittivity; the corresponding analytical relations yield;[11]

$$\begin{cases} -\Delta V=4\pi\rho_M & \text{inside } C \\ L_e V=0 & \text{outside } C \\ [V]=0 & \text{on } \Sigma \\ [\partial_L V]=0 & \text{on } \Sigma \end{cases}, \tag{11}$$

where V indicates the electrostatic potential and Σ is the cavity surface. The operator L_e as well as the last two boundary conditions (continuity of the potential and of its gradient across the surface) have to be differently defined for each particular system we are considering (isotropic dielectrics, anisotropic dielectrics or ionic solutions). For more details on the formalism exploited in Eq. (11), see Ref. 11.

The solution of the general system [Eq. (11)] is quite a hard task. Until now, each particular system has been solved by exploiting a different approach. On the contrary, the integral equation formalism (IEF)[11] we shall exploit in the this article manages to treat all the different solvation systems listed above within a common integral equation-like approach. In this framework, by defining the proper Green functions one can transform the first two equations in Eq. (11) into integral equations on the surface Σ, that can be solved with standard numerical methods. The solution of system Eq. (11) is thus reduced to a sum of two electrostatic potentials, one produced by ρ_M and the other due to a surface charge distribution σ, placed on the interface, arising from the polarization of the dielectric medium:

$$V(x)=V_M(x)+V_\sigma(x)=\int_{R^3}\frac{\rho_M(y)}{|x-y|}dy+\int_\Sigma\frac{\sigma(s)}{|x-s|}ds, \tag{12}$$

where the integral on the first term is taken over the entire three-dimensional space.

The problem is then shifted to the definition of the apparent surface charge (ASC) density $\sigma(s)$ for the various systems. The mathematical derivation of the method as well as the details of its computational implementation can be found in the source Ref. 11; here we shall report only the final expressions.

In the computational practice, use is made of a cavity formed by interlocking spheres centered on selected points (generally the solute nuclei). The surface of such cavity is then partitioned in small regions, called tesserae, with known area a_k. In the limit of a sufficiently accurate mapping of the cavity surface, one can always approximate the continuum charge $\sigma(s)$ on each tessera as a single-value quantity so as to define the equivalent sets of point-like charges as $q(s_k)=\sigma(s_k)a_k$ where s_k indicates the location of the single tessera k at which we calculate the representative value of $\sigma(s)$.

In the IEF framework the equations defining the point charges, collected in the vector $\mathbf{q}$ take on the following matrix form:

$$\mathbf{Aq}=-\mathbf{g}. \tag{13}$$

Here we do not report the definitions of the matrix **A** and vector **g**, to be specified in a different way for each dielectric (isotropic, anisotropic, or ionic solutions) but we recall that in all systems **A** only depends on the geometrical parameters defining the cavity and its tessellation, and on the form of the permittivity, while $\mathbf{g}_k$ also contains terms related to the electrostatic potential and field generated by the solute charge (electronic plus nuclear contributions) at the point s_k.

For standard isotropic solvents as that chosen for the numerical applications we shall present in the following, **g** can be reduced to the solute potential matrix **V** and **A** is transformed in the following matrix **Q**:

$$\mathbf{Q}=(\mathbf{I}/2-\mathbf{D})^{-1}\left(\frac{\mathbf{I}}{2}\frac{\epsilon+1}{\epsilon-1}-\mathbf{D}\right)\mathbf{S}, \tag{14}$$

with **I** indicating the unit matrix. Details on the derivation of this simplified, but still exact, model as well as definitions of the geometrical matrices **S** and **D** can be found in Ref. 20.

From the partition of the potential matrix **V** into electronic and nuclear components a corresponding partition in ''electron'' and ''nuclei-induced'' charges: $\mathbf{q}=\mathbf{q}^e+\mathbf{q}^N$ can be introduced, where $\mathbf{q}^x=-\mathbf{Q}^{-1}\mathbf{V}^x$.

Going back to the quantum mechanical scheme reported at the beginning of the section, here limited to the standard approximation of the molecular orbitals as combinations of atomic basis set $\{\chi_\mu\}$, the reaction potential $\hat{V}_p^R+\hat{V}_{wf}^R$ to be introduced in the effective Hamiltonian of Eq. (10) becomes the sum of two matrices **j** and $\mathbf{X}(\mathbf{R})$, corresponding to the constant and the wave function dependent term, respectively (**R** is the one-electron density matrix). In the limit of isotropic solvents described within the framework of Eqs. (13) and (14) we have

$$\mathbf{j}_{\mu\nu}=-\sum_k \mathbf{V}_{\mu\nu}(s_k)q^N(s_k);\quad \mathbf{q}^N=-\mathbf{Q}^{-1}\mathbf{V}^N, \tag{15}$$

$$\mathbf{X}(\mathbf{R})_{\mu\nu}=-\sum_k \mathbf{V}_{\mu\nu}(s_k)q^e(s_k);\quad \mathbf{q}^e(\mathbf{R})=-\mathbf{Q}^{-1}\mathbf{V}^e(\mathbf{R}), \tag{16}$$

where $\mathbf{V}_{\mu\nu}(s_k)$ are the potential integrals on the atomic basis set computed at the collocation points s_k of the tesserae and the electron-induced $\mathbf{q}^e$ charges depend on the solute density matrix **R** through the potential $\mathbf{V}^e(\mathbf{R})$ due to the electronic charge of the solute. The *constant* term **j** is computed in terms of the nuclei-induced charge $\mathbf{q}^N$.

By reducing the description of solvent effects to a couple of one-electron operators (**X** is actually a product of two of these operators), the PCM-IEF method can be straightforwardly applied to many levels of the QM description, and modeled to include various concepts and approaches provided by the general QM theory. The important new aspect to be taken into account is the introduction of an additional nonlinear character not present in isolated systems; the apparent charges $\mathbf{q}^e$ depend in fact on the solute electronic charge distribution (here represented in terms of **R**) which they contribute to modify through the related interaction matrix **X**. In the following we shall see how this deeply affects the implementation of CI-type procedures like that exploited in the CIPSI algorithm.

B. Dynamical aspects

Until now, the model has been implicitly limited to describe static systems; when dynamical aspects have to be taken into account further refinements are necessary. In the following, we shall focus on a specific dynamical phenomenon, namely vertical electronic excitations in molecular solutes, in which the involved time intervals are so short (i.e., the excitation can be assumed as instantaneous) that the solvent response cannot completely follow the fast changes in the solute charge distribution.

