To Improve Accuracies of the HITRAN H_2O Database with Theoretically Developed Tools

Q. Ma

Goddard Institute for Space Studies & Department of Applied Physics and Applied Mathematics, Columbia University
2880 Broadway, New York, NY 10025, USA

Contents

1. **Sun and Earth's Radiations**

2. **Atmospheric Molecules**

3. **Molecular Absorptions**

4. **HITRAN Database**

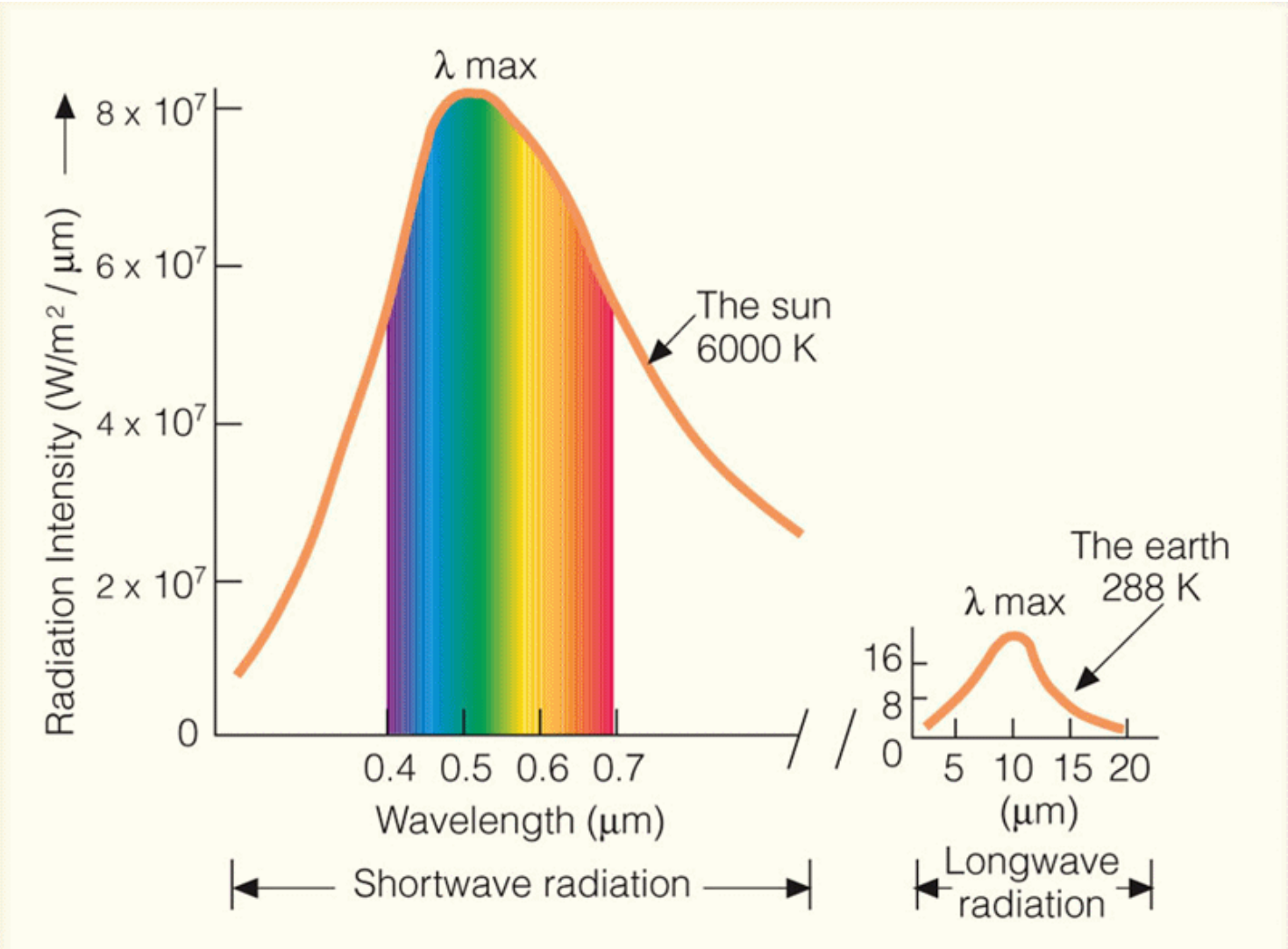
5. **Why and How the Spectroscopic Parameters vary with H_2O Lines**

6. **Screening HITRAN H_2O Database with the Rules**

7. **Identifying Missing Lines**

*** Some of slides are copied from web sites**.

I. Sun and Earth Radiations



I. Sum and Earth Radiations
Main Features

- Sun with a surface temperature T = 6000 K emits radiation that peaks at relatively short wavelength around 0.5 μm.

- Earth with a surface temperature T = 288 K (15^0C) emits peak radiation at longer wavelengths around 10 μm.

- The total energy flux emitted across all wavelengths is proportional to T^4.

II. Atmospheric Molecules

Important gases in the Earth's Atmosphere
(Note: Influence not necessarily proportional to % by volume, such as H_2O!)

• TABLE 1.1

Composition of the Atmosphere Near the Earth's Surface

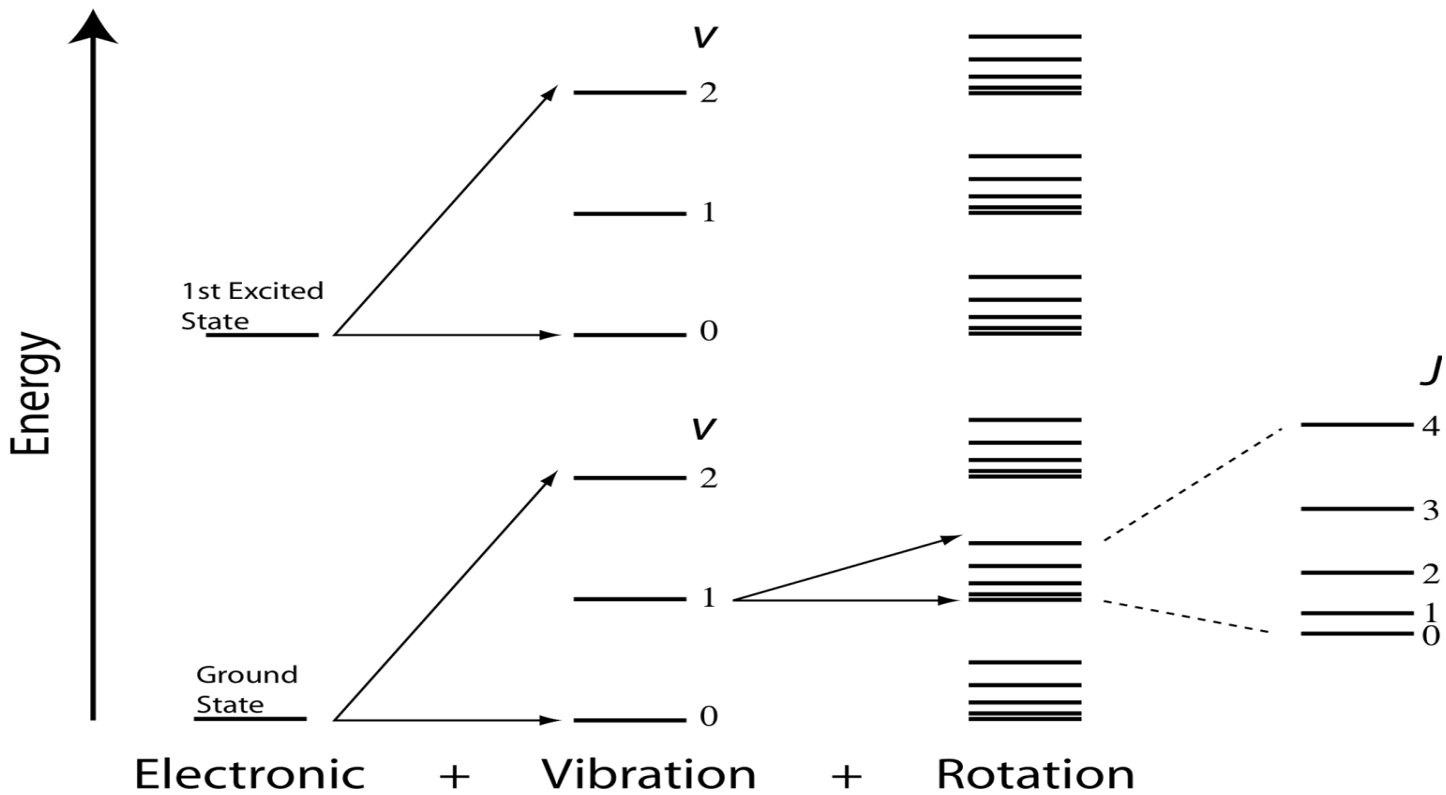
PERMANENT GASES			VARIABLE GASES			
Gas	Symbol	Percent (by Volume) Dry Air	Gas (and Particles)	Symbol	Percent (by Volume)	Parts per Million (ppm)*
Nitrogen	N_2	78.08	Water vapor	H_2O	0 to 4	
Oxygen	O_2	20.95	Carbon dioxide	CO_2	0.038	380*
Argon	Ar	0.93	Methane	CH_4	0.00017	1.7
Neon	Ne	0.0018	Nitrous oxide	N_2O	0.00003	0.3
Helium	He	0.0005	Ozone	O_3	0.000004	0.04†
Hydrogen	H_2	0.00006	Particles (dust, soot, etc.)		0.000001	0.01–0.15
Xenon	Xe	0.000009	Chlorofluorocarbons (CFCs)		0.00000002	0.0002

*For CO_2, 380 parts per million means that out of every million air molecules, 380 are CO_2 molecules.
†Stratospheric values at altitudes between 11 km and 50 km are about 5 to 12 ppm.

© 2007 Thomson Higher Education

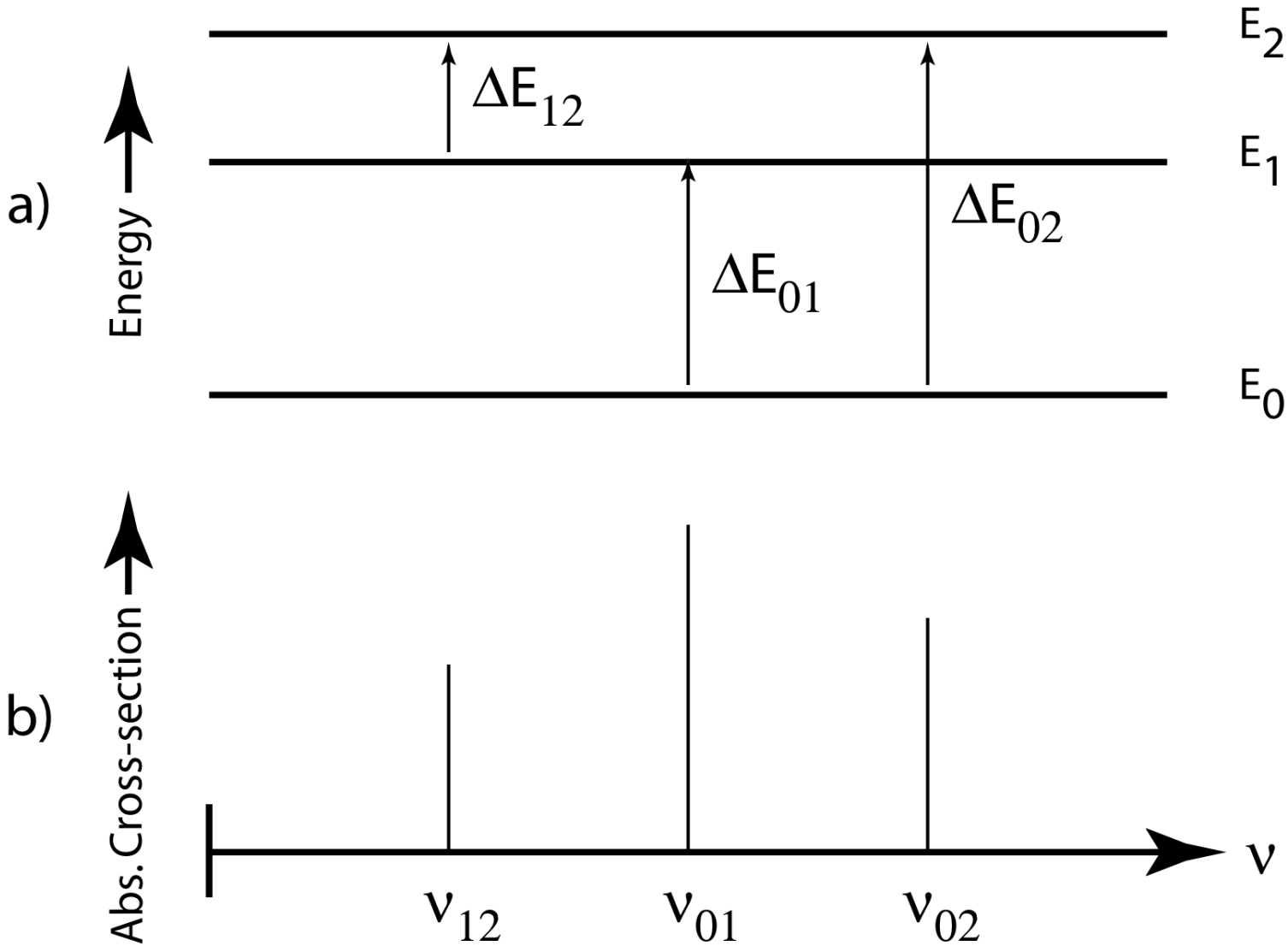
III. Molecular Absorptions

Energy Levels of Molecular States



III. Molecular Absorptions

Transitions of Molecular States by Absorbing Radiations



III. Molecular Absorptions
Line Broadening

Natural Broadening:
Mechanism: Heisenberg uncertainty Principle
Line Shape: Lorentz
Character: Independent of the environment
Doppler Broadening:
Mechanism: Molecular Thermal Motion
Line Shape: Gaussian

