









# Detecting Weak Spectral Lines in Interferometric Data through Matched Filtering

Ryan A. Loomis<sup>1</sup> , Karin I. Öberg<sup>1</sup> , Sean M. Andrews<sup>1</sup> , Catherine Walsh<sup>2,3</sup> , Ian Czekala<sup>1,4,5</sup> , Jane Huang<sup>1</sup> , and Katherine A. Rosenfeld<sup>1</sup>

<sup>1</sup>Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA; [rloomis@cfa.harvard.edu](mailto:rloomis@cfa.harvard.edu)

<sup>2</sup>Leiden Observatory, Leiden University, P.O. Box 9531, 2300 RA Leiden, The Netherlands

<sup>3</sup>School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK

<sup>4</sup>Stanford University, 452 Lomita Mall, Stanford, CA, 94305, USA

Received 2017 July 19; revised 2018 March 6; accepted 2018 March 8; published 2018 April 5

## Abstract

Modern radio interferometers enable observations of spectral lines with unprecedented spatial resolution and sensitivity. In spite of these technical advances, many lines of interest are still at best weakly detected and therefore necessitate detection and analysis techniques specialized for the low signal-to-noise ratio (S/N) regime. Matched filters can leverage knowledge of the source structure and kinematics to increase sensitivity of spectral line observations. Application of the filter in the native Fourier domain improves S/N while simultaneously avoiding the computational cost and ambiguities associated with imaging, making matched filtering a fast and robust method for weak spectral line detection. We demonstrate how an approximate matched filter can be constructed from a previously observed line or from a model of the source, and we show how this filter can be used to robustly infer a detection significance for weak spectral lines. When applied to ALMA Cycle 2 observations of CH<sub>3</sub>OH in the protoplanetary disk around TW Hya, the technique yields a  $\approx 53\%$  S/N boost over aperture-based spectral extraction methods, and we show that an even higher boost will be achieved for observations at higher spatial resolution. A Python-based open-source implementation of this technique is available under the MIT license at <https://github.com/AstroChem/VISIBLE>.

**Key words:** methods: data analysis – protoplanetary disks – radio lines: general – submillimeter: planetary systems – techniques: interferometric – techniques: spectroscopic

## 1. Introduction

The rich spatio-kinematic information that radio interferometric data sets can provide for molecular spectral lines is crucial for studying the astrophysical and chemical processes occurring in host sources. The broadband capabilities of modern interferometers allow many spectral lines to be observed in a single correlator setup, enabling astronomers to simultaneously trace multiple astrophysical phenomena, or undertake unbiased line surveys to work toward complete molecular inventories (e.g., Jørgensen et al. 2011; Coutens et al. 2016; Müller et al. 2016). Within these data sets, many scientifically interesting lines may be weak or not detected due to low column densities or intrinsically low line strengths. Finding these lines and robustly assessing their strength is key to achieving many science goals.

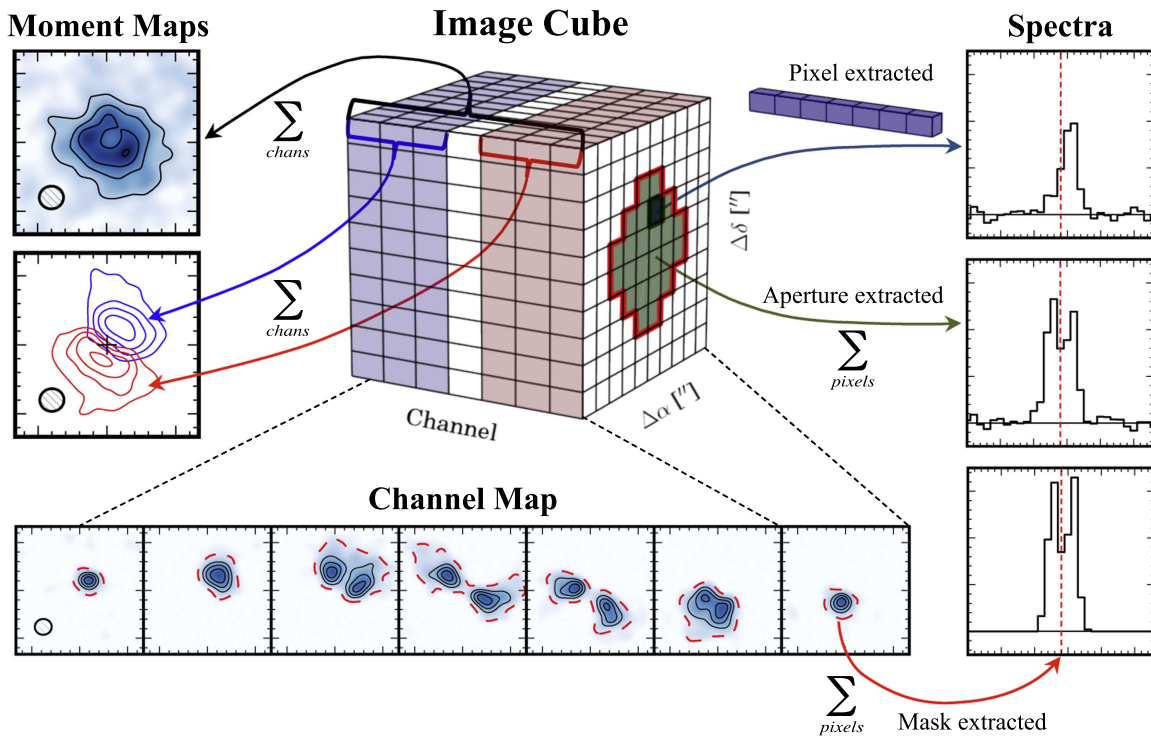
Resolved interferometric observations pose special challenges to detecting weak spectral lines. Radio interferometers measure visibilities, samples of the Fourier transform of the distribution of emission intensities from an astrophysical source at discrete spatial and spectral frequencies. These visibilities are then Fourier inverted and deconvolved with a routine such as CLEAN (Högbom 1974) to create an image cube. As shown in Figure 1, this image cube consists of a series of images (channel maps) of the emission intensity distribution in distinct spectral frequency bins, which correspond to projected radial velocity bins. In the simplest line detection scenario, emission is directly observed in these channel maps.

When emission is too weak to be directly visible in the channel maps, the image cube might be manipulated in a variety of ways to increase the signal-to-noise ratio (S/N).

Spectra can be extracted from the cube, and moment maps can be generated by collapsing the cube along the spectral axis, illustrated in Figure 1. All spectral extraction approaches incorporate a spatial mask. If the source is unresolved, a single pixel-extracted spectrum will contain all available information. In cases where the emission is extended and spatially resolved, the simplest mask that contains all emission is an aperture drawn around the source. Such a mask rarely results in spectra with optimal S/N, however. In sources with complex spatio-kinematic patterns, e.g. due to bulk rotation, emission may “move” across the channel maps. The aperture mask is then larger than the emitting area in any given channel, adding noise to the extracted spectrum. To combat this, a spatio-kinematic mask specifically tailored to the structure of the source may be used to reduce the amount of added noise (e.g., Dutrey et al. 2007; Loomis et al. 2015; Öberg et al. 2015; Yen et al. 2016).

The application of spatio-kinematic masks to spectral image cubes has already enabled new science, but there are both computational and interpretive challenges when attempting to extend this technique to detect weak lines. First, the observed visibilities must be imaged, a non-trivial computational cost for high-resolution observations or spectral surveys with large bandwidths. Second, when the visibilities are Fourier inverted, the point-spread function is oversampled with pixels to reduce imaging artifacts. This introduces a spatial covariance between pixels on the scale of the beam, making statistical interpretations of extracted spectra difficult. Finally, tailored spatio-kinematic masks reduce added noise but sacrifice a meaningful spectral baseline, making robust weak line detection difficult unless more complicated bootstrapping approaches are taken to establish a false positive rate (Barenfeld et al. 2016).

<sup>5</sup> KIPAC Postdoctoral Fellow.



**Figure 1.** Multiple ways of viewing an image cube. Counter-clockwise from the top left: velocity-integrated moment maps, made by integrating slices of the cube along the frequency axis; channel maps, where each panel corresponds to a channel of the cube; spectra, generated from top to bottom from a single pixel, integrated over an aperture, and integrated using a matched spatio-kinematic mask (dashed red contours in channel maps). The synthesized beam is shown in the lower left of the moment and channel maps.

These obstacles can be overcome while retaining the benefits of the spatio-kinematic approach by applying a matched filter directly to the observed visibilities. When the shape of a signal is known, the optimal linear filter for signal extraction is a matched filter, equivalent to the known signal with a normalization constant. Cross-correlating a noisy signal with this filter maximizes the output S/N. This approach is used extensively in digital signal processing; prominent examples include RADAR (e.g., Woodward 1953; Cumming & Wong 2005, and references therein), source detection in imaging surveys (e.g., Bertin & Arnouts 1996; Bertin 2001; Meillier et al. 2016; Herenz & Wisotzki 2017; Zackay & Ofek 2017), gravitational wave detection (e.g., Owen & Sathyaprakash 1999; Schutz 1999; Abbott et al. 2016), and exoplanet detection through direct imaging (Ruffio et al. 2017).

Because matched filters are simply cross-correlated with the data, they can be easily applied in the Fourier domain to provide a fast and unbiased approach to weak line detection over broad bandwidths. An image-plane spectral extraction mask reduces unnecessary noise contributions by incorporating estimated spatio-kinematic (and, correspondingly, interferometric phase) information into the extracted spectrum. Similarly, matched filtering quantitatively combines both amplitude and phase information of the observed visibilities into a robust detection probability. Line detection directly in the visibilities both avoids the high computational expense of fully imaging wide-bandwidth data sets and retains a straightforward statistical interpretation of detection significance.

In this paper, we describe how to construct and apply a matched filter to interferometric spectral line data and demonstrate the method on observations from the Atacama Large Millimeter/sub-millimeter Array (ALMA). In Section 2,

we provide an overview of matched filtering and detail the steps of the method. In Section 3, we apply the technique to ALMA Cycle 2 observations from Walsh et al. (2016) of CH<sub>3</sub>OH in a protoplanetary disk. In Section 4, we discuss how much S/N boost might be expected for a given data set, compare the technique to other methods, and suggest applications where matched filtering may prove useful. A summary is given in Section 5. Formulas to approximate the expected S/N boost are derived in Appendices A and B, derivations of noise covariance matrices for correlated channels are presented in Appendix C, and details of the example filters used in the paper are given in Appendix D.

## 2. Method

In this section we first present a brief overview of the principles behind matched filtering, introducing the one-dimensional matched filter. The one-dimensional approach is then easily extended to higher-dimensional problems such as searching for signals within an image or image cube (e.g., Bertin & Arnouts 1996; Feng et al. 2017; Herenz & Wisotzki 2017; White & Padmanabhan 2017). We present here a novel method to apply matched filters in the native measurement space of interferometric data, the incompletely sampled Fourier ( $u, v$ ) plane. After defining interferometric visibilities and their noise properties, we provide detailed instruction and examples for each of the steps in the method:

1. Generation of a ( $u, v$ ) plane filter which approximates the true emission pattern.
2. Cross-correlation of this filter with the measured visibilities.
3. Spectrum normalization and detection inference.