In a very common and used approach, this incomplete response of the solvent is represented in terms of a polarization vector depending on time, $P(t)$, which is a more general expression of the solvent response we have previously described through the apparent surface charges. In its largest definition, $P(t)$ should account for all the degrees of freedom of the solvent molecules; however, in the specific case of fast phenomena, a reliable assumption is the following:[18]

$$P(t)=P_{\text{fast}}+P_{\text{slow}},$$

where P_{fast} represents the almost instantaneous polarization of solvent molecules due to their electronic motions and P_{slow} the analog contribution due to nuclear degrees of freedom and related motions: vibrational, rotational, and translational. Since the relaxation times of the two components of $P(t)$ are very different (i.e., $\tau_{\text{fast}}\ll\tau_{\text{slow}}$, by at least one or two orders of magnitude if one assumes a Debye-type response for the slow polarization) the corresponding behaviors can be very different. For example, after a sudden change from an equilibrium solute–solvent state, P_{fast} is still in equilibrium with the new solute charge distribution, while if the same change is rapid enough, P_{slow} remains fixed at the value corresponding to the solute charge distribution of the initial state. The different behavior of the two components of $P(t)$ gives origin to a state of nonequilibrium between solute and solvent which then will relax to the standard full equilibrium condition common to all static systems.

In the PCM-like framework such model is realized defining fast and slow surface charges related to different values of the dielectric constant.[19] From a physical point of view, we can always define the equations giving the fast surface charge by considering a system of equations as that reported in Eq. (11) this time defined in terms of the optical value ϵ_∞ of the dielectric constant. The slow surface charge is then obtained as difference of the total and the fast charges.

Such a scheme applies well to a vertical electronic transition from the solute ground state (GS) to a given excited state (ex); this excitation, in fact, can be safely assumed as much faster than solvent nuclear motions (exactly as in the standard Franck–Condon approximation for the solute nuclei). By applying the fast–slow model we obtain that during the excitation process the slow part of apparent charges remains fixed to the value corresponding to the molecular charge distribution of the initial state of full equilibrium. As a consequence in the final state only the fast charges will be in equilibrium with the new solute charge distribution ρ_M^{ex} describing the excited state. In this context of nonequilibrium σ_{fast} will be defined by a new system of equations, namely:

$$\begin{cases} -\Delta V = 4\pi\rho_M^{\mathrm{ex}} & \text{inside } C \\ -\epsilon_\infty \Delta V = 0 & \text{outside } C \\ V_i - V_e = 0 & \text{on } \Sigma \\ (\partial V/\partial n)_i - \epsilon_\infty (\partial V/\partial n)_e = 4\pi\sigma_{\mathrm{slow}} & \text{on } \Sigma \end{cases} . \tag{17}$$

The second jump condition on the gradient of the total electrostatic potential V has here a different form with respect to that of the equilibrium situation [see Eq. (11)], due to the presence of the constant slow charge σ_{slow} on the cavity surface.

Following the IEF approach described for the static case, the system Eq. (17) is solved by evaluating the related fast surface charge σ_{fast}, or better its discretized analog given by the set of point charges $\mathbf{q}_{\mathrm{fast}}$, through the following equation:[20]

$$\mathbf{q}_f = -\mathbf{Q}_f^{-1}(\mathbf{V}^{\mathrm{ex}} + \mathbf{V}_s), \tag{18}$$

where $\mathbf{Q}_f$ is defined as in Eq. (14) but with ϵ_∞ instead of ϵ. $\mathbf{V}^{\mathrm{ex}}$ is the potential due to the solute charge distribution in its excited state, whereas the additional potential $\mathbf{V}_s$ is due to the set of constant slow charges $\mathbf{q}_s$ computed in a previous calculation on the fully equilibrated ground state system.

In the nonequilibrium case as well the definition of the electrostatic free energy has to be changed; in particular, the contribution of the slow polarization is seen as an external fixed field which does not follow the variationally derived rules given before; the proper equation becomes:

$$G^{\mathrm{neq}} = G_f + E_{es}^{\mathrm{neq}} + E_{Ns}^{\mathrm{neq}} + \tfrac{1}{2} E_{fs}^{\mathrm{neq}} \tag{19}$$

where G_f is obtained in the standard way but with solvent induced matrices defined in terms of the fast charges only, $\mathbf{j}_f$ and $\mathbf{X}_f(\mathbf{R})$. The last three quantities in the right-hand side of Eq. (19) represent the interaction energies of the constant slow apparent charges with solute electrons, solute nuclei, and fast apparent charges, respectively. In addition, if we assume as reference solvent state the unperturbed, i.e., unpolarized, pure liquid S, as in the static model, some further contributions have to be considered. Eq. (19) has been derived with the slow part of the solvent polarization already switched on; hence in order to get a reliable comparison between equilibrium and nonequilibrium energies we have to take into account the energy needed to create the slow charge. This can be done by substracting to G^{neq} of Eq. (19) $1/2(E_{es}^{\mathrm{eq}} + E_{Ns}^{\mathrm{eq}} + E_{fs}^{\mathrm{eq}})$, where each term has the usual meaning but this time the solute charge distribution and the fast apparent charges are computed in the initial full equilibrium state. In this way all the possible aspects related to static and dynamic electrostatic interactions induced by solvent molecules have been properly introduced in the QM scheme.

C. Repulsion interactions

At the beginning of this section we have quoted other types of interactions possibly acting between solute and solvent molecules. In fact, even if often largely dominant, the electrostatic forces cannot saturate solvent effects. When excited state with more diffuse charge distributions are involved, it can be interesting to consider also repulsion effects between solute and solvent. Let us see how. In the most accurate theories, the modeling of repulsion interactions in solution is based either on a discrete molecular description of the liquid or on a continuum dielectric model.[9] Going back to the discrete approach, the repulsion contribution to the solvation free energy can be defined by the general expression

$$G_{\mathrm{rep}} = \rho_S \int V_{\mathrm{rep}}^{MS}(r) g_{MS}(r) dr. \tag{20}$$

Here the label M refers to the solute, S to the solvent, r is an appropriate set of coordinates defining the internal geometry of the complex MS, ρ_S is the number density and g_{MS} is a correlation function which is 0 inside the solute cavity ($r \in C$) and 1 outside ($r \notin C$). Generally this equation is applied to calculated or estimated potentials V_{rep}^{MS} available from the literature. Recently a new approach has been suggested[21] in which V_{rep}^{MS} is substituted with a suitable expression taken from the theory of intermolecular forces recalling that as mainly originated from the Pauli exclusion principle, the repulsion forces between two interacting molecules increase with the overlap of the two distributions and are strictly related to the density of electrons with the same spin.