$$f(\omega - \omega_{fi}) = \frac{1}{\sqrt{\pi}\Delta\omega} \cdot e^{-\left(\frac{\omega-\omega_{fi}}{\Delta\omega}\right)^2} \qquad \Delta\omega \propto \frac{\omega_{fi}}{c}\sqrt{\frac{2KT}{m}}$$

Characters: Half-width $\Delta\omega$ depends on ω_{fi} and T. It is important in the mesosphere.

Pressure Broadening:
Mechanism: Interactions between molecules
Line Shape: Lorentz

$$f(\omega - \omega_{fi}) = \frac{1}{\pi} \cdot \frac{\gamma_{fi}}{(\omega - \omega_{fi} - \delta_{fi})^2 + \gamma_{fi}^2}$$

Characters: The half-width δ_{fi} and shift γ_{fi} are proportional to the pressure and depend on T.

$$\gamma = \gamma_0 P \left(\frac{T_0}{T}\right)^n$$

It is very important for the troposphere and lower stratosphere.

III. Molecular Absorptions
Line Broadening

In summary,

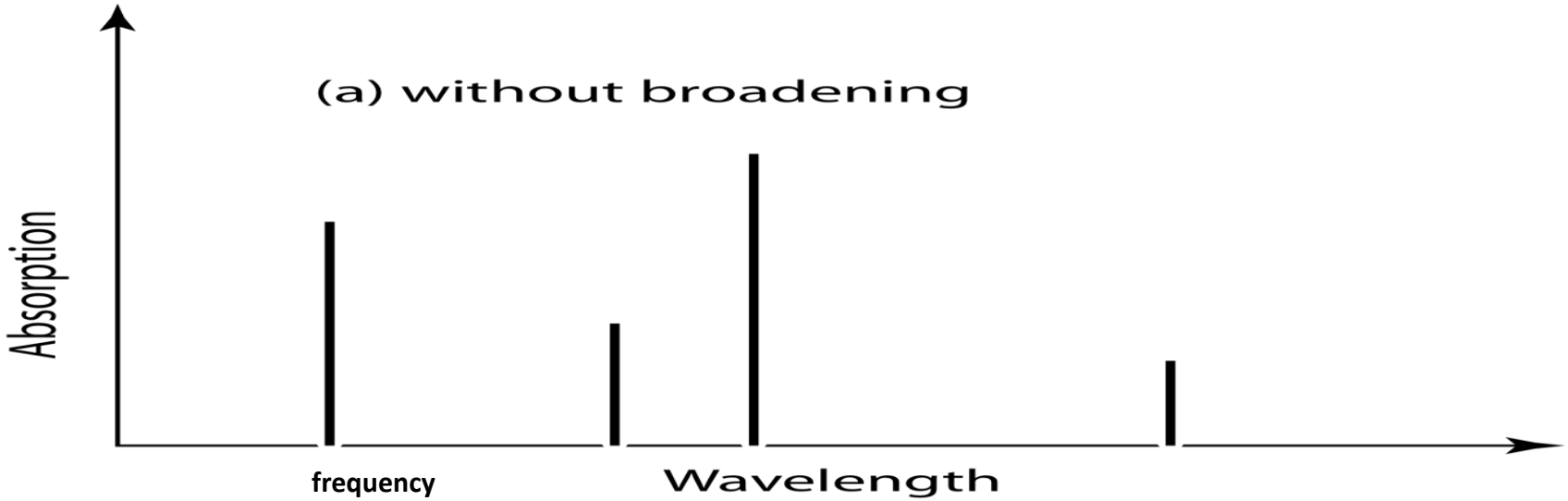
(1) The natural broadening is negligible.

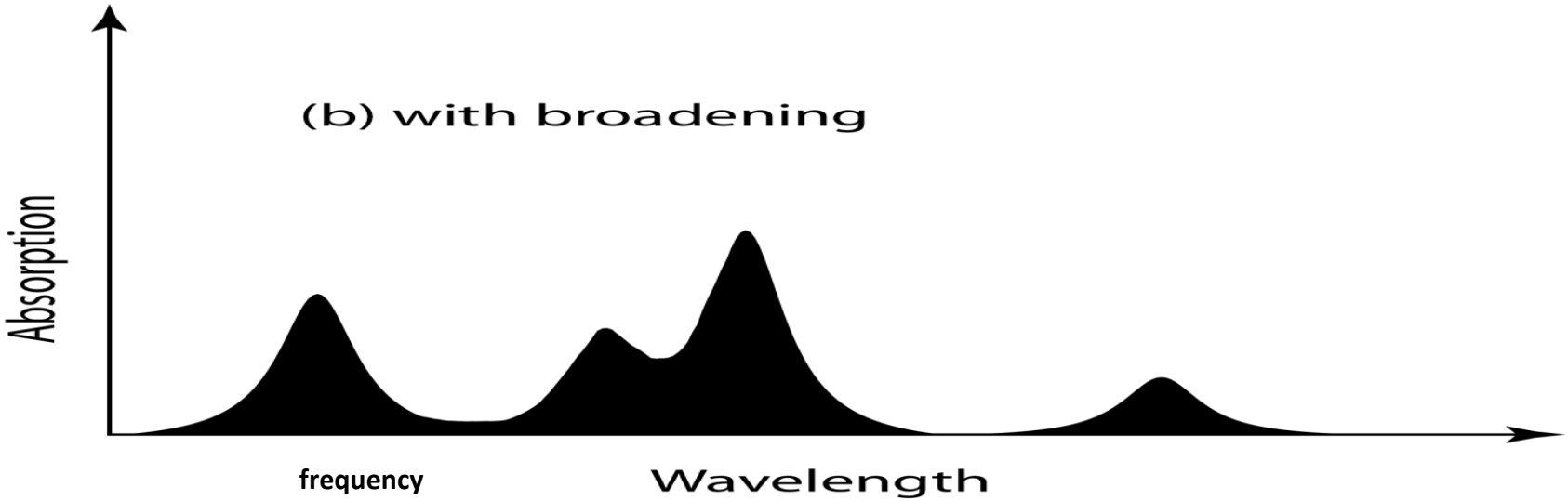
(2) The pressure broadening is dominated for a lower frequency region and at lower altitude.

(3) When both the Doppler and pressure broadenings have to be considered, the Voigt line shape is introduced.

$$f_{voi}(\omega, \omega_{fi}) = \int_{-\infty}^{\infty} f_L(\omega, \omega') f_D(\omega', \omega_{fi}) d\omega' .$$

III. Molecular Absorptions
Line Broadening

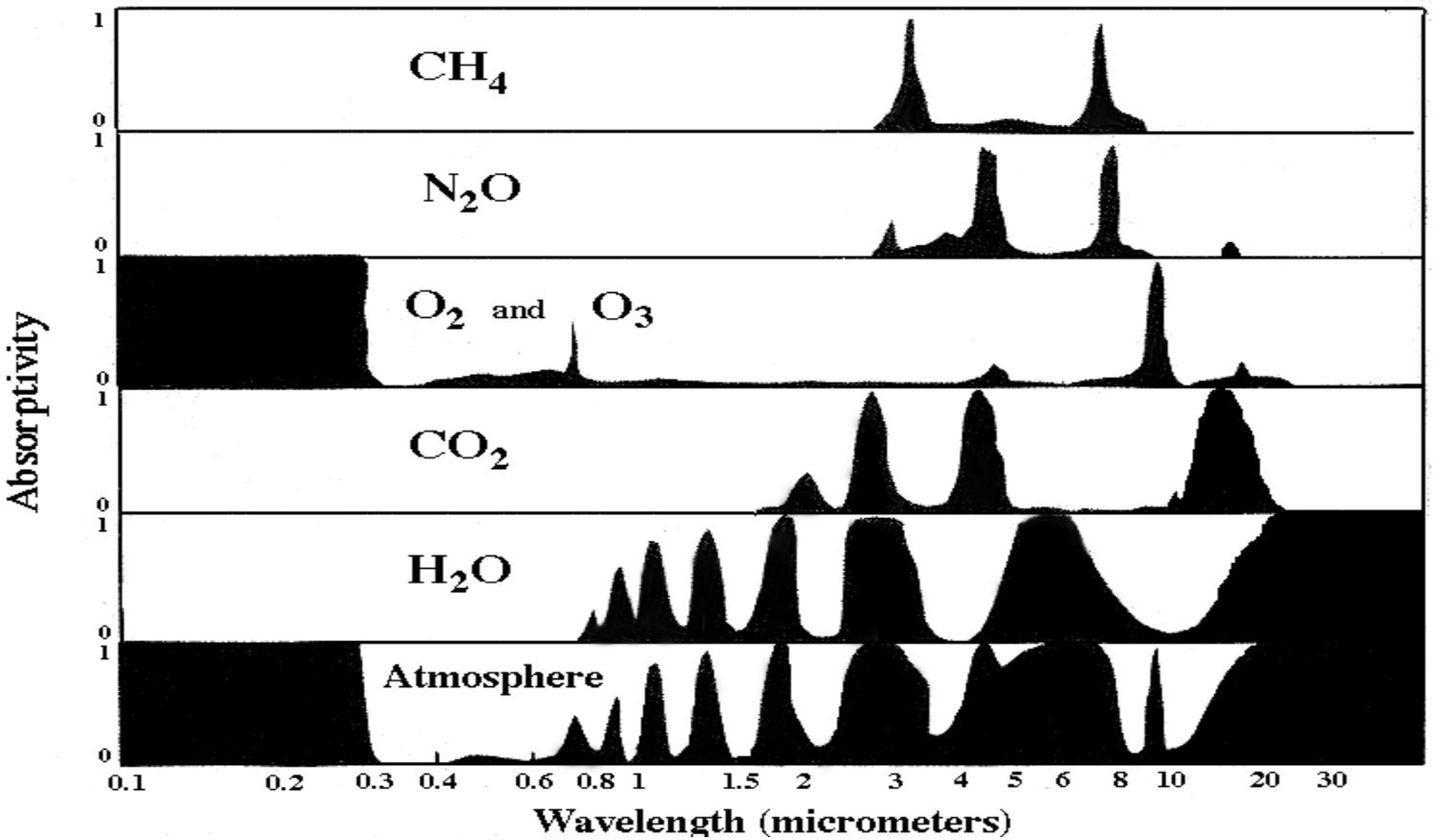




III. Molecular Absorptions

Line Intensities $S_{fi}(T)$, Positions ω_{fi}, and Spectra $\alpha(\omega)$

$$\alpha(\omega) = \sum_k \alpha_k(\omega) = \sum_k \sum_{fi} S_{fi}(T) f(\omega - \omega_{fi})$$



IV. HITRAN DATABASE

HITRAN is an acronym for **hi**gh-resolution **tran**smission molecular absorption database. HITRAN is a compilation of spectroscopic parameters that use to predict and simulate the transmission and emission of light in the atmosphere. The current version contains 7,400,447 spectral lines for 47 different molecules, incorporating 120 isotopologues.

The HITRAN database has been used by the radiation group at GISS for decades.

