RAS Techniques and Instruments

RASTAI **3,** 565–583 (2024) https://doi.org/10.1093/rasti/rzae037
Advance Access publication 2024 August 30

A hybrid approach to generating diatomic line lists for high resolution studies of exoplanets and other hot astronomical objects: updates to ExoMol MgO, TiO, and VO line lists

Laura K. McKemmish [1]★ Charles A. Bowesman [2] Kyriaki Kefala [2] Armando N. Perri [1,2] Anna-Maree Syme,[1] Sergei N. Yurchenko [2] and Jonathan Tennyson [2]★

[1] *School of Chemistry, University of New South Wales, 2052 Sydney, Australia*
[2] *Department of Physics and Astronomy, University College London, Gower Street, WC1E 6BT London, UK*

Accepted 2024 August 15. Received 2024 August 7; in original form 2024 May 6

ABSTRACT

The best molecular line lists for astrophysical applications today require both high accuracy of line positions for strong lines as well as high overall completeness. The former is required to enable, for example, molecular detection in high-resolution cross-correlation observations of exoplanets, while completeness is required for accurate spectroscopic and radiative properties over broad temperature and spectral ranges. The use of empirical energies generated with the MARVEL procedure is a standard way to improve accuracy; here we explore methods of extending the use of these levels using predicted shifts and isotopologue extrapolation, as well as augmenting the levels from other sources such as effective Hamiltonian studies. These methods are used to update ExoMol line lists for the main $^{24}Mg^{16}O$ and $^{48}Ti^{16}O$ isotopologues, as well as for $^{24}Mg^{17}O$, $^{24}Mg^{18}O$, $^{25}Mg^{16}O$, $^{26}Mg^{16}O$, $^{46}Ti^{16}O$, $^{47}Ti^{16}O$, $^{49}Ti^{16}O$, and $^{50}Ti^{16}O$; new MARVEL results for $^{51}V^{16}O$ are also presented.

Key words: Data Methods; – molecular data – techniques: spectroscopic.

1 INTRODUCTION

The desire to characterize the atmospheres of exoplanets, and other hot astronomical objects, has led to the creation of molecular line lists (sets of energy levels and the intensity of transitions between these levels) which can be used for such studies, see the recent review by Tennyson & Yurchenko (2022). The ExoMol project was conceived to produce line lists with high completeness at temperatures appropriate for the studies of exoplanet atmospheres (Tennyson & Yurchenko 2012); see also Yurchenko et al. (2014). Typically, such activities require spectra with a resolving power R ($= \lambda/\Delta\lambda$) of a few hundred rising to perhaps 3 000 for the *JWST*.

More recently, a new means to identify molecular species in exoplanetary atmospheres has emerged that uses ground-based high-resolution spectroscopy taking advantage of cross-correlation of observed spectra with very high-resolution template molecular spectra (Birkby 2018; Mollière & Snellen 2019). This method currently provides the most secure detections of molecules in exoplanets. To be successful the cross-correlation method requires very high resolution of strong spectral lines which are known accurately enough to model observations obtained with R values in the typical range 70 000 to 150 000 used for these studies. Even for diatomic species, line lists originally developed for space-based transit spectroscopy are often not suitable for these applications, as discussed by Hoeijmakers et al. (2015) and de Regt et al. (2022).

The methodologies for constructing line lists have been adapting to meet these dual needs, crucially including explicit experimentally derived energy levels within a high-completeness modelled line list. Details of how this is accomplished are important to enable production of the best possible data for astronomers within the limitations of current experimental data availability and accuracy of ab initio electronic structure. Different methods have been developed in an ad-hoc manner to address the needs of individual molecules; it is timely now to consolidate these methodology developments and standardize terminology and nomenclature to allow comparison between different line lists. Therefore, in Section 2, we articulate and standardize the general approach whereby the ExoMol data base is updated to use the best-known energy for that quantum state from any data source (see Table 1 for common sources). This expands the now standard dihybrid MARVELized variational line list approach, which incorporates experimentally derived energy levels obtained from the MARVEL (measured active rotational-vibration energy levels) procedure, to a more comprehensive methodology where multiple different sources of energy levels can be incorporated into a single line list to optimize its accuracy.

In Sections 3 to 5, we demonstrate these modern approaches to line list construction for three different molecules, MgO, TiO, and VO, providing important improvements to their existing line list data, while showcasing how the approach taken differs between molecules based on existing line list qualities and spectral data availability. For both MgO and VO, good variational line lists existed (McKemmish, Yurchenko & Tennyson 2016b; Li, Tennyson & Yurchenko 2019) ensuring completeness, but the accuracy of the strong line positions needed improvement; hence, the primary work in creating a better line list was in the collation of experimental data and a MARVEL analysis, which was subsequently used to produce a new hybrid line

★ E-mails: l.mckemmish@unsw.edu.au (LKM); j.tennyson@ucl.ac.uk (JT)

Table 1. Source type abbreviations consistently applied to diatomic hybrid line lists discussed in this paper.

Abbr.	Meaning	Description	Strength	Limitation	Rel. Acc.
Direct predictions					
Ca	Calculated	Usually variational from an energy spectroscopic model, e.g. using DUO	Most extensive coverage, physically consistent extrapolation	Lower accuracy than other approaches	Low
Ma	MARVEL	Derived from collation of assigned experimental transitions	Highest accuracy	Limited scope, relies on experimental data availability	High
EH	Effective Hamiltonian	Fitted to experimental data	Good interpolation	Lower accuracy extrapolation especially to higher vibrations	Mid
Mo	MOLLIST	Usually effective Hamiltonian, but see original paper	–	Method not consistent across all molecules	Mid
HI	HITRAN	Mostly experimentally derived with careful validation process, see original paper	–	–	High
Corrections					
PS	Predicted shift	Ma-Ca difference (or similar) interpolated or propagated to higher rotational states within a vibronic level	Good way to correct for limitations in the spectroscopic model, e.g. missing couplings	–	Mid
IE	Isotopologue-extrapolation	Ma-Ca difference (or similar) in main isotopologue propagated to other isotopologues	Accuracy obtained without additional experiments on each isotopologue	–	Mid

list. For TiO, the heavy demand by the astronomical community, the complexity of the molecule, and new experimental data warrants an update of the Toto line list (McKemmish et al. 2019) despite the original study using largely similar methodologies to the ones considered here.

2 HYBRID LINE LIST CONSTRUCTION METHODOLOGY

2.1 Overview

In this paper, we argue that the best approach for modern line list construction is an energy-hybridized line list constructed through the following procedure:

(i) Produce a high-completeness line list based on a spectroscopic model fit to experimental data; within standard ExoMol format (Tennyson et al. 2014, 2020) a line list should have two main components, a `.states` file which stores energy levels plus associated metadata for each state (such as degeneracy and quantum numbers) and a `.trans` file which stores Einstein A coefficients for transitions between rovibronic states. Transition wavenumbers are computed from the energy levels in the `.states` file.

(ii) Update the energies of the `.states` file to provide the most accurately predicted energy of each state. A variety of data sources may be used, such as MARVEL (Furtenbacher, Császár & Tennyson 2007) experimentally derived or effective Hamiltonian energy levels.

The first step is well-established, with the most popular methodology involving fitting a spectroscopic model to experimental data then producing a variational line list using a nuclear motion program such as DUO (Yurchenko et al. 2016a) for diatomics, see Tennyson (2012) and Tennyson & Yurchenko (2017) for further details.

The second step is the key innovation over recent years, and the focus of this paper. We focus on diatomic molecules here for simplicity, but the methodology is general for rovibronic and rovibrational line lists for small molecules.

We use the term 'source type' to reference the source of the energy level prediction. There are two main classes of predicted energy levels, variational calculations, and empirical such as from MARVEL, plus corrections that directly improve predicted energy levels by taking note of observed uncertainties between high-accuracy and low-accuracy direct predictions. Though different terminology has been used in the literature, we standardize this in Table 1 and in particular have updated line lists on the ExoMol website (www.exomol.com) to provide consistent two-letter abbreviations for the source of energy levels within the `.states` files.

Note that, regardless of the source type for the energy level, transition intensities are calculated using dipole moment curves that together with a particular model for the energy levels allow the Einstein A coefficient of each transition to be calculated. Dipole moment curves are almost exclusively derived from ab initio calculations, which have been shown to be highly accurate in favourable cases (Polyansky et al. 2015; Bielska et al. 2022) and their use compensates for unreliability or unavailability of absolute intensity measurements for most of the molecules considered in the ExoMol data base. However, in some cases measurements of excited state lifetimes can be used to scale these ab initio dipoles (Yurchenko et al. 2024a; Bowesman et al. 2024b).

2.2 Source types for energy levels

In this section, we detail the source types for energy levels included in the three line lists updated in this paper: MgO, TiO, and VO. Of particular importance is the consideration of uncertainties.

2.2.1 Experiment energy levels (often collated by MARVEL analysis)

Description: The experimental methodology relies on measurements obtained astronomically or in the laboratory. By the late twentieth century, this approach led to an abundance of experimental studies which were largely left as independent sources of data, which are extremely challenging for astronomers to utilize.

Furtenbacher et al. (2007) developed the MARVEL procedure to facilitate the provision of empirical energy levels with experimental accuracy. This procedure involves the collection of assigned, high-

resolution experimental transition frequencies from all experimental studies in the literature. Using the MARVEL procedure, the full breadth of experimental studies can be successfully collated into a self-consistent spectroscopic network of energies and transition frequencies with reliable uncertainties. The MARVEL algorithm inverts measured frequencies, with specified uncertainties, to give the empirical energy levels of the molecule. This is achieved using a weighted least-squares fit and produces a self-consistent spectroscopic network that accounts for all observed transitions (Császár & Furtenbacher 2011). The MARVEL algorithm employs graph theory (Arendas, Furtenbacher & Császár 2020) and thus makes no model assumptions but relies exclusively on the rovibronic transition labels provided.

Discussion: The experimental methodology almost certainly produces the most accurate spectral frequencies data for a molecule where available (though absolute intensities are far less reliable).

However, the time and cost associated with producing experimental spectral data is very high and the number of lines measured tends to be very small compared to the full number of transitions that are important for describing molecular opacity, particularly at higher temperatures.

Use of the MARVEL procedure to obtain high accuracy energy levels can enable a small number of transitions to be used to predict the transition frequencies of a much larger number of transitions while retaining the low experimental uncertainties.

Utilization: The MARVEL procedure has proved indispensable in producing a single comprehensive experimental line list from the plethora of experimental data available in the literature. There are currently 18 diatomic molecules of astronomical importance for which MARVEL data sets are available: C_2 (Furtenbacher et al. 2016; McKemmish et al. 2020), TiO (McKemmish et al. 2017), AlH (Yurchenko et al. 2018a, 2024b), SiO (Yurchenko et al. 2022), AlO (Bowesman et al. 2021), NO (Qu, Yurchenko & Tennyson 2021; Furtenbacher et al. 2022), BeH (Darby-Lewis et al. 2018), O_2 (Furtenbacher et al. 2019), CN (Syme & McKemmish 2020), NH (Darby-Lewis et al. 2019), ZrO (McKemmish et al. 2018), VO (Bowesman et al. 2022), CH (Furtenbacher et al. 2022), OH (Furtenbacher et al. 2022), SiN (Semenov, Tennyson & Yurchenko 2022), SO (Brady et al. 2024), YO (Yurchenko et al. 2024a), and CH^+ (Pearce, Yurchenko & Tennyson 2024).

Uncertainties: A central component of the MARVEL procedure is the propagation of errors from the measured transition frequencies to errors in the empirical energy levels. This means that the MARVEL procedure automatically determines the uncertainty of energy levels based on the input uncertainties of transitions. Self-consistency in the spectroscopic network is forced through the validation process in MARVEL, helping to ensure that unrealistically low transition uncertainties are corrected.

The method of predicting uncertainties can change between MARVEL versions. In this work we employed the latest version of MARVEL, MARVEL4.1, which uses a bootstrap method to determine the uncertainties associated with each final energy level (Tennyson et al. 2024a).

For most molecules, MARVEL uncertainties have been adopted in ExoMol line lists without modification as an estimate of the uncertainty of that energy level. This approach makes the implicit assumption that the uncertainties of the initial ($\Delta\tilde{E}_i$) and final ($\Delta\tilde{E}_f$) state energy value can be used to give the uncertainty of the observable transition frequency ($\Delta\tilde{\nu}$) as

$$\Delta\tilde{\nu} = \sqrt{(\Delta\tilde{E}_i)^2 + (\Delta\tilde{E}_f)^2}. \quad (1)$$

2.2.2 Calculated variational energy levels (best compiled by DUO for coupled electronic states)

Description: The variational methodology of line list construction relies on the numerical solution of the nuclear motion Schrödinger equation. This requires a preliminary solution to the electronic structure problem which yields an electronic spectroscopic model comprising a set of potential energy, coupling and dipole moment curves that vary with the geometric parameters of the molecule. For diatomic molecules, this model is simply a function of bond length. Although the energy spectroscopic model can be obtained purely from ab initio electronic structure calculations, the resulting ab initio variational line list often deviates significantly from an experimental line list. This is especially true for open-shell electronic states (Tennyson et al. 2016). Therefore, it is essential to replace ab initio curves with empirical curves that are fitted to experimental data whenever possible. This fitting procedure can be a lengthy process that requires many iterations and some spectroscopic intuition.

Discussion: The variational methodology provides the most complete line list that extrapolates well to high energy. However, even after fitting to experimental data, variational line lists do not typically achieve spectroscopic accuracy and are typically less accurate than perturbative methods in regions where effective Hamiltonians perform well.

Utilization: The ExoMol data base (Tennyson & Yurchenko 2012; Tennyson et al. 2020) is the most comprehensive source of variational spectroscopic data. This data base contains spectroscopic models and variational line lists for over 70 diatomic molecules. These data are largely generated using the DUO variational nuclear motion program (Yurchenko et al. 2016a), although LEVEL (Le Roy 2017) has also been used for some (largely closed shell) molecules.

