

# Theoretical investigation of the electronic spectrum of pyrazine

Clemens Woywod · Attila Papp · Gábor J. Halász ·  
Ágnes Vibók

Received: 9 February 2009 / Accepted: 27 October 2009 / Published online: 18 November 2009  
© Springer-Verlag 2009

**Abstract** The description of Rydberg states by the complete active space self-consistent field (CASSCF) electronic structure method is known to be a difficult topic. In particular, two problems are frequently encountered: (a) the simultaneous presence of valence and Rydberg excited states in the same energy region can potentially lead to artificial valence–Rydberg mixing in the electronic wave functions. (b) Rydberg states have a tendency to be difficult to converge. We have implemented an approach for the consistent description of both valence and Rydberg excited states within the CASSCF electronic structure model. By employing the multiconfigurational second- and third-order perturbation theory (CASPT2/3) methods based on CASSCF reference wave functions, the procedure is verified by comparison with spectroscopic results for the example molecule pyrazine. Vertical excitation energies and other properties have been calculated for various electronic states. Basis sets and active spaces were selected to provide accurate results. Two combinations of aug-cc-pVTZ level basis sets complemented by Rydberg functions

have been employed to calculate estimates for the properties of 19 singlet excited states of pyrazine. While many of the assignments made in previous studies could be confirmed, there are also several new aspects emerging from the present investigation.

**Keywords** Pyrazine · Electronic spectrum · Vertical excitations · Rydberg states · CASSCF · CASPT2 · CASPT3

## 1 Introduction

The electronic spectrum of pyrazine has been the subject of several experimental [1–5] and theoretical [6–10] studies. This molecule has been of interest for a long time not only because it provides insight into the molecular dynamics of simple aromatic systems [11–15], but also because it is a fundamental parent system for numerous biologically active compounds, such as nicotinic acid and the nucleotides cytosine, uracil, and thymine, etc. [16–18]. Innes et al. [1] published an important review article covering experimental and theoretical studies on the electronic states of azabenzene up to 1988. Extensive theoretical work on this subject has been performed by Fülischer and Roos [6, 7]. They used the complete active space self-consistent field (CASSCF) and complete active space multiconfigurational second-order perturbation theory (CASPT2 or short PT2) methods to investigate the dependence of computed valence excitation energies and transition moments on the basis sets. DelBene et al. [8] performed a series of equation of motion coupled-cluster (EOM-CC) calculations and compared the obtained results to available experimental and theoretical data. Not only for the singlet but also for the triplet states of this molecule several ab initio and density

---

Dedicated to Professor Sandor Suhai on the occasion of his 65th birthday and published as part of the Suhai Festschrift Issue.

---

C. Woywod (✉)  
Theoretical Chemistry (CTCC), Chemistry Department,  
University of Tromsø, 9037 Tromsø, Norway  
e-mail: woywod@ch.tum.de

A. Papp · Á. Vibók  
Department of Theoretical Physics, University of Debrecen,  
PO Box 5, 4010 Debrecen, Hungary  
e-mail: vibok@phys.unideb.hu

G. J. Halász  
Department of Information Technology,  
University of Debrecen, PO Box 12, 4010 Debrecen, Hungary

functional calculations were reported by Weber and Reimers [9]. Recently, theoretical results for the vertical excitation energies of pyrazine have been presented by Li et al. [10]. They have studied the ground and excited states of this molecule and compared the calculated observables to available theoretical and experimental data. There are many inconsistencies among the results obtained in these studies. For example, the ordering of the vertical excitation energies of the different states has been quite controversial, the estimated values strongly depend on the electronic structure methods and basis sets.

By using the CASSCF, PT2 and PT3 (=CASPT3) methods, one of the goals of the present work is to find suitable basis sets and active spaces not only for the description of valence excited states but also of Rydberg states. We compare the computational results to experimental data, to values obtained by Fülcher and Roos [7] and to recent calculations employing the iterative coupled cluster approach including connected triples (CC3) by Schreiber et al. [19].

One perspective of the present work is to use the experience obtained on this system to later extend our studies in order to calculate accurate potential energy surfaces for those states, to localize conical intersections between different electronic states and to eventually perform dynamical calculations.

This paper is organized as follows: the computational details of the calculations are explained in Sect. 2. In Sect. 3, the results are presented and discussed. The summary and conclusions are given in Sect. 4.

## 2 Methods and numerical details

### 2.1 Information on programs and basis sets

The calculations were carried out partly at the Jülich Supercomputing Centre on an IBM p6 575 Cluster JUMP and partly at the Leibniz Rechenzentrum of the Bavarian Academy of Sciences on a Linux Cluster. The MRCI (only used for the evaluation of transition dipole moments), CASSCF, PT2 and PT3 calculations were performed with the MOLPRO program package [20].

Two different combinations of atomic basis sets have been employed for this study. Since  $\pi\pi^*$  and  $n\pi^*$  valence states of unsaturated systems are known to be significantly more diffuse than for example the ground state, in particular along the axis perpendicular to the molecular plane [21–23], Dunning's augmented correlation consistent aug-cc-pVTZ (AVTZ) basis set has been selected as the key element of the atomic orbitals for the electronic structure calculations reported in this study [24, 25]. More precisely, the smaller of the two basis sets, referred to as basis 1 in the

following, is composed of AVTZ spd functions on the C and N atoms as well as of cc-pVTZ sp functions on H. In this case, the AVTZ and cc-pVTZ basis sets as implemented in MOLPRO have been applied [24]. In addition, four and three sets of molecule centered diffuse s and p,d functions, respectively, have been included with basis 1 for a better description of Rydberg states, leading to a total of 256 contracted basis functions.

The larger basis set (basis 2) is assembled from AVTZ spdf functions on C and N plus AVTZ spd orbitals on H. This version of the AVTZ basis is supplemented by one set of diffuse s,p and d functions, again located at the center of the molecule. In total, 377 contractions are resulting for basis 2.

The equilibrium geometry of the ground state ( $S_0$  ( $^1A_g$ )) of pyrazine has been determined at the PT2 level of theory for both basis sets 1 and 2. The optimized structures are employed as reference geometries for the calculation of excited states.

### 2.2 General strategy for selection of active orbitals

The quality of the results strongly depends on the specification of the active orbitals that are selected for the description of the various electronic states. The active space configurations are obtained by forming all spin- and symmetry-allowed excitations involving  $n$  electrons distributed over  $m$  orbitals. We will identify an active space by the pair  $(n, m)$ .

For the estimation of properties of a set of electronic states like excitation energies, dipole moments, transition dipole moments and second moments of the charge distribution ( $\langle x^2 \rangle$ , etc.) at a certain nuclear geometry two different strategies can be pursued: (i) an approach ensuring maximum consistency or (ii) a state specific scheme, i.e., the concept of tailoring the electronic structure model to best describe each individual state.

In a type (i) scenario, the goal would be to reach an optimized balance between the approximations of the various wave functions. If the CASSCF method is employed, then this will imply that the same active space is defined for the evaluation of all electronic states of interest.

In order to describe a set of electronic states exactly on an equal footing, each root should moreover be included with identical weight into the CASSCF energy functional. In this way, not only the active space but also the complete set of molecular orbitals would be common to all roots and differences between the states would be localized exclusively in the CI part of the electronic wave functions.

In principle, there are no restrictions to the number and symmetries of states that can be described by such a CASSCF state-averaging scheme. However, in practice two major difficulties frequently arise when the

optimization of several roots in a single CASSCF calculation is attempted. The first problem is that convergence may not be reached. This could possibly be resolved by reducing or increasing the number of states and by modifying the weighting pattern. The other downside is a potentially inaccurate approximation of the individual wave functions. Even if the active space may be appropriate for the description of all selected roots, the condition of a single one-electron basis shared by all states significantly limits the variational flexibility.

An alternative to the state-averaging ansatz that would still provide a high consistency and at the same time would imply a higher accuracy of individual states is to relax the constraint of a common set of orbitals and to instead converge each root separately, while continuing to employ the same comprehensive active space for the state-specific optimizations.

In this context, it is important to refine the terminology of an active space and of active orbitals. In a strict sense, a truly common set of active orbitals can only be realized by simultaneous optimization of several roots. This notion of a common set of active orbitals is unambiguous, but the idea of a common active space for a combination of CASSCF calculations of selected electronic states is much less clear-cut. If the inactive, active and external orbitals are allowed to vary freely in the CASSCF iterations for each state, then obviously individually adapted sets of orbitals are computed. The perception of a common active space therefore refers to the view that orbitals with certain physical characteristics, e.g., of valence  $\pi$  nature, can be identified in each of the active orbital sets resulting from the different wave functions optimizations.

The term common active space thus can be useful, for example, if the computation of a choice of valence  $\pi\pi^*$  states of a polyene is addressed, since unique valence  $\pi$  and  $\pi^*$  orbitals that constitute the active space should be straightforwardly distinguishable even if each wave function has been converged individually. However, there are also situations when the classification of a certain orbital, say as a  $\pi$  orbital, will become difficult if not impossible as a result of orbital variations performed during wave function optimization. In such a case, only a detailed analysis of the active orbitals produced by the CASSCF calculations for each of the electronic states of interest can reveal if is justified to invoke the characterization of an active space as common or if this expression would be misleading or actually incorrect.