The HITRAN compilation, and its analogous database HITEMP (high-temperature spectroscopic absorption parameters), are now being developed at the Atomic and Molecular Physics Division, Harvard-Smithsonian Center for Astrophysics under the continued direction of Dr. Laurence S. Rothman.

IV. HITRAN DATABASE

Samples of Molecules

Diatomics

Molecule	ν_1
CO (carbon monoxide)	2143
O_2 (oxygen)	1556
NO (nitric oxide)	1875
OH (hydroxyl radical)	3569
HF (hydrogen fluoride)	3961
HCl (hydrogen chloride)	2886
HBr (hydrogen bromide)	2559
HI (hydrogen iodide)	2230
ClO (chlorine monoxide)	884
N_2 (nitrogen)	2330
NO^+ (NO cation)	2344

Triatomics

Molecule	ν_1	ν_2	ν_3
H_2O (water)	3657	1595	3756
CO_2 (carbon dioxide)	1388	667	2349
O_3 (ozone)	1103	701	1042
N_2O (nitrous oxide)	2224	589	1285
SO_2 (sulfur dioxide)	1152	518	1362
NO_2 (nitrogen dioxide)	1320	750	1617
OCS (carbonyl sulfide)	859	520	2062
HOCl (hypochlorous acid)	3609	1239	724
HCN (hydrogen cyanide)	3311	713	2097
H_2S (hydrogen sulfide)	2615	1183	2626
HO_2 (hydroperoxyl radical)	3436	1392	1098
HOBr (hydrobromous acid)	3615	1163	620

IV. HITRAN DATABASE

Molecule	Isotopo-logue[a]	Abundance	HITRAN2012 Spectral Coverage (cm^{-1})	HITRAN2008 Spectral Coverage (cm^{-1})	HITRAN2012 Number of Transitions	HITRAN2008 Number of Transitions
(1) **H_2O**	161	$9.973\ 10^{-1}$	**0 - 25711**	0 - 25233	**142 045**	37 432
	181	$1.999\ 10^{-3}$	**0 - 19918**	0 - 14519	**39 903**	9 753
	171	$3.719\ 10^{-4}$	**0 - 19946**	10 - 14473	**27 544**	6 992
	162	$3.107\ 10^{-4}$	**0 - 22708**	0 - 22708	**13 237**	13 238
	182	$6.230\ 10^{-7}$	**0 - 3825**	0 - 3825	**1 611**	1 611
	172	$1.158\ 10^{-7}$	**1234 - 1599**	1234 - 1599	**175**	175
(2) **CO_2**	626	$9.842\ 10^{-1}$	**345 - 12785**	352 - 12785	**169 292**	128 170
	636	$1.106\ 10^{-2}$	**406 - 12463**	438 - 12463	**70 611**	49 777
	628	$3.947\ 10^{-3}$	**0 - 9558**	0 - 11423	**116 482**	79 958
	627	$7.340\ 10^{-4}$	**0 - 9600**	0 - 8271	**72 525**	19 264
	638	$4.434\ 10^{-5}$	**489 - 6745**	489 - 6745	**26 737**	26 737
	637	$8.246\ 10^{-6}$	**583 - 6769**	583 - 6769	**2 953**	2 953
	828	$3.957\ 10^{-6}$	**491 - 8161**	491 - 8161	**7 118**	7 118
	827	$1.472\ 10^{-6}$	**626 - 5047**	626 - 5047	**821**	821
	727[b]	$1.368\ 10^{-7}$	**535 - 6933**	c	**5187**	c
	838[b]	$4.446\ 10^{-8}$	**4599 - 4888**	4599 - 4888	**121**	121
(3) **O_3**	666	$9.929\ 10^{-1}$	**0 - 6997**	0 - 5787	**261 886**	249 456
	668	$3.982\ 10^{-3}$	**0 - 2768**	0 - 2768	**44 302**	44 302
	686	$1.991\ 10^{-3}$	**1 - 2740**	1 - 2740	**18 887**	18 887
	667	$7.405\ 10^{-4}$	**0 - 2122**	0 - 2122	**65 106**	65 106
	676	$3.702\ 10^{-4}$	**0 - 2101**	0 - 2101	**31 935**	31 935

IV. HITRAN DATABASE
Line-by-line Parameters

Parameter	Field size	Definition
Mol	I2	**Molecule number**
Iso	I1	**Isotopologue no.**(1 = most abundant, 2 = second most abundant, …)
ν_{if}	F12.6	**Transition wavenumber in vacuum** [cm^{-1}]
S_{if}	E10.3	**Intensity** [cm^{-1}/(molecule·cm^{-2}) @ 296K]
A_{if}	E10.3	**Einstein *A*-coefficient** [s^{-1}]
γ_{air}	F5.4	**Air-broadened half-width** (HWHM) [cm^{-1}/atm @ 296K]
γ_{self}	F5.4	**Self-broadened half-width** (HWHM) [cm^{-1}/atm @ 296K]
E''	F10.4	**Lower-state energy** [cm^{-1}]
n_{air}	F4.2	**Temperature-dependence coefficient of γ_{air}**
δ_{air}	F8.6	**Air pressure-induced shift** [cm^{-1}/atm @ 296K]
v', v''	2A15	**Upper and Lower "global" quanta**
q', q''	2A15	**Upper and Lower "local" quanta**
ierr	6I1	**Uncertainty indices for ν_{if}, S_{if}, γ_{air}, γ_{self}, n_{air}, δ_{air}**
iref	6I2	**Reference pointers for ν_{if}, S_{if}, γ_{air}, γ_{self}, n_{air}, δ_{air}**
*	A1	**Flag for line-coupling algorithm**
g', g''	2F7.1	**Upper and Lower statistical weights**

160-character total

IV. HITRAN DATABASE
Uncertainties of the Spectroscopic Parameters

1. **Line Position: Excellent**. It has been claimed that in the pure rotational band, 22% of lines are within [0.001 cm^{-1}, 0.01cm^{-1}]; 18% within [0.0001 cm^{-1}, 0.001cm^{-1}]; 60 % within [0.00001 cm^{-1}, 0.0001cm^{-1}].
2. **Line Strength**: **Good**. It has been claimed that 33 % of lines are within [10 %, 20 %]; 67% within [5 %, 10 %].
3. **Air-Broadened Half-width**: **Reasonable Good**. It has been claimed that 34 % of lines are within [10 %, 20 %] and 60 % within [5 %, 10 %].
4. **Self-Broadened Half-width**: **Not good**. It has been claimed that 88 % of lines are from average or estimation.
5. **Temperature exponent**: **Poor**. It has been claimed that 33 % of lines are larger than 20 %; 64 % are within [10 %, 20 %].
6. **Induced Shift**: **Poor**. It has been claimed that 17 % of lines are larger than 1.0 cm^{-1} or unreported. 80% are within [0.001 cm^{-1}, 0.01cm^{-1}].

IV. HITRAN DATABASE

Uncertainties of the Width, Temperature Exponent, and Shift

For the width, temperature exponent, and shift, their accuracies can't meet users' requirements. These situations would remain for a while.

Measurements:

It is difficult to perform reliable measurements and there are too many lines.

Theories:

(1) Full quantum calculations are not feasible except for simpler molecular systems. After **Dr. S. Green at GISS** suddenly passed away in 1995, there are lack of progresses in developing the **MOLSCAT** code.

(2) Semi-classical calculations have played an important role in providing these parameters. However, the theory is based on several assumptions whose applicability have not been thoroughly justified. For linear molecular systems where the full quantum calculations are available, the semi-classically calculated half-widths are poor.

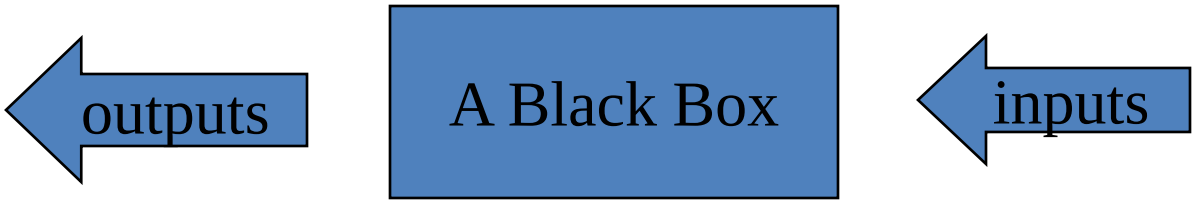
(3) Empirical formulas are widely used to provide these data, but there are lack of physically sound guides in developing these formulas.

V. Why and How the Spectroscopic Parameters vary with H_2O Lines

Basic Idea in answering these questions

- **How to improve the accuracies of these parameters? One solution is to understand why and how the spectroscopic parameters vary with the H_2O lines of interest.**

- **By considering a whole system consisting of one absorber H_2O molecule, bath molecules, and electromagnetic fields as a black box, its inputs and outputs are the H_2O lines of interest and their spectroscopic parameters.**



- **Main assumptions: (1) The outputs depend on the inputs. (2) Identical inputs should yield identical outputs. (3) Similar inputs should yield similar outputs.**

- **The inputs are the energy levels and wave functions associated with the initial and final states of the H_2O lines.**

V. Why and How the Spectroscopic Parameters vary with H_2O Lines

Properties of the energy levels and wave functions of H_2O states

(1) The H_2O states $J_{K_aK_c}$ can be divided as the A or C type. The A type states are k_a values are closer to J and the C type states are k_c are closer to J.
In cases of k_a and k_c are comparative, there is a bias to consider the states as the A type because H_2O is closer to a prolate top.

(2) Two states of $J_{J-\Delta,\Delta}$ and $J_{J-\Delta,\Delta+1}$ in the A type are paired and two states of $J_{\Delta,J-\Delta}$ and $J_{\Delta+1,J-\Delta}$ in the C type are paired. For each of the A and C types, **pairs are categorized into different groups according to the Δ values**.
Main Features: The paired states have almost identical energy levels. For different pairs within the same groups, their energy levels vary smoothly as J varies and these variation patterns are well organized.

(3) The H_2O wave functions are given in terms of expansion coefficients $U^J_{K\tau}$ over the symmetric top wave functions |JKM>, $|J\tau M>=\sum_K U^J_{K\tau}|JKM>$.
Main Features: The coefficients of two paired states with high J have almost identical magnitudes. For different pairs within the same groups, patterns of their coefficients are very similar and the patterns shift smoothly as J varies.