As in the continuum approach here exploited the electron density of the solvent is not given, some approximations are necessary; namely that solvent valence electrons are localized in pairs described by a corresponding density matrix spreading in the volume around the molecular cavity C. The simplest way of defining such a density, each pair being localizable, is through a Gaussian representation of localized orbitals so that, after some algebra, the repulsion contribution to free energy becomes[21]

$$G_{\mathrm{rep}} = \kappa_r \int_{r \notin C} dr\, P_M(r), \tag{21}$$

$$\int_{r \notin C} dr\, P_M(r) = n_{el} + \int_\Sigma E_M(s) n\, ds, \tag{22}$$

where κ_r is a factor depending on solvent density ρ_S, number of valence electrons and molecular weight. In Eq. (22) we have exploited Gauss theorem to evaluate the portion of solute electronic charge outside the cavity [$E_M(s) \cdot n$ represents the component of the solute electrostatic field perpendicular to the cavity surface].

What is worth stressing of Eq. (21) is that going back to the approximation of the MOs in terms of atomic basis set, the repulsion effect can be described by a new one-electron matrix:

$$[\mathbf{h}_{\mathrm{rep}}]_{\mu\nu} = \kappa_r \left(S_{\mu\nu} - \frac{1}{4\pi} \int_\Sigma E_{\mu\nu}(s) ds \right), \tag{23}$$

to be added to the previous solvent-induced $\mathbf{j}$ and $\mathbf{X}$ matrices; here $\mathbf{S}$ is the overlap matrix on the atomic basis and $\mathbf{E}$ collects the electric field integrals on the same basis.

IV. IEF-CIPSI

In the two previous sections we have presented the basic aspects of the CIPSI algorithm and the PCM-IEF model (with possible inclusion of repulsion effects). Here we shall

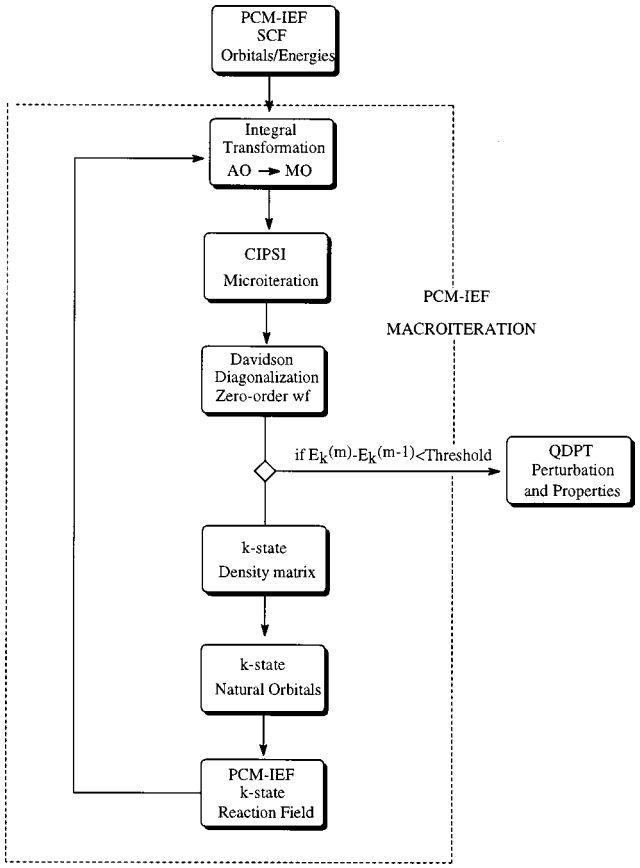

FIG. 2. Flow chart of IEF-CIPSI algoritm.

report and comment on the strategy we have formulated to merge the two approaches in a procedure coherently linking correlation and solvent effects. In particular, the analysis will be focused on the study of vertical electronic transitions of molecular solutes.

The fundamental element to be stressed from the very beginning is the peculiar aspect of such merging, i.e., the introduction of a nonlinear system into a complex CI-perturbative scheme as that characterizing the CIPSI algorithm.

In the present implementation the nonlinearity due to the dependence of the solvent reaction potential on the solute wave function is solved through an iterative procedure (from now on indicated as macroiteration) to be nested to that described in Sec. I for the general CIPSI algorithm [the microiteration resumed in points (i)–(iv)]. Let us see how this is realized (see the flow chart in Fig. 2).

From the step (i) of the first microiteration, the solvent-induced terms are added to the Hamiltonian to be diagonalized as one-electron operators $\mathbf{j}$, $\mathbf{X}$, and possibly $\mathbf{h}_{\mathrm{rep}}$, defined in Eqs. (15), (16), and (23), respectively; for the first two contributions the solution of the matrix system [Eq. (13)] defining the apparent charges is required. No formal and practical problems arise for both $\mathbf{j}$ and $\mathbf{h}_{\mathrm{rep}}$ as constant quantities; the same is not true for $\mathbf{X}$ which changes according to variations of the wave function. The strategy we have chosen to take into account such dependency is to repeat the microiteration schemes (i)–(iv), and the following CI diagonalization and evaluation of the density matrix [points (v)–(vi)], so that at each step m a new set of electron-induced charges $\mathbf{q}^e$ computed from the density matrix of step $m-1$ are calculated and used to evaluate the current $\mathbf{X}$ matrix. The procedure is repeated until convergency on the final energy is obtained (such scheme constitutes the previously quoted macroiteration).

The necessity of introducing an iterative scheme has also led us to improve the standard CIPSI algorithm adding a further degree of freedom in the sense that the molecular orbitals defining the CI basis are renewed at each step of the macroiteration in terms of the natural orbitals computed at point (vii); only in this way in fact solute wave function and solvent charges can completely readjust in a self-consistent scheme.

The main consequence of a correct application of the scheme outlined above is that each electronic state (ground and excited) requires a separated calculation involving a macroiteration optimized on the specific state of interest. This is obtained in the following way; in each macroiteration step the natural orbitals defining the CI basis and the density matrix used to evaluate the solvent-induced component of the effective Hamiltonian are those corresponding to the specific state under scrutiny so that in the following step both the solute wave function and the solvent charges better describe the real system.

