Amoeba: Automated Molecular Excitation Bayesian Line-fitting Algorithm

Anita Hafner1,2, J. R. Dawson1,2,3, and Mark Wardle1,2

1 Department of Physics and Astronomy, Macquarie University, NSW 2109, Australia; anita.hafner@csiro.au
2 Research Centre in Astronomy, Astrophysics and Astrophotonics, Macquarie University, NSW 2109, Australia
3 Australia Telescope National Facility, CSIRO Space & Astronomy, PO Box 76, Epping, NSW 1710, Australia

Received 2021 July 25; revised 2021 September 9; accepted 2021 October 11; published 2021 December 28

Abstract

The hyperfine transitions of the ground-rotational state of the hydroxyl radical (OH) have emerged as a versatile tracer of the diffuse molecular interstellar medium. We present a novel automated Gaussian decomposition algorithm designed specifically for the analysis of the paired on-source and off-source optical depth and emission spectra of these OH transitions. In contrast to existing automated Gaussian decomposition algorithms, AMOEBA (Automated Molecular Excitation Bayesian line-fitting Algorithm) employs a Bayesian approach to model selection, fitting all four optical-depth and four emission spectra simultaneously. AMOEBA assumes that a given spectral feature can be described by a single centroid velocity and full width at half maximum, with peak values in the individual optical-depth and emission spectra then described uniquely by the column density in each of the four levels of the ground-rotational state, thus naturally including the real physical constraints on these parameters. Additionally, the Bayesian approach includes informed priors on individual parameters that the user can modify to suit different data sets. Here we describe AMOEBA and establish its validity and reliability in identifying and fitting synthetic spectra with known (but hidden) parameters, finding that the code performs very well in a series of practical tests. AMOEBA’s core algorithm could be adapted to the analysis of other species with multiple transitions interconnecting shared levels (e.g., the 700 MHz lines of the first excited rotational state of CH). Users are encouraged to adapt and modify AMOEBA to suit their own use cases.

Unified Astronomy Thesaurus concepts: Interstellar medium (847); Bayesian statistics (1900); Hydroxyl sources (772); Radio astronomy (1338)

1. Introduction

The gas and dust between the stars—the Galactic interstellar medium (ISM)—plays an important role in the dynamic evolution of our Galaxy and others. The molecular gas of the ISM (the vast majority of which is H2) is key to several important stages of that evolution. Unfortunately, molecular hydrogen does not have readily accessible energy levels at the low temperatures and densities seen in the majority of the molecular ISM, rendering it effectively invisible. Consequently, other molecules with accessible states in this low-energy environment that are expected to coexist with H2 are instead observed as tracers of the molecular gas. The hydroxyl radical (OH) is one such tracer.

The ground-rotational state of OH is split into 4 levels by Δ-doubling and hyperfine splitting, with 4 allowed transitions between those levels at 1612.231, 1665.402, 1667.359, and 1720.530 MHz. These levels and transitions are illustrated in Figure 1. OH was the first molecule discovered in the ISM (Weinreb et al. 1963) and observations of the hyperfine transitions within its ground-rotational state have been used extensively to trace the distribution, properties, and dynamics of ISM gas (e.g., Robinson & McGee 1967; Goss 1968; Heiles 1969; Turner 1979; Liszt & Lucas 1996; Ebisawa et al. 2015; Li et al. 2018; Rugel et al. 2018; Busch et al. 2019; Ebisawa et al. 2019; Petzler et al. 2020). Crucially, OH has been demonstrated both theoretically (Wolfire et al. 2010) and through observations (Barriault et al. 2010) to correlate with the molecular hydrogen of the ISM across a wide range of number densities and column densities (see also Nguyen et al. 2018 and references therein).

This link has led to an increase in the use of OH as a supplementary tracer of molecular hydrogen (e.g., Dawson et al. 2014; Li et al. 2018; Nguyen et al. 2018; Rugel et al. 2018; Engelke & Allen 2019), particularly in diffuse regions (Wannier et al. 1993; Liszt & Lucas 1996; Barriault et al. 2010; Allen et al. 2012, 2015). This is in contrast to observations of the low rotational transitions of CO, which have been shown to underestimate the mass of molecular gas in diffuse regions (e.g., Reach et al. 1994; Grenier et al. 2005; Planck Collaboration et al. 2011; Remy et al. 2018).

A key advantage of observing OH is the sensitivity of its ground-rotational state level populations to local gas conditions such as kinetic temperature, number density, and radiation field (Elitzur 1976; Elitzur et al. 1976; Gübert et al. 1978; Petzler et al. 2020). In particular, the so-called “satellite lines” at 1612 and 1720 MHz often display anomalous excitation, implying significant departures from local thermodynamic equilibrium (LTE, e.g., Elitzur 1976; Elitzur et al. 1976; Dawson et al. 2014; Petzler et al. 2020). These non-thermal excitation patterns are very sensitive to the local environment. Therefore, by measuring the distribution of OH molecules across the four levels of the ground-rotational state, non-LTE molecular excitation modeling can then be used to constrain these local conditions (Ebisawa et al. 2015, 2019; Petzler et al. 2020) which in turn can inform a wide range of theories of ISM phenomena. Though this distribution can, in principle, be determined directly from observable quantities for a single ISM cloud along a given line of sight, most observations will capture several Doppler-shifted and potentially blended spectral features representing several distinct ISM clouds. This blending necessitates an initial step where these blended features are decomposed. However, this presents several potential complications of relevance to this work. Chief among these is the question of the number and velocity position of individual components. As outlined in Section 2.2 and discussed...
in detail in Section 3, our approach calculates Bayes factors of competing Gaussian decomposition models to find the “best” decomposition (given the a priori and likelihood distributions). Another key question concerns the treatment of the four transitions during this process of Gaussian decomposition: namely, whether the spectra should be decomposed independently or if they should be linked in some way. As we discuss in Section 3, we have chosen a novel approach to decompose all four spectra of the ground-rotational state transitions simultaneously. We take advantage of the intrinsic relationships between the parameters of the Gaussian-shaped features in each of the four transitions’ spectra (i.e., centroid velocity, full width at half maximum, and peak values in each transition) and the underlying excitation properties of the OH molecules (i.e., the column densities in each level of the ground-rotational state). This approach allows us to take advantage of the available data: by using each transition as additional information about the same gas, we are able to focus only on models of the OH excitation that are completely self-consistent to the limit of our underlying assumptions (discussed in depth in Section 3).

In this paper we introduce AMOeba: an open-source Automated Molecular Excitation Bayesian line-fitting Algorithm for the simultaneous decomposition of all four OH ground-rotational spectra. AMOeba is intended to be used for large OH data sets such as the H1/OH/recombination line survey of the Milky Way (THOR; Beuther et al. 2016), or the upcoming Galactic Australian Square Kilometre Array Pathfinder Survey (GASKAP Dickey et al. 2013). These large-scale surveys would take full advantage of AMOeba’s automated nature in that it takes spectra as an input and outputs fully decomposed spectra fit with scientifically useful parameters that can then be used with molecular excitation modeling to determine or constrain the local conditions of the gas. AMOeba’s core algorithm could also be modified to facilitate analysis of other species.

In Section 2 we discuss the OH ground-rotational state and the observable quantities related to the distribution of molecules across the levels of that state. Section 2 also discusses the place of this work in the context of other Gaussian decomposition algorithms. In Section 3 we introduce and describe AMOeba, and in Section 4 we describe its performance in a series of tests on synthetic and real data. In Section 5 we conclude.