Pair Identity and Smooth Variation of the Energy Levels of H_2O

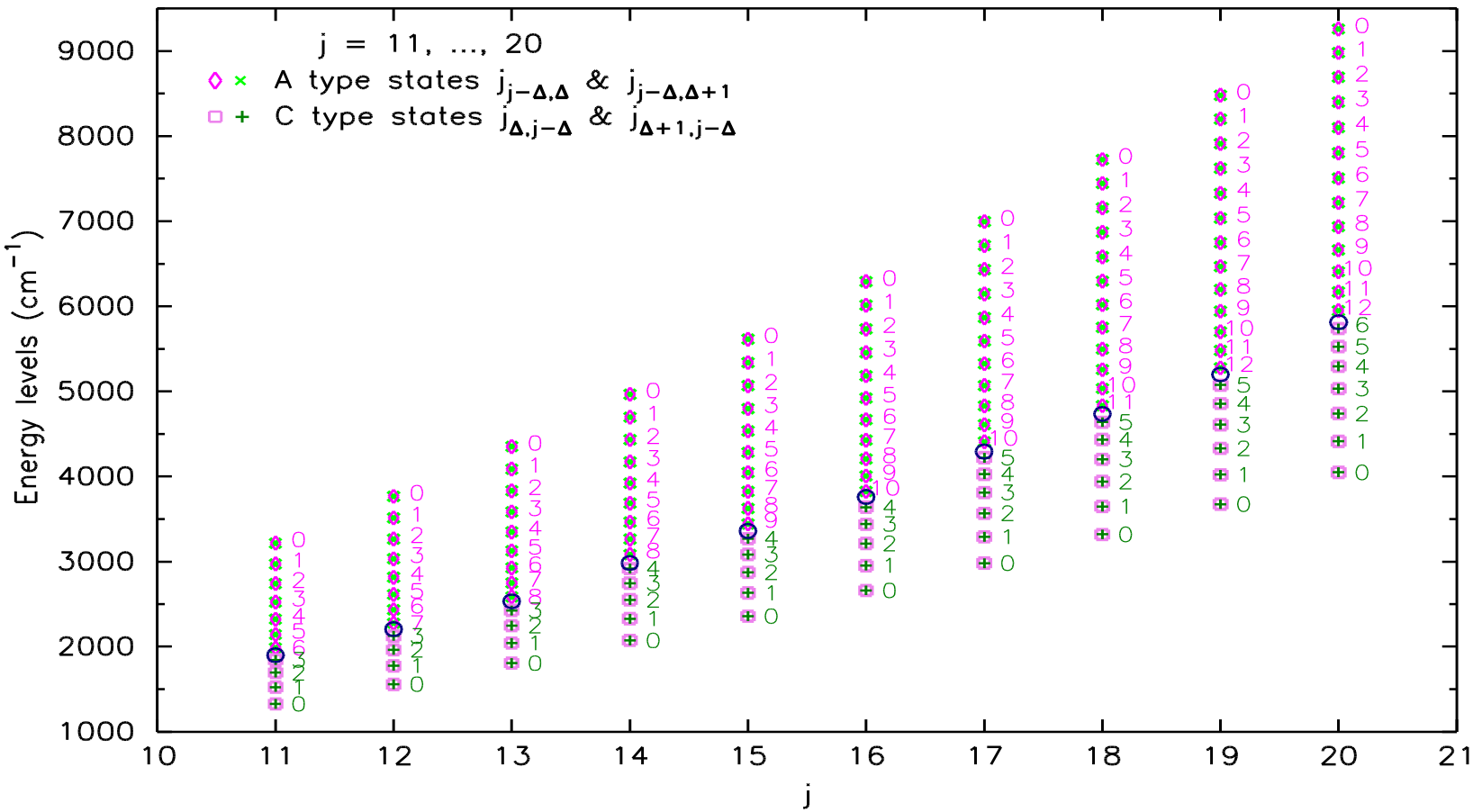
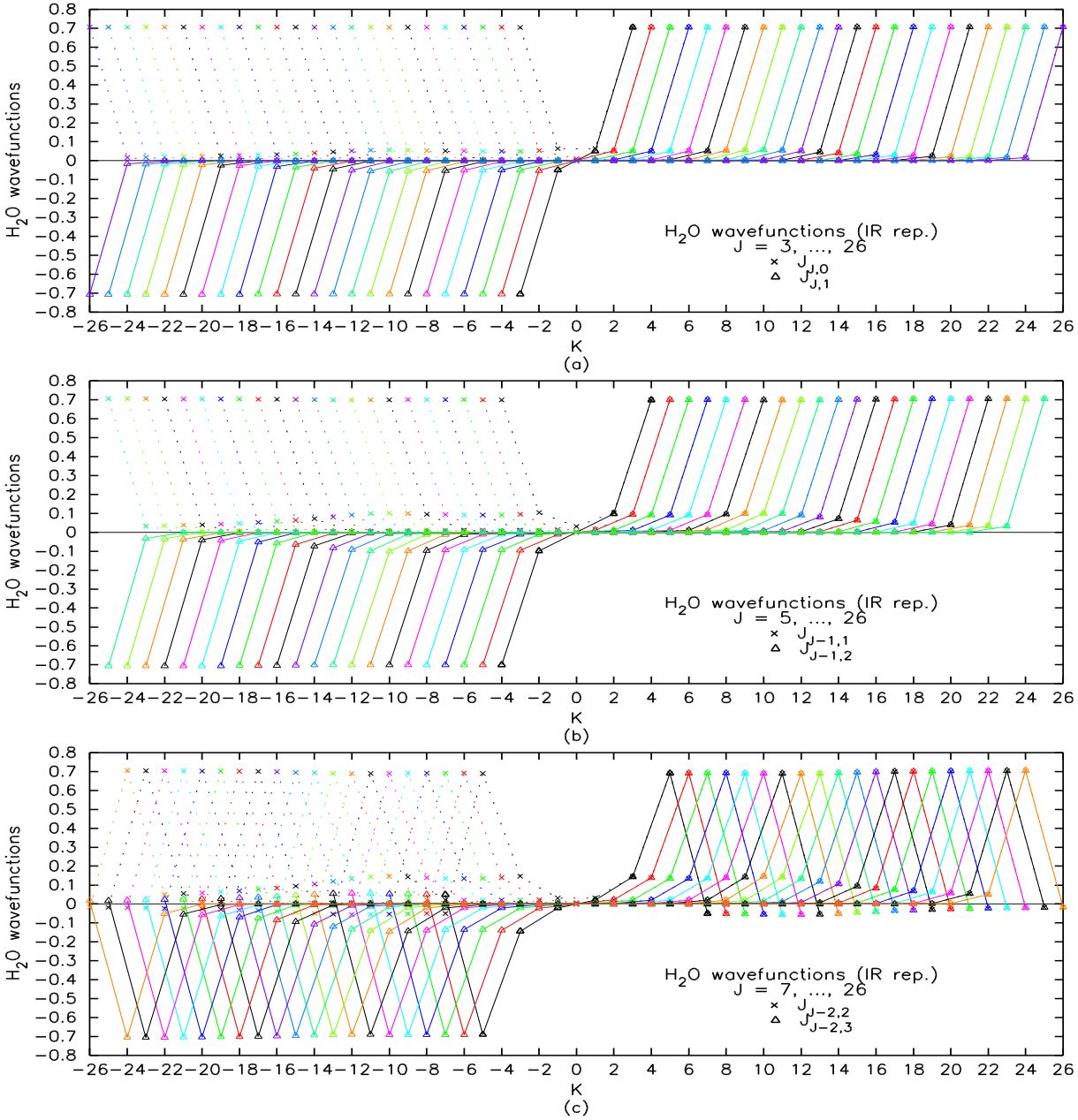


Fig. 1 A plot to show energy levels of H_2O states with j = 11, ⋯, 20 in the vibrational ground state. For the groups of $j_{j-\Delta,\Delta}$ and $j_{j-\Delta,\Delta+1}$ in the A type with Δ = 0, 1, ⋯, their energy levels are plotted by ◊ and × and for the groups of $j_{\Delta,j-\Delta}$ and $j_{\Delta+1,j-\Delta}$ in the C type with Δ = 0, 1, ⋯, their energy levels are plotted by □ and +. Meanwhile, their corresponding values of Δ are presented on the right side of the symbols.

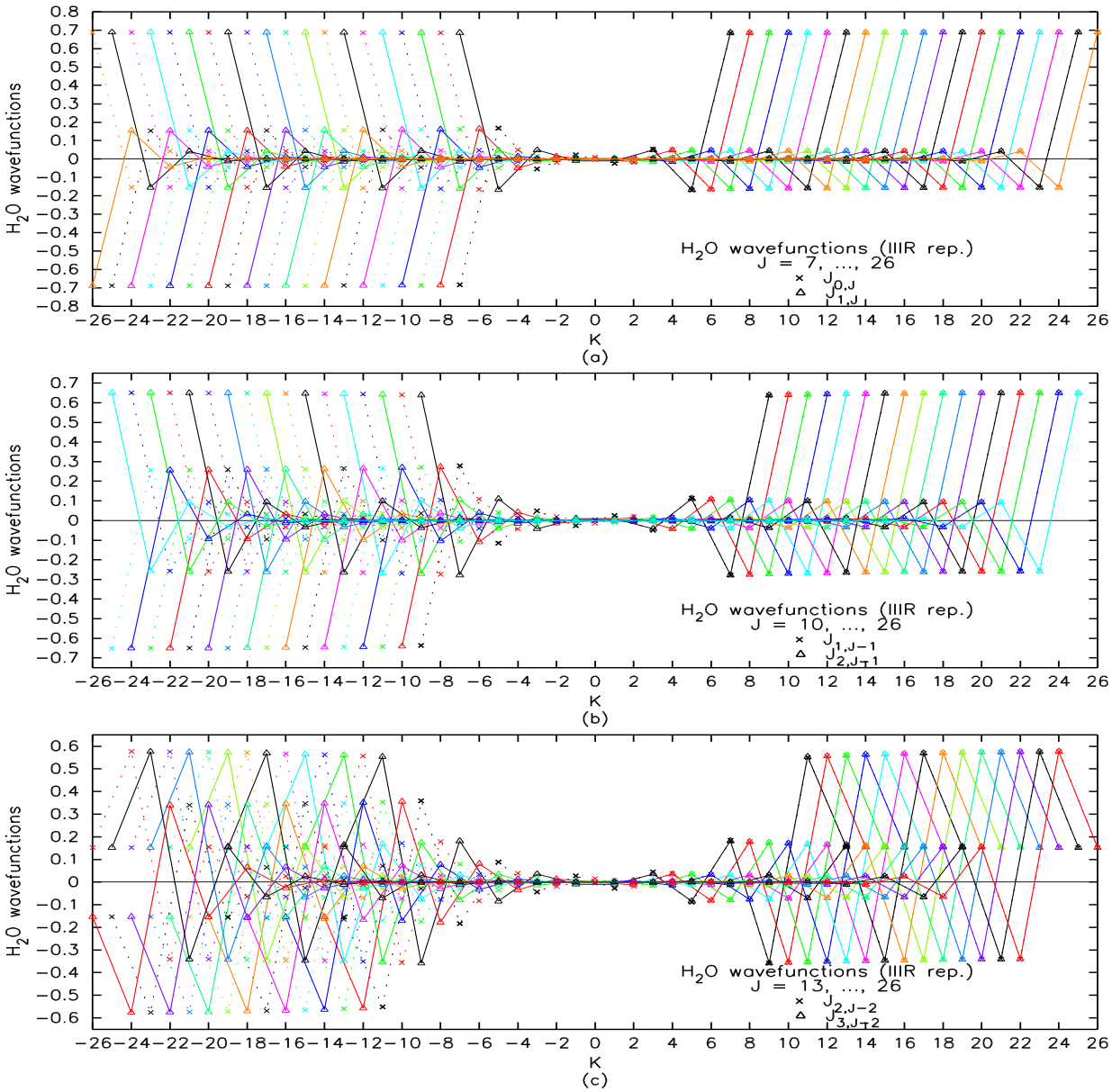
VI. Screening HITRAN H_2O Database with the Rules



Pair Identity and Smooth Variation of the Wave Functions in A type

Fig. 2 A plot to show properties of H_2O wave functions in the IR representation for three sets of paired states with the A type : $\{J_{J,0},J_{J,1}\}$, $\{J_{J-1,1},J_{J-1,2}\}$, and $\{J_{J-2,2},J_{J2,3}\}$.

VI. Screening HITRAN H_2O Database with the Rules



Pair Identity and Smooth Variation of the Wave Functions in C type

Fig. 3 A plot to show properties of H_2O wave functions in the III R representation for three sets of pairs of states with the C type: $\{J_{0,J},J_{1,J}\}$, $\{J_{1,J-1},J_{2,J-1}\}$, and $\{J_{2,J-2},J_{3,J-2}\}$.

:.

Applicability Boundaries of the Two Rules for Different Groups

Gr.	$J_{J,0}$ $J_{J,1}$ A	$J_{J-1,1}$ $J_{J-1,2}$ A	$J_{J-2,2}$ $J_{J-2,3}$ A	$J_{J-3,3}$ $J_{J-3,4}$ A	$J_{J-4,4}$ $J_{J-4,5}$ A	$J_{J-5,5}$ $J_{J-5,6}$ A	$J_{J-6,6}$ $J_{J-6,7}$ A	…	$J_{4,J-4}$ $J_{5,J-4}$ C	$J_{3,J-3}$ $J_{4,J-3}$ C	$J_{2,J-2}$ $J_{3,J-2}$ C	$J_{1,J-1}$ $J_{2,J-1}$ C	$J_{0,J}$ $J_{1,J}$ C
J_{bd}	3	5	7	9	10	12	14	…	19	16	13	10	7
ε (%)	0.63	0.66	0.49	0.32	1.08	0.60	0.32	…	1.2	0.55	1.13	1.00	0.60

In each of the groups, the pair identity and the smooth variation hold for J are above certain boundaries.

By introducing a numerical measure ε defined by

$$\varepsilon = \sum_K \left| |U^J_{K\tau_2}|^2 - |U^J_{K\tau_1}|^2 \right| / \sum_K |U^J_{K\tau_1}|^2,$$

where J, τ_1 and J, τ_2 are paired states and $\tau = K_a - K_c$.

The higher the J, the smaller the ε. By choosing ε is about 1 %, one can determine a boundary J_{bd} for each of the sets. The results is listed in Table. With the same Δ, states with the A type have lower boundaries.

