A simple non-parametric method for resolving merged doublet lines: insights into complex kinematics and outflows

Carolin Villforth,[★] Tom L. Underwood,[★] Mark Tolson and Nikhil Modha
Department of Physics, University of Bath, Claverton Down, Bath BA27AY, UK

ABSTRACT

Doublet line emission and absorption are common in astronomical sources (e.g. [O III], [O II], NaD, Mg II). In many cases, complex kinematics in the emitting source can cause the doublet lines to merge, making characterization of the source kinematics challenging. Here, we present a non-parametric method for resolving merged doublet lines when the line ratio and wavelength difference are known. The method takes as input only the line ratio and wavelength difference, using these quantities to resolve the components of the doublet without resorting to fitting (e.g. using multiple Gaussians) or making any assumptions about the components' line profiles (save that they are the same for both components). The method is simple, fast, and robust. It is also ideal for visualization. We show that the method recovers line profiles of merged emission lines in simulated data. We also show, using simulated data and mathematical analysis, that the method does not significantly increase noise levels in the extracted lines, and is robust to background contamination. We demonstrate the strength of the method by applying it to strongly merged [OIII] 5007/4959 Å in Active Galactic Nuclei. A PYTHON implementation of the method is provided in the appendix.

Key words: line: profiles – methods: analytical – stars: kinematics and dynamics – ISM: jets and outflows – ISM: kinematics and dynamics.

1 INTRODUCTION

Complex kinematics are common in astronomical objects (e.g. Mullaney et al. 2013; Zakamska & Greene 2014; Zakamska et al. 2016), evidenced by extremely asymmetric and non-Gaussian line profiles. A non-parametric classification of such kinematics from emission or absorption lines using, for example, skew, kurtosis, or per centiles can be used to quantify the kinematic structure of the emitting gas.

In many cases, the analysis of emission or absorption lines is complicated by the fact that many common atomic lines are observed as doublets, e.g. [O III] 5007/4959 Å, [O II] 3727/3730 Å, Mg II 2796/2803 Å, NaD 5892/5898 Å, CaH/K 3935/3970 Å. High velocities can cause the line profiles to merge (see e.g. Greene, Zakamska & Smith 2012; Zakamska & Greene 2014; Zakamska et al. 2016). And while simple kinematic structures can be fit using Gaussian mixture models (e.g. Mullaney et al. 2013), reconstructing extremely complex line profiles requires a large number of free parameters. Moreover without careful fitting procedures, continuum emission can throw off fits using multiple components, especially in noisy data. Fitting can therefore be challenging when applied to large samples.

A non-parametric approach to decompose doublet lines is therefore valuable for recovering the underlying line profiles for further

analysis. Methods for decomposing doublet line profiles by exploiting symmetry in the different lines are known in the literature (e.g. Junkkarinen, Burbidge & Smith 1983). However no general mathematical framework for this approach is known.

In this paper we focus on the general problem of merged doublet lines, and introduce and test a simple non-parametric method for resolving merged doublet line profiles when the line ratio and wavelength difference is known. The method is a fast, simple alternative to parametric fitting techniques and requires as an input only the line ratio and wavelength separation. No estimate of the line shape is required. The method is introduced in Section 2; as an example application, the method is applied to strongly merged [O III] 5007/4959 Å doublets in Active Galactic Nuclei (AGN) in Section 3 and compared to a commonly used fitting approach; we discuss practicalities associated with using the method in Section 4; and a summary of our findings, as well as possible use cases, and caveats and limitations of the method, are provided in Section 5. An implementation of the method in PYTHON is given in Appendix A, and discussion and derivations supplementary to the main text are given in Appendix B.

2 METHOD

Here, we describe the method used for doublet decomposition. The basic idea is that, since the line profiles are identical, one of the lines can be removed by subtracting the profile of the other. The

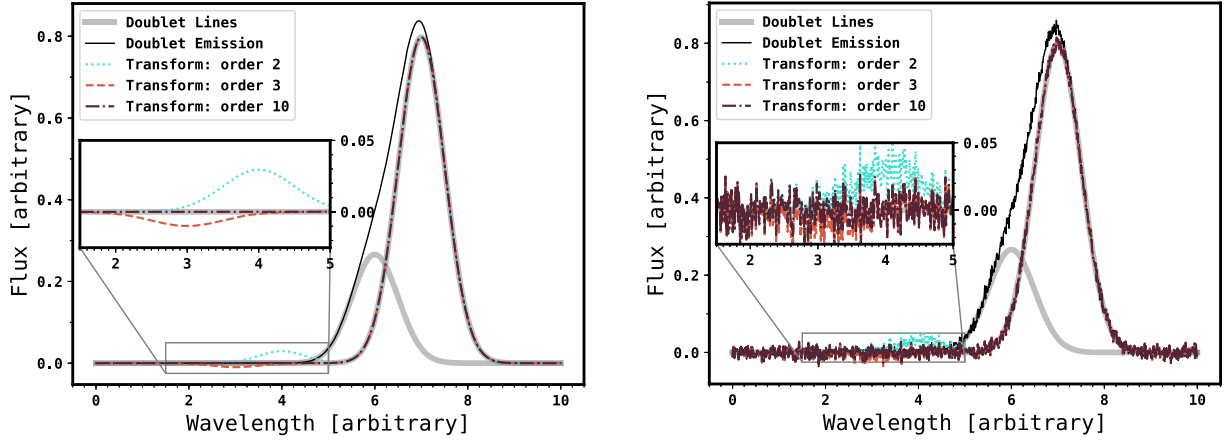★ E-mail: c.villforth@bath.ac.uk (CV); t.l.underwood@bath.ac.uk (TLU)

Figure 1. Simulations showing the performance of the method. Left-hand panel: Application of the method to noise-free merged Gaussian emission lines for different numbers of iterations (order of the transform). Individual lines are shown in grey, and the fully merged emission line is shown in dashed grey. The transformed data from orders 2, 3, and 10 are shown as cyan, orange, and dark red dashed lines, respectively. The transform is well converged to the underlying line profile at order 10. Right-hand panel: same as upper left, but with noise added to data. In both cases, the line ratio of the simulated doublet is $R = 1/3$, the separation between the lines is $\delta = 1$ and the width of each line is $\sigma = 0.5$. The noise added in the right-hand panel is Gaussian with $\sigma_{\rm noise} = 0.01$.

method shifts the full line profile, multiplied by the line ratio R, by the wavelength difference δ and subtracts it from the data. This is repeated for a number of iterations. The continued iterations effectively correct for the fact that the full line profile, rather than an individual line profile, is used for subtraction. We will show below that this simple procedure of shifting and subtracting the data recovers the underlying line profile correctly. A visualization of the method is shown in Fig. 1 (left-hand panel).

We emphasize that the method makes no assumption regarding the underlying line profile, and hence is valid for any line profile. Furthermore, we will show later in Section 4.2 that continuum emission which can be approximated as linear across the wavelength range of interest does not affect the performance of the method. Emission lines not associated with the doublet, however, should be removed before the method is applied. Nevertheless we have examined the effect of line contamination, and discuss this topic later in Section 4.5.

