Draft version November 27, 2017
Preprint typeset using LaTeX style emulateapj v. 12/16/11

DETECTING WEAK SPECTRAL LINES IN INTERFEROMETRIC DATA THROUGH MATCHED FILTERING

Ryan A. Loomis[1,2], Karin I. Öberg[2], Sean M. Andrews[2], Catherine Walsh[3,4], Ian Czekala[2,5,6] Jane Huang[2], Katherine A. Rosenfeld[2],

Draft version November 27, 2017

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 intrinsically weak and therefore necessitate detection and analysis techniques specialized for the low signal-to-noise ratio (SNR) 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 SNR 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 model of the source and 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_3OH in the protoplanetary disk around TW Hya, the technique yields a ≈53% SNR 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 at https://github.com/AstroChem/VISIBLE.

1. INTRODUCTION

The rich spatio-kinematic information that radio interferometric datasets 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 datasets, 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 Fig. 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 SNR. Spectra can be extracted from the cube, and moment maps can be generated by collapsing the cube along the spectral axis, illustrated in Fig. 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 spectrum with optimal SNR, however. In sources with spatio-kinematic patterns, due to e.g. 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; Öberg et al. 2015; Loomis 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 nontrivial computational cost for high resolution observations or spectral surveys with large bandwidths. Second, when the visibilities are Fourier inverted, the PSF 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

[1] Corresponding author: rloomis@cfa.harvard.edu
[2] Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138
[3] Leiden Observatory, Leiden University, P.O. Box 9531, 2300 RA Leiden, The Netherlands
[4] School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK
[5] Stanford University, 452 Lomita Mall, Stanford, CA, 94305
[6] KIPAC Postdoctoral Fellow

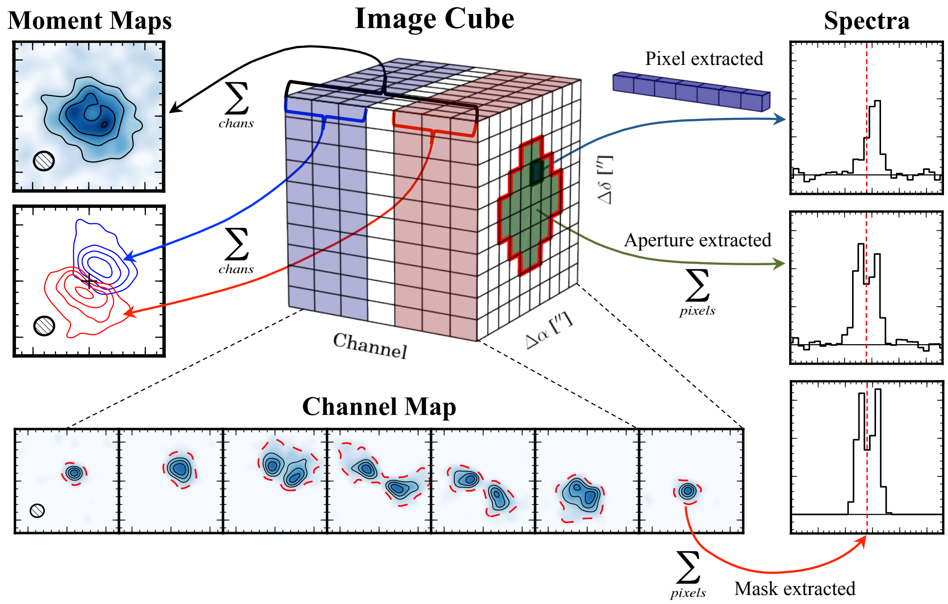

Figure 1. A diagram illustrating the multiple ways of viewing an image cube. Counter-clockwise from the top-left: Velocity-integrated moment maps, made by integrating slices of the cube along the frequency axis; channel maps, where each panel corresponds to a channel of the cube; spectra, generated from 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).

(Barenfeld et al. 2016).

The characteristics that make detecting spectral lines difficult in the image plane prove to be an asset, however, when operating in the native Fourier domain. 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, an ideal line detection method would quantitatively combine both amplitude and phase information of the observed visibilities into a robust detection probability, with a clear baseline or false positive rate against which to compare. Line detection directly in the visibilities both avoids the high computational expense of fully imaging wide-bandwidth datasets and retains a straightforward statistical interpretation of detection significance.

This task can be accomplished through application of 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 SNR. This approach is used extensively in digital signal processing; prominent examples include RADAR (e.g. Woodward 1953; Cumming & Wong 2005, and references therein), 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.

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 §2, we provide an overview of matched filtering and detail the steps of the method. In §3, we successfully apply the technique to ALMA Cycle 2 observations from Walsh et al. (2016) of CH_3OH in a protoplanetary disk. In §4, we discuss how much SNR boost might be expected for a given dataset, compare the technique to other methods, and suggest applications where matched filtering may prove useful. A summary is given in §5.

2. METHOD

In this section we first present a brief overview of the principles behind matched filtering and interferometric visibilities. We then provide detailed instruction and examples for each of the steps in the method:

1. Generation of a filter which approximates the true emission pattern.
2. Cross-correlation of this filter with the data.
3. Spectrum normalization and detection inference.
4. Line stacking (where applicable).

2.1. *Matched Filtering*

A signal $\boldsymbol{s}$ might be corrupted by additive white noise $\boldsymbol{v}$, yielding an observation $\boldsymbol{x} = \boldsymbol{s} + \boldsymbol{v}$. To maximize the SNR of this signal by applying a linear filter $\boldsymbol{h}$, we can first write the SNR (using the definition of signal power/noise power) as

$$\mathrm{SNR} = \frac{\boldsymbol{h}^* \boldsymbol{s}\boldsymbol{s}^* \boldsymbol{h}}{\boldsymbol{h}^* \boldsymbol{R_v} \boldsymbol{h}}, \tag{1}$$

where * denotes the conjugate transpose and $\boldsymbol{R_v} = E[\boldsymbol{v}\boldsymbol{v}^*]$ is a covariance matrix of the noise $\boldsymbol{v}$, where $E[\,]$ is the expectation operator. Under these conditions, the filter $\boldsymbol{h}$ which maximizes SNR is

$$\boldsymbol{h} = [\frac{1}{\sqrt{\boldsymbol{s}^* \boldsymbol{R_v}^{-1} \boldsymbol{s}}}]\boldsymbol{R_v}^{-1}\boldsymbol{s} = C\boldsymbol{R_v}^{-1}\boldsymbol{s}, \tag{2}$$

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

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

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

This impulse response spectrum loses any 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 SNR of the signal is increased (see Appendix A for a mathematical discussion of this SNR boost). Because the filter is linear, the Gaussian nature of the noise is preserved. The impulse response spectrum can be easily examined for evidence of $\boldsymbol{s}$, with a detection threshold set to some multiple of the standard deviation of $\boldsymbol{T}$ (e.g. 4σ), or a false positive rate scaled with the variance of $\boldsymbol{T}$.