V. Why and How the Spectroscopic Parameters vary with H_2O Lines

Categorizations of lines and discovery of two rules

- **With knowledge about how to categorize H_2O states, one categorizes the H_2O lines based on four combinations of their initial and final states' types (i.e., A-A, A-C, C-C, and C-A). The categorization is carried out for each of the P, Q, and R branches. The categorization formulas are listed in the table.**

- **After categorization, within individually defined groups, the inputs of lines have identity and similarity properties. Then, one should expects the outputs have similar properties too.**

- **Thus, two rules are established within individual groups. The pair identity rule: two paired lines whose J values are above certain boundaries have almost identical spectroscopic parameters. The smooth variation rule: for different pairs in the same groups, values of their spectroscopic parameters vary smoothly as their J values vary.**

- **The rules are valid with certain accuracy tolerance. The second rule becomes applicable more earlier than the first rule.**

- **By screening databases with these rules, one can pick up outliers (i.e., mistakes) and provide better results from smooth procedure. Thus, by full exploiting the current resources, one is able to significantly improve the accuracies.**

V. Why and How the Spectroscopic Parameters vary with H_2O Lines
The Line Categorization procedure

- Table 1 Categorizations of lines in the 000-000, 010-000, and 100-000 bands

Cases	The P and R branches	The Q branch
A-A	$\mathbf{j'}_{j'-\Delta',\Delta'+1} \leftarrow \mathbf{j''}_{j''-\Delta'',\Delta''}$ $\mathbf{j'}_{j'-\Delta',\Delta'} \leftarrow \mathbf{j''}_{j''-\Delta'',\Delta''+1}$ (Δ' & Δ'' in the same parity)	$\mathbf{j'}_{j'-\Delta',\Delta'} \leftarrow \mathbf{j''}_{j''-\Delta'',\Delta''}$ $\mathbf{j'}_{j'-\Delta',\Delta'+1} \leftarrow \mathbf{j''}_{j''-\Delta'',\Delta''+1}$ (Δ' & Δ'' in the opposite parity)
A-C	$\mathbf{j'}_{j'-\Delta',\Delta'+1} \leftarrow \mathbf{j''}_{\Delta'',j''-\Delta''}$ $\mathbf{j'}_{j'-\Delta',\Delta'} \leftarrow \mathbf{j''}_{\Delta''+1,j''-\Delta''}$	$\mathbf{j'}_{j'-\Delta',\Delta'} \leftarrow \mathbf{j''}_{\Delta'',j''-\Delta''}$ $\mathbf{j'}_{j'-\Delta',\Delta'+1} \leftarrow \mathbf{j''}_{\Delta''+1,j''-\Delta''}$
C-C	$\mathbf{j'}_{\Delta'+1,j'-\Delta'} \leftarrow \mathbf{j''}_{\Delta'',j''-\Delta''}$ $\mathbf{j'}_{\Delta',j'-\Delta'} \leftarrow \mathbf{j''}_{\Delta''+1,j''-\Delta''}$ (Δ' & Δ'' in the same parity)	$\mathbf{j'}_{\Delta',j'-\Delta'} \leftarrow \mathbf{j''}_{\Delta'',j''-\Delta''}$ $\mathbf{j'}_{\Delta'+1,j'-\Delta'} \leftarrow \mathbf{j''}_{\Delta''+1,j''-\Delta''}$ (Δ' & Δ'' in the opposite parity)
C-A	$\mathbf{j'}_{\Delta'+1,j'-\Delta'} \leftarrow \mathbf{j''}_{j''-\Delta'',\Delta''}$ $\mathbf{j'}_{\Delta',j'-\Delta'} \leftarrow \mathbf{j''}_{j''-\Delta'',\Delta''+1}$	$\mathbf{j'}_{\Delta',j'-\Delta'} \leftarrow \mathbf{j''}_{j''-\Delta'',\Delta''}$ $\mathbf{j'}_{\Delta'+1,j'-\Delta'} \leftarrow \mathbf{j''}_{j''-\Delta'',\Delta''+1}$

V. Why and How the Spectroscopic Parameters vary with H_2O Lines
The Line Categorization procedure

- Table 2 Categorizations of lines in the 001-000 and 011-000 bands

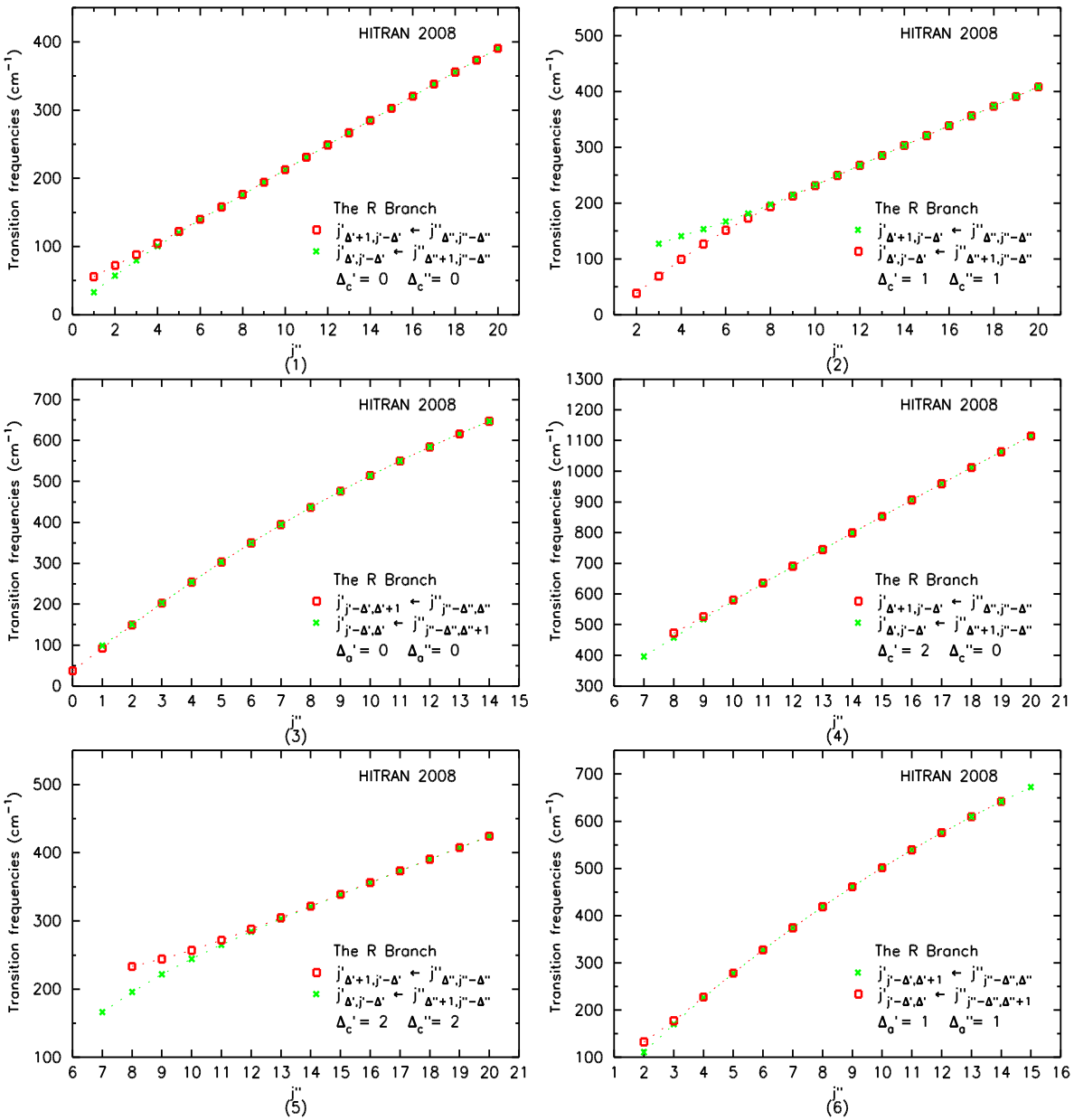
Cases	The P and R branches	The Q branch
A-A	$j'_{j'-\Delta',\Delta'} \leftarrow j''_{j''-\Delta'',\Delta''}$ $j'_{j'-\Delta',\Delta'+1} \leftarrow j''_{j''-\Delta'',\Delta''+1}$ (Δ' & Δ'' in the opposite parity)	$j'_{j'-\Delta',\Delta'+1} \leftarrow j''_{j''-\Delta'',\Delta''}$ $j'_{j'-\Delta',\Delta'} \leftarrow j''_{j''-\Delta'',\Delta''+1}$ (Δ' & Δ'' in the same parity)
A-C	$j'_{j'-\Delta',\Delta'} \leftarrow j''_{\Delta'',j''-\Delta''}$ $j'_{j'-\Delta',\Delta'+1} \leftarrow j''_{\Delta''+1,j''-\Delta''}$	$j'_{j'-\Delta',\Delta'+1} \leftarrow j''_{\Delta'',j''-\Delta''}$ $j'_{j'-\Delta',\Delta'} \leftarrow j''_{\Delta''+1,j''-\Delta''}$
C-C	$j'_{\Delta',j'-\Delta'} \leftarrow j''_{\Delta'',j''-\Delta''}$ $j'_{\Delta'+1,j'-\Delta'} \leftarrow j''_{\Delta''+1,j''-\Delta''}$ (Δ' & Δ'' in the same parity)	$j'_{\Delta'+1,j'-\Delta'} \leftarrow j''_{\Delta'',j''-\Delta''}$ $j'_{\Delta',j'-\Delta'} \leftarrow j''_{\Delta''+1,j''-\Delta''}$ (Δ' & Δ'' in the opposite parity)
C-A	$j'_{\Delta',j'-\Delta'} \leftarrow j''_{j''-\Delta'',\Delta''}$ $j'_{\Delta'+1,j'-\Delta'} \leftarrow j''_{j''-\Delta'',\Delta''+1}$	$j'_{\Delta'+1,j'-\Delta'} \leftarrow j''_{j''-\Delta'',\Delta''}$ $j'_{\Delta',j'-\Delta'} \leftarrow j''_{j''-\Delta'',\Delta''+1}$

V. Why and How the Spectroscopic Parameters vary with H_2O Lines Groups in main H_2O Bands

- Table 3 Groups of lines in the P, Q, and R branches for several bands

Bands HITRAN 2008	Total # lines	The P branch		The Q branch		The R branch	
		# lines	# group	# lines	# group	# lines	# group
000-000	1639	207	53	544	84	888	110
010-000	1903	586	106	661	121	656	113
100-000	1326	477	92	460	92	389	79
001-000	1626	557	98	527	102	542	102
011-000	1221	378	71	408	80	435	82

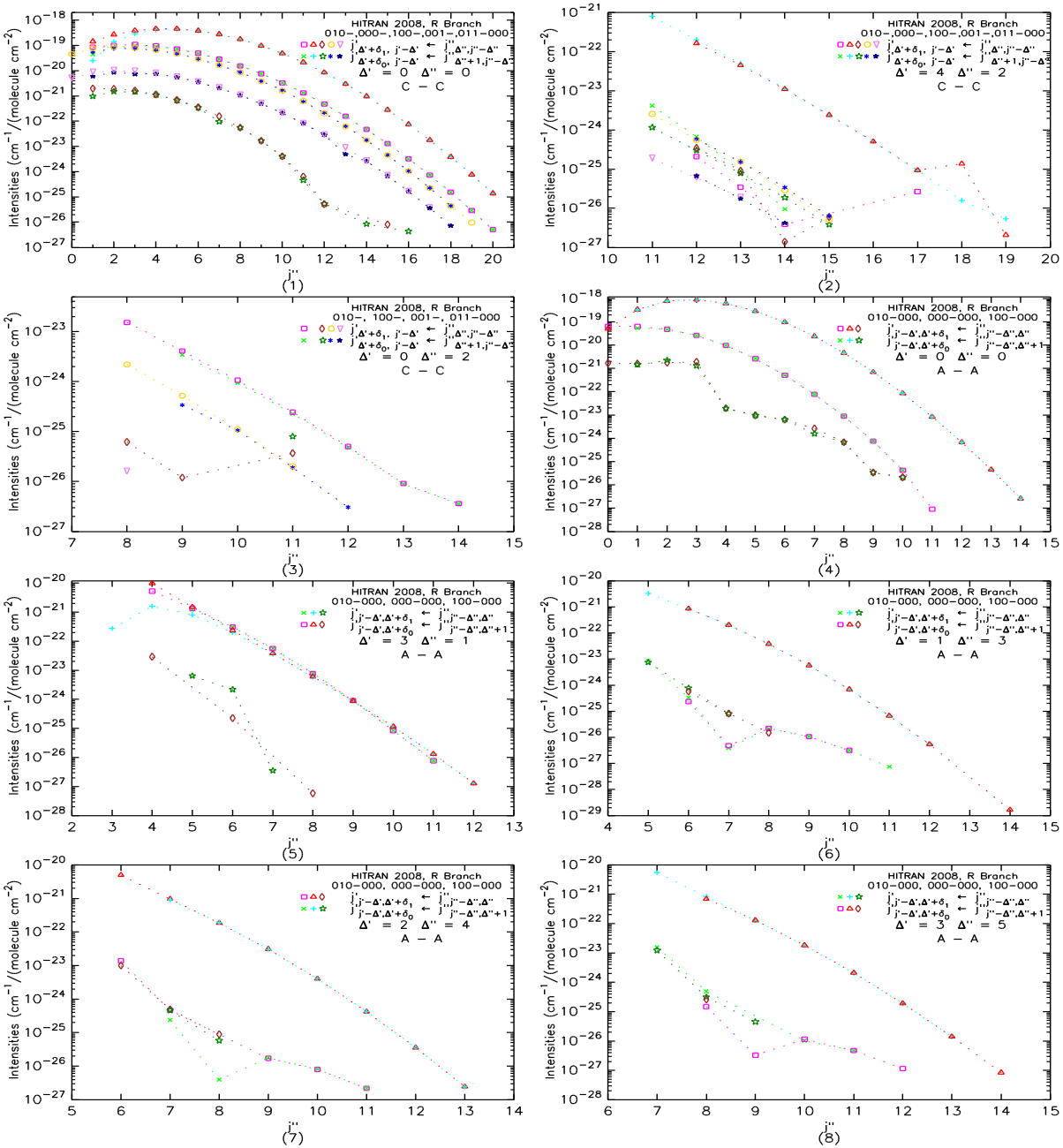
VI. Screening HITRAN H_2O Database with the Rules



