I. INTRODUCTION

Azabenzenes are important parent molecular systems for numerous compounds such as biologically active nicotinic acid and the nucleotides cytosine, uracil, and thymine, and therefore have been the subject of extensive experimental and theoretical studies.1–17 Azabenzenes are isoelectronic with benzene. The nitrogen atoms introduce perturbations to the benzene energy levels and give rise to new transitions due to excitation of their lone-pair electrons. Their lower electronic excitation spectra in the near and vacuum ultraviolet regions are due to both n → π* and π → π* valence transitions, and the lowest-lying transitions are of n → π* origin; these are the fundamental differences in the photophysical behavior of azabenzenes and benzene.

A good review of experimental and theoretical studies on the electronic structure of azabenzenes up to 1988 has been given by Innes et al.1 Recently, a series of studies have been reported by Walker and co-workers,2–5 which include vacuum ultraviolet absorption (VUV) spectra, electron energy loss (EEL) spectra, and the results of multi-reference multi-root configuration-interaction (MRD-CI) calculations. In addition, a comprehensive collection of VUV spectra and theoretical results has also been presented by Bolovinos et al.6 Pyridine is the simplest azabenzene because only one CH group is replaced by nitrogen atom. In fact, the optical spectrum of pyridine is dominated by three bands at about 5.0, 6.4, and 7.2 eV, which are associated to the π → π* valence excitations from the comparison with the three π → π* valence excitations bands of benzene at 4.9, 6.2, and 7.0 eV, respectively. The valence nature of these transitions was confirmed by their existence in the liquid phase.7 The lowest-lying singlet absorption of pyridine contains an n → π* excitation on the rising shoulder of the first optical absorption band.2,6 The lowest Rydberg-excited state, n → 3s, has been observed at 6.28 eV by multi-photon ionization.8,9 Other Rydberg series have been assigned from the fine structure at 8–9 eV in the optical spectrum.2,6

Optical absorption data for pyridine have been complemented by electron energy-loss spectra. With incident electrons of relatively high energies, the electron-impact excitation spectrum resembles the optical absorption spectrum.9 When the incident electron energy is reduced to near-threshold values,10 additional transitions emerge, including spin-forbidden ones. Ion-impact (He+) excitation of pyridine reveals a singlet-triplet excitation, with an onset energy of 3.8 eV and E_max=4.1 eV.12 This is the lowest-lying excited state to have been definitively located in the gas phase. Other singlet-triplet excitations at 4.84 and 5.4 eV were observed in the experimental studies of Walker et al.2

The first band in the photoelectron spectrum of pyridine is a composite of two ionization potentials involving an N nonbonding (n) electron and a π electron. High-resolution studies and comparison with the deuterated, D5, molecule show that the first ionization potential (IP) is nonbonding (n,1a1).13,14 This is consistent with the observation that the lowest Rydberg state has A1 symmetry (n,3s).8,9 Recently, a high-resolution synchrotron photoelectron spectroscopy (PES) study, including the outer- and inner-valence regions,
was reported by Moghaddam et al.\textsuperscript{15} In their report, 1\textalpha\textsubscript{2} was assigned to the first IP using an algebra diagrammatic construction accurate to the third order in electron–electron interaction ADC(3) calculations.

On the other hand, in contrast to this abundant experimental information regarding pyridine, there have been only a few \textit{ab initio} studies (at the electron correlation level) on the excited and ionized states of pyridine: symmetry-adapted cluster (SAC) and SAC configuration interaction (SAC-CI) calculations\textsuperscript{16} in 1988, MRD-CI calculations\textsuperscript{2} in 1989, complete active space second-order perturbation (CASPT2) calculations\textsuperscript{17,18} in 1992 and 1995, and ADC(3) calculations\textsuperscript{15} in 1996. The SAC/SAC-CI study,\textsuperscript{13} a pioneering work, but due to the limitations in the computational resources available at that time, the rather limited basis sets and relatively small active space impaired the quality of the results. CASPT2\textsuperscript{17} has been used to systematically study the singlet valence $n \rightarrow \pi^\ast$ and $\pi \rightarrow \pi^\ast$ excitations of azabenzenes including pyridine, pyrimidine, pyrazine, pyridazine, and $s$-triazine, and the singlet Rydberg transitions ($n=3$, $l=0,1,2$), the triplet excited states and the first three ionized states of pyridine.\textsuperscript{18}