2.2. *Interferometric Visibilities*

Matched filtering can be easily extended to higher dimensional problems such as searching for signal within an image or image cube (e.g. Feng et al. 2017; White & Padmanabhan 2017). We consider here the native measurement space of interferometric data, or the Fourier (u, v) plane.

Visibilities are measured at a series of discrete frequencies or channels, and the projected baseline distances between antennae define locations on the (u, v) plane (see e.g. Thompson et al. 2017). Each visibility V_i is associated with a unique weight $w_i = 1/\sigma_i^2$, where σ_i^2 encodes the variance of V_i. Cross-correlation in this discretely sampled three dimensional $(u, v, channel)$ space is computationally awkward, but the dataset can be reshaped to a two-dimensional dataset of size (n_{uv}, n_c), where each visibility row corresponds to a unique location on the (u, v) plane. Visibilities are stored this way in both the UVFITS and Measurement Set (MS) 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 `simobserve` 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 and is faster than either `uvmodel` or `simobserve`. This package builds on an algorithm in the DiskJockey package (Czekala et al. 2015; Czekala 2016), and uses the spheroidal gridding function approximations described by Schwab (1984). [7]

2.3. *Filter Kernel Generation*

The principle assumption of a matched filter analysis is that the shape of the signal 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. 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.

A simple example of the first approach is to use a basic model of the source spatio-kinematic structure as the filter kernel. For objects like protoplanetary disks or galaxies, the source inclination and position angle are often well-known and the gas velocity pattern can be easily approximated. A demonstration is shown in the top panels of Fig. 2, where we have generated a simple Keplerian mask for molecular emission from the protoplanetary disk around TW Hya (with an inclination of 7° and PA of 155°; e.g., Qi et al. 2004; Andrews et al. 2012). The Keplerian velocity field is calculated as

$$v_k = \sqrt{\frac{GM_*}{r}}, \tag{4}$$

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

A more detailed filter kernel cab 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 Fig. 2, generated from the parametric CH_3OH abundance model in Walsh et al. (2016). We use their 'fiducial' model, with CH_3OH constrained to a vertical layer z/r < 0.1 between radii of 30-100 AU. From this abundance structure, an emission profile was calculated for the 3_{12}-3_{03} transition using LIME. As seen in Fig. 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_3OH model has an inner disk depletion (as CH_3OH is mainly formed through hydrogenation of CO on grain surfaces outside the CO snowline), which is not present in the simple Keplerian model.

[7] In addition to its utility for filter kernel generation, we note that `vis_sample` may be useful for visibility fitting of modern interferometric datasets (e.g. MacGregor et al. 2016). `vis_sample` is publicly available at https://github.com/AstroChem/vis_sample or in the Anaconda Cloud at https://anaconda.org/rloomis/vis_sample

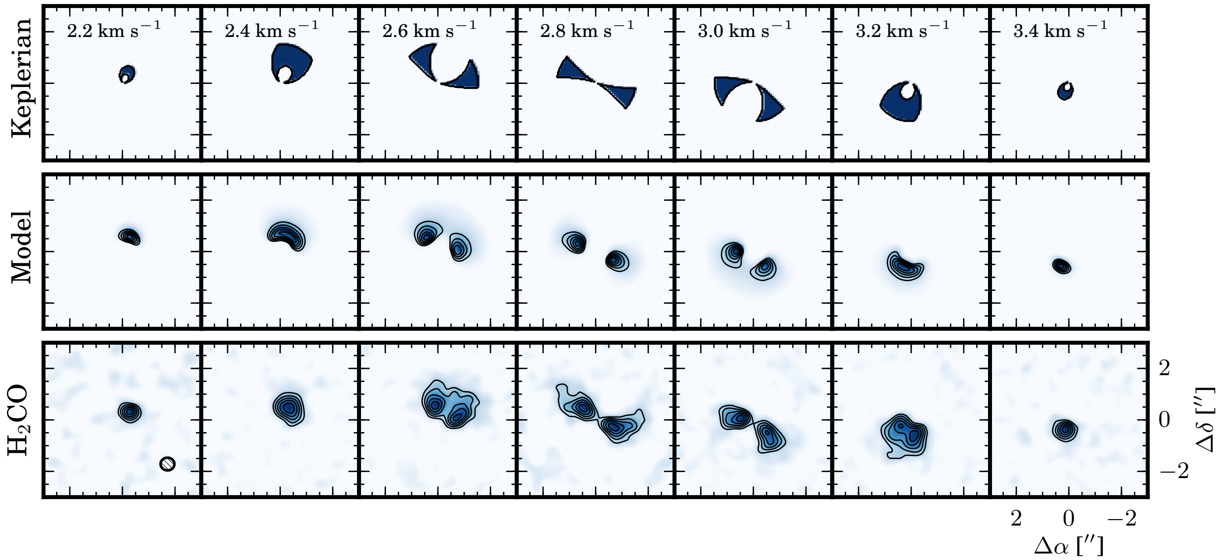

Figure 2. Three examples of filter kernels for emission from the protoplanetary disk around TW Hya. *Top:* a simple kernel based on Keplerian rotation. *Middle:* a kernel based on a parametric model of CH_3OH from Walsh et al. (2016). *Bottom:* a data-driven kernel generated from H_2CO observations from Öberg et al. (2017). All kernels have 0.2 km s^{-1} channels and are normalized by their peak intensities.

In contrast to these modeled kernels, a data-driven approach makes the assumption 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 molecular species. In Carney et al. (2017), we used this approach to detect weak H_2CO lines in HD 163296, using a stronger H_2CO line as a data-driven filter (see §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 Fig. 2 present observations of H_2CO around TW Hya (Öberg et al. 2017) which should be possible to use as a filter for CH_3OH, due to their linked formation pathways (e.g., Cuppen et al. 2009; Qi et al. 2013; Walsh et al. 2014; Loomis et al. 2015). The H_2CO data were imaged in CASA using `CLEAN` with natural weighting, yielding a high SNR image cube. After imaging, all noise below 3σ and any emission outside of a $3''$ radius were masked out, yielding an approximation of the true H_2CO emission.

