

UNCERTAINTY OF SPECTROSCOPIC DATA

Molecular
Spectroscopy

Quantum
Chemistry

Attila G. Császár

Laboratory of Molecular Structure and Dynamics &
MTA-ELTE Research Group on Complex Chemical Systems
& COST Action CM1405, MOLIM: Molecules in Motion
ELTE, Budapest, Hungary

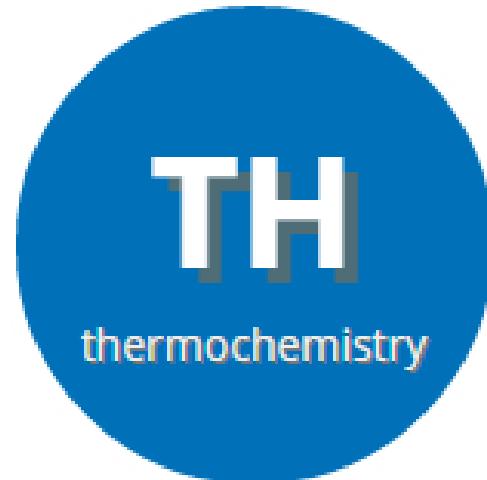
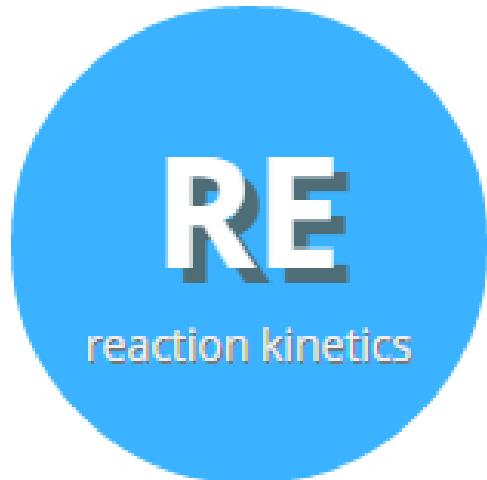
COST Training School on the
Analysis of Combustion Systems
Budapest, July 4-7, 2016

DOPI

DEWE

ReSpecTh:

a joint reaction kinetics, spectroscopy, and thermochemistry information system



<http://respecth.hu>

Reaction kinetics

- Respect Kinetics Data Format
- Database of combustion experiments
- Collection of Chemkin-format mechanisms
- Programs for the analysis of reaction mechanisms

Spectroscopy

- **Experimental databases**

The experimentally measured, assigned transitions validated by MARVEL: nine major water isotopologues, H_3^+ , H_2D^+ , D_2H^+ , NH_3 , ketene, C_2

- **MARVEL databases**

MARVEL energy levels with uncertainties

- **First-principles databases**

BT2 line list of H_2^{16}O , VTT line list of HD^{16}O , own D_2^{16}O energy levels, own energy levels of H_3^+ , H_2D^+ , and D_2H^+

Thermochemistry

- **NEAT database**

NEAT, standing for a “network of computed reaction enthalpies leading to atom-based thermochemistry”, is a simple and fast, weighted, linear least-squares refinement protocol and code for inverting the information contained in a network of quantum chemically computed 0 K reaction enthalpies.

- **Burcat's thermochemical data**

A mirror of Burcat's thermodynamic data collection

**Rovibronic (high-resolution)
molecular spectroscopy:**

the traditional experimental paradigm
for accuracy and precision

**First-principles variational
molecular spectroscopy:**

a new computational paradigm
for accuracy

Experiment vs. Theory

<u>Experiment</u>		<u>Theory</u>
	<i>Accuracy</i>	
<i>in Spec:</i> $(10^{-4}-10^{-6} \text{ cm}^{-1})$	>>	$10^{-1}-10^{-3} \text{ cm}^{-1}$
<i>in Th:</i> 1 kJ mol^{-1}	~	1 kJ mol^{-1}
	<i>Completeness</i>	
<i>in Spec:</i> ~1-10%	<<	~100%
<i>in Th:</i> ~100%	~	~100%
	<i>Results</i>	
<i>Implicit</i>		<i>Explicit</i>

Conclusion: use experiment and theory together and connect them with '*inversion tools*'

Inverse problems

Measured quantities



refinement procedure
(inversion tools)



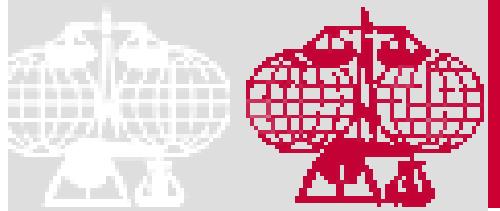
parameters of the physical model(s)

Why measure bound molecular states (spectra) of molecules?

- ⦿ **Modeling** in many scientific and engineering applications (e.g., star formation models, atmospheric modeling, including the greenhouse effect, and combustion) need detailed, precise, T -dependent, line-by-line information, usually deposited in old-fashioned databases
- ⦿ Most detailed information about the **structure** and **dynamics** of molecules
- ⦿ **Resonances** (quasibound states) and **tunneling** are also of great potential interest, partly for advanced reaction dynamics but also for scientific and engineering applications

Why compute bound molecular states (spectra) of molecules?

- **Test** theoretical methods against wealth of experimental results (including testing of potential energy (PES) and dipole moment (DMS) hypersurfaces)
- **Combine** theory and experiment to obtain the maximum information in order to determine complete spectra
- **Unravel** complicated spectra, basis for new assignments
- **Predict** experimentally not easily accessible or even unaccessible spectral regions or features
- **Bridge** (overtone) spectroscopy (anharmonicity and resonances) and dynamics (*e.g.*, IVR, vibrational adiabaticity, quantum ergodicity)



I U P A C

Number: 2004-035-1-100

Title: A database of water transitions from experiment and theory

Task Group

Chairman: J. Tennyson (U.K.)

Members: P. Bernath (Canada), A. Campargue (France),
M. R. Carleer (Belgium), **A. G. Császár (Hungary)**,
R. Gamache (Belgium), J. Hodges (U.S.A.),
A. Jenouvrier (France), O. Naumenko (Russia),
O. L. Polyansky (Germany), L. Rothman (U.S.A.),
R. A. Toth (U.S.A.), A. C. VanDaele (Belgium),
and N. F. Zobov (Russia)

<http://www.iupac.org/projects/2003/2004-035-1-100.html>

Question: would experiment or theory provide more useful data for scientific and engineering applications?

Question: would experiment or theory provide more useful data for the scientific and engineering applications?

Answer: basically neither, the best approach involves both and takes advantage of the strengths of the two complementary approaches while minimizing their weaknesses. This is true for high-resolution spectroscopy and thermochemistry, as well.

Spectroscopic networks (SN): large, finite, weighted, undirected, and rooted graphs

- energy levels: **vertices** (nodes, with given, theoretical or experimental, uncertainties)
- allowed transitions: **edges** (links)
- transition intensities: **weights**

A. G. Császár, T. Furtenbacher, *J. Mol. Spectrosc.* **2011**, *266*, 99-103
T. Furtenbacher, A. G. Császár, *J. Mol. Struct.* **2012**, *1009*, 123