The procedure briefly sketched above is valid both for the ground and the excited state fully equilibrated with the solvent; the inclusion of the nonequilibrium needs further refinements. First, we have to perform a ground state calculation from which the slow apparent charges, $\mathbf{q}_s$, and the related energies, $\frac{1}{2}(E_{es}^{\mathrm{eq}}+E_{Ns}^{\mathrm{eq}}+E_{fs}^{\mathrm{eq}})$, are obtained and then saved in a file for the successive calculation on the excited state. The latter is performed exactly in the same way as it is for the ground state but with an interaction potential composed by three components. The first two are constant as determined by the fixed slow charges $\mathbf{q}_s$ of the previous calculation, and the current nuclear charges $\mathbf{q}^N$, respectively. The third changes during the iteration procedure; in particular, the corresponding operator is defined in terms of the fast charges $\mathbf{q}_f$ as obtained from Eq. (18) with the potential matrix $\mathbf{V}^{\mathrm{ex}}$ determined by the charge distribution of the solute excited state.

At convergence the resulting CI energy of the excited state, $E_{\mathrm{ex}}=\langle\Psi_{\mathrm{ex}}|\hat{H}|\Psi_{\mathrm{ex}}\rangle$, must be corrected to obtain the free energy analog by adding a term similar to that used for the ground state, $-\frac{1}{2}\langle\hat{V}_{wf}^{R}\rangle$, but this time referred only to the fast polarization, $-\frac{1}{2}\langle\hat{V}_{wf}^{\mathrm{fast}}\rangle$, and some new terms accounting for the work done in the slow polarization of the solvent. Following the formalism exploited in Eq. (19) we can write the expression for the free energy of the excited state as

$$G_{\mathrm{ex}}^{\mathrm{neq}}=E_{\mathrm{ex}}-\tfrac{1}{2}\langle\Psi_{\mathrm{ex}}|V_{wf}^{\mathrm{fast}}|\Psi_{\mathrm{ex}}\rangle-\tfrac{1}{2}E_{\mathrm{es}}^{\mathrm{eq}}+\tfrac{1}{2}(E_{fs}^{\mathrm{neq}}-E_{fs}^{\mathrm{eq}}), \quad (24)$$

where the last two terms represent the difference in the interaction energy between fast and slow apparent charges going from the description of the ground state to that of the excited state. In Eq. (24) we have skipped the nuclear terms, $\frac{1}{2}E_{NN}+V_{NN}$, collecting both the solute nuclei repulsion and their interaction with the total nuclei-induced charges $\mathbf{q}^N$.

An important feature of the procedure briefly illustrated above, seldom present in the various theoretical models proposed to treat solvent effects, is that all the CI matrix elements, including the off-diagonal elements, are explicitly modified by the solute–solvent interaction. The present iterative scheme eliminates the nonlinear character of the wavefunction-dependent contribution of the solvent operator, which is now computed from the density matrix obtained in the previous step. In addition, the optimization on each single electronic state enables us to properly define the solvent operators acting on both diagonal and off-diagonal CI terms.

Until now no references have been made to the perturbative part of the algorithm; this applies both to the selection of determinants to be added to the initial $\mathcal{S}$ space during the microiteration and to the final diagrammatic evaluation of the second-order energy. In passing we recall that for solvated systems involving an additional macroiteration scheme the diagrammatic calculation is performed only once, namely when the zero-order energy has converged.

Concerning formal aspects related to the inclusion of a nonlinear Hamiltonian as that containing solvent terms into perturbative schemes, important studies have been done in the past.[22] However, only very recently a rigorous analysis focused on PCM-like models has been presented.[23] Even if this study explicitly refers to the second-order Moller–Plesset perturbation scheme for one-determinantal wavefunctions, it reports many points which can be straightforwardly extended to multireference wave functions as those exploited here. In particular, the fundamental aspect we can extract from such analysis is that one correctly introduces solvent effects in the second-order perturbation scheme by evaluating them from the reference zero-order wavefunction.

In this framework the perturbative Eqs. (3)–(7) are also still valid for solvated systems but they all require a redefinition of the involved quantities. The first difference is in the definition of the unperturbed part of the Hamiltonian [Eq. (3)], which for solvated systems includes also solvent operators [see Eq. (10)] computed on the final zero-order wave function obtained at convergency of the macroiteration scheme. The same considerations apply to the Fock operator exploited in the MPB partition Eq. (6), while the analog EN scheme is automatically modified by the different definition of $\hat{H}_0$.

V. NUMERICAL RESULTS: THE EXCITATION SPECTRUM OF FORMAMIDE

In the previous sections we have reported the basic formalism of both CIPSI and PCM-IEF algorithms and the way to merge them; however, many important computational aspects have been omitted there for brevity's sake and thus reported here.

To merge PCM-IEF and CIPSI, the standard CIPSI code has been interfaced to a development version of GAMESS package[24] in which the last version of the PCM-IEF model has been implemented. Referring to the flow charts of Figs. 1 and 2, GAMESS is exploited both in the initial SCF step and in recomputing all the data starting from the new natural orbitals and density matrix obtained at the end of the CIPSI microiteration. When solvent effects are included, GAMESS is also exploited to compute reaction field operators and integrals.

TABLE I. Exponents of the diffuse functions.

Type	Atom	Exponents	
s	C	0.04370	0.01752
s	N	0.05320	0.02100
s	O	0.06080	0.02400
p	C	0.01575	0.03990
p	N	0.04750	0.01875
p	O	0.05320	0.02100

The whole strategy has been applied to the study of the valence electronic spectra of formamide both *in vacuo* and in solution. Following the suggestions of previous theoretical gas-phase investigations, we have exploited a 6-31G** basis with two diffuse *s* functions and two diffuse *p* functions centered on each of the heavy atoms with exponents appropriate for the Rydberg states (the exponents of the diffuse functions are given in Table I).

As concerns solvent-related aspects, we recall that the molecular cavity enveloping the solute is built of interlocking spheres centered on the solute nuclei with radii R_k equal to 1.2 times the corresponding van der Waals values R_k^{vdw}; namely, we have: $R_H = 1.44$ Å, $R_C = 2.04$ Å, $R_O = 1.80$ Å, and $R_N = 1.92$ Å. The continuum medium has been described in terms of two dielectric constants, $\epsilon = 78.5$ and $\epsilon_\infty = 1.776$, corresponding to the static and the optical dielectric constants of liquid water at 298 K. Some results referring to acetonitrile as solvent are also reported, in this case values of 36.64 and 1.806 are exploited for ϵ and ϵ_∞, respectively. In solvation calculations including repulsion effects, values of 1.0, and 0.7899 are used for the numeral density of water and acetonitrile, respectively.