2. Background

2.1. Observable Quantities

If gas containing OH molecules lies between the observer and a compact background source of continuum with brightness $T_c$ and the spatially extended background continuum (due to Galactic synchrotron and the cosmic microwave background) $T_{bg}$, the observed continuum-subtracted line brightness temperature is described by:

$$T_b = (T_{ex} - T_c - T_{bg})(1 - e^{-\tau}),$$

(1)

where the excitation temperature $T_{ex}$ is a reparameterization of the population of the upper and lower levels of the given transition within the cloud, defined by the Boltzmann factor:

$$\frac{N_u}{N_l} = \frac{g_u}{g_l} \exp \left[ -\frac{h\nu_0}{k_b T_{ex}} \right],$$

(2)

where $N_u$ and $N_l$ are the column densities in the upper and lower levels, respectively, $g_u$ and $g_l$ are the degeneracies of the upper and lower levels, respectively, and $\nu_0$ is the rest frequency of the transition. The excitation temperatures of the four ground-rotational transitions are not independent, and are related via the excitation temperature sum rule:

$$\frac{\nu_{1612}}{T_{ex}(1612)} + \frac{\nu_{1720}}{T_{ex}(1720)} = \frac{\nu_{1665}}{T_{ex}(1665)} + \frac{\nu_{1667}}{T_{ex}(1667)},$$

(3)

which is derived from the definition of excitation temperature (Equation (2)) and the energy difference between the four levels and will therefore hold in all environments. The optical depth $\tau_\nu$ of an isothermal homogeneous cloud is given by:

$$\tau_\nu = \frac{c^2}{8\pi\nu_0^2 g_l N_l A_{1ul}} \left( 1 - \exp \left[ -\frac{h\nu_0}{k_b T_{ex}} \right] \right) \phi(\nu),$$

(4)

where $\phi(\nu)$ is the line profile. In environments where $|T_{ex}| > h\nu_0/k_B = 0.08$, the bracketed term in Equation (4) is well-represented by $h\nu_0/k_B T_{ex}$, and the peak optical depths in the four ground-rotational transitions are then related via the optical-depth sum rule (Robinson & McGee 1967):

$$\tau_{peak}(1612) + \tau_{peak}(1720) = \frac{\tau_{peak}(1665)}{5} + \frac{\tau_{peak}(1667)}{9}.$$

(5)

The shape of the line profile $\phi(\nu)$ will be determined by the dominant broadening mechanisms within the cloud, which in the molecular ISM is assumed to be turbulence. We further assume that this turbulence is on a small scale which would result in a Gaussian-shaped profile via the central limit theorem. Thus far we have assumed that a given gas cloud can be accurately described by a single $T_{ex}$ value in each OH transition despite the fact that ISM clouds are known to have clumpy and turbulent structure on all observable scales (e.g., Hennebelle et al. 2008; Benincasa et al. 2013). This assumption implies that these $T_{ex}$ and $N_{OH}$ values—which would in truth only describe a completely homogeneous cloud—actually represent a quasi-“averaging” of this ensemble of smaller structures. Therefore when we ascribe a single set of $T_{ex}$ and $N_{OH}$ values to a Gaussian feature—and by extension, to an ISM cloud along the line of sight—we acknowledge that this is a simplification.

$T_{ex}$ and $\tau_\nu$ are uniquely described by the column densities in each level of the ground-rotational state $N_1 - N_4$, which in turn are determined by the properties and local conditions of the cloud. There are, however, two key complications in characterizing $N_1 - N_4$. First, Equation (1) implies that $T_{ex}$ and $\tau_\nu$ cannot both be uniquely characterized from a single observation. Second, it is very common for multiple clouds
to lie along a single line of sight, so the observed $T_e$ at a given spectral channel may represent flux from more than one cloud, which must be disentangled before $T_{ex}$ and $\tau_v$ can be determined. The second complication is readily resolved by decomposing the observed spectra into Gaussian components, then relating the parameters of those Gaussians to the target quantities $T_{ex}$ and $\tau_v$, and therefore $N_1$–$N_4$.

The first complication can be solved through choice of observing strategy, two examples of which are relevant to this work. First, a single-dish telescope with a sufficiently narrow beam (e.g., the Arecibo Radio Telescope) can observe both directly “on” a background continuum source $T_{ex}^on = (T_e - T_c - T_{bg})(1 - e^{-\tau_v})$ and several “off” continuum points surrounding the “on” position $T_{ex}^{off} = (T_e - T_{bg})(1 - e^{-\tau_v})$ (e.g., Heiles & Troland 2003; Li et al. 2018; Nguyen et al. 2018, 2019). This approach assumes that $T_{ex}$, $T_{bg}$, and $\tau_v$ are the same in the “on” and “off” positions and also that none of the compact background continuum is present in the “off” positions. Such observations can then be converted to optical-depth spectra:

$$
\frac{T_{ex}^{on} - T_{ex}^{off}}{T_c} = e^{-\tau_v} - 1.
$$

In this case, the “off-source” observations then represent the brightness temperature that would be expected if the compact background continuum source could be switched off. However, in the case where the compact background continuum source is resolved and overlaps with the “off” source positions (e.g., observations toward radio-bright HII regions within the Galaxy), the “on” source brightness temperature is described by $T_{ex}^{on} = (T_e - T_c - T_{bg})(1 - e^{-\tau_v})$ and the “off” source brightness temperature is $T_{ex}^{off} = (T_e - T_{off} - T_{bg})(1 - e^{-\tau_v})$. The optical-depth spectra are then obtained from:

$$
\frac{T_{ex}^{on} - T_{ex}^{off}}{T_c} = e^{-\tau_v} - 1.
$$

In practice the value of $T_{ex}^{off}$ would likely differ between “off” source positions, requiring care when generating optical-depth spectra.

In both cases, following the convention of Heiles & Troland (2003) we can define an “expected brightness temperature” $T_{exp}$ as the spectrum expected if the compact background continuum source could be switched off:

$$
T_{exp} = (T_e - T_{bg})(1 - e^{-\tau_v}),
$$

from which $T_{ex}$ and therefore $N_1$–$N_4$ can be determined uniquely. AMOEBA has been designed to take as its primary input a set of these optical depth ($\tau_v$) and expected brightness temperature versus velocity spectra. Such data sets are referred to as “on–off spectra” in this paper.

The degeneracy between $\tau_v$ and $T_{ex}$ can also be broken by observing a bright background continuum source with an interferometer, in which case the $T_{ex}$ and $T_{bg}$ terms in Equation (1) become insignificant. This assumes that either the OH cloud (and hence the emission from it) is larger on the sky than the interferometer fringes and is therefore not detected due to spatial filtering, or that $T_c \gg T_{ex}$. (Where this latter condition may also be fulfilled for some single-dish observations with a sufficiently small beam and sufficiently bright background source.) This then allows the measured $T_{ex}$ spectra to be directly converted to optical-depth spectra. Though this method does not allow the excitation temperatures of the four transitions and therefore the populations of the levels of the ground-rotational state to be determined directly, it does allow for useful constraints to be placed on these populations (Petzler et al. 2020). AMOEBA has therefore also been designed to fit a set of optical-depth spectra only.

2.2. Gaussian Decomposition

The prime difficulty in Gaussian decomposition is the question of the number and location of Gaussian components present in a spectrum. While the statistics community is broadly in agreement about the best approach to this problem—which is similar to the approach we take in this work, the astrophysics community has somewhat fallen behind. In astrophysics, the human eye and informed personal judgment are often relied upon to determine the number and position of Gaussian components, then a least-squares minimization technique is employed to optimize the parameters of those Gaussians (e.g., Heiles & Troland 2003; Li et al. 2018; Rugel et al. 2018, etc.). Several approaches exist in the astrophysics literature to automate the process of identifying the number and position of Gaussian features, mostly in the context of the notoriously blended spectra of H1. One such approach is that described by Haud (2000), where first a single Gaussian is placed and optimized, and then another, etc. until a threshold residual is reached. That number of components is then reduced by removing features deemed to be unreasonable or merging close features.

Other methods take a more analytical approach (as opposed to the simple sequential approach of Haud 2000) to determine the number and positions of features, such as GAUSSPY (Lindner et al. 2015) which identifies the locations of potential Gaussian features using derivative spectroscopy informed by a trained machine-learning model. Several more recent methods also insist on spatial coherence of Gaussian components, such as GAUSSPY+ (Riener et al. 2019), ROHSA (Marchal et al. 2019) and SCOUSEPY (Henshaw et al. 2019). Though some of these methods employ elements of Bayesian statistics (i.e., SCOUSEPY uses the Akaike Information Criterion to resolve discrepancies between spatially neighboring decompositions) most ultimately rely on a frequentist assessment of the goodness of a particular fit. Therefore, these methods are not able to fully incorporate prior information or directly compare the posterior probabilities of competing decomposition models: the main advantages of Bayesian algorithms.

An example of a fully Bayesian approach to Gaussian decomposition is described in Allison et al. (2012). The aim of this method is to identify extragalactic H1 absorption in radio data cubes such as those obtained by the Australian Square Kilometre Array Pathfinder (ASKAP). Their method first identifies bright continuum sources in the two on-sky dimensions, then it searches the frequency (equivalently, the velocity or redshift) domain for absorption. They utilize MULTINEST8 (Feroz & Hobson 2008; Feroz et al. 2009), a Bayesian Monte Carlo multi-nested sampling algorithm. Though nested sampling methods in general are very powerful and robust (see Buchner 2021, for a comprehensive review), they are computationally expensive. The best way to ameliorate this shortcoming is to have very restrictive prior distributions, i.e., to restrict the initial parameter space as much as possible. A primary goal for

---

8 https://github.com/farhanferoz/MultiNest
AMOEBA was for it to be completely automated, so the fine-tuning of the prior distributions required to make a nested sampling approach feasible was deemed prohibitive, and AMOEBA instead applies a Markov Chain Monte Carlo algorithm.