CHEMISTRY AND ALGEBRA

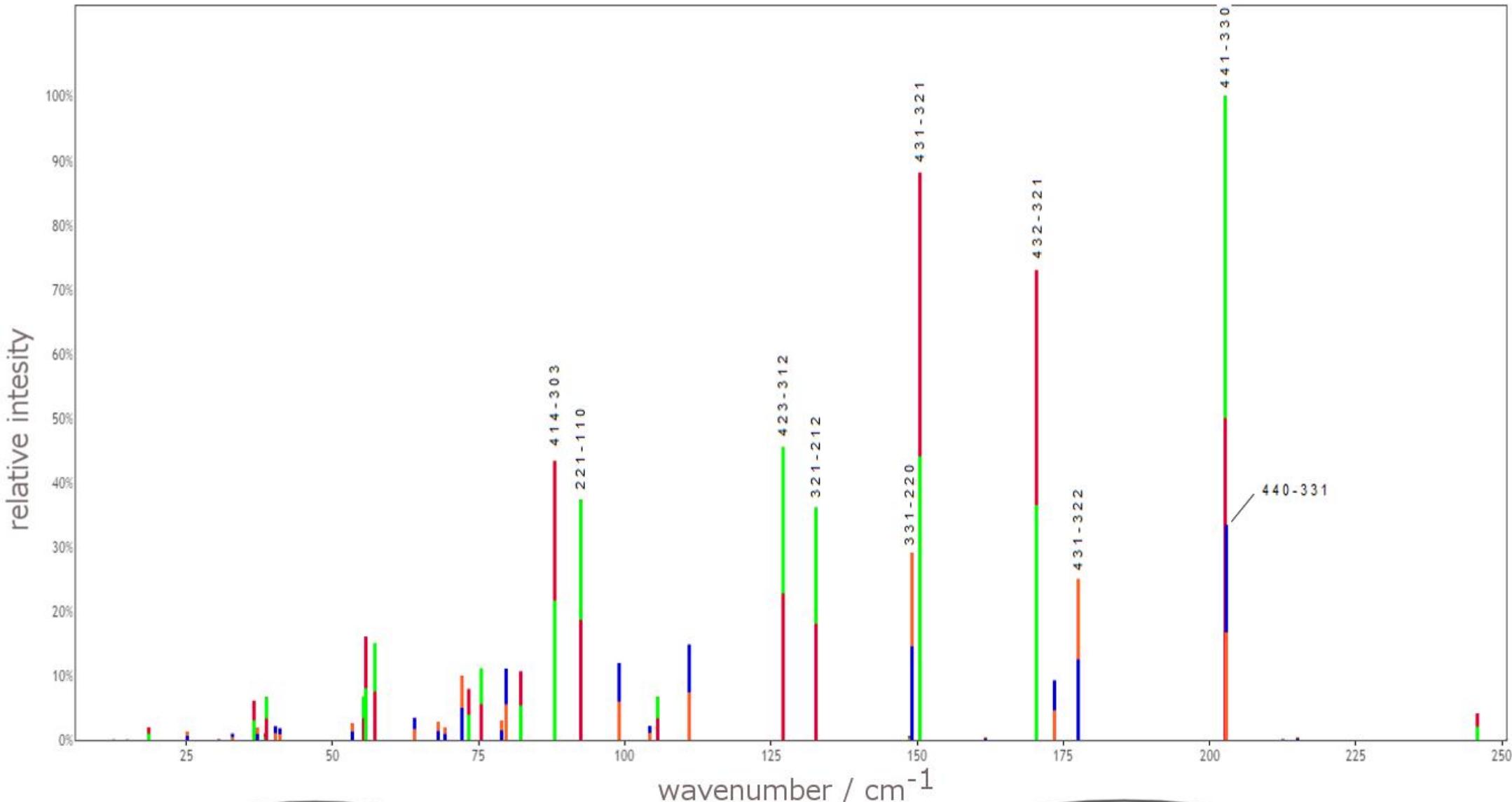
IT may not be wholly without interest to some of the readers of NATURE to be made acquainted with an analogy that has recently forcibly impressed me between branches of human knowledge apparently so dissimilar as modern chemistry and modern algebra. I have found it of great utility in explaining to non-mathematicians the nature of the investigations which algebraists are at present busily at work upon to make out the so-called *Grundformen* or irreducible forms appurtenant to binary quantics taken singly or in systems, and I have also found that it may be used as an instrument of investigation in purely algebraical inquiries. So much is this the case that I hardly ever take up Dr. Frankland's exceedingly valuable "Notes for Chemical Students," which are drawn up exclusively on the basis of Kekulé's exquisite conception of *valence*, without deriving suggestions for new researches in the theory of algebraical forms. I will confine myself to a statement of the grounds