The Line Position

Fig. 4 The Line Positions for six groups in the R branch of the pure rotational band. For paired lines, their positions listed in HITRAN 2008 are represented by □ and ×. The same symbols are connected by color lines.

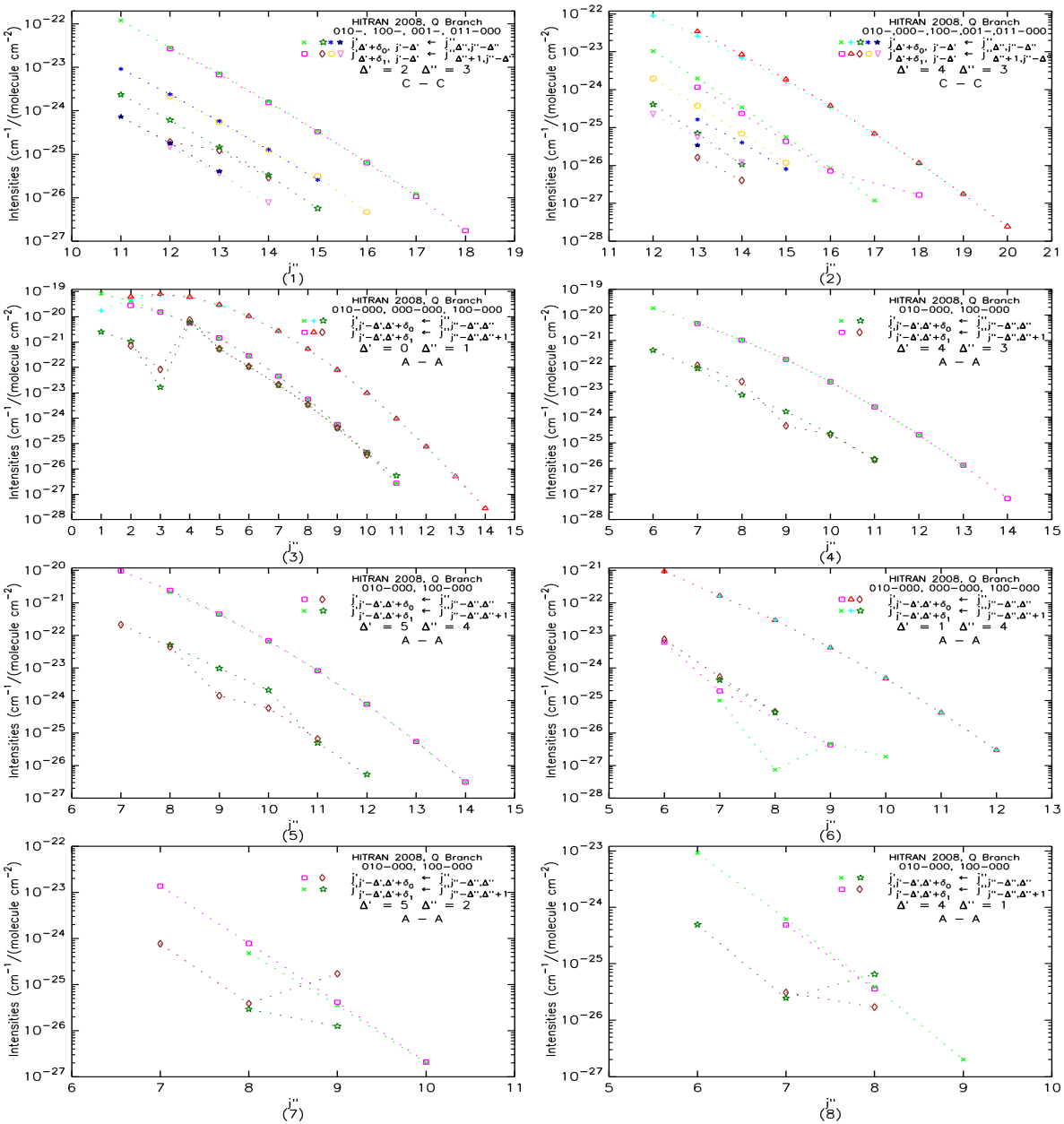
VI. Screening HITRAN H_2O Database with the Rules



The Line Intensity In the R Branch

Fig. 5 The intensities of paired lines in eight groups in the five main H_2O bands of 000-000, 010-000, 100-000, 001-000, and 011-000 are plotted. By excluding the spin degeneracy factor, for each of the two paired lines, their intensity values in the five bands are represented by five different symbols.

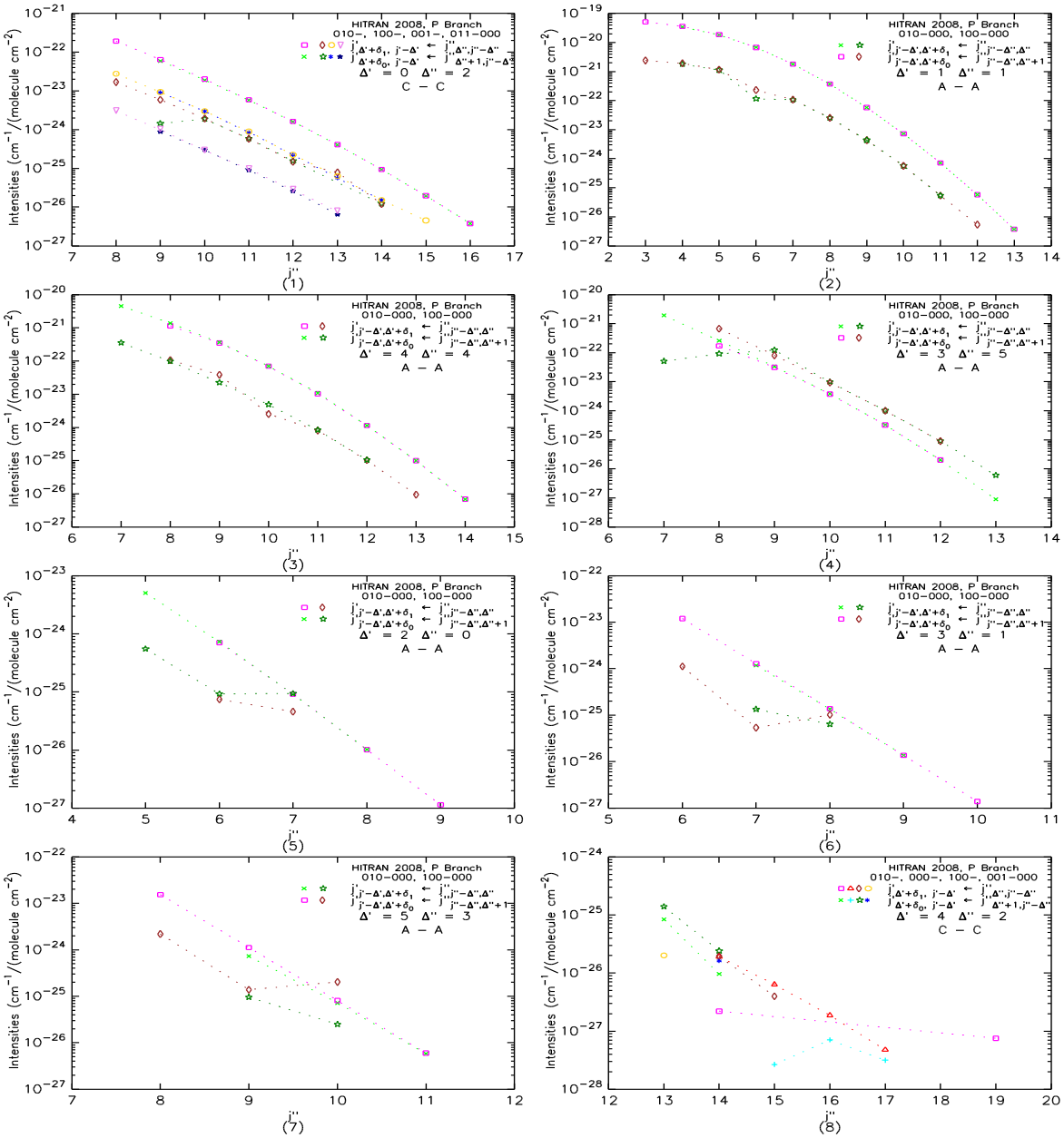
VI. Screening HITRAN H_2O Database with the Rules



The Line Intensity In the Q Branch

Fig. 6 The intensities of paired lines in eight groups in the five main H_2O bands of 000-000, 010-000, 100-000, 001-000, and 011-000 are plotted. By excluding the spin degeneracy factor, for each of the two paired lines, their intensity values in the five bands are represented by five different symbols.

VI. Screening HITRAN H_2O Database with the Rules



The Line Intensity In the P Branch

Fig. 7 The intensities of paired lines in eight groups in the five main H_2O bands of 000-000, 010-000, 100-000, 001-000, and 011-000 are plotted. By excluding the spin degeneracy factor, for each of the two paired lines, their intensity values in the five bands are represented by five different symbols.