2.3. A Bayesian Approach to the Gaussian Decomposition of OH Spectra

AMOEBA employs a fully Bayesian approach to Gaussian decomposition, allowing the incorporation of prior information and a direct analytical comparison of the posterior probabilities of competing decomposition models. These are related via Bayes’ Theorem:

\[ P(\theta|d, M) = \frac{P(d|\theta, M) P(\theta|M)}{P(d|M)}. \tag{9} \]

Bayes’ Theorem describes the process by which the prior knowledge of the distribution of probability across the possible parameters \( \theta \) of a given model \( M \)—defined by the a priori probability distribution \( P(\theta|M) \)—is modified by the presence of the data \( d \), the likelihood of which given the possible parameters of the model is described by the distribution \( P(d|\theta, M) \). This modification of the prior knowledge into the posterior knowledge is via the “evidence” \( P(d|M) \) which represents the probability of the data \( d \) given the model \( M \) (Kass & Raftery 1995). The ratio of the evidence of two competing models is then equivalent to the odds in favor of one model over the other (Kass & Raftery 1995).

The key novelty of AMOEBA is that it simultaneously fits all four ground-rotational transitions of OH, taking advantage of the dependence of directly observable quantities (\( \tau \) and \( T_{\text{exp}} \)) on the underlying distribution of OH molecules across the levels of the ground-rotational state, in contrast to typical methods of identifying features in OH spectra.

Some past approaches to the analysis of OH spectra are summarized as follows:

1. Gaussian features are identified and fit in the main-line transitions only with independent centroid velocities \( v \) and full widths at half maximum \( \Delta v \) (e.g., Engelke & Allen 2018; Li et al. 2018; Nguyen et al. 2018; Rugel et al. 2018).
2. Gaussian features are identified and fit in all four ground-rotational transitions with independent \( v \) and \( \Delta v \), but with peak values of \( \tau \) or \( T_{\text{b}} \) linked using a statistical equilibrium code such as RADEX (van der Tak et al. 2007) or their own molecular excitation code (e.g., Ebisawa et al. 2015; Xu et al. 2016).
3. Gaussian features are identified and fit in all four ground-rotational transitions with the same \( v \) and \( \Delta v \), and the molecular excitation code is used to predict the peak values of \( \tau \) or \( T_{\text{b}} \) based on local cloud properties such as kinetic temperature, number density, and local radiation field (Ebisawa et al. 2019).

AMOEBA takes advantage of the reasonable assumption that if an OH-containing cloud lies along the line of sight toward a compact continuum source, in the absence of obscuring instrument noise one would expect to observe a Gaussian-shaped feature at the same centroid velocity and with the same FWHM in all four ground-rotational state transitions. It also takes advantage of the fact that the peak values of those features \( (T_{\text{exp}}, \tau_{\text{peak}}) \) are not independent, but ultimately depend on just 5 parameters: the FWHM of the line profile, and the column density in each of the four levels of the ground-rotational state \( N_1 - N_4 \). Therefore, while a single Gaussian feature in a set of on–off spectra may be described by up to 24 parameters (if all 8 spectra are fit independently), AMOEBA will describe the same feature with 6 independent parameters.

3. Amoeba

AMOEBA takes as its primary input a set of on–off spectra at each of the OH ground-rotational state transition frequencies. The spectra should not be smoothed or binned beyond their native frequency resolution. The posterior probability distribution of successively complex models (i.e., with increasing number of Gaussian components) is then sampled until a point is reached where an additional feature does not sufficiently improve the evidence. A model \( M_N \) with \( N \) components with the \( i \)th component having parameters \( \theta_i = [v_i, \log_{10} \Delta v_i, \log_{10} N_{i1}, T_{\text{ex}}^{-1}(1612), T_{\text{ex}}^{-1}(1665), T_{\text{ex}}^{-1}(1677)] \), where \( v \) is the centroid velocity of the Gaussian feature, \( \Delta v \) is its FWHM and \( N_1 \) is the column density in the lowest hyperfine level. \( T_{\text{ex}}^{-1}(1720) \) is then not a free parameter, but is determined from Equation (3). These log and inverse parameters (i.e., \( \log_{10} \Delta v \), \( \log_{10} N_1, T_{\text{ex}}^{-1} \)) were chosen rather than \( \Delta v, N_1, N_2, N_3, N_4 \) as their relationship to observable quantities is more linear. From these quantities, the observable quantities \( \tau_{\text{peak}} \) and \( T_{\text{exp}} \) are computed from Equations (2), (4), and (8).

Alternatively, AMOEBA can take as input a set of OH optical-depth versus velocity spectra only (hereafter referred to as “optical-depth spectra”). In this case the \( i \)th component of the model will have parameters \( \theta_i = [v_i, \log_{10} \Delta v_i, \tau_{\text{peak}}^{-1}(1612), \tau_{\text{peak}}^{-1}(1665), \tau_{\text{peak}}^{-1}(1677), \tau_{\text{peak}}^{-1}(1720)] \) where \( \tau_{\text{peak}}^{-1} \) are the Gaussian heights at 1612, 1665, 1667, and 1720 MHz. Table 1 outlines AMOEBA’s user-controlled parameters.

The process of Gaussian decomposition is outlined in Figure 2. The first step is to divide the spectra into velocity segments to be processed sequentially. This step serves to reduce computation time and can either be done automatically (via the process explained below) or with velocity ranges provided by the user. Any velocity range extending over a user-defined lower limit of channels (default value is three channels) for which the data in all transitions fall below a user-defined “detection limit” is assumed to contain no signal and the spectra are cut at the center of the velocity range where the detection limit is a multiple of the rms noise level of each of the provided spectra. This process is illustrated in Figure 3. In this example the detection limit is set at \( 1.4 \times \text{rms noise} \). Any velocities for which three consecutive channels in any of the spectra are above the detection limit are flagged as having significant signal. These are indicated by the red points along the velocity axis in the figure. Any time two ranges of velocities with significant signal are separated by the user-defined lower limit of channels without significant signal, the spectra are all cut at the center of that range. In the example shown in Figure 3 the cut point is indicated by the vertical blue line. The resulting velocity segments are assumed to be independent (i.e., they are not expected to include any significant signal from features centered in another velocity segment) and are therefore processed consecutively. The detection limit can be tuned to be quite conservative to avoid the unlikely scenario that a small amount of flux from a Gaussian component may be missed due to the process of cutting the spectra. This process only cuts the spectra into segments; it does not discard any of the data. In the case where
Table 1
The Input and User-controlled Parameters of AMOEBA

| Parameter          | Description                                                                 | Default values |
|--------------------|-----------------------------------------------------------------------------|----------------|
| source_name        | Sightline designation used for reporting results.                           | None           |
| vel_axes           | List containing four velocity axes that correspond to the optical-depth and expected brightness temperature spectra. | None           |
| tau_spectra        | List containing four optical-depth spectra ($\tau_\nu$ vs. v).                | None           |
| Texp_spectra       | List containing four expected brightness temperature spectra ($T_{\text{exp}}$ vs. v). | Optional       |
| tau_rms            | List of four rms noise levels corresponding to the spectra provided in tau_spectra. | Optional       |
| Texp_rms           | List of four expected brightness temperature spectra ($T_{\text{exp}}$ vs. v). | Optional       |
| Thg                | List of four background brightness temperature ($T_b$) values. Only required when fitting on-off spectra. | None           |
| sigma_tolerance    | Detection limit (as a multiple of the spectral rms noise) used for separating input spectra into velocity segments. | 1.4            |
| num_chan           | Minimum number of consecutive channels that must be below sigma_tolerance in all provided spectra in order for AMOEBA to cut the spectra at the center of those velocity channels (see text for details). | 3              |
| sig_vels           | Rather than allowing AMOEBA to divide the spectra into velocity segments the user may provide velocity ranges in the form $\{\nu_{\text{start}}, \nu_{\text{end}}\}, \ldots$. | Optional       |
| Bayes_threshold    | Threshold factor of increase in evidence required to accept a new model.     | 10             |
| extra_gaussians    | The number of extra Gaussian features to attempt if a model does not improve the evidence by the Bayes_threshold. | 3              |
| fwhm_mean          | Mean of the log$_{10}\Delta\nu$ prior probability distribution in km s$^{-1}$ (Gaussian in shape). | 0.05           |
| fwhm_sig           | Standard deviation of the log$_{10}\Delta\nu$ prior probability distribution in km s$^{-1}$. | 0.2            |
| IN1_mean           | Mean of the log$_{10}N_1$ prior probability distribution in cm$^{-2}$ (Gaussian in shape). | 12.8           |
| IN1_sig            | Standard deviation of the log$_{10}N_1$ prior probability distribution in cm$^{-2}$. | 0.7            |
| tau_mean           | Mean(s) of the $\tau_{\text{peak}}$ prior probability distribution. AMOEBA allows for an arbitrarily complex-shaped prior defined by a list of Gaussians, but it is the same shape for all four transitions. | [0.0075, 0.0005] |
| tau_sig            | Standard deviations(s) of the $\tau_{\text{peak}}$ prior probability distribution. | [0.012, 0.0015] |
| tau_weight         | Weight(s) of the Gaussian components of the $\tau_{\text{peak}}$ prior probability distribution. | [0.7, 0.3]     |
| sumrule_sig        | Standard deviation of the expected residual of the optical-depth sum rule (Equation (5)). | Optional       |
| invTex_mean        | Mean(s) of the $T_{\text{exp}}^{-1}$ prior probability distribution in K$^{-1}$, same as the format of tau_mean. | [0.35]         |
| invTex_sig         | Standard deviation of the $T_{\text{exp}}^{-1}$ prior probability distribution in K$^{-1}$. | [0.4]          |
| invTex_weight      | Weight(s) of the $T_{\text{exp}}^{-1}$ prior probability distribution. | [1]            |

Note. The default values given in square brackets imply that the input should be in the form of an array.