In this study, we re-examined the singlet and triplet valence and Rydberg excitations of pyridine using an extended basis set and a larger active space by SAC/SAC-CI single- and double-operator (SD-R) calculations.\textsuperscript{19–27} Furthermore, to provide a theoretical interpretation of correlation (satellite) peaks characterized by a multi-electron process in the PES spectrum,\textsuperscript{15} the SAC/SAC-CI general-R method\textsuperscript{25} was also used to accurately calculate the inner-valence region of the ionization spectrum. In Sec. II, we outline the computational details. Calculated results and discussions are presented in Sec. III. A summary is given in Sec. IV.

## II. Computational Details

Details of the SAC/SAC-CI method for calculating ground, excited, and ionized states of molecules have been presented elsewhere.\textsuperscript{19–27} An experimental ground-state equilibrium geometry of pyridine with $C_{2v}$ symmetry\textsuperscript{28} was used, and the pyridine ring was put on the $yz$-plane with $z$ as the principle axis. Therefore, the calculated excitation energies are vertical in nature.

One of the major sources of errors in the excitation energies is from incompleteness in the basis set. Most standard basis sets are optimized for ground-state calculations. Since excited states are often considerably more diffuse than ground states, extended basis sets are required. Therefore, we use well-extended basis sets to perform SAC/SAC-CI SD-R calculations to overcome this problem: Dunning’s augmented correlation consistent basis set AUG-cc-pVTZ\textsuperscript{29} was used for nitrogen and carbon atoms with one $f$ polarization function removed for all carbon atoms, and cc-pVTZ\textsuperscript{30} was used for H atoms. Additionally, a set of molecule-centered diffuse function (5s5p5d) selected from the studies of Kaufmann et al.\textsuperscript{31} was placed on the N atom, since many $n$-Rydberg states are targets in pyridine. For convenience, the above basis set is referred to as basis set A. In the SAC/SAC-CI general-R calculations for the ionized states of pyridine, we use basis set B, in which AUG-cc-pVDZ was used for nitrogen and carbon atoms, and cc-pVDZ was used for H atoms, since the computing dimension is considerably increased in the general-R method, and these were then augmented by the molecule-centered diffuse functions (2s2p2d) above the molecular center of gravity. All SCF calculations were performed with the \textsc{gaussian} 98 package.\textsuperscript{32}

The present SAC/SAC-CI calculations were performed with the local version of the SAC/SAC-CI module.\textsuperscript{27} The active space consists of the complete molecular orbital space except that 1$s$ core orbitals were frozen. In the SAC ground-state calculation, all single-excitation and selected double-excitation operators $S_I$ were included in the linked term. The energy threshold $\lambda_g$ for perturbation selection\textsuperscript{24} (PS) was $1.0 \times 10^{-5}$ a.u. For the unlinked term, we included only the products of the double-excitation operators $S_I^* S_I^\dagger$, when the coefficients $C_I$ and $C_{I'}$ estimated by SD-CI in practice, were larger than $1.0 \times 10^{-3}$.

In the SAC-CI SD-R calculations of excited states, all single-excitation operators and selected double-excitation operators were included in the linked operators $R_K^\pm$. Perturbation selection was performed as follows. First, we selected the main reference configurations from SE-CI with a coefficient greater than 0.1, and then selected the double-excitation operator whose second-order perturbation matrix element with one of the main reference configurations is greater than the threshold $\lambda_g (1.0 \times 10^{-6}$ a.u.). In the unlinked term $\Sigma_{K^\pm} d_{K^\pm} C_{K^\pm} S_{I^\pm}^\dagger | 0 \rangle$, we included double-excitation operators $\{S_I^\pm\}$ whose coefficients $(C_I)$ were larger than $1.0 \times 10^{-3}$ in the SAC ground-state calculation, and as the $R_K^\pm$ operator, we included single- and double-excitation operators whose coefficients (taken from SD-CI in practice) were greater than 0.05.