We will now explain the method in detail. Let $y(x)$ denote the flux observed at wavelength x (i.e. the signal). This can be expressed as

$$y(x) = f_1(x) + f_2(x), \qquad (1)$$

where $f_1(x)$ and $f_2(x)$ are the two lines that make up the doublet, and we have ignored contributions such as noise and interference from the tails of other lines near the doublet – we discuss such complications later. The assumption that underpins our method is that the two lines which make up the doublet are *similar* in the geometrical sense, i.e. line 1 can be obtained from line 2 by rescaling the former, then shifting the rescaled line along the x-axis by some amount. This can be expressed mathematically as follows: the two lines, $f_1(x)$ and $f_2(x)$, are defined to be similar if there exists an R and δ such that

$$f_2(x) = R f_1(x + \delta). \qquad (2)$$

Note that R is the line ratio of line 2 relative to line 1, and δ is the x position of peak 1 relative to peak 2 (i.e. the wavelength difference between the two lines); if peak 1 is centred at $x = 0$, then peak 2 will be centred at $x = -\delta$. In other words R is the line ratio and δ

is the wavelength difference of the two peaks which make up the doublet.[1]

Our method is to apply a transformation to $y(x)$ that yields $f_1(x)$. Once $f_1(x)$ has been obtained in this manner, $f_2(x)$ can then be obtained trivially by applying equation (2). The shape of each single line can then be analysed in detail. The transformation is as follows:

$$\tilde{g}(x) \equiv \sum_{m=0}^{\infty} (-1)^m R^m g(x + m\delta), \qquad (3)$$

where $g(x)$ is an arbitrary function and $\tilde{g}(x)$ denotes its transform (we use the notation $\tilde{g}$ to denote the transform of a function g throughout this work). Note that the transformation involves superimposing copies of the function $g(x)$, scaled by factors 1, $-R$, R^2, $-R^3, \ldots$ (which correspond to $m = 0, 1, 2, 3, \ldots$), at respective positions $x = 0, \delta, 2\delta, \ldots$; we emphasize that R and δ enter into the transformation. Proof that applying the transformation to $y(x)$ yields $f_1(x)$, i.e. that

$$\tilde{y}(x) = f_1(x), \qquad (4)$$

is provided in Appendix B1.

With the above in mind we can restate our key result. The transformation, equation (3), can be used to extract the two lines that make up a doublet, assuming the doublet is known to be composed of two similar peaks (equation 2), and the line ratio and wavelength difference are known. Note that the method requires shifting spectra in wavelength; if the sampling of the data does not match the shift applied, interpolation of data points is required.

This is demonstrated in Fig. 1 (left-hand panel), where the method is used to resolve a simulated doublet with $R = 1/3$, $\delta = 1$, and a Gaussian underlying line shape with width $\sigma = 0.5$. We use this simulated doublet to test the performance of the method later in Section 4.

---

[1]Note that that R and δ reflect the intrinsic properties of the two lines, and not their manifestations in the doublet $y(x)$. For example if the doublet $y(x)$ exhibits two local maxima, then R does not necessarily correspond to the relative heights of these maxima, nor does δ necessarily correspond to their separation.

We will now apply the method to astronomical data (Section 3) and then discuss considerations for the use of the method in Section 4.

3 EXAMPLE APPLICATION: OUTFLOWS IN TYPE 2 AGN

To show the strength of the method, we now apply the method to astronomical data. However, as outlined in Section 2, the method is generally applicable to doublet lines: the method is not just limited to the specific application considered in this section. For instance, here we apply the method to emission lines, although it is also applicable to absorption.

As a test case, we chose AGN with asymmetric [O III] 5007/4959 Å emission. The lines are separated by ∼3000 km s^{-1}. The [O III] emission in AGN traces the narrow-line region (NLR). The NLR represents the extended gas ionized by emission from the AGN (Antonucci 1993; Urry & Padovani 1995). The NLR size correlates with the luminosity of the AGN and has a size between 10 and 10^4 pc (e.g. Liu et al. 2013; Schmitt et al. 2003). Outflows are common in the NLR and are observed through blueshift with respect to the host galaxy, strong asymmetries in the emission lines (e.g. Crenshaw et al. 2010; Bae & Woo 2014), and detailed kinematic modelling of spatially resolved NLRs (e.g. Rupke & Veilleux 2011; Crenshaw et al. 2014; Carniani et al. 2015). Asymmetries in the observed narrow lines are common since dust in the NLR obscures part of the emission from the outflow, causing line profiles with blueshifted wings (e.g. Mullaney et al. 2013; Zakamska et al. 2016). These outflows form part of a multiphase outflow seen in different gas phases (e.g. Rupke, Veilleux & Sanders 2005; Rupke & Veilleux 2011; Aalto et al. 2012; Cicone et al. 2012; Liu et al. 2013; Rupke & Veilleux 2013). The velocity and radial extent of these outflows can be used to estimate outflow rates that can in many cases be well in excess of star formation rates in the host galaxy (e.g. Carniani et al. 2015). The physics of the large-scale outflows seen in AGN is still under investigation (e.g. Fabian 2012; Faucher-Giguere & Quataert 2012), but it is widely believed that such outflows can have a profound effect on the host galaxy by removing gas or terminating star formation in the host galaxy through so-called AGN feedback (e.g. Di Matteo, Springel & Hernquist 2005; Hopkins et al. 2008; Fabian 2012, and references therein). Studying NLR kinematics, especially at the highest velocities most affected by merged doublet lines (Zakamska & Greene 2014; Zakamska et al. 2016), is therefore of great interest for galaxy evolution studies. NLR emission in AGN is therefore a test case in which line profiles show complex kinematics and cannot be easily approximated by parametric fits, and the wings of the lines tracing the highest velocities are of great interest. The method presented here is therefore ideally suited for this case since it allows decomposition of merged doublets irrespective of the shape of the two lines, even when they are merged (see Fig. 1).

A detailed review of the physics of outflows from both AGN and starburst is beyond the scope of this paper; we refer the reader to the literature for a detailed treatment of the physics of outflows (e.g. Heckman, Armus & Miley 1990; Fabian 2012; Faucher-Giguere & Quataert 2012; Heckman & Thompson 2017, and references therein)

Here, we apply the method to AGN with a range of narrow line kinematics. Spectra are taken from the Sloan Digital Sky Survey (SDSS) (Abazajian et al. 2009) covering the rest frame wavelength range 3800–9200 Å and have a spectral resolution of $R_{\text{spec}} \approx 1850$–2200. We selected Type 2 AGN to avoid contamination from iron

emission (e.g. Vestergaard & Wilkes 2001; Kovacevic, Popovic & Dimitrijevic 2010) or broad Hβ emission, since a contaminating line is a complication which the method was not designed to account for. We discuss the effect of a contaminating line on the results of the method in more detail in Section 4.5.

We applied the method to AGN for which [O III] is in the SDSS spectral window (z ≤ 0.7) and a significant detection of [O III] is present, leaving >1000 spectra, depending on the cut-off used for [O III] detection. Of those, a large fraction were visually inspected to check for obvious residuals; the transform was found to be robust. Here we present, for the sake of brevity, results for only four of the AGN we considered, chosen to exhibit a range of line profiles, thus illustrating that the method works irrespective of the underlying line shape. Specifically, we choose three sources showing the strongly skewed emission lines to illustrate the performance of the transform for merged emission lines. The chosen objects show a range of S/N, illustrating the performance of the transform under different levels of noise. Additionally, we show a double-peaked [O III] emitter to demonstrate that the method recovers even complex line profiles without residuals.

Before applying the method, continuum emission was removed by locally fitting a linear function, and then subtracting the continuum. Having a (flat) non-zero continuum contribution does not adversely affect recovery of the emission line shapes (see Section 4.2, Fig. 4); the transform was applied to non-continuum subtracted data as well and shown to be reliable. We show the continuum subtracted data to allow better assessment of the residuals.

The aforementioned four [O III] doublets, as well as the results of applying the transform to them, are shown in Fig. 2. The [O III] doublets are significantly merged in some of the cases shown here. The transform separates the two lines, revealing the underlying line shape. Note that the transform does not alter the peak of the line. The transform allows recovery of the tail of the line, revealing the high-velocity structure in the outflow. The example of the double-peaked emission line (bottom left panel) shows that the transform can recover even complex line shapes reliably. The method therefore enables study of the complex kinematics in powerful outflows without having to rely on fitting models with a large number of free parameters.

For comparison, the results of applying a conventional fitting procedure to resolve the doublets are also shown in Fig. 2. The fitting procedure we employed is similar to those commonly employed in the field, and involved modelling each line of the doublet as either two (similar to e.g. Mullaney et al. 2013) or three Gaussians, with the line ratio and wavelength difference of the lines coupled and fixed. The standard `scipy` function `curve_fit` was used for minimization.