VI. Screening HITRAN H_2O Database with the Rules

The Air-Broadened Half-width

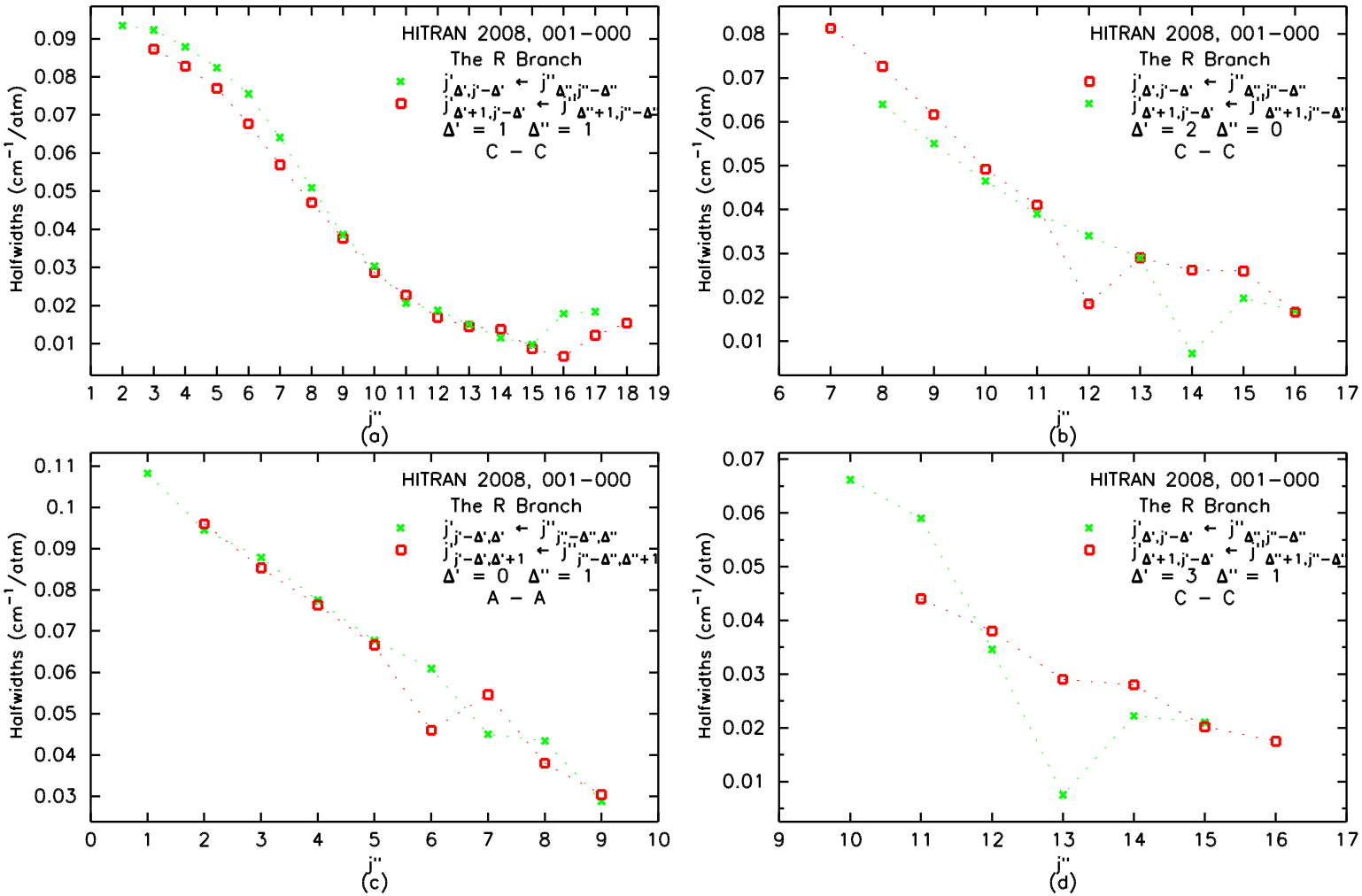


Fig. 8 Air-broadened half-widths (in units of cm^{-1}/atm) with the four groups in the R branch of the 001-000 band are plotted. Their values are represented by □ and ×.

VI. Screening HITRAN H_2O Database with the Rules

The Self-Broadened Half-width

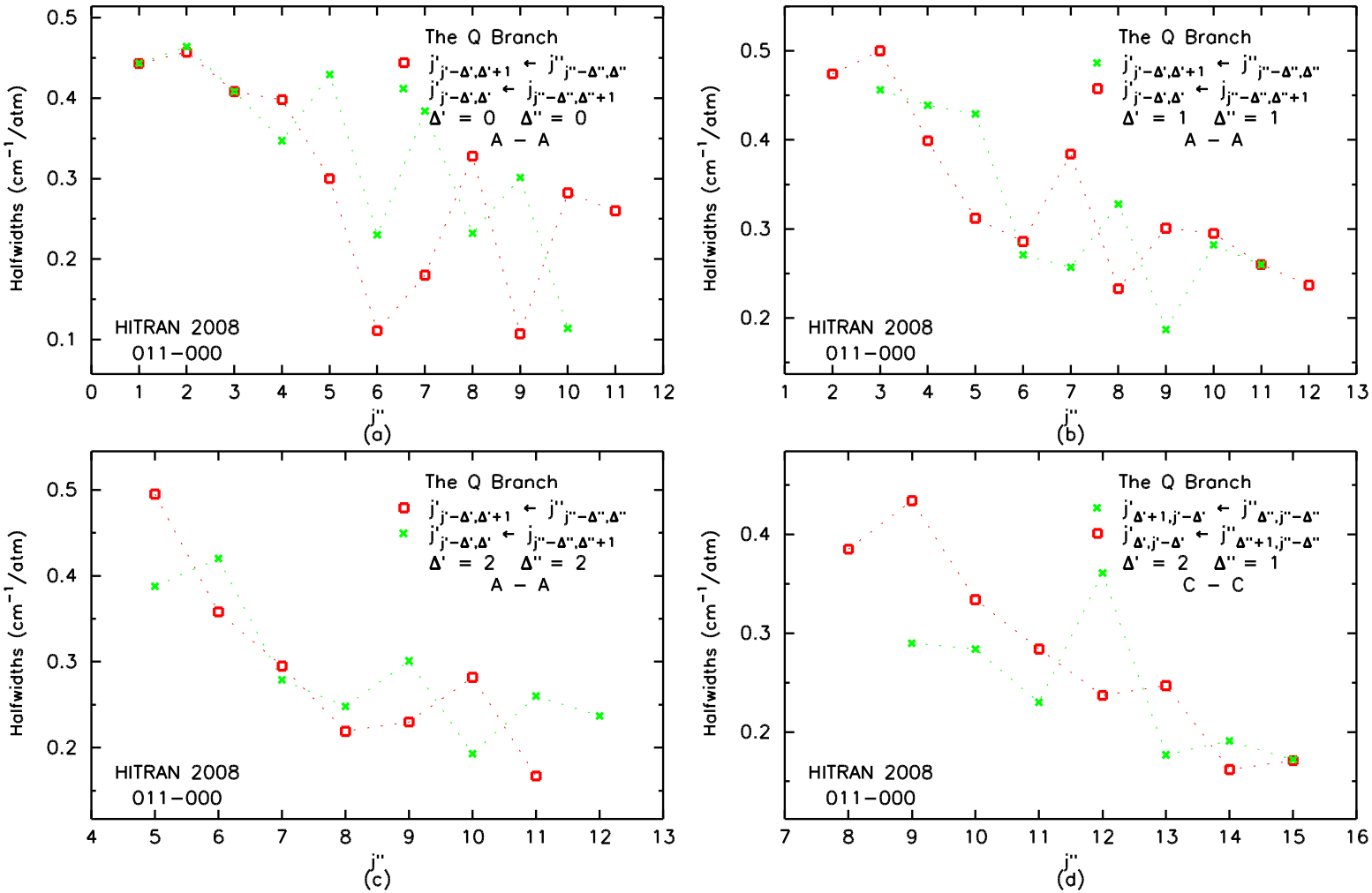
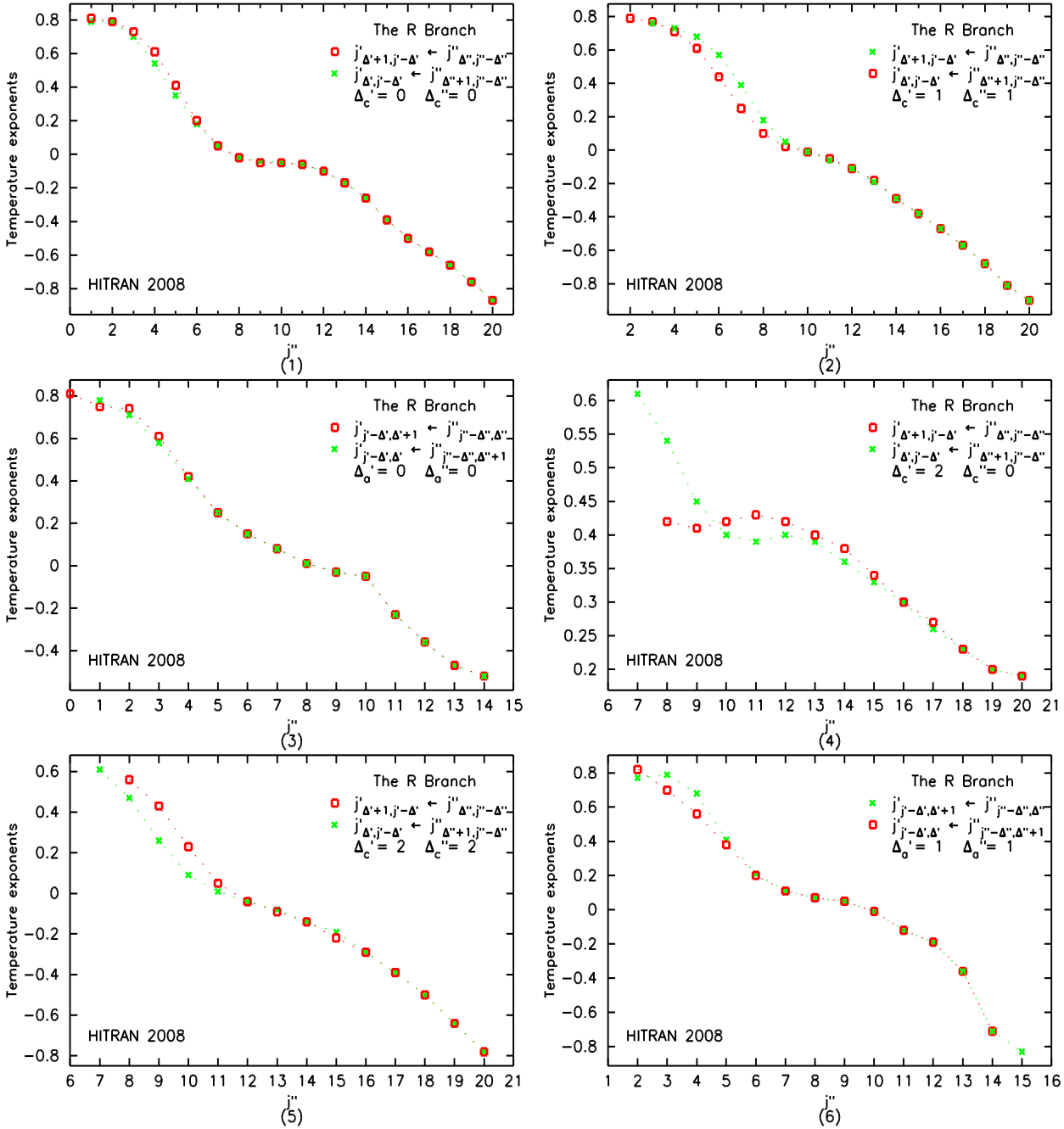


Fig. 9 Self-broadened half-widths (in units of cm^{-1}/atm) with the four groups in the R branch of the 011-000 band are plotted. Their values are represented by □ and ×.

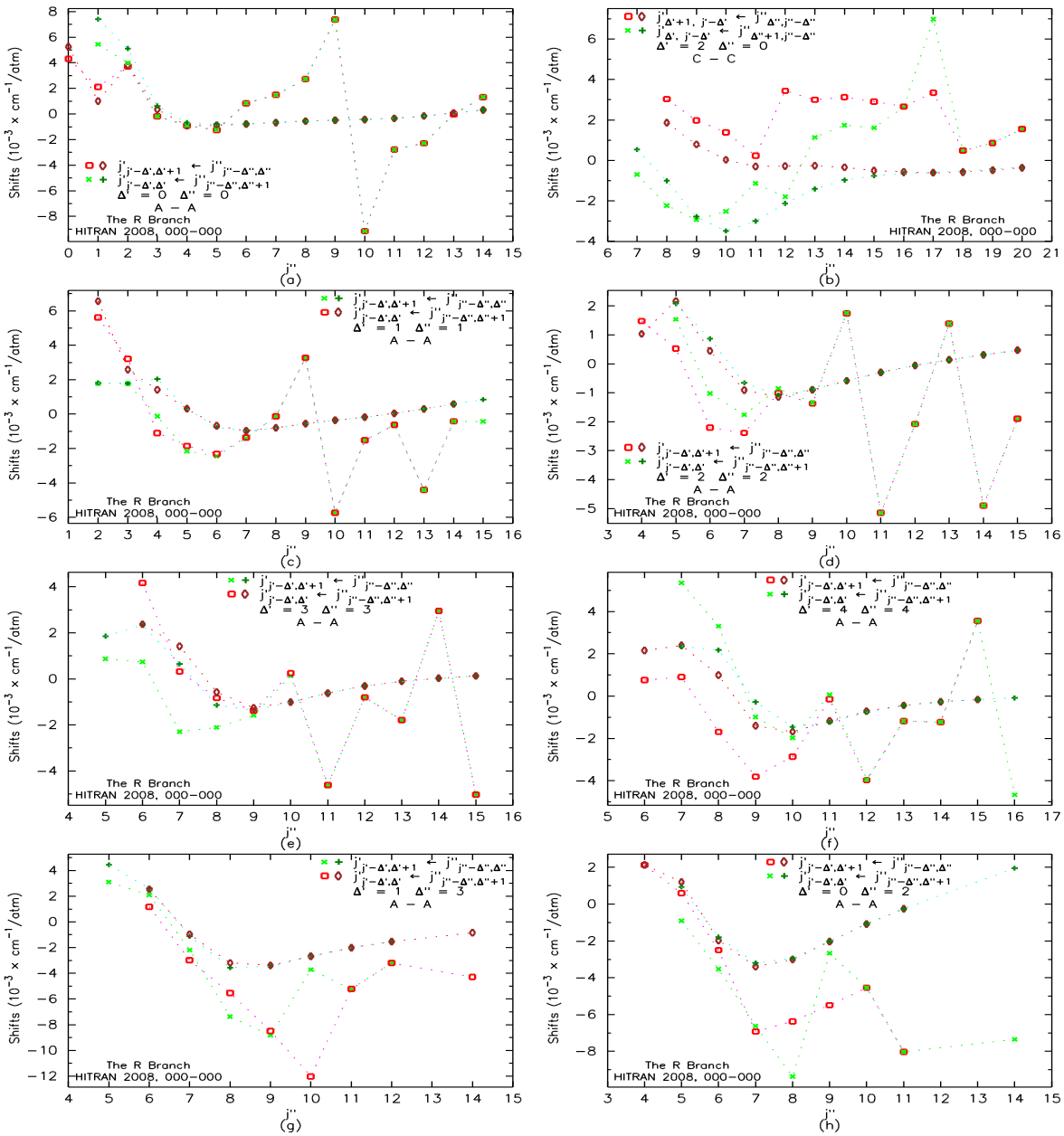
VI. Screening HITRAN H_2O Database with the Rules



The Temperature Exponent

Fig. 10 The same as Fig. 8 except for the temperature exponent in the 000-000 band.