This deliberation indicates that it is not always possible to achieve a highly consistent description of an assembly of electronic states by definition of a common active space and by separately converging the different wave functions. In fact, the present task of a balanced approximation of valence and Rydberg excitations of pyrazine represents such a case.

Independent of the question if a maximum consistency scheme, that is, performance of one single CASSCF calculation encompassing all states with equal weights, or a downgraded version of this approach that would relax the constraint of a common set of active orbitals and would instead rely on the less stringent model of a common active space, is applied, one problematic feature would be inherent to both strategies.

As a consequence of the presence of valence as well as Rydberg transitions in the section of the electronic spectrum of pyrazine under consideration, only a part of the orbitals included in the active space for the representation of a given wave function will actually be required from a physical perspective, since the nature of the electronic excitations varies from state to state. This increased computational effort resulting from a maximum consistency scheme (i) or also from related calculations with the relaxed condition of a common active space would not constitute a major obstacle for CASSCF calculations on pyrazine as the system is sufficiently small.

However, there is another more serious issue that makes the definition of a common active space for the relevant states of pyrazine very difficult: the necessary integration of both valence and Rydberg orbitals in the active space. While the valence orbitals that are included in the active space form a stable block and do not show any tendency for a rotation into the inactive or external orbital subspaces, this is not true for the Rydberg functions.

In fact, we observe a strong affinity for a replacement of diffuse active orbitals by more compact ones during the CASSCF iterations. A given Rydberg function can only be stabilized in the active space if configuration state functions (CSFs) that correspond to an excitation of electrons into this particular orbital are substantially contributing to the CI expansion of the electronic state.

Because this property of diffuse orbitals makes the definition of a consistent active space for a representation of both valence and Rydberg states impossible if the MO coefficients are allowed to vary as in a CASSCF calculation, we are resorting to the second best option regarding consistency and add only those diffuse functions that are critical for the formation of a certain Rydberg state to the valence set of active orbitals in the sense of the state-specific proposal (ii).

More precisely, the computational procedure implemented for the present study can be interpreted as a compromise between the extreme goals of achieving the best balance between all states [guideline (i)] on one side and the purely state specific focus of adapting the computational approach to the individual states [guideline (ii)] on the other. This compromise is realized by definition of an essential block of active orbitals that is specifically supplemented by respective key orbitals for representatives of

a subset of states, namely for the Rydberg excitations. In this way, a maximum intersection of the active spaces employed for the individual wave function optimizations can be reached.

### 2.3 Screening procedure for identification of key excitations

A collection of valence orbitals is involved in excitations that are of relevance for practically all wave functions of interest in the context of this study and is therefore referred to as core active space. This set is formed by two non-bonding orbitals on the nitrogen atoms as well as by three  $\pi$  and three  $\pi^*$  valence orbitals.

In addition to transitions within this valence domain, we are also considering 3s and 3p Rydberg states. Consequently, 8 valence and 4 Rydberg-type orbitals (diffuse s,  $p_x$ ,  $p_y$  and  $p_z$  functions) are representing the pool for the definition of the active space. We will focus on the two lowest states in each of the eight irreducible representations of the  $D_{2h}$  point group. In addition, the  $3^1A_g$ ,  $4^1A_g$ ,  $3^1B_{2u}$  and  $3^1B_{1u}$  states are calculated.

After the specification of the nature of the target electronic wave functions, in the next step, information on the characteristic excitations which are contributing to a particular electronic state must be provided. For this purpose, a so-called screening procedure is performed. All calculations described in the following are carried out at the reference geometries obtained with basis sets 1 and 2, respectively. Initially, for each state a CASSCF wave function  $\psi_{\text{core}}^{\text{CASSCF}}$  defined by the core set of active orbitals is optimized:

$$\psi_{\text{core}}^{\text{CASSCF}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{el}}) = \sum_{k=1}^{N_{\text{core}}} C_k^{\text{core}} \phi_k(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{el}}), \quad (1)$$

where

$$\phi_k(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{el}}) = (\chi_1(\mathbf{r}_1), \chi_2(\mathbf{r}_2), \dots, \chi_{N_{el}}(\mathbf{r}_{N_{el}})). \quad (2)$$

The functions, variables and parameters in Eqs. 1 and 2 are specified as follows:  $\mathbf{r}_n$  is the coordinate vector of electron  $n$ ,  $\phi_k$  denotes a configuration state function (CSF) composed of a set of  $N_{el}$  occupied CASSCF pseudocanonical or natural spin orbitals  $\chi_n(\mathbf{r}_n)$  and  $N_{\text{core}}$  corresponds to the length of the CASSCF CI expansion with coefficients  $C_k^{\text{core}}$ .

$\psi_{\text{core}}^{\text{CASSCF}}$  is preferably obtained by selectively optimizing the desired root and if this is not successful, then a state averaging scheme is employed with the largest possible weight assigned to the state of interest.

The calculation of  $\psi_{\text{core}}^{\text{CASSCF}}$  provides the set of orbitals  $\chi_n(\mathbf{r}_n)$  for the actual screening step: the core active space is

extended by the four Rydberg functions to form a complete set of active orbitals and a full CI calculation within this space of 12 orbitals is conducted for each electronic state, yielding:

$$\psi_{\text{complete}}^{\text{CI}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{el}}) = \sum_{k=1}^{N_{\text{complete}}} C_k^{\text{complete}} \phi_k(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{el}}), \quad (3)$$

with  $N_{\text{complete}}$  and  $C_k^{\text{complete}}$  being the number and coefficients of the terms in the full CI expansion, respectively. Since the CI method implies that the orbitals are not allowed to vary, it is ensured that the Rydberg functions remain included in the configuration space during the CI iterations and are not rotated out of the active space. The roots are optimized individually throughout in the CI calculations.

Subsequently, the resulting wave functions  $\psi_{\text{complete}}^{\text{CI}}$  are analyzed to determine the most important excitations for a given electronic state. Based on this information, the active orbitals required for the description of a certain electronic state can be determined and the screening process accomplished.

It should be mentioned that orbitals  $\chi_n(\mathbf{r}_n)$  derived from a CASSCF optimization with core active space for the  $m$ th root represent just one possible choice of an one-electron basis in the screening CI calculation for this particular root. Alternatively, the CASSCF orbitals of the electronic ground state could for example be employed for the screening CI calculations of all excited states throughout. In fact, many different orbital sets could be used in order to produce information on the nature of the excitations that are characteristic for a certain electronic state. The selection of orbitals for the screening procedure in this study is based on the assumption that the  $\chi_n(\mathbf{r}_n)$  are well suited as starting orbitals for the following steps since they have already been pre-optimized for a certain root.

### 2.4 CASSCF optimization of the Rydberg states

Because of the often high sensitivity of the CASSCF convergence in particular for Rydberg states to the quality of the starting guess, a second full CI step is joined up in circuit. The reason for this is that while the orbitals  $\chi_n(\mathbf{r}_n)$  resulting from the valence-space CASSCF calculation of  $\psi_{\text{core}}^{\text{CASSCF}}$  (Eq. 1) will usually serve well for the definition of the one-electron part of the initial state vector that will be supplied for the CASSCF calculation of a Rydberg state, neither the coefficients  $C_k^{\text{core}}$  and  $C_k^{\text{complete}}$  can be directly employed in the CI part of the starting wave function, since a specially tailored active space is required for the description of a given Rydberg excitation. Therefore, a second CI calculation is performed by using again the

orbitals  $\chi_n(\mathbf{r}_n)$  but a CI expansion that now corresponds to the active space adapted to a particular Rydberg state. Individual optimization of the target root leads to:

$$\psi_{\text{Ryd}}^{\text{CI}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{\text{el}}}) = \sum_{k=1}^{N_{\text{Ryd}}} C_k^{\text{Ryd}} \phi_k(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{\text{el}}}), \quad (4)$$

where  $C_k^{\text{Ryd}}$  and  $N_{\text{Ryd}}$  are the coefficients and length of the CI expansion, respectively.

$\psi_{\text{Ryd}}^{\text{CI}}$  can be considered as a suitable starting wave function for the final step, the CASSCF calculation of the Rydberg state:

$$\psi_{\text{Ryd}}^{\text{CASSCF}}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{\text{el}}}) = \sum_{k=1}^{N_{\text{Ryd}}} \tilde{C}_k^{\text{Ryd}} \tilde{\phi}_k(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{\text{el}}}), \quad (5)$$

where

$$\tilde{\phi}_k(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_{\text{el}}}) = (\tilde{\chi}_1(\mathbf{r}_1), \tilde{\chi}_2(\mathbf{r}_2), \dots, \tilde{\chi}_{N_{\text{el}}}(\mathbf{r}_{N_{\text{el}}})) \quad (6)$$

The CASSCF optimization of the Rydberg excitation provides the set of expansion coefficients  $\tilde{C}_k^{\text{Ryd}}$  for the CSFs  $\tilde{\phi}_k$ , which are represented in terms of the new set of orbitals  $\tilde{\chi}_n(\mathbf{r}_n)$ . Again, other roots are only included in the CASSCF energy functional if this is necessary to converge  $\psi_{\text{Ryd}}^{\text{CASSCF}}$ .