The top row in Fig. 2 corresponds to the most strongly merged doublets. It is not obvious what the 'true' underlying line shape is for these cases. Hence we first discuss weakly merged doublets in the figure (bottom row), where the underlying line shape is more 'obvious'. Here we see that the line shapes obtained from the transform appear correct: they are in excellent agreement with those of the 5007 Å line. By contrast, the two-Gaussian-component fits fail to reproduce key features of the line shape. While the three-component fits reproduce the line shape better, significant convergence issues start to emerge; all four doublets required the starting parameters to be modified to achieve converged fits. With regards to the strongly merged doublets (top row), similar to the weakly merged doublets, the transform yields a line shape which closely resembles that of the 5007 Å line, while obtaining reasonable-looking fits with the fitting procedure is more difficult. Additionally, for the
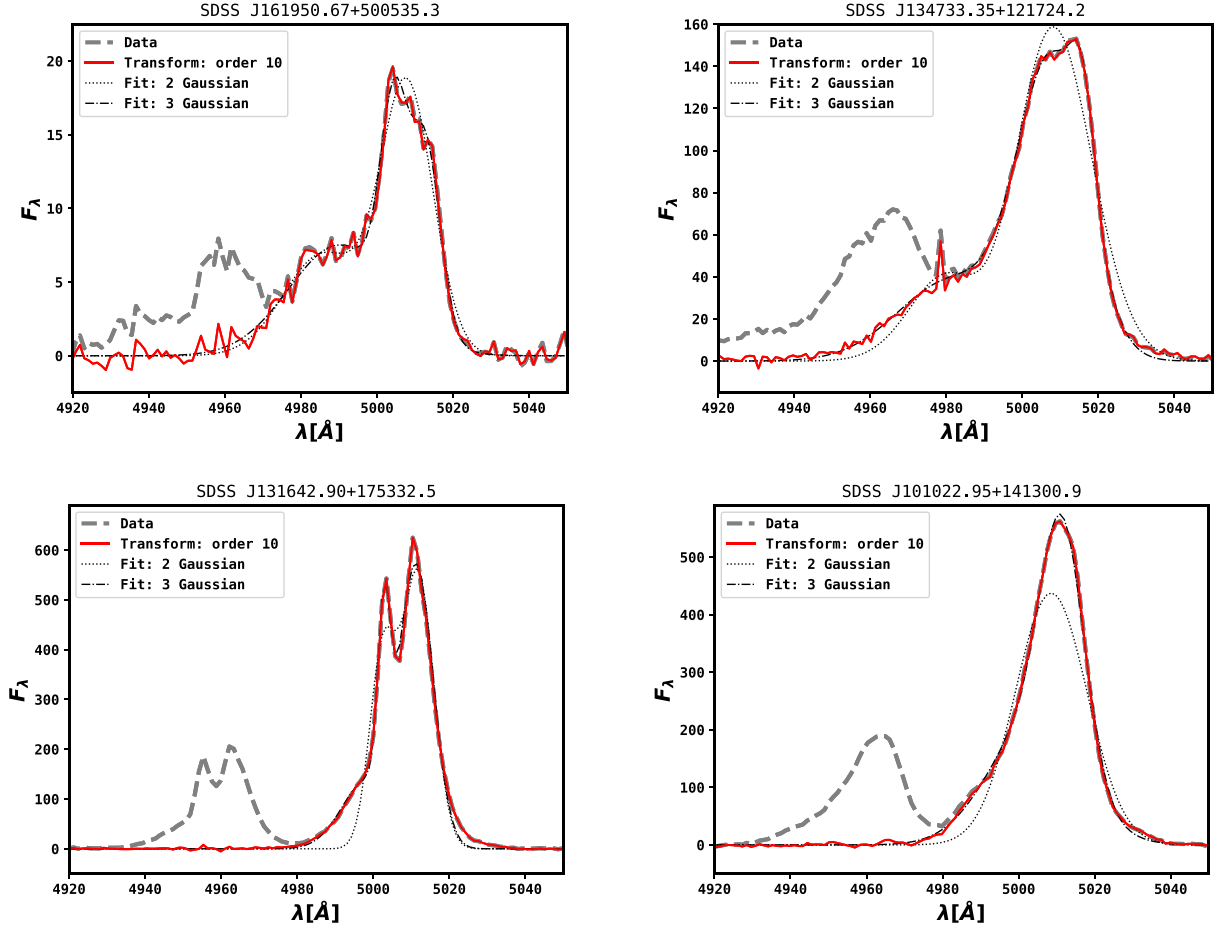
Figure 2. Application of the transform to merged [O III] emission for a number of SDSS AGN with strongly merged [O III] emission. The data were background-subtracted before applying the transform. We assumed a line ratio $R = 1/3$. For comparison, fits with two and three pairs of Gaussians are shown as dotted and dash dotted black lines, respectively. For the fit, we show the solution for the main line, rather than the full doublet for better comparison with the result from the transform; note that fits are performed for the full doublet.

noisier line profile (upper left), the three-component fitting procedure required considerable modification of starting parameters to achieve a good fit. By contrast the transform performs equally well for lower S/N data. This can be difficult to achieve in large samples.

This comparison highlights the key strength of the transform over conventional fitting procedures: it reproduces line shapes without the need for carefully choosing input parameters or dealing with issues such as convergence or choice of initial conditions which be-devil conventional fitting procedures. The transform also remains robust for lower S/N data that can be challenging for modelling. Moreover the transform involves only a single operation on the data array per order, regardless of the complexity of the line shape. Hence the transform will outperform fitting procedures that involve optimization over many free parameters – we note that using many free parameters are necessary to capture complex line shapes.

All further applications of the transform in this work are to simulated data, which has the advantage over real data that the correct solution is known. Simulated data thus affords a more rigorous test of the transform. Specifically, we test the performance of the transform in the presence of continuum contamination, line contamination, noise, and when the line ratio has an error.

4 CONSIDERATIONS FOR USE OF METHOD

The transformation defined in equation (3) can be applied to a doublet $y(x)$, comprised of two merged, geometrically similar lines, in order to extract the individual lines $f_1(x)$ and $f_2(x)$. We showed in Fig. 1 that this method recovers ideal, noise-free data. In this section, we will discuss the numerical implementation of the method, as well as the effect of issues in real data: specifically, contamination from noise, background, and errors in the line ratio R and peak separation δ.

It should be noted that the spectra this method will be applied to are given either as a function of wavelength or frequency. Earlier, we explicitly assumed that the lines were geometrically similar in wavelength. The emission lines are, however, similar in velocity v, not wavelength λ. Now, for a source with rest wavelength λ_{rest}, the mapping between observed wavelength λ and source velocity v is given by (for $v \ll c$)

$$\lambda = \left(\frac{v}{c} + 1\right) \lambda_{\mathrm{rest}}, \tag{5}$$

where c is the speed of light. Hence, since the two lines have different values of λ_{rest}, there is no mapping between λ and v common to both lines. A result of this is that if the two lines are geometrically similar in velocity then they are not geometrically similar in wavelength, and vice versa. Hence our earlier assumption of similarity
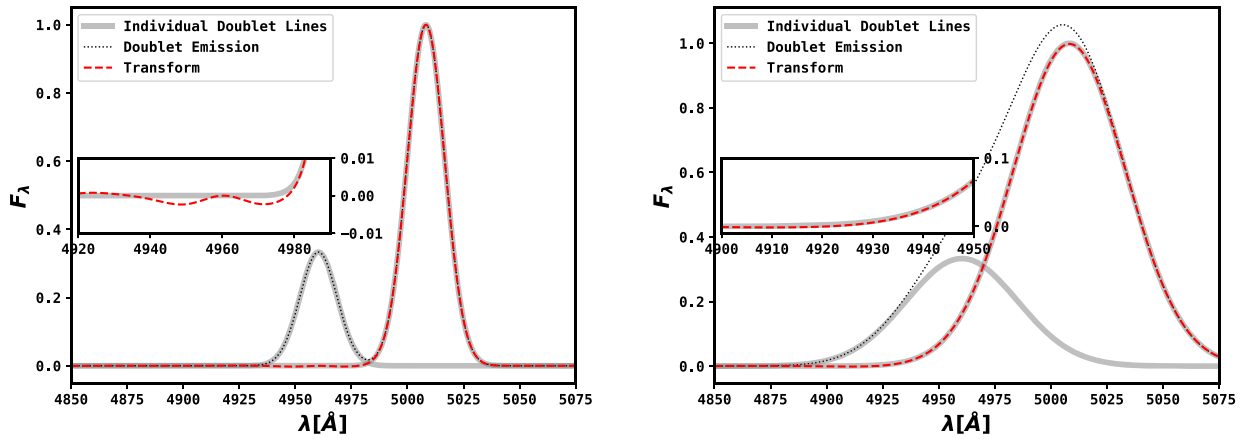
Figure 3. Effect of the fact that the wavelength does not directly map onto the velocity. The example simulates typical [O III] 5007/4959 doublets, the test case for this method. The left-hand panel shows a line width of 500 km s^{-1}. Minimal residuals in the tail of the line are apparent. The right-hand panel shows a well-merged doublet, where no residuals are seen as a result of the use of wavelength instead of velocity. The line ratio of the simulated doublet is 1/3, the separation between the lines is $\delta = 1$ and the width of each line is $\sigma = 0.5$.