Baltimore, January 1

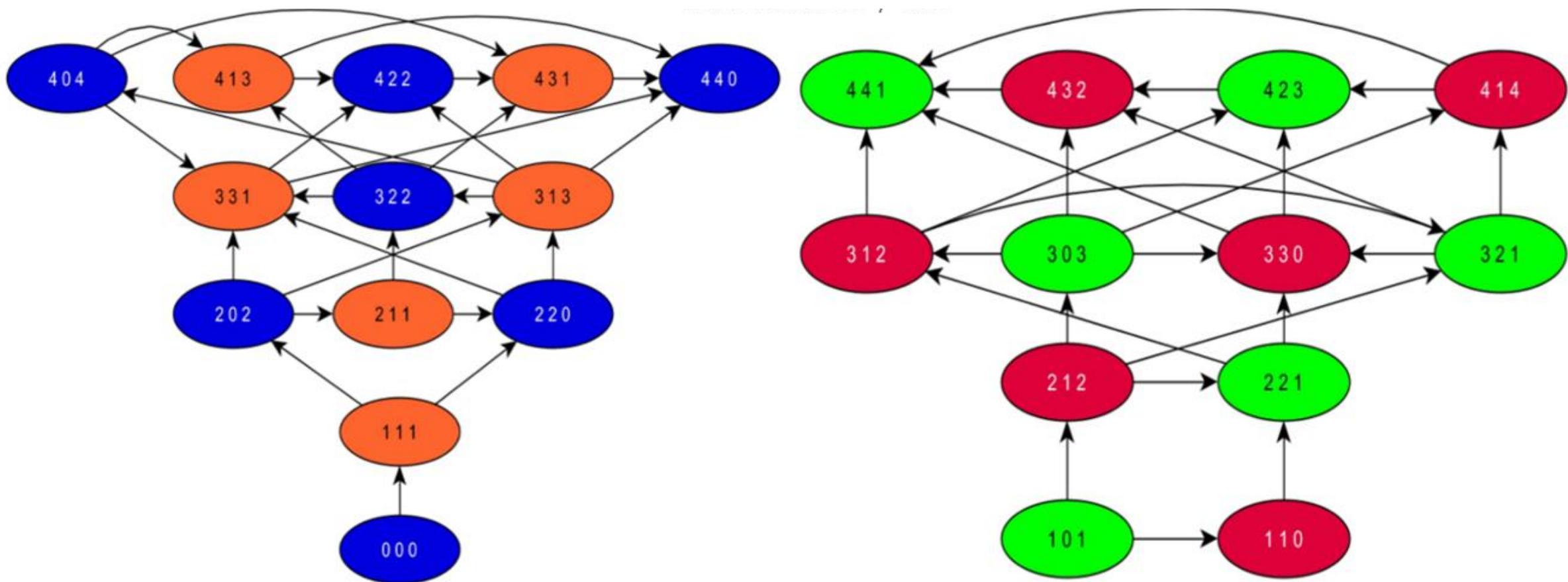
J. J. SYLVESTER

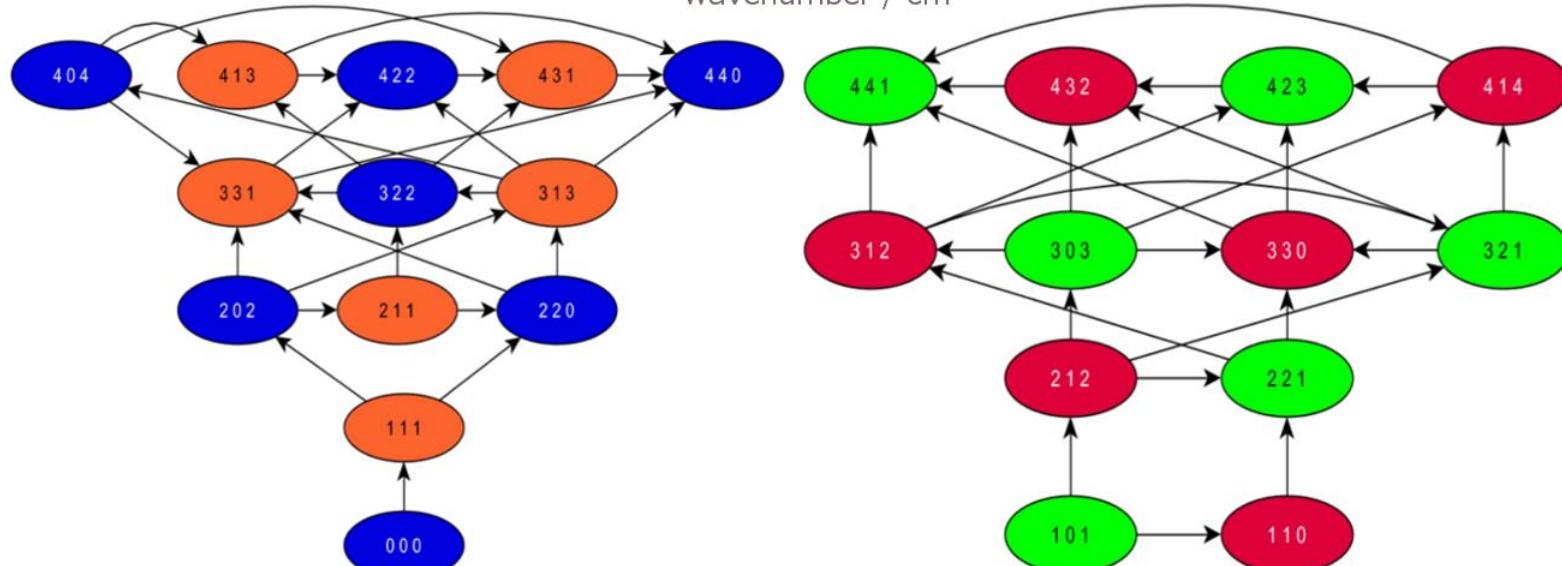
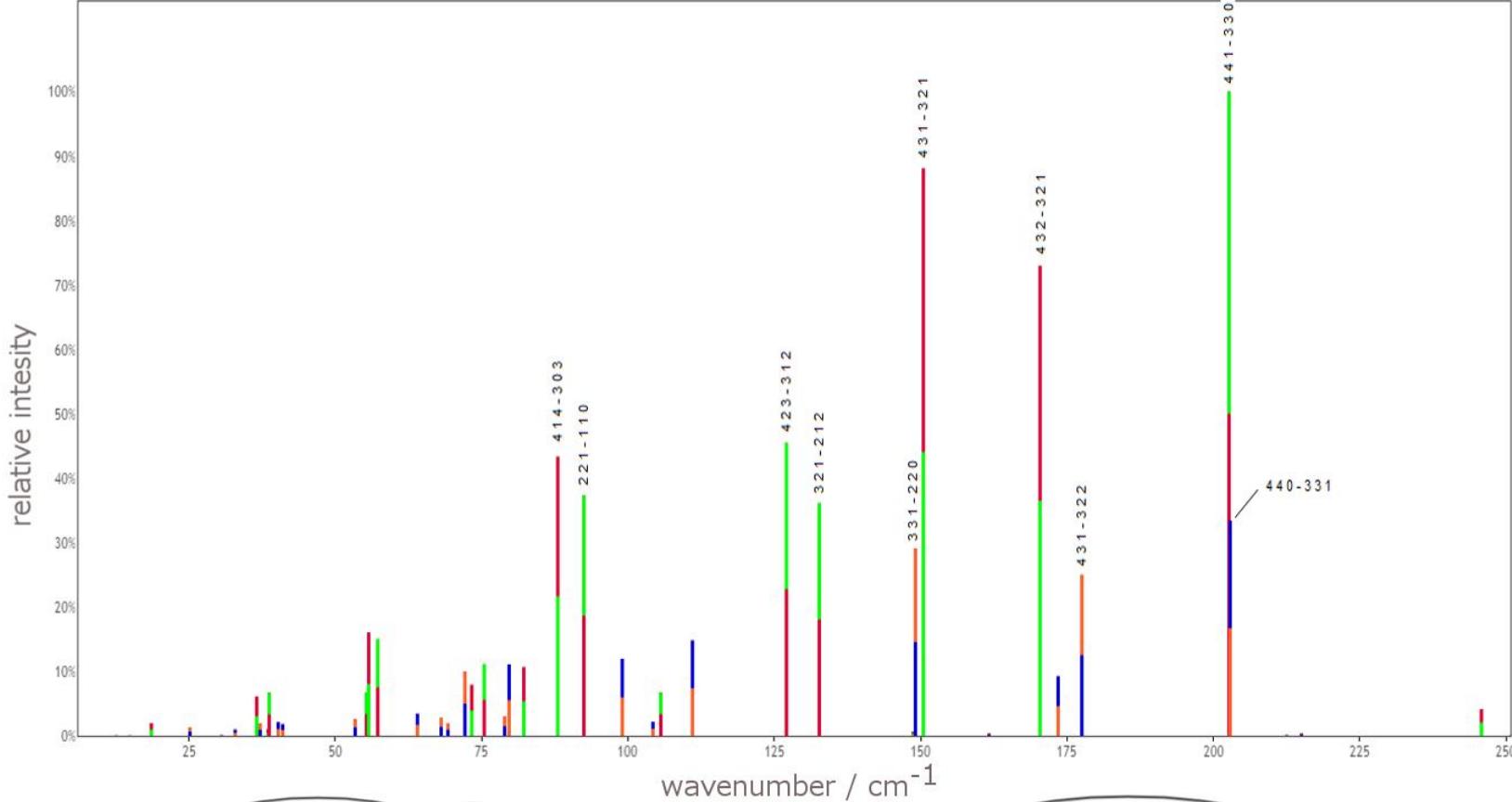
The weight of an invariant is identical with the number of the bonds in the chemicograph of the analogous chemical substance, and the weight of the leading term (or basic differentiant) of a co-variant is the same as the number of bonds in the chemicograph of the analogous compound radical. Every invariant and covariant thus becomes expressible by a *graph* precisely identical with a Kekuléan diagram or chemicograph. But not every chemicograph is an algebraical one. I show that by an application of the algebraical law of reciprocity every algebraical graph of a given invariant will represent the constitution in terms of the roots of a quantic of a type reciprocal to that of the given invariant of an invariant belonging to that reciprocal type. I give a rule for the geometrical multiplication of graphs, *i.e.* for constructing a *graph* to the product of in- or co-variants whose separate graphs are given. I have also ventured upon a hypothesis

Pure absorption rotational spectrum of the ground vibrational state of H_2^{16}O

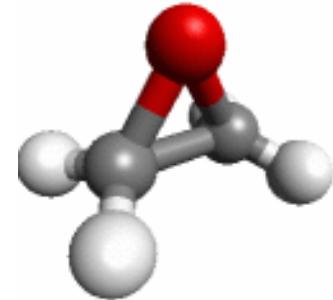
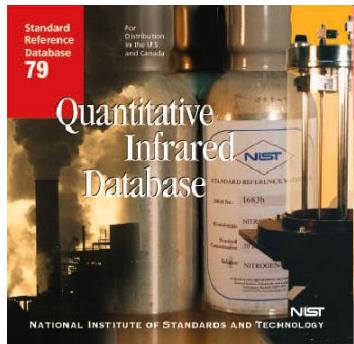


Pure absorption spectroscopic network of the ground vibrational state of H_2^{16}O





Spectroscopic databases



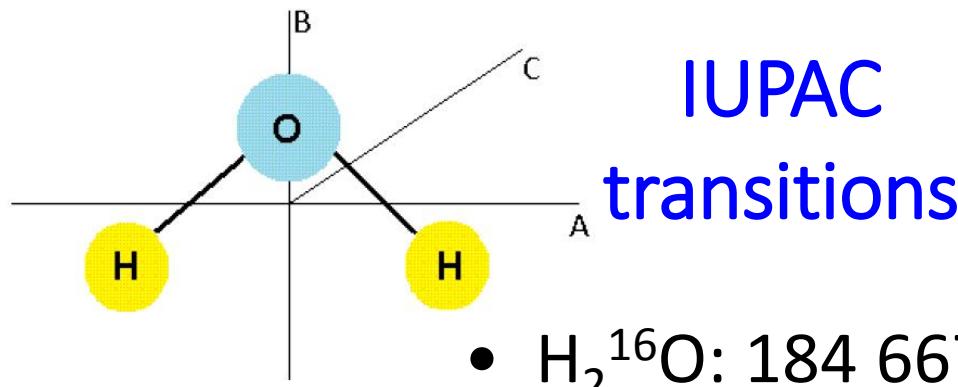
ASA – High-resolution Atmospheric
Atmos. Transmission Radiation
Spectr. Molecular Analysis
Appl. Absorption (GEISA)
Database
(Harvard-
Smithsonian)

Quantitative Infrared Database (NIST)
(CDMS)

IUPAC database of water isotopologues:
JQSRT 110 (2009) 573, 111 (2010) 2160,
117 (2013) 29, and 142 (2014) 93.