Generally, the frequently observed difficulty to calculate Rydberg excitations with the CASSCF method is related to the topology of the potential energy surface for a given CI root that corresponds to a Rydberg solution in the space spanned by CI and MO variables. The problem is that the energy minima of Rydberg states are frequently protected by energy barriers. Even though the Rydberg minimum may be pronounced and possibly the global one on the energy surface, it is not easy to reach it.

There are two main screws that can be adjusted in order to direct the CASSCF optimization towards the Rydberg minimum: (i) It is necessary to provide a good starting vector, e.g., a wave function produced by a pilot CI calculation as described in this section. (ii) If the state cannot be described individually, then a suitable state averaging scheme can often help to achieve convergence. For example, the presence of a valence and a Rydberg solution of similar energy may be treated by simultaneous optimization of both roots in the CASSCF calculation.

## 2.5 PT2 and PT3 calculations

The CASSCF wave functions  $\psi_{\text{core}}^{\text{CASSCF}}$  and  $\psi_{\text{Ryd}}^{\text{CASSCF}}$  are then employed as reference functions for PT2 and PT3 calculations of valence and Rydberg states, respectively, in order to take dynamic electron correlation into account. The PT2 and PT3 versions of multiconfigurational

perturbation theory implemented in MOLPRO have been selected [20]. The g0 Fock operator is chosen in the zeroth-order Hamiltonian [26]. In general, the smallest level shift required for convergence of the PT2 iterations for a given state has been applied.

If the CASSCF reference function has been obtained by individual optimization of the  $m$ th root, then single-state PT2 has been applied. In the case of a simultaneous optimization of two or more roots in the CASSCF calculation the multi-state PT2 version with mixed treatment of the reference states has been employed.

No core orbitals have been defined in the calculations with basis set 2. In the case of basis set 1, the 1s orbitals on carbon and nitrogen have been frozen in the PT2 and PT3 calculations. It has been verified that the effect of this approximation on the electronic properties of interest is negligible.

## 3 Results and discussion

### 3.1 Nature of the electronic states

The vertical excitation energies and oscillator strengths are calculated at the reference geometries for each basis set. In Table 1, the two geometries are compared to the experimental values. The bond lengths and bond angle parameters illustrate that both theoretical ground state equilibrium structures are of satisfactory quality.

The nature of the various electronic states treated in this study will be discussed next. The lowest excited singlet state ( $1^1B_{3u}$ ) is of  $n\pi^*$  character. The most important contribution to this electronic state is the ( $6a_g \rightarrow 2b_{3u}$ ) excitation. According to the energetical sequence obtained at the PT2 level of theory (Table 2), the second ( $1^1A_u$ ) and third ( $1^1B_{2u}$ ) excited singlet states at the ground state equilibrium geometry are of singly  $\pi\pi^*$  excited character. The PT3 and CC3 methods predict an inverse ordering of the  $1^1A_u$  and  $1^1B_{2u}$  states.

The two most important configurations in the CI expansion for the bright  $1^1B_{2u}$  state are the ( $1b_{1g} \rightarrow 2b_{3u}$ ) and ( $1b_{2g} \rightarrow 1a_u$ ) excitations.

**Table 1** Ground-state equilibrium geometry of pyrazine: comparison of PT2 results with experiment [1]

|             | CC (Å)   | CN (Å)   | CH (Å)   | ∠ NCC    | ∠ CNC    | ∠ NCH     |
|-------------|----------|----------|----------|----------|----------|-----------|
| Basis set 2 | 1.395    | 1.339    | 1.081    | 122.2    | 115.6    | 117.0     |
| Basis set 1 | 1.398    | 1.343    | 1.081    | 122.4    | 115.2    | 116.9     |
| Expt.       | 1.403(4) | 1.339(2) | 1.115(4) | 122.2(4) | 115.6(4) | 113.9(10) |

Other valence states of interest are:  $2^1A_g$  (double  $\sigma\pi^*$  and single  $\pi\pi^*$ ),  $3^1A_g$  (two important single  $\pi\pi^*$  and two important double  $\pi\pi^*$  excitations),  $4^1A_g$  (three important double  $\pi\pi^*$  excitations),  $1^1B_{1g}$  (single  $\pi\pi^*$ ),  $1^1B_{1u}$  (single  $\pi\pi^*$ ),  $2^1B_{1u}$  (single  $\pi\pi^*$ ),  $2^1B_{2u}$  (single  $\pi\pi^*$ ),  $1^1B_{2g}$  (single  $\sigma\pi^*$  and  $n\pi^*$ ),  $2^1B_{2g}$  (single  $n\pi^*$  and combination of two single  $n\pi^*/\sigma\pi^*$ ),  $1^1B_{3g}$  (three important single  $\pi\pi^*$  excitations) and  $2^1B_{3g}$  (single  $\pi\pi^*$  and two combined single  $\pi\pi^*$  excitations).

The second moments of the charge distribution obtained with basis sets 1 and 2 for the valence and Rydberg states can be found in Table 3. It can be seen that the absolute values of the second moments of the valence states do not exceed  $33.6 a_0^2$ .

Since various definitions of the second and higher moments of the molecular charge distribution can be found in the literature [29], we want to precisely define the quantities compiled in Table 3. The electronic structure models applied in the framework of this study are restricted to the solution of the electronic Schrödinger equation and the nuclei are approximated as fixed point charges  $Q_k$ . We are therefore concerned with a discrete and a continuous charge distribution arising from the nuclei and from the electrons, respectively. The values given in Table 3 correspond to the sum of the nuclear ( $\sum_{k=1}^{N_{nuc}} Q_k x_k^2$  etc.) and electronic contributions ( $-\int \rho(\mathbf{r})x^2 d\tau$ , where  $\rho(\mathbf{r})d\tau$  corresponds to the change of the charge density in volume element  $d\tau$ ). For the 14 valence states, the electronic contributions to the second moments can accordingly be written as  $\langle \psi_{core}^{CASSCF} | x^2 | \psi_{core}^{CASSCF} \rangle$ , etc.

Several Rydberg excitations are included in the scope of this study. The  $2^1B_{3u}$  state corresponds to a single ( $n \mapsto 3p_x$ ) transition, assuming that the molecule is placed in the  $yz$  plane with the nitrogen atoms located along the  $z$  axis. The  $3^1B_{2u}$  and  $3^1B_{1u}$  wave functions can be described as single ( $\pi \mapsto 3p_x$ ) excitations. The promotion of an electron into the diffuse  $3p_x$  orbital common to the  $2^1B_{3u}$ ,  $3^1B_{2u}$  and  $3^1B_{1u}$  states is reflected in large second moments  $\langle \psi_{Ryd}^{CASSCF} | x^2 | \psi_{Ryd}^{CASSCF} \rangle$  (Table 3).

The  $2^1B_{1g}$  and  $2^1A_u$  states represent a single ( $\pi \mapsto 3s$ ) and two combined single ( $\pi \mapsto 2s$ ,  $\pi \mapsto 3p_z$ ) transitions, respectively. Accordingly, the second moments are large but relatively isotropic for the  $2^1B_{1g}$  state while the  $2^1A_u$  state is very diffuse along the  $z$  axis (Table 3).

### 3.2 Computational details

The CASSCF, PT2 and PT3 vertical excitation energies  $T_v$  obtained with basis sets 1 and 2 for the 19 excited states of pyrazine are presented in Table 2. The energies calculated in the framework of this study are compared to experimental data for  $T_v$  and also for the adiabatic transition

energies  $T_{00}$ . Included are also CASSCF/PT2 results published by Fülischer and Roos [7] as well as recent CC3 data and so-called “best estimate” values derived from ab initio calculations with large basis sets [19]. In addition, MRCI  $T_v$  values for the  $1^1B_{3u}$  and  $1^1B_{2u}$  states are given [27].

At this point, a short comment on the basis sets and ground state equilibrium geometries employed by Refs. [7, 19, 27] is in order. For the calculations described in Refs. [7, 19], atomic natural orbital and TZVP [30] basis sets, respectively, have been applied, both comparable in size to the AVTZ basis set. The MRCI computations published in Ref. [27] have been carried out with DZP basis functions.

While Ref. [7] uses a structure determined by experiment, Refs. [19, 27] are relying on geometries optimized at the MP2/DZP level of theory.

The lack of Rydberg-type basis functions implies that only valence states were considered by Refs. [7, 19, 27]. It is also worth noting, that the  $1^1A_g$  reference structures are not entirely consistent with the methods utilized for the evaluation of excitation energies according to the three articles. The effects of such incoherences on the values of  $T_v$  can be considerable.

The results under subcolumns (a) and (b) have been derived by performing CASSCF, PT2 and PT3 calculations, that is, by using the fully orbital optimized wave functions  $\psi_{core}^{CASSCF}$  ((10/8) active space) and  $\psi_{Ryd}^{CASSCF}$  ((10/9) active space) for valence and Rydberg states, respectively.

In order to obtain the  $T_v$  values collected in subcolumns (c), Fülischer and Roos [7] defined two active orbital sets, tailored for the calculation of  $\pi\pi^*$  and  $n\pi^*$  valence states, respectively. Both orbital selections described in Ref. [7] differ from the (10/8) active space adopted for the representation of valence states in the present work and also in the CASSCF/MRCI calculations outlined in Ref. [27].