In all the excited state calculations, the solute geometry (and the molecular cavity for solvated systems) is kept fixed to the value optimized for the ground state both *in vacuo* and in solution; this means to assume a vertical transition in which the solute nuclei do not have enough time to relax towards the new electronic charge distribution describing the excited state. In particular, *in vacuo* and in water geometries are obtained at MP2 level using the standard 6-31+G** basis set; a selection of numerical values are reported in Table II. We recall that for these calculations we have used a development version of the GAUSSIAN code,[25] where PCM-IEF MP2 derivatives have been recently implemented.[23]

As in any CI procedure, also in the CIPSI algorithm, the fundamental step is the choice of molecular orbitals. Here many different options have been checked and compared. In particular, for formamide in which $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions are searched, the main difficulties arise from the use of diffuse functions in the atomic basis set; these in fact largely affect the standard description in terms of n, π, and π^* orbitals making the choice of the initial $\mathcal{S}_n$ space difficult (see Figs. 1 and 2). The best compromise between accuracy and meaningful chemical description has been given by a

TABLE II. MP2/6-31+G** geometrical parameters of formamide calculated *in vacuo* and in water. Bond distances r in Å, angles in degrees.

	Vacuo	Water
r_{CN}	1.3617	1.3447
r_{CO}	1.2287	1.2424
r_{NH_1}	1.0075	1.0095
r_{NH_2}	1.0050	1.0074
r_{CH}	1.0992	1.0951
$\angle$NCO	124.5303	124.6416
$\angle CNH_1$	119.2463	120.2838
$\angle CNH_2$	121.3586	120.5959
$\angle$NCH	112.8284	113.5135

TABLE III. Summary of theoretical gas-phase calculations. Transition energies are in eV.

	$n\pi^*$	R	$\pi\pi^*$
GVB[a]	5.65		
SCF[b]	5.7		8.5
CI[b]			6.96–8.21
CIPSI[c]	5.85	6.09	7.56
CASPT2[d]	5.85		7.67
MRCI[e]	5.86	6.14	7.94
MCSCF[f]	5.66		
CASSCF[g]	8.30	7.47	8.83
CASPT2[g]	5.61	6.52	7.41
EOM-CCSD[h]	5.71		7.66

[a]Reference 30(a).
[b]Reference 30(b).
[c]Reference 30(c).
[d]Reference 30(d).
[e]Reference 30(f).
[f]Reference 30(g).
[g]Reference 30(h).
[h]Reference 30(i).

new procedure[26] indicated as ''floating occupation;'' details on such technique are given in the following section.

A. Floating occupation technique

The basic idea is to allow electronic charge to distribute not only on the orbitals with energy less than Fermi energy but also on some virtual orbitals. The latter are no more a simple orthogonal complement to doubly occupied orbitals as in a pure SCF closed shell technique but they allow a better description of the electronic states in which they are significantly occupied. The procedure we have exploited is the following: at each SCF iteration we associate a Gaussian function $f_\xi(E)=\sqrt{(2/\pi\omega^2)}\exp[-(E-E_\xi)/2\omega^2]$, to each eigenvalue, E_ξ, of the Fock matrix, so that:

$$\int_{-\infty}^{+\infty} f_\xi(E)dE=2.$$

The integral from $-\infty$ to the Fermi energy represents the occupation of the ξth orbital:

$$n_\xi=\int_{-\infty}^{E_{\text{Fermi}}} f_\xi(E)dE.$$

The Fermi energy is defined so that a number of electrons equal to the total number N of electrons of the molecule is distributed on all K orbitals:

$$N=\sum_{\xi=1}^{K}\int_{-\infty}^{E_{\text{Fermi}}} f_\xi(E)dE.$$

Deeply inner orbitals will have occupation 2 while valence orbitals will have less and less electrons as their energy approaches the Fermi value; the first virtual orbitals will be occupied in such an amount as to complete the Gaussian tail until Fermi level. Higher-energy virtual orbitals will be empty. Once such occupations have been obtained, the new density matrix becomes:

$$\rho_{ij}=\sum_\xi C_{i\xi}n_\xi C_{j\xi}.$$

Such a matrix is then exploited in the following SCF iteration, and the whole cycle is repeated until convergency is reached. In this procedure the band width parameter, ω, can be chosen so as to distribute a given portion of charge on the virtual orbitals.

The ''floating orbital'' technique introduces two main advantages. First, the description of the lower virtual orbitals is significantly improved; in fact, from the first SCF iteration, they contain a certain portion of electrons and can get a lower energy in the following step. The quality of virtual orbitals can be quantified as follows. If we build excited electronic states in two different ways, namely from floating orbitals or from standard SCF RHF orbitals, it is found that within a given threshold of accuracy, in the floating scheme a smaller number of determinants is required. The second advantage is that those orbitals which are equivalent by symmetry are automatically degenerate and with the same electronic occupation.

The only evident disadvantage introduced by such a technique is that Fock energies loose their intrinsic meaning: they will depend on the portion of electrons distributed on virtual orbitals, and then on the width ω. This drawback requires the orbital energies to be recomputed before the CI calculation assuming double occupations.

B. Results

The electronic absorption spectrum of formamide *in vacuo* has been measured by several groups.[27,28] In the spectrum five absorption bands have been identified.[28] The first three bands, which are of interest in this article, are a weak transition corresponding to the $n\pi^*$ excitation at 5.65 eV (labeled the W band), a Rydberg excitation at 6.7 eV (the R_1 band), and an intense $\pi\pi^*$ excitation at 7.32 eV, termed the V_1 band. Recent measurements[29] give a value of at most 5.8 eV for the $n\pi^*$ transition and of 7.36 eV for the $\pi\pi^*$. The experimental oscillator strength for the $\pi\pi^*$ transition ranges from 0.31 to 0.37.[28,29]

Several theoretical studies have been performed so far on gas-phase electronic spectra of formamide;[30] a summary of the computed transition energies is reported in Table III with the related references.

The previously mentioned studies shed light on the necessity of performing suitable calculations in order to reach a

TABLE IV. Experimental and CIPSI/6-31+G** transition energies (eV) and dipole moments (a.u.) of formamide *in vacuo*. The computed energies refer to variational CI, perturbative MP, and EN levels of theory.

	var	MP	EN	exp[a]	μ
GS					1.647
$n\pi^*$	5.62	5.59	5.65	5.65	0.779
R	7.44	7.04	7.20	6.7	3.786
$\pi\pi^*$	7.80	7.41	7.56	7.32	2.271

[a]Reference 28.

realistic description of the excited states and then to obtain reliable results of the excitation energies. In particular, these studies have remarked that it is necessary to use an adequate basis set, to use a set of molecular orbitals suitable for the whole of the excited states, and not favoring a particular state. Finally, it has been shown that it is necessary to properly take into account the contribution arising from the electron correlation.