## 4. Line stacking (where applicable).

## 2.1. Matched Filtering

The matched filter can be derived in a number of ways. Here, we introduce its derivation by maximizing the S/N of a signal, but it can equivalently be interpreted as a least-squares estimator (see e.g., Schwartz & Shaw 1975; Vio & Andreani 2016). In general, a signal  $s$  may be corrupted by additive white noise  $\mathbf{v}$ , yielding an observation  $\mathbf{x} = \mathbf{s} + \mathbf{v}$ . To maximize the S/N of this signal by applying a linear filter  $\mathbf{h}$ , we can first write the S/N (using the definition of signal power/noise power) as

$$\text{S/N} = \frac{\mathbf{h}^* \mathbf{s} \mathbf{s}^* \mathbf{h}}{\mathbf{h}^* \mathbf{R}_v \mathbf{h}}, \quad (1)$$

where  $*$  denotes the conjugate transpose and  $\mathbf{R}_v = E[\mathbf{v}\mathbf{v}^*]$  is a covariance matrix of the noise  $\mathbf{v}$ , where  $E[\ ]$  is the expectation operator. Under these conditions, the filter  $\mathbf{h}$  which maximizes S/N is

$$\mathbf{h} = \left[ \frac{1}{\sqrt{\mathbf{s}^* \mathbf{R}_v^{-1} \mathbf{s}}} \right] \mathbf{R}_v^{-1} \mathbf{s} = C \mathbf{R}_v^{-1} \mathbf{s}, \quad (2)$$

i.e., the original signal  $s$  multiplied by the data weights  $\mathbf{R}_v^{-1}$  and a normalization constant  $C = 1/\sqrt{\mathbf{s}^* \mathbf{R}_v^{-1} \mathbf{s}}$  (e.g., Woodward 1953; North 1963; Cumming & Wong 2005).

A simple application of such a filter is locating a signal within a one-dimensional data set such as an emission spectrum. In this case, a short signal  $s$  of length  $n_s$  is embedded within a longer noisy observed spectrum  $\mathbf{x}$  of length  $n_x$ , with the location of  $s$  within  $\mathbf{x}$  unknown. As long as the shape of  $s$  is known, a filter kernel  $\mathbf{h}$  can be calculated using Equation (2) and cross-correlated with  $\mathbf{x}$  to locate  $s$ . This cross-correlation is often thought of as a sliding dot product, and yields a one-dimensional impulse response spectrum  $\mathbf{T}$ , of length  $n_x - n_s + 1$ . Each element of  $\mathbf{T}$  at a position  $i_0$  will then be:

$$T_{i_0} = \sum_{i=i_0}^{i_0+N_s-1} x_i h_{i-i_0}; \quad i_0 \in [0, n_x - n_s]. \quad (3)$$

This impulse response spectrum loses physical significance that the original observed spectrum held (it is no longer in units of power or flux). It instead encodes the degree of similarity between the observations and the filter at any given point in the observed spectrum. By projecting the noisy observations in this way, the total noise is decreased and the S/N of the signal is increased. Because the filter is linear, the Gaussian nature of the noise is preserved. The impulse response spectrum can be easily examined for evidence of  $s$ , with a detection threshold set to some multiple of the standard deviation of  $\mathbf{T}$  (e.g.,  $4\sigma$ ), or a false positive rate scaled with the variance of  $\mathbf{T}$ .

## 2.2. Interferometric Visibilities

Each interferometric visibility,  $V_i$ , is measured as the complex product of the output of the first antenna of a baseline pair in the array with the complex conjugate of the output of the second antenna (see, e.g., Thompson et al. 2017). The projected baseline distance between the two antennas then defines the location of the visibility on the  $(u, v)$  plane. Each visibility  $V_i$  is

associated with a unique weight  $w_i = 1/\sigma_i^2$ , where  $\sigma_i^2$  encodes the variance of  $V_i$  (see Appendix C for more details on data weights in interferometric data sets). In addition to being measured at discrete spatial frequencies, the visibilities are also measured at a series of spectral frequencies (channels). Cross-correlation in this discretely sampled three-dimensional  $(u, v, \text{channel})$  space is computationally awkward, but the data set can be reshaped to a two-dimensional data set of size  $(n_{uv}, n_c)$ , where each visibility row corresponds to a unique location on the  $(u, v)$  plane in units of distance. Both the complex visibilities and their corresponding real weights are stored this way in the UVFITS<sup>6</sup> and Measurement Set (MS)<sup>7</sup> formats of the Common Astronomy Software Applications package (CASA).

Transforming between visibility space and image space requires a gridding and deconvolution routine, such as CLEAN, in one direction and a visibility sampling routine in the other direction, such as uvmodel in MIRIAD, or simobserve in CASA. As using the full simobserve task is relatively slow and uvmodel is not easily interfaced with Python, we have written a Python-based visibility sampling routine, vis\_sample, which is able to interface with CASA MS and UVFITS formats. This package builds on an implementation of the sampling algorithm in the DiskJockey package (Czekala et al. 2015; Czekala 2016) identical to that used in uvmodel and simobserve and uses the spheroidal gridding function approximations described by Schwab (1984). As identical algorithms and gridding functions are used, output from vis\_sample is identical to output from uvmodel and simobserve.<sup>8</sup>

## 2.3. Filter Kernel Generation

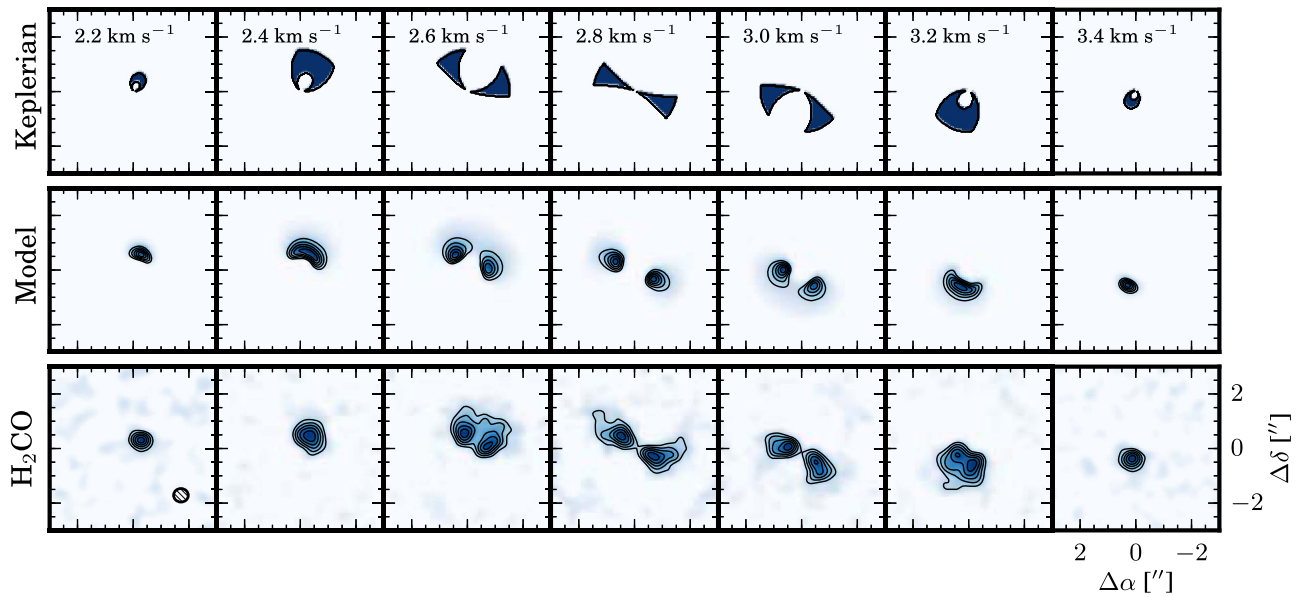
The principal assumption of a matched filter analysis is that the shape of the signal  $s$  is known, or can be reasonably approximated. In traditional applications, such as RADAR, the outbound signal is user-generated and therefore the exact form is known. In astronomical applications, however, the ideal matched filter kernel is unknown and must be approximated. As it is unknown how closely the filter approximates the true signal, any derived detection significance will be a lower limit. The method is relatively robust to choice of filter, however, as long as the filter is a reasonable approximation of the source spatio-kinematic structure.

We suggest two possible approaches: (1) calculating a kernel from a model of the source (model-driven), or (2) calculating a kernel from prior observations of strong emission lines (data-driven). In both cases the kernel is first constructed in the image plane and then Fourier transformed and visibility sampled to match the  $(u, v)$  coverage of the observations. The approximated signal  $\mathbf{f}$  and the inverted noise covariance matrix  $\mathbf{R}_v^{-1}$  (calculated from the observational data weights, see Appendix C) are then used to compute the full filter kernel,

<sup>6</sup> The UVFITS format definition can be found in AIPS Memo #114 at <http://www.aips.nrao.edu/aipsmemo.html>.

<sup>7</sup> The MS format definition can be found at <https://casa.nrao.edu/Memos/229.html>.

<sup>8</sup> In addition to its utility for filter kernel generation, we note that vis\_sample may be useful for visibility fitting of modern interferometric data sets (e.g., MacGregor et al. 2016; Loomis et al. 2017). vis\_sample is publicly available under the MIT license at <https://github.com/AstroChem/VISIBLE> or in the Anaconda Cloud at [https://anaconda.org/rloomis/vis\\_sample](https://anaconda.org/rloomis/vis_sample).



**Figure 2.** Three examples of channel maps used to generate filter kernels for emission from the protoplanetary disk around TW Hya. Top: a simple kernel based on Keplerian rotation. Middle: a kernel based on a parametric model of CH<sub>3</sub>OH from Walsh et al. (2016). Bottom: a data-driven kernel generated from H<sub>2</sub>CO observations from Öberg et al. (2017). All kernels have 0.2 km s<sup>-1</sup> channels and are normalized by their peak intensities.

including the normalization prefactor:

$$\mathbf{h} = \left[ \frac{1}{\sqrt{\mathbf{f}^* \mathbf{R}_v^{-1} \mathbf{f}}} \right] \mathbf{R}_v^{-1} \mathbf{f}. \quad (4)$$

Figure 2 presents three examples of the different filter kernel estimation approaches. First, for objects such as protoplanetary disks or galaxies, the source inclination and position angle are often well-known and a spatio-kinematic model of the gas can be approximated. In the top panels of Figure 2 we have generated a Keplerian mask for molecular emission from the protoplanetary disk around TW Hya. Alternatively, a more detailed filter kernel can be generated from an astrochemical model of the source, with emission calculated using a radiative transfer code such as RADMC-3D (Dullemond 2012) or LIME (Brinch & Hogerheijde 2010). An example is shown in the middle panels of Figure 2, generated from the parametric CH<sub>3</sub>OH abundance model in Walsh et al. (2016). Details of the model are presented in Appendix D.

In the data-driven approach, an assumption is made that an observed molecular transition shares its spatio-kinematic pattern with the desired weak line. The filter will be most effective when the template and target lines have well-matched spatial distributions, e.g., if the two lines are a strong and a weak line, respectively, of the same molecule. In Carney et al. (2017), we used this approach to detect weak H<sub>2</sub>CO lines in HD 163296, using a stronger H<sub>2</sub>CO line as a data-driven filter (see Sections 4.1 and 4.4). Similarly, lines of a known species can be used as a filter for an undetected but chemically related molecule that is presumed to be co-spatial. The bottom row panels of Figure 2 present observations of H<sub>2</sub>CO around TW Hya (Öberg et al. 2017) which could be used as a filter for CH<sub>3</sub>OH emission, due to their linked formation pathways (e.g., Cuppen et al. 2009; Qi et al. 2013; Walsh et al. 2014; Loomis et al. 2015). Details of the observations and kernel generation are presented in Appendix D.