in wavelength, which underpins the transform, is invalid. However, the transform remains valid for all intents and purposes if $\lambda_1 \approx \lambda_2$. This is shown in Fig. 3, where doublets, geometrically similar in velocity, have been transformed into the wavelength frame before the transform is applied to extract the underlying line shape. It can be seen from the figure that the resulting residuals are minimal if the velocity separation of the line is comparable to the width of the line (right-hand panel), and only become noticeable for line widths much smaller than the separation (left-hand panel). Even in this case, the errors are negligible ($\sim$1 per cent for a realistic case shown in Fig. 3) and affect only the tail of the lines. These results imply that if the method is to be used for doublets comprised of well-separated lines, each line should separately be transformed to the velocity frame. However, the case of very well separated lines is not the intended use case of this transform, and discussion of this is beyond the scope of this paper.

We now discuss how limitations such as noise, background, and errors in the assumed line ratio, affect the performance of the transform.

4.1 Truncation and convergence

As can be seen from its definition, equation (3), the transform involves an infinite number of iterations. Each term m in the sum corresponds to a shifted and scaled version of $g(x)$, with R^m the scale factor and $m\delta$ the shift applied to $g(x)$. In numerical implementations of the transformation, it is of course necessary to truncate the infinite summation. We refer to the term $m = M$ at which the summation is truncated as the *order*, i.e. terms in equation (3) with $m > M$ are omitted. Thus, in numerical implementations equation (3) becomes

$$\tilde{g}(x) \equiv \sum_{m=0}^{M} (-1)^m R^m g(x + m\delta). \tag{6}$$

We henceforth assume that the transform is defined by the above equation as opposed to equation (3).

For application to the data, we therefore need to know by which M the calculation should be truncated so that convergence is achieved. Recall that convergence and the order of the transform have been demonstrated in Fig. 1 (left-hand panel). This doublet was designed to be representative of real-world [O III] doublets (see Section 3). We

use the same doublet when we examine the effects of background, noise, and uncertainty in R and δ below.

Fig. 1 shows the working of the method, as described in Section 2: the data is shifted and subtracted from itself, the increasing iterations correcting for the oversubtraction (which gives rise to the residuals shown in the inset figure). As can be seen from the figure, the transform corrects for residuals of the previous iterations further and further from the peak of the main line. Choosing a particular order M means that the transform is converged over a wavelength range $\approx M\delta$ downstream from the main line (where downstream means in direction of the minor line). The order should therefore be chosen keeping in mind the wavelength range of interest. Similarly, to avoid edge effects, if applying the transform to the order of M, a background of width $M\delta$ should be included (this is discussed in detail in Section 4.2).

Note that the transform converges only for $R < 1$. The main line should therefore always be chosen so that the line ratio is smaller than one, which can be achieved by reordering the lines. $R = 1$ constitutes a special case. Here, the transform does not converge since, unlike for the case of $R < 1$, successive iterations always perform corrections of the same magnitude. However, the transform still yields converged results over a wavelength range $M\delta$ from the peak of the main line. Choosing a suitable M, as described generally above, is therefore of greater importance for $R = 1$.

In summary, the order M should be chosen keeping in mind the wavelength range $\Delta\lambda$ of interest so that $\Delta\lambda \approx M\delta$ ($M\delta$ both up- and downstream of the main line).

We emphasize that due to the simplicity of the method, computation time is minimal, irrespective of order. Furthermore, note that while we have modelled the line shape here as a Gaussian, we emphasize that our method is general in that it works for *any* line shape; the method makes no assumption about the underlying line shape, e.g. that it is a Gaussian.

4.2 Continuum contamination

Generally, spectral data will have a non-zero continuum. This can be fit for and subtracted, but residuals are likely to remain. Here, we discuss how continuum contamination affects the performance of the transform. In the following, we refer to continuum contamination from any source as background.
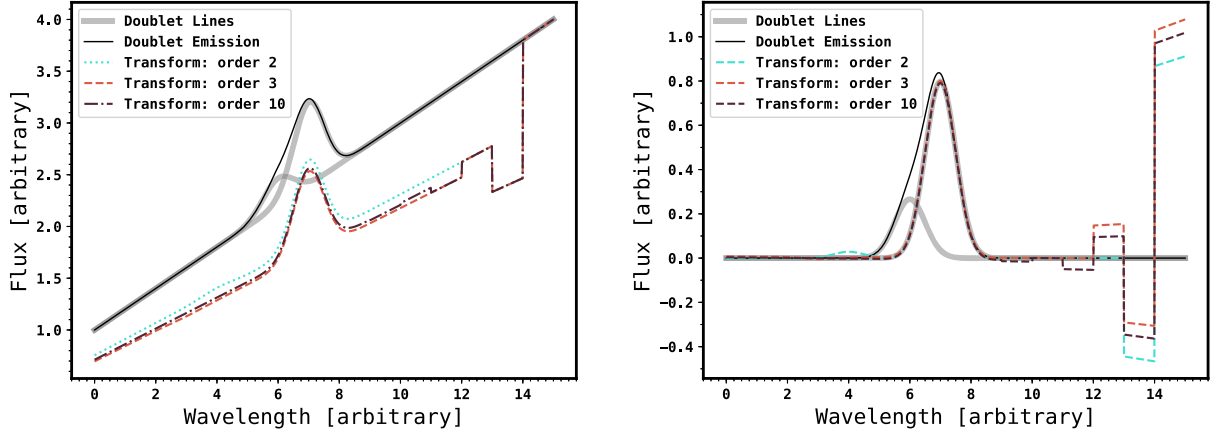
Figure 4. Simulations showing the performance of the transform for imperfect background subtraction. Left-hand panel: A linear background has been added to the signal (the model doublet, described in the main text). Right-hand panel: Data to which the transform has been applied, identical to the data in the left, but with the background fit and subtracted. The extracted line shapes are correct, regardless of the presence of a linear background in the signal. The line ratio of the simulated doublet is 1/3, the separation between the lines is $\delta = 1$, and the width of each line is $\sigma = 0.5$.

To account for noise or a background, equation (1) can be generalized to

$$y(x) = f_1(x) + f_2(x) + \eta(x), \qquad (7)$$

where $\eta(x)$ describes the contribution to the observed doublet due to noise and a background. We now investigate how the addition of $\eta(x)$ affects the ability of the transform to extract the singlet $f_1(x)$ (equation 4).

We begin with a general result. Applying the transform to equation (7), and exploiting equation (4) and the linearity of the transform (see Appendix B2), it follows that

$$\tilde{y}(x) = f_1(x) + \tilde{\eta}(x). \qquad (8)$$

Hence if the doublet exhibits noise or a background, then applying the transform still obtains the singlet $f_1(x)$ but with the transform of the background, $\tilde{\eta}(x)$, added to it.