For a given velocity segment, the evidence of the null model is calculated, which is simply the likelihood of a Gaussian probability distribution of the model containing a single Gaussian feature. The number of data points is given by $n$, the data points are denoted by $d_i$, and the corresponding model points are denoted by $m_i$. The likelihood distribution is calculated using the relation $P(d|\theta, M) = \prod_i P(d_i|\theta, M)$ which follows from the fact that both sides of Equation (9) integrate to one over all parameter space. Therefore, the following two subsections outline how the likelihood distribution $P(d|\theta, M)$ and the a priori distribution $P(\theta|\mathcal{M})$ are quantified.

3.1. The Likelihood Distribution

As outlined in Section 1, the joint probability distribution is composed of the likelihood $P(d|\theta, M)$ and the prior $P(\theta|\mathcal{M})$. The likelihood is defined as

$$P(d|\theta, M) = \frac{1}{\sigma_{\text{noise}}^n \sqrt{(2\pi)^n}} \exp \left[ -\frac{\sum_i (d_i - m_i)^2}{2\sigma_{\text{noise}}^2} \right],$$

where the data noise is assumed to be Gaussian, uncorrelated, and defined by a standard deviation $\sigma_{\text{noise}}$, which is assumed to be constant across a given spectrum but may vary between transitions. The number of data points is given by $n$, and the data points are denoted by $d_i$ with corresponding model points $m_i$. AMOEBA can obtain the value of $\sigma_{\text{noise}}$ from the input spectra by measuring the standard deviation of several regularly spaced, overlapping segments of each spectrum, then taking the median of these standard deviations. This approach assumes that the majority of these segments are signal-free.
3.2. The a priori Distributions

AMOEBA assumes a uniform a priori probability distribution over the centroid velocity parameter. Additionally, it requires that all centroid velocities fall within the given velocity segment and in the case of models with more than one Gaussian component, that the centroid velocities are in ascending order. The provision that the centroid velocities be in ascending order is to avoid the consideration of degenerate models that only differ in the ordering of components. Thus the centroid velocity prior distribution for a model $\mathcal{M}_N$ with $N$ Gaussian components takes the form:

$$P(\theta | \mathcal{M}_N) = \frac{N!}{\prod_{i=1}^{N} a_i} \prod_{i=1}^{N} \frac{a_i}{v_{\text{max}} - v_{\text{min}}} ,$$

where $a_i = \begin{cases} 1 & \text{if } v_{\text{min}} \leq v_i \leq v_{\text{max}} \text{ and } v_i > v_{i-1} \\ 0 & \text{otherwise} \end{cases}$.

The default a priori distributions for $\log_{10} \Delta \nu$ and $\log_{10} N_1$ were loosely based on the values found from the Millennium Survey by Li et al. 2018 (hereafter L18), which used the Arecibo telescope to observe the 1665 and 1667 MHz OH transitions and included sightlines both in and off the Galactic Plane. This survey was chosen for its large sample size, quality of data, and derived parameters. The priors are Gaussian in shape with means $\mu = -0.5 \text{ km s}^{-1}$ and $\mu = 12.5 \text{ cm}^{-2}$, respectively, and standard deviations $\sigma = 0.2 \text{ km s}^{-1}$ and $\sigma = 0.75 \text{ cm}^{-2}$, respectively. These distributions, along with values from L18 are shown in Figure 4. The default a priori distribution of $\log_{10} N_1$ was chosen to be more conservative than the distribution suggested by the results of L18 and is therefore broader than their results.

The a priori distributions for the remaining parameters of the Gaussian models have defaults informed primarily by the expected distributions given a “reasonable” range of cloud characteristics (outlined in Table 2) as determined from non-LTE molecular excitation modeling. This was performed using the models of Petzler et al. (2020) by choosing random values of each parameter from the distributions noted in Table 2, for a total of $10^6$ models. The default model-based values were cross-checked with the observed distributions from L18.

The default a priori distributions for both $T_{\text{ex}}^{-1}$ and $\tau_{\text{peak}}$ are guided by the distributions suggested by our non-LTE modeling and the results of L18, which are shown in Figure 5. The a priori distribution for $\tau_{\text{peak}}$ also includes a weak penalty for models that violate the optical-depth sum rule (Equation (5)).

Given that AMOEBA is intended as an automated algorithm for large data sets, we chose default a priori distributions somewhat simpler than the distributions suggested by our modeling and the results of L18, primarily to avoid unacceptable biases from overtuned priors. This aim led us to choose a Gaussian-shaped prior for $T_{\text{ex}}^{-1}$ with mean $\mu = 0.35 \text{ K}^{-1}$ and standard deviation $\sigma = 0.4 \text{ K}^{-1}$. The distribution for $\tau_{\text{peak}}$, seen in our modeling and L18 had a narrower peak and wider, less symmetric wings than that seen in $T_{\text{ex}}^{-1}$, so we chose a double-Gaussian prior for $\tau_{\text{peak}}$ with means $\mu = 0.01, 0.0025$; standard deviations $\sigma = 0.01, 0.0015$; and relative weights of 0.6 and 0.4, respectively. These default a priori distributions are shown in Figure 5.

Though care was taken when choosing these default priors, we urge users of AMOEBA to assess whether they are appropriate for their intended use case. For instance, for data

---

6 L18 reported $N_{\text{OH}}$ which is the total OH column density. We converted these to $N_1$ by multiplying their reported value by a factor of 3/16, which assumes a roughly equal population distribution across the sublevels of each of the ground-rotational state levels, which is true in nearly all cases.
Figure 3. Illustration of the process by which AMOEBA cuts spectra into independent velocity ranges using a set of synthetic data. The user-defined “detection limit” (1.4 × the rms noise in this example) is indicated by the dark red lines in each plot. The red points indicate all velocities for which at least three consecutive channels in any of the spectra exceeded the detection limit. The velocity cut point is indicated by the vertical blue line.

Figure 4. The prior probability distributions for $D_{\nu \log_{10}}$ (left) and $N_{1 \log_{10}}$ (right) plotted over histograms of their values of $N_{1 \log_{10}}$ as determined from 1665 MHz and from 1667 MHz as reported by Li et al. 2018 (L18). The prior for $D_{\nu \log_{10}}$ has a default mean of $\mu = -0.5$ km s$^{-1}$ and a default standard deviation of $\sigma = 0.2$ km s$^{-1}$, while the prior for $N_{1 \log_{10}}$ has a default mean of $\mu = 12.5$ cm$^{-2}$ and a default standard deviation of $\sigma = 0.75$ cm$^{-2}$. 
toward the Central Molecular Zone or even some parts of the Inner Galaxy, one might expect much larger FWHM or higher optical depths. In that case, the default priors will introduce significant bias and these parameters will tend to be underestimated. We illustrate this in Section 4.2 by showing that for synthetic spectra generated with parameters that have a very low prior probability, the fitted parameters will show a small but systematic shift (clearly seen for optical depth in Figure 8) toward the values preferred by the priors. The user can avoid this by choosing a more appropriate or wider prior distribution for these parameters.

3.3. Speed

The speed of AMOEBA will of course vary based on the complexity of features in the spectra. As a test case, the synthetic spectra shown in Figure 3 were decomposed on a typical laptop computer in less than one hour. However, the most expensive spectra tested during the development of this code took nearly 20 hr on the OzStar supercomputer (Swinburne University, Melbourne) to identify 14 Gaussian features, 7 of which were in the same velocity segment. We therefore suggest that the user first identify individual velocity segments in their spectra, either manually or by using the “findranges” function within AMOEBA, then utilize the “sig_vels” keyword (outlined in Table 1) to parallelize each velocity range on a separate core. We utilized this approach for a set of 83 on–off observations (the subject of an upcoming paper) using the OzStar supercomputer and the entire set was fully decomposed in approximately 10 hr.