VI. Screening HITRAN H_2O Database with the Rules



The Shift

Fig. 11 The induced shifts for eight specified groups in the R branch of the pure rotational band. The values listed in HITRAN 2008 are represented by □ and ×. Meanwhile, our calculated results are given by ◊ and +.

VI Screening HITRAN H_2O Database with the Rules

Our comments on the shifts listed in HITRAN 2008 and a outcome

Comments:

(1) The two rules are derived and established from the properties of the energy levels and wave functions of H_2O states. Thus, all the spectroscopic parameters involving high J states must follow the rules whether they are measured data, or represent theoretically calculated values.

(2) Unless one has made mistakes in deriving energy levels and wave functions or made inconsistent errors somewhere else, calculated results from any self-consistent theories should automatically follow these rules.

(3) Given the fact that most of shift values in HITRAN 2008 are theoretically calculated result, the severe violations demonstrated in Fig. 11 definitely mean that the formalism itself contains fatal mistakes.

Outcome:

The corrections have been made for the shifts in HITRAN 2012 such that all the theoretically calculated values in the pure rotational band of HITRAN 2008 have been replaced and among all the bands, only 0.02 % of values in HITRAN 2012 come from these calculations.

VII. Identifying Missing Lines

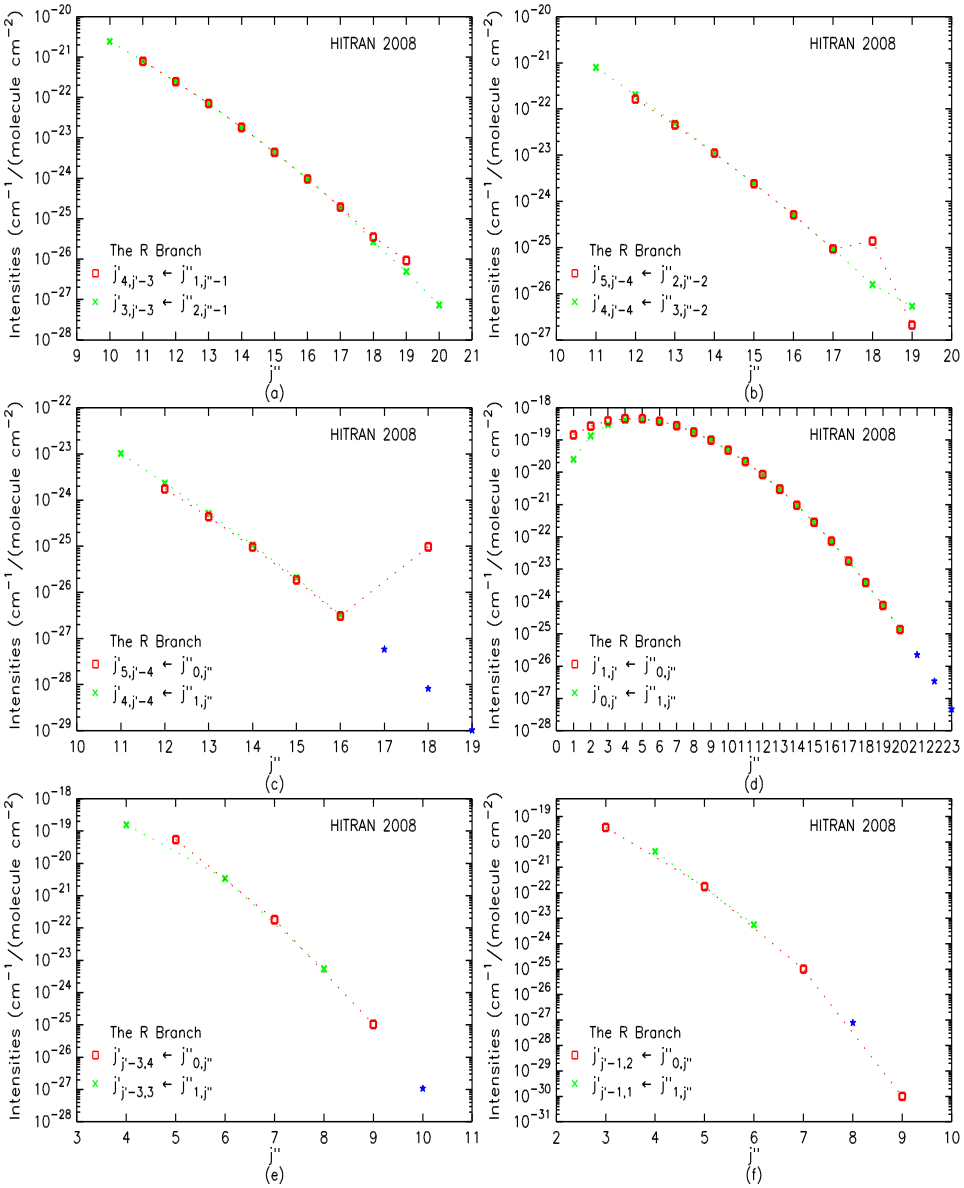


Fig. 11 The line intensities for the six groups in the pure rotational band.

Comments:

(1) There are violations of the two rules in Figs. (a)-(c). Some lines intensity data in these three groups are wrong. The intensity of the line $19_{5,15} \leftarrow 18_{0,18}$ in Fig. (c) is completely wrong.

(2) There are 7 missing lines whose intensities are above the threshold in Figs. (d)-(f).

They are $22_{1,22} \leftarrow 21_{0,21}$, $22_{0,22} \leftarrow 21_{1,21}$, $23_{1,23} \leftarrow 22_{0,22}$, $23_{0,23} \leftarrow 22_{1,22}$, $24_{1,24} \leftarrow 23_{0,23}$ in Fig. (d);

$11_{8,3} \leftarrow 10_{1,10}$ in Fig. (e);

$9_{8,1} \leftarrow 8_{1,8}$ in Fig. (f).

By going through all the groups in the pure rotational band, there are at least 81 missing lines in HITRAN 2008. The complementary list can be derived.

VII. Identifying Missing Lines

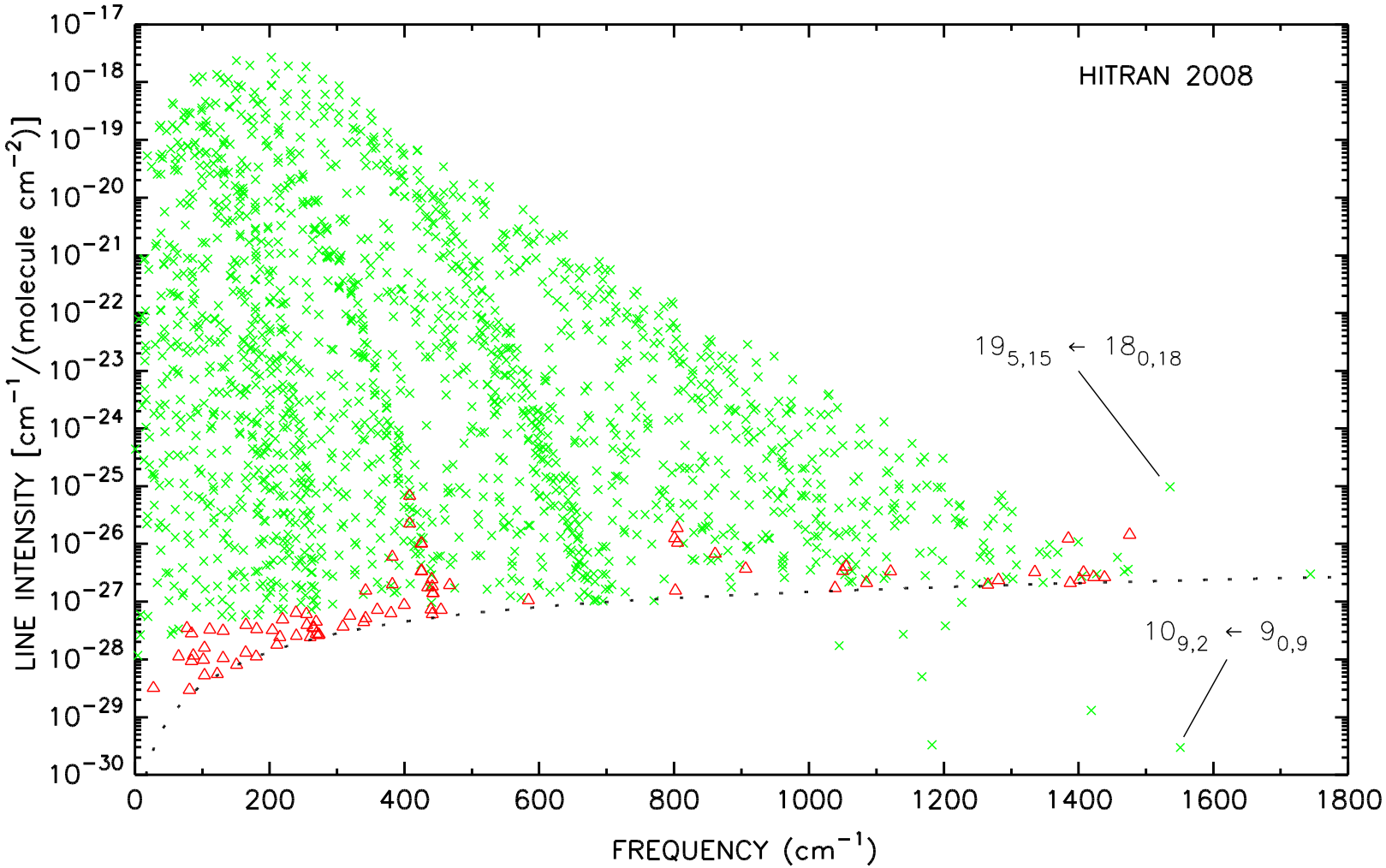


Fig. 12 The distribution of 81 missing H_2O lines in the pure rotational band over the transition positions Their predicted intensities are plotted by Δ. The original 1639 H_2O lines listed in HITRAN 2008 are also plotted by ×. The intensity threshold is plotted by a dotted line.

Conclusions

- **The spectroscopic parameters vary with the H_2O lines because their initial and final H_2O states have different energy levels and wave functions. Within individually defined groups, the variations of the spectroscopic parameters can be well monitored by the two rules.**

- **The two rules bear two characters: they are natural, but they are local. Local here means the rules work within individual groups and are valid for lines above certain boundaries.**

- **By screening the databases with the rules, one is able to identify mistakes and to provide more reliable values. In addition, one can find missing lines from line lists and develop a supplementary list. Thus, by applying these two rules, one can make significant contributions to improve the accuracy of the databases.**

- **The idea of the present work is simple and general. One can applied it for other important molecules in the atmosphere.**