When possible, the wave functions  $\psi_{core}^{CASSCF}$  and  $\psi_{Ryd}^{CASSCF}$  have been optimized individually with both basis sets. Exceptions are the  $2^1B_{3u}$  [optimized together with  $1^1B_{3u}$  state, weights 0.05 ( $1^1B_{3u}$ ), 0.95 ( $2^1B_{3u}$ )], the  $2^1B_{1g}$  [weights 0.307 ( $1^1B_{1g}$ ), 0.693 ( $2^1B_{1g}$ ) with basis 1 and 0.70 ( $1^1B_{1g}$ ), 0.30 ( $2^1B_{1g}$ ) with basis 2],  $3^1B_{1u}$  (weights 0.30 ( $1^1B_{1u}$ ), 0.70 ( $2^1B_{1u}$ )),  $2^1B_{3g}$  [weights 0.20 ( $1^1B_{3g}$ ), 0.80 ( $2^1B_{3g}$ )]; state averaging only applied in case of basis 1), and  $2^1A_u$  (weights 0.10 ( $1^1A_u$ ), 0.90 ( $2^1A_u$ )) wave functions.

In the PT2 calculations, the following level shifts had to be set in order to reach convergence:  $1^1A_g$  (basis 2: 0.0; basis 1: 0.0),  $2^1A_g$  (0.5; 0.5),  $3^1A_g$  (0.5; 0.5),  $4^1A_g$  (0.7; 0.7),  $1^1B_{3u}$  (0.3; 0.3),  $2^1B_{3u}$  (0.3; 0.3),  $1^1B_{2u}$  (0.3; 0.3),  $2^1B_{2u}$  (0.3; 0.2),  $3^1B_{2u}$  (0.3; 0.35),  $1^1B_{1g}$  (0.3; 0.3),  $2^1B_{1g}$  (0.5; 0.5),  $1^1B_{1u}$  (0.0; 0.0),  $2^1B_{1u}$  (0.3; 0.2),  $3^1B_{1u}$  (0.3; 0.3),  $1^1B_{2g}$  (0.3; 0.3),  $2^1B_{2g}$  (0.5; 0.5),  $1^1B_{3g}$  (0.5; 0.5),  $2^1B_{3g}$  (0.3; 0.4),  $1^1A_u$  (0.3; 0.3),  $2^1A_u$  (0.3; 0.3).

### 3.3 Analysis of the consistency of CASSCF, PT2 and PT3 results

The agreement between the CASSCF results shown under subcolumns (a), (b) and (c) in Table 2 is generally quite good, the largest difference of 0.58 eV is found between the CASSCF energies of the  $2^1B_{1u}$  valence state listed in subcolumns (a) and (c). Typically, the variation of the three CASSCF- $T_v$  entries is much smaller for a given electronic state.

This indicates that the identification of different active orbitals made in Ref. [7] and in the present study has no significant effect at the CASSCF level of theory. Application of the PT2 method, however, leads to a pronounced differentiation between the three combinations of active orbitals and basis sets under consideration. Deviations of energies given in subcolumns (a) and (b) are exclusively due to the utilization of basis sets 2 and 1, respectively.

For most states, the agreement of the PT2 estimates obtained with basis 1 and 2 is excellent and within 0.1 eV, but there are four notable exceptions from this rule. Since the main difference between the two basis sets is in the Rydberg

part, it is not too surprising that a substantial disagreement of 1.19 eV is obtained for  $T_v$  of the  $2^1B_{1g}$  Rydberg state.

This difference of  $T_v(2^1B_{1g})$  obtained with basis sets 1 and 2 is only slightly reduced to 0.80 eV at the PT3 level of theory. It is interesting to note that the PT3 predictions of  $T_v$  calculated with both basis sets coincide even better on average than the PT2 values and agree to within 0.1 eV, the only runaway case besides  $T_v(2^1B_{1g})$  is  $T_v(1^1B_{3g})$ .

However, the discrepancies of the PT2 estimates for  $T_v(2^1B_{1u})$  (0.24 eV),  $T_v(2^1B_{2u})$  (0.44 eV) and in particular for  $T_v(1^1B_{3g})$  (1.04 eV) resulting with basis 1 and 2 come unexpected. Since all three states are consistently described as valence excitations by CASSCF calculations with both basis sets (Table 3), it appears that details of the approximations to the compact wave functions are critical in those cases. The different diffuse components of both basis sets are not likely to be responsible for the inconsistent predictions of  $T_v$  by the PT2 method.

A comparison to the experimental data shows that the approach PT2 / basis set 2 yields values for  $T_v(2^1B_{1u})$  and  $T_v(2^1B_{2u})$  that are closer to experiment than those calculated with basis set 1. This makes sense since basis 2 is

**Table 2** Vertical excitation energies ( $T_v$ ) (in eV), for 19 excited singlet states of pyrazine

|                       | $T_v(\text{CASSCF})$ |       |      | $T_v(\text{PT2})$ |       |      | $T_v(\text{PT3})$ |       | $T_v(\text{MRCI})$<br>[27] | $T_v(\text{CC3})$<br>[19] | $T_v(\text{BE})$<br>[19] | $T_v(\text{exp})$ | $T_{00}(\text{exp})$ | $T_{00}(\text{exp})$<br>[1] |
|-----------------------|----------------------|-------|------|-------------------|-------|------|-------------------|-------|----------------------------|---------------------------|--------------------------|-------------------|----------------------|-----------------------------|
|                       | (a)                  | (b)   | (c)  | (a)               | (b)   | (c)  | (a)               | (b)   |                            |                           |                          |                   |                      |                             |
| $2^1A_g(\text{V})$    | 8.48                 | 8.39  | 8.40 | 8.69              | 8.64  | 8.22 | 8.92              | 8.90  |                            | 8.69                      |                          |                   |                      | (6.30)                      |
| $3^1A_g(\text{V})$    | 11.74                | 11.66 |      | 10.00             | 9.97  |      | 10.77             | 10.76 |                            |                           |                          |                   |                      |                             |
| $4^1A_g(\text{V})$    | 12.28                | 12.18 |      | 11.81             | 11.75 |      | 12.40             | 12.37 |                            |                           |                          |                   |                      |                             |
| $1^1B_{3u}(\text{V})$ | 4.84                 | 4.87  | 5.19 | 3.86              | 3.93  | 3.85 | 4.53              | 4.60  | 4.22                       | 4.24                      | 3.95                     |                   | 3.83 [4]             | 3.83                        |
| $2^1B_{3u}(\text{R})$ | 7.22                 | 7.28  |      | 7.24              | 7.19  |      | 7.57              | 7.58  |                            |                           |                          |                   |                      | (6.75)                      |
| $1^1B_{2u}(\text{V})$ | 5.02                 | 4.96  | 5.10 | 4.81              | 4.79  | 4.76 | 5.25              | 5.20  | 5.05                       | 5.02                      | 4.64                     | 4.81 [4]          | 4.69 [4]             | 4.69–4.81                   |
| $2^1B_{2u}(\text{V})$ | 9.86                 | 9.83  | 9.55 | 7.06              | 6.62  | 7.74 | 7.95              | 7.85  |                            | 8.05                      | 7.60                     | 7.67 [4]          | 7.53 [10]            |                             |
| $3^1B_{2u}(\text{R})$ | 7.24                 | 7.22  |      | 8.07              | 8.03  |      | 7.94              | 7.95  |                            |                           |                          |                   |                      |                             |
| $1^1B_{1g}(\text{V})$ | 7.19                 | 7.23  |      | 6.30              | 6.36  |      | 6.93              | 6.97  |                            | 6.75                      | 6.60                     | 6.10 [10]         |                      |                             |
| $2^1B_{1g}(\text{R})$ | 7.44                 | 7.37  |      | 7.58              | 6.39  |      | 7.70              | 6.90  |                            |                           |                          | 7.13 [28]         | 7.12 [28]            |                             |
| $1^1B_{1u}(\text{V})$ | 8.35                 | 8.35  | 8.51 | 6.37              | 6.39  | 6.69 | 7.04              | 7.03  |                            | 7.07                      | 6.58                     | 6.51 [4]          | 6.31 [4]             | 6.30                        |
| $2^1B_{1u}(\text{V})$ | 10.23                | 10.21 | 9.65 | 7.12              | 6.88  | 7.53 | 8.02              | 7.97  |                            | 8.06                      | 7.72                     | 7.67 [4]          | 7.53 [10]            |                             |
| $3^1B_{1u}(\text{R})$ | 9.05                 | 9.03  |      | 9.18              | 9.07  |      | 9.25              | 9.18  |                            |                           |                          |                   |                      |                             |
| $1^1B_{2g}(\text{V})$ | 5.85                 | 5.87  |      | 5.48              | 5.54  |      | 5.92              | 5.97  |                            | 5.74                      | 5.56                     | 6.10 [4]          |                      | 5.45                        |
| $2^1B_{2g}(\text{V})$ | 9.12                 | 9.12  |      | 8.08              | 8.13  |      | 8.68              | 8.74  |                            |                           |                          |                   |                      |                             |
| $1^1B_{3g}(\text{V})$ | 11.24                | 11.19 |      | 6.70              | 7.74  |      | 9.05              | 9.21  |                            | 8.77                      |                          |                   |                      |                             |
| $2^1B_{3g}(\text{V})$ | 8.38                 | 8.30  | 8.53 | 8.61              | 8.56  | 8.31 | 8.97              | 8.91  |                            |                           |                          |                   |                      |                             |
| $1^1A_u(\text{V})$    | 5.95                 | 6.01  | 6.29 | 4.52              | 4.58  |      | 5.32              | 5.37  |                            | 5.05                      | 4.81                     | 4.72 [10]         |                      |                             |
| $2^1A_u(\text{R})$    | 7.47                 | 7.45  |      | 8.22              | 8.10  |      | 8.13              | 8.07  |                            |                           |                          |                   |                      |                             |