Uncertainties: Uncertainties for variational line list data are very dependent on the molecule and energy level. Typically, uncertainties are best evaluated by comparing variational data with experimentally derived energy levels (like those from MARVEL) whenever possible and then extrapolating to similar rovibronic states with no available experimental data, increasing uncertainties with v and J.

Electronic states for which there is no available experimental data have the largest uncertainty as this relies on the cancellation of errors in the quantum mechanical treatment between states usually of different spin symmetry. For 3d transition metal diatomics with modern quantum chemistry approaches, the review by Tennyson et al. (2016) found that errors in excess of 1000 cm^{-1} are common and thus purely computational predictions are clearly unsuitable for direct comparisons with experimental spectra. Typically, these states are included to increase the accuracy of the partition functions (which are not as sensitive to errors) and to understand the scale of the missing absorption to direct future experiments.

The uncertainty in the position of excited vibrational states purely known from experiment is smaller, with fundamental frequencies typically predicted to an accuracy of around 20 cm^{-1} or better theoretically (Tennyson et al. 2016) even for complex 3d transition metals and much better for smaller molecules. Nevertheless, this is often not accurate enough to model absorption bands at the resolution needed for *JWST* and thus having experimental data at least for the $v = 0, 1$ and ideally the $v = 0, 1, 2$ levels is a strong advantage. We note that, for well-studied molecules, typically the available experimental data is for the strongest absorption bands as these are the easiest experiments.

Uncertainties are estimated for variationally calculated levels assuming that the errors grow linearly with vibrational excitation but

quadratically with rotational excitation; these assumptions accord with our previous experience for a variety of this species, as first adopted in Yurchenko et al. (2020). We therefore estimated uncertainties using the expression:

$$\Delta\tilde{E}_{\rm Ca} = a \cdot v + b \cdot J(J+1) + c, \quad (2)$$

where a, b, and c are constants in units of cm^{-1}. c is the initial minimum uncertainty determined for each electronic state. a and b quantify the uncertainty scales for the vibrational v and rotational $J(J+1)$ parameters, respectively.

2.2.3 Predicted shift correction

Description: Levels in variational line lists are generally computed to a higher maximum total angular momentum quantum number J than are available from experimental data. While ab initio calculations and refinement to experimental energies often yield highly accurate energies, extrapolations must still be made to levels outside the observed range. It is also often the case that experiments do not provide complete level coverage up to the highest observed J level. An improved estimate of the energies of the unobserved rotational states can be obtained by applying a shift to the calculated energy levels based on the trends in the observed minus calculated (obs.-calc.) energy differences in that band, where 'obs.' refers to the experimentally derived MARVEL energy levels and 'calc.' to the DUO variationally computed energy levels. This is done by fitting to the obs.-calc. trends in a given spin-orbit and parity component of a vibronic state, defined by the electronic state, vibrational quantum number v, Ω, and rotationless parity, and predicting synthetic obs.-calc. values for the missing levels.

Discussion: Applying these synthetic energy shifts to the variationally calculated energies allows for a more extensive provision of high-accuracy energy levels, particularly when interpolating energy shifts for levels within the J range of experimental data. With this method, conservative energy shifts can be predicted for levels outside the observed J range, extrapolating to the maximum J in the variational calculation. This also allows for predictions to be made for levels with lower values of J than are observed in experiment, which is particularly useful in cases where low-J transitions lie in clustered band heads and cannot be assigned. Similarly, this approach is effective at correcting the calculated energies of levels near electronic perturbations and resonances, which can often be difficult to accurately model in a spectroscopic model. Applying this method removes any potential discontinuities between MARVELized and calculated energy levels, which is particularly important for states that have a poor fit to experimental data.

Utilization: The predicted shift methodology has been employed in the construction of high-resolution line lists for AlO (Bowesman et al. 2021) and VO (Bowesman et al. 2024b). This AlO line list has been successfully utilized in the detection of near-infrared electronic bands in the atmosphere in an eruptive young stellar object (Guo et al. 2024).

Uncertainties: The uncertainties for PS levels for every state within the J range for which we have experimental data were calculated the same way as the energy values (see Section 2.2.3), as a fit of the MARVEL uncertainties. When extrapolating to lower J values, the PS level uncertainties were estimated to be simply equal to the standard deviation σ of the obs.-calc.

Outside this range, the PS energy uncertainties were estimated as a function of the standard deviation of the known obs.-calc. data point in the band, σ. This σ value was used as a starting minimum uncertainty for the extrapolation; uncertainties were scaled for a level based on how much larger its J value was than the maximum J that occurred for the MARVEL data in that band, $J^{\rm Ma}_{\rm max}$. Taking $J_{\rm ext} = J - J^{\rm Ma}_{\rm max}$, the following equation was used:

$$\Delta\tilde{E}_{\rm PS} = a \cdot J_{\rm ext}(J_{\rm ext}+1) + \sigma, \quad (3)$$

where a and $\tilde{E}_{\rm PS}$ are determined for each individual molecule.

2.2.4 Isotopologue-extrapolation correction

Description: A spectroscopic model is often only optimized for the main isotopologue of a given molecule due to the lack of assigned experimental spectral data for other isotopologues. Computationally, however, a variational line list for the other isotopologues can easily be produced by changing the mass(es) of the atoms in the main spectroscopic model. Consequently, an improved pseudo-hybrid line list for an isotopologue can be generated by shifting each energy with reference to the corresponding energies in the main variational and hybridized line lists (Polyansky et al. 2017) using

$$\tilde{E}^{\rm iso}_{\rm IE} = \tilde{E}^{\rm iso}_{\rm Ca} + \left(\tilde{E}^{\rm main}_{\rm hybrid} - \tilde{E}^{\rm main}_{\rm Ca}\right), \quad (4)$$

where $\tilde{E}^{\rm main}_{\rm Ca}$ is the calculated variational energy of the main isotopologue, $\tilde{E}^{\rm main}_{\rm hybrid}$ is its hybridized energy, $\tilde{E}^{\rm iso}_{\rm Ca}$ is the calculated variational energy of a given isotopologue, and $\tilde{E}^{\rm iso}_{\rm IE}$ is its isotopologue-extrapolated (IE) energy. In other words, it is assumed that the energy residuals between the hybridized and variational line lists are constant for all isotopologues.

Discussion: Isotopologue-extrapolation is a very effective way to provide improved accuracy for isotopologue spectra, relying on the fact that the impact of nuclear mass on vibrational energy levels is well understood and that the underlying potential energy surface is to a high level of approximation identical.

Utilization: The usefulness of the isotopologue methodology can perhaps best be illustrated by considering TiO, for which the dominant isotopologue typically only comprises 73 per cent of abundance but is the subject of almost all experimental studies. By using the isotopologue-extrapolated method, line lists for ^{46}Ti^{16}O, ^{47}Ti^{16}O, ^{49}Ti^{16}O, and ^{50}Ti^{16}O (McKemmish et al. 2019) were successfully utilized by Pavlenko et al. (2020) to obtain Ti isotopic abundances in a number of spectral regions for two M-dwarf stars. They have also been employed in the successful detection of TiO in the atmospheres of several hot Jupiter exoplanets (Edwards et al. 2020; Serindag et al. 2021; Prinoth et al. 2022; Edwards & Changeat 2024).

Note that the isotopologue-extrapolation method has been in use since at least Polyansky et al. (2017), who describe the procedure as pseudo-experimental, with various terminology. This paper intends to standardize this for clear communication (McKemmish 2021).

Uncertainties: The uncertainties for IE levels are typically scaled by a constant with respect to those of the main isotopologue. This constant may be determined through comparison to experimental data if available.

For Ca uncertainties, equation (2) should be applied for each isotopologue with modified a, b, and c parameters. This is because isotopologues with a reduced mass greater than that of the main isotopologue will in general have a higher density of states and it thus may not always be possible to match all energy levels by assignment.

2.3 Construction of hybrid line lists

2.3.1 Approach

Our goal is to produce a line list that is as complete and as accurate as possible, which we achieve by combining multiple sources of data. Specifically, the aforementioned experimental, perturbative, and variational line lists are collated into a hybrid line list with energy levels then modified using predicted shifts; isotopologue line list accuracy can be enhanced by using the isotopologue-extrapolation method described above. The hybrid line lists produced in this way exploit the unique advantages of each methodology and ameliorate the limitations in order to produce the most accurate and comprehensive source of spectroscopic data for a given molecule.

To construct the final `.states` file for a hybridized line list, we start with a set of `.states` files in a standard and consistent ExoMol format (Tennyson et al. 2014, 2020), each containing predicted energies with uncertainty estimates for all quantum states of interest. The length of each of these `.states` files is expected to vary with the highest quality data (usually MARVEL) typically having the fewest energy levels and providing the desired high-accuracy spectral information, while the lowest quality data (usually variational) incorporates all energy levels considered within the line list and provides the completeness for the final line list. Using a loop over all quantum states, a final hybridized `.states` file is constructed by using the highest quality energy level prediction available for each state and assigning the appropriate label and predicted uncertainties. The highest quality data can be identified either by comparison of the compiled uncertainties or through a pre-determined hierarchy (e.g. experimental data is more accurate than perturbative data which is more accurate than variational data).

2.3.2 Challenges with hybridization

At this stage, significant deviations of the predicted energies from different data sources for the same quantum state should be flagged and manually reviewed. Misassignments of experimental transitions can be uncovered. However, more commonly, there is a misalignment between the sources in determining the quantum numbers used to describe each rovibronic state. Both issues are more common in complex diatomic molecules where there is strong mixing between electronic states. Anecdotally, differences in assignment of quantum numbers between DUO computationally assigned variational and model Hamiltonian assigned experimental states becomes quite common for large J especially when electronic states are coupled like for transition metal diatomic molecules with complex electronic structures. On occasion it has proved necessary to perform a thorough reassignment of the experimental transitions to obtain a consistent solution, see Yurchenko et al. (2016b) for example.

The origin of this challenge is that most of the quantum numbers used to describe states are not rigorous and thus must be defined approximately manually or by computer programs such as DUO. Formally, as DUO diagonalizes the total rovibronic Hamiltonian using Hund's case (a) functions, a rovibrational state with a given electronic state can be uniquely defined (in the absence of hyperfine couplings) using its total parity, vibrational v quantum number, total angular momentum excluding nuclear spin J quantum number, and its projection along the internuclear axis Ω ($= \Lambda + \Sigma$) quantum number. Aside from the rigorously defined J and parity, v and Ω are simply convenient labels that are assigned based on dominant contribution to the total wave function. From experience, the misassignment of these two quantum numbers becomes common for sufficiently large J, especially in cases where mixing arises from large spin-orbit or other couplings. Due to the overall completeness of the variational methodology, however, the rovibronic states population can be trusted (i.e. all states are calculated without any gaps), although the energy of each state may not be consistent with experimental data. For this reason, v and Ω may be simply reassigned based on energy orderings.

2.3.3 Line list data product

At the conclusion of the energy-hybridized line list construction process, a final `.states` file is constructed that typically looks similar to Table 2 (extract from the $^{24}Mg^{16}O$ line list constructed in Section 3 of this paper).

As stated above, the final hybrid line list retains the `.trans` file from the variational line list.

Each ExoMol line list is given a name and version number (the date in YYYYMMDD format) which appears in the definitions (`.def`) file for each isotopologue/line list combination (Tennyson et al. 2024b). By convention updates to the `.states` file retain the line list name but generate a new version number. Any changes to `.trans` file, in particular to the Einstein A coefficients, would result in a new name for the line list. Due to energy changes arising from the hybridization procedure, transitions may yield negative frequencies after updating an existing `.states` file. These transitions have been removed herein without changing the line list name.

2.4 Analysis: suitability of line list for molecule detection using high-resolution cross-correlation techniques

In order to ensure effective utilization of ground-based telescopes, astronomers need to know where each line list can be trusted for high-resolution studies. Similarly, to support these efforts, experimentalists need to know the most important gaps in existing data to target to improve coverage and accuracy.

To support these goals, here we describe two relatively new powerful visualizations and how these should be interpreted; these visualizations have already been produced for some recent molecular line lists including ZrO (Perri, Taher & McKemmish 2023) and NH (Perri & McKemmish 2024).

Both visualizations aim to understand the suitability of a line list for high-resolution cross-correlation studies (HRCC) by looking at the transitions whose frequencies are predicted to high accuracies. The simplest approach, which we adopt in this paper, is to consider all MARVEL energy levels to be sufficiently accurate for HRCC studies and also consider the predicted shift energy levels to be potentially sufficiently accurate. This is only an approximation, however, as some experimental energy levels might have higher uncertainties; for example, for MgO, only 21.4 per cent of the transitions between the MARVEL energy levels have an accuracy which corresponds to a resolving power $R = \lambda/\Delta\lambda$ larger than 100 000.