In the SAC-CI general-R calculations of the outer- and inner-valence ionization potentials of pyridine, a preliminary SD-CI calculation was performed to select operators for the exponential generation (EG) scheme\textsuperscript{26} by which we can construct important triple- and quadruple-excitation operators required for a quantitative description of the multi-electron process in the ionization spectrum. The generated higher-order excitation operators were further selected by the per-
turbation selection procedure with an energy threshold of $1.0 \times 10^{-6}$ hartree. Therefore, all single (ionization) operators and selected higher operators (SDTQ) were included in the linked term of the SAC-CI general-R equation. In the unlinked term, we included all of the double-excitation operators $S_{12}$ in the SAC ground-state calculation, and as the $R_K$ operator, we included all single operators and selected higher-order operators up to quadruple.

### III. RESULTS AND DISCUSSIONS

Oscillator strengths and excited-state properties, such as second moments, are important for characterization of the excited electronic states and give information that is useful for the qualitative assignment of the excited states, for example in terms of valence and Rydberg states. These points have been illustrated in our previous calculations on five-membered ring compounds. The triplet excitation energies and second moments calculated by the SAC-CI SD-R method are shown in Table I, and compared with the experimental values and the CASPT2 and MRD-CI results. Table II shows the singlet excitation energies, oscillator strengths, and second moments calculated by the SAC-CI SD-R method, together with the experimental values, and the CASPT2 and MRD-CI results. Figure 1 shows a comparison of the experimental VUV spectrum and the theoretical SAC-CI spectrum of pyridine. The ionization potentials calculated by the SAC-CI general-R method are listed in Table III, together with the PES data and other theoretical values.
are observed at 4.44(1B₁) and 4.99 eV(1B₂).² Our present computed results gave the 1B₁ state at 4.59 eV with a weak oscillator strength of 0.0054, and the 1B₂ state at 4.85 eV with an oscillator strength of 0.04. The 1B₁ state is n → π*, while the 1B₂ state is π → π* in nature. MRD-CI calculations gave these at 4.50(1B₁) and 4.76 eV(1B₂), respectively, while CASPT2 calculations put them at 4.91(1B₁) and 4.84 eV(1B₂), respectively. Note that the CASPT2 calculations¹⁸ suggested the lowest n → π* transition corresponded to an adiabatic transition.

The maximum of the second optically allowed absorption band (5.8–6.8 eV) is observed at 6.38 eV, which is assigned to valence π → π* excitation (1A₁).²,⁶ The present calculations gave the 31A₁ state at 6.26 eV with an oscillator strength of 0.02. The second moment (93.9) and main configurations clearly defined it as a valence π → π* excitation. CASPT2 calculations gave this at 6.42 eV with a weak oscillator strength of 0.005, while MRD-CI calculations put it at 6.99 eV. Close to the maximum, the Rydberg n–s excitations is determined by two-photon absorption.⁸,⁹ at 6.28 eV. Our present results gave the 21A₁(n–3s) Rydberg state at 6.17 eV with a weak oscillator strength of 0.0008. MRD-CI gave a calculated value of 6.19 eV, while CASPT2 gave this state at 6.70 eV.