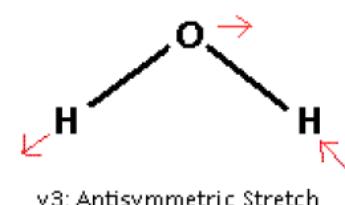
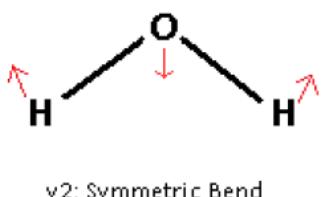
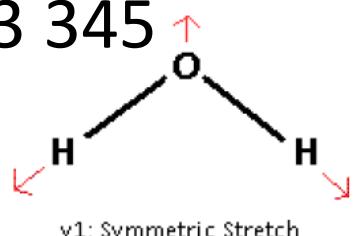
IUPAC energy levels

- H_2^{16}O : 18 486
- H_2^{17}O : 2 723
- H_2^{18}O : 5 131
- HD^{16}O : 8 818
- HD^{17}O : 162
- HD^{18}O : 1 864
- D_2^{16}O : 12 288
- D_2^{17}O : 338
- D_2^{18}O : 3 345



IUPAC transitions

- H_2^{16}O : 184 667
- H_2^{17}O : 9 169
- H_2^{18}O : 32 325
- HD^{16}O : 54 740
- HD^{17}O : 485
- HD^{18}O : 8 729
- D_2^{16}O : 53 532
- D_2^{17}O : 600
- D_2^{18}O : 12 167



MARVEL: An inverse, Hamiltonian-free approach to highly accurate rovibrational energy levels



**Measured Active
Rotational-Vibrational
Energy Levels**

T. Furtenbacher, A. G. Császár, J. Tennyson, *J. Mol. Spectrosc.* 245, 115 (2007)
T. Furtenbacher, A. G. Császár, *J. Quant. Spectr. Rad. Transfer* 113, 929 (2012)

Database of observed
transition wavenumbers
 ν_{ij} with assignments and
uncertainties

The ν_{ij} are determined by
term values E_i, E_j, \dots

$$\nu = a \times E$$

Diagram illustrating the relationship between the observed wavenumber database and the term value matrix:

- The database is represented as a matrix with columns labeled v_{ij} and rows labeled *Assignment i,j*.
- The term value matrix is represented as a column vector with entries E_i, E_j, \dots .
- The assignment row (i,j) is highlighted with a red box.
- The equation $\nu = a \times E$ is shown below the matrices.

Solve for E (in a least-squares sense, taking into account the experimental uncertainties of the ν_{ij}) to obtain experimentally derived term values E_i, E_j, \dots

Building of the spectroscopic network

Several graph-theoretical algorithms:

- Floyd–Warshall
- Dijkstra
- **DFS (depth-first search)**
- BFS (breadth-first search)

MARVEL

MARVEL: Measured Active Rotational
Vibrational Energy Levels

Database: v_{ij} w_{ij} and unique labels for upper
and lower energy levels

Basic equation (Ritz): $v_{ij} = E_i - E_j$

$$\begin{bmatrix} v_1 \\ \vdots \\ v_{ij} \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} \cdots & 0 & +1 & \cdots & -1 & 0 & \cdots \end{bmatrix} \begin{bmatrix} E_1 \\ \vdots \\ E_i \\ \vdots \\ E_j \\ \vdots \\ E_m \end{bmatrix}$$

MARVEL

MARVEL: Measured Active Rotational
Vibrational Energy Levels

Database: v_{ij} w_{ij} and unique labels for upper
and lower energy levels

Least-squares solution:

$$\begin{pmatrix} \sum_k w_{1k} & -w_{12} & -w_{13} & \dots & -w_{1n_T} \\ -w_{12} & \sum_k w_{2k} & -w_{23} & \dots & -w_{2n_T} \\ \dots & \dots & \dots & \dots & \dots \\ -w_{1n_T} & -w_{2n_T} & \dots & \sum_k w_{n_T k} & \end{pmatrix} \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{pmatrix} = \begin{pmatrix} \sum_k w_{1k} v_{1k} \\ \sum_k w_{2k} v_{2k} \\ \vdots \\ \sum_k w_{n_T k} v_{n_T k} \end{pmatrix}$$

MARVEL

Obtaining the energy levels and the associated uncertainties is **not** a simple inversion process

We employ an iterative method to adjust the uncertainties to get a self-consistent database

Method: **Robust reweighting**

$$w_{ij} = \frac{1}{\delta_{ij}^2 + \alpha \Delta_{ij}^2} \quad \text{where } \Delta_{ij} = v_{ij} - (E_i - E_j)$$

while $\sum_i \frac{g_i \Delta_{ij}^2}{N_t - N_\ell} - 1 < \varepsilon$

MARVEL

Solving the system of linear equations

A) Exact uncertainties:

$$\varepsilon_j = \sqrt{A_{jj}^{-1}}$$

Sparse-adaptive **LDL^T** decomposition
(Cholesky)

B) Approximate uncertainties:

$$A_{jj}^{-1} = \frac{1}{A_{ii}} + \frac{1}{A_{\text{sum}}}$$

conjugate gradient method

Experimental information systems

Species	Ref.	Measured transitions		Energy levels	No. of sources
		Identified	Validated		
$^{12}\text{C}_2$	55	23 343	22 949	5 699	39
H_3^+	52	1 610	1 410	652	26
H_2D^+	53	195	185	109	13
D_2H^+	53	154	136	104	9
H_2^{16}O	42	184 667	182 156	18 486	93
H_2^{18}O	40	32 325	31 705	5131	48
H_2^{17}O	40	9169	9028	2723	33
HD^{16}O	41	54 740	53 291	8818	74
HD^{18}O	41	8729	8634	1864	18
HD^{17}O	41	485	478	162	3
D_2^{16}O	43	63 050	62 372	12 301	74
D_2^{18}O	43	12 163	12 018	3351	18
D_2^{17}O	43	600	583	338	3
$^{14}\text{NH}_3$	54	29 450	28 530	4 961	56
$\text{H}_2^{12}\text{C}^{12}\text{C}^{16}\text{O}$	51	3 982	3 194	1 722	12

IUPAC water database II. HD¹⁶O, HD¹⁷O and HD¹⁸O

	HD ¹⁶ O	HD ¹⁷ O	HD ¹⁸ O
No. of transitions collected	54 740	485	8 729
Maximum <i>J</i>	30	11	18
Highest VBO (cm ⁻¹)	22 625	1 399	9 930
No. of energy levels	8 818	162	1 864
No. of sources	74	3	18

J. Tennyson, P. F. Bernath, L. R. Brown, A. Campargue, A. G. Császár, L. Daumont, R. R. Gamache, J. T. Hodges, O. V. Naumenko, O. L. Polyansky, L. S. Rothman, R. A. Toth, A. C. Vandaele, N. F. Zobov, S. Fally, A. Z. Fazliev, T. Furtenbacher, I. F. Gordon, S.-M. Hu, S. N. Mikhailenko, and B. Voronin, IUPAC Critical Evaluation of the Rotational-Vibrational Spectra of Water Vapor. Part II. Energy Levels and Transition Wavenumbers for HD¹⁶O, HD¹⁷O, and HD¹⁸O, *J. Quant. Spectr. Rad. Transfer* **2010**, *111*, 2160-2184.