Following such suggestions we have repeated the study exploiting the CIPSI algorithm described in the previous sections. In particular, the computational strategy we have used starts with a first SCF calculation using the ''floating-occupation'' technique presented above and then, at the end of the CIPSI microiteration, the resulting natural orbitals are used as new CI basis in a second microiteration scheme (see Fig. 1, dashed line). At the end, the second-order perturbation energies of all the states are computed with the diagrammatic code. The numerical results for zero- and second-order calculations (in both MP and EN perturbative schemes) are reported in Table IV. The geometrical parameters used are shown in Table II.

The analysis of the molecular orbitals involved has led to identify the three lowest excited states as $n\pi^*$, ns Rydberg, and $\pi\pi^*$ transitions corresponding to the W, R_1, and V_1 bands of the experimental spectrum, respectively; the parallel analysis of higher states has not been done. The computed transition energies agree well with experimental analogs. Perturbative corrections seem not to significantly improve the results as concerns the first two transitions. On the contrary, the too high value computed for the $\pi\pi^*$ excitation at the CI-variational level is largely improved by such corrections, which reduce the discrepancy with respect to the experiment to 0.09 eV (at MP level).

Passing to data in solution, to the best of our knowledge the only available experimental values regarding formamide are those reported by Basch for acetonitrile as solvent:[28] the reported $n\pi^*$ excitation energy is 5.58 eV and the $\pi\pi^*$ is 6.81 eV. However it is worth noticing that the $n\pi^*$ value is not fully reliable, since in the spectrum this band is weak and partially covered by the $\pi\pi^*$. Other available experimental data in solution regard a similar molecular system, acetamide. In this case the $n\pi^*$ and $\pi\pi^*$ excitation energies are, respectively, near 5.5 and 6.96 eV in cyclohexane and near 5.9 and 6.81 eV in water (both the $n\pi^*$ values are quite uncertain).[31] Using the last set of data, it can then be evaluated in water: a blueshift of $\sim$0.25 (eV) for the $n\pi^*$ transition and a redshift of 0.51 eV for the $\pi\pi^*$.

Regarding previous calculations in solution, there are only a few available studies. Sobolewski[32] investigated the $n\pi^*$ and $\pi\pi^*$ states of a complex of formamide and one molecule of water, finding transition energies of 6.03 and 7.52 eV, respectively. Krauss and Webb[33] found an energy of 6.14 eV for the $n\pi^*$ transition in water, using the effective fragment potential model.[34] Finally Besley and Hirst[35] using the Onsager model found transition energies of 5.51 and 6.87 eV in cyclohexane for the $n\pi^*$ and $\pi\pi^*$ bands, respectively. Passing from a nonpolar to a polar solvent (water) they did not find a clear shift in the two previously mentioned transitions, obtaining 5.48 and 6.86 eV, respectively.

TABLE V. CIPSI/6-31+G** transition energies (eV) and dipole moments (a.u.) of formamide in water obtained with three different solvation models.

	var	MP	EN	μ
Equilibrium				
GS				2.224
$n\pi^*$	5.82	5.83	6.00	0.962
$\pi\pi^*$	7.18	6.89	7.22	2.661
Nonequilibrium				
$n\pi^*$	5.83	5.86	6.04	1.098
$\pi\pi^*$	7.23	7.02	7.28	2.646
neq+repulsion				
GS				2.198
$n\pi^*$	5.88	5.81	5.97	1.096
$\pi\pi^*$	7.45	7.15	7.39	2.174

The results obtained in this article by exploiting the merging of CIPSI and PCM-IEF methods are reported in Table V. The three sets of numbers refer to the two electrostatic-only solvation models (assuming a complete, equilibrium, or a partial, nonequilibrium, response of solvent molecules), and to the nonequilibrium electrostatic +repulsion scheme, respectively. The exploited geometry is reported in Table II.

The first point to be stressed is that in all three solvation models the Rydberg state found in gas phase is shifted to higher energies, above the two valence $n\pi^*$ and $\pi\pi^*$ states. In addition, taking into account the specific characteristics of these two states it is also immediate to see that all models give the right trend, showing a blue- and a redshift with respect to gas-phase results for the transition to the less polar $n\pi^*$ and the more polar $\pi\pi^*$ state, respectively.

Passing now to a more detailed analysis, we can note that the more sensitive state is the $\pi\pi^*$, the only one showing not-negligible differences in the three different models. In particular, focusing on nonequilibrium effects, we obtain an averaged increase of 0.1 eV in the transition energy passing from a fully relaxed solvent (the equilibrium model) to a partially frozen one (the nonequilibrium model). This difference can be explained by observing that the nonequilibrium scheme applies only to the excited state (ground state is by definition characterized by a full equilibrium between solute and solvent) and that such a state feels a smaller stabilizing effect from the partially frozen solvent with respect to the case of a complete response. This, as a result, increases the energy of the excited state leading to a higher transition en-

TABLE VI. CIPSI/6-31+G** transition energies (eV) and dipole moments (a.u.) of formamide in acetonitrile (including nonequilibrium and repulsion effects). The geometry is that optimized in water.

	var	MP	EN	μ
GS				2.045
$n\pi^*$	5.55	5.63	5.83	1.001
$\pi\pi^*$	7.37	7.11	7.34	2.506

ergy. The same effect only slightly modifies the $n\pi^*$ state, where the difference between the two models is of the order of 0.01–0.04 eV (passing from variational to perturbative values).

In Sec. III, we have recalled that even if the electrostatic component is the most important contribution to the solvation effect; consideration of other terms can give a more complete description of the phenomenon. In particular, the Pauli repulsion between solute and solvent is expected to give not-negligible contributions when considering excited states with quite diffuse charge distributions.[36] The computed results confirm this prediction. Once again the more sensitive $\pi\pi^*$ state shows the largest effect; the related transition energy increases by 0.31 and 0.24 eV (at the variational level) with respect to the two electrostatic-only models. This important change is also reflected in the dipole value which is largely decreased ($\sim$0.5 a.u.) due to repulsion effects. Very small modifications are, on the contrary, induced on the $n\pi^*$ state as shown by both transition energy and dipole values.