The maximum of the strongest optical absorption band is experimentally observed at 7.22 eV.²,⁶ This absorption corresponds to the benzene 1E₂u excitation, which becomes 1A₁+1B₂ in pyridine. Our present calculations gave two nearly degenerate singlet excited states 1A₁(π → π*) and 3B₂(π → π*) at 7.18 and 7.27 eV, with large oscillator strengths of 0.62 and 0.41, respectively, which are certainly responsible for the strongest absorption at 7.22 eV. The second moment of the 41A₁ state is 97.6 while that of the 31B₂ state 114.9, which shows a somewhat diffuse character due to mixing of the Rydberg state, particularly in the 31B₂ state. Vice versa when the Rydberg state is strongly influenced by valence excitations, the former will borrow some intensity from the latter. This is why the calculated Rydberg 21B₂ state had a relatively intense absorption (0.21). MRD-CI gave these two valence excitations at 8.18 eV(1A₁) and 7.70 eV(1B₂), respectively, while CASPT2 gave them at 7.23 eV(1A₁), with a CASSCF oscillator strength of 0.82, and 7.48 eV(1B₂), with a CASSCF oscillator strength of 0.64.

In this energy region, two other weak absorptions are visible at 6.93 and 7.6 eV which were associated with Rydberg states by Walker et al. In the higher energy region of 8–9 eV, a broad optical absorption band is observed, which is caused by transitions to Rydberg states. Our present results gave many Rydberg excitations in these energy regions, and an assignment is given in Table II.

B. Ionized states and assignment of the photoelectron spectrum

The ground-state valence electronic configurations of pyridine (C₃ᵥ) in order of increasing orbital energy is 5a₁, 6a₁, 3b₂, 7a₁,4b₂, 8a₁, 5b₂, 9a₁, 6b₂, 10a₁, 1b₁, 7b₂, 11a₁, 2b₁,1a₂.
TABLE III. Calculated ionization potentials (IP, in eV), monopole intensities (M.I.), and main configurations with the basis set B for pyridine.