2.4. *Computing the Impulse Response Function*

Fig. 3 schematically diagrams 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 sampled to produce a (u, v) plane filter kernel $\boldsymbol{f}$ of size (n_{uv}, n_k). This kernel is cross-correlated with the data $\boldsymbol{V}$, 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 is calculated by taking the inner product of the windowed data (multiplied by the data weights $\boldsymbol{w}$) with the kernel. The resultant impulse response spectrum $\boldsymbol{T}$ is therefore:

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

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) dataset in memory or non-sequential memory access, preventing speed increases through vectorization.[8] 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 5, but with the order of the summations switched,

$$T_{i_0} = \sum_{j=0}^{n_{uv}} \sum_{i=i_0}^{i_0+n_k-1} V_{i,j} w_{i,j} f_{i-i_0,j}; \quad i_0 \in [0, n_c - n_k], \quad (6)$$

yielding a final impulse response spectrum identical to that from the sliding window method shown in Fig. 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 (up to 3840 channels) can be filtered very quickly on a desktop (e.g. a few seconds on a quad-core 3.3GHz processor).

[8] Using FFT cross-correlation is even slower for typical interferometric dataset sizes.

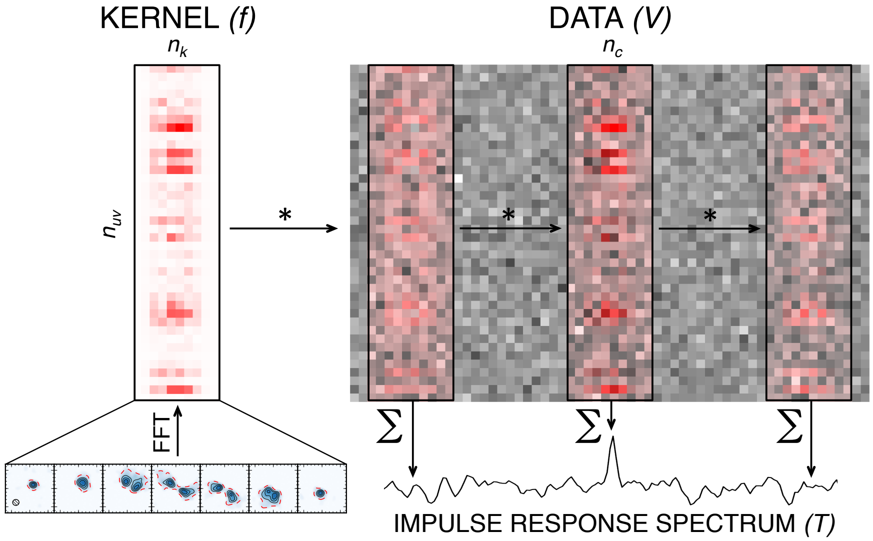

Figure 3. A diagram describing how the filter kernel is cross-correlated with the data to produce a spectrum of the impulse response to the filter. The kernel is shown on the left, with dimensions of n_k horizontally and n_{uv} vertically (not shown to scale). Several representative channels of the kernel are shown imaged. The values of the kernel have been binned and pixelated to be visually intuitive. The data shown in gray-scale is also binned and pixelated, and has an identical number of (u, v) points, 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 response is summed over the entire region to produce the corresponding impulse response for each channel, with the line detected in the central channel.

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, and any unsubtracted continuum emission will result in a constant offset in the response spectrum. The raw response spectrum can therefore be normalized by first subtracting off any constant offset, and then dividing by the standard deviation of the spectrum (excluding any obvious signal). The resulting spectrum will have units of standard deviations (σ) with the RMS noise level normalized to unity.

Peaks in this response spectrum can be evaluated against a detection threshold, set at some number of standard devations corresponding to an acceptable false alarm rate. It is important to note, however, that the detection significance is a lower limit, as it is unknown how closely the filter approximates the ideal matched filter. Additionally, if multiple filters are tested on the data, the detection statistic must be corrected for multiple hypothesis testing after assessing the degree of independence between the different filters. More details about false alarm rates in matched filtering can be found in Vio & Andreani (2016) and Vio et al. (2017).

2.6. *Comparison to Image-Plane Spectral Extraction*

As a proof of concept, we apply the method to synthetic observations of CH_3OH emission in TW Hya and compare to an aperture-based spectral extraction in the image plane. The modeled emission from the middle panels of Fig. 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 §3.

A synthetic measurement set of observations was created from the CH_3OH emission cube described in §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 mJy bm^{-1} across each 0.15 km/s channel, equivalent to the Walsh et al. (2016) observations. The noiseless and noisy measurement sets 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 SNR of the images. Only the noiseless measurement set was `CLEAN`ed; the noisy measurement set was dirty imaged to prevent bias from over-`CLEAN`ing, as the emission is practically at the noise limit in any given channel. Integrated intensity (moment-0) maps of the noiseless and noisy 3_{12}-3_{03} transitions are shown in Fig. 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 ∼3.3 mJy bm^{-1} km s^{-1} and the peak integrated flux is ∼13.2 mJy bm^{-1} km s^{-1}, yielding a SNR of ∼4. A spectrum was extracted from the noisy image cube using an aperture 3″ in diameter, equivalent to the extent of the CH_3OH emission (Fig. 4, panel c). The spectrum has a peak flux of ∼11.4 mJy and a rms noise of ∼3.2 mJy, yielding a SNR of ∼3.5σ. The rms noise level of the noisy spectrum was estimated from all channels without significant emission (i.e., excluding a velocity range of ± 1.5 km s^{-1} around the systemic velocity of 2.8 km s^{-1}).