IUPAC water database. III. H_2^{16}O

	H_2^{16}O
No. of transitions collected	184 667
Maximum J	42
Highest VBO (cm^{-1})	41 121
No. of energy levels	18 486
No. of sources	93

J. Tennyson, P. F. Bernath, L. R. Brown, A. Campargue, A. G. Császár, L. Daumont, R. R. Gamache, J. T. Hodges, O. V. Naumenko, O. L. Polyansky, L. S. Rothman, A. C. Vandaele, N. F. Zobov, A. R. Al Derzi, C. Fábri, A. Z. Fazliev, T. Furtenbacher, I. E. Gordon, L. Lodi, and I. I. Mizus, IUPAC Critical Evaluation of the Rotational-Vibrational Spectra of Water Vapor. Part III. Energy Levels and Transition Wavenumbers for H_2^{16}O , *J. Quant. Spectrosc. Rad. Transfer* 117 (2013) 29-58.

Summary of MARVEL analyses

property	H_3^+	H_2D^+	HD_2^+
no. of sources	26	13	9
range (cm^{-1})	7 – 16 506	5 – 7 105	23 – 6 581
no. of transitions (A/V)	1610 / 1410	194 / 185	154 / 136
energy levels – ortho	259	63	52
– para	393	46	52
– floating	105	14	27
no. of VBOs	13	7	6

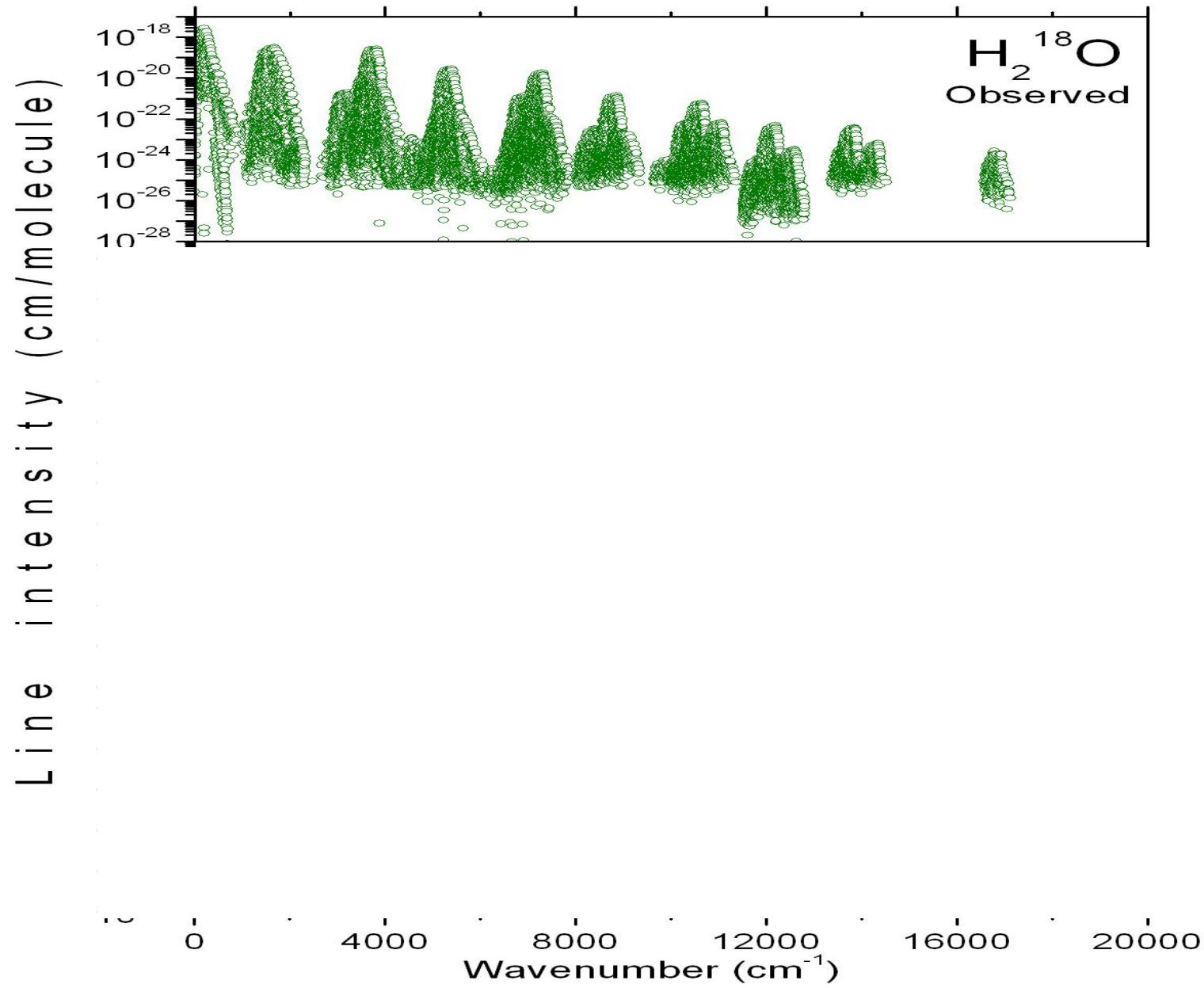
T. Furtenbacher, T. Szidarovszky, C. Fábri, A. G. Császár, Phys. Chem. Chem. Phys. 15, 10181 (2013).