Ideally the background would be perfectly subtracted before the transform is applied to the doublet. If it is not, we can generally approximate the background as linear across the doublet. For this reason we now consider how the presence of a linear background affects the results. A linear background corresponds to $\eta(x) = Ax + C$, where A and C are constants, and in Appendix B3 it is shown that its transform is given by

$$\tilde{\eta}(x) = A'x + C', \qquad (9)$$

where

$$A' = A \frac{1 - (-R)^{M+1}}{1 + R} \qquad (10)$$

and

$$C' = C \frac{1 - (-R)^{M+1}}{1 + R}$$
$$- A\delta \frac{R}{(1 + R)^2}[1 + M(-R)^{M+1} - (M + 1)(-R)^M]. \qquad (11)$$

Recall that M is the order of the transform; see equation (6). Hence a linear background becomes a linear background in the transform, albeit shifted and rescaled such that the gradient A' and intercept C' differ from the background in the raw data.

Crucially, the nature of the line obtained from the transform is unaffected by the presence of the linear background; the shifted and

rescaled background is simply added to the singlet $f_1(x)$ extracted from the transform. Thus $f_1(x)$ (and hence $f_2(x)$ via equation 2) can be extracted from $\tilde{y}(x)$ by subtracting the transformed background (see equation 8).

Fig. 4 shows the effect of applying the transform to our model doublet in the presence of a linear background. As can be seen from the left-hand panel of Fig. 4, applying the transform to the data yields the single line plus a linear background with a different gradient and intercept, as predicted by the above equations. Note, however, that there are edge effects due to the fact that the simulated array is not of infinite length; for increasing orders of the transform, we observe edge effects due to the fact that the shifted data contains no information beyond the limit of the data, and this boundary is shifted further and further left. This effect is visible only when a non-negligible background is present. This should be taken into account when choosing the wavelength range as well as the order M of the transform (see also Section 4.1). Specifically, if applying the transform with order M, a background of width $M\delta$ should be included upstream to avoid the edge effect adversely affecting features of interest.

In the right-hand panel of the figure the linear background in the transformed data has been fit and removed, leaving a single line which is indistinguishable from the 'true' line. Imperfect background subtraction therefore does not affect the transform. The transform is robust to background contamination.

4.3 Noise

Noise will be present in all astronomical data sets. We now consider the effect of noise. Specifically, we consider *Gaussian additive noise*: $\eta(x)$ is a random variable with mean 0 and variance σ^2, i.e. $\langle \eta(x) \rangle = 0$ and $\langle \eta(x)^2 \rangle = \sigma^2$ for all x. In this case it can be shown (see Appendix B4) that, for all x, the mean and variance in $\tilde{\eta}(x)$ are given by

$$\langle \tilde{\eta}(x) \rangle = 0 \qquad (12)$$

and

$$\langle \tilde{\eta}(x)^2 \rangle = \sigma^2 \left[\frac{1 - R^{2(M+1)}}{1 - R^2} \right], \qquad (13)$$
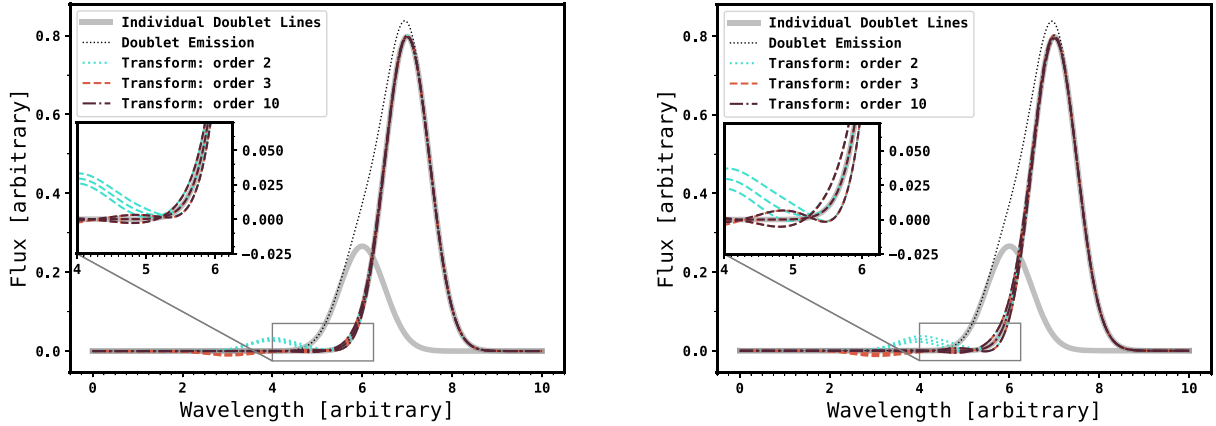
Figure 5. Effect of applying the transform with an incorrect line ratio R. The left-hand panel corresponds to a 5 per cent error in R, the right one a 10 per cent error. In both cases the 'true' line ratio of the simulated doublet is 1/3, the separation between the lines is $\delta = 1$, and the width of each line is $\sigma = 0.5$.

respectively, where we have assumed that $R < 1$. Equations (12) and (8) imply that Gaussian additive noise does not 'distort' the singlet obtained from the transform:

$$\langle \tilde{y}(x) \rangle = f_1(x). \tag{14}$$

However equation (13) implies that the fluctuations in the noise are always magnified by the transformation, with larger magnifications as M is increased (with a limiting value of $1/(1 - R^2)$ as $M \to \infty$) or as R approaches 1. Though, as mentioned above, this does not affect the shape of the single line obtained from the transform, and thus Gaussian additive noise will not cause biases in line shape analysis. To restate, the transformation will recover the correct line shape in the presence of Gaussian additive noise (and assuming no other complications such as contamination from other emission lines, e.g. contamination from other emission lines) regardless of the signal-to-noise ratio.

For $R = 1$ equation (12) still applies, but $\langle \tilde{\eta}(x)^2 \rangle$ is instead given by (see Appendix B4)

$$\langle \tilde{\eta}(x)^2 \rangle = \sigma^2 (M + 1). \tag{15}$$

This diverges as $M \to \infty$. However, this problem is sidestepped since, as discussed in Section 4.1, in practice we consider only a finite wavelength range, and hence it is sufficient to use a finite M.

However, while equations (13) and (15) imply that the magnitude of the noise in the transform is the same for all x, these equations have been derived assuming the presence of noise at all wavelengths. In practice we would be considering a finite wavelength range, and the obvious course of action is to assume that there is no noise out with the considered range. In this case noise is not present at all wavelengths, and equations (13) and (15) break down. The result is that $\langle \tilde{\eta}(x)^2 \rangle$ becomes x-independent. This is an edge effect that could significantly complicate the task of correctly fitting the underlying line shape. This problem can, however, be sidestepped by using a sufficiently large wavelength range so that the line shape is converged before the distance $M\delta$ is comparable to the distance from the main peak to the edge.

In Fig. 1 (right-hand panel) the transform is applied to our model doublet, but with noise with standard deviation $\sigma_{\text{noise}} = 0.01$ added to the data. Crucially, the underlying line shape is still recovered reliably despite the noise. Moreover, as in the noise-free data, $M = 10$ is sufficient to obtain converged results. As mentioned above,

the magnitude of the noise will be magnified by the transform. For the R, δ, M, and σ_{noise} in the simulated data, equation (13) predicts that the noise in the transformed data will have standard deviation $\tilde{\sigma}_{\text{noise}} = 0.0106$, which corresponds to a 6 per cent magnification. This increase is too small to be noticeable in the figure. Further results for the case $R = 1$, where the magnification in the noise is substantial, are presented in Appendix B5.

4.4 Uncertainty in line ratio and peak separation

The appropriate values of R and δ to use in the transform (equation 6) might not be known exactly. Hence it is desirable to know the effect of applying the transform with slightly incorrect values of R and δ. In practice the line separation will be known to a high accuracy; it is the line ratio that has the dominant source of uncertainty. Hence we focus on the effect of uncertainty in R here.