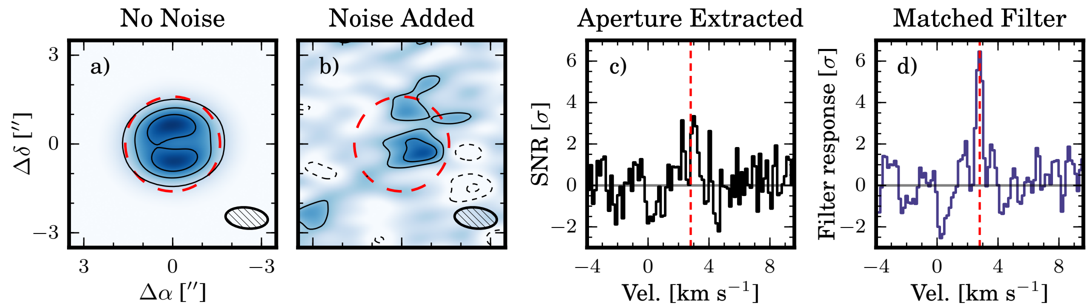

Figure 4. Comparison of the ideal matched filter with conventional spectral extraction through aperture masking. *Panel a:* Moment-0 map of simulated, noiseless CH_3OH 3_{12}-3_{03} emission. The synthesized beam is shown in the lower left. Contours are [-3, -1.5, 1.5, 3]×3.3 mJy bm^{-1} km s^{-1}, corresponding to 1σ in panel b. *Panel b:* Moment-0 map of simulated and noise-corrupted CH_3OH 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 σ, defined as the rms filter response in channels with no emission.

We cross-correlate the CH_3OH filter kernel with the synthetic observations, as described in §2.4, generating the filter response shown in Fig. 4 panel d. The peak value of the filter response, 6.5σ, is the maximum SNR extractable from the data and represents a ∼60% and ∼85% improvement over the moment-0 and spectral detections, respectively. This already corresponds to an increase in effective observing time of 2-4, but as discussed in both §4.1 and Appendices A & B the level of possible SNR improvements will be higher for data sets that are better-resolved.

2.7. *Stacking*

Stacking is a common method of SNR 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,

$$\boldsymbol{T_s} = \sum_{i=0}^{n_s} \boldsymbol{T_i} w_i, \tag{7}$$

where the stacked spectrum $\boldsymbol{T_s}$ is generated by summing n_s individual spectra $\boldsymbol{T_i}$ multiplied by weights w_i, proportional to the SNR of each $\boldsymbol{T_i}$. 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 SNR 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 §2.6 for three CH_3OH transitions: 2_{11}-2_{02}, 3_{12}-3_{03}, and 4_{13}-4_{04}, with relative strengths of 1.8:1.3:1. Moment-0 maps of the emission from each of these transitions are shown in Fig. 5, panels a, b, and c, with peak integrated fluxes of 11.7, 13.2, and 8.5 mJy km s^{-1} and corresponding SNRs of 3.5, 4, and 2.6σ, respectively. The individual filter responses are shown in Fig. 5, panels d, e, and f, with peak SNRs of 7.9, 6.5, and 4.3σ, respectively. The filter responses were stacked using a weighted average, yielding the spectrum shown in Fig. 5, panel g, with a peak SNR of 11.4σ. The

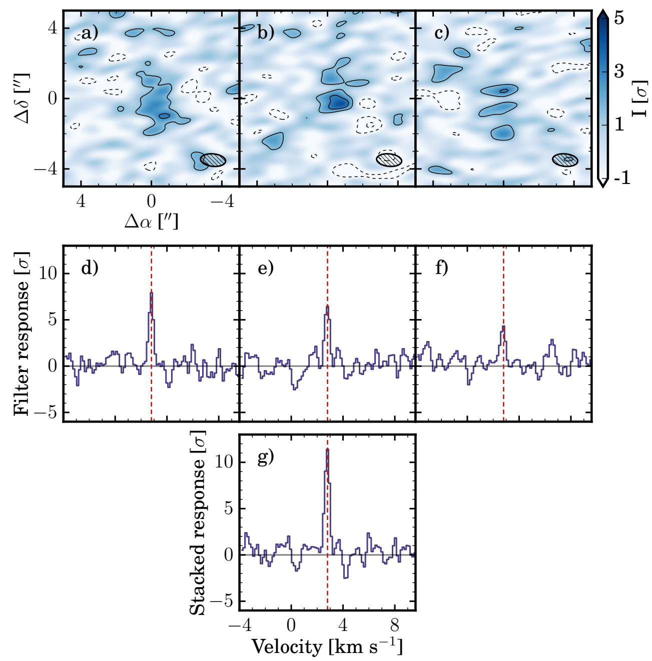

Figure 5. Demonstration of line stacking on synthetic CH_3OH emission. *Panel a-c:* Moment-0 maps of the three simulated and noise-corrupted transitions, 2_{11}-2_{02}, 3_{12}-3_{03}, and 4_{13}-4_{04}. Contours are [-3, -1.5, 1.5, 3]×σ, σ=3.3 mJy bm^{-1} km s^{-1}. The synthesized beam is shown in the lower left. *Panel d-f:* Ideal matched filter response spectra. Peak SNRs are 7.9, 6.5, and 4.3σ. *Panel g:* Filter response spectrum created by stacking the individual spectra from panels d-f. Each spectrum was weighted by its SNR, and the resultant spectrum has a SNR of 11.4.

ratio of the filter responses (1.8:1.5:1) very well recovers the flux ratio of the input models (1.8:1.3:1), even though the 2_{11}-2_{02} transition is weaker than would be expected in the imaged data (likely due to random noise fluctuations in the inherently more noisy moment-0 maps). This 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. However, its utility is less clear when the filter can only be approximated. To explore this, we have applied the method to real ALMA Band 7 observations of CH_3OH toward TW Hya (project

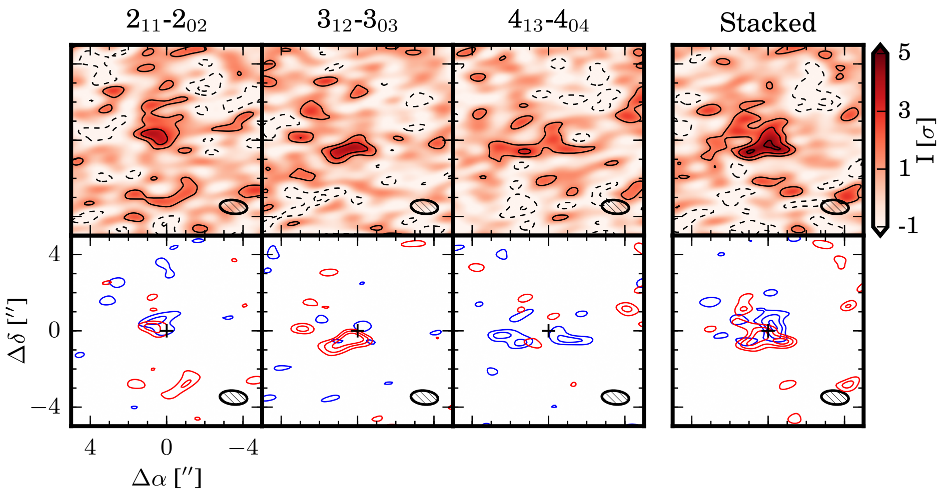

Figure 6. CH_3OH observations toward TW Hya. *Top:* CH_3OH emission from the 2_{11}-2_{02}, 3_{12}-3_{03}, and 4_{13}-4_{04} transitions, and all three stacked. Contours are [-3, -1.5, 1.5, 3, 4.5]$\times\sigma$, σ=$\sim$3.6 mJy bm^{-1} km s^{-1} for the individual transitions and $\sim$2.3 mJy bm^{-1} km s^{-1} for the stacked image. *Bottom:* same as top, but for 1 km s^{-1} velocity bins around the source velocity, showing the disk rotation.