#### 2.4. Computing the Impulse Response Function

Figure 3 schematically describes how these filter kernels would be applied to the data to produce an impulse response spectrum. First, the image plane kernel is Fourier transformed and visibility sampled to produce a complex  $(u, v)$  plane kernel  $\mathbf{f}$  of size  $(n_{uv}, n_k)$ . The inverted noise covariance matrix  $\mathbf{R}_v^{-1}$  is calculated from the data weights  $\mathbf{w}$  and combined with  $\mathbf{f}$  using Equation (4) to produce the full filter kernel  $\mathbf{h}$ . This kernel is then cross-correlated with the data  $\mathbf{V}$ , of size  $(n_{uv}, n_c)$ . The kernel and the data both have the same number of visibilities,  $n_{uv}$ , but different numbers of channels, and the kernel slides through the data along the spectral axis. At each channel, the filter impulse response spectrum  $\mathbf{T}$  is calculated by taking the complex inner product of the windowed data with the kernel:

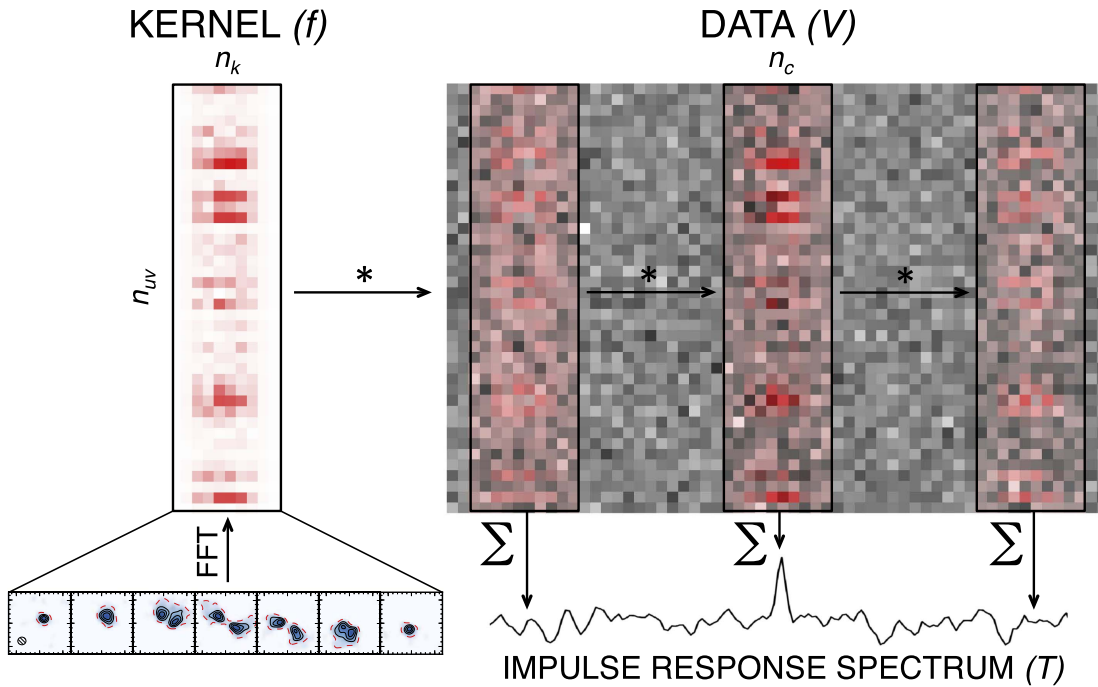
$$T_{i_0} = \sum_{i=i_0}^{i_0+n_k-1} \sum_{j=0}^{n_{uv}} V_{i,j} h_{i-i_0,j}; \quad i_0 \in [0, n_c - n_k]. \quad (5)$$

As the data are complex, the filter response spectrum is also complex with a normalized total noise power. Signal power will leak from the real to the imaginary portion of the response, however, if there is a phase misalignment between the sky locations of the filter and the source. Thus if the filter has been properly phase shifted to be aligned with the source, the resultant impulse response spectrum  $\mathbf{T}$  can be written as:

$$T_{i_0} = \text{Re} \left[ \sqrt{2} \sum_{i=i_0}^{i_0+n_k-1} \sum_{j=0}^{n_{uv}} V_{i,j} h_{i-i_0,j} \right]; \quad i_0 \in [0, n_c - n_k], \quad (6)$$

with the factor of  $\sqrt{2}$  introduced to normalize the noise power in the real portion of the spectrum.

This method of calculating the cross-correlation is conceptually simple, but computationally inefficient. Computing inner products of the windowed data requires either manipulation of the (very large) data set in memory or non-sequential memory



**Figure 3.** How the filter kernel is cross-correlated with the data to produce a spectrum of the impulse response to the filter. The kernel is shown on the left, with dimensions of  $n_k$  channels horizontally and  $n_{uv}$  visibilities vertically (not shown to scale). Several representative channels of the kernel are shown imaged. The amplitudes of the complex kernel have been binned and pixelated to be visually intuitive. The complex data shown in gray-scale are also binned and pixelated, and have an identical number of visibilities, but  $n_c \gg n_k$ . The filter is applied to the data as a sliding inner product, and three illustrative regions are shown to visualize the cross-correlation at various points. Within these regions, a stronger red color signifies a stronger correlation with the corresponding kernel value, and the real portion of the response is summed over the entire region to produce the corresponding impulse response for each channel, with the line detected in the central channel.

access, preventing speed increases through vectorization.<sup>9</sup> There is no restriction, however, on the order of operations in which the inner products are internally calculated. We use this to our advantage and treat the partial two-dimensional cross-correlation as a series of  $n_{uv}$  one-dimensional cross-correlations along the spectral axis, yielding  $n_{uv}$  individual impulse response curves. The UVFITS and MS data formats store visibilities in a row-major order such that these one-dimensional cross-correlations quickly access data sequentially in memory. The resulting impulse response curves are then summed along the spatial frequency dimension, identical to Equation (6), but with the order of the summations switched,

$$T_{i_0} = \text{Re} \left[ \sqrt{2} \sum_{j=0}^{n_{uv}} \sum_{i=i_0}^{i_0+n_k-1} V_{i,j} h_{i-i_0,j} \right]; \quad i_0 \in [0, n_c - n_k], \quad (7)$$

yielding a final impulse response spectrum identical to that from the sliding window method shown in Figure 3. The  $n_{uv}$  one-dimensional cross-correlations are independent, making parallelization trivial. Using this approach, the full bandwidth of a typical ALMA spectral window (e.g., 3840 channels over a 1 hr integration with 43 antennas) can be filtered very quickly on a desktop (e.g., a few seconds on a quad-core 3.3 GHz processor).

### 2.5. Normalization and Detection Inference

Assessing the probability of a line detection from the filter impulse response requires understanding the noise properties of the response spectrum, which no longer holds the same

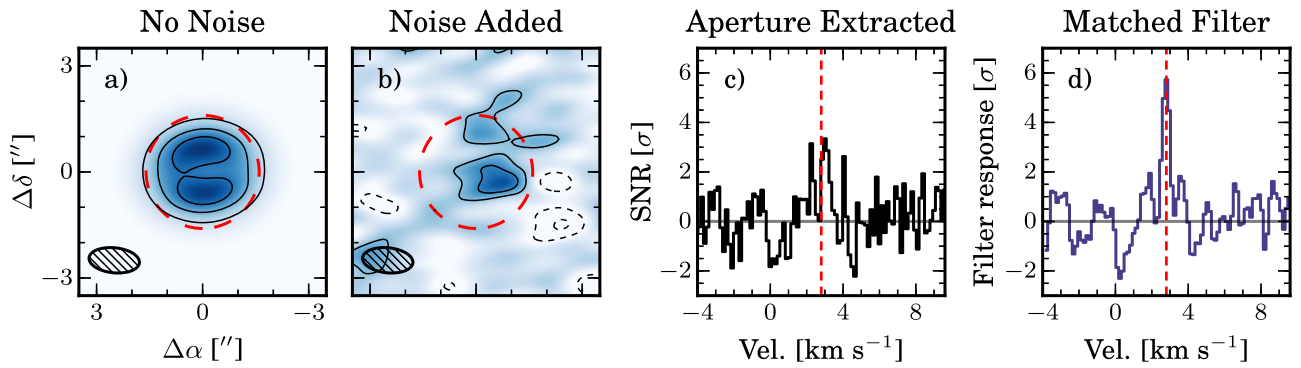
physical significance as an emission spectrum. The response spectrum at a given frequency now represents how closely the data correspond to the filter, rather than the flux at that frequency. As the filter is linear, uncorrelated thermal noise in the visibilities manifests as Gaussian noise in the filter response. If the data weights are properly calibrated (see Appendix C), the prefactor in Equation (4) will normalize the filter response such that the spectrum has units of standard deviations ( $\sigma$ ) with a rms noise level of unity.

In practice, however, we have found that the calculated data weights are often only accurate to  $\sim 20\%$  compared to the actual variance of the visibilities. Thus we strongly recommend comparing the data weights and visibility scatter and recalculating the weights using a task such as `statwt` in CASA if there is a discrepancy. Alternatively, the filter response itself can be manually normalized by dividing by the standard deviation of the spectrum (excluding any channels with obvious signal).

Additionally, the linear nature of the filter means that any unsubtracted continuum emission will result in a constant offset in the response spectrum. A model of the continuum can either be subtracted in the  $(u, v)$  plane as a pre-processing step prior to filtering using a task such as `uvcontsub` in CASA, or be subtracted after filtering by subtracting the mean of signal-free channels in the impulse response spectrum. Subtraction after filtering removes the possibility of small inaccuracies in the  $(u, v)$ -subtracted model, to which the filter will be sensitive. Conversely, subtraction prior to filtering may be more convenient when a large bandwidth is covered for a source with a non-zero spectral index.

Once the response spectrum is normalized and any offsets are removed, peaks can then be evaluated against a detection threshold, set at some number of standard deviations

<sup>9</sup> Using FFT cross-correlation is even slower for typical interferometric data set sizes.



**Figure 4.** Comparison of the ideal matched filter with conventional spectral extraction through aperture masking. Panel (a): moment-0 map of simulated, noiseless CH<sub>3</sub>OH 3<sub>12</sub>–3<sub>03</sub> emission. The synthesized beam is shown in the lower left. Contours are  $[-3, -1.5, 1.5, 3] \times 3.3 \text{ mJy } \text{bm}^{-1} \text{ km s}^{-1}$ , corresponding to  $1\sigma$  in panel (b). Panel (b): moment-0 map of simulated and noise-corrupted CH<sub>3</sub>OH emission. Panel (c): spectrum of the noise-corrupted emission, extracted using an aperture 3'' in diameter. Panel (d): ideal matched filter response to the noisy emission, with units of  $\sigma$ .

corresponding to an acceptable false alarm rate. We stress, however, that the detection significance is a lower limit, as it is unknown how closely the filter approximates the ideal matched filter.

### 2.6. Comparison to Image-plane Spectral Extraction

As a proof of concept, we apply the method to synthetic observations of CH<sub>3</sub>OH emission in TW Hya and compare to an aperture-based spectral extraction in the image plane. The modeled emission from the middle panels of Figure 2 is used to generate both the observations (with noise added) and the filter kernel. As this is a true matched filter, it provides a useful benchmark for comparison with the approximate filter results as presented in Section 3.