2.4.1 Transition source type plots

The transition source type plots Figs 3(b), 6(a) and (b) show the cumulative density of transitions as a function of intensity figures. These plots should be read by looking at vertical slices – these visually compare the data sources used to predict the frequencies of all transitions with intensity larger than the x axis value. For example, taking a vertical slice at $x = 10^{-20}$ cm per molecule in the TiO figure, we are considering *all* transitions in $^{48}Ti^{16}O$ with intensities *above* 10^{-20} cm per molecule at 2000 K (i.e. not just those transitions binned in this intensity range). The colours and y axis shows the origin of the energy of the upper and lower quantum state for these transitions; e.g. 20 per cent of the transitions are 'Ma-Ma' meaning the upper and

Table 2. A part of the final `.states` file for $^{24}Mg^{16}O$. The same structure is used for TiO and VO line lists.

i	$\tilde{E}$	g_{tot}	J	unc	τ	g	$p_{+/-}$	$p_{e/f}$	State	v	Λ	Σ	Ω	Label	$\tilde{E}_{Ca}$
1	0.000000	1	0.0	0.000000	NaN	0.000000	+	e	X(1SIGMA+)	0	0	0.0	0.0	Ma	0.000000
2	774.738739	1	0.0	0.001420	1.3596E+00	0.000000	+	e	X(1SIGMA+)	1	0	0.0	0.0	Ma	774.737704
3	1539.174552	1	0.0	0.001730	6.3331E-01	0.000000	+	e	X(1SIGMA+)	2	0	0.0	0.0	Ma	1539.136871
4	2292.687940	1	0.0	0.003610	3.9854E-01	0.000000	+	e	X(1SIGMA+)	3	0	0.0	0.0	Ma	2293.035126
5	2621.208578	1	0.0	0.004809	2.5844E-02	0.000000	+	e	a(3PI)	0	1	−1.0	0.0	PS	2621.207376
6	3037.070194	1	0.0	0.040000	2.2228E-01	0.000000	+	e	X(1SIGMA+)	4	0	0.0	0.0	Ca	3037.070194
7	3265.064334	1	0.0	0.043377	2.0867E-02	0.000000	+	e	a(3PI)	1	1	−1.0	0.0	PS	3264.977473
8	3770.264143	1	0.0	0.050000	9.7014E-02	0.000000	+	e	X(1SIGMA+)	5	0	0.0	0.0	Ca	3770.264143
9	3900.848866	1	0.0	0.030000	1.8289E-02	0.000000	+	e	a(3PI)	2	1	−1.0	0.0	Ca	3900.848866
10	4478.700487	1	0.0	0.040000	2.3189E-02	0.000000	+	e	a(3PI)	3	1	−1.0	0.0	Ca	4478.700487
11	4543.125919	1	0.0	0.060000	2.7738E-02	0.000000	+	e	X(1SIGMA+)	6	0	0.0	0.0	Ca	4543.125919

[a] i: State counting number.
[b] $\tilde{E}$: Energy (in cm^{-1}).
[c] g_{tot}: Total degeneracy.
[d] J: Total angular momentum.
[e] unc ($\Delta\tilde{E}$): Uncertainty (in cm^{-1}).
[f] τ: Lifetime (in s; calculated in ExoCross (Yurchenko, Al-Refaie & Tennyson 2018b)).
[g] g: Lande g-factor.
[h] $p_{+/-}$: Total parity.
[i] $p_{e/f}$: Kronig rotationless parity.
[j] State: Electronic state.
[k] v: Vibrational quantum number.
[l] Λ: Projection of electronic orbital angular momentum on the internuclear axis.
[m] Σ: Projection of electronic spin angular momentum on the internuclear axis.
[n] Ω: Projection of the total angular momentum excluding nuclear spin along the internuclear axis.
[o] Label (see Table 1): Ma for experimental MARVEL energies, PS for energies from Predicted Shift method, Ca for unchanged calculated energy by Li et al. (2019).
[p] $\tilde{E}_{Ca}$: DUO calculated energy by Li et al. (2019).

lower state energies are both obtained from a MARVEL analysis and thus highly reliable, while approximately 30 per cent are 'Ma-Ca' meaning one energy level is from MARVEL and one is from DUO.

In the transition source type plots, the frequencies for transitions with source 'Ma-Ma' are highly accurate (and thus suitable for molecule detection using HRCC techniques), while other transitions with other sources have decreasing accuracy in frequency down the legend (e.g. 'Ca-Ca' have least accurate transition frequencies).

Typically, the strongest transitions have far more highly accurate 'Ma-Ma' transitions than weaker transitions. This occurs because experimentally more intense transitions are easier to observe. Fortunately, strong transitions are also the transitions for which high accuracy in frequencies is most important in cross-correlation analyses.

2.4.2 *High-res versus total cross-section plots*

The second type of plot that is extremely useful for the analysis of a line list's suitability for HRCC is the high-resolution (high-res) versus total cross-section plots, shown in Figs 3(d), 5(c), and 7(b). The dark black line shows the total cross-section for the line list while the coloured line shows the cross-section as a function of wavenumber (or equivalently wavelength on the upper horizontal axis) obtained only using transitions with frequencies known to very high precision (usually MARVEL energies for both the upper and lower state). For the temperature considered, in those spectral regions where the coloured and black line visually overlap, the current line list is entirely suitable for HRCC analysis; in contrast regions where the coloured line is much lower in intensity than the black line are not suitable for cross-correlation analysis because many intense transitions are not known to very high experimental accuracy.

It is important to note that these high-res versus total cross-section plots will change substantially at different temperatures. However, these plots can be produced at a given temperature using ExoCross using filtering based on the source type of the upper and lower energy levels.

3 MGO: INCREASING ACCURACY BY ADDING MARVEL ENERGY LEVELS

3.1 Background

Due to the cosmic abundance of magnesium and oxygen and the strength of the bond, diatomic magnesium oxide (MgO) in the gas phase is known to be present in the upper atmosphere of Earth (Correira et al. 2008) and Mercury (Sarantos et al. 2011; Stockstill-Cahill et al. 2012), where it is believed to be produced by micrometeoroid impact. However, searches for gaseous MgO in the interstellar medium have thus far been unsuccessful (Turner & Steimle 1985; Sakamoto et al. 1998).

In its solid form, magnesium oxide is thought to be one of the most abundant rocks in the interiors of planets (Pan et al. 2023) and to be a component of interstellar dust (Maclean & Duley 1982; Mogren Al Mogren et al. 2015; Yoneda & Grossman 1995; Nozawa et al. 2003; Rietmeijer 2009); certainly it is an important part of Earth's mantle (Correira et al. 2008), but the chemical identity of solid species is difficult to confirm remotely.

However, in ultra hot rocky exoplanets like hot rocky super-Earths, often referred to as lava or magma worlds, the higher temperatures in the atmosphere are thought to favour the production of MgO vapour (Tennyson & Yurchenko 2017; Schaefer, Lodders & Fegley 2012; Zilinskas et al. 2022; Falco et al. 2024). With high-quality spectral line lists, these hypotheses can be experimentally tested.

Table 3. Overview of the MgO MARVEL compilation energy levels (EL) and comparison against DUO variational calculated energy levels (|Ma-Ca|), in cm^{-1}.

State	v range	J range	#ELs	Unc. range	Avg. Unc.	EL range	|Ma-Ca| range	Mean |Ma-Ca|
X $^1\Sigma^+$	0-4	0-61	125	0-0.25	0.06	0-3042.13	0-1.54	0.07
A $^1\Pi$	0-3	1-62	368	0.001-0.25	0.09	3504.32-7055.19	0.0003-0.1	0.02
B $^1\Sigma^+$	0-1	0-42	71	0.003-0.14	0.06	20003.59-21387.62	0.00004-0.06	0.02
a $^3\Pi$	0-1	1-43	256	0.001-0.15	0.07	2551.29-4129.27	0.00006-0.33	0.03

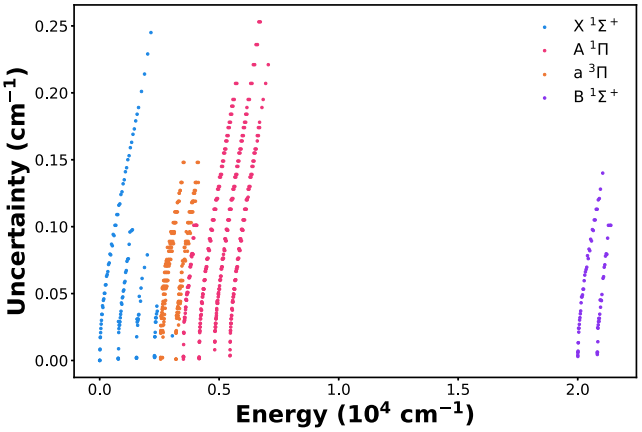

Figure 1. The MARVEL (Ma) energy uncertainties for $^{24}Mg^{16}O$ as a function of energy.

Li et al. (2019) produced the LiTY variational line list covering the isotopologues $^{24}Mg^{16}O$, $^{24}Mg^{17}O$, $^{24}Mg^{18}O$, $^{25}Mg^{16}O$, and $^{26}Mg^{16}O$. While this line list was heavily tuned to experimental data, no formal MARVEL project was undertaken. This line list thus has the completeness and accuracy necessary for molecular detection in moderate resolution observations, but not the accuracy needed for high-resolution cross-correlation studies.

The existing LiTY spectral data is likely already sufficient for enabling detection with space-based telescopes like *JWST*. However, more definitive detection is obtained with ground-based telescopes using high-resolution cross-correlation techniques; this is enabled by the updates of this paper to the LiTY line list data specifically the explicit inclusion of MARVEL energies.

Here, we facilitate the detection of MgO in high-resolution studies by performing a MARVEL analysis of $^{24}Mg^{16}O$ and updating the line list to get spectroscopic accuracy for many of the strong spectral lines.

3.2 Data sources for energy levels

3.2.1 Experimental energy levels from a new MgO MARVEL analysis

A MARVEL analysis for the spectroscopic data of $^{24}Mg^{16}O$ was performed with results summarized in Table 3. We used MARVEL4.1 and produced the MARVEL energy uncertainties with the bootstrapping method as described in Tennyson et al. (2024a). The MARVEL uncertainties are shown in Fig. 1.

A total of 1181 transitions were collected from nine sources (Azuma et al. 1984; Steimle, Azuma & Carrick 1984; Törring & Hoeft 1986; Ip et al. 1991; Civis, Hedderich & Blom 1991; Kagi et al. 1994; Mürtz et al. 1994, 1995; Kagi & Kawaguchi 2006). 1169 of these transitions are validated through the MARVEL procedure and inverted to determine 820 empirical energy levels and uncertainties for the four lowest electronic states (X $^1\Sigma^+$, a $^3\Pi$, A $^1\Pi$, B $^1\Sigma^+$) and vibrational levels $v = 0 - 4$, with the highest rotational quantum number being $J_{max} = 62$.

The relevant details can be found within the supporting information, including a full list of all the experimental data used in this analysis, a segment file connecting the sources to the wavenumber units used, a justification of the uncertainties considered for the line positions, a list of our MARVEL energy levels, and a list of papers considered but not used in this work as well as the reasons for their exclusion.

For energy levels with MARVEL replacement, uncertainties were taken from the MARVEL procedure (see Fig. 1).

3.2.2 Variational energy levels

Here we retain the spectroscopic model used for the LiTY line list for $^{24}Mg^{16}O$ in Li et al. (2019). However, the availability of high-quality experimental data allows an assessment of the accuracy of the spectroscopic model and thus a improves the quality of our predicted uncertainties. For the four lowest lying electronic states we find a high accuracy, with an average error of 0.029 cm^{-1} and maximum value 0.2 cm^{-1} within the states for which we had experimental data and applied the PS method (see Table 3 for J and v ranges). For all other cases, the errors exceed this value. For calculated levels, the errors are within 2 cm^{-1} for J values up to 140 and increase to a maximum of 10 cm^{-1} for the highest J values ($J > 300$).

The uncertainties of the calculated levels (Ca) were considered to have a linear dependence on v and a quadratic dependence on J following the relation given in equation (2). Here, we adopted values of $a = 0.0001$ cm^{-1} and $b = 0.01$ cm^{-1} for all electronic states, $c = 0$ cm^{-1} for the ground state, and $c = 0.01$ cm^{-1} for all excited electronic states to estimate the uncertainties, with a maximum cutoff of 10 cm^{-1}. Unfortunately, due to a lack of experimental data for the b $^3\Sigma^+$ state we have little information to base the uncertainties of the energy levels in this state on; we estimated an uncertainty of 10 cm^{-1} for all energy levels in this state.

3.2.3 Predicted shift energy levels

The predicted shift methodology was used to calculate energy levels. Where necessary, uncertainties were extrapolated using $a = 0.0001$ cm^{-1} (see equation 3) and a maximum cutoff for $\Delta\tilde{E}_{PS}$ of 10 cm^{-1}.

The way in which the predicted shift levels interpolate and extrapolated from the MARVEL energies to the different rotational levels of the same vibronic state is visualized most clearly in Fig. 2.

3.2.4 Isotopologue-extrapolation energy levels

The line lists for the $^{24}Mg^{17}O$, $^{24}Mg^{18}O$, $^{25}Mg^{16}O$, and $^{26}Mg^{16}O$ isotopologues were updated using the isotopologue extrapolation

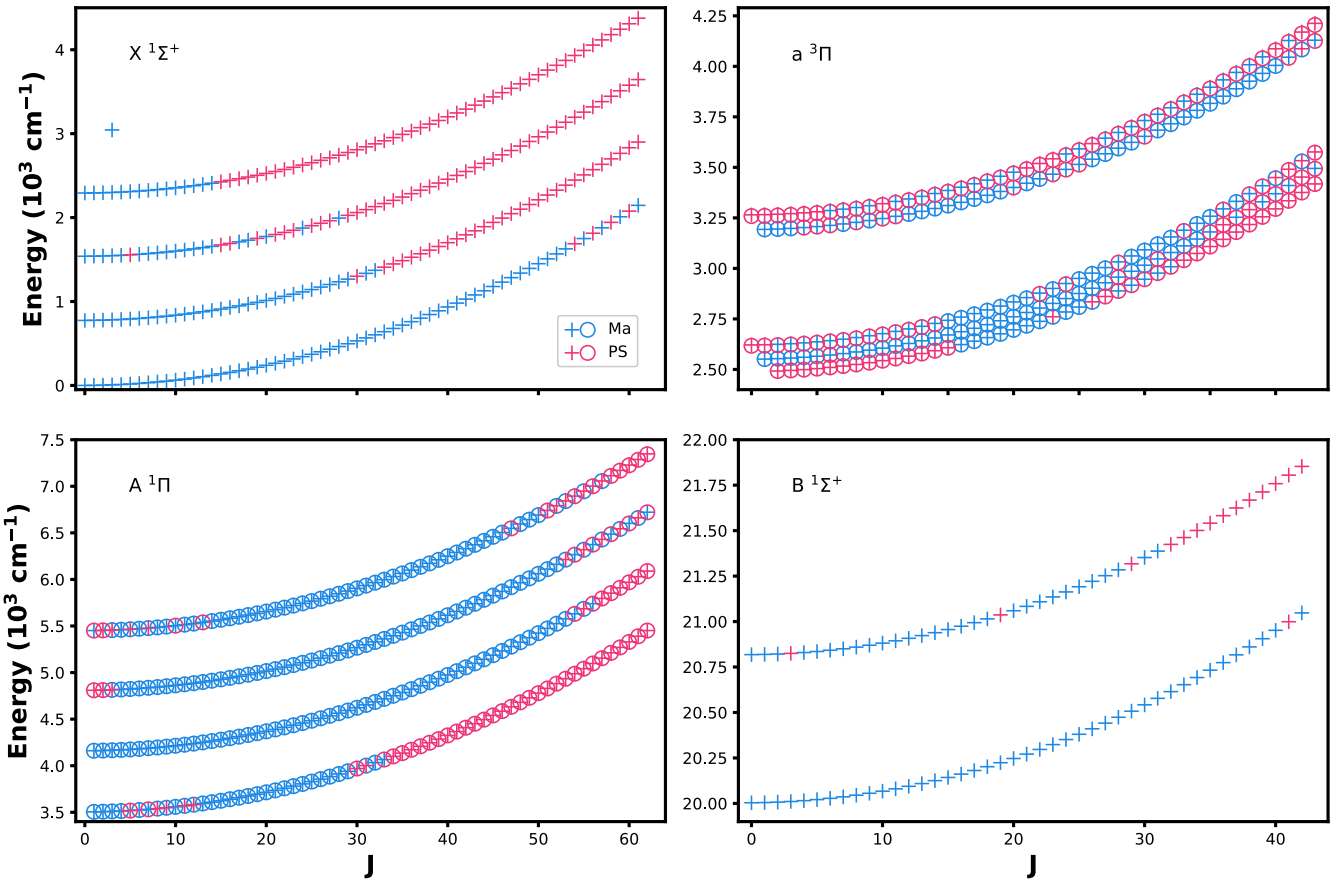

Figure 2. The updated MARVEL (Ma) and predicted shift (PS) energies for the X $^1\Sigma^+$, A $^1\Pi$, a $^3\Pi$, B $^1\Sigma^+$ electronic states of $^{24}Mg^{16}O$ as a function of the J quantum number. The 'o' symbol indicates f parity and the '+' symbol indicates e parity.