T. Furtenbacher, T. Szidarovszky, E. Mátyus, C. Fábri, A.G. Császár, J. Chem. Theory Comput. 9, 5471 (2013).

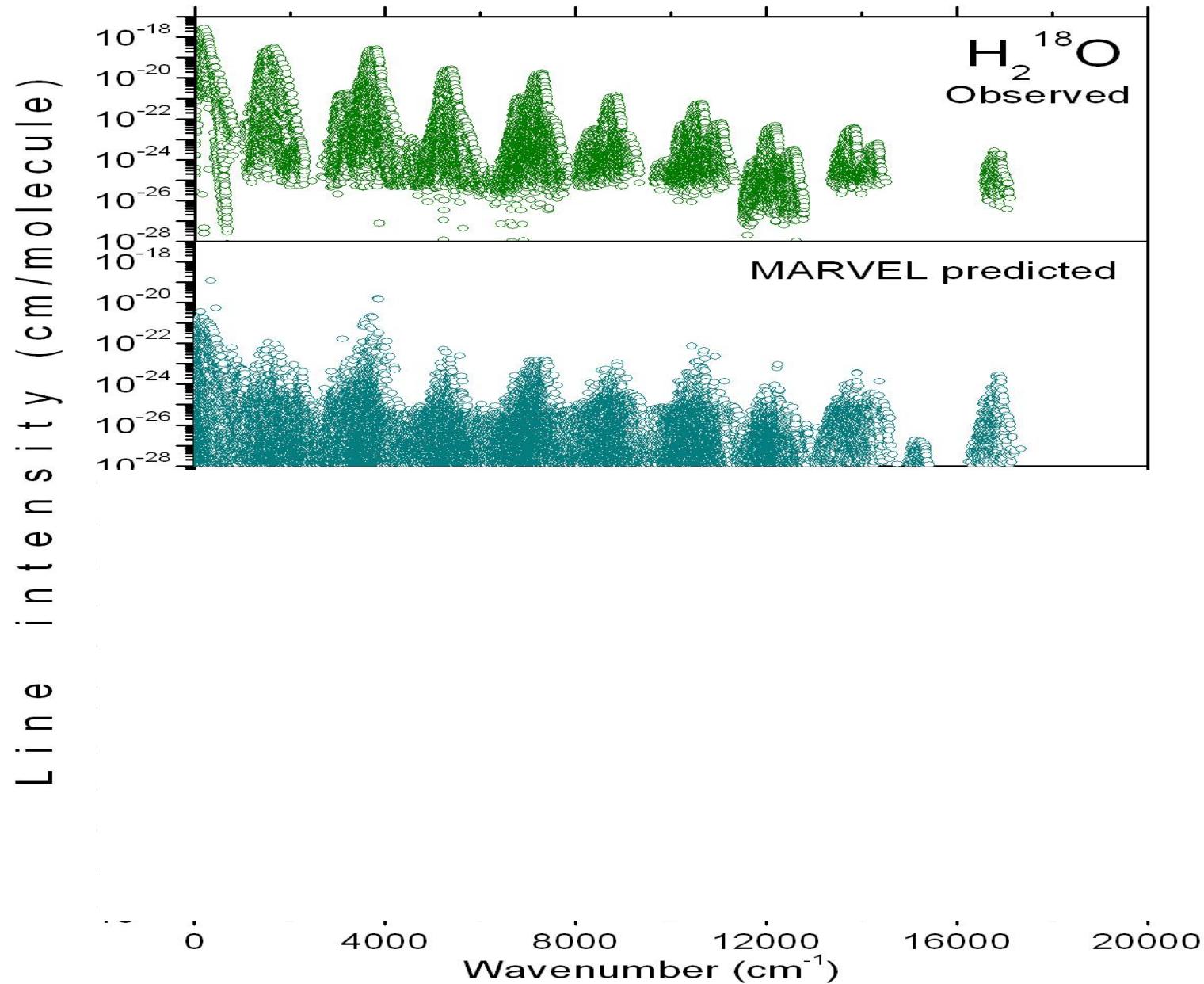
MARVEL vibrational band origins of D₂H⁺

$\nu_1\nu_2\nu_3$	VBO	Unc.	RL
0 0 0	0.0000 ^b	0	17
0 1 0	1968.1620	100	20
0 0 1	2078.4300	100	18
1 0 0	—		22
0 2 0	—		4
0 0 2	4042.7722	90	6
0 1 1	4060.7884	140	5
1 2 0	—		1
1 0 2	—		6
1 1 1	6616.0270	50	1