A synthetic MS of observations was created from the CH<sub>3</sub>OH emission cube described in Section 2.3 by visibility sampling at baselines corresponding to the observations from Walsh et al. (2016) using `vis_sample`. The complex visibilities were then noise corrupted such that the rms noise was  $5 \text{ mJy } \text{bm}^{-1}$  across each  $0.15 \text{ km s}^{-1}$  channel, equivalent to the Walsh et al. (2016) observations. The noiseless and noisy MSs were imaged in CASA using the CLEAN task, with a CLEAN mask generated from the LIME output emission profile and a circular 1'' FWHM Gaussian taper applied in the Fourier plane to increase the S/N of the images. Only the noiseless MS was CLEANed; the noisy MS was dirty imaged to prevent bias from over-CLEANing, as the emission is practically at the noise limit in any given channel. Integrated intensity (moment-0) maps of the noiseless and noisy 3<sub>12</sub>–3<sub>03</sub> transitions are shown in Figure 4, panels (a) and (b), respectively, and were generated by integrating all channels with emission. No clipping threshold was used. In the noisy case, the moment-0 map rms is  $\sim 3.3 \text{ mJy } \text{bm}^{-1} \text{ km s}^{-1}$  and the peak integrated flux is  $\sim 13.2 \text{ mJy } \text{bm}^{-1} \text{ km s}^{-1}$ , yielding a S/N of  $\sim 4$ . A spectrum was extracted from the noisy image cube using an aperture 3'' in diameter, equivalent to the extent of the CH<sub>3</sub>OH emission (Figure 4, panel (c)). The spectrum has a peak flux of  $\sim 11.4 \text{ mJy}$  and a rms noise of  $\sim 3.2 \text{ mJy}$ , yielding an S/N of  $\sim 3.5\sigma$ . The rms noise level of the noisy spectrum was estimated from all channels without significant emission (i.e., excluding a velocity range of  $\pm 1.5 \text{ km s}^{-1}$  around the systemic velocity of  $2.8 \text{ km s}^{-1}$ ).

We cross-correlate the CH<sub>3</sub>OH filter kernel with the synthetic observations, as described in Section 2.4, generating the filter response shown in Figure 4, panel (d). The peak value

of the filter response,  $5.7\sigma$ , is the maximum S/N extractable from the data and represents a  $\sim 40\%$  and  $\sim 60\%$  improvement over the moment-0 and spectral detections, respectively. This already corresponds to a factor of 2–3 increase in effective observing time but, as discussed in both Section 4.1 and Appendices A and B, the level of possible S/N improvements will be higher for data sets that are better-resolved.

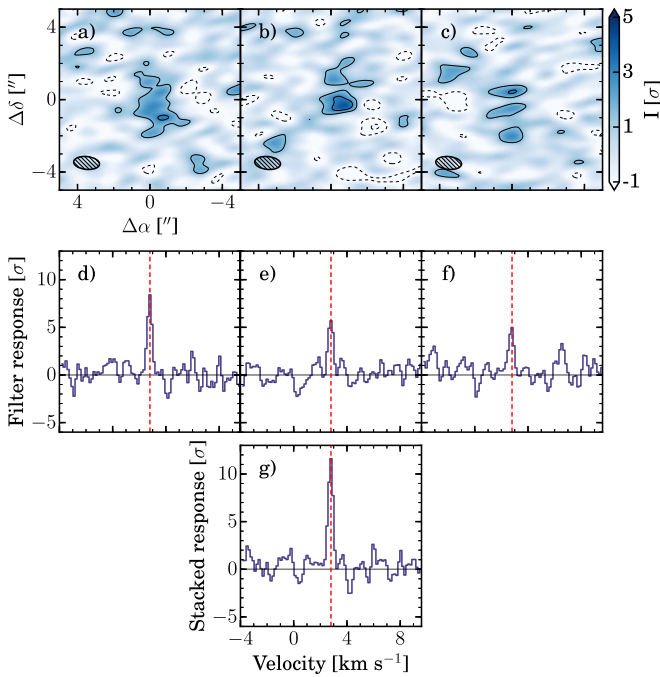
### 2.7. Stacking

Spectral stacking is a common method of S/N improvement for observations of multiple transitions of the same molecule (e.g., Langston & Turner 2007; Kalenskii & Johansson 2010; Loomis et al. 2016; Walsh et al. 2016). If the excitation conditions of two or more transitions are similar and their rest frequencies are well known, then the signals can be combined through weighted averaging,

$$\mathbf{T}_s = \sum_{i=0}^{n_s} \mathbf{T}_i w_i, \quad (8)$$

where the stacked spectrum  $\mathbf{T}_s$  is generated by summing  $n_s$  individual spectra  $\mathbf{T}_i$  multiplied by weights  $w_i$ , proportional to the S/N of each  $\mathbf{T}_i$ . Knowledge of the relative strengths of each transition is therefore important to gain the most signal improvement. Application of a matched filter results in an estimated S/N for each transition, which can be used as a proxy for their relative strengths. The resultant impulse response spectra are then easily stacked to generate an appropriately weighted stacked spectrum.

To illustrate this process, we have repeated the simulations and filtering described in Section 2.6 for three CH<sub>3</sub>OH transitions: 2<sub>11</sub>–2<sub>02</sub>, 3<sub>12</sub>–3<sub>03</sub>, and 4<sub>13</sub>–4<sub>04</sub>, with relative strengths of 1.8:1.3:1.0. Moment-0 maps of the emission from each of these transitions are shown in Figure 5, panels (a)–(c), with peak integrated fluxes of 11.7, 13.2, and 8.5 mJy km s<sup>-1</sup> and corresponding S/Ns of 3.5, 4, and 2.6σ, respectively. The individual filter responses are shown in Figure 5, panels (d)–(f), with peak S/Ns of 8.4, 5.7, and 4.9σ, respectively. The filter responses were stacked using a weighted average, yielding the spectrum shown in Figure 5, panel (g), with a peak S/N of 11.6σ. The ratio of the filter responses  $\sim (1.7:1.2:1)$  recovers the flux ratio of the input models (1.8:1.3:1.0) fairly well, even though the 2<sub>11</sub>–2<sub>02</sub> transition appears weaker than would be expected in the imaged data (likely due to random noise fluctuations in the inherently more noisy moment-0 maps). This



**Figure 5.** Demonstration of line stacking on synthetic CH<sub>3</sub>OH emission. Panels (a)–(c): moment-0 maps of the three simulated and noise-corrupted transitions,  $2_{11}-2_{02}$ ,  $3_{12}-3_{03}$ , and  $4_{13}-4_{04}$ . Contours are  $[-3, -1.5, 1.5, 3] \times \sigma$ ,  $\sigma = 3.3$  mJy  $\text{km}^{-1}$   $\text{beam}^{-1}$ . The synthesized beam is shown in the lower left. Panels (d)–(f): ideal matched filter response spectra. Peak S/Ns are 8.4, 5.7, and 4.9 $\sigma$ . Panel (g): filter response spectrum created by stacking the individual spectra from panels (d)–(f). Each spectrum was weighted by its S/N, and the resultant spectrum has an S/N of 11.6.

highlights one of the advantages of applying the matched filter in the Fourier plane.

### 3. Application to Real ALMA Data

Matched filtering provides clear benefits when the ideal filter kernel is known. To explore its utility when the filter is approximated, we have applied the method to real ALMA Band 7 observations of CH<sub>3</sub>OH toward TW Hya (project 2013.1.00902.S, P.I. C. Walsh), using all three kernels shown in Figure 2. Details of the CH<sub>3</sub>OH observations are presented by Walsh et al. (2016). They reported that the three observed CH<sub>3</sub>OH transitions ( $2_{11}-2_{02}$ ,  $3_{12}-3_{03}$ , and  $4_{13}-4_{04}$ ) were not conclusively detected in any of the individual data cubes, and therefore only presented the stacked imaging data with a 5.1 $\sigma$  detection in an aperture extracted spectrum. Moment-0 maps of the three observed CH<sub>3</sub>OH transitions are presented in Figure 6, with peak S/Ns of 4.3, 4.3, and 2.9 $\sigma$ , respectively. A stacked moment-0 map is shown on the far right with a peak S/N of 4.8 $\sigma$ . The lower set of panels in Figure 6 show binned red and blueshifted emission, highlighting the disk rotation. Rotation about the principal axis is seen for the stacked emission, and hinted at for two of the individual transitions.

Each of the three filter kernels from Figure 2 were cross-correlated with the visibilities of each of the observed CH<sub>3</sub>OH transitions, producing the filter responses shown in Figure 7. All three filters detect the individual lines and show similar S/N boosts, demonstrating that the method is robust to the choice of filter. The H<sub>2</sub>CO filter yields the strongest responses for the individual lines (4.3, 6.0, and 3.4 $\sigma$ , respectively). Stacking these spectra together, the H<sub>2</sub>CO filter yields a robust detection of 7.8 $\sigma$ , a 53% improvement over the 5.1 $\sigma$  detection reported

in Walsh et al. (2016). The Keplerian and CH<sub>3</sub>OH model filters produce stacked responses of 7.4 $\sigma$  and 7.7 $\sigma$ , respectively.

## 4. Discussion

We have presented a formulation of matched filtering for interferometric spectral line data and shown that this technique can improve S/N and therefore line detectability in both synthetic and real test cases. We now discuss how much of an S/N boost one might expect for a given data set, compare to alternative techniques, and suggest potential further applications of this method.

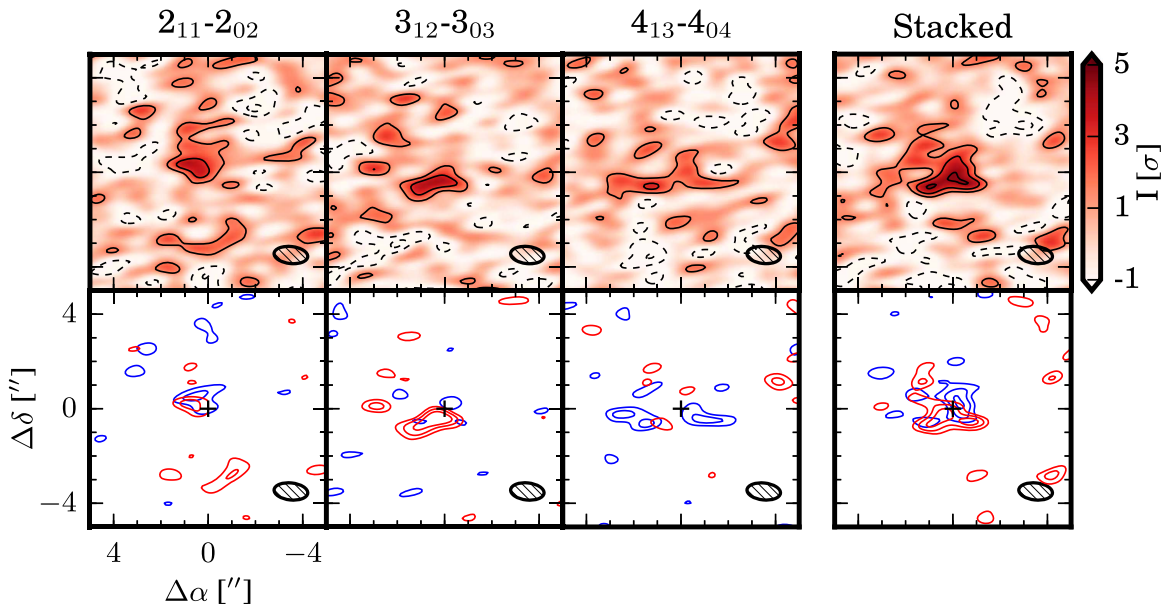
### 4.1. Factors Affecting S/N Boost

Compared with traditional line detection methods, the matched filter approach offers an improved S/N. The degree of S/N boost depends on both the accuracy of the approximated kernel as well as the specific properties of the data (particularly the spatial resolution). In the synthetic and real data test cases presented in Section 3, we found that application of a matched filter could increase S/N by  $\sim 60\%$ . The method was found to be similarly effective when applied to real data ( $\sim 53\%$  versus  $\sim 60\%$  boost), demonstrating that it is robust to the choice of approximated filter.