4. Performance and Discussion

AMOEBA’s performance was assessed in a series of tests that examined the validity of its core Bayesian algorithm (see Section 4.1), its ability to recover “correct” parameters (see Section 4.2), its sensitivity to noise (see Section 4.3), and its ability to resolve closely blended features (see Section 4.4). These tests were performed on synthetic data with a range of parameters outlined in Table 3. In all but the last test, the synthetic data were generated with a single Gaussian feature, and in the last test, two identical Gaussian features were present. Noise was added to these spectra and they were then decomposed.

In the tests discussed in Sections 4.2 to 4.4 we use slightly different criteria to assess the fits, which we summarize here for clarity. Our aim in Section 4.2 was to illustrate the range of parameters returned by AMOEBA as a function of the “actual” feature parameters when only a single feature was present in

![Figure 5](image-url)

Figure 5. The prior probability distribution for $T_{ex}^{-1}$ (left) and $\tau_{peak}$ (right) plotted over histograms of their values at 1665 MHz and at 1667 MHz as reported by Li et al. 2018 (L18) and those found from non-LTE molecular excitation modeling as described in the text. The average 1σ $\tau_{peak}$ detection limit of L18 is indicated by the gray shaded area. The edges of the histograms were chosen so that this range is a single bin centered at 0. The dashed histograms represent the distribution of $T_{ex}^{-1}$ parameters estimated. We illustrate this in Section 4.2 by showing that for signifient parameters.

| Parameter | Values | Distribution |
|-----------|--------|--------------|
| $T_{ex}$(K) | 10–200 | uniform |
| $T_{gas}$(int)(K) | 10–50 | uniform |
| $T_{gas}$(ext)(K) | 10–100 | uniform |
| $\log_{10} N_{H}$ (cm$^{-2}$) | $\mu = 13.2, \sigma = 1$ | normal |
| $\log_{10} \Delta v$ (km s$^{-1}$) | $\mu = -0.18, \sigma = 0.23$ | normal |
| $A_{v}$(mag) | 0.1, 0.3, 1 | choice |
| ($\log_{10} n_{H}$ (cm$^{-3}$), $X_{v}$) | (2, –4), (3, –5), (4, –7) | choice |

![Table 2](image-url)

Table 2: Parameters of “Typical” ISM Clouds Input into Non-LTE Molecular Excitation Modeling (Petzler et al. 2020) Used to Inform AMOEBA’s Peak Optical Depth and Inverse Excitation Temperature a priori Probability Distributions.
All values were chosen from a uniform distribution in the stated range. Therefore, we only considered results where AMOEBA identified a single feature, and categorized results where AMOEBA identified zero features as a false negative, and results where AMOEBA identified two or more features as a false positive. This differs from the definition of false positive we employed in Section 4.3, where we also include results where AMOEBA returned a single feature but where none of the 99.95% credibility intervals of the returned parameters contained the correct parameters. We chose not to apply this stricter definition in Section 4.2 so that we could show the full range of parameters returned by AMOEBA: the definition used in Section 4.3 would have artificially improved the results reported in Section 4.2. In Section 4.4 we have the further complication that the actual spectra contain two features. In this case (following the pattern in Sections 4.2 and 4.3), we characterize any results where less than two features were identified as false negatives and any where more than two were identified as false positives. However, we also had to account for cases where two features were identified in the results, but where one corresponds to an actual feature and the other does not. We define a feature as having been not fit if none of the returned features’ 99.95% credibility intervals contained the actual parameters. In this case, where AMOEBA identifies two features but where one or both of the actual features was not fit, we categorize the result as a false positive. We made this choice (i.e., to not categorize such results as false negatives) because we felt that returning false positives was the more problematic behavior. Further, these tests were conducted under conditions where false negatives were expected (i.e., at low signal-to-noise or small feature separation), so we wanted to distinguish this different behavior.

All tests were performed with both on-off spectra and optical-depth spectra only. AMOEBA performed well in these tests, as is discussed further in this section.

### 4.1. Validity of Bayesian Algorithm

The validity of AMOEBA’s Bayesian inference algorithm was tested using “simulation-based calibration” (SBC) following the method described by Talts et al. (2018). Though Bayesian analysis is itself straightforward in its application of Bayes Theorem (Equation (9)), algorithms built to handle complex models and data sets can easily fail to reach satisfactory conclusions if the algorithm lacks certain self-consistencies. For example, the choice of a given parameter of a model—say the FWHM \( \Delta v \) of a Gaussian-shaped feature—may introduce unintended biases to the algorithm, whereas a different choice—i.e., its logarithm \( \log_{10} \Delta v \)—may eliminate this bias. SBC aims to identify and diagnose these issues through a straightforward test, the results of which are described in Section 4.1.

In this test, a set of parameters representing a single Gaussian-shaped profile feature:

\[
\theta = [v, \log_{10} \Delta v, \log_{10} N L_\text{ex}^{-1}(1612), T_\text{ex}^{-1}(1665), T_\text{ex}^{-1}(1667)]
\]

for on-off spectra, or

\[
\tilde{\theta} = [v, \log_{10} \Delta v, \tau_{\text{peak}}(1612), \tau_{\text{peak}}(1665), \tau_{\text{peak}}(1667), \tau_{\text{peak}}(1720)]
\]

for optical depth only, were drawn from the prior distribution \( P(\theta|M) \) and a sample data set was constructed by drawing from the likelihood distribution \( P(d|\theta, M) \). AMOEBA was then used to sample the posterior probability distribution, returning the converged Markov chains. If AMOEBA’s Bayesian inference algorithm is valid, the originally drawn set of parameters \( \theta \) will have an equal probability of falling at any percentile rank within the converged Markov chains (see Talts et al. 2018, for a detailed proof). Therefore, if this process of selecting a set of parameters from the prior distribution, simulating data by drawing from the likelihood distribution, and then sampling the posterior probability distribution is repeated many times, the distribution of the percentile ranks of those drawn parameters within the converged Markov chains will show a flat distribution. The distribution of these percentile ranks for 2000 trials is shown in Figure 6.

Figure 6 shows ranks from 5 to 95, as the edge bins will tend to be overpopulated due to auto-correlation in the sample set. This auto-correlation is found in all samples drawn using MCMC and indicates limitations in the simulation-based calibration test rather than the Bayesian algorithm (Talts et al. 2018). The shaded regions of Figure 6 indicate the area that should contain 99.7% of the data points shown if their distribution was truly uniform: only 1 point should lie outside the shaded region. For the method using on-off spectra, 4 points fell outside this shaded region, while for optical-depth spectra, 2 points fell outside. Overall 99% of the points fall within the required range, implying no significant biases in our Bayesian algorithm.

#### 4.2. Ability to Recover “Correct” Parameters

The purpose of this test was to quantify the ability of AMOEBA to recover true parameters from data by fitting synthetic data with known parameters. It also served to test the sensitivity of the parameter recovery rate to the a priori distributions. In this test, synthetic spectra were generated with a range of parameters (see Table 3) with noise added such that each spectrum had a signal-to-noise ratio of 5. This signal-to-noise ratio is intended to represent a reasonably strong signal given the weakness of most diffuse OH features (e.g., Li et al. 2018; Rugel et al. 2018). AMOEBA was then used to recover the parameters. From 5000 trials none produced a false-negative result, and only 7 of the on-off spectra returned a false-positive result. Here we define a false negative as a result where the null model was preferred (i.e., where AMOEBA did not identify any features in the spectrum), and a false positive as a result where AMOEBA identified 2 or more features in the spectra.

Figures 7 and 8 then show the difference between the final fitted parameters returned by AMOEBA to the original parameters for the remaining spectra. The shaded regions in these figures show the range covered by the inner 20%, 50%,
and 90% of tests within each bin of 100 samples; the outer line shows the full 100% of samples.

Testing over an FWHM range of 0.3–5.0 km s\(^{-1}\), we find that 50% of the synthetic on–off spectra and optical-depth spectra of the centroid velocity was within 1 channel (0.1 km s\(^{-1}\)) of the correct location, and in both cases was within 2 channels of the correct location for 90% of the spectra. We note, however, that the 68% credibility intervals reported by AMOEBA for the centroid velocity parameter tends to not reflect this spread: the typical 68% credibility interval covers \(\approx 0.01\) km s\(^{-1}\). This is a reflection of our treatment of the data as an incomplete sampling of a continuous Gaussian distribution rather than considering the individual channels as bins of that continuous function. The overall effect of this simplification is that the credibility interval for centroid velocity (and FWHM, discussed further below) will underestimate the true uncertainty. We therefore suggest users of AMOEBA adopt a more conservative estimate of the uncertainty of these parameters: \(\approx \pm 1\) channel width is likely a more accurate reflection of the effective uncertainty at this signal-to-noise level.