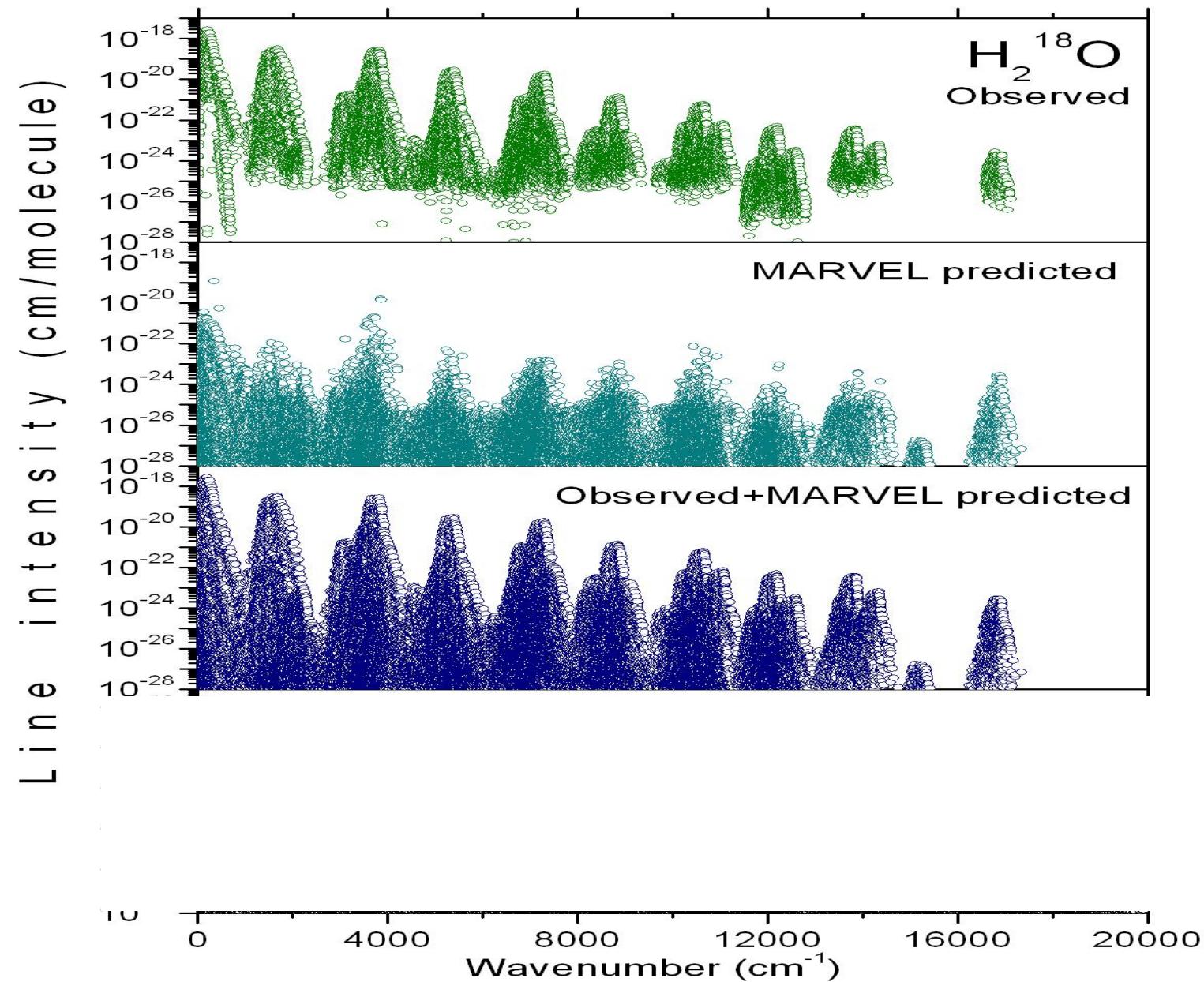
Predictive power of MARVEL



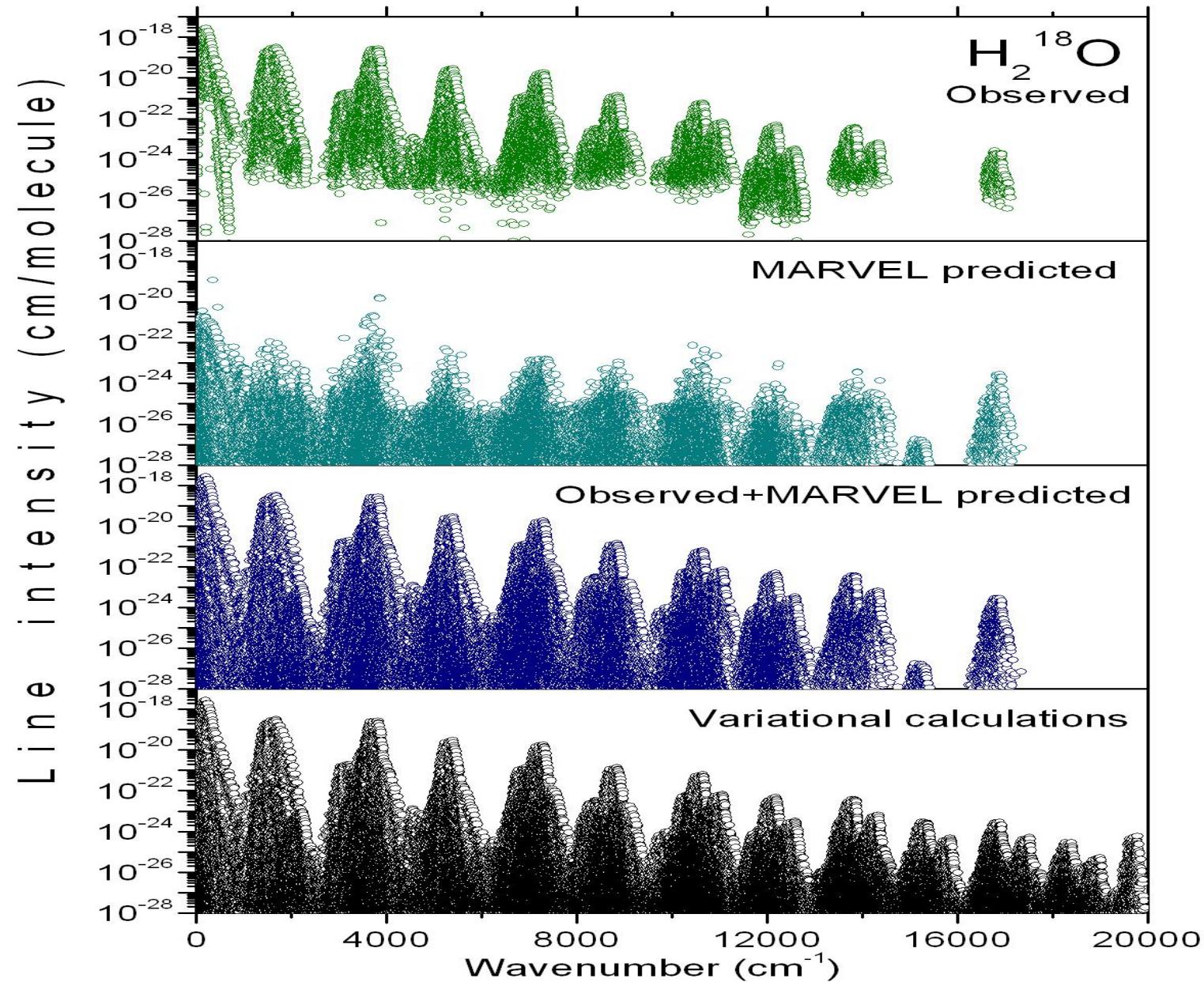
Predictive power of MARVEL



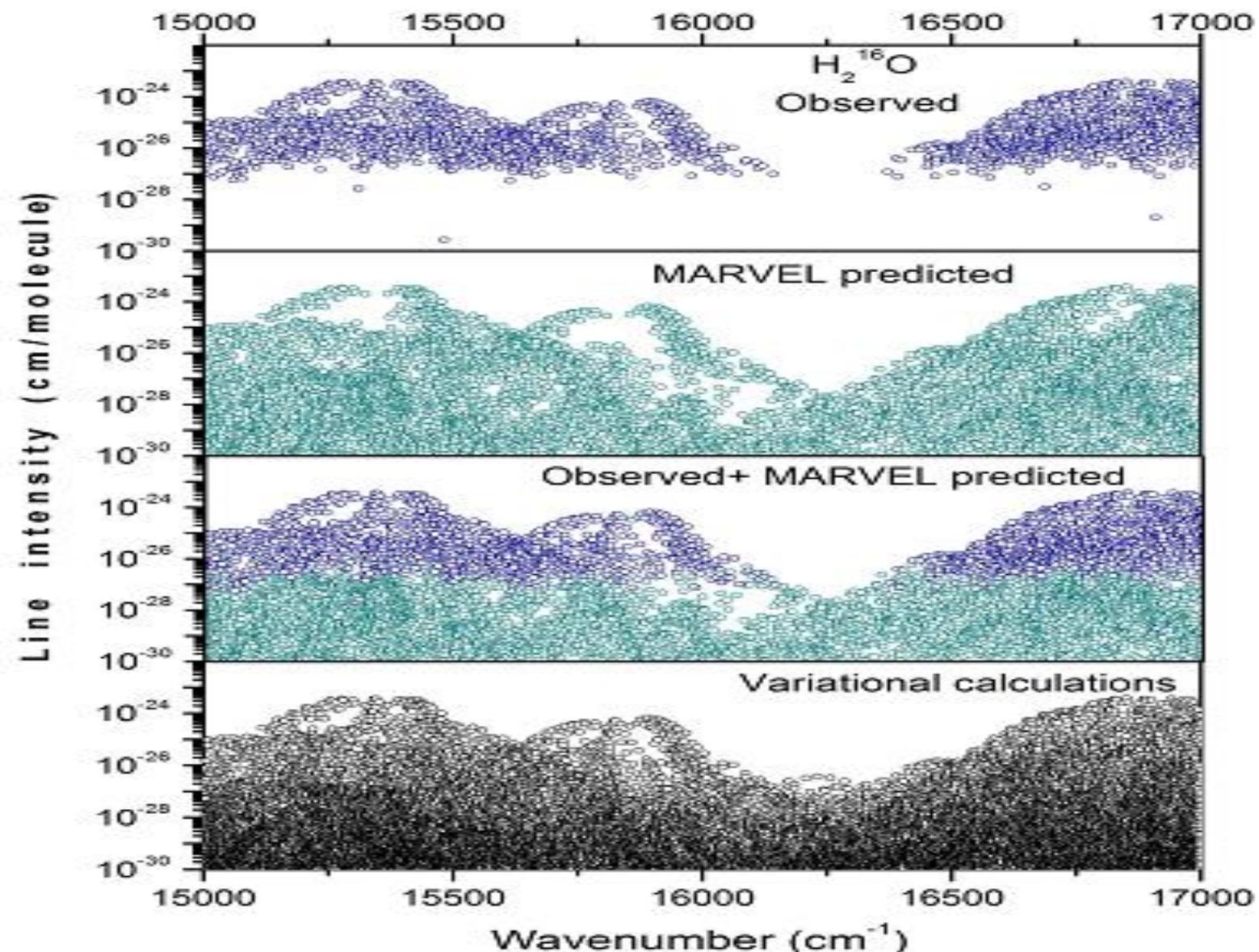
Predictive power of MARVEL



Predictive power of MARVEL



Predictive power of MARVEL



New Journal of Physics

The open-access journal for physics

New Journal of Physics 12 (2010) 103006 (10pp)

Frequency metrology in the near-infrared spectrum of H₂¹⁷O and H₂¹⁸O molecules: testing a new inversion method for retrieval of energy levels

A Gambetta¹, E Fasci², A Castrillo², M Marangoni¹,
G Galzerano³, G Casa², P Laporta³ and L Gianfrani^{2,4}

vibrational state. The measurement of the sub-Doppler line-center frequencies, also extended to the H₂¹⁷O spectrum, has been performed with an overall uncertainty of ~ 30 kHz, i.e. about three orders of magnitude lower than the HITRAN data set. The retrieved energy separations agree, by less than 80 kHz, with recent findings provided by the so-called MARVEL procedure for spectral data inversion, thus yielding a very stringent test of its accuracy.

New calibration standards: H_2^{16}O , H_2^{18}O , D_2^{16}O (15-170 cm⁻¹)

- only para-water transitions
- measured with internal precision better than 33 kHz
- underlying energy levels are involved in at least three measured transitions

H_2^{16}O	H_2^{18}O	D_2^{16}O
752,033.104(13)	203,407.502(30)	643,247.288(37)
916,171.405(13)	322,465.170(30)	692,243.579(36)
987,926.743(16)	517,181.960(30)	714,087.313(36)
1,207,638.714(13)	745,320.142(36)	743,563.526(36)