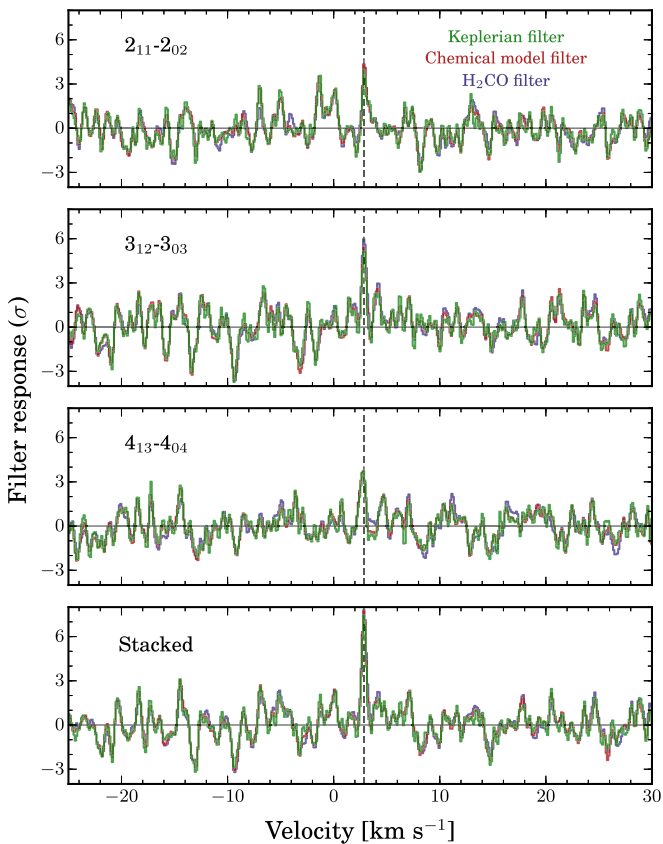
Intuitively, the S/N boost and spatial resolution of the emission should be coupled. By definition, a spatially unresolved signal encodes no spatio-kinematic information, and in this limit the matched filter technique will provide no increase in S/N other than the boost from spectral averaging. As emission is spatially resolved, S/N will decrease roughly with the square of the degree of spatial resolution (source width/beam size), with additional losses due to spatial filtering (see, e.g., Crane & Napier 1986). With appropriate knowledge of the velocity structure, a matched filter essentially negates this effect, and thus the S/N boost scales directly with the spatial resolution of the signal (see Yen et al. 2016 for a detailed image-plane derivation of this S/N boost). Figure 8 illustrates this effect, with an identical simulation to that shown in Figure 4, but with a higher spatial resolution of  $\sim 0''.3$ . The data were noise corrupted to reach a similar  $\sim 4\sigma$  detection in the moment map, although the S/N in the extracted spectrum is now  $\sim 6\sigma$ , highlighting how ineffective moment maps are at high spatial resolutions. The filter response is also now much larger (S/N = 23.6 $\sigma$ ), yielding an S/N boost over the aperture extracted spectra of  $\sim 400\%$ , compared to  $\sim 60\%$  in Figure 4. Similarly, application of matched filtering to higher-resolution observations of H<sub>2</sub>CO (Carney et al. 2017) produced an S/N gain of over 500%, confirming in practice the relationship between S/N boost and spatial resolution.

### 4.2. Comparison to Other Methods

Recently Matrà et al. (2015) and Marino et al. (2016) introduced an image-plane line detection technique (also independently introduced and formalized by Yen et al. 2016) that provides some similar benefits to the matched filter technique. In their approaches, pixels from a dirty image are adjusted for an assumed velocity offset (from a source model), and the velocity-corrected spectra are then stacked. In many ways, this can be seen as an image-plane analog to a Fourier plane matched filter, and it should yield comparable increases in S/N (see Appendix B for more details). This is confirmed by comparing the results of the matched filtering technique to



**Figure 6.** CH<sub>3</sub>OH observations toward TW Hya. Top: CH<sub>3</sub>OH emission from the 2<sub>11</sub>-2<sub>02</sub>, 3<sub>12</sub>-3<sub>03</sub>, and 4<sub>13</sub>-4<sub>04</sub> transitions, and all three stacked. Contours are  $[-3, -1.5, 1.5, 3, 4.5] \times \sigma$ ,  $\sigma = \sim 3.6 \text{ mJy } \text{bm}^{-1} \text{ km s}^{-1}$  for the individual transitions and  $\sim 2.3 \text{ mJy } \text{bm}^{-1} \text{ km s}^{-1}$  for the stacked image. Bottom: same as top, but for 1  $\text{km s}^{-1}$  velocity bins around the source velocity, showing the disk rotation.



**Figure 7.** Filter response spectra for each CH<sub>3</sub>OH transition. The impulse responses to the Keplerian, CH<sub>3</sub>OH model, and H<sub>2</sub>CO filters are shown in green, red, and blue, respectively.

detect H<sub>2</sub>CO in HD 163296 (Carney et al. 2017) with those obtained on the same data set by Yen et al. (2016). In both cases, an S/N boost of  $\sim 500\%$  is achieved.

Several subtle differences between matched filtering and pixel stacking, however, may motivate their use in a synergistic

fashion. First, application of a matched filter in the  $(u, v)$ -plane requires no imaging of the data, and is therefore much faster and more robust than image-plane spectral stacking. Second, the matched filter technique allows for a more accurate emission model than simple Keplerian rotation to be applied to the data (i.e., the spatial distribution of molecular signal can be properly used for weighting). On the other hand, stacking pixels in the image plane makes extracting a flux measurement for the line much simpler and allows for a radial profile to be estimated (Yen et al. 2016). The two techniques could therefore be used sequentially to exploit the benefits of each, with a matched filter first used to quickly identify and confirm line detections in a data set and pixel stacking then used to better characterize the lines.

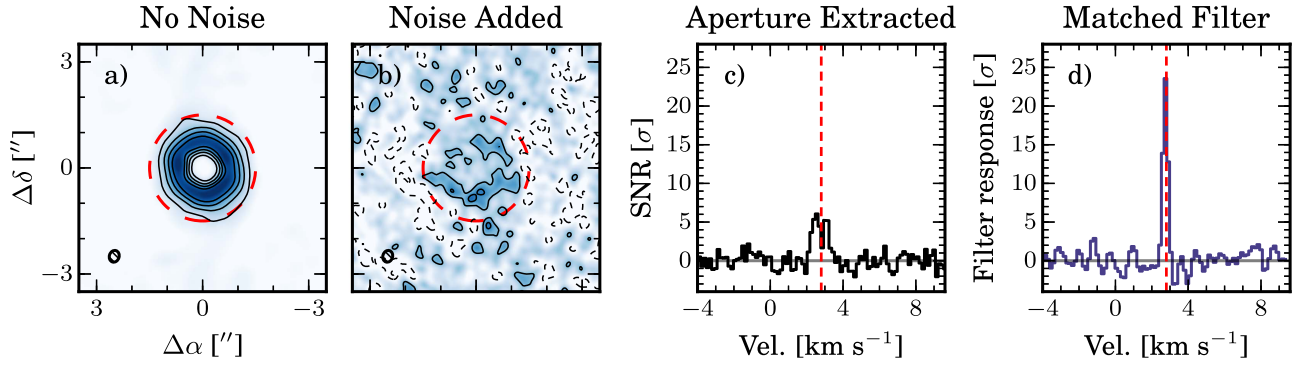
#### 4.3. Line Flux Estimation

The main utility of matched filtering when applied to interferometric spectral line data is in the rapid detection of weak lines, rather than their detailed characterization. Once a line is identified, it might be further characterized through careful imaging, spectral stacking, or model-fitting to the visibilities. For the weakest lines, however, detailed characterization will likely require additional observations. Matched filtering provides useful predictive utility when planning these observations, robustly confirming weak lines which might be desirable targets.

In particular, after a weak line is identified, the matched filter method can be used to estimate a line flux if the emission is too weak to be seen directly in the image cube. When a data-driven approach is taken, the responses of the target and template lines to the filter can be compared. If the two lines have a similar emission morphology, the ratio of their responses will be similar to the ratio of their fluxes, with the flux of the strong line being easy to measure.

This can be proven by considering a modified version of Equation (1), writing down the S/N using the signal/rms





**Figure 8.** Comparison of matched filtering with traditional methods, as in Figure 4 but at higher spatial resolution. Panel (a): moment-0 map of simulated noiseless  $\text{CH}_3\text{OH } 3_{12-3_{03}}$  emission. The synthesized beam is shown in the lower left. Contours are  $[-3, -1.5, 1.5, 3] \times 5.6 \text{ mJy } \text{bm}^{-1} \text{km s}^{-1}$ , corresponding to  $1\sigma$  in panel (b). Panel (b): moment-0 map of simulated and noise-corrupted  $\text{CH}_3\text{OH}$  emission. Panel (c): spectrum of the noise-corrupted emission, extracted using an aperture  $3''$  in diameter. Panel (d): ideal matched filter response to the noisy emission. Units are  $\sigma$ , defined as the rms filter response in channels with no emission.

definition:

$$S/N_s = \sqrt{\frac{\mathbf{h}^* \mathbf{s} \mathbf{s}^* \mathbf{h}}{\mathbf{h}^* \mathbf{R}_v \mathbf{h}}} \quad (9)$$

We can treat the two lines as signals  $\mathbf{y}$  and  $\mathbf{s}$ , where  $\mathbf{y}$  differs only from  $\mathbf{s}$  by an arbitrary constant  $\alpha$ , i.e.,  $\mathbf{y} = \alpha \mathbf{s}$ . The  $S/N$  of  $\mathbf{y}$  after filter application is:

$$S/N_y = \sqrt{\frac{\mathbf{h}^* \alpha \mathbf{s} \alpha \mathbf{s}^* \mathbf{h}}{\mathbf{h}^* \mathbf{R}_v \mathbf{h}}} \quad (10)$$

which reduces to:

$$S/N_y = \alpha \sqrt{\frac{\mathbf{h}^* \mathbf{s} \mathbf{s}^* \mathbf{h}}{\mathbf{h}^* \mathbf{R}_v \mathbf{h}}} \quad (11)$$

$$S/N_y = \alpha S/N_s. \quad (12)$$

Therefore we can use the filter impulse response ratio to roughly estimate the flux, with the accuracy dependent on the closeness of the filter kernel to the true emission distribution. It is important to note that this estimate will always be a lower limit. An upper limit can additionally be derived from the imaged weak line, bounding the flux measurement. This approach was used by Carney et al. (2017) to determine the flux ratio of multiple detected  $\text{H}_2\text{CO}$  lines, enabling them to constrain the  $\text{H}_2\text{CO}$  excitation temperature.

#### 4.4. Application to Line Surveys

In addition to aiding the detection of specific known weak lines, interferometric matched filtering provides substantial benefits for the processing of spectral line surveys where source locations and approximate spatio-kinematic structures are known. Imaging the full bandwidth of these large data sets at their native spectral resolution is a time-consuming process, often taking many hours or even days. Because much of the information in these data sets is contained in spatio-kinematic patterns of the spectral lines, decreasing spectral resolution through channel averaging is typically not a viable option and can result in signal loss. A choice must therefore be made between imaging only small targeted windows of the broadband data set, or spending time and computing resources on imaging the full bandwidth. For sparsely populated line surveys (e.g., of protoplanetary disks), imaging the entire data set is inherently inefficient, since most of the channels do not contain

signal. Conversely, selective imaging reduces the likelihood of serendipitously detecting weak species, and conflicts with the motivations of an unbiased survey.