(IE) correction methodology. For the energy levels of the main isotopologue which were updated either with a MARVEL energy (Ma) or with the predicted shifts (PS), we updated the respective level of the other four isotopologues applying the same final pseudo-experimental correction according to equation (4).

In total, 7017 energy levels were updated for each of the five isotopologues for the four lowest electronic states using the IE method. The energy uncertainties for the four isotopologues were estimated as twice the uncertainty of the main isotopologue. For the cases without matches to the main (labelled as Ca in the `.states` file), equation (2) was used with a maximum cutoff of 10 cm^{-1}, and with $a = 0.0001$ cm^{-1} $b = 0.01$ cm^{-1} for all electronic states, and c as the average shift for each electronic state.

For MgO, we are able to directly verify the quality of our isotopologue-extrapolation correction by comparing against experimental data. Specifically, we assessed a small part of the updated line lists for the isotopologues $^{25}Mg^{16}O$ and $^{26}Mg^{16}O$ for which we found experimental data recorded by Kagi et al. (1994) and Törring & Hoeft (1986): Kagi et al. (1994) recorded transitions in the A $^1\Pi$ – X $^1\Sigma^+$ electronic band and the $v_A - v_X = 1 - 0$ vibrational band with $J = 6 - 41$ for $^{25}Mg^{16}O$ and $J = 6 - 42$ for $^{26}Mg^{16}O$, and Törring & Hoeft (1986) published six microwave transitions with $J = 0 - 7$ for $^{26}Mg^{16}O$ with $v = 0$.

For $^{26}Mg^{16}O$ the average residues in positions against experiment for four rotational X $^1\Sigma^+$-X $^1\Sigma^+$ transitions are 1.2×10^{-4} cm^{-1} without IE correction and 6.7×10^{-6} cm^{-1} with IE correction, while for the rovibronic A $^1\Pi$-X $^1\Sigma^+$ band, the IE correction had a more modest reduction in errors from 0.027 to 0.021 cm^{-1}. The experimental data mostly agreed with our line list energies within their mutual estimated uncertainties.

Similar comparisons for $^{25}Mg^{16}O$ rovibronic A $^1\Pi$ –X $^1\Sigma^+$ transitions found that for the P and Q branches, applying the IE correction led to a reduction in average residual from 0.017 to 0.012 cm^{-1}. However, the residuals associated with the R branch transition were notably larger (0.05 cm^{-1} average) and did not significantly change with applying the IE correction; the likely cause of this difference is incorrect assignments of the experimental data.

3.3 Line list

3.3.1 Construction

The LiTY line list (Li et al. 2019) for $^{24}Mg^{16}O$ was updated by incorporating 820 MARVEL (Ma), 6197 predicted shift (PS) energy levels into the line list, with 179 825 variationally calculated (Ca) energy levels retained for vibronic levels without any experimental data. The distribution of energy levels within each electronic state depending on their source is shown in Fig. 3(a). The pattern that emerges of MARVEL energy levels focused on the low lying rovibronic states in each electronic state complemented by the predicted shift and calculated levels is common for most molecules.

The `.states` files for the $^{24}Mg^{17}O$, $^{24}Mg^{18}O$, $^{25}Mg^{16}O$, and $^{26}Mg^{16}O$ isotopologues were updated using the isotopologue extrapolation (IE) correction methodology as detailed in Section 3.2.4.

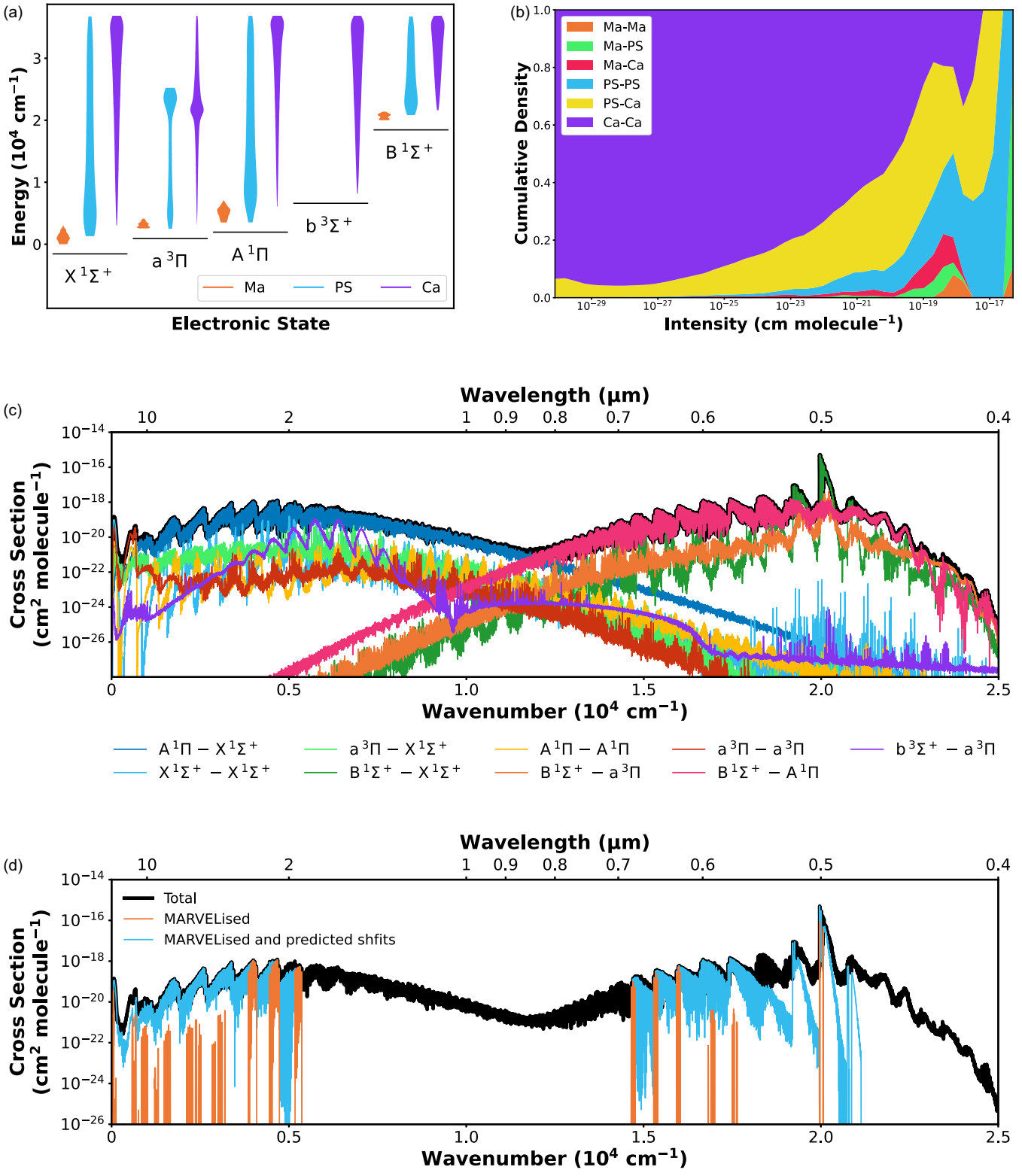

Figure 3. Details of updated LiTY line list for $^{24}Mg^{16}O$. (a) The energy distribution for the rovibronic states of $^{24}Mg^{16}O$ as a function of energy source type in each electronic state. (b) The transition source type for $^{24}Mg^{16}O$, i.e. the cumulative density of transitions as a function of intensity depending on energy source type, computed at 2000 K using the program ExoCross (Yurchenko et al. 2018b). The $^{24}Mg^{16}O$ absorption cross-section computed at 2000 K using the program ExoCross with Gaussian line profiles of 1.0 cm^{-1} half-width half-maximum. The uppermost (black) cross-section shows all transitions in the new 2024 Toto line list with decomposition into dominant electronic bands. (d) The upper (black) cross-section shows all transitions in the line list, whereas the lower (orange) cross-section shows only MARVEL (Ma) experimental transitions (with variational intensities), and the intermediate (blue) cross-section shows all possible transitions between MARVEL (Ma) and predicted shift (PS) energy levels (with variational intensities).

The `.trans` files for all isotopologues were cleaned to remove negative transition frequencies. Specifically, 92, 136, 139, 131, and 127 transitions were removed for the $^{24}Mg^{16}O$ $^{24}Mg^{17}O$, $^{24}Mg^{18}O$, $^{25}Mg^{16}O$, and $^{26}Mg^{16}O$ isotopologues, respectively.

The absorption cross-section between the several electronic bands computed at 2000 K is illustrated in Fig. 3(c). The strongest transitions around 500 nm are the B $^1\Sigma^+$ - X $^1\Sigma^+$ band, but the hot band B $^1\Sigma^+$ - A $^1\Pi$ transition dominates for most of the visible spectra with the A $^1\Pi$ - X $^1\Sigma^+$ transition dominating the near IR region. The spin forbidden B $^1\Sigma^+$ - a $^3\Pi$ band displays significant intensity due to a $^3\Pi$ - A $^1\Pi$ spin-orbit coupling that mixes their wavefunctions. The experimentally unobserved band b $^3\Sigma^+$-a $^3\Pi$ is confirmed to be less intense than the rovibronic A $^1\Pi$ - X $^1\Sigma^+$ manifold across the full spectral region.

3.3.2 Suitability for high-resolution cross-correlation

The transition source type plot (described in Section 2.4) for MgO is shown in Fig. 3(b). This peak in Ma and PS involved transitions near $10^{-18/19}$ is atypical and indicates that some of the weaker spectral bands have been more heavily studied than the spectral band that is strongest at 2000 K. The MARVEL data for this molecule is reasonably modest with the predicted shift data is crucial for dramatically extending the number of transitions whose wavenumbers are known to reasonably high accuracy.

The high-res versus total cross-section of MgO is shown in Fig. 3(d) at 2000 K. In spectral regions where the MARVEL and total cross-sections overlap, then the line list is very suitable for HRCC. In the case of MgO, due to the limited experimental data, this region is very small (a few narrow windows between around 1.9 and 3 μm, and between 640–690 nm). In the absence of perturbations, predicted shift may be sufficiently reliable for HRCC especially for high signal-to-noise observations; in this case, the suitable spectral windows for HRCC in MgO extend to approximately 580–690 nm and 1.9 to 3 μm).

While cross-correlation techniques are most sensitive to the strongest features in spectra, good coverage across the wavelength region being targeted for cross-correlation is desirable to reduce noise and obtain an unambiguous detection.

Note that this analysis has been done at 2000 K; cooler temperatures would likely increase the suitability of the line list for HRCC analysis since the experimental data is more complete at lower rotational quantum numbers that would be more populated at lower temperatures.

3.4 Future work

Figs 3(b) and (d) show clearly that more high-resolution experimental data for the four lowest electronic states of MgO are high priority to enable more robust searches for this molecule using high-resolution cross-correlation techniques. In particular, the experimental data was limited to modest rotational quantum numbers, meaning that the band peak and full range couldn't be predicted for the hot 2000 K environments. Nevertheless, the PS methodology allowed extrapolation to higher J than the experimental data and may be sufficiently reliable to enable HRCC detections of MgO.

Predicted shift methodology is only feasible if some rotational states of a particular vibronic level are known. Data for the $v = 2, 3$ B $^1\Sigma^+$ state and any levels of b $^3\Sigma^+$ electronic states would be welcome.

4 TiO: DETAILS AND UPDATES ARE CRITICAL

4.1 Background

Diatomic titanium oxide (TiO) has long been known to be a critical molecular absorber in cool stellar atmospheres with temperatures around 1700–2500 K (Allard, Hauschildt & Schwenke 2000). After initial controversial detections (Nugroho et al. 2017; Sedaghati et al. 2017) and some unexpected non-detections (Merritt et al. 2020), it is now generally accepted to also be present in at least some hot Jupiter exoplanets including WASP-189 b (Prinoth et al. 2022) with further tentative detections in WASP-69 b (Ouyang et al. 2023) and potentially WASP-33 b (Serindag et al. 2021). The presence of TiO in these planets' atmospheres is thought to contribute to the formation of temperature inversions (Piette et al. 2020), where the upper atmosphere is hotter than the lower atmosphere. To understand these complex atmospheric processes, reliable measurements of TiO abundance are required. Given there has been significant controversy surrounding TiO presence in some hot Jupiter exoplanets (Serindag et al. 2021), verifying the completeness and accuracy of the TiO spectral data is crucial to enable robust examination of modelling results. New experimental data filling existing gaps should be incorporated into the line list as quickly as feasible (McKemmish 2021).