<table>
<thead>
<tr>
<th>Sym.</th>
<th>SAC-CI general-R</th>
<th>Expt.*</th>
<th>CASPT² ADC(3)*</th>
<th>MRD-Cl²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Main configuration</td>
<td>M.I.</td>
<td>IP  (eV)</td>
<td>IP</td>
</tr>
<tr>
<td>2A₁</td>
<td>0.91(1a₁) - 0.12(2b₁)</td>
<td>0.8476</td>
<td>9.23</td>
<td>9.60(1)</td>
</tr>
<tr>
<td></td>
<td>0.90(1a₂) + 0.14(2b₁)</td>
<td>0.8385</td>
<td>13.56</td>
<td>13.8(6)</td>
</tr>
<tr>
<td></td>
<td>0.89(9a₁) - 0.17(1a₂a₂;1a₁)</td>
<td>0.8018</td>
<td>15.70</td>
<td>15.9(9)</td>
</tr>
<tr>
<td></td>
<td>0.0221</td>
<td>15.79</td>
<td>0.0135</td>
<td>16.50</td>
</tr>
<tr>
<td></td>
<td>- 0.87(8a₁) + 0.16(2b₁)</td>
<td>0.7687</td>
<td>17.10</td>
<td>17.4(10)</td>
</tr>
<tr>
<td></td>
<td>0.0309</td>
<td>17.92</td>
<td>0.1175</td>
<td>19.82</td>
</tr>
<tr>
<td></td>
<td>- 0.63(7a₁) + 0.31(10a₁a₂;1a₂)</td>
<td>0.4022</td>
<td>20.05</td>
<td>19.8(11)</td>
</tr>
<tr>
<td></td>
<td>0.0447</td>
<td>20.54</td>
<td>0.0956</td>
<td>20.21</td>
</tr>
<tr>
<td></td>
<td>0.0101</td>
<td>21.66</td>
<td>0.0129</td>
<td>21.24</td>
</tr>
<tr>
<td></td>
<td>0.0124</td>
<td>22.59</td>
<td>0.0102</td>
<td>21.97</td>
</tr>
<tr>
<td></td>
<td>0.0172</td>
<td>22.71</td>
<td>0.01954</td>
<td>23.27</td>
</tr>
<tr>
<td>2A₂</td>
<td>- 0.41(6a₁) - 0.24(1a₂a₂;10a₁) - 0.19(10a₁b₁;2b₁)</td>
<td>0.1954</td>
<td>23.27</td>
<td>23.4(13)</td>
</tr>
<tr>
<td>2B₁</td>
<td>0.94(1a₂)</td>
<td>0.8989</td>
<td>9.36</td>
<td>9.75(2)</td>
</tr>
<tr>
<td></td>
<td>0.94(2h₁) + 0.10(2b₁b₁;1b₁)</td>
<td>0.8907</td>
<td>10.12</td>
<td>10.51(3)</td>
</tr>
<tr>
<td></td>
<td>0.87(1b₁) - 0.28(1a₂b₁;1a₂) + 0.22(2b₁b₁;2b₁)</td>
<td>0.7636</td>
<td>13.06</td>
<td>13.1(5)</td>
</tr>
<tr>
<td>2B₂</td>
<td>0.93(7b₂)</td>
<td>0.8806</td>
<td>12.48</td>
<td>12.6(4)</td>
</tr>
<tr>
<td></td>
<td>0.91(6b₂) - 0.16(1a₂b₁;11a₁)</td>
<td>0.8490</td>
<td>14.14</td>
<td>14.57</td>
</tr>
<tr>
<td></td>
<td>- 0.86(5b₂) - 0.31(1a₂b₁;11a₁)</td>
<td>0.7610</td>
<td>15.60</td>
<td>15.9(8)</td>
</tr>
<tr>
<td></td>
<td>0.1187</td>
<td>16.84</td>
<td>0.0307</td>
<td>20.66</td>
</tr>
<tr>
<td></td>
<td>0.0203</td>
<td>24.07</td>
<td>0.0161</td>
<td>21.94</td>
</tr>
<tr>
<td></td>
<td>0.0134</td>
<td>24.51</td>
<td>0.0134</td>
<td>24.51</td>
</tr>
<tr>
<td></td>
<td>- 0.87(4b₂) - 0.16(1a₂b₁;11a₁)</td>
<td>0.7629</td>
<td>20.43</td>
<td>20.6(12)</td>
</tr>
<tr>
<td></td>
<td>0.0236</td>
<td>24.30</td>
<td>0.0294</td>
<td>21.72</td>
</tr>
<tr>
<td></td>
<td>- 0.52(3b₂) - 0.27(1a₂a₁;2b₁) - 0.24(2b₁a₁;1a₂)</td>
<td>0.2792</td>
<td>24.35</td>
<td>24.3(14)</td>
</tr>
<tr>
<td></td>
<td>0.0180</td>
<td>24.09</td>
<td>0.0110</td>
<td>22.76</td>
</tr>
</tbody>
</table>

*aReferences 2 and 6, the order of peaks (binding energies) is given in parenthesis.
*bReference 18.
*cReference 15.
*dReference 2.

Synchrotron PES experiments with pyridine have recently been conducted by Moghaddam et al.¹⁵ up to 25 eV, which includes both outer- and inner-valence regions (Fig. 2). Table III shows the SAC-CI general-R theoretical ionization energies (binding energies), monopole intensities, and assignments of the ionization peaks observed in PES. The present SAC-CI general-R calculations are in good agreement with the experimental values. In Fig. 2, the experimental PES and the present theoretical spectrum calculated by the SAC-CI general-R method are compared.

Most early assignments in the pyridine photoelectron spectrum were based on Koopman's theorem, i.e., |IP| = - |e_i|, where e_i is the orbital energy. Koopmans' values deviate strongly from the experimental values except for those in the two states associated with 1a₂ and 2b₁. As previously mentioned (Sec. I), the first band in PES contains two ionizations, IP order 11a₁ < 1a₂ (π), which have been separated in high-resolution studies.¹³¹⁴ Koopmans' results put the IP order 1a₂(π) < 11a₁(n). Our present SAC-CI general-R calculations computed these two states at 9.23[11a₁(n)] and 9.36 eV[1a₂(π)], respectively, which are effectively degenerate compared to the Koopmans' values: 9.46[1a₂(π)] and 11.42[11a₁(n)] eV. Our present results support the experimental studies.⁸,⁹,¹³,¹⁴ MRD-Cl gave 8.96[11a₁(n)] and 9.38 eV[1a₂(π)]. However, ADC(3) gave a different IP order: 9.44[1a₂(π)] and 9.78[11a₁(n)].