2013.1.00902.S, P.I. C. Walsh), using all three kernels shown in Fig. 2. Details of the CH_3OH observations are presented by Walsh et al. (2016). They reported that the three observed CH_3OH 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 5.1σ detection in an aperture extracted spectrum. Moment-0 maps of the three observed CH_3OH transitions are presented in Fig. 6, with peak SNRs of 4.3, 4.3, and 2.9σ, respectively. A stacked moment-0 map is shown on the far right with a peak SNR of 4.8σ. The lower set of panels in Fig. 6 show binned red and blue-shifted 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 Fig. 2 were cross-correlated with the visibilities of each of the observed CH_3OH transitions, producing the filter responses shown in Fig. 7. All three filters detect the individual lines and show a SNR boost, demonstrating that the method is relatively robust to the choice of filter. Slight differences in the final stacked SNR, however, may provide further insight to the CH_3OH distribution. The simple Keplerian filter performs poorest, likely due to an inaccurate emission morphology, as line broadening and radially decreasing column density and temperature were not taken into account. The CH_3OH model and H_2CO data-driven filters perform better, with the H_2CO filter yielding the strongest responses for the individual lines (4.4, 6.2, and 3.4σ, respectively). Stacking these spectra together, the H_2CO filter yields a robust detection of 7.8σ, a 53% improvement over the 5.1σ detection reported in Walsh et al. (2016).

As the data responds more strongly to the H_2CO filter than the CH_3OH model filter, we can infer that the true CH_3OH distribution is likely more similar to H_2CO than the model, perhaps due to a smaller inner radius of the CH_3OH emission ring than was reported from the model fit in Walsh et al. (2016), where a monotonically decreasing column density with radius was enforced. Higher resolution observations of the CH_3OH emission at higher SNR will be able to probe the CH_3OH inner radius and test this prediction.

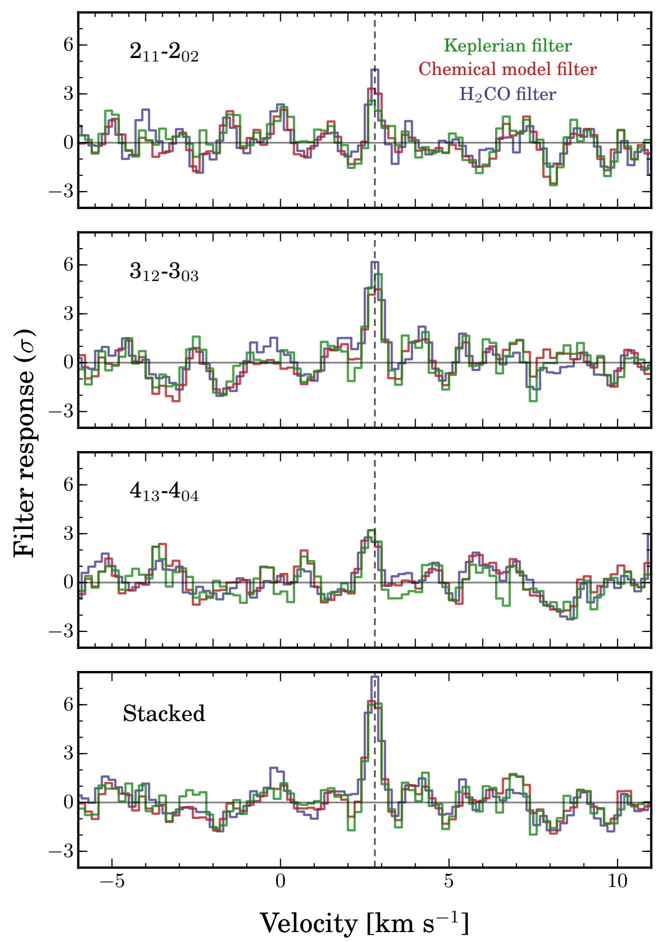

Figure 7. Filter response spectra for each CH_3OH transition. The impulse responses to the Keplerian, CH_3OH model, and H_2CO filters are shown in green, red, and blue, respectively.

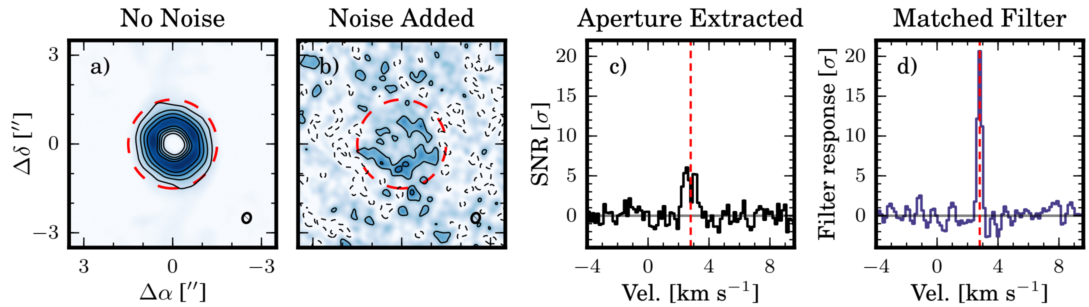

Figure 8. Comparison of matched filtering with traditional methods, as in Fig. 4 but at higher spatial resolution. *Panel a:* Moment-0 map of simulated noiseless CH_3OH 3_{12}-3_{03} emission. The synthesized beam is shown in the lower left. Contours are [-3, -1.5, 1.5, 3]×5.6 mJy bm^{-1} km s^{-1}, corresponding to 1σ in panel b. *Panel b:* Moment-0 map of simulated and noise-corrupted CH_3OH 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 σ, defined as the rms filter response in channels with no emission.