Subcolumns labeled by (a) and (b) present the CASSCF/PT2/PT3 results calculated with basis sets 2 and 1, respectively, in combination with active spaces (10/8) for valence (V) and (10/9) for Rydberg states (R). The data in subcolumns designated by (c) are taken from Table 7 in Ref. [7] and correspond to CASSCF/PT2 calculations with a (4s3p2d/3s2p) ANO basis set. For selected valence states, CC3 estimates of  $T_v$  and so-called “best estimate” values are given [19]. In addition, values for  $T_v$  of the  $1^1B_{3u}$  and  $1^1B_{2u}$  states obtained by MRCI/DZP calculations [27] are provided. As far as available, experimental information on  $T_v$  and also on the adiabatic transition energies ( $T_{00}$ ) of the various excited states is included as a reference

**Table 3** The sum of the electronic and nuclear contributions to the second moments of the charge distribution (see text for details), given in units of  $a_0^2$ , for 20 singlet states of pyrazine

|                | Basis set 2            |                        |                        | Basis set 1            |                        |                        |
|----------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
|                | $-\langle x^2 \rangle$ | $-\langle y^2 \rangle$ | $-\langle z^2 \rangle$ | $-\langle x^2 \rangle$ | $-\langle y^2 \rangle$ | $-\langle z^2 \rangle$ |
| $1^1A_g(V)$    | 26.9                   | 20.0                   | 30.9                   | 27.0                   | 20.0                   | 30.9                   |
| $2^1A_g(V)$    | 26.8                   | 21.5                   | 30.8                   | 26.9                   | 21.5                   | 30.8                   |
| $3^1A_g(V)$    | 27.2                   | 19.9                   | 33.0                   | 27.3                   | 19.8                   | 33.0                   |
| $4^1A_g(V)$    | 27.2                   | 21.2                   | 32.0                   | 27.2                   | 21.2                   | 32.0                   |
| $1^1B_{3u}(V)$ | 28.1                   | 20.7                   | 27.8                   | 28.2                   | 20.7                   | 27.8                   |
| $2^1B_{3u}(R)$ | 85.5                   | 32.9                   | 41.0                   | 86.2                   | 33.3                   | 41.2                   |
| $1^1B_{2u}(V)$ | 26.9                   | 20.2                   | 31.5                   | 26.9                   | 20.2                   | 31.5                   |
| $2^1B_{2u}(V)$ | 28.4                   | 21.7                   | 31.0                   | 28.5                   | 21.7                   | 31.0                   |
| $3^1B_{2u}(R)$ | 82.3                   | 30.8                   | 44.7                   | 83.3                   | 31.3                   | 45.0                   |
| $1^1B_{1g}(V)$ | 28.2                   | 22.8                   | 26.9                   | 28.2                   | 22.7                   | 26.9                   |
| $2^1B_{1g}(R)$ | 50.1                   | 38.6                   | 44.1                   | 49.8                   | 38.8                   | 43.9                   |
| $1^1B_{1u}(V)$ | 28.9                   | 21.1                   | 32.7                   | 29.0                   | 21.1                   | 32.8                   |
| $2^1B_{1u}(V)$ | 27.8                   | 20.9                   | 31.5                   | 27.9                   | 20.9                   | 31.6                   |
| $3^1B_{1u}(R)$ | 76.0                   | 31.8                   | 42.0                   | 75.5                   | 31.7                   | 41.9                   |
| $1^1B_{2g}(V)$ | 28.0                   | 21.9                   | 27.1                   | 28.1                   | 21.8                   | 27.1                   |
| $2^1B_{2g}(V)$ | 28.0                   | 21.6                   | 28.0                   | 28.1                   | 21.5                   | 27.9                   |
| $1^1B_{3g}(V)$ | 29.7                   | 21.3                   | 33.5                   | 29.9                   | 21.3                   | 33.6                   |
| $2^1B_{3g}(V)$ | 26.6                   | 20.6                   | 32.3                   | 26.6                   | 20.5                   | 32.3                   |
| $1^1A_u(V)$    | 28.3                   | 21.7                   | 27.4                   | 28.4                   | 21.7                   | 27.4                   |
| $2^1A_u(R)$    | 47.5                   | 38.9                   | 92.2                   | 48.1                   | 38.9                   | 95.5                   |

The expectation values have been obtained by the CASSCF method with basis sets 1 and 2 in combination with active spaces (10/8)/(10/9)

composed of a more extensive version of the AVTZ basis (cf. Sect. 2.1), providing a higher variational flexibility for modelling the compact wave functions.

The agreement of the PT3 energies obtained with both basis sets is significantly better, the differences are 0.05 eV ( $T_v(2^1B_{1u})$ ) and 0.10 eV ( $T_v(2^1B_{2u})$ ), and is in line with the observation that the  $T_v$  values computed by the PT3 method are generally less dependent on the choice of the basis set.

The large difference between  $T_v(1^1B_{3g})$  obtained by switching from basis set 1 to 2 at the PT2 level of theory is puzzling. In contrast to the  $2^1B_{1u}$  and  $2^1B_{2u}$  states, in this case, the energy predicted with basis set 2 is lower. This ordering is confirmed at the PT3 level of theory, but the PT3 calculations with both basis sets are in much better agreement (within 0.16 eV).

Unfortunately, no measurements of  $T_v(1^1B_{3g})$  are available. The  $T_v$  estimates for this state evaluated by the PT3 and CC3 methods indicate that the higher PT2 energy of 7.74 eV computed with basis set 1 may be the more realistic solution. A straightforward explanation why the combination PT2 / basis set 2 ( $T_v(1^1B_{3g}) = 6.70$  eV) would have difficulties to describe the  $1^1B_{3g}$  state does not seem

to be available. It is possible that the larger set of diffuse functions included with basis set 1 would offer some advantage in approximating this particular valence excitation although this is unlikely if one takes into account that the second moments obtained with both basis sets are very similar (Table 3). This argument would also not explain why the CC3 estimate of  $T_v(1^1B_{3g})$  is closer to the value calculated with the combination PT2 / basis set 1 because no diffuse functions have been added to the basis set employed by Ref. [19].

An interesting aspect in this context is that the  $1^1B_{3g}$  wave function is distinguished from the two other ‘problematic’ valence states ( $2^1B_{1u}$  and  $2^1B_{2u}$ ) because of the multiconfigurational nature (cf. Sect. 3.1). This observation is important because it may point to possibly inaccurate CC3 excitation energies for this state. To decide which of the two PT2 results for  $T_v$  of the  $1^1B_{3g}$  state is more exact would seem to be premature given these uncertainties.

A comparison of the current PT2 results [subcolumns (a), (b)] to previous calculations (subcolumn (c), from Ref. [7]) reveals significant deviations of  $T_v$  only for the  $2^1B_{2u}$  [up to 1.12 eV between (b) and (c)] and for the  $2^1B_{1u}$  [up to 0.65 eV between (b) and (c)] states. Disagreements for  $T_v$  of a given state between subcolumns (a), (b) on one side and (c) on the other can be due to both different choices of basis sets and active orbitals. The  $2^1B_{2u}$  and  $2^1B_{1u}$  wave functions can be abstracted as single  $\pi\pi^*$  excitations (cf. Sect. 3.1). In Ref. [7], an active space including 12  $\pi$  and  $\pi^*$  orbitals is defined for  $\pi\pi^*$  excited states, that is, 6 additional non-valence  $\pi^*$  orbitals are included compared to the (10/8) valence active space employed for the present study. The differences between the basis sets utilized for obtaining the results summarized in subcolumns (a), (b) and (c) are certainly responsible for a part of the  $T_v$  variations, but since both states are rather compact all three basis sets can be considered adequate for representing the  $2^1B_{2u}$  and  $2^1B_{1u}$  wave functions. It is striking that the addition of 6 active  $\pi^*$  orbitals [subcolumn (c)] leads to a substantial increase of  $T_v$  for both states as compared to PT2 calculations with the (10/8) active space. The experimental data indicate that the PT2 estimates for  $T_v(2^1B_{2u})$  and  $T_v(2^1B_{1u})$  given in Ref. [7] are more accurate than the PT2 energies obtained for this work. Obviously, excitations into the 6 additional  $\pi^*$  orbitals that are taken into account in the CASSCF reference functions as described in Ref. [7] are playing an important role at the PT2 level of theory.

### 3.4 Comparison of PT2 and PT3 excitation energies to reference data

We will now focus on an assessment of the accuracy of the PT2 and PT3 predictions for  $T_v$  calculated in the framework of this study via comparison with experiment and



with the alternative computational results compiled in Table 2.