The ExoMol TiO line list, Toto, was published by McKemmish et al. (2019) and contained 301 245 Ca (variational DUO) and 17 802 Ma (experimentally derived MARVEL) energy levels. In mid-2021, the Toto line list for TiO was updated to correct an oversight in the original 2019 release in which the a $^1\Delta$ and e $^1\Sigma^+$ electronic states were initially not MARVELized. This 2021 update replaced the original DUO variational energy levels with experimentally derived MARVEL energy levels for 660 energy levels in the a $^1\Delta$ and 98 energy levels of the e $^1\Sigma^+$ state. This update improved the high-resolution accuracy of the line list, particularly in the spectral regions dominated by transitions involving these states, particularly 2000–2500 cm^{-1} (4–5 μm), 3000–3500 cm^{-1} (2.8–3.3 μm), 9000–9500 cm^{-1} (1.05–1.10 μm), 10 000–10 500 cm^{-1} (0.95–1.00 μm), and 11 000–11 500 cm^{-1} (0.87–0.91 μm).

Here, we provide significant updates to these original line lists for all isotopologues. With the release of this paper, the 2021 update will be superseded by the 2024 update of the Toto TiO line list. All the advantages of the 2021 update will be retained in this 2024 update.

Specifically, in this 2024 TiO update, we construct a hybridized line list by combining the calculated variational DUO energy levels (Ca) from the 2019 Toto spectroscopic model (McKemmish et al. 2019), an updated MARVEL compilation (Ma) and new predicted shift energy levels (PS). This update also provides uncertainty estimates for all energies, which were not present in the original 2019 line list. The spectroscopic model for TiO has not been refit for this update and thus the `.trans` file is unchanged from the 2019 or 2021 line lists. As new experimental data becomes available, this is certainly a worthwhile exercise though beyond the scope of this paper.

4.2 Data sources for energy levels

4.2.1 The 2024 MARVEL update

A MARVEL compilation of experimental transition frequencies into a self-consistent spectroscopic network is an ongoing process that should continue in line with the publication of new assigned exper-

Table 4. An overview of the 2024 TiO MARVEL energy levels (EL) and comparison against DUO variational calculated energy levels (|Ma-Ca|) in cm^{-1}. It should be noted that parity is not explicitly considered in the TiO MARVEL compilation for Δ and Φ states and is instead added when hybridizing the full line list.

State	v range	J range	#ELs	Unc. range	Avg. Unc.	EL range	|Ma-Ca| range	Mean |Ma-Ca|
$X\,^3\Delta$	0–5	1–162	5000	0.0000–7.92	0.40	0–14878	0.0000–5.50	0.19
$a\,^1\Delta$	0–3	2–100	638	0.0215–2.02	0.12	3446–8978	0.0001–7.45	0.12
$d\,^1\Sigma^+$	0–5	0–92	402	0.0273–0.36	0.06	5661–12 259	0.0002–0.28	0.05
$E\,^3\Pi$	0–1	0–61	386	0.0100–0.10	0.04	11 838–13 969	0.0022–1.82	0.69
$A\,^3\Phi$	0–5	2–163	5152	0.0002–8.06	0.48	14 021–28 825	0.0000–3.19	0.36
$b\,^1\Pi$	0–4	1–100	806	0.0245–1.20	0.08	14 717–20 507	0.0002–1.05	0.15
$B\,^3\Pi$	0–2	0–148	2275	0.0080–3.62	0.16	16 224–27 727	0.0011–68.2	4.25
$C\,^3\Delta$	0–7	1–158	4756	0.0010–6.32	0.28	19 341–31 462	0.0001–11.1	0.50
$c\,^1\Phi$	0–3	3–101	608	0.0245–2.31	0.13	21 290–26 685	0.0001–8.27	0.20
$f\,^1\Delta$	0–2	2–72	302	0.0245–0.15	0.05	22 515–25 321	0.0004–0.63	0.08
$e\,^1\Sigma^+$	0–1	1–59	98	0.0582–0.11	0.07	29 960–32 515	0.0107–3.57	0.77

imental data (McKemmish 2021). For $^{48}Ti^{16}O$, the MARVEL energy levels were first compiled in McKemmish et al. (2017) and updated for McKemmish et al. (2019) (as part of the 2019 Toto line list construction).

Herein, we use MARVEL4.1 to expand the TiO MARVEL data set to incorporate new experimental data for the $E\,^3\Pi - X\,^3\Delta$ (Bernath & Cameron 2020) (659 transitions, all validated) and $B\,^3\Pi - X\,^3\Delta$ bands (Cameron & Bernath 2022) (5506 transitions, all validated), as well as new experimental transitions for the $X\,^3\Delta$ rovibrational band were included from 19BrWaTu (Breier et al. 2019) (16 transitions, all validated) and 21WiBrDo (Witsch et al. 2021) (514 transitions, all validated).

There were some validation issues; in these cases, the newer data was preferred, and inconsistent older data removed from the final spectroscopic network [most notably, 179 lines from the 73Phillips (Phillips 1973) data were invalidated].

The 2024 TiO MARVEL compilation contains 12 164 energy levels from 61 509 validated (62 935 total) transitions compared to 2019 compilation of 10 761 energy levels from 51 547 validated (56 240 total) transitions. The scope of the data is summarized in Table 4. The most significant difference from the 2019 MARVEL data set is that $B\,^3\Pi$, $v = 2$ lines are available for the first time; there has also been a significant increase in the J range of data for the $E\,^3\Pi$ and $B\,^3\Pi$ states as expected. Furthermore, 515 out of 517 lines by Kobayashi et al. (2002) and Simard & Hackett (1991) in the band $E\,^3\Pi - X\,^3\Delta$ with unresolved upper state parities were assigned e and f parities using the MARVEL energy levels. No significant changes in the $X\,^3\Delta$ state were identified.

We use MARVEL4.1 to produce the MARVEL energy uncertainties shown in Fig. 4. MARVEL4.1 allows for the uncertainties in the energies to be calculated using a bootstrapping approach, as described by Tennyson et al. (2024a). This method accounts for inconsistencies between multiple transitions to or from a given level by applying an increase to the level's final uncertainty. This is achieved through randomly increasing the uncertainties of all transitions by a factor between 2 and 10 and assessing the variance in the final energy values obtained over multiple iterations. Accordingly, the uncertainties in the MARVEL energy levels present in the 2024 update that were also present in the 2019 data set have uncertainties on average 14 per cent larger. These uncertainty changes have a large standard deviation of 87 per cent however, with uncertainties in the extreme cases being up to 200 times smaller or 50 times larger in the 2024 MARVEL data. These uncertainties are incorporated into the 2024 TiO line list update for the MARVELized energy levels.

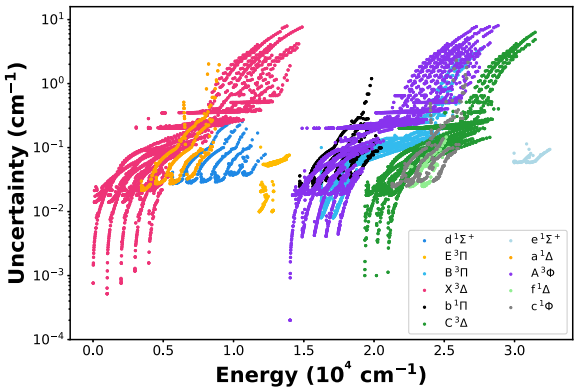

Figure 4. The MARVEL (Ma) energy uncertainties for $^{48}Ti^{16}O$ as a function of energy.

4.2.2 *Calculated variational data*

The spectroscopic model for TiO has not been updated from the 2019 model, so we refer readers to McKemmish et al. (2019) for details. The new MARVEL experimental data offers the opportunity to assess the accuracy of the 2019 TiO spectroscopic model for states against which it was not explicitly fit, offering new insight into the quality of the spectroscopic model's extrapolation. Most notably, we have 2394 new MARVEL energy levels, including 216 in the E state (in $v = 0$) and 1783 in the B state (477, 607, and 699 in v=0, 1, 2, respectively). We note that there are small differences in other electronic states as a result of the MARVEL procedure and correction of quantum numbers (see below).

The $E\,^3\Pi$, $v = 0$ state is very well predicted with residuals less than 1.82 cm^{-1} and uncertainties less than 0.11 cm^{-1} even up to $J = 61$. These errors are well within the anticipated accuracy of the DUO spectroscopic model and are not of concern (indeed this is one of the key reasons for including MARVELized energy levels). While the $B\,^3\Pi$, $v = 0$ and 1 states possess comparable residuals for similar J, the $v = 2$ state was found to diverge from the new MARVEL energy levels even for low J. The calculated deviation from MARVEL energies for $v = 2$ is expected as the initial fitting data only included data from $v = 0$ and 1, and only one fitting parameter was employed in the exponent of the extended Morse oscillator potential. The B state has a maximum residual of 68.19 cm^{-1} for $J = 130$ (the largest residual across the whole line list; 11.08 cm^{-1} outside the B state). A refit of the full spectroscopic model could be merited based on this new data, but this is left to future work.

Uncertainties for the calculated energies of all five isotopologues of TiO were estimated using equation (2), taking $a = 0.0001$ cm^{-1} and $b = 0.05$ cm^{-1}. c values were determined for each electronic state and in the primary isotopologue these were based on the mean obs.-calc. value for the MARVELized energies in that state. For states with no MARVEL data, the mean obs.-calc. value for all MARVEL data was taken, in an attempt to quantify a global uncertainty for the calculations. The same approach was used to quantify uncertainties in the four isotopically substituted species of TiO, where the uncertainty estimates were based on the shifts seen in the isotopically extrapolated energies.

4.2.3 Predicted shift data

Due to the good energy level coverage provided by the MARVEL data for ^{48}Ti^{16}O, it was found that the use of predicted shifts was well suited for estimating the energy residuals in the small number of missing states in the fine structure components of each vibronic branch. To obtain the extrapolated predicted shift uncertainties from equation (3), we employed $a = 0.0001$ cm^{-1}, and σ values that ranged from 0.0009 cm^{-1} for the X $^3\Delta_1$, $v = 0$ state to 5.0 cm^{-1} for the A $^3\Phi_4$, $v = 2$ state.

In this context, the inclusion of predicted shift (PS) energy levels over effective Hamiltonian (EH) data should be evaluated. Notably, Bernath & Cameron (2020) and Cameron & Bernath (2022) recently produced perturbative line lists for the E $^3\Pi$ – X $^3\Delta$ and B $^3\Pi$ – X $^3\Delta$ bands, respectively, based on an effective Hamiltonian fit of these states. Typically, perturbative energy levels (EH) like these are assumed to be less accurate than the experimental MARVEL data (Ma), but more accurate than the variational (Ca) data.

When fit over small gaps in the known energy levels, however, the predicted shift methodology can interpolate with accuracy close to that of the MARVEL energies. This is particularly favourable as predictions can be made for perturbed levels without direct consideration of the underlying couplings or resonances, which would need to be explicitly included in the construction of an effective Hamiltonian.

For spin vibronic states with experimental data, the PS method (see Section 4.2.3) was found to yield high accuracy with an average uncertainty of 0.5 cm^{-1} for all J values.

4.3 Line list

4.3.1 Construction

The initial step in constructing the hybridized line list was matching the Ca variational energy levels with the Ma experimentally derived energy levels. A naive energy-hybridization was extremely problematic, revealing very stark differences in quantum numbers assigned from experiment compared to automatically by DUO with some errors up to 3750 cm^{-1} due to misassignments. The strategies discussed in Section 2 were very important in ensuring accurate hybridization for this molecule. Overall, more than 40 per cent of rovibronic states have incorrect v and Ω assignments; most of these are higher energy states and thus unlikely to be MARVELized, but correcting these assignments is also crucial if energy levels extrapolated using the predicted shift or model Hamiltonians are to be used. After application of the predicted shift correction methodology outlined in Section 2.2.3, the final energy distribution by source type for each electronic state can be seen in Fig. 5(a).

The `.states` files for the ^{46}Ti^{16}O, ^{47}Ti^{16}O, ^{49}Ti^{16}O, and ^{50}Ti^{16}O isotopologues were updated using the isotopologue extrapolation (IE) correction methodology.

The `.trans` files for all isotopologues were cleaned to remove negative transition frequencies. Specifically, 1147, 1155, 1151, 1690, and 1902 transitions were removed for the ^{46}Ti^{16}O, ^{47}Ti^{16}O, ^{48}Ti^{16}O, ^{49}Ti^{16}O, and ^{50}Ti^{16}O isotopologues, respectively.

4.3.2 Suitability for high-resolution cross-correlation techniques

Fig. 5(c) shows the high-res versus total cross-section for ^{48}Ti^{16}O with the new 2024 Toto line list, as well as the partial cross-sections for transitions between only MARVEL (Ma) states using the 2024 and 2021 Toto line lists. The MARVEL coverage of the strong spectral lines for most of the visible and near-IR region is extremely high for TiO, meaning a wide spectral window is possible for HRCC studies, enhancing its sensitivity and thus the ability for the line list to detect lower TiO concentrations. This new 2024 update added experimental data supporting HRCC between 590 and 630 nm, which corresponds to the B $^3\Pi$ – X $^3\Delta$ rovibronic transitions.

Fig. 6(a) shows the transition source type plot (described in Section 2.4) as a useful complementary perspective. The diagram confirms that the most intense transitions (above 10^{-18} cm per molecule) are more than 90 per cent determined very precisely by two MARVEL experimental levels. However, looking at slightly weaker transitions above 10^{-19} cm per molecule, approximately half of all transitions are between two experimental (Ma) energy levels, where the other half have only one Ma energy level. The widths of the red and purple curves demonstrate how important the calculated DUO energy levels are for line list completeness, accurate partition functions, and total opacity calculations. The predicted shift energy levels have a small impact but far less than in MgO.