For values of \(\sim 1\) km s\(^{-1}\) (typical of the L18 sample and hence well matched to our chosen a priori distribution), the FWHM recovered was within one channel width of the correct value for 90% of both the synthetic on–off spectra and the optical-depth spectra. For larger line widths, the accuracy of the recovered value is reduced due to the influence of the a priori distribution. We therefore advise if wide features are expected (i.e., \(\Delta v > 4\) km s\(^{-1}\)) that the a priori distribution be adjusted accordingly.

Testing over a range of \(-20\) to \(+20\) K, we find that the recovered excitation temperatures were within \(\approx 10\%\) of the actual values down to an accuracy of \(\pm 1\) K for 90% of on–off spectra. We note that Figure 7 implies a reduction in accuracy for \(T_{\text{ex}}(1720)\) when \(|T_{\text{ex}}(1720)| > 5\) K. This is a reflection of the fact that \(T_{\text{ex}}^{-1}(1720)\) is not a free parameter in our model and therefore the values of \(T_{\text{ex}}(1720)\) did not have a uniform distribution in this test.

Testing over a \(\log N_1\) range of 11 to 15 (cm\(^{-2}\)), the recovered column density was within \(\approx 10\%\) of the correct value for 90% of the synthetic on–off spectra. This is a very encouraging result when compared to the uncertainties found by L18 which

Figure 6. Histograms of the ranks obtained through simulation-based calibration of AMOEBA for synthesized data with both on- and off-source observations (top) and for synthesized data with only optical-depth spectra (bottom). The peaks of each histogram bin are plotted as points on a line for visual clarity. Each line represents an individual model parameter as described in the text. A well-calibrated Bayesian algorithm is expected to have a uniform distribution across all ranks. The shaded region indicates the area expected to contain 99.7% of the data points if the distribution was uniform.
Figure 7. Difference between the final parameters recovered by AMOEBA and the “actual” parameters for “on–off” spectra synthesized with a signal-to-noise ratio of 5 and a channel width of 0.1 km s$^{-1}$. The shaded regions represent the inner 90%, 50%, and 20% range of all trials, with each bin on the x-axis containing 100 trials. The region enclosing all trial results (the 100% range) is indicated by the pale blue line. For the tests shown, there were no false positives or false negatives.
tended toward 100%, or Nguyen et al. (2018) which tended toward 40%.

In optical depth, we test a range of $-0.05$ to $+0.05$, and in this range, 60% of recovered values were within $\pm 10\%$ of the actual values. However, this range is considerably wider than the expected distribution as quantified by our a priori distribution (see Figure 5). We can therefore see in Figure 8 that for those more extreme values (i.e., $|\tau_{\text{peak}}| > 0.02$) the fitted results are systematically shifted to lower values. This highlights the importance of modifying the priors if higher magnitude optical depths are expected in the data. If we look at the range that is more consistent with that of the priors (i.e., $-0.02 < \tau_{\text{peak}} < 0.04$), over 90% of recovered values were within $\pm 10\%$ of the actual values.

Finally, for all of these test cases, we also computed the sum of the squared residuals of the fitted model and the synthetic spectra ($\sum_i (d_i - m_i)^2$) and compared this to the sum of the squared residuals of the actual model and the synthetic spectra ($\sum_i (d_i - a_i)^2$, where $a$ is the model constructed from the actual parameters $\theta$). This resulted in two sets of distributions (i.e., the fitted and actual models for optical depth and expected brightness temperature spectra in each of the four transitions) which we then compared using the Kolmogorov–Smirnov (KS) test. If the fitted models are a good representation of their corresponding synthetic spectra these distributions should be indistinguishable, in which case the KS statistic will be small and the $p$-test will approach 1. In each case the distribution of summed squared residuals for the fitted model were indistinguishable from those of the actual model. The KS statistics and two-sided $p$-values (expressed as $1 - p$-test as the $p$-test values all approached or were equal to 1) are given in Table 4.

Figure 8. Same as Figure 7 but for synthetic optical-depth spectra. For the tests shown there was a false-positive rate of 0.04% and there were no false negatives.
4.3. Effect of Signal-to-noise Ratio

The purpose of this test was to quantify the signal-to-noise ratio over which AMOEBA could be relied upon to recover low-signal features. In this test, spectra with varying noise levels and feature parameters were generated and AMOEBA was then used to recover those parameters. The aim of this test was to determine the relationship between signal-to-noise ratio and AMOEBA’s ability to produce “good-quality” fits to the data.

In order to describe the quality of the fits returned by AMOEBA, we define four broad categories that we use in this and the next test: well fit, poorly fit, false negative, and false positive. The fits were placed in these categories by the following method. Any fits where fewer than the correct number of features were identified were immediately categorized as false negatives. Similarly, any fits where more than the correct number of features were identified were immediately categorized as false positives. Then all fits with the correct number of features were checked to see if all the true parameters fell within the 99.95% credibility interval as approximated by three times the 68% credibility interval. We found that in general the posterior probability distribution was Gaussian, implying that three times the 68% credibility interval is a good approximation of the 99.95% credibility interval. (Typical corner plots are shown in Figures 17 and 18 in the Appendix).

If all parameters of a fit were within the 99.95% credibility interval, it was categorized as well fit. If at least one but not all were in the 99.95% credibility interval it was categorized as poorly fit, otherwise it was categorized as a false positive. False positives identified in this way would also technically be false negatives, but we felt that false positives were the more concerning behavior and therefore categorized them as such.

This seemingly wide range in the credibility interval was chosen upon examination of a selection of the fits which showed that using the same steps with a narrower credibility interval threshold would categorize a significant number of visually good fits as poor or false positives. Even with our broad credibility interval threshold most fits categorized as “poor” are not obviously “wrong” to the eye (see an example of a “poor” fit in Figure 9). This behavior was mostly a reflection of the narrow credibility intervals often returned for some parameters, notably the FWHM. Indeed, as can be seen from the results of the previous test outlined in Section 4.2, although there is a spread in the returned parameter values compared to the actual values, there are few if any extreme outliers at a signal-to-noise level of 5 (see Figures 7 and 8), i.e., even “poor” fits in the present test are unlikely to correspond to pathological deviations in the true physical parameters.

The results of this test for on–off synthetic spectra as well as optical-depth-only synthetic spectra are illustrated in Figure 10. As expected, at signal-to-noise ratios below 1 the majority of spectra resulted in a false-negative result: the noise was too high to justify the acceptance of a feature. In both cases, this false-negative rate dropped quickly with increased signal-to-noise ratio and the number of “well-fit” spectra reached 90% at a signal-to-noise ratio of 2 for the synthetic on–off spectra and 3 for the synthetic optical-depth spectra. Both had a similar rate of “poorly fit” results. The ability of AMOEBA to produce reliable fits to the majority of on–off spectra even at the 2σ level is a reflection of the power of fitting all four transitions simultaneously and represents a significant improvement over traditional decoupled fitting.

4.4. Effect of Feature Separation

The purpose of this test was to quantify how far apart in velocity two identical features must be for AMOEBA to be able to reliably distinguish them. In this test, spectra with two identical Gaussian features and a signal-to-noise ratio of 5 were generated and AMOEBA was used to fit the spectra. The quality of the fits were judged to be “well” or “poorly fit” using the same criteria as the previous test. A fit was categorized as a “false negative” if one or no features were identified. A fit that identified three or more features, or a fit with two features where one feature’s 99.95% credibility intervals did not contain any of the correct values was categorized as a “false positive.” This category was further subdivided into those where both, one, or neither of the two features were “well fit.” The results of this test are illustrated in Figure 11.

As expected, at separations less than the FWHM, the majority of spectra returned a false-negative result: the two features could not be distinguished. In both methods, this false-negative rate dropped off quickly and ≈90% of spectra were well fit at separations more than 2 times the FWHM for both the synthetic on–off spectra and optical-depth spectra. However, this false-negative rate is still high given that 2 identical features separated by the FWHM should be resolvable. In exploring the origin of this behavior, we identified that the features returning a false negative at these higher separations (>1.5 × Δν) were skewed toward features with small FWHMs.

In Figure 12 we plot the FWHM versus separation as a multiple of FWHM for all tests that returned a false-negative result. This clearly shows a dramatic decrease in false negatives at separations greater than the FWHM and that features with a greater separation only returned a false negative if they were quite narrow, i.e., Δν < 1 km s⁻¹, or equivalently Δν < 10 channels where the channel width is 0.1 km s⁻¹. We attribute this to the lower number of data points across these features: the improvement to the likelihood distribution of fitting a narrow feature that only covers a few channels is much less than that of fitting a wide feature (see Equation (10)).