An inspection of Table 2 shows that PT3 and CC3 always place  $T_v$  of valence states equal to or higher than the corresponding PT2 estimates. For Rydberg states and PT3, the pattern is not uniform. In the case of the  $2^1B_{1g}$ ,  $2^1B_{3u}$  and  $3^1B_{1u}$  states, the PT3 values are equal to or higher than the corresponding PT2 values, for  $T_v(3^1B_{2u})$  the PT3 energies are lower. In the case of the  $2^1A_u$  Rydberg excitation the PT2 and PT3 results are fully consistent, varying within only 0.15 eV.

Because of the symmetry-induced exclusion of one-photon transitions almost no successful measurements of energies of the totally symmetric excited states of pyrazine are reported in the literature. The only exception is the value  $T_{00} = 6.30$  eV for a state of  $^1A_g$  symmetry obtained by two-photon absorption [1]. All calculations of  $^1A_g$  states presented in Table 2 refer to valence excitations and place  $T_v$  at 8.39 eV or higher. Since the interpretation of two-photon amplitudes is frequently afflicted with large error margins the assignment made in Ref. [1] may be questioned. Furthermore, the excited state of  $^1A_g$  symmetry identified in Ref. [1] is characterized as a ( $n \rightarrow 3s$ ) Rydberg excitation. We have therefore placed the low value for the adiabatic transition energy of the  $2^1A_g$  state in brackets in Table 2. No Rydberg states of  $^1A_g$  symmetry have been investigated in Ref. [31]. The agreement of the PT2, PT3 and CC3 estimates on the value of  $T_v(2^1A_g)$  is very good.

No computational or experimental reference data are available for  $T_v$  of the  $3^1A_g$  and  $4^1A_g$  valence states, therefore no statements can be made regarding the quality of the PT2 and PT3 computations, but we find that the results are reasonably consistent, lying within 0.80 eV ( $T_v(3^1A_g)$ ) and 0.65 eV ( $T_v(4^1A_g)$ ). The choice of basis set 1 or 2 is almost irrelevant for the PT2 and PT3 energies of both states. It should be noted that the  $T_v$  values estimated for both states are well beyond the lowest ionization potential of pyrazine (9.28 eV) [31].

The PT3, MRCI and CC3 approaches certainly overestimate  $T_v(1^1B_{3u})$ , in the case of PT3 the error is even quite large, while the PT2 values are slightly too low but closer to experiment. The “best estimate” of  $T_v = 3.95$  eV (Ref. [19]) coincides well with parameters derived from explicit simulations of the  $1^1B_{3u}$  absorption band [16, 27].

References [1, 31] give an experimental value of  $T_{00}(2^1B_{3u}) = 6.75$  eV, Ref. [32] remarks that no clear evidence for this state could be detected. Ref. [4] questions the assignments for Rydberg states made in Ref. [31]. If the interpretation of the observed “weak sharp band system” (Ref. [1]) is correct, then it is plausible to assume that the PT2 estimates for  $T_v(2^1B_{3u})$ , which are by ca. 0.4 eV lower than the corresponding PT3 guesses, are more realistic.

Nonadiabatic effects dominate the theoretically well investigated  $1^1B_{2u}$  valence band. The lower PT2 estimates for  $T_v(1^1B_{2u})$  coincide again much better with experiment than the values computed by the PT3 formalism. The MRCI and CC3 methods also overestimate  $T_v(1^1B_{2u})$  like PT3, but only slightly. In this particular case, the “best estimate” of  $T_v = 4.64$  eV given by Ref. [19] does not make any sense as it is even below  $T_{00}$  for this state.

While being disappointingly inaccurate for several other valence states, the PT3 guesses of  $T_v = 7.85$  eV (7.95 eV) and  $T_v = 7.97$  eV (8.02 eV) obtained with basis 1 (basis 2) for the very intense  $2^1B_{2u}$  and  $2^1B_{1u}$  valence bands, respectively, match better with experiment than the PT2 entries given in subcolumns (a) and (b). As mentioned in Sect. 3.3, the PT2 energies for both states improve significantly upon supplementing the active space by 6 non-valence  $\pi^*$  orbitals. One can infer that the additional amount of dynamic electron correlation taken into account in PT2 calculations by replacing the (10/8) active space with the set of 3  $\pi + 9 \pi^*$  active orbitals selected by Ref. [7] for the optimization of the CASSCF reference functions is analogous to the effect achieved by upgrading the level of theory from PT2 to PT3 while leaving the (10/8) active space invariant.

Since  $2^1B_{2u}$  and  $2^1B_{1u}$  are to a good approximation singly  $\pi\pi^*$  excited states, the CC3 method can be expected to yield very accurate estimates of  $T_v$  in both cases. The CC3 and PT3 values for  $T_v(2^1B_{2u})$  and  $T_v(2^1B_{1u})$  agree to within 0.2 and 0.1 eV, respectively. Ref. [4] deduces a common  $T_v$  value of 7.67 eV for both states from experiment and thus places them ca. 0.2–0.4 eV below the CC3 and PT3 predictions. The “best estimates” of  $T_v(2^1B_{2u})$  and  $T_v(2^1B_{1u})$  provided by Ref. [19] lie within 0.07 eV of the values given by Ref. [4].

We note that it is frequently very difficult to determine  $T_v$  values from experiment since information on the shape of the potential energy functions of a particular electronic state can be critical for a correct interpretation of the vibronic structure of an absorption band. In the case of the  $2^1B_{2u}$  and  $2^1B_{1u}$  excitations an additional difficulty for extracting  $T_v$  of each state from the observed spectrum is imposed by the fact that both bands are essentially overlapping. In general, a simulation of the intensity and energy distribution of the vibrational levels, for example based on ab initio potential energy surfaces, can be very useful for delimiting  $T_v$  of an electronic transition.

This detailed analysis of available estimates for  $T_v(2^1B_{2u})$  and  $T_v(2^1B_{1u})$  is partly motivated by the surprisingly large deviations between the experimental values taken from Ref. [4] and the CC3 estimates given by Ref. [19]. Even taking into account large tolerances for the spectroscopically determined  $T_v$  values for the reasons outlined above, it appears that CC3 overestimates  $T_v$  of

both states by ca. 0.3–0.4 eV, an unusually large error for single excitations. It is possible that a significant part of this discrepancy can be attributed to the inconsistency between the methods employed for the ground state geometry optimization (MP2) and excitation energy calculation (CC3).

We conclude that of the computational results collected in Table 2 the best approximations to  $T_v(2^1B_{2u})$  and  $T_v(2^1B_{1u})$  are obtained by PT2 calculations with inclusion of non-valence  $\pi^*$  orbitals into the active space presented in Ref. [7]. Ordered according to decreasing accuracy, these estimates are followed by the values computed with the PT3, CC3 and PT2 (with (10/8) active space) approaches.

The PT2 and PT3 calculations performed for this work congruently locate  $T_v(3^1B_{2u})$  close to 8 eV. Refs. [1, 31] report a ( $n \rightarrow 3p_y$ ) Rydberg transition of  $^1B_{2u}$  symmetry with  $T_{00} = 6.84\text{eV}$  and, for a Rydberg state, an extremely large oscillator strength of 0.25. Again, this assignment is doubted by Ref. [4]. Reference [32] tentatively identifies a ( $n \rightarrow 3p_y$ ) excitation with  $T_{00} = 7.04\text{eV}$ .

We cannot provide reference data for the postulated ( $n \rightarrow 3p_y$ ) Rydberg excitation. As has been mentioned above, the  $3^1B_{2u}$  transition considered in the present study is of ( $\pi \rightarrow 3p_x$ ) character. Unfortunately, we did not find experimental evidence for the energy of this state.

For  $T_v$  of the  $1^1B_{1g}$  and  $1^1B_{1u}$  valence states, the present PT2 calculations [subcolumns (a) and (b)] reach better agreement with experiment than PT3 and even than CC3. The quite different scenario of a strong dependence of the PT2 and PT3 energies on the choice of the basis set encountered in the case of the  $2^1B_{1g}$  Rydberg state has already been pointed out. This example shows that the additional effort required for the PT3 calculation can pay off as a good approximation of  $T_v(2^1B_{1g})$  is obtained, but obviously the accuracy is very sensitive to the Rydberg functions since a satisfactory result is only obtained with basis set 1, which provides the more extended version of diffuse functions (cf. Sect. 2.1).

The description of the  $3^1B_{1u}$  Rydberg state by the PT2 and PT3 schemes is consistent, the spread of the values for  $T_v(3^1B_{1u})$  is limited to a width of 0.18 eV.

Comparison to experiment reveals that the evaluation of  $T_v(1^1B_{2g})$  is another successful application of the PT3 approach. The computational results obtained for the  $1^1B_{2g}$  valence state, composed primarily of two single excitations, are therefore reminiscent of the accuracies found for  $T_v(2^1B_{2u})$  and  $T_v(2^1B_{1u})$ . PT3 works better than CC3, which could again be due to the multiconfigurational nature of this state, and the PT2 method yields a relatively low precision. This state has not been considered by Ref. [7], but it would be interesting to check if an extension of the

(10/8) active space by 6 non-valence  $\pi^*$  active orbitals could improve the performance of PT2 in analogy to  $T_v(2^1B_{2u})$  and  $T_v(2^1B_{1u})$ .