Fig. 5 shows the effect of applying the transform to our model doublet, but using slightly 'wrong' values of R, namely 5 and 10 per cent above and below the true value of R. It is clear from the figure that the effect of using slightly incorrect values of R here is small. To elaborate, here, when the order of the transform is sufficiently high to yield convergence (i.e. by order $M = 10$), using a value of R which differs from the true value by 10 per cent yields errors with magnitude $\lesssim 2$ per cent of the maximum in the doublets. However, such small errors may not result for all doublets. To facilitate further investigation of this, equations quantifying the effect of using incorrect R and δ in the transform are provided in Appendix B6.

4.5 Contamination from emission lines

Doublet emission might be contaminated by unassociated line emission, either from the astrophysical source itself or from sky emission. The resulting transform will consist of the transform of the doublet, with the transform applied to any contaminating line emission superimposed on top of it (this follows from the linearity of the transform; see Appendix B2). In other words, the contaminating line will be shifted and subtracted from the data set, resulting in residuals in the transform with separation δ and amplitude R times the original amplitude of the contaminating line.
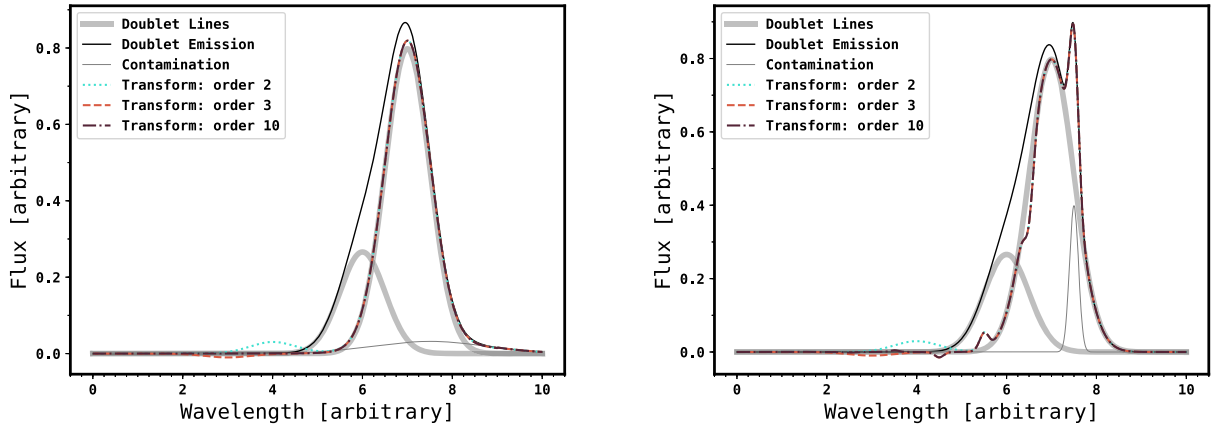
Figure 6. Simulations showing the effect of line contamination on the transform. As in Fig. 1, $R = 1/3$, $\sigma = 0.5$, and $\delta = 1$. We show two extreme cases, the left-hand panel shows contamination from a line much broader than the doublet line ($\sigma = 1.25$) and the right-hand panel shows contamination from a narrow emission line ($\sigma = 0.1$). Both contaminating lines have 10 per cent of the flux in the main line. We see that the broader emission line does not strongly affect the transform, with the line shape recovered well. The narrow emission line results in residuals with decreasing amplitude with separations δ, as the transform shifts and subtracts the contaminating line from the contaminated doublet.

This is illustrated in Fig. 6, for cases of broad (left-hand panel) and narrow (right-hand panel) contaminating lines with 10 per cent flux of the main line. It can be seen that the residuals resulting from the broad contaminating line are minimal; while the transform does not remove the contaminating line, it does not increase the error resulting from contamination further. By contrast the narrow contaminating line introduces noticeable residuals at intervals of δ from the location of the contaminating line, which decrease in amplitude with distance from it. As for the broad line, the transform does not remove the narrow contaminating lines. The transform will therefore result in a residual in the presence of line contamination, but the amplitude of this is always smaller or equal to the initial line.

All of the above observations should be borne in mind if the transform is to be applied in the presence of contaminated line emission.

5 SUMMARY AND DISCUSSION

Here, we have presented a method (Section 2) that allows the non-parametric separation of merged doublet emission lines when the line ratio is known. The method is generally robust for noisy data and recovers the emission line with high precision without introducing significant noise (Section 4). The transform is also robust to background contamination.

More specifically, we have presented mathematical derivations and simulation results which show that the method is robust in the presence of a linear background (Section 4.2), and does not inflate noise considerably (Section 4.3) unless the line ratio is close to 1. The method is also numerically fast, enabling it to be easily applied to large data sets for further processing without the need to fit emission lines. Another appealing feature of the method is that it is easily implemented; PYTHON source code is included in Appendix A. We have also applied the method to extreme AGN-driven outflows (see Section 3, Fig. 2) and show that it allows line shapes to be recovered, including the scientifically interesting high-velocity wings that trace the most high-velocity structures in the outflow.

To conclude this paper we summarize the requirements for reliable use of the method, and suggest possible future applications.

5.1 Requirements for use of method

The method presented here can be a powerful tool for decomposing doublet emission. However the following limitations should be kept in mind when applying the method.

(i) Sampling: The data needs to be sampled appropriately both for any features to be recovered as well as the wavelength difference δ.

(ii) Noise levels and S/N: For low values of R, noise is only mildly increased by the transform. The noise levels in the transform diverge as R approaches 1 for infinite number of iterations ($M \rightarrow \infty$). However, since the wavelength range $\Delta\lambda$ used in practice is limited, and therefore the number of iterations required for convergence of the transform is limited to $M \approx \Delta\lambda/\delta$, the noise level does not reach infinity. Noise levels should be considered for low signal-to-noise data and line ratios $R \sim 1$. The increase in noise level is discussed in Section 4.3 and in more general terms in Appendix B4. There is therefore no formal signal-to-noise requirement for the use of the method, and while the transform will increase noise levels for typical cases ($R < 1$, $M < 10$), the increase is small, only by a few per cent.

(iii) Continuum background subtraction: Flat as well as linear backgrounds do not affect the transform; see Fig. 4 as well as Section 4.2 and Appendix B3.

(iv) Line ratios: The line ratio needs to be known; if the assumed line ratio is incorrect, the transform does not return the underlying line shape. The error introduced is discussed in Section 4.4 and in Appendix B6. For small errors ~5 per cent in the line ratio, the resulting errors in the transform are small (<1 per cent; see Fig. 5).

(v) Special case – line ratio $R = 1$: In the case of line ratio 1, the transform does not converge on an infinitely large wavelength range. However, the transform still converges over a limited wavelength range. In this case M should be chosen carefully; see Section 4.1 as well as 4.3.

(vi) Contamination by other emission lines (e.g. Fe II or H β in case of the [O III] doublet) cannot be accounted for using the method. However, this limitation equally applies to any other method that can be used to decompose merged doublets. The resulting residuals are outlined in Section 4.5.

Under these conditions, the method allows decomposition of a doublet for further non-parametric characterization, such

as per centiles or skew measurements. The method also allows separation of stacked spectra, which might not follow ideal parametric functions. It allows the direct comparison of different doublet lines, such as [O III] and [O II], without requiring a parametrization. The method is also ideally suited for visualization purposes for strongly merged doublets, as shown in Section 3, Fig. 2. The strength of the method is its speed and simplicity, which makes it ideal for application to large samples.

5.2 Future applications

As explained above, the method can be used for any doublet emission with suitable data. Here we list a few scientific cases for which the data would be suitable.

(i) [O II] 3727/3730 Å: This doublet is heavily merged, with a separation of only ∼250 km s^{-1}. [O II] however is used for kinematic studies at redshifts for which [O II] is shifted outside of the optical range (Patricio et al. 2018). The line ratio is insensitive to density above a density threshold of 10^4 cm^{-3} and so can be assumed to be fixed for most cases (Draine 2011).

(ii) Mg II 2796/2803 Å is seen in absorption in outflows from AGN. With a velocity separation of only ∼750 km s^{-1}, merged emission lines can be common. If the emission can be assumed to be either optically thick or thin, the line ratio is known (Kovacevic et al. 2010).