Recalibration of measurements

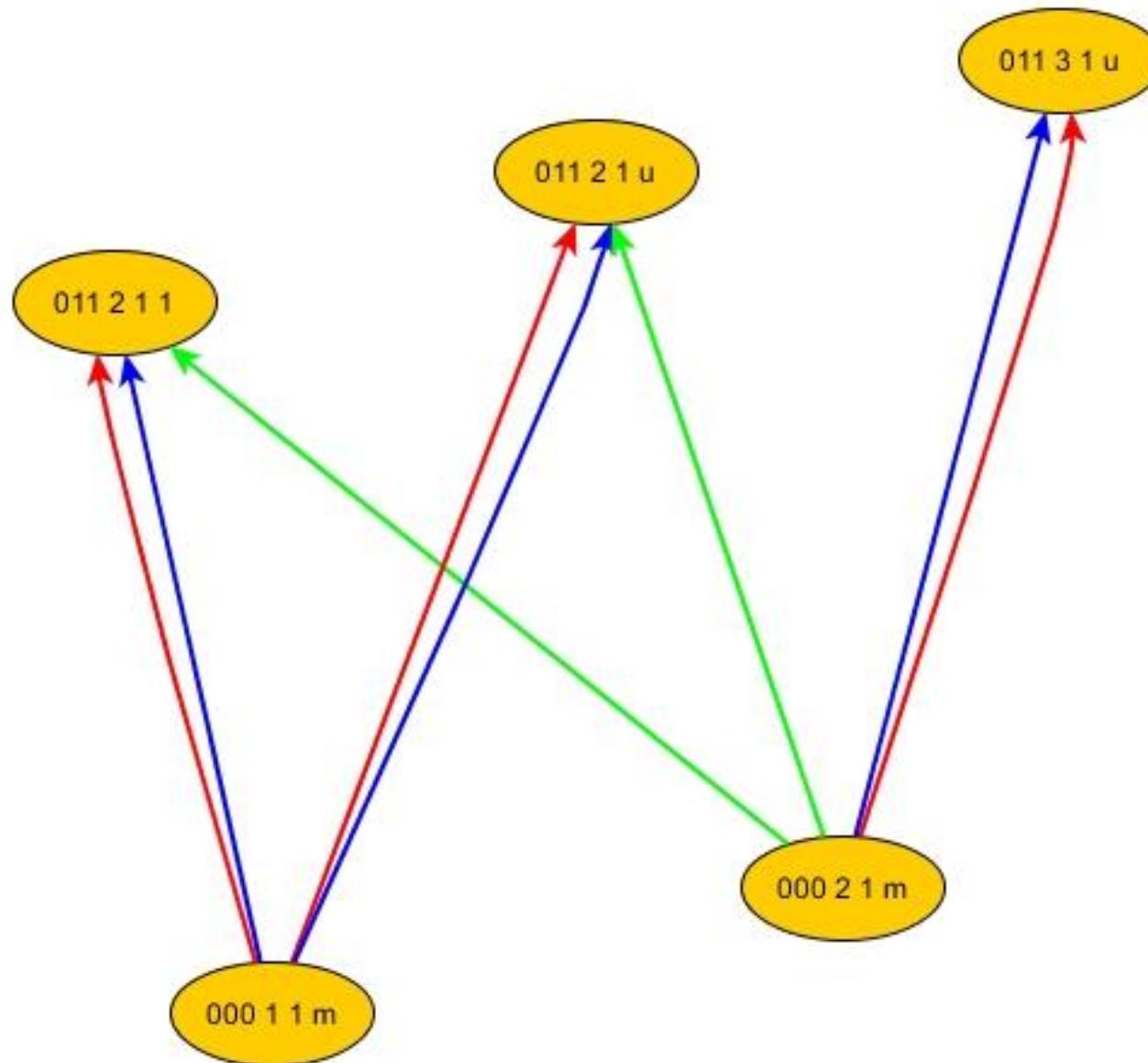
Source	Range (cm^{-1})	Calib. factor
78KaKaKy [70]	33–714	0.999 996 07(7)
95PaHo [105]	177–520	1.0
05HoAnAlPi [138]	212–595	1.0
96PoBuGuZh [107]	407–921	0.999 999 66(2)
82KajHo [79]	501–714	0.999 998 84(2)
83Guelachv [81]	1066–2583	0.999 999 77(4)
73CaFlGuAm [62]	2933–4251	0.999 999 93(2)
94Tothb [103]	5750–7988	1.0
80CaFlMa [74]	5934–6400	0.999 999 80(3)
05ToTe [141]	7423–9595	1.0
88MaChFlCa [91]	8057–9482	0.999 999 86(4)
75FlCaNaCh [66]	8060–9367	1.0
05ToNaZoSh [142]	9250–25 224	1.0
08ToTe [155]	9502–14 495	0.999 999 89(3)
89ChMaFlCa [92]	9603–11 481	0.999 999 83(3)
02BrToDu [126]	9676–11 383	0.999 999 90(2)
99CaJeVaBe [116]	13 185–21 389	0.999 999 96(1)
98PoZoViTe [115]	13 239–15 995	1.0

H₃⁺ measurements

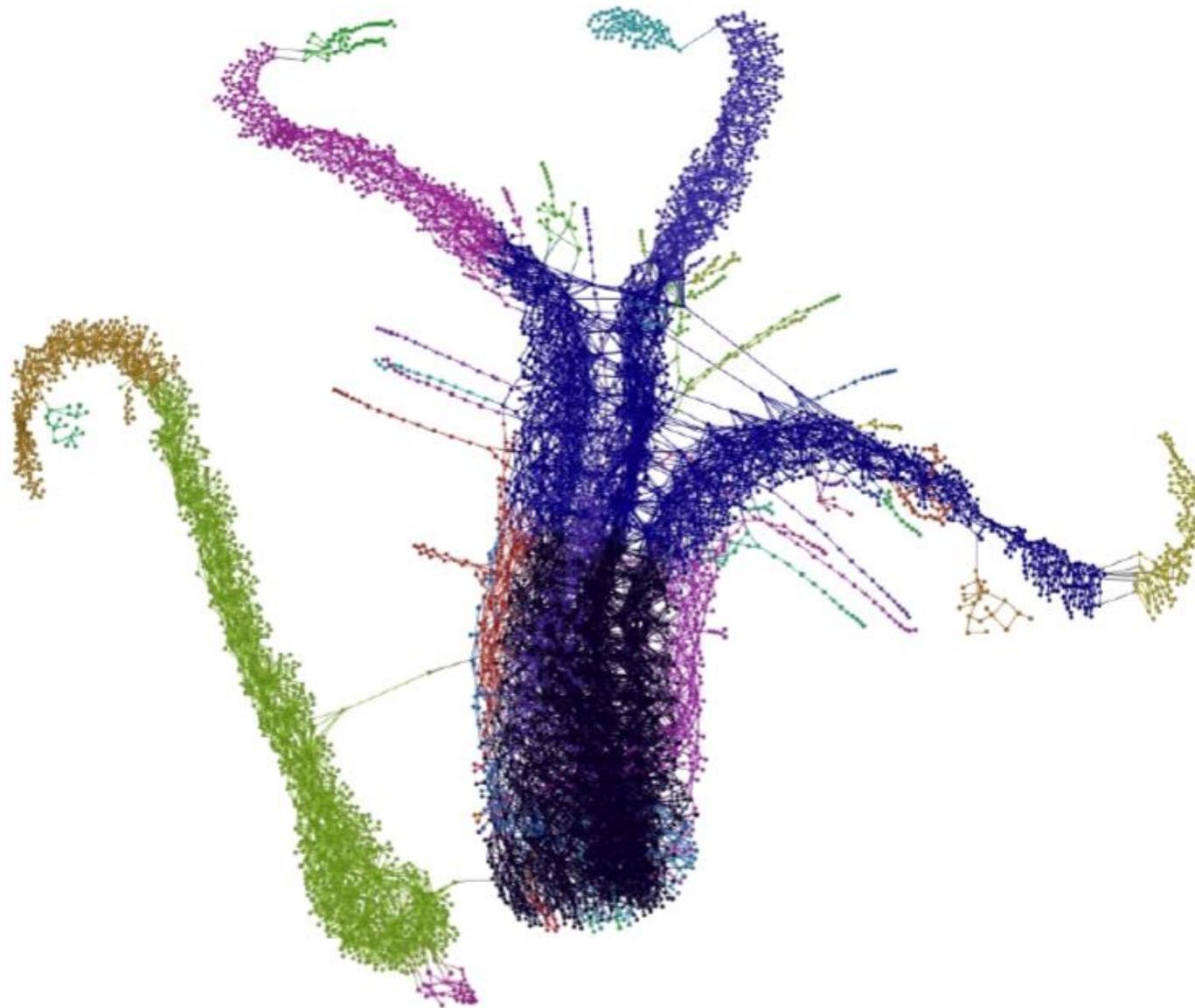
Table 5. Selected high-quality measured transitions for H₃⁺ from 13HoPeJeSi,¹⁰⁴ 13WuLiLiLi,¹⁰⁵ and 16JuKoScAs,¹⁰⁶ showing considerable disagreement between the former two measurements

Transition	13HoPeJeSi ¹⁰⁴	13WuLiLiLi ¹⁰⁵	Diff / 10 ⁻⁵ cm ⁻¹	16JuKoScAs ¹⁰⁶
R(1,1) ^l	2691.44239(2)	2691.44305(33)	66	2691.442718(5)
R(1,1) ^u	2726.21965(1)	2726.22025(66)	60	2726.220011(7)
R(2,1) ^u	2826.11628(1)	2826.11683(33)	55	

H_3^+ rovibrational measurements



C_2 rovibronic measurements



Summary

- Spectroscopic networks (SN) are weighted, undirected graphs, whereby energy levels are the nodes, allowed transitions are the links, and weights are provided by intensities.
- SNs provide highly useful concepts for high-resolution molecular spectroscopy and for database management.
- MARVEL, based on SNs, allows to turn scattered information into focused knowledge.
- A large number of spectroscopic-quality energy levels can be determined via MARVEL.