4. DISCUSSION

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

4.1. *Factors Affecting SNR Boost*

Compared with traditional line detection methods, the matched filter approach offers an improved SNR. The degree of SNR 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 §3, we found that application of a matched filter could increase SNR up to 85%. The decreased efficacy in the application to real data (53% vs 85%) can likely be attributed to imperfect filters.

Intuitively, the SNR 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 SNR other than the boost from spectral averaging. As emission is spatially resolved, SNR 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 SNR boost scales directly with the spatial resolution of the signal (see Yen et al. 2016, for a detailed image-plane derivation of this SNR boost). Fig. 8 illustrates this effect, with a simulation similar to that shown in Fig. 4, but with higher spatial resolution. The data were noise corrupted to reach a similar $\sim$4σ detection in the moment map, although the SNR in the extracted spectrum is now $\sim$6σ, highlighting how ineffective moment maps are at high spatial resolutions. The filter response is also now much larger (SNR=20.6σ), yielding a SNR boost over the aperture extracted spectra of $\sim$340%, compared to 85% in Fig. 4. Application of matched filtering to higher resolution observations of H_2CO (Carney et al. 2017) produced a SNR gain of over 500%, confirming in practice the relationship between SNR 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 SNR (see Appendix B for more details). This is confirmed by comparing the results of the matched filtering technique to detect H_2CO in HD 163296 (Carney et al. 2017) with those obtained on the same dataset by Yen et al. (2016). In both cases, a SNR 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 uv-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). As shown by the different filter responses in Figure 7, this can have a significant effect on the SNR boost. On the other hand, the image-plane nature of the pixel stacking 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 dataset and pixel stacking then used to better characterize the lines.

4.3. *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. Imaging the full bandwidth of these large datasets at their native spectral resolution is a time consuming process, often taking many hours or even days. Because much of the information in these datasets 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 dataset, 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.

Tools are currently being developed to aid in windowing out spectral lines from broadband datasets (e.g., ADMIT, Friedel et al. 2015), but they rely on a fully imaged datacube. 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 dataset in a small fraction of the time it would take to image that same dataset. The resulting full-band impulse response spectrum then provides a convenient first look at the dataset, 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.

4.4. *Line Flux Estimation*

As previously discussed, the main utility of matched filtering when applied to interferometeric 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 SNR using the signal/rms definition:

$$\mathrm{SNR_s} = \sqrt{\frac{\boldsymbol{h}^*\boldsymbol{s}\boldsymbol{s}^*\boldsymbol{h}}{\boldsymbol{h}^*\boldsymbol{R_v}\boldsymbol{h}}} \tag{8}$$

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

$$\mathrm{SNR_y} = \sqrt{\frac{\boldsymbol{h}^*\alpha\boldsymbol{s}\alpha\boldsymbol{s}^*\boldsymbol{h}}{\boldsymbol{h}^*\boldsymbol{R_v}\boldsymbol{h}}} \tag{9}$$

which reduces to:

$$\mathrm{SNR_y} = \alpha\sqrt{\frac{\boldsymbol{h}^*\boldsymbol{s}\boldsymbol{s}^*\boldsymbol{h}}{\boldsymbol{h}^*\boldsymbol{R_v}\boldsymbol{h}}} \tag{10}$$

$$\mathrm{SNR_y} = \alpha\ \mathrm{SNR_s} \tag{11}$$

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 H_2CO lines, enabling them to constrain the H_2CO excitation temperature.

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 at https://github.com/AstroChem/VISIBLE. As a case study, we have focused on the detection of CH_3OH in protoplanetary disks, but our approach is applicable to spectral line data of any astronomical source with a coherent velocity field. We find that when applied to real data, the method results in large sensitivity increases, ranging from 53% for CH_3OH in Walsh et al. (2016), to $\sim$500% for H_2CO 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 dataset. Finally, the method works synergistically with the methods presented in (Matrà et al. 2015) and Yen et al. (2016) and tools such as ADMIT, making a formidable suite of analysis techniques for spectral lines in large interferometric datasets.

We would like to thank Andrew Vanderburg and Peter Williams for productive discussions and assistance with software package management and distribution. RAL and JH gratefully acknowledge funding from National Science Foundation Graduate Research Fellowships (Grant No. DGE-1144152). RAL also acknowledges funding from the NRAO Student Observing Support Program. KIO acknowledges funding from the Alfred P. Sloan Foundation and the David and Lucile Packard Foundation. CW 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.

REFERENCES

Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., Adhikari, R. X., & et al. 2016, Physical Review Letters, 116, 061102

Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., Rosenfeld, K. A., Öberg, K. I., Birnstiel, T., Espaillat, C., Cieza, L. A., Williams, J. P., Lin, S.-Y., & Ho, P. T. P. 2012, ApJ, 744, 162

Barenfeld, S. A., Carpenter, J. M., Ricci, L., & Isella, A. 2016, ApJ, 827, 142

Briggs, D. S. 1995, in Bulletin of the American Astronomical Society, Vol. 27, American Astronomical Society Meeting Abstracts, 1444

Brinch, C. & Hogerheijde, M. R. 2010, A&A, 523, A25

Carney, M. T., Hogerheijde, M. R., Loomis, R. A., Salinas, V. N., Öberg, K. I., Qi, C., & Wilner, D. J. 2017, ArXiv e-prints

Coutens, A., Jørgensen, J. K., van der Wiel, M. H. D., Müller, H. S. P., Lykke, J. M., Bjerkeli, P., Bourke, T. L., Calcutt, H., Drozdovskaya, M. N., Favre, C., Fayolle, E. C., Garrod, R. T., Jacobsen, S. K., Ligterink, N. F. W., Öberg, K. I., Persson, M. V., van Dishoeck, E. F., & Wampfler, S. F. 2016, A&A, 590, L6

Crane, P. C. & Napier, P. J. 1986, in Synthesis Imaging, ed. R. A. Perley, F. R. Schwab, & A. H. Bridle, 87–108

Cumming, I. G. & Wong, F. H. 2005, Artech house, 1, 3

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

Czekala, I., Andrews, S. M., Jensen, E. L. N., Stassun, K. G., Torres, G., & Wilner, D. J. 2015, ApJ, 806, 154

Dullemond, C. P. 2012, RADMC-3D: A multi-purpose radiative transfer tool, Astrophysics Source Code Library