All the analyses above can be summarized in a comparison of solvatochromic shifts with respect to gas-phase results. If we can assume that there are not large differences between spectra of acetamide and formamide, the computed results show a quite good agreement with experiments, in particular for nonequilibrium model (0.21 and -0.57 eV for $n\pi^*$ and $\pi\pi^*$ transitions, respectively); larger discrepancies appear in full equilibrium calculation but this could be easily expected from the rough approximations of the physical model. This in fact assumes that the nuclear degrees of freedom of solvent molecules can instantaneously relax exactly as the electronic counterpart. More delicate is the analysis of changes induced by repulsion; here in fact two different interactions act at the same time combining their effects. The final result is reflected in an increase in the $n\pi^*$ shift (0.26 eV) and a parallel decrease in the $\pi\pi^*$ (-0.35 eV). The modifications induced by perturbative corrections are quite small for the lowest transition in all the solvation models, while more evident effects can be observed for the $\pi\pi^*$ analog.

All the solvation data discussed above are obtained at a geometry optimized in water. This means to introduce two different solvent effects, one indirect due to geometry changes and the other direct due to modifications in the charge distribution. To split this combined action to get more hints on the direct contribution of easier use for interpretative purposes, we have repeated the calculations in another solvent (acetonitrile) but this time keeping the gas-phase geometry. The computed results are reported in Table VI. Here, the only solvation model we have considered includes both nonequilibrium and repulsion effects.

From the comparison of these data with the parallel ones obtained in water (see third set of results in Table V) we can obtain some useful information. First, the well-known electrostatic effect due to the reduced polarity of the solvent is completely confirmed by the smaller value computed for the $n\pi^*$ transition energy and the related shift with respect to the gas phase. This purely electrostatic phenomenon which is dominant for this lowest transition for which repulsion interactions have shown to be negligible (see calculations in water), is on the contrary, effectively opposed by repulsion in the other transition. Here, by considering only electrostatic interactions one should expect a higher transition energy in acetonitrile than in water, due to the less polar character of the former solvent. However, the computed results show that in this case the opposite action of repulsion when applied to the two different systems reverses the predicted trend.

VI. CONCLUSIONS

In this article we have reported and discussed a method joining the potentialities of the CIPSI algorithm with those of PCM-IEF model, possibly including solvent repulsion effects. Such a method enables to achieve both an accurate description of the solute wave function and a good estimate of the solvent effects in a self-consistent way as confirmed by the promising numerical results obtained in the study of excitation spectra of formamide. However, the numerical aspects do not have to hide the main purpose of the paper, namely giving a further contribution to the inclusion of solvation models into accurate quantum mechanical descriptions. Following previous examples in this direction given by our group, much attention has been devoted to the theoretical aspects of the merging, in particular to the analysis of formal problems arising from the inclusion of nonlinear operators into CI and perturbative schemes. We recall here that an almost equivalent analysis could be easily extended to other closely related schemes, such as CASPT2 since no evident difficulties arise from the theoretical point of view.

Parallel to the methodological study, we have also formulated and implemented a new computational strategy interfacing a largely used package as GAMESS with the more special CIPSI algorithm. The resulting tool which is not limited to solvation studies can be exploited in many different fields by anyone looking for an accurate description of molecular systems. As concerns specific applications to solvated systems, the strategy presented in the previous sections represents a quite refined application of the general methodology. The specific phenomenon under scrutiny in fact required a detailed description of all the possible aspects of solvation and the way they affect solute electronic charge. However, for more general studies simpler schemes can be formulated still keeping the characteristics of accuracy of the method. In this direction one could think of applying the same strategy for including the solvation effects into other more widely used approaches than the CIPSI algorithm; the previously mentioned CASPT2 is just one of the possible examples. Concerning this point, perhaps it is worth saying something more.

In our limited experience, a rather well described wave function is required to evaluate reliable solvent interactions; in particular, as concerns the specific system presented in this article, we found out that a CAS-SCF description was not good enough but that a more accurate wave function, and the related density matrix, as those given by the CIPSI algorithm were necessary. Obviously this is just a single case and further examples should be considered before drawing any stated conclusion. However even these first limited results seem to show that CIPSI scheme presents some real advantages with respect to other related techniques. We have just stressed the high quality of the resulting wave function but also the use of selected spaces which reduce the dimension of the problem, and the possibility of exploiting perturbed density matrices can represent good reasons to prefer it. In addition, the quite promising performances we have obtained by interfacing CIPSI to the IEF-PCM model, strongly suggest that further applications to more complex molecular systems and different chemical phenomena should be considered.

ACKNOWLEDGMENTS

The authors would like to thank Professor J. Tomasi for his continuous help both in the formulation of the method and in the preparation of the manuscript. We are also grateful to Professor M. Persico for useful discussions and revising the manuscript.