Numerous tools have been developed to aid in identifying spectral line emission in broadband data sets from current and future instruments such as ALMA, ASKAP, VLA, SKA, and the ngVLA (e.g., Koribalski 2012; Whiting 2012; Whiting & Humphreys 2012; Friedel et al. 2015; Serra et al. 2015), but these methods often rely on a fully imaged datacube as input. In instances where the locations of the sources being targeted are known a priori, our described method of matched filtering can help streamline this process by quickly and robustly identifying lines in the native visibilities. Then only these lines need be imaged and analyzed. In sources with a single dominant velocity pattern, a strong line could be imaged, converted to a filter kernel, and cross-correlated through the entire data set in a small fraction of the time it would take to image that same data set. The resulting full-band impulse response spectrum then provides a convenient first look at the data set, guiding the observer as to which sections of the data are worth windowing out for further imaging and analysis. In particular, matched filtering will highlight weak lines that the observer would have missed even in a careful CLEAN of the data.

We note that our  $(u, v)$ -plane method could likely be extended to full 3D searches in blind surveys, where source locations are not known a priori, but such an implementation is beyond the scope of this paper. It is not immediately clear whether the large speed benefits of  $(u, v)$ -plane analysis over full survey imaging would be maintained in such a 3D search space, and we encourage further research in this area.

## 5. Conclusion

We have shown that the technique of matched filtering can be easily implemented for analyzing interferometric observations of spectral lines, significantly improving sensitivity when searching for weak lines. An open-source Python-based implementation is freely available under the MIT license at <https://github.com/AstroChem/VISIBLE>. As a case study, we have focused on observations of protoplanetary disks with ALMA, but our approach is applicable to spectral line data of any astronomical source with a spatio-kinematic pattern that can be used to generate a filter kernel, and will likely be beneficial for spectral line observations from a wide range of current and future instruments (e.g., the SKA and the ngVLA).

We find that, when applied to real data, the method results in large sensitivity increases, ranging from 53% for CH<sub>3</sub>OH in Walsh et al. (2016), to ~500% for H<sub>2</sub>CO in Carney et al. (2017). The degree of sensitivity boost is proportional to the spatial resolution of the observations. These sensitivity increases are equivalent to factors of 2–25 in effective observing time, allowing observers to better leverage limited telescope resources. Additionally, the speed of the technique is beneficial when analyzing large-bandwidth line surveys, robustly identifying all lines in a spectrum in a small fraction of the time it would take to image the same data set. Finally, the method works synergistically with the methods presented in Matrà et al. (2015) and Yen et al. (2016) and tools such as ADMIT, forming a comprehensive suite of analysis techniques for spectral lines in large interferometric data sets.

We would like to thank Andrew Vanderburg and Peter Williams for productive discussions and assistance with software package management and distribution. We also thank the anonymous referee for providing comments that greatly improved the quality of the manuscript. R.A.L. and J.H. gratefully acknowledge funding from National Science Foundation Graduate Research Fellowships (Grant No. DGE-1144152). R.A.L. also acknowledges funding from the NRAO Student Observing Support Program. K.I.Ö. acknowledges funding from the Alfred P. Sloan Foundation and the David and Lucile Packard Foundation. C.W. acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO, grant 639.041.335) and start-up funds from the University of Leeds, UK. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2013.1.00902.S and ADS/JAO.ALMA#2013.1.00114.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

*Software:* Astropy (Astropy Collaboration et al. 2013), CASA (McMullin et al. 2007), casa-python, DiskJockey (Czekala et al. 2015; Czekala 2016), LIME (Brinch & Hogerheijde 2010), Matplotlib (Hunter 2007), NumPy (Jones et al. 2001), SciPy (van der Walt et al. 2011), VISIBLE (Loomis et al. 2018a, 2018b).

### Appendix A Calculating S/N Boost for a Matched Filter

S/N (using the definition of signal-power/noise-power) can be written for an arbitrary signal  $s$  and filter  $h$  as:

$$S/N = \frac{h^* s s^* h}{h^* R_v h}. \quad (13)$$

As discussed in Section 2.1, North (1963) showed that a linear matched filter of form:

$$h = \left[ \frac{1}{\sqrt{s^* R_v^{-1} s}} \right] R_v^{-1} s = C R_v^{-1} s, \quad (14)$$

maximizes the output S/N. The natural question is then: how much is the S/N boosted by applying such a filter? This can be

analytically calculated for a given signal by comparing the S/N after applying a matched filter with the S/N from applying a flat filter 1 (e.g., a unity matrix of all ones). We start by calculating the S/N after applying the matched filter:

$$S/N_{\text{mf}} = \frac{h^* s s^* h}{h^* R_v h}. \quad (15)$$

We then substitute for  $h$ , noting that that  $R_v^{-1}$  is Hermitian and therefore  $R_v^{-1*} = R_v^{-1}$ :

$$S/N_{\text{mf}} = \frac{R_v^{-1} s^* s s^* R_v^{-1} s}{R_v^{-1} s^* R_v R_v^{-1} s}. \quad (16)$$

Under the assumption of uncorrelated noise (which is reasonable for the case of independent interferometric visibilities), there are no off-diagonal terms in  $R_v$  and we can reduce this equation to:

$$S/N_{\text{mf}} = \sum_i^N \|s_i\|^2 R_{ii}^{-1}, \quad (17)$$

where there are  $N$  elements of the signal  $s$  (which can be summed in multiple dimensions or flattened as shown here). Similarly, if we write the S/N of the flat filter as:

$$S/N_{\text{flat}} = \frac{1^* s s^* 1}{1^* R_v 1}, \quad (18)$$

then we find it reduces to:

$$S/N_{\text{flat}} = \frac{\left( \sum_i^N \|s_i\| \right)^2}{\text{tr}[R_v]}. \quad (19)$$

So the ratio of these two S/Ns, or the total S/N boost from a matched filter, is:

$$\text{boost} = \frac{S/N_{\text{mf}}}{S/N_{\text{flat}}} = \frac{\left( \sum_i^N \|s_i\|^2 R_{ii}^{-1} \right) \text{tr}[R_v]}{\left( \sum_i^N \|s_i\| \right)^2}. \quad (20)$$

### Appendix B Calculating S/N Boost in Comparison to Image-plane Measurements

The boost value in Equation (20) can be analytically calculated for a given filter kernel, but is not particularly useful at this point as it has not been related to the image-plane S/Ns discussed throughout the paper. Thus the fundamental problem is how to relate the visibilities to an image-plane S/N. We start by writing down the definition of S/N in the dirty image  $I^D$ , or the raw discrete Fourier transform of the visibilities (i.e., not deconvolved):

$$S/N = \frac{I^D}{\Delta I^D} = \frac{\sum_k W_k V_k}{\left( \sum_k W_k^2 \sigma_k^2 \right)^{1/2}}, \quad (21)$$

where there are  $k$  visibilities in the data set, each with a source visibility contribution  $V_k$ , total weight  $W_k$  (including any taper weights, density weights, and the variance weights  $w_k$  discussed in the main text), and thermal noise  $\sigma_k$ . Notation is borrowed from Briggs (1995), which contains a detailed

discussion of image-plane S/Ns and their relation to the measured visibilities. If the data are gridded into cells  $\{p, q\}$  and the noise properties of all visibilities within each cell are similar, an approximate form can be written:

$$S/N = \frac{\sum_{p,q} W_{pq} V_{pq}}{\left(\sum_{p,q} W_{pq}^2 \sigma_{pq}^2\right)^{1/2}}. \quad (22)$$

In particular, we are interested in the S/N at a particular location in the dirty map, e.g., the peak pixel in a given channel. If this pixel is phase shifted to the map center, the S/N can be written as:

$$S/N'(0, 0) = \frac{\sum_{p,q} W_{pq} |V_{pq}|}{\left(\sum_{p,q} W_{pq}^2 \sigma_{pq}^2\right)^{1/2}}. \quad (23)$$

If we consider a moment-0 map of a resolved source, however, the S/N at the center of the moment map is:

$$S/N_{\text{mom0}} = \frac{\sum_{c,p,q} W_{cpq} \text{Re}(V_{cpq})}{\left(\sum_{c,p,q} W_{cpq}^2 \sigma_{cpq}^2\right)^{1/2}}, \quad (24)$$

and only the projected real component of each visibility will contribute signal. The S/N will then decrease as the ratio of the emission size to the resolution element increases, as discussed in Crane & Napier (1986) and Yen et al. (2016). Compounding this issue, if the source has a strong spatio-kinematic signature and peak emission moves throughout the dirty map, then the projected real component will vary as a function of channel  $c$ . Applying these properties to Equation (20), we can estimate the S/N boost of the matched filter over a peak moment-0 value as:

$$\text{boost} = \frac{S/N_{\text{mf}}}{S/N_{\text{mom0}}} = \frac{\left(\sum_i^N \|V_i\|^2 R_{ii}^{-1}\right) \text{tr}[\mathbf{R}_v]}{\left(\sum_i^N \text{Re}(V_i)\right)^2}. \quad (25)$$

Aligning the signal in the image plane through pixel shifting and stacking (e.g., as in Matrà et al. 2015 or Yen et al. 2016) is analogous to phase shifting the individual visibilities to the map center.  $\text{Re}(V_{pq})$  can then be replaced by  $|V_{pq}|$ , and the S/N after applying a pixel shifting method is roughly:

$$S/N_{\text{ps}} = S/N'_{\text{mom0}} = \frac{\sum_{c,p,q} W_{cpq} |V_{cpq}|}{\left(\sum_{c,p,q} W_{cpq}^2 \sigma_{cpq}^2\right)^{1/2}}. \quad (26)$$

Returning to Equation (20) and applying this logic, we can write the boost as:

$$\text{boost} = \frac{S/N_{\text{mf}}}{S/N_{\text{ps}}} = \frac{\left(\sum_i^N \|V_i\|^2 R_{ii}^{-1}\right) \text{tr}[\mathbf{R}_v]}{\left(\sum_i^N |V_i|\right)^2}, \quad (27)$$

which defines the additional benefit matched filtering provides over a pixel stacking approach.

Equations (24) and (26) can be applied to any visibility sampled filter kernel to calculate these boosts. We note, however, that due to the line-broadening of most astronomical signals, true phase alignment from a pixel shifting approach is

not possible and therefore the boost formulas are only approximations.