In a similar result to the previous test, the majority of fits categorized as “poorly fit” according to our criteria were not visually unreasonable, with typical examples appearing similar to those in Figure 9. However, for both on–off and optical-depth-only spectra there were a significant number of false-positive
Figure 9. Typical examples of synthetic spectra “poorly fit” by AMOEBA while testing the effect of signal-to-noise ratio. The first two columns show a single set of synthetic on–off spectra with optical depth at the left and expected brightness temperature at the right. The third column shows a different set of synthetic optical-depth spectra only. The top row of each column show the 1612 MHz synthetic data (gray), the noiseless “actual” feature (green) and the fit returned by AMOEBA (black). The second, third, and fourth rows then show the 1665, 1667, and 1720 MHz transitions, respectively. The bottom row shows the residuals of each transition. For both the on–off spectra (left) and the optical-depth spectra (right), the actual values of the spectra parameters did not fall within the 99.95% credibility interval of the sampled posterior probability distribution.

Figure 10. Fraction of synthetic “on–off” spectra (left) and optical-depth spectra (right) for which the 99.95% credibility interval returned by AMOEBA contained all the correct parameters (“well fit”), contained at least one but not all of the correct parameters (“poorly fit”), where AMOEBA failed to identify any Gaussian features (“false negative”), or where AMOEBA recovered more than one feature or where the returned credibility intervals did not contain any of the correct parameters (“false positive”), plotted against the signal-to-noise ratio of all synthesized spectra.
results. For synthetic on–off spectra, the majority of these did successfully identify the two correct features but then detected additional features. For synthetic optical-depth spectra, approximately half of the total false positives successfully identified the two correct features (as well as one or more additional features), while the other half failed to identify either. In the vast majority of cases the spurious features had very low column density (log$_{10}$N$_{\text{c}}$ ≤ 10) or optical depth (∥T$_{\text{peak}}$∥ ≤ 0.001). We therefore suggest that the user take care when interpreting the identification of such low column density or optical-depth features as they may indeed be false positives.

5. Conclusions

Here we have introduced AMOEBA, an automated algorithm for the Bayesian Gaussian decomposition of hydroxyl spectra. AMOEBA takes a novel approach to the analysis of hydroxyl spectra by fitting on–off observations of all four ground-state transitions simultaneously. This approach takes full advantage of the expected relationships between the features of these spectra, namely that components arising from the same ISM cloud should share the same centroid velocity and FWHM and their peak values in each of the 8 spectra should be described uniquely by the column densities in the four levels of the ground-rotational state. It is also versatile in its ability to take as input both on–off spectra or optical-depth spectra only.

We tested AMOEBA extensively, demonstrating the validity of its core Bayesian algorithm, its ability to recover the parameters of the synthetic spectra, as well as the impact of signal-to-noise ratio and feature separation on its ability to identify features. As summarized in Figures 7 and 8, for 90% of trials using unblended features with a signal-to-noise ratio of 5, AMOEBA is accurate to within 10% when returning the centroid velocity, FWHM, excitation temperature, log column density, and optical depth of synthetic spectra. When fitting “on–off” spectra, it returns accurate parameters (i.e., where the “true” parameters are contained within the 99.95% credibility interval; see Section 4.3 for a detailed explanation of this criterion) in over 90% of trials above a signal-to-noise ratio of 2, and for optical-depth spectra above a signal-to-noise level of 3 (with ≈80% accuracy at a signal-to-noise ratio of 2). In the case of two blended, identical features present in synthetic on–off spectra, AMOEBA is able to correctly identify both features in ≈90% of trials at a separation of 1.5 times the FWHM and at 1.75 times the FWHM for optical-depth spectra. However, in the case of blended spectra with separations between 1 and 2 times the FWHM AMOEBA yielded false positives in a significant number of trials: ≈15% for on–off spectra and ≈20% for optical-depth spectra.

AMOEBA is a powerful algorithm for the statistically robust Gaussian decomposition of hydroxyl spectra. By fitting all four
ground-rotational state transitions simultaneously, AMOEBA is able to take full advantage of available data and produce self-consistent results with good precision. In fitting these spectra, AMOEBA returns centroid velocity, FWHM, log column density in the lowest hyperfine level, and inverse excitation temperature in the 1612, 1665, and 1667 MHz transitions. These parameters, in conjunction with a non-LTE molecular excitation code, can then be used to find local parameters of the cloud, such as gas temperature, number density, and local radiation field. AMOEBA is completely automated and therefore an excellent tool for large-scale surveys in hydroxyl, such as THOR (Beuther et al. 2016).

Finally, it is worth noting that AMOEBA’s core algorithm could be adapted to the analysis of other species. This would be most straightforward for directly analogous systems—i.e., in which Δ-doubling and hyperfine splitting results in four levels connected by three or four transitions (e.g., the 700 MHz lines of the first excited rotational state of CH). More generally, the code could be usefully adapted to any multi-line system in which the observed transitions share energy levels to a sufficient degree that the model’s excitation parameters must obey internal constraints. Users are encouraged to adapt and modify AMOEBA to suit their own use cases.

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

A.H. would like to acknowledge the help and advice received from members of the Astrostatistics FaceBook group (https://www.facebook.com/groups/astro.r) during the development of AMOEBA. A.H. is the recipient of an Australian Government Research Training Program (RTP) stipend and tuition fee offset scholarship. J.R.D. is the recipient of an Australian Research Council (ARC) DECRA Fellowship (project number DE170101086), which partially supported this work. This work was performed on the OzSTAR national facility at Swinburne University of Technology. The OzSTAR program receives funding in part from the Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) allocation provided by the Australian Government.

Appendix A

Fitting a Sample Spectrum

Here we outline the process by which AMOEBA trials progressively more complex models using a set of synthetic on–off spectra (the same as in Figure 3) containing three Gaussian-shaped profile features. The values of the parameters for these three Gaussian features are given in Table 5.

This set of synthetic spectra was fit with AMOEBA a total of 26 times as a test of reliability. Though there was a systematic shift in the values found for some parameters, the standard deviation of the values found across these trials for each parameter was always significantly less (90–95%) than the 68% confidence interval, and all were assessed as “well fit” according to the definition used in Section 4. During these 26 trials AMOEBA was set to only trial a single “extra” Gaussian feature after identifying one that did not improve the evidence. Therefore, included in the figures below will be two “over-fit” models for each velocity range.

Each panel in Figures 13 and 14 shows the synthetic optical-depth spectra at left and the expected brightness temperature at right in gray, and the velocity at which the spectra are cut with a vertical blue line. The “current” model is shown in black, with the individual Gaussian features that make up that model shown in red with peaks indicated by a red circle. The residuals of the current model are shown in the bottom subplot, with 1612 MHz in red, 1665 MHz in gold, 1667 MHz in cyan, and 1720 MHz in blue. The parameters of the models shown are outlined in Table 6.

After fitting a null model to the spectra from −10 to 1.5 km s$^{-1}$ (resulting in log$_{10}$ null evidence $= -725$), AMOEBA then sampled the posterior probability distribution of a model with a single Gaussian component (“Step 1” in Figure 13). This model had log$_{10}$ evidence = 1019 and was therefore accepted over the null model. AMOEBA then sampled the probability distribution of a model with two Gaussian components (“Step 2” in Figure 13). This model had log$_{10}$ evidence = 1087 and was therefore accepted over the model with a single Gaussian. This model was found to have the highest evidence of those trialed and was accepted as the final model for this velocity range, and is included in the “Final” plot in Figure 16.

After accepting the model with two Gaussian components, AMOEBA then sampled the posterior probability distribution of a model with three Gaussian components (“Step 3” in Figure 14). In this case, AMOEBA identified the same two components it had accepted in the previous model, as well as a third component. For the sake of illustration, an arbitrary example of this fit from the 26 trials is shown in Figure 14, but the median position in parameter space of the converged Markov chains varied widely (the third Gaussian was placed anywhere from −9 to −1 km s$^{-1}$). However, the average measured evidence of this model (log$_{10}$ evidence = 1050) did not vary significantly, and the model was rejected each time. The posterior probability distribution of a model with 4 Gaussian components was then sampled (“Step 4” in Figure 14). Again, the components from the previously accepted model were identified, along with two additional features. While the median positions of these extra features again varied widely, the evidence was relatively constant (log$_{10}$ evidence $= 1000$) and the model was rejected each time.