[1] B. Huron, J.-P. Malrieu, and P. Rancurel, J. Chem. Phys. **58**, 5745 (1973).
[2] S. Evangelisti, J.-P. Daudey, and J.-P. Malrieu, Chem. Phys. **75**, 91 (1983).
[3] F. Spiegelmann and J.-P. Malrieu, J. Phys. B **17**, 1235 (1984).
[4] R. Cimiraglia, J. Chem. Phys. **83**, 1746 (1985).
[5] R. Cimiraglia and M. Persico, J. Comput. Chem. **8**, 39 (1987).
[6] For MCSCF calculation: (a) K. V. Mikkelsen, A. Cesar, H. Ågren, and H. J. A. Jensen, J. Chem. Phys. **103**, 9010 (1995); (b) R. R. Pappalardo, M. Reguero, M. A. Robb, and M. Frisch, Chem. Phys. Lett. **212**, 12 (1993); (c) L. Serrano-Andrés, M. P. Fülscher, and G. Karlström, Int. J. Quantum Chem. **65**, 167 (1997).
[7] For DFT calculations: (a) M. F. Ruiz-Lopez, F. Bohr, M. T. C. Martinis Costa, and D. Rinaldi, Chem. Phys. **221**, 109 (1994); (b) J. L. Chen, A. Noodleman, D. A. Case, and D. Bushford, J. Phys. Chem. **98**, 11059 (1994); (c) T. N. Truong and E. V. Stefanovich, Chem. Phys. Lett. **240**, 253 (1995); (d) J. Andzelm, C. Kolmel, and A. Klamt, J. Chem. Phys. **103**, 9312 (1995); (e) T. Zhu, J. Li, C. J. Cramer, and D. G. Truhlar, *ibid.* **109**, 9117 (1998).
[8] S. Miertus, E. Scrocco, and J. Tomasi, J. Chem. Phys. **55**, 117 (1981).
[9] J. Tomasi and M. Persico, Chem. Rev. **84**, 2027 (1994).
[10] R. Cammi and J. Tomasi, J. Comput. Chem. **16**, 1449 (1995).
[11] (a) E. Cancès and B. Mennucci, J. Math. Chem. **23**, 309 (1998); (b) E. Cancès, B. Mennucci, and J. Tomasi, J. Chem. Phys. **107**, 3032 (1997); (c) B. Mennucci, E. Cancès, and J. Tomasi, J. Phys. Chem. B **101**, 10506 (1997).
[12] For CI and/or MCSCF: (a) B. Mennucci, R. Cammi, and J. Tomasi, J. Chem. Phys. **109**, 2798 (1998); (b) M. A. Aguilar, F. J. Olivares del Valle, and J. Tomasi, *ibid.* **98**, 7375 (1993); (c) C. Amovilli, B. Mennucci, and F. M. Floris, J. Phys. Chem. B **102**, 3023 (1998); (d) M. Cossi, V. Barone, and M. A. Robb, J. Chem. Phys. (submitted). For DFT: (e) A. Fortunelli and J. Tomasi, Chem. Phys. Lett. **231**, 34 (1994); (f) M. Cossi, V. Barone, R. Cammi, and J. Tomasi, *ibid.* **255**, 327 (1996). For MP2: (g) R. Cammi, B. Mennucci, and J. Tomasi, J. Phys. Chem. A (in press).
[13] C. Angeli, R. Cimiraglia, M. Persico, and A. Toniolo, Theor. Chem. Acc. **98**, 57 (1997).
[14] C. Angeli and M. Persico, Theor. Chem. Acc. **98**, 117 (1997).
[15] C. Angeli, R. Cimiraglia, and M. Persico, Theor. Chem. Acc. **100**, 324 (1998).
[16] R. Cimiraglia, Int. J. Quantum Chem. **60**, 167 (1996).
[17] J.-P. Malrieu and F. Spiegelmann, Theor. Chim. Acta **52**, 55 (1979).
[18] (a) S. Lee and J. T. Hynes, J. Chem. Phys. **88**, 6853 (1988); **88**, 6863 (1988); (b) H. J. Kim and J. T. Hynes, *ibid.* **93**, 5194 (1990); (c) A. M. Berezhkosvskii, Chem. Phys. **164**, 331 (1992); (d) M. V. Basilevski and G. E. Chudinov, *ibid.* **144**, 155 (1990); (e) R. Marcus, J. Phys. Chem. **96**, 1753 (1992); (d) reference (a) in Ref. 6.
[19] (a) R. Bonaccorsi, R. Cimiraglia, and J. Tomasi, J. Comput. Chem. **4**, 567 (1983); (b) M. A. Aguilar, F. J. Olivares del Valle, and J. Tomasi, J. Chem. Phys. **98**, 7375 (1993); (c) M. L. Sánchez, M. A. Aguilar, and F. J. Olivares del Valle, J. Phys. Chem. **99**, 15758 (1995); (d) R. Cammi and J. Tomasi, Int. J. Quantum Chem., Quantum Chem. Symp. **29**, 465 (1995).
[20] Reference (a) in Ref. 12.
[21] C. Amovilli and B. Mennucci, J. Phys. Chem. B **101**, 1051 (1997).
[22] (a) J. G. Ángyán, Int. J. Quantum Chem. **47**, 469 (1993); (b) F. J. Olivares del Valle, M. A. Aguilar, and S. Tolosa, J. Mol. Struct.: THEOCHEM **279**, 223 (1993); (c) A. Willetts and J. E. Rice, J. Chem. Phys. **99**, 426 (1993), just to quote the most representative ones.
[23] Reference (g) in Ref. 12.
[24] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).
[25] GAUSSIAN 99, Development Version (Revision A.10); M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, C. S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, A. G. Baboul, J. Cioslowski, B. B. Stefanov, G. Liu, C. A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, C. M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, and J. A. Pople (Gaussian, Inc., Pittsburgh, PA, 1998).
[26] M. Persico and A. Toniolo (unpublished)
[27] (a) H. D. Hunt and W. T. Simpson, J. Am. Chem. Soc. **75**, 4540 (1953); (b) D. L. Peterson and W. T. Simpson, *ibid.* **79**, 2375 (1957); (c) K. Kaya and S. Nagakura, Theor. Chim. Acta **7**, 117 (1967).
[28] H. Basch, M. B. Robin, and N. A. Kuebler, J. Chem. Phys. **49**, 5007 (1968).
[29] J. M. Gingell, N. J. Mason, H. Zhao, I. C. Walker, and M. R. F. Siggel, Chem. Phys. **220**, 191 (1997).
[30] (a) L. B. Harding and W. A. Goddard, J. Am. Chem. Soc. **97**, 6300 (1975); (b) L. Z. Stenkamp and E. R. Davidson, Theor. Chim. Acta **44**, 405 (1977); (c) E. Oliveros, M. Riviere, C. Teichteil, and J. P. Malrieu, Chem. Phys. Lett. **57**, 220 (1978); (d) A. L. Sobolewski, J. Photochem. Photobiol. **89**, 89 (1995); (e) S. L. Price, J. S. Andrews, C. W. Murray, and R. D. Amos, J. Am. Chem. Soc. **114**, 8268 (1992); (f) J. D. Hirst, D. M. Hirst, and C. L. Brooks III, J. Phys. Chem. **100**, 13487 (1996); (g) M. Krauss and S. P. Webb, J. Chem. Phys. **107**, 5771 (1997); (h) L. Serrano-Andrés and M. P. Fülscher, J. Am. Chem. Soc. **118**, 12190 (1996); (i) P. Szalay and G. Fogarasi, Chem. Phys. Lett. **270**, 406 (1997).
[31] E. B. Nielsen and J. A. Schellman, J. Phys. Chem. **71**, 2297 (1967).
[32] Reference (d) in Ref. 30.
[33] Reference (g) in Ref. 30.
[34] W. Chen and M. S. Gordon, J. Chem. Phys. **105**, 11081 (1996).
[35] N. A. Besley and J. D. Hirst, J. Phys. Chem. A **102**, 10791 (1998).
[36] B. Mennucci, C. Amovilli, and J. Tomasi, Chem. Phys. Lett. **286**, 221 (1998).