No experimental or computational reference data are available for  $T_v(2^1B_{2g})$  and  $T_v(2^1A_u)$ . The PT3 estimates for  $T_v$  of the  $2^1B_{2g}$  valence state exceed the PT2 predictions by  $\sim 0.6\text{eV}$ . As already indicated above, the agreement of the PT2 and PT3 methods on  $T_v$  of the  $2^1A_u$  Rydberg state is excellent, on the other hand.

The complicated case of the divergent predictions for  $T_v(1^1B_{3g})$  by the PT2, PT3 and CC3 electronic structure models has been discussed in Sect. 3.1. The current PT2 and PT3 estimates for  $T_v(2^1B_{3g})$  can only be compared to a PT2 result given in Ref. [7]. The  $2^1B_{3g}$  wave function can be described as a multiconfigurational  $\pi\pi^*$  excitation and accordingly Ref. [7] maintains that a  $3\pi + 9\pi^*$  active space has been defined for the CASSCF optimization. In contrast to the significant dependence of the PT2 results for the singly excited  $2^1B_{2u}$  and  $2^1B_{1u}$  states on the set of active orbitals, almost no variation of  $T_v(2^1B_{3g})$  is obtained upon replacing the (10/8) active space by the 12 orbitals outlined in Ref. [7].

The effect of switching from the PT2 to the PT3 approach on  $T_v(2^1B_{3g})$  is again an increase of the excitation energy, but the change of  $T_v(2^1B_{3g})$  by 0.35 eV is much less pronounced than in the case of  $T_v(2^1B_{2u})$  and  $T_v(2^1B_{1u})$ .

The computation of  $T_v(1^1A_u)$  is characterized by almost no influence of the choice of basis set 1 or 2 on the result but by a strong impact of selecting either the PT2 or PT3 methods. The PT3 scheme overestimates  $T_v(1^1A_u)$  by  $\sim 0.6\text{eV}$  while PT2 obtains a good agreement with experiment, the values are by 0.15–0.20 eV too low. The CC3 prediction exceeds experiment (4.72 eV) by  $\sim 0.3\text{eV}$  and is thus better than the PT3 calculation but a higher accuracy for a singly  $\pi\pi^*$  excited state would be expected.

### 3.5 Comments on the energetical ordering, the convergence and degeneracies of electronic states

Finally, we address the dependence of the energetical ordering of the various electronic states on the electronic structure model. The PT2 and PT3 calculations performed for this study as well as the CC3 results given in Ref. [19] all find that the  $1^1B_{3u}$  state corresponds to the lowest singlet excitation of pyrazine. However, more than one suggestion is already obtained for the  $S_2$  state. The PT2 method [subcolumns (a) and (b)] sequences the  $1^1A_u$  and  $1^1B_{2u}$  states as  $S_2$  and  $S_3$ , respectively, while PT3 and CC3 calculations arrive at the opposite succession.

The  $1^1B_{2g}$  state is unambiguously identified by the PT2, PT3 and CC3 approaches as  $S_4$ . The PT2 entries in

subcolumns (a) and (b) further agree on the  $S_5(1^1B_{1g})$ ,  $S_6(1^1B_{1u})$ ,  $S_8(2^1B_{2u})$ ,  $S_9(2^1B_{1u})$ ,  $S_{10}(2^1B_{3u})$ ,  $S_{12}(3^1B_{2u})$ ,  $S_{15}(2^1B_{3g})$ ,  $S_{16}(2^1A_g)$ ,  $S_{17}(3^1B_{1u})$ ,  $S_{18}(3^1A_g)$  and  $S_{19}(4^1A_g)$  assignments. The most noticeable dissent between the orderings obtained by the PT2 calculations concerns the  $1^1B_{3g}$  ( $S_7$  with basis 2,  $S_{11}$  with basis 1) and  $2^1B_{1g}$  ( $S_7$  with basis 1,  $S_{11}$  with basis 2) states.

For excitations from  $S_5$  and higher, the PT3 method proposes a different state allocation than PT2, coincidence is only reached again for the mapping of the  $S_{18}$  and  $S_{19}$  states.

Another interesting aspect is the correlation between CASSCF and PT2 energies. For valence states, the CASSCF estimates of  $T_v$  are usually higher than the corresponding PT2 predictions. The only exception in Table 2 is the  $2^1B_{3g}$  state. This state is also special because the inclusion of dynamic electron correlation via the PT2 method changes  $T_v$  only by  $\sim 0.25$  eV. Typically,  $T_v$  reacts much stronger to the consideration of higher lying excitations via the PT2 approach.  $T_v(1^1B_{3g})$  offers an extreme counter example: the transition energy drops by a huge 4.54 eV upon adding a PT2 calculation on top of the CASSCF optimization (with basis 2). As a result, the state ordering is reversed when the PT2 method is applied. The situation becomes even more complex when the state ordering is again turned around at the PT3 level of theory.

If Rydberg states are considered, no clear pattern with respect to the effect of dynamic electron correlation on  $T_v$  can be extracted from Table 2. Almost no modification of  $T_v$  is observed when CASSCF is replaced by PT2 in the case of the  $2^1B_{3u}$  and  $3^1B_{1u}$  states. The same procedure leads to an increase of  $T_v$  by  $\sim 0.8$  eV for the  $3^1B_{2u}$  and  $2^1A_u$  transitions. The  $2^1B_{1g}$  state finally is unique in that inclusion of dynamic electron correlation via the PT2 approach induces an increase by 0.14 eV with basis 2 and a decrease by 0.98 eV with basis 1.

It was very difficult to converge the  $2^1B_{1g}$  wave function in the CASSCF with basis 1. The first two roots of  $1^1B_{1g}$  symmetry are nearly degenerate at this level of theory and valence–Rydberg mixing is very likely to occur in this situation. The CASSCF algorithm tends either to place the Rydberg state below the valence state or to completely lose the Rydberg character of one root and return two valence states, depending on the state averaging scheme. According to the CASSCF approach both roots are within 0.140 eV and PT2 even places the Rydberg excitation only 0.03 eV above the valence state (cf. Table 2).

Provided that the electronic structure calculations are accurate this result would imply that the  $1^1B_{1g}$  and  $2^1B_{1g}$  states constitute an interesting case of accidental degeneracy of a valence and a Rydberg state that transform according to the same irreducible representation, in contrast to the symmetry induced degeneracy in a Jahn–Teller system.

We also note that PT2 calculations with basis 1 predict a degeneracy of the  $2^1B_{1g}$  and  $1^1B_{1u}$  states. The vibrational modes  $\nu_{16a}$  and  $\nu_{17a}$  of  $A_u$  symmetry should therefore be able to induce strong vibronic coupling between the first two roots of  $1^1B_{1g}$  symmetry and the  $1^1B_{1u}$  state.

### 3.6 Oscillator strengths

In Table 4, oscillator strengths  $f$  computed by the same procedures as employed for the energies compiled in Table 2 are compared to experimental values for the symmetry allowed transitions. The PT2 oscillator strengths are calculated by combining the transition dipole moment values obtained at the CASSCF level of theory with the vertical excitation energies predicted by the PT2 method.

We obtain a very good consistency of the  $f$  values obtained with basis sets 1 and 2 for this study [subcolumns

**Table 4** Oscillator strengths  $f$  for eight excited states of pyrazine

|                | $f(\text{CASSCF})$ |       |       | $f(\text{PT2})$ |       |       | $f(\text{CCSD})$ | $f(\text{exp})$ | $f(\text{exp})$ |
|----------------|--------------------|-------|-------|-----------------|-------|-------|------------------|-----------------|-----------------|
|                | a                  | b     | c     | a               | b     | c     | [19]             | [4]             | [1]             |
| $1^1B_{3u}(V)$ | 0.013              | 0.014 | 0.014 | 0.011           | 0.011 | 0.010 | 0.008            | 0.006           | 0.006           |
| $2^1B_{3u}(R)$ | 0.002              | 0.002 |       | 0.002           | 0.002 |       |                  |                 |                 |
| $1^1B_{2u}(V)$ | 0.055              | 0.053 | 0.073 | 0.052           | 0.052 | 0.068 | 0.067            | 0.062           | 0.100/0.060     |
| $2^1B_{2u}(V)$ | 0.949              | 0.942 | 0.817 | 0.679           | 0.634 | 0.423 | 0.662            | 0.720           |                 |
| $3^1B_{2u}(R)$ | 0.011              | 0.011 |       | 0.012           | 0.013 |       |                  |                 | (0.250)         |
| $1^1B_{1u}(V)$ | 0.191              | 0.193 | 0.103 | 0.146           | 0.148 | 0.081 | 0.074            | 0.100           | 0.150/0.100     |
| $2^1B_{1u}(V)$ | 1.021              | 1.027 | 0.966 | 0.710           | 0.692 | 0.754 | 0.458            | 0.720           | 1.000/0.700     |
| $3^1B_{1u}(R)$ | 0.001              | 0.001 |       | 0.001           | 0.001 |       |                  |                 |                 |

Subcolumns labeled by (a) and (b) present the CASSCF/PT2 results calculated with basis sets 2 and 1, respectively, in combination with active spaces (10/8)/(10/9). The data in subcolumns designated by (c) are taken from Table 7 in Ref. [7] and correspond to CASSCF/PT2 calculations with a (4s3p2d/3s2p) ANO basis set. Also included are CCSD estimates of  $f$  [19]. As far as available, experimental information on  $f$  is included as a reference

(a) and (b)]. With respect to the agreement between the CASSCF and PT2 results calculated with both basis sets, it is clear that any deviations of  $f$  must be due to different values of  $T_v$  for these excitations obtained at the CASSCF and PT2 levels of theory, respectively, since only CASSCF transition dipole moment values are employed (Table 2).