Dutrey, A., Henning, T., Guilloteau, S., Semenov, D., Piétu, V., Schreyer, K., Bacmann, A., Launhardt, R., Pety, J., & Gueth, F. 2007, A&A, 464, 615

Feng, L., Vaulin, R., Hewitt, J. N., Remillard, R., Kaplan, D. L., Murphy, T., Kudryavtseva, N., Hancock, P., Bernardi, G., Bowman, J. D., Briggs, F., Cappallo, R. J., Deshpande, A. A., Gaensler, B. M., Greenhill, L. J., Hazelton, B. J., Johnston-Hollitt, M., Lonsdale, C. J., McWhirter, S. R., Mitchell, D. A., Morales, M. F., Morgan, E., Oberoi, D., Ord, S. M., Prabu, T., Udaya Shankar, N., Srivani, K. S., Subrahmanyan, R., Tingay, S. J., Wayth, R. B., Webster, R. L., Williams, A., & Williams, C. L. 2017, AJ, 153, 98

Friedel, D., Looney, L., Xu, L., Pound, M. W., Teuben, P. J., Rauch, K. P., Mundy, L., & Kern, J. S. 2015, in 70th International Symposium on Molecular Spectroscopy

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

Jørgensen, J. K., Bourke, T. L., Nguyen Luong, Q., & Takakuwa, S. 2011, A&A, 534, A100

Kalenskii, S. V. & Johansson, L. E. B. 2010, Astronomy Reports, 54, 295

Langston, G. & Turner, B. 2007, ApJ, 658, 455

Loomis, R. A., Cleeves, L. I., Öberg, K. I., Guzman, V. V., & Andrews, S. M. 2015, ApJ, 809, L25

Loomis, R. A., Shingledecker, C. N., Langston, G., McGuire, B. A., Dollhopf, N. M., Burkhardt, A. M., Corby, J., Booth, S. T., Carroll, P. B., Turner, B., & Remijan, A. J. 2016, MNRAS, 463, 4175

MacGregor, M. A., Lawler, S. M., Wilner, D. J., Matthews, B. C., Kennedy, G. M., Booth, M., & Di Francesco, J. 2016, ApJ, 828, 113

Marino, S., Matrà, L., Stark, C., Wyatt, M. C., Casassus, S., Kennedy, G., Rodriguez, D., Zuckerman, B., Perez, S., Dent, W. R. F., Kuchner, M., Hughes, A. M., Schneider, G., Steele, A., Roberge, A., Donaldson, J., & Nesvold, E. 2016, MNRAS, 460, 2933

Matrà, L., Panić, O., Wyatt, M. C., & Dent, W. R. F. 2015, MNRAS, 447, 3936

Müller, H. S. P., Belloche, A., Xu, L.-H., Lees, R. M., Garrod, R. T., Walters, A., van Wijngaarden, J., Lewen, F., Schlemmer, S., & Menten, K. M. 2016, A&A, 587, A92

North, D. O. 1963, Proceedings of the IEEE, 51, 1016

Öberg, K. I., Guzmán, V. V., Furuya, K., Qi, C., Aikawa, Y., Andrews, S. M., Loomis, R., & Wilner, D. J. 2015, Nature, 520, 198

Öberg, K. I., Guzmán, V. V., Merchantz, C. J., Qi, C., Andrews, S. M., Cleeves, L. I., Huang, J., Loomis, R. A., Wilner, D. J., Brinch, C., & Hogerheijde, M. 2017, ApJ, 839, 43

Owen, B. J. & Sathyaprakash, B. S. 1999, Phys. Rev. D, 60, 022002

Qi, C., Ho, P. T. P., Wilner, D. J., Takakuwa, S., Hirano, N., Ohashi, N., Bourke, T. L., Zhang, Q., Blake, G. A., Hogerheijde, M., Saito, M., Choi, M., & Yang, J. 2004, ApJ, 616, L11

Qi, C., Öberg, K. I., & Wilner, D. J. 2013, ApJ, 765, 34

Ruffio, J.-B., Macintosh, B., Wang, J. J., Pueyo, L., Nielsen, E. L., De Rosa, R. J., Czekala, I., Marley, M. S., Arriaga, P., Bailey, V. P., Barman, T., Bulger, J., Chilcote, J., Cotten, T., Doyon, R., Duchêne, G., Fitzgerald, M. P., Follette, K. B., Gerard, B. L., Goodsell, S. J., Graham, J. R., Greenbaum, A. Z., Hibon, P., Hung, L.-W., Ingraham, P., Kalas, P., Konopacky, Q., Larkin, J. E., Maire, J., Marchis, F., Marois, C., Metchev, S., Millar-Blanchaer, M. A., Morzinski, K. M., Oppenheimer, R., Palmer, D., Patience, J., Perrin, M., Poyneer, L., Rajan, A., Rameau, J., Rantakyrö, F. T., Savransky, D., Schneider, A. C., Sivaramakrishnan, A., Song, I., Soummer, R., Thomas, S., Wallace, J. K., Ward-Duong, K., Wiktorowicz, S., & Wolff, S. 2017, ApJ, 842, 14

Schutz, B. F. 1999, Classical and Quantum Gravity, 16, A131

Schwab, F. R. 1984, in Indirect Imaging. Measurement and Processing for Indirect Imaging, ed. J. A. Roberts, 333–346

Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2017, Interferometry and Synthesis in Radio Astronomy, 3rd Edition (Springer)

Vio, R. & Andreani, P. 2016, A&A, 589, A20

Vio, R., Verges, C., & Andreani, P. 2017, ArXiv e-prints

Walsh, C., Loomis, R. A., Öberg, K. I., Kama, M., van 't Hoff, M. L. R., Millar, T. J., Aikawa, Y., Herbst, E., Widicus Weaver, S. L., & Nomura, H. 2016, ApJ, 823, L10

Walsh, C., Millar, T. J., Nomura, H., Herbst, E., Widicus Weaver, S., Aikawa, Y., Laas, J. C., & Vasyunin, A. I. 2014, A&A, 563, A33

White, M. & Padmanabhan, N. 2017, ArXiv e-prints

Woodward, P. 1953, Probability and Information Theory: With Applications to Radar, Electronics and Waves No. v. 3 (Elsevier Science & Technology)

Yen, H.-W., Koch, P. M., Liu, H. B., Puspitaningrum, E., Hirano, N., Lee, C.-F., & Takakuwa, S. 2016, ApJ, 832, 204