4.4 Future work

The very high MARVEL coverage for TiO between 450 nm and 1.5 μm means that further improvements to this line list are likely low priority compared to other molecules.

With this new update, there are only very small windows of weaker TiO transitions between 450 nm and 1.5 μm where HRCC for TiO at 2000 K may be unreliable; probably most notable is the 565–580 nm region which would require experimental data for transitions involving B $^3\Pi$, $v = 3$ levels.

Improvements to the Toto spectroscopic model using the new MARVEL B $^3\Pi$, $v = 2$ energy levels would improve the quality of the calculated energies for this state significantly by allowing anharmonicity in this state to be modelled based on experimental data for the first time. This improvement would enhance the accuracy of 565–580 nm spectral region, the modelling of some weak transitions, and the overall partition function.

5 VO: COMBINING MARVEL AND DUO VARIATIONAL CALCULATED ENERGY LEVELS FOR COMPLETE AND ACCURATE LINE LIST

5.1 Overview

Much like TiO, VO is of astrophysical importance in cool stars and hot exoplanets. It has been observed in the atmospheres of a variety of M-type stars, ranging from subdwarfs (Kirkpatrick, Henry &

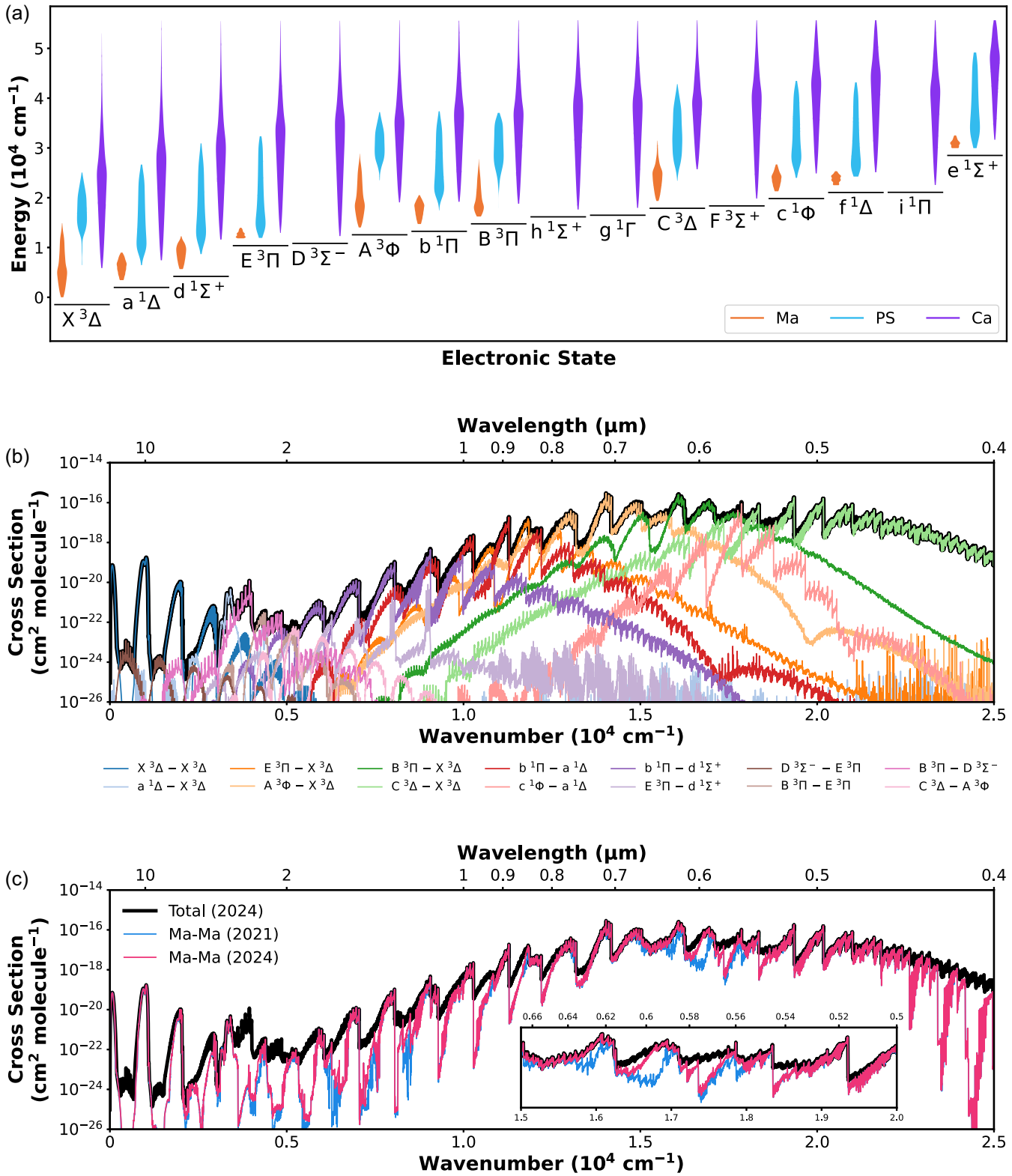

Figure 5. Details of the updated Toto line list for $^{48}Ti^{16}O$. (a) The energy distribution for the rovibronic states of $^{48}Ti^{16}O$ as a function of energy source type in each electronic state. (b) The $^{48}Ti^{16}O$ absorption cross-section computed at 2000 K using the program ExoCross (Yurchenko et al. 2018b) with Gaussian line profiles of 1.0 cm^{-1} half-width half-maximum. The uppermost (black) cross-section shows all transitions in the new 2024 Toto line list with decomposition into dominant electronic bands. (c) The darker (black) cross-section shows all transitions in the new 2024 Toto line list, whereas the lighter (pink) cross-section shows only Marvel (Ma) experimental transitions (with variational intensities). The lower (blue) cross-section shows the Ma – Ma cross-section produced from the 2021 Toto line list (with variational intensities).

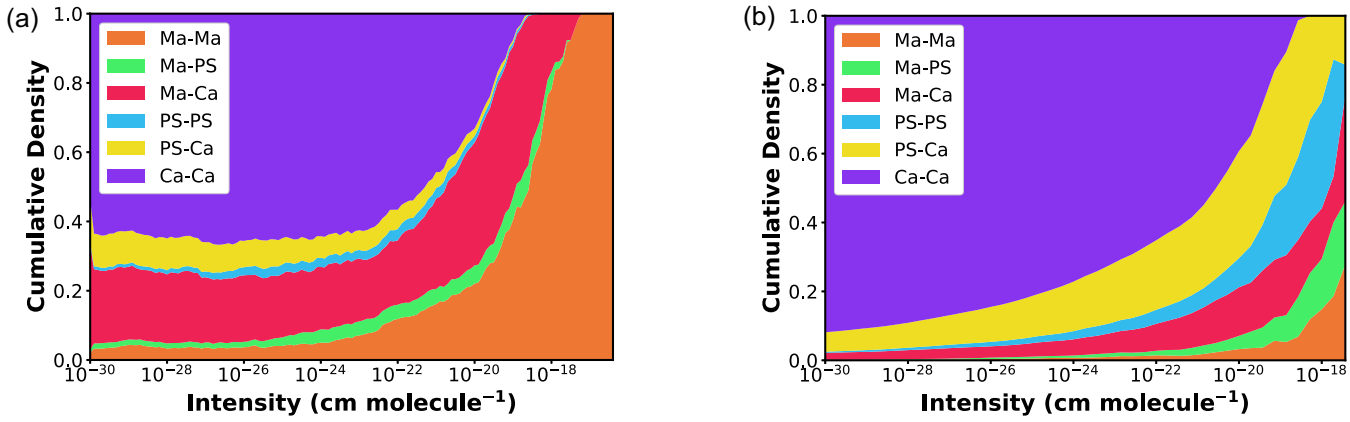

Figure 6. Comparing transition source type plots for $^{48}Ti^{16}O$ from updated Toto and $^{51}V^{16}O$ from HyVO line list. (a) The transition source type plot for $^{48}Ti^{16}O$ using the updated Toto line list, i.e. the cumulative density of transitions as a function of intensity depending on energy source type, computed at 2000 K using the program ExoCross (Yurchenko et al. 2018b). (b) The transition source type plot for $^{51}V^{16}O$ using the HyVO line list, i.e. the cumulative density of transitions as a function of intensity depending on energy source type, computed at 2000 K using the program ExoCross (Yurchenko et al. 2018b).

Liebert 1993) and main sequence dwarfs (Spinrad & Younkin 1966; Wing, Spinrad & Kuhi 1967) to post-main sequence supergiants and hypergiants (Wallerstein 1971; Fawley 1977). In these atmospheres, the strengths of the VO bands are often used to define the spectral classification of the star (Bernath 2009). In exoplanets like hot and ultra-hot Jupiters, VO is believed to be a strong source of opacity due to the absorption bands in the visible and near-infrared. This in turn is believed to drive atmospheric thermal inversions and stratospheric heating (Showman et al. 2009).

The main isotopologue of VO, $^{51}V^{16}O$, is a special case. As a result of the large nuclear spin, $I = \frac{7}{2}$, of the ^{51}V atom and the large associated nuclear quadrupole moment, the rovibronic spectrum of $^{51}V^{16}O$ shows large splittings which cannot be ignored if one wishes to model it for resolving powers greater than roughly $R = 25\,000$. Through comparisons made by Bowesman et al. (2022) between the $^{51}V^{16}O$ MARVEL energies and the calculated DUO energies of McKemmish et al. (2016b), it became apparent that simply updating the existing spectroscopic model (McKemmish et al. 2016a) would be insufficient to produce a high-resolution line list. As is apparent in the unusually large number of hyperfine-resolved transition measurement for $^{51}V^{16}O$, an accurate description of the hyperfine effects in the spectra of $^{51}V^{16}O$ is necessary for a line list intended for use in high-resolution studies (de Regt et al. 2022). As a result, significant developments have been undertaken to allow the treatment of hyperfine-resolved rovibronic spectra of $^{51}V^{16}O$. These includes allowing for a full (i.e. non-perturbative) treatment of hyperfine effects in DUO (Qu, Yurchenko & Tennyson 2022a) plus studies on how to best build models which include the explicit treatment of hyperfine effects (Qu, Yurchenko & Tennyson 2023). Following on from the work of Qu, Yurchenko & Tennyson (2022b) to construct a hyperfine-resolved ground state model for VO, and the subsequent 11 electronic state, hyperfine-unresolved model of Qu et al. (2023), a new 15 electronic state hyperfine-resolved model was produced (Bowesman, Yurchenko & Tennyson 2024a). This model refined all of the potential energy curves and the majority of coupling terms present in previous models and crucially introduced a selection of hyperfine coupling terms.

A new, fully hyperfine-resolved line list for VO called HyVO was produced using this new spectroscopic model (Bowesman et al. 2024b), which is based in part on the updated MARVEL networks described below. The HyVO line list supersedes the widely used VOMYT line list of McKemmish et al. (2016b) which neglected hyperfine splittings and was not MARVELized.

5.2 Data sources for energy levels

5.2.1 MARVEL energy levels

A MARVEL analysis of $^{51}V^{16}O$ was recently conducted by Bowesman et al. (2022), resulting in two distinct spectroscopic networks as half of the spectroscopic data available was hyperfine resolved. Accordingly, separate spectroscopic networks were constructed for the hyperfine-resolved and hyperfine-unresolved data sets. Further, the hyperfine-resolved transitions were deperturbed and added to the hyperfine-unresolved network to supplement its energy level coverage and verify the transitions. While the hyperfine-resolved network only covered the $X\,^4\Sigma^-$, $B\,^4\Pi$, $C\,^4\Sigma^-$, and $1\,^2\Sigma^+$ electronic states, the hyperfine-unresolved network was additionally connected to the $A\,^4\Pi$, $A'\,^4\Phi$, $D\,^4\Delta$, $1\,^2\Delta$, $1\,^2\Phi$, $1\,^2\Pi$, $2\,^2\Pi$, $2\,^2\Delta$, and $3\,^2\Delta$ states. The $1\,^2\Sigma^+$ state is only known via perturbations to the $B\,^4\Pi$ state. The rest of the doublet states were finally connected to the quartet network and hence defined relative to the zero-energy level of the network thanks to the assignment of a spin-forbidden $2\,^2\Pi$–$X\,^4\Sigma^-$ band.

A new study (Döring et al. 2022) presented 1 439 hyperfine-resolved ground state rovibrational transitions in the $X\,^4\Sigma^-$–$X\,^4\Sigma^-$ (1,0) and (2,1) bands. These transitions were added to the existing set of 6 643 hyperfine-resolved transitions and passed through MARVEL where they were all validated. This new data increases the number of hyperfine-resolved MARVEL energy levels from 4 402 to 5 702. The primary increase comes from new levels in the previously missing $X\,^4\Sigma^-$ $v = 1, 2$ bands, contributing 865 and 413 levels, respectively. The extended level coverage also allowed for previously unconnected components of the network to be joined up, adding an additional two levels in the $B\,^4\Pi$ state and six in the $C\,^4\Sigma^-$ state.

The hyperfine-resolved transitions of Döring et al. (2022) were deperturbed into 204 hyperfine-unresolved transitions. The deperturbation procedure has been used during the MARVELization of the AlO line list (Bowesman et al. 2021); the details of the procedure and how the uncertainties in individual hyperfine transitions are propagated are given in Bowesman et al. (2022). The newly deperturbed transitions were added to the existing network of 9140 hyperfine-

unresolved and deperturbed transitions. All of the new transitions were successfully validated by MARVEL. This brings the number of hyperfine-unresolved, empirical energy levels up to 4 813 from 4 712.

5.3 Line list

5.3.1 Construction

The new hyperfine resolved line list for $^{51}V^{16}O$ was presented by Bowesman et al. (2024b) and comprises 3 410 598 energy levels and 58 904 173 243 transitions. Of these levels, 28 760 are MARVELized, 105 566 were updated based on predicted shifts and the remaining 3 276 272 used the term energies calculated by DUO; the distribution of the energy levels as a function of source type can be seen in Fig. 7(a).