(iii) Velocity delay maps are commonly used to study the kinematics of the broad line region (BLR) of AGN (Welsh & Horne 1991). BLR lines show complex profiles; for Mg II, where the doublet has a velocity separation of ∼750 km s^{-1}, the doublet causes significant smearing. If the line is optically thick, the method presented here can be used to decompose BLR emission since the line ratio is known.

(iv) Absorption line kinematics: While we have discussed and visualized the method in terms of emission lines, the method is equally suitable for absorption-line doublets (such as Mg II or NaD outflows). As for the case of emission lines, contamination from other lines will affect the transform and result in residuals, see Section 4.5 for details.

(v) The transform is also ideal for quick visualization of merged doublet emission without requiring fitting.

Naturally, this list is not complete. Moreover, since the problem of separating doublets comprised of identically shaped singlets is encountered in many fields of science, there is the prospect of the method being applied generally outside of astrophysics.

In summary, we have presented a simple non-parametric method for resolving merged doublet emission. The method is robust to linear background contamination. The method can be used for a wide range of science cases as well as visualization. PYTHON code for implementation is given in Appendix A.

ACKNOWLEDGEMENTS

REFERENCES

Aalto S., Garcia-Burillo S., Muller S., Winters J. M., van der Werf P., Henkel C., Costagliola F., Neri R., 2012, A&A, 537, A44
Abazajian K. N. et al., 2009, ApJS, 182, 543
Antonucci R., 1993, ARA&A, 31, 473
Bae H. J., Woo J. H., 2014, ApJ, 795, 30
Carniani S. et al., 2015, A&A, 580, A102
Cicone C., Feruglio C., Maiolino R., Fiore F., Piconcelli E., Menci N., Aussel H., Sturm E., 2012, A&A, 543, A99
Crenshaw D. M., Schmitt H. R., Kraemer S. B., Mushotzky R. F., Dunn J. P., 2010, ApJ, 708, 419
Crenshaw D. M., Kraemer S. B., Schmitt H. R., Fischer T. C., Gagne J., 2014, A&AS, 223, 251.12
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Draine B. T., 2011, Physics of the Interstellar and Intergalactic Medium. Princeton University Press, Princeton
Fabian A. C., 2012, ARA&A, 50, 455
Faucher-Giguere C. A., Quataert E., 2012, MNRAS, 425, 605
Greene J. E., Zakamska N. L., Smith P. S., 2012, ApJ, 746, 86
Heckman T. M., Thompson T. A., 2017, Handbook of Supernvovea, Springer International Publishing AG, p. 2431, ISBN:978-3-319-21845-8
Heckman T. M., Armus L., Miley G. K., 1990, ApJS, 74, 833
Hopkins P. F., Hernquist L., Cox T. J., Keres D., 2008, ApJS, 172, 356
Junkkarinen V. T., Burbidge E. M., Smith H. E., 1983, ApJ, 265, 51
Kovacevic J., Popovic L. C., Dimitrijevic M. S., 2010, ApJS, 189, 15
Liu G., Zakamska N. L., Greene J. E., Nesvadba N. P. H., Liu X., 2013, MNRAS, 430, 2327
Mullaney J. R., Alexander D. M., Fine S., Goulding A. D., Harrison C. M., Hickox R. C., 2013, MNRAS, 433, 622
Patrício V. et al., 2018, MNRAS, 477, 18
Rupke D. S. N., Veilleux S., 2011, ApJ, 729, L27
Rupke D. S. N., Veilleux S., 2013, ApJ, 768, 75
Rupke D. S., Veilleux S., Sanders D. B., 2005, ApJ, 632, 751
Schmitt H. R., Donley J. L., Antonucci R. R. J., Hutchings J. B., Kinney A. L., Pringle J. E., 2003, ApJ, 597, 768
Urry C. M., Padovani P., 1995, PASP, 107, 803
Vestergaard M., Wilkes B. J., 2001, ApJS, 134, 1
Welsh W. F., Horne K., 1991, ApJ, 379, 586
Zakamska N. L. et al., 2016, MNRAS, 459, 3144
Zakamska N. L., Greene J. E., 2014, MNRAS, 442, 784

APPENDIX A: PYTHON IMPLEMENTATION OF ALGORITHM

The function to implement the transform in PYTHON is presented below. As an input, the function takes the wavelength array, the flux array, δ, R, and optionally the order M (denoted in the code below as `delta`, `R` and `order`, respectively). `flux.scipy.interpolate` is used to interpolate the spectrum and allow evaluation of the spectrum after the wavelength shift is applied. However such interpolation is not an essential aspect of the method, and it would be possible to write a function which did not rely on this.

```
import numpy
import scipy.interpolate


def twin_peaks(wl, flux, delta, R, order=10):
    ''''''
    A function to separate doublet lines.
    The outflows are not what they seem.
    x: input wavelength array
    y: input flux array
    delta: offset between the two lines,
    same unit as x (positive if shorter
    wavelength line has lower flux)
```

```
R: ratio (<1) between emission lines,
order: order to which to correct for,
default is 10, note that higher order
have no effect if wavelength range is
shorter than order*offset
returns: flux after transform
'''''''
#Interpolation to allow evaluation of the
#spectrum at different wavelength after
#shifting the array
rawf = scipy.interpolate.interp1d(wl, flux,
 bounds_error=False, fill_value=0)
outf = numpy.zeros_like(flux)
outf += flux
for i in range(1, order+1):
 #Spectrum is shifted and added
 #to original dataset.
 outf += (-1)**i * R**(i)
 * rawf(wl + i * delta)
 return outf
```

We now outline how this function can be applied. As a first step, the spectrum needs to be loaded. The function requires a wavelength and flux array with sufficient wavelength resolution. No error data are required. The wavelength and flux arrays can be assigned in the usual manner, e.g. `wl, flux = ...`,

where `...` signifies code specific to the problem at hand. In this example, we are applying the transform to an [O III] doublet with a wavelength difference of 47.945 Å and a line ratio of 1/3, as in the test cases shown in Section 3. Since the stronger line has a longer wavelength, the wavelength difference `delta` must be given as *positive*. For a doublet where the shorter wavelength line is stronger, the wavelength difference `delta` would be given as *negative*. The line ratio `R` should always be <1. Accordingly, for this case `delta=47.945` and `R=1./3` in the function. The only other information required by the transform is the order. In this example, we choose 10, which will apply the transform over `10*delta≈500` Å. Hence transform in this example is applied as follows:

```
transform = twin_peaks(wl=wl, flux=flux,
 delta=47.945,
 R=1./3,
 order=10)
```

The function returns the transformed spectra, `transform`, with the same array shape as the input spectrum. A comparison between the original spectrum and transform could be plotted as follows:

```
import pylab
pylab.plot(wl, flux)
pylab.plot(wl, transform)
```

APPENDIX B: FURTHER RESULTS

B1 Derivation of equation (4)

To derive equation (4), we first substitute equation (2) into equation (1). This gives

$$y(x) = f_1(x) + R f_1(x + \delta), \tag{B1}$$

which becomes

$$f_1(x) = y(x) - R f_1(x + \delta) \tag{B2}$$

after rearranging. Now, substituting this equation into itself recursively gives

$$f_1(x) = y(x) - R\left[y(x + \delta) - R f_1(x + 2\delta)\right] \tag{B3}$$

$$= y(x) - R y(x + \delta) + R^2 f_1(x + 2\delta) \tag{B4}$$

$$= y(x) - R y(x + \delta) + R^2\left[y(x + 2\delta) - R f_1(x + 3\delta)\right] \tag{B5}$$

$$= y(x) - R y(x + \delta) + R^2 y(x + 2\delta) - R^3 f_1(x + 3\delta) \tag{B6}$$

$$= y(x) - R y(x + \delta) + R^2 y(x + 2\delta) - R^3\left[y(x + 3\delta)\right. \tag{B7}$$

$$\left. - R f_1(x + 4\delta)\right] \tag{B8}$$

$$= \ldots \tag{B9}$$

$$= y(x) - R y(x + \delta) + R^2 y(x + 2\delta) - R^3 y(x + 3\delta) \tag{B10}$$

$$+ R^4 y(x + 4\delta) - R^5 y(x + 5\delta) - \ldots. \tag{B11}$$

This can be expressed concisely as

$$f_1(x) = \sum_{m=0}^{\infty} (-1)^m R^m y(x + m\delta), \tag{B12}$$

which yields equation (4) after comparing the right-hand side with equation (3).