AMOEBA then moved to the next velocity range from 1.5 to 10 km s$^{-1}$ which had log$_{10}$ null evidence $= 807$. AMOEBA then sampled the posterior probability distribution of progressively more complex models in the same way as for the previous velocity range and the models represented by the median positions of the converged Markov chains are illustrated in “Step 5” and “Step 6” in Figure 15 and “Step 7” in Figure 16. The model with a single Gaussian component was found to have log$_{10}$ evidence $= 834$ and was therefore accepted over the null model. When the posterior probability distributions of models with two and three Gaussian components were sampled, the median position of the converged Markov chains varied, but the log$_{10}$ evidence values were relatively constant at 806 and 764, respectively, and these models were rejected in all 26 trials. The lower panel of Figure 16 (“Final”) then shows the final model accepted by AMOEBA.
Figure 13. Illustration of the process by which AMOEBA fits a set of synthetic on–off spectra. The top and bottom panels each show the synthetic spectra in gray, with the optical-depth spectra on the left and the expected brightness temperature spectra on the right. These are each shown for the 1612, 1665, 1667, and 1720 MHz transitions of the ground-rotational state of OH. The spectra are first divided into separate velocity ranges as described by the text, indicated in these plots by the vertical blue line at 1.5 km s$^{-1}$. In the first step (top panel), a single Gaussian feature is placed in the first velocity range. This feature improved on the evidence of the null model (see Table 6), so a model with two Gaussian features was then trialed in step 2 (bottom panel). This model improved the evidence over the model with a single Gaussian. The process continues in Figure 14.
Figure 14. Continuing from Figure 13, in step 3 (top panel) a model with three Gaussian features was trialed, but this model did not improve the evidence over the model with two Gaussian features (see Table 6). Nonetheless, AMOEBA trialed a model with four Gaussian features, which also did not improve the evidence over the model with two Gaussians (step 4, bottom panel). Therefore, the model with two Gaussians was accepted and AMOEBA moved on to the next velocity range, continued in Figure 15.
Figure 15. Continuing from Figure 14, a model with one Gaussian feature was trialed in the second velocity range (step 5, top panel), which improved on the evidence of the null model (see Table 6). The model with two Gaussian features was then trialed (step 6, bottom panel), but it did not improve the evidence over the model with one Gaussian. The process continues in Figure 16.
Step 7:

Continuing from Figure 15, a model with three Gaussian features was trialed in the second velocity range (step 7, top panel), which also failed to improve the evidence over the model with a single Gaussian (see Table 6). Therefore the model with a single Gaussian was accepted, and along with the model with two Gaussians from the previous velocity range, form the final result shown in the bottom panel.

Final:

Figure 16. Continuing from Figure 15, a model with three Gaussian features was trialed in the second velocity range (step 7, top panel), which also failed to improve the evidence over the model with a single Gaussian (see Table 6). Therefore the model with a single Gaussian was accepted, and along with the model with two Gaussians from the previous velocity range, form the final result shown in the bottom panel.
Appendix B

Typical Corner Plots

Typical corner plots showing the joint distributions of the model parameters for a single Gaussian feature are shown in Figure 17 for on–off spectra, and in Figure 18 for optical-depth-only spectra. No significant correlations between the parameters are seen. Some weak correlations between pairs of parameters are seen occasionally (i.e., between log$N_1$ and log $\Delta v$ in Figure 17), but these were not systematic.
Figure 17. A typical corner plot from a single test of AMOEBA on a set of synthetic on–off spectra with a single Gaussian feature and a signal-to-noise ratio of 5. The black dotted lines on each histogram indicate the median and the 68% credibility interval of each parameter. The true parameter values are indicated with red dotted lines on each histogram and with red crosses on each scatter plot. Only minor correlations between some pairs of parameters are seen in this example. This plot was generated using the Corner package in PYTHON (Foreman-Mackey 2016).
Figure 18. Same as Figure 17 but for synthetic optical-depth-only spectra. No significant correlations are seen in this example.
Anita Hafner ORCID iDs https://orcid.org/0000-0001-6179-0606
J. R. Dawson ORCID iDs https://orcid.org/0000-0003-0235-3347
Mark Wardle ORCID iDs https://orcid.org/0000-0002-1737-0871

References

Allen, R. J., Hogg, D. E., & Engelke, P. D. 2015, AJ, 149, 123
Allen, R. J., Ivette Rodríguez, M., Black, J. H., & Booth, R. S. 2012, AJ, 143, 97
Allison, J. R., Sadler, E. M., & Whiting, M. T. 2012, PASA, 29, 221
Barriault, L., Joncas, G., Lockman, F. J., & Martin, P. G. 2010, MNRAS, 407, 2645
Benincasa, S. M., Tasker, E. J., Pudritz, R. E., & Wadsley, J. 2013, ApJ, 776, 23
Beuther, H., Birr, S., Rugel, M., et al. 2016, A&A, 595, A32
Buchner, J. 2021, arXiv:2101.09675
Busch, M. P., Allen, R. J., Engelke, P. D., et al. 2019, ApJ, 883, 158
Dawson, J. R., Walsh, A. J., Jones, P. A., et al. 2014, MNRAS, 439, 1596
Dickey, J. M., McClure-Griffiths, N., Gibson, S. J., et al. 2013, PASA, 30, e003
Ebisawa, Y., Inokuma, H., Sakai, N., et al. 2015, ApJ, 815, 13
Ebisawa, Y., Sakai, N., Menten, K. M., & Yamamoto, S. 2019, ApJ, 871, 89
Elitzur, M. 1976, ApJ, 203, 124
Elitzur, M., Goldreich, P., & Scoville, N. 1976, ApJ, 205, 384
Engelke, P. D., & Allen, R. J. 2018, ApJ, 858, 57
Engelke, P. D., & Allen, R. J. 2019, ApJ, 874, 49
Feroz, F., & Hobson, M. P. 2008, MNRAS, 384, 449
Feroz, F., & Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Foreman-Mackey, D. 2016, JOSS, 1, 24
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Goss, W. M. 1968, ApJS, 15, 131
Grenier, I. A., Casandjian, J.-M., & Terrier, R. 2005, Sci, 307, 1292
Guilbert, J., Rieu, N. Q., & Elitzur, M. 1978, A&A, 66, 395
Haud, U. 2000, A&A, 364, 83
Heiles, C. 1969, ApJ, 157, 123
Heiles, C., & Troland, T. H. 2003, ApJS, 145, 329
Hennebelle, P., Banerjee, R., Vázquez-Semadeni, E., Klessen, R. S., & Audit, E. 2008, A&A, 486, L43
Henshaw, J. D., Ginsburg, A., Haworth, T. J., et al. 2019, MNRAS, 485, 2457
Jeffreys, H. 1961, Theory of probability (3rd ed.; Oxford: Clarendon), 447
Kass, R. E., & Raftery, A. E. 1995, J. Am. Stat. Assoc., 90, 773
Li, D., Tang, N., Nguyen, H., et al. 2018, ApJS, 235, 1
Lindner, R. R., Vera-Ciro, C., Murray, C. E., et al. 2015, AJ, 149, 138
Liszt, H., & Lucas, R. 1996, A&A, 314, 917
Marchal, A., Miville-Deschênes, M.-A., Orme, J., & et al. 2019, A&A, 626, A101
Nguyen, H., Dawson, J. R., Lee, M.-Y., et al. 2019, ApJ, 880, 141
Nguyen, H., Dawson, J. R., Miville-Deschênes, M. A., et al. 2018, ApJ, 862, 49
Petzler, A., Dawson, J. R., & Wardle, M. 2020, MNRAS, 497, 4066
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2011, A&A, 536, A19
Reach, W. T., Koo, B.-C., & Heiles, C. 1994, ApJ, 429, 672
Renn, Q., Grenier, I. A., Marshall, D. J., & Casandjian, J. M. 2018, A&A, 611, A51
Riener, M., Kainulainen, J., Henshaw, J. D., et al. 2019, A&A, 628, A78
Robinson, B. J., & McGee, R. X. 1967, ARA&A, 5, 183
Rugel, M. R., Beuther, H., Birr, S., et al. 2018, A&A, 618, A159
Talts, S., Betancourt, M., Simpson, D., Vehtari, A., & Gelman, A. 2018, arXiv:1804.06788
Turner, B. E. 1979, A&AS, 37, 1
van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J., & van Dishoeck, E. F. 2007, A&A, 468, 627
Wannier, P. G., Andersson, B. G., Federman, S. R., et al. 1993, ApJ, 407, 163
Weinreb, S., Barrett, A. H., Meeks, M. L., & Henry, J. C. 1963, Natur, 200, 829
Wollf, M. G., Hollenbach, D., & McKee, C. F. 2010, ApJ, 716, 1191
Xu, D., Li, D., Yue, N., & Goldsmith, P. F. 2016, ApJ, 819, 22