5.3.2 Suitability for high-resolution cross-correlation techniques

The high-res versus total cross-section shown in Fig. 7(b) confirms that the MARVEL coverage of HyVO is more extensive than that of $^{24}Mg^{16}O$ shown in Fig. 3(d) but significantly less than the extensive coverage in the updated $^{48}Ti^{16}O$ ToTo line list seen in Fig. 5(c). This conclusion is supported also by the transition source type plot for VO in Fig. 6(b) compared to MgO and TiO. Less than 20 per cent of the strong lines (above 10^{-18} cm per molecule for VO have frequencies predicted by transitions between two MARVELized levels.

The fully MARVELized transitions span a large spectral range and typically cover the strongest part of each spectral band. However, for much of the spectral band, the MARVELized transitions are often an order of magnitude or more less intense than the overall VO cross-section; this might be due to inclusion of only some hyperfine levels.

Like for MgO, the use of the predicted shift (PS) methodology considerably increases the spectral range for which the cross-section of VO can be used for HRCC measurements reliably; this can be seen in Figs 6(b), 7(b), and (c). Transition frequencies obtained by the predicted shift methodology will be of lower accuracy than fully MARVELized transitions. However, for the specific case of VO, many of the predicted shift energy levels were to unmeasured hyperfine levels within the same rovibronic state; the errors associated with this prediction are likely much lower than for other extrapolations as the underlying physics is well understood and predictable. Conversely, due to the J coverage of the MARVEL data, the extrapolation to high J is much larger for VO than TiO, which increases position uncertainty.

Overall, though some caution is warranted when using predicted shift levels (rather than experimentally derived MARVEL levels) for HRCC studies, it is still of sufficiently high quality that detection using HyVO should be possible for sufficiently high molecular concentration of VO. High-resolution cross-correlation studies targeting the most intense bands, such as those in the visible, are likely to produce the most consistent results. For the case of VO with the HyVO line list, we recommend that astronomers focus on the strongest C $^4\Sigma^-$–X $^4\Sigma^-$ bands between 520 and 650 nm or the strongest B $^4\Pi$–X $^4\Sigma^-$ between 730 and 900 nm, avoiding the 650–730 nm region which contains higher vibrational bands for both electronic transitions unconstrained by MARVEL data.

However, since the publication of the HyVO line list, it has been successfully utilized by Maguire et al. (2024) in a HRCC study of the atmosphere of WASP-76b through transit spectra over the wavelength range 370–790 nm. This study obtained an 8σ detection of VO, confirming a previous detection in this ultra-hot Jupiter exoplanet by Pelletier et al. (2023) which utilized the older VOMYT line list, though obtaining stronger signals with the HyVO data.

5.4 Future work

The above analysis shows that the suitability of the HyVO line list for HRCC measurements depends on the accuracy of the predicted shift energy levels. Additional experimental data for VO would be highly desirable to reducing this dependence and enhancing the accuracy of the predicted shift levels by reducing the need for extrapolation. We recommend first focusing on extending the coverage of assigned transitions for the X $^4\Sigma^-$, B $^4\Pi$, and C $^4\Sigma^-$ states at hyperfine resolution to high J.

If the predicted shift energy levels prove to be highly reliable in the initial studies and/or the main spectral bands are well characterized, then new measurements should move to characterizing higher vibrational levels for these main states in order to broaden the spectral window for VO HRCC measurements, e.g. into the 650–730 nm region. These additional measurements of transitions between higher vibrational bands would similarly allow for better constraints on the electronic state potentials and hence more accurate extrapolation to higher vibrational bands, useful for ensuring the accuracy of the partition function and modelling the full cross-section for non-HRCC applications.

6 CONCLUSIONS

While line list intensities will continue to be produced using variational approaches based on energy and intensity spectroscopic models, the future of line list energies is looking to be far more diverse with many data sources combined to take advantage of their strengths. This updated approach is crucial to meet the current needs of the astronomical community in terms of high completeness and high position accuracy of strong spectral lines, particularly with new telescopes (e.g. *JWST*, large ground-based instruments) and new techniques (most notably high-resolution cross-correlation).

Here, we consolidate recent developments in line list construction methodology into a single framework, which we call the hybrid approach. In brief, this approach involves constructing a variational line list based on an energy spectroscopic model (usually ab initio data fitted to experiment) and intensity model (usually purely ab initio) which produces a `.states` file with rovibronic energies and `.trans` file with transition intensities. The calculated energy levels within `.states` file are then updated with other data sources (e.g. MARVEL experiment, model Hamiltonians, HITRAN, MOLLIST) or corrections (e.g. predicted shifts, isotopologue-extrapolation) to produce the best currently known unified source of rovibronic energy levels for that molecule. We select a single terminology and associated acronym for each data source or correction that we apply uniformly across all ExoMol-hosted diatomic line list to this new format to ensure consistency and clear communication. This approach is extensible; new data sources and corrections to energy levels can readily be added. Note that the ExoMol format and this hybrid approach ensure that transition positions arise solely from the energy levels and cannot be updated independently.

Adding uncertainties for each data source is crucial and indeed is the way that the ‘best’ data source is selected for each energy level (this process can sometimes be implicit but making it explicit is helpful). However, this is always, of course, an estimate and

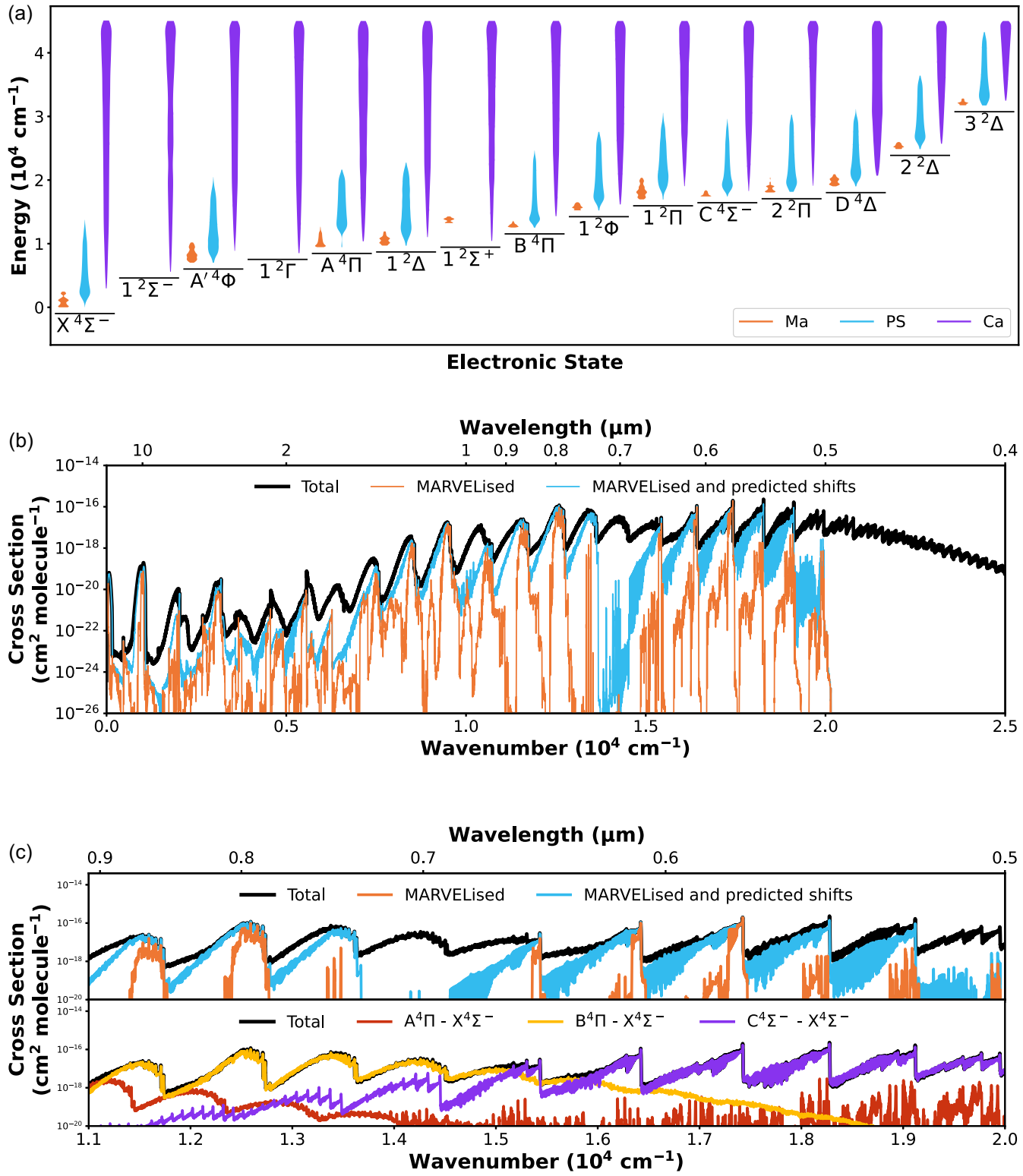

Figure 7. Details of HyVO line list for $^{51}V^{16}O$. (a) Energy distribution for the rovibronic states of $^{51}V^{16}O$ as a function of energy source type in each electronic state. (b) $^{51}V^{16}O$ absorption cross-sections computed at 2000 K using the program ExoCross (Yurchenko et al. 2018b) with Gaussian line profiles of 1.0 cm^{-1} half-width half-maximum. The upper (black) cross-section shows all transitions in the line list, whereas the lower (orange) cross-section shows only MARVEL (Ma) experimental transitions (with variational intensities), and the intermediate (blue) cross-section shows all possible transitions between MARVEL (Ma) and predicted shift (PS) energy levels (with variational intensities). (c) Closeup of the visible cross-section of $^{51}V^{16}O$ over the regions recommended for analysis with HRCC techniques, using the HyVO line list. The upper panel shows the contributions from MARVEL levels, as well as levels with energies corrected with predicted shifts. The lower panel shows which electronic bands contribute to the total cross-section in this region. As with TiO in Fig. 5(b), the total cross-section here is comprised of many different electronic bands, with the $A\,^4\Pi$–$X\,^4\Sigma-$, $B\,^4\Pi$–$X\,^4\Sigma-$, and $C\,^4\Sigma-$–$X\,^4\Sigma-$ bands dominating from roughly 6000 cm^{-1} onwards.

there are unexpected challenges that can arise. The most notable example here is when uncertainties of energy levels (e.g. MARVEL's global uncertainties) substantially overestimate the uncertainty of the observable transition position; manual judgement is currently required here.

We demonstrate this new approach using MgO, TiO, and VO as exemplar molecules through analysis of the aforementioned energy source types with careful treatment of uncertainties in all cases. Of particular importance is the suitability of these updated line lists in HRCC studies, which take advantage of the very high spectral resolving power of ground-based telescopes to make molecular detections despite poor signal-to-noise in observations.

For MgO, we have updated the existing variational line list with experimental MARVEL energy levels, and the predicted shifts and isotopologue-extrapolation corrections. While the previous line list was completely unsuitable for HRCC, this new line list can be used with reasonable confidence for HRCC applications in the strong transition around 500 nm and in the broad spectral region 580–680 nm. We note that additional experimental data at higher J and v in measured electronic states would be useful to verify the quality of the predicted shift levels as there is significant extrapolation from experimentally measured data for this molecule.

For TiO, we have updated the previous MARVEL compilation through invalidation of old experimental data and inclusion of new experimental data for the X $^3\Delta$ rovibrational, B $^3\Pi$ – X $^3\Delta$ and E $^3\Pi$ – X $^3\Delta$ bands. This updated MARVEL data set was then used to rehybridize the 2021 Toto line list for each isotopologue with novel predicted shift calculations. In this case, the high quantity of experimental data meant that predicted shift was only needed to essentially fill gaps in observations, giving it very high reliability. However, to hybridize this line list, a very careful treatment of quantum numbers was required to align the DUO-computed quantum numbers to those derived from energy ordering arguments; intrinsically this reflects the approximate nature of quantum numbers for heavily coupled states but practically this can be a significant challenge when hybridizing energy states files. Furthermore, for the B $^3\Pi$ electronic state, comparison of the original variational spectroscopic model (which was fitted on only $v = 0, 1$) with the new experimental data (including $v = 0, 1, 2$) did reveal that refitting of the B $^3\Pi$ electronic state in the spectroscopic model will be desirable in the future. Regardless, the current 2024 Toto line list has extremely high coverage of experimental data for TiO strong lines across almost all of the visible and near-IR region accessible for ground-based HRCC, with only minor improvements available near 570 nm. Compared to 2021 Toto, this updated data set provides better experimental coverage in the visible region around 600 nm (17000 cm^{-1}).

Importantly, for TiO, the very high quality of the available line list in terms of both completeness and accuracy of strong lines (from experiment) means there should be no concerns about spectral line list quality when analysing *JWST* observations or conducting HRCC studies; if TiO is present with absorption above the noise limit of the observation, it should be detected with current data in most visible spectral regions (Fig. 5c).

For VO, we outline a new MARVEL compilation and hyperfine-resolved variational line list. New figures are presented that enable additional insight into the suitability of this line list for HRCC applications. We find that careful selection of spectral region is much more important for VO than TiO as only the very strongest bands around in the near IR and those around 520–650 nm are suitable for reliable HRCC studies. The spectral lines with frequencies obtained by predicted shift are also very important for providing a sufficiently large spectral window for effective HRCC measurements. Additional experimental data considering vibronic transitions involving the B $^4\Pi$ state with $v \geq 2$ would be the most beneficial for enhancing the HRCC potential of the VO line list, opening up the spectral region 650–730 nm for analysis. Further, experimental assignments of higher J populations for some VO spectral bands would reduce the need to use predicted shifts or validate the accuracy of this approach more closely.