## Appendix C

### Data Weights and Noise Covariance Matrices

#### C.1. Interferometric Data Weights

In general, each visibility  $V_i$  in an interferometric data set corresponding to an antenna pair  $(m, n)$  will have a characteristic variance  $\sigma_{mn}^2$ , which is often assumed to be dominated by the system noise (see e.g., Chapter 6 of Thompson et al. 2017). When the system noise dominates,  $\sigma_{mn}$  can be written in units of Jy as:

$$\sigma_{mn}(\text{Jy}) = \frac{2k}{n_q n_c A_{\text{eff}}} \sqrt{\frac{T_{\text{sys},m} T_{\text{sys},n}}{2\Delta\nu\Delta t}} \times 10^{26}, \quad (28)$$

where  $k$  is Boltzmann's constant,  $n_q$  and  $n_c$  are the quantization and correlator efficiencies,  $A_{\text{eff}}$  is the effective antenna area,  $T_{\text{sys},m}$  and  $T_{\text{sys},n}$  are system temperatures for antennas  $m$  and  $n$ , respectively,  $\Delta\nu$  is the effective channel bandwidth, and  $\Delta t$  is the integration time. For identical antennas and system temperatures, this can be simplified as:

$$\sigma_{mn}(\text{Jy}) = \frac{2kT_{\text{sys}}}{n_q n_c A_{\text{eff}}} \sqrt{\frac{1}{2\Delta\nu\Delta t}} \times 10^{26}. \quad (29)$$

The data weight for each visibility is then calculated as  $w_i = 1/\sigma_i^2$ . For instruments which record channelized system temperatures (e.g., ALMA), the weights will also be channelized, and are recorded in CASA as a “weight spectrum” for each visibility.<sup>10</sup>

#### C.2. Noise Covariance Matrices

As discussed in Section 2.1, a matched filter can be written as:

$$\mathbf{h} = \left[ \frac{1}{\sqrt{\mathbf{s}^* \mathbf{R}_v^{-1} \mathbf{s}}} \right] \mathbf{R}_v^{-1} \mathbf{s} = \mathbf{C} \mathbf{R}_v^{-1} \mathbf{s}, \quad (30)$$

where  $\mathbf{R}_v$  is the noise covariance matrix. When the  $n_c$  channels in an interferometric data set are fully independent,  $\mathbf{R}_v$  can be written for an individual visibility  $V_i$  as a  $n_c \times n_c$  diagonal matrix:

$$\mathbf{R}_v = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_{n_c}^2 \end{bmatrix} \quad (31)$$

<sup>10</sup> For more details on how weights are handled in different versions of CASA, see <https://casa.nrao.edu/casadocs/casa-5.1.0/reference-material/data-weights>.

and  $\mathbf{R}_v^{-1}$  is then:

$$\mathbf{R}_v^{-1} = \begin{bmatrix} \frac{1}{\sigma_1^2} & & & \\ & \frac{1}{\sigma_2^2} & & \\ & & \ddots & \\ & & & \frac{1}{\sigma_{n_c}^2} \end{bmatrix} \quad (32)$$

i.e., a diagonal matrix initialized to the channelized data weights. When the weights are not channelized, or can be approximated as equal across channels ( $w_j \approx w_i \forall j$ ),  $\mathbf{R}_v^{-1}$  can be written as:

$$\mathbf{R}_v^{-1} = w_i \mathbf{I}_{n_c}, \quad (33)$$

where  $\mathbf{I}_{n_c}$  is an  $n_c \times n_c$  identity matrix.

### C.2.1. Correlated Channels

In the case of fully independent channels, the filter and normalization prefactor are simple to calculate due to the lack of off-diagonal elements in the noise covariance matrix. In practice, however, channels are often correlated. In particular, astronomical interferometric data sets are generally correlated due to the use of a window function (e.g., Hann, Hamming, etc.) to reduce the ringing effect introduced by the finite maximum lag time of the correlator hardware.<sup>11</sup> As the Hann function is a popular choice of window function (and applied directly in the time domain for the ALMA data described in this paper), we derive here the appropriate noise covariances matrices for Hanning-smoothed data. Similar results could be calculated for other choices of window function.

In the frequency/channel domain, the Hann window applied to an observation  $\mathbf{x}$  can be written as:

$$x'_i = \frac{1}{4}x_{i-1} + \frac{1}{2}x_i + \frac{1}{4}x_{i+1}. \quad (34)$$

For an observation which can be linearly decomposed into signal and additive white Gaussian noise ( $\mathbf{x} = \mathbf{s} + \mathbf{v}$ ), we can calculate the noise covariance matrix of the smoothed data,  $\mathbf{R}'_v$ , which now contains off-diagonal elements:

$$R'_v[j, k] = E[v'_j v'^*_k], \quad (35)$$

where

$$v'_j = \frac{1}{4}v_{j-1} + \frac{1}{2}v_j + \frac{1}{4}v_{j+1}. \quad (36)$$

Noting that in the uncorrelated case

$$E[v_j v_k^*] = \begin{cases} \sigma_j^2, & \text{if } j = k \\ 0, & \text{otherwise} \end{cases} \quad (37)$$

we find that for diagonal elements of  $R'_v$ , Equation (35) reduces to

$$R'_v[j, j] = \frac{\sigma_{j-1}^2}{16} + \frac{\sigma_j^2}{4} + \frac{\sigma_{j+1}^2}{16}, \quad (38)$$

and for the populated off-diagonal elements it reduces to

$$R'_v[j, j \pm 1] = \frac{\sigma_j^2}{8} + \frac{\sigma_{j \pm 1}^2}{8}, \quad (39)$$

and

$$R'_v[j, j \pm 2] = \frac{\sigma_{j \pm 1}^2}{16}. \quad (40)$$

Under the assumption that the channelized weights are all approximately equal for a given visibility ( $w_j \approx w_i \forall j$ ),  $\mathbf{R}'_v$  can be written as

$$\begin{aligned} \mathbf{R}'_v &\approx \sigma_i^2 \begin{bmatrix} \frac{3}{8} & \frac{1}{4} & \frac{1}{16} & & & \\ \frac{1}{4} & \frac{3}{8} & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \ddots & \\ & & & \frac{3}{8} & \frac{1}{4} & \\ & & & \frac{1}{16} & \frac{1}{4} & \frac{3}{8} \end{bmatrix} \\ &= \sigma_i'^2 \begin{bmatrix} 1 & \frac{2}{3} & \frac{1}{6} & & & \\ \frac{2}{3} & 1 & \ddots & \ddots & & \\ \frac{1}{6} & \ddots & \ddots & \ddots & \ddots & \\ & \ddots & \ddots & 1 & \frac{2}{3} & \\ & & & \frac{1}{6} & \frac{2}{3} & 1 \end{bmatrix} = \sigma_i'^2 [\mathbf{M}] \end{aligned} \quad (41)$$

where  $\sigma'_i = \sqrt{3/8} \sigma_i$ . As tasks such as `statwt` in CASA do not consider effective channel bandwidth,  $w'_i = 1/\sigma_i'^2$  is likely what will be reported as the data weights of the observations.<sup>12</sup> The assumption of equal weights across channels is reasonable when  $T_{\text{sys}}$  is stable across channels, which will likely be true unless there were issues with data calibration or there are strong water lines in the data. Under this assumption, the matrix inversion only needs to be computed once:

$$\mathbf{R}'_v{}^{-1} \approx \frac{1}{\sigma_i'^2} [\mathbf{M}]^{-1} = w'_i [\mathbf{M}]^{-1}. \quad (42)$$

When the data are initially Hanning smoothed, the edge channels are clipped. Thus for the purposes of preventing boundary effects during inversion, the covariance matrix can be treated as describing an infinitely wide data set. In practice, we extend the matrix to be several times larger than  $n_c$ , and then window a  $n_k$  sized portion from the center of  $\mathbf{R}'_v{}^{-1}$ . In the unbinned case, however,  $\mathbf{R}'_v$  is an ill-conditioned matrix (the condition number is  $>10^{10}$  for a typical ALMA spectral window), making the matrix inversion numerically unstable. The issue can be avoided by channel binning the data by a factor of two, as discussed in the following section.

### C.2.2. Averaged Correlated Channels

As Hanning smoothing inflates the effective channel width by a factor of 8/3, it is very common to bin data across

<sup>11</sup> A full description of these effects and the choice of window function for ALMA can be found at [https://safe.nrao.edu/wiki/pub/Main/ALMAWindowFunctions/Note\\_on\\_Spectral\\_Response\\_V2.pdf](https://safe.nrao.edu/wiki/pub/Main/ALMAWindowFunctions/Note_on_Spectral_Response_V2.pdf).

<sup>12</sup> See <https://casaguides.nrao.edu/index.php/DataWeightsAndCombination> for details about the absolute accuracy of data weights in CASA.

channels by a factor of 2; this is the default for ALMA observations. Here we calculate the appropriate covariance matrices for binning factors of 2, 3, and 4. A similar method can be followed for other binning factors, but we note that past a binning factor of 4, the covariance matrix and its inverse rapidly approach the uncorrelated case and the channel correlation can likely be neglected without significant adverse effect.

With a binning factor of 2 applied to data that have been Hanning smoothed, the data (indexed by channel in the averaged data set) can be described as:

$$\begin{aligned} x'_{i,\text{bin}\times 2} &= \frac{\left(\frac{1}{4}x_{i-0.5} + \frac{1}{2}x_i + \frac{1}{4}x_{i+0.5}\right) + \left(\frac{1}{4}x_i + \frac{1}{2}x_{i+0.5} + \frac{1}{4}x_{i+1}\right)}{2} \\ &= \frac{1}{8}x_{i-0.5} + \frac{3}{8}x_i + \frac{3}{8}x_{i+0.5} + \frac{1}{8}x_{i+1}, \end{aligned} \quad (43)$$

and the noise component is therefore:

$$v'_{i,\text{bin}\times 2} = \frac{1}{8}v_{i-0.5} + \frac{3}{8}v_i + \frac{3}{8}v_{i+0.5} + \frac{1}{8}v_{i+1}. \quad (44)$$

Substituting this into Equation (35) and applying Equation (37), we find that for diagonal elements of  $\mathbf{R}'_v$

$$R'_{v,\text{bin}\times 2}[j, j] = \frac{\sigma_{j-0.5}^2}{64} + \frac{9\sigma_j^2}{64} + \frac{9\sigma_{j+0.5}^2}{64} + \frac{\sigma_{j+1}^2}{64}, \quad (45)$$

and the only populated off-diagonal elements are

$$R'_{v,\text{bin}\times 2}[j, j \pm 1] = \frac{3\sigma_{j\pm 0.5}^2}{64} + \frac{3\sigma_{j\pm 1}^2}{64}. \quad (46)$$

Assuming the weights are roughly equivalent across channels for a given visibility,  $\mathbf{R}'_v$  can be approximated as

$$\begin{aligned} \mathbf{R}'_{v,\text{bin}\times 2} &\approx \sigma_i^2 \begin{bmatrix} \frac{5}{16} & \frac{3}{32} & & & \\ \frac{3}{32} & \frac{5}{16} & \ddots & & \\ & \ddots & \ddots & \frac{3}{32} & \\ & & & \frac{3}{32} & \frac{5}{16} \end{bmatrix} \\ &= \sigma_i'^2 \begin{bmatrix} 1 & \frac{3}{10} & & & \\ \frac{3}{10} & 1 & \ddots & & \\ & \ddots & \ddots & \frac{3}{10} & \\ & & & \frac{3}{10} & 1 \end{bmatrix} = \sigma_i'^2 [\mathbf{M}_{\text{bin}\times 2}] \end{aligned} \quad (47)$$

where  $\sigma'_i = \sqrt{5/16}\sigma_i$ . Correspondingly,

$$\mathbf{R}'_{v,\text{bin}\times 2}{}^{-1} \approx \frac{1}{\sigma_i'^2} [\mathbf{M}_{\text{bin}\times 2}]^{-1} = w'_i [\mathbf{M}_{\text{bin}\times 2}]^{-1}. \quad (48)$$

where the weights  $w'_i$  were calculated from the binned data using a task such as `statwt`. In contrast to the unbinned case,  $\mathbf{R}'_v$  is now tridiagonal and well-conditioned for inversion.