The third IP peak is observed at 10.51 eV. Our present calculations gave this state at 10.12 eV with a monopole intensity of 0.89. The nature of this state is, from the main configurations shown in Table III, dominated by a one-electron process of the orbital 2b₁. Note that CASPT2 gave...
the first three ionization potentials at 9.54 eV[11a1(n)], 9.65 eV[1a2(p)], and 10.37 eV(2b1).

Following the third IP peak, there are four sequential ionization peaks (4th, 5th, 6th, and 7th) observed at 12.61, 13.1, 13.8, and 14.5 eV. Our present results gave the corresponding ionized states at 12.48, 13.06, 13.56, and 14.41 eV, with monopole intensities of 0.88, 0.76, 0.84, and 0.85, respectively. From the main configurations of each ionized state, the 12.61 eV peak was associated with the ionization from the orbital 7b2, and dominated by a one-electron process. The 13.1 eV peak was assigned to the ionization from the orbital 1b1, and was affected by a two-electron process. The 13.8 eV peak was assigned to ionization from the orbital 10a1, While the 14.5 eV peak was associated with the ionization from the orbital 6b2, which is dominated by a one-electron process. Note that Koopmans’ order of ionizations is maintained within these states.

The next two ionization peaks (8th and 9th) are observed at 15.9 eV, which overlap with each other. Our present calculations gave two almost degenerate states at 15.60 and 15.70 eV with monopole intensities of 0.76 and 0.80, respectively. We assigned the 15.9 eV peak to ionizations from the orbitals 5b2 and 9a1. Note that Koopmans’ order broke down again in this energy region.

The 10th ionization peak is located at 17.4 eV. Our present results associated it with ionization from the orbital 8a1, which was computed at 17.10 eV with a monopole intensity of 0.77. In our present calculations, some satellite peaks occurred in this peak region. The detailed numerical results are listed in Table III.

In the energy region of 20 to 25 eV, two groups of broad peaks with satellite characteristics (11th, 12th, and 13th, 14th) are observed. These ionized states are strongly influenced by multi-electron processes. The detailed numerical results, the nature of the ionized states, and the corresponding assignments are shown in Table III. Generally, our present results agree well with all PES peaks up to 25 eV.

IV. CONCLUSION

In the present study, singlet and triplet excited states and ionized states of pyridine were re-investigated by the SAC/SAC-CI method with an extended basis set and a large active orbital space. The 48 low-lying singlet and triplet excited states were calculated and a reliable assignment for VUV and EEL spectra up to 8 eV was made using the calculated excitation energies and oscillator strengths. The valence n-π*, π-π*, and Rydberg excited states were discussed separately in detail. Furthermore, the SAC-CI general-R method, along with both the exponential generation and perturbation selection schemes for including higher-order excitation operators, was applied to the ionized states of pyridine up to 25 eV. The experimental PES spectrum and our present calculations show good agreement in terms of the binding energies and peak intensities in both the outer- and inner-valence regions.

ACKNOWLEDGMENTS

We are grateful to K. Toyota and R. Fukuda for their helpful discussions on this subject. This research was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Culture and Sports.

27 H. Nakatsuji, M. Hada, M. Ehara, J. Hasegawa, T. Nakajima, H. Nakai, O. Kitao, and K. Toyota, SAC/SAC-CI program system (SAC-CI96) for calculating ground, excited, ionized, and electron attached states and singlet to septet spin multiplicities.
32 GAUSSIAN 98 (Revision A.1), M. J. Frisch, G. W. Trucks, H. B. Schlegel et al. (Gaussian, Inc., Pittsburgh, PA, 1998).