The significantly lower estimates for  $T_v(2^1B_{2u})$ ,  $T_v(1^1B_{1u})$  and  $T_v(2^1B_{1u})$  derived by the PT2 method as compared to the CASSCF predictions are consequently reflected in the the corresponding  $f$  values.

The CASSCF and PT2 data for  $f$  made available in Ref. [7] are found to coincide well with the present calculations. The  $1^1B_{1u}$  valence state is a notable exception, Ref. [7] suggests that  $f$  should be by almost a factor two smaller [subcolumns (c)] than the corresponding entries in subcolumns (a) and (b). Two experimental figures of 0.15 and 0.10 are given for  $f(1^1B_{1u})$  in Ref. [1], the first number is closer to the current values while the second supports the notion published in Ref. [7]. Reference [4] only quotes an experimental value of 0.10 for  $f(1^1B_{1u})$ . The CCSD-based estimate for  $f(1^1B_{1u})$  is even lower, 0.074 [19].

Another relatively large disagreement between the earlier PT2 calculations described in Ref. [7] and the current results can be seen for  $f(2^1B_{2u})$ . The estimate given in Ref. [7] is again lower (0.423). In this case, the agreement of the present PT2 results [subcolumns (a) and (b)] with experiment [4] and with the recent CCSD computation [19] is clearly better.

In general, a good match between the PT2 [subcolumns (a), (b) and (c)], CCSD and experimental estimates for  $f$  of the various excited states can be established.

## 4 Conclusions

In this study, we have investigated the electronic spectrum of pyrazine by carrying out CASSCF, PT2 and PT3 calculations. We present and apply a procedure for the evaluation of accurate wave functions for Rydberg excitations within the CASSCF electronic structure model. The key idea of this approach is the definition of sets of active orbitals that are adapted to the physical nature of each state (state-specific scheme). The selection of the active space is performed with the help of a screening step. This concept reduces the problem of artificial valence–Rydberg mixing and allows for a highly consistent evaluation of electronic properties. The difficult convergence of Rydberg states in the CASSCF optimization is accelerated by constructing suitable starting vectors. The critical point in this context is to lock Rydberg-type functions within the active space. This is accomplished by performing full CI calculations for the set of active orbitals.

The vertical excitation energies of 19 states of pyrazine have been calculated by employing two AVTZ-level basis

sets extended by one set (basis 2), respectively, by four s-type and three p,d-type sets (basis 1) of molecule-centered diffuse Rydberg functions. Estimates of vertical excitation energies have been obtained with quite non-uniform accuracy, depending strongly on the nature of a particular state. The identification of the different states and the ordering of the vertical excitation energies has been discussed quite controversially by spectroscopists and theoreticians. From the theory point of view, this depends to some extent on the electronic structure methods and basis sets used for the calculations. Our results support most of the assignments made in Refs. [1, 4].

However, a few notable disagreements are also obtained. For example, we cannot confirm the existence of a low lying ( $n \rightarrow 3s$ ) Rydberg excitation as implied by Ref. [1].

Regarding the comparison of the present calculations of  $T_v$  for various excited states of pyrazine to experimental data and to alternative computational results, we obtain a quite incoherent picture. The accuracy of the PT2, PT3 and CC3 estimates of  $T_v$  appears to depend strongly on the nature of a particular electronic state and questions remain regarding the reasons for these large error oscillations. For example, it is well known that the CC3 approach may have difficulties in describing multiply excited states. However, we find that CC3 significantly overestimates  $T_v$  of the singly  $\pi\pi^*$  excited  $1^1B_{1g}$  state (Table 2).

Another topic is the accuracy of the PT3 approach. Recently, significant attention has been paid to the development of this and similar methods (see, e.g., Refs. [33, 34]). Our results do not reveal a unequivocal pattern that advancing from the PT2 to the PT3 method will likely lead to an improvement of the  $T_v$  approximations. Instead, we find that PT3 has a tendency to significantly overestimate  $T_v$  of valence excited states. It appears as if the divergence problem of the single reference MPn series [35] may similarly affect the multiconfigurational PTn counterpart. This question should be investigated in more detail. For example, we have applied the PT3 method in this work to evaluate  $T_v$  values at geometries that have been optimized at the PT2 level of theory. The effect of this inconsistency on the  $T_v$  predictions obtained by the PT3 formalism could lead to a substantial distortion of the results.

The experience obtained by working on this system will be used to continue studies on the excited states of pyrazine. We are planning to calculate accurate potential energy surfaces for different electronic states including Rydberg transitions with the goal to localize conical intersections that may be relevant for nonadiabatic effects on the spectroscopy of this system.

**Acknowledgments** Á.V. acknowledges the OTKA Grant No. T67923 and the computational resources provided by the John-von-

Neumann Institute, Research Centre Jülich (Project ID ehu01). Thanks go also to Wolfgang Domcke (TU München) and to the Leibniz Supercomputing Centre in München for providing computational resources on the Linux Cluster. We also thank Kenneth Ruud (U Tromsø) for stimulating discussions. G.J.H. and Á.V. acknowledge the DFG Grant (Project ID SU244/7-1) for partly supporting this study. C.W. would like to thank the Centre for Theoretical and Computational Chemistry (CTCC) at the University of Tromsø and the Research Council of Norway (Grant Nr. 177558/V30) for continued support.

## References

1. Innes KK, Ross IG, Moomaw WR (1988) *J Mol Spectr* 132:492
2. Yamazaki I, Murao T, Yamanaka T, Yoshihara K (1983) *Faraday Discuss Chem Soc* 75:395
3. Kommandeur J, Majewski WA, Meerts WL, Pratt DW (1987) *Ann Rev Phys Chem* 38:433
4. Bolovinos A, Tsekeris P, Philips J, Pantos E, Anditsopoulou G (1984) *J Mol Spectr* 103:240
5. Palmer MH, Walker IC (1991) *Chem Phys* 157:187
6. Fülischer MP, Andersson K, Roos BO (1992) *J Phys Chem* 96:9204
7. Fülischer MP, Roos BO (1994) *Theor Chim Acta* 87:403
8. DelBene JE, Watts JD, Bartlett RJJ (1997) *J Chem Phys* 106:6051
9. Weber P, Reimers JR (1999) *J Phys Chem A* 9821:103
10. Li Y, Wan J, Xu X (2007) *J Comp Chem* 28:1658
11. Worth GA, Meyer H-D, Cederbaum LS (1996) *J Chem Phys* 105:4412
12. Raab A, Worth GA, Meyer H-D, Cederbaum LS (1999) *J Chem Phys* 110:936
13. Puzari P, Swathi RS, Sarkar B, Adhikari SJ (2005) *J Chem Phys* 123:134317
14. Puzari P, Sarkar B, Adhikari SJ (2005) *J Chem Phys* 125:194316
15. Domcke W, Yarkony DR, Köppel H (2004) *Conical intersections: electronic structure, dynamics and spectroscopy*. World Scientific, Singapore
16. Stock G, Woywod C, Domcke W, Swinney T, Hudson BS (1995) *J Chem Phys* 103:6851
17. Sobolewski AL, Domcke W (1991) *Chem Phys Lett* 180:381
18. RongXing H, ChaoYuan Z, Chih-Hao C, Sheng-Hsien L (2008) *Sci China Ser B Chem* 51:1166
19. Schreiber M, Silva-Junior MR, Sauer SPA, Thiel W (2008) *J Chem Phys* 128:134110
20. Werner H-J et al (2006) *Molpro*, version 2006.1, a package of ab initio programs
21. Müller T, Dallos M, Lischka H (1999) *J Chem Phys* 110:7176
22. Dallos M, Lischka H (2004) *Theor Chem Acc* 112:16
23. Woywod C, Livingood WC, Frederick JH (2000) *J Chem Phys* 112:613
24. Dunning TH Jr (1989) *J Chem Phys* 90:1007
25. Kendall RA, Dunning TH Jr, Harrison RJ (1992) *J Chem Phys* 96:6796
26. Andersson K (1995) *Theor Chim Acta* 91:31
27. Woywod C, Domcke W, Sobolewski AL, Werner H-J (1994) *J Chem Phys* 100:1400
28. Oku M (2008) *J Phys Chem A* 112:2293
29. Hinchliffe A (2003) *Molecular modelling for beginners*. Wiley, West Sussex
30. Schäfer A, Horn H, Ahlrichs R (1992) *J Chem Phys* 97:2571
31. Scheps R, Florida D, Rice SA (1972) *J Mol Spectr* 44:1
32. Song JK, Tsubouchi M, Suzuki T (2001) *J Chem Phys* 115:8810
33. Angeli C, Cimraglia R (2002) *Theor Chem Acc* 107:31
34. Angeli C, Bories B, Cavallini A, Cimraglia R (2006) *J Chem Phys* 124:054108
35. Halkier A, Larsen H, Jørgensen P (1999) *J Chem Phys* 110:7127