Repeating these calculations for binning factors of 3 and 4, we find that:

$$\begin{aligned} \mathbf{R}'_{v,\text{bin}\times 3} &\approx \sigma_i^2 \begin{bmatrix} \frac{1}{4} & \frac{1}{24} & & & \\ \frac{1}{24} & \frac{1}{4} & \ddots & & \\ & \ddots & \ddots & \frac{1}{24} & \\ & & & \frac{1}{24} & \frac{1}{4} \end{bmatrix} \\ &= \sigma_i'^2 \begin{bmatrix} 1 & \frac{1}{6} & & & \\ \frac{1}{6} & 1 & \ddots & & \\ & \ddots & \ddots & \frac{1}{6} & \\ & & & \frac{1}{6} & 1 \end{bmatrix} = \sigma_i'^2 [\mathbf{M}_{\text{bin}\times 3}] \end{aligned} \quad (49)$$

where  $\sigma'_i = \sqrt{1/4}\sigma_i$ ,

$$\mathbf{R}'_{v,\text{bin}\times 3}{}^{-1} \approx \frac{1}{\sigma_i'^2} [\mathbf{M}_{\text{bin}\times 3}]^{-1} = w'_i [\mathbf{M}_{\text{bin}\times 3}]^{-1}, \quad (50)$$

and

$$\begin{aligned} \mathbf{R}'_{v,\text{bin}\times 4} &\approx \sigma_i^2 \begin{bmatrix} \frac{13}{64} & \frac{3}{128} & & & \\ \frac{3}{128} & \frac{13}{64} & \ddots & & \\ & \ddots & \ddots & \frac{3}{128} & \\ & & & \frac{3}{128} & \frac{13}{64} \end{bmatrix} \\ &= \sigma_i'^2 \begin{bmatrix} 1 & \frac{3}{26} & & & \\ \frac{3}{26} & 1 & \ddots & & \\ & \ddots & \ddots & \frac{3}{26} & \\ & & & \frac{3}{26} & 1 \end{bmatrix} = \sigma_i'^2 [\mathbf{M}_{\text{bin}\times 4}] \end{aligned} \quad (51)$$

where  $\sigma'_i = \sqrt{13/64}\sigma_i$ ,

$$\mathbf{R}'_{v,\text{bin}\times 4}{}^{-1} \approx \frac{1}{\sigma_i'^2} [\mathbf{M}_{\text{bin}\times 4}]^{-1} = w'_i [\mathbf{M}_{\text{bin}\times 4}]^{-1}. \quad (52)$$

## Appendix D Filter Kernel Generation

The Keplerian filter kernel presented in Figure 2 was calculated for the viewing geometry of the protoplanetary disk around T Tauri star TW Hya (with an inclination of  $7^\circ$  and PA of  $155^\circ$ ; e.g., Qi et al. 2004; Andrews et al. 2012). The Keplerian velocity field is calculated as

$$v = \sqrt{\frac{GM_*}{r}}, \quad (53)$$

with an assumed stellar mass of  $0.8 M_\odot$  (Hughes et al. 2011). Using this field, we computed the emitting region of the disk for channels with  $0.2 \text{ km s}^{-1}$  spacing.

The parametric model filter kernel was calculated from the ‘‘fiducial’’ model in Walsh et al. (2016), where  $\text{CH}_3\text{OH}$  around TW Hya is constrained to a vertical layer  $z/r < 0.1$  between radii of 30 and 100 au. From this abundance structure, an emission profile was calculated for the  $3_{12-3_{03}}$  transition using

LIME. As seen in Figure 2, the emission from this model tapers radially due to decreasing column density and temperature, in contrast to the Keplerian mask which simply has a hard outer radius cutoff. Additionally, the CH<sub>3</sub>OH model has an inner disk depletion (as CH<sub>3</sub>OH is mainly formed through hydrogenation of CO on grain surfaces outside the CO snowline), which is not present in the simple Keplerian model.

Finally, the data-driven filter kernel was generated from observations of H<sub>2</sub>CO around TW Hya (Öberg et al. 2017). The data were imaged in CASA using CLEAN with natural weighting, yielding a high-S/N image cube. After imaging, all noise below  $3\sigma$  and any emission outside of a  $3''$  radius were masked out, creating a mostly noiseless approximation of the true H<sub>2</sub>CO emission distribution. ( $u$ ,  $v$ )-plane kernels were generated from each of these image cubes using `vis_sample`, as described in Sections 2.2, 2.3, and 2.4.

### ORCID iDs

Ryan A. Loomis  <https://orcid.org/0000-0002-8932-1219>  
 Karin I. Öberg  <https://orcid.org/0000-0001-8798-1347>  
 Sean M. Andrews  <https://orcid.org/0000-0003-2253-2270>  
 Catherine Walsh  <https://orcid.org/0000-0001-6078-786X>  
 Ian Czekala  <https://orcid.org/0000-0002-1483-8811>  
 Jane Huang  <https://orcid.org/0000-0001-6947-6072>

### References

- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *PhRvL*, **116**, 061102  
 Andrews, S. M., Wilner, D. J., Hughes, A. M., et al. 2012, *ApJ*, **744**, 162  
 Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33  
 Barenfeld, S. A., Carpenter, J. M., Ricci, L., & Isella, A. 2016, *ApJ*, **827**, 142  
 Bertin, E. 2001, in *Mining the Sky: Proc. of the MPA/ESO/MPE Workshop*, ed. A. J. Bandy, S. Zaroubi, & M. Bartelmann (Berlin: Springer-Verlag), 353  
 Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393  
 Briggs, D. S. 1995, AAS Meeting 187 Abstracts, 112.02  
 Brinch, C., & Hogerheijde, M. R. 2010, *A&A*, **523**, A25  
 Carney, M. T., Hogerheijde, M. R., Loomis, R. A., et al. 2017, arXiv:1705.10188  
 Coutens, A., Jørgensen, J. K., van der Wiel, M. H. D., et al. 2016, *A&A*, **590**, L6  
 Crane, P. C., & Napier, P. J. 1986, in *Synthesis imaging in radio astronomy, A collection of Lectures from the Third NRAO Synthesis Imaging Summer School*, ed. R. A. Perley, F. R. Schwab, & A. H. Bridle (San Francisco, CA: ASP), 87  
 Cumming, I. G., & Wong, F. H. 2005, *Digital Processing of Synthetic Aperture Radar Data: Algorithm and Implementation* (Boston, MA: Artech House)  
 Cuppen, H. M., van Dishoeck, E. F., Herbst, E., & Tielens, A. G. G. M. 2009, *A&A*, **508**, 275  
 Czekala, I. 2016, *DiskJockey: Protoplanetary Disk Modeling for Dynamical Mass Derivation*, Astrophysics Source Code Library, ascl:1603.011  
 Czekala, I., Andrews, S. M., Jensen, E. L. N., et al. 2015, *ApJ*, **806**, 154  
 Dullemond, C. P. 2012, *RADMC-3D: A Multi-purpose Radiative Transfer Tool*, Astrophysics Source Code Library, ascl:1202.015  
 Dutrey, A., Henning, T., Guilloteau, S., et al. 2007, *A&A*, **464**, 615  
 Feng, L., Vaulin, R., Hewitt, J. N., et al. 2017, *AJ*, **153**, 98  
 Friedel, D., Looney, L., Xu, L., et al. 2015, in *Admit: Alma Data Mining Toolkit*, 70th Int. Symp. Molecular Spectroscopy, **WII5**  
 Herenz, E. C., & Wisotzki, L. 2017, *A&A*, **602**, A111  
 Högbom, J. A. 1974, *A&AS*, **15**, 417  
 Hughes, A. M., Wilner, D. J., Andrews, S. M., Qi, C., & Hogerheijde, M. R. 2011, *ApJ*, **727**, 85  
 Hunter, J. D. 2007, *CSE*, **9**, 90  
 Jones, E., Oliphant, T., Peterson, P., et al. 2001, *SciPy: Open Source Scientific Tools for Python*, <http://www.scipy.org/>  
 Jørgensen, J. K., Bourke, T. L., Nguyen Luong, Q., & Takakuwa, S. 2011, *A&A*, **534**, A100  
 Kalenskii, S. V., & Johansson, L. E. B. 2010, *ARep*, **54**, 295  
 Koribalski, B. S. 2012, *PASA*, **29**, 359  
 Langston, G., & Turner, B. 2007, *ApJ*, **658**, 455  
 Loomis, R., Öberg, K., Andrews, S., et al. 2018a, *VISIBILITY: VISIBILITY Based Line Extraction*, v0.1.0, Zenodo, <https://doi.org/10.5281/zenodo.1174549>  
 Loomis, R., Öberg, K., Andrews, S., et al. 2018b, *VISIBILITY: VISIBILITY Based Line Extraction*, Astrophysics Source Code Library, ascl:1802.006  
 Loomis, R. A., Cleeves, L. I., Öberg, K. I., Guzman, V. V., & Andrews, S. M. 2015, *ApJL*, **809**, L25  
 Loomis, R. A., Öberg, K. I., Andrews, S. M., & MacGregor, M. A. 2017, *ApJ*, **840**, 23  
 Loomis, R. A., Shingledecker, C. N., Langston, G., et al. 2016, *MNRAS*, **463**, 4175  
 MacGregor, M. A., Lawler, S. M., Wilner, D. J., et al. 2016, *ApJ*, **828**, 113  
 Marino, S., Matrà, L., Stark, C., et al. 2016, *MNRAS*, **460**, 2933  
 Matrà, L., Panić, O., Wyatt, M. C., & Dent, W. R. F. 2015, *MNRAS*, **447**, 3936  
 McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in *ASP Conf. Ser. 376, Astronomical Data Analysis Software and Systems XVI*, ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco, CA: ASP), 127  
 Meillier, C., Chatelain, F., Michel, O., et al. 2016, *A&A*, **588**, A140  
 Müller, H. S. P., Belloche, A., Xu, L.-H., et al. 2016, *A&A*, **587**, A92  
 North, D. O. 1963, *Proceedings of the IEEE*, **51**, 1016  
 Öberg, K. I., Guzmán, V. V., Furuya, K., et al. 2015, *Natur*, **520**, 198  
 Öberg, K. I., Guzmán, V. V., Merchant, C. J., et al. 2017, *ApJ*, **839**, 43  
 Owen, B. J., & Sathyaprakash, B. S. 1999, *PhRvD*, **60**, 022002  
 Qi, C., Ho, P. T. P., Wilner, D. J., et al. 2004, *ApJL*, **616**, L11  
 Qi, C., Öberg, K. I., & Wilner, D. J. 2013, *ApJ*, **765**, 34  
 Ruffio, J.-B., Macintosh, B., Wang, J. J., et al. 2017, *ApJ*, **842**, 14  
 Schutz, B. F. 1999, *CQGra*, **16**, A131  
 Schwab, F. R. 1984, in *Indirect Imaging. Measurement and Processing for Indirect Imaging*, ed. J. A. Roberts (Cambridge: Cambridge Univ. Press), 333  
 Schwartz, M., & Shaw, L. 1975, *Signal Processing: Discrete Spectral Analysis, Detection, and Estimation* (New York: McGraw-Hill)  
 Serra, P., Westmeier, T., Giese, N., et al. 2015, *MNRAS*, **448**, 1922  
 Thompson, A. R., Moran, J. M., & Swenson, G. W., Jr. 2017, *Interferometry and Synthesis in Radio Astronomy* (3rd ed.; Berlin: Springer)  
 van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, *CSE*, **13**, 22  
 Vio, R., & Andreani, P. 2016, *A&A*, **589**, A20  
 Walsh, C., Loomis, R. A., Öberg, K. I., et al. 2016, *ApJL*, **823**, L10  
 Walsh, C., Millar, T. J., Nomura, H., et al. 2014, *A&A*, **563**, A33  
 White, M., & Padmanabhan, N. 2017, arXiv:1705.09669  
 Whiting, M., & Humphreys, B. 2012, *PASA*, **29**, 371  
 Whiting, M. T. 2012, *MNRAS*, **421**, 3242  
 Woodward, P. 1953, *Probability and Information Theory: With Applications to Radar* (Amsterdam: Elsevier)  
 Yen, H.-W., Koch, P. M., Liu, H. B., et al. 2016, *ApJ*, **832**, 204  
 Zackay, B., & Ofek, E. O. 2017, *ApJ*, **836**, 187