We note that experimental efforts are beginning to be undertaken which are explicitly designed to improve the quality of the MARVEL networks for certain key molecules, thus far for water (Tobias et al. 2020) and H_2CO (Germann et al. 2024). This activity, and in particular extending it to rovibronic spectra of diatomics such as those considered here, would be beneficial for the overall quality of the available line lists. Finally, the whole ExoMol data base is vast, in the region of 10^{13} individual lines (Tennyson et al. 2024b). A new data base, ExoMolHR, has been built (Zhang et al., inpreparation) which extracts energy levels determined with low uncertainty ($\Delta\tilde{E} \leq 0.01$ cm^{-1}) and provides data on lines predicted to high resolution ($R \geq 100\,000$).

ACKNOWLEDGEMENTS

This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI Australia), an NCRIS enabled capability supported by the Australian Government. The work performed at UCL and the ExoMol project is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme through Advance Grant 883830 (ExoMolHD).

The authors declare no conflicts of interest.

DATA AVAILABILITY

The most accurate line list data for the diatomic molecules discussed here will always be available on the ExoMol website (https://exomol.com/).

The MARVEL energy and transitions files for $^{24}Mg^{16}O$, $^{48}Ti^{16}O$, and $^{51}V^{16}O$ are included as supporting information. The DUO spectroscopic models for these molecules are unchanged from previously published data and are available on the ExoMol website and the original publications.

REFERENCES

Allard F., Hauschildt P. H., Schwenke D., 2000, ApJ, 540, 1005
Arendas, P., Furtenbacher, T., Császár. A. G., 2020, Sci. Rep., 10, 19489
Azuma Y., Dyke T. R., Gerke G. K., Steimle T. C., 1984, J. Mol. Spectrosc., 108, 137
Bernath P. F., 2009, Int. Rev. Phys. Chem., 28, 681
Bernath P., Cameron W. D., 2020, ApJ, 904, 24
Bielska K. et al., 2022, Phys. Rev. Lett., 129, 043002
Birkby J.-L., 2018, preprint (arXiv:1806.04617)
Bowesman C. A., Shuai M., Yurchenko S. N., Tennyson J., 2021, MNRAS, 508, 3181
Bowesman C. A., Akbari H., Hopkins S., Yurchenko S. N., Tennyson J., 2022, J. Quant. Spectrosc. Radiat. Transfer, 289, 108295
Bowesman C. A., Yurchenko S. N., Tennyson J., 2024a, Mol. Phys., 122, e2255299
Bowesman C. A., Qu Q., McKemmish L. K., Yurchenko S. N., Tennyson J., 2024b, MNRAS, 529, 1321
Brady R. P., Yurchenko S. N., Tennyson J., Kim G.-S., 2024, MNRAS, 527, 6675

Breier A. A., Waßmuth B., Fuchs G. W., Gauss J., Giesen T. F., 2019, J. Mol. Spectrosc., 355, 46
Cameron W. D., Bernath P., 2022, ApJ, 926, 39
Civis S., Hedderich H. G., Blom C. E., 1991, Chem. Phys. Lett., 176, 489
Correira J., Aikin A. C., Grebowsky J. M., Pesnell W. D., Burrows J. P., 2008, Geophys. Res. Lett., 35, L06103
Császár A. G., Furtenbacher T., 2011, J. Mol. Spectrosc., 266, 99
Darby-Lewis D., Tennyson J., Lawson K. D., Yurchenko S. N., Stamp M. F., Shaw A., Brezinsek S., JET Contributor, 2018, J. Phys. B: At. Mol. Opt. Phys., 51, 185701
Darby-Lewis D. et al., 2019, J. Mol. Spectrosc., 362, 69
de Regt S., Kesseli A. Y., Snellen I. A. G., Merritt S. R., Chubb K. L., 2022, A&A, 661, A109
Döring E., Fuchs G. W., Giesen T. A. B. A., Blum L., 2022, in International Symposium on Molecular Spectroscopy. University of Illinois, Urbana-Champaign, IL, p. FG11
Edwards B., Changeat Q., 2024, ApJ, 962, L30
Edwards B. et al., 2020, ApJ, 160, 8
Falco A., Tremblin P., Charnoz S., Ridgway R. J., Lagage P.-O., 2024, A&A, 683, A194
Fawley W. M., 1977, ApJ, 218, 181
Furtenbacher T., Császár A. G., Tennyson J., 2007, J. Mol. Spectrosc., 245, 115
Furtenbacher T., Szabó I., Császár A. G., Bernath P. F., Yurchenko S. N., Tennyson J., 2016, ApJS, 224, 44
Furtenbacher T., Horváth M., Koller D., Sólyom P., Balogh A., Balogh I., Császár A. G., 2019, J. Phys. Chem. Ref. Data, 48, 023101
Furtenbacher T., Hegedus S. T., Tennyson J., Császár A. G., 2022, Phys. Chem. Chem. Phys., 24, 19287
Germann M. et al., 2024, J. Quant. Spectrosc. Radiat. Transfer, 312, 108782
Guo Z. et al., 2024, MNRAS, 528, 1769
Hoeijmakers H. J., de Kok R. J., Snellen I. A. G., Brogi M., Birkby J. L., Schwarz H., 2015, A&A, 575, A20
Ip P. C., Cross K. J., Field R. W., Rostas J., Bourguignon B., McCombie J., 1991, J. Mol. Spectrosc., 146, 409
Kagi E., Kawaguchi K., 2006, J. Mol. Spectrosc., 795, 179
Kagi E., Hirano T., Takano S., Kawaguchi K., 1994, J. Mol. Spectrosc., 168, 109
Kirkpatrick J. D., Henry T. J., Liebert J., 1993, ApJ, 406, 701
Kobayashi K., Hall G. E., Muckerman J. T., Sears T. J., Merer A. J., 2002, J. Mol. Spectrosc., 212, 133
Le Roy R. J., 2017, J. Quant. Spectrosc. Radiat. Transfer, 186, 167
Li H.-Y., Tennyson J., Yurchenko S. N., 2019, MNRAS, 486, 2351
Maclean S., Duley W. W., 1982, ApJ, 252, L25
Maguire C., Gibson N. P., Nugroho S. K., Fortune M., Ramkumar S., Gandhi S., de Mooij E., 2024, A&A, 687, A49
McKemmish L. K., 2021, WIREs Comput. Mol. Sci., 11, e1520
McKemmish L. K., Yurchenko S. N., Tennyson J., 2016a, Mol. Phys., 114, 3232
McKemmish L. K., Yurchenko S. N., Tennyson J., 2016b, MNRAS, 463, 771
McKemmish L. K. et al., 2017, ApJS, 228, 15
McKemmish L. K. et al., 2018, ApJ, 867, 33
McKemmish L. K., Masseron T., Hoeijmakers J., Pérez-Mesa V. V., Grimm S. L., Yurchenko S. N., Tennyson J., 2019, MNRAS, 488, 2836
McKemmish L. K., Syme A. M., Borsovszky J., Yurchenko S. N., Tennyson J., Furtenbacher T., Csaszar A. G., 2020, MNRAS, 497, 1081
Merritt S. R. et al., 2020, A&A, 636, A117
Mogren Al Mogren M., Ajili Y., Almania S., Ben Abdallah D., Hochlaf M., 2015, MNRAS, 452, 1561
Mollière P., Snellen I. A. G., 2019, A&A, 622, A139
Mürtz P., Richter S., Pfelzer C., Thümmel H., Urban W., 1994, Mol. Phys., 82, 989
Mürtz P., Thümmel H., Pfelzer C., Urban W., 1995, Mol. Phys., 86, 513
Nozawa T., Kozasa T., Umeda H., Maeda K., Nomoto K., 2003, ApJ, 598, 785
Nugroho S. K., Kawahara H., Masuda K., Hirano T., Kotani T., Tajitsu A., 2017, ApJ, 154, 221
Ouyang Q. et al., 2023, MNRAS, 521, 5860
Pan S. et al., 2023, Nat. Commun., 14, 1165
Pavlenko Y. V., Yurchenko S. N., McKemmish L. K., Tennyson J., 2020, A&A, 42, A77
Pearce O., Yurchenko S. N., Tennyson J., 2024, MNRAS, 527, 10726
Pelletier S. et al., 2023, Nature, 619, 491
Perri A. N., McKemmish L. K., 2024, MNRAS, 531, 3023
Perri A. N., Taher F., McKemmish L. K., 2023, MNRAS, 524, 4631
Phillips . G., 1973, ApJS, 26, 313
Piette A. A. A., Madhusudhan N., McKemmish L. K., Gandhi S., Masseron T., Welbanks L., 2020, MNRAS, 496, 3870
Polyansky O. L., Bielska K., Ghysels M., Lodi L., Zobov N. F., Hodges J. T., Tennyson J., 2015, Phys. Rev. Lett., 114, 243001
Polyansky O. L., Kyuberis A. A., Lodi L., Tennyson J., Ovsyannikov R. I., Zobov N., 2017, MNRAS, 466, 1363
Prinoth B. et al., 2022, Nat. Astron., 6, 449
Qu Q., Yurchenko S. N., Tennyson J., 2021, MNRAS, 504, 5768
Qu Q., Yurchenko S. N., Tennyson J., 2022a, J. Chem. Theory Comput., 18, 1808
Qu Q., Yurchenko S. N., Tennyson J., 2022b, J. Chem. Phys., 157, 124305
Qu Q., Yurchenko S. N., Tennyson J., 2023, J. Mol. Spectrosc., 391, 111733
Rietmeijer F. J. M., 2009, ApJ, 705, 791
Sakamoto S., White G. J., Kawaguchi K., Ohishi M., Usuda K. S., Hasegawa T., 1998, MNRAS, 301, 872
Sarantos M., Killen R. M., McClintock W. E., Bradley E. T., Vervack Jr. R. J., Benna M., Slavin J. A., 2011, Planet Space Sci., 59, 1992
Schaefer L., Lodders K., Fegley B., 2012, ApJ, 755, 41
Sedaghati E. et al., 2017, Nature, 549, 238
Semenov M., Tennyson J., Yurchenko S. N., 2022, Phys. Chem. Chem. Phys., 516, 1158
Serindag D. B., Nugroho S. K., Mollière P., de Mooij E. J. W., Gibson N. P., Snellen I. A. G., 2021, A&A, 645, A90
Showman A. P., Fortney J. J., Lian Y., Marley M. S., Freedman R. S., Knutson H. A., Charbonneau D., 2009, ApJ, 699, 564
Simard B., Hackett P. A., 1991, J. Mol. Spectrosc., 148, 128
Spinrad H., Younkin R. L., 1966, PASP, 78, 65
Steimle T. C., Azuma Y., Carrick P. G., 1984, ApJ, 277, L21
Stockstill-Cahill K. R., McCoy T. J., Nittler L. R., Weider S. Z., Hauck II S. A., 2012, J. Geophys. Res.: Planets, 117, E00L15
Syme A.-M., McKemmish L. K., 2020, MNRAS, 499, 25
Tennyson J., 2012, WIREs Comput. Mol. Sci., 2, 698
Tennyson J., Yurchenko S. N., 2012, MNRAS, 425, 21
Tennyson J., Yurchenko S. N., 2017, Mol. Astrophys., 8, 1
Tennyson J., Yurchenko S. N., 2022, Front. Astron. Space Sci., 8, 795040
Tennyson J. et al., 2014, Pure Appl. Chem., 86, 1931
Tennyson J., Lodi L., McKemmish L. K., Yurchenko S. N., 2016, J. Phys. B: At. Mol. Opt. Phys., 49, 102001
Tennyson J. et al., 2020, J. Quant. Spectrosc. Radiat. Transfer, 255, 107228
Tennyson J., Furtenbacher T., Yurchenko S. N., Császár A. G., 2024a, J. Quant. Spectrosc. Radiat. Transfer, 316, 108902
Tennyson J. et al., 2024b, J. Quant. Spectrosc. Radiat. Transfer, 326, 109083
Tobias R. et al., 2020, Nat. Commun., 11, 1708
Törring T., Hoeft J., 1986, Chem. Phys. Lett., 126, 477
Turner B. E., Steimle T. C., 1985, ApJ, 299, 956
Wallerstein G., 1971, ApJ, 169, 195
Wing R. F., Spinrad H., Kuhi L. V., 1967, ApJ, 147, 117
Witsch D., Breier A. A., Döring E., Yamada K. M. T., Giesen T. F., Fuchs G. W., 2021, J. Mol. Spectrosc., 377, 111439
Yoneda S., Grossman L., 1995, Geochim. Cosmochim. Acta, 59, 3413
Yurchenko S. N., Tennyson J., Bailey J., Hollis M. D. J., Tinetti G., 2014, Proc. Nat. Acad. Sci., 111, 9379
Yurchenko S. N., Lodi L., Tennyson J., Stolyarov A. V., 2016a, Comput. Phys. Commun., 202, 262

Yurchenko S. N., Blissett A., Asari U., Vasilios M., Hill C., Tennyson J., 2016b, MNRAS, 456, 4524
Yurchenko S. N., Williams H., Leyland P. C., Lodi L., Tennyson J., 2018a, MNRAS, 479, 1401
Yurchenko S. N., Al-Refaie A. F., Tennyson J., 2018b, A&A, 614, A131
Yurchenko S. N., Mellor T. M., Freedman R. S., Tennyson J., 2020, MNRAS, 496, 5282
Yurchenko S. N. et al., 2022, MNRAS, 510, 903
Yurchenko S. N., Brady R. P., Tennyson J., Smirnov A. N., Vasilyev O. A., Solomonik V. G., 2024a, MNRAS, 527, 4899
Yurchenko S. N., Szajna W., Hakalla R., Semenov M., Sokolov A., Tennyson J., Pavlenko Y., Schmidt M. R., 2024b, MNRAS, 527, 9736
Zilinskas M., van Buchem C. P. A., Miguel Y., Louca A., Lupu R., Zieba S., van Westrenen W., 2022, A&A, 661, A126

SUPPORTING INFORMATION

Supplementary data are available at *RASTAI* online.

The MARVEL energy, transitions, and segment files for $^{24}Mg^{16}O$, $^{48}Ti^{16}O$, and $^{51}V^{16}O$ are included as supporting information.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.