APPENDIX

A. CALCULATING SNR BOOST FOR A MATCHED FILTER

SNR (using the definition of signal-power/noise-power) can be written for an arbitrary signal $\boldsymbol{s}$ and filter $\boldsymbol{h}$ as:

$$\mathrm{SNR} = \frac{\boldsymbol{h}^* \boldsymbol{s} \boldsymbol{s}^* \boldsymbol{h}}{\boldsymbol{h}^* \boldsymbol{R_v} \boldsymbol{h}}. \tag{A1}$$

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

$$\boldsymbol{h} = [\frac{1}{\sqrt{\boldsymbol{s}^* \boldsymbol{R_v}^{-1} \boldsymbol{s}}}]\boldsymbol{R_v}^{-1}\boldsymbol{s} = C\boldsymbol{R_v}^{-1}\boldsymbol{s}, \tag{A2}$$

maximizes the output SNR. The natural question is then how much is the SNR boosted by applying such a filter? This can be analytically calculated for a given signal by comparing the SNR after applying a matched filter with the SNR from applying a flat filter **1** (e.g. a unity matrix of all ones). We start by calculating the SNR after applying the matched filter:

$$\mathrm{SNR_{mf}} = \frac{\boldsymbol{h}^* \boldsymbol{s}\boldsymbol{s}^* \boldsymbol{h}}{\boldsymbol{h}^* \boldsymbol{R_v} \boldsymbol{h}}. \tag{A3}$$

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

$$\mathrm{SNR_{mf}} = \frac{\boldsymbol{R_v}^{-1}\boldsymbol{s}^* \boldsymbol{s}\boldsymbol{s}^* \boldsymbol{R_v}^{-1}\boldsymbol{s}}{\boldsymbol{R_v}^{-1}\boldsymbol{s}^* \boldsymbol{R_v} \boldsymbol{R_v}^{-1}\boldsymbol{s}}. \tag{A4}$$

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

$$\mathrm{SNR_{mf}} = \sum_i^N \|s_i\|^2 R_{ii}^{-1}, \tag{A5}$$

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

$$\mathrm{SNR_{flat}} = \frac{\boldsymbol{1}^* \boldsymbol{s}\boldsymbol{s}^* \boldsymbol{1}}{\boldsymbol{1}^* \boldsymbol{R_v} \boldsymbol{1}}, \tag{A6}$$

then we find it reduces to:

$$\mathrm{SNR_{flat}} = \frac{(\sum_i^N \|s_i\|)^2}{\mathrm{tr}[\boldsymbol{R_v}]}. \tag{A7}$$

So the ratio of these two SNRs, or the total SNR boost from a matched filter, is:

$$\mathrm{boost} = \frac{\mathrm{SNR_{mf}}}{\mathrm{SNR_{flat}}} = \frac{(\sum_i^N \|s_i\|^2 R_{ii}^{-1})\mathrm{tr}[\boldsymbol{R_v}]}{(\sum_i^N \|s_i\|)^2}. \tag{A8}$$

B. CALCULATING SNR BOOST IN COMPARISON TO IMAGE-PLANE MEASUREMENTS

The boost value in equation A8 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 SNRs discussed throughout the paper. Thus the fundamental problem is how to relate the visibilities to an image-plane SNR. We start by writing down the definition of SNR in the dirty image $\boldsymbol{I^D}$, or the raw discrete Fourier transform of the visibilites (i.e. not deconvolved):

$$\mathrm{SNR} = \frac{\boldsymbol{I^D}}{\boldsymbol{\Delta I^D}} = \frac{\sum\limits_k W_k V_k}{(\sum\limits_k W_k^2 \sigma_k^2)^{1/2}}, \tag{B1}$$

where there are k visibilities in the dataset, 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 σ_k. Notation is borrowed from Briggs (1995), which contains a detailed discussion of image-plane SNRs and their relation to the measured visibilities. If the data is gridded into cells $\{p, q\}$ and the noise properties of all visibilities within each cell are similar, an approximate form can be written:

$$\mathrm{SNR} = \frac{\sum\limits_{p,q} W_{pq} V_{pq}}{(\sum\limits_{p,q} W_{pq}^2 \sigma_{pq}^2)^{1/2}}. \tag{B2}$$

In particular, we are interested in the SNR 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 SNR can be written as:

$$\mathrm{SNR}'(0,0) = \frac{\sum\limits_{p,q} W_{pq}|V_{pq}|}{(\sum\limits_{p,q} W_{pq}^2\sigma_{pq}^2)^{1/2}}. \tag{B3}$$

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

$$\mathrm{SNR}_{\mathrm{mom0}} = \frac{\sum\limits_{c,p,q} W_{cpq}\mathrm{Re}(V_{cpq})}{(\sum\limits_{c,p,q} W_{cpq}^2\sigma_{cpq}^2)^{1/2}}, \tag{B4}$$

and only the projected real component of each visibility will contribute signal. The SNR 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 A8, we can estimate the SNR boost of the matched filter over a peak moment-0 value as:

$$\mathrm{boost} = \frac{\mathrm{SNR}_{\mathrm{mf}}}{\mathrm{SNR}_{\mathrm{mom0}}} = \frac{(\sum_i^N \|V_i\|^2 R_{ii}^{-1})\mathrm{tr}[\boldsymbol{R_v}]}{(\sum_i^N \mathrm{Re}(V_i))^2}. \tag{B5}$$

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. $\mathrm{Re}(V_{pq})$ can then be replaced by $|V_{pq}|$, and the SNR after applying a pixel shifting method is roughly:

$$\mathrm{SNR}_{\mathrm{ps}} = \mathrm{SNR}'_{\mathrm{mom0}} = \frac{\sum\limits_{c,p,q} W_{cpq}|V_{cpq}|}{(\sum\limits_{c,p,q} W_{cpq}^2\sigma_{cpq}^2)^{1/2}}. \tag{B6}$$

Returning to equation A8 and applying this logic, we can write the boost as:

$$\mathrm{boost} = \frac{\mathrm{SNR}_{\mathrm{mf}}}{\mathrm{SNR}_{\mathrm{ps}}} = \frac{(\sum_i^N \|V_i\|^2 R_{ii}^{-1})\mathrm{tr}[\boldsymbol{R_v}]}{(\sum_i^N |V_i|)^2}, \tag{B7}$$

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

Equations B4 and B6 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.