B2 Linearity of the transform

A pleasing aspect of the transformation is that it is linear, i.e. for a constant C and arbitrary functions g and h,

$$\mathcal{T}[Cg] = C\mathcal{T}[g] \tag{B13}$$

and

$$\mathcal{T}[g + h] = \mathcal{T}[g] + \mathcal{T}[h], \tag{B14}$$

where $\mathcal{T}[g] \equiv \tilde{g}(x)$ denotes the result of applying the transform to $g(x)$. These equations can be derived straightforwardly from equation (3), or, for the case of finite order M, equation (6).

B3 Derivation of equation (9)

Consider equation (8) with $\eta(x) = Ax + C$. From the definition of the transformation with finite order M, equation (6), we have

$$\tilde{\eta}(x) = \sum_{m=0}^{M} (-1)^m R^m [A(x + m\delta) + C] \tag{B15}$$

$$= (Ax + C) \sum_{m=0}^{M} (-R)^m + A\delta \sum_{m=0}^{M} m(-R)^m. \tag{B16}$$

This becomes equation (9) after applying the well-known equation

$$\sum_{m=0}^{M} r^m = \frac{1 - r^{M+1}}{1 - r}, \tag{B17}$$

to the first term, and an analogous equation (obtained from differentiating the above equation)

$$\sum_{m=0}^{M} m r^m = \frac{r}{(1-r)^2}[1 - (M+1)r^M + M r^{M+1}] \tag{B18}$$

to the second term. Note that both these equations are only valid for $r \neq 1$. Hence in deriving equation (9) from equation (B15) we have tacitly assumed that $(-R) \neq 1$, which is inconsequential since we are only interested in $R > 0$.

B4 Derivation of equations (12) and (13)

Consider if $\eta(x)$ is a random variable corresponding to additive Gaussian noise with mean and variance $\langle \eta(x) \rangle = 0$ and $\langle \eta(x)^2 \rangle = \sigma^2$ for all x. Taking the expected value of $\tilde{\eta}(x)$ yields

$$\langle \tilde{\eta}(x) \rangle = \sum_{m=0}^{M} (-1)^m R^m \langle \eta(x + m\delta) \rangle, \tag{B19}$$

where we have used equation (6) and exploited the linearity of the expected value (i.e. $\langle (a+b) \rangle = \langle a \rangle + \langle b \rangle$). Equation (12) follows from this since $\langle \eta(x) \rangle = 0$ for all x.

Now, since $\langle \tilde{\eta}(x) \rangle = 0$, the variance in $\tilde{\eta}(x)$ is given by $\langle \tilde{\eta}(x)^2 \rangle$. Noting that

$$\tilde{\eta}(x)^2 = \sum_{m,n=0}^{M} (-1)^m R^m \eta(x + m\delta)(-1)^n R^n \eta(x + n\delta) \tag{B20}$$

$$= \sum_{m,n=0}^{M} (-1)^{m+n} R^{m+n} \eta(x + m\delta)\eta(x + n\delta), \tag{B21}$$

it follows that

$$\langle \tilde{\eta}(x)^2 \rangle = \sum_{m,n=0}^{M} (-1)^{m+n} R^{m+n} \langle \eta(x + m\delta)\eta(x + n\delta) \rangle, \tag{B22}$$

where again we have exploited the linearity of the expected value.

To proceed further we assume that there are no correlations in the noise over length-scales δ in x, in which case $\langle \eta(x + m\delta)\eta(x + n\delta) \rangle = 0$ for $n \neq m$. With this, all terms $n \neq m$ in the above equation vanish, leaving

$$\langle \tilde{\eta}(x)^2 \rangle = \sum_{m=0}^{M} (-1)^{2m} R^{2m} \langle \eta(x + m\delta)^2 \rangle = \sum_{m=0}^{M} (-1)^{2m} R^{2m} \sigma^2, \tag{B23}$$

where in the last equality we have used the fact that $\langle \eta(x)^2 \rangle = \sigma^2$ for all x. Noting that $(-1)^{2m} = 1$, this can in turn be re-expressed as

$$\langle \tilde{\eta}(x)^2 \rangle = \sigma^2 \sum_{m=0}^{M} (R^2)^m, \tag{B24}$$

which yields equation (13) after using equation (B17), where note that equation (B17) is only valid if $r \neq 1$ and hence equation (13) only applies if $R^2 \neq 1$, or equivalently $R \neq \pm 1$. For the special case $R = 1$ the above equation can be easily shown to be equivalent to equation (15).

B5 Divergence in noise if $R = 1$

Fig. B1 (left-hand panel) shows the effect of the transform on a doublet with line ratio $R = 1$ in the presence of noise. It is clear that the noise is magnified with increasing order of the transformation,

as discussed in Section 4.3. Analogous results for $R = 1/3$ are shown in the right-hand panel of Fig. B1.

The transform can still be used for line ratios $R = 1$, but great care should be taken; the number of iterations applied should be carefully considered.

B6 Uncertainty in R and δ

Here we present an expression that describes the effect of applying the transformation with incorrect values of R and δ. We use R_t and δ_t to denote the 'true' line ratio and peak separation for the doublet under consideration, and R and δ to denote the line ratio and peak separation actually used in the transform. Assuming that R and δ are 'close' to the true values, i.e. that $\Delta R \equiv (R - R_t)$ and $\Delta \delta \equiv (\delta - \delta_t)$ are small, the error in the transform of $y(x)$, given by equations (1) and (2), from the true result $\tilde{y}(x) = f_1(x)$ (equation 4), is as follows:

$$\Delta \tilde{y}(x) = \left[\tilde{f}_1(x; R_t, \delta_t) - f_1(x) \right] \frac{\Delta R}{R_t} + \left[\tilde{f}'_1(x; R_t, \delta_t) - f'_1(x) \right] \Delta \delta, \tag{B25}$$

where $\tilde{g}(x; R, \delta)$ denotes the transform of $g(x)$ using R and δ; and $g'(u)$ denotes the value of the derivative of $g(x)$ with respect to x, evaluated at $x = u$. Note that in the above $\tilde{f}'_1(x; R_t)$ is the transform of the derivative of $f_1(x)$ (and not the derivative of the transform). Note also that this equation applies for $M \to \infty$, i.e. $\tilde{g}(x; R, \delta)$ is defined as the right-hand side of equation (3).

The derivation of this equation is as follows. Expanding $\tilde{y}(x)$ as a Taylor series about $R = R_t$ and $\delta = \delta_t$ gives

$$\Delta \tilde{y}(x) \equiv \tilde{y}(x; R, \delta) - \tilde{y}(x; R_t, \delta_t)$$

$$= \frac{\partial \tilde{y}(x; R, \delta)}{\partial R} \Delta R + \frac{\partial \tilde{y}(x; R, \delta)}{\partial \delta} \Delta \delta, \tag{B26}$$

where the derivatives in the above are evaluated at $(R, \delta) = (R_t, \delta_t)$, and we have ignored quadratic and higher-order terms in ΔR and $\Delta \delta$. The task now is to evaluate derivatives.

We begin with $\partial \tilde{y}(x; R, \delta)/\partial R$. Taking the derivative of equation (3) with respect to R gives

$$\frac{\partial \tilde{y}(x; R, \delta)}{\partial R} = \sum_{m=0}^{\infty} (-1)^m m R^{m-1} y(x + m\delta), \tag{B27}$$

which becomes

$$\frac{\partial \tilde{y}(x; R, \delta)}{\partial R} = R_t^{-1} \sum_{m=0}^{\infty} (-1)^m m R_t^m \Big[f_1(x + m\delta_t)$$

$$+ R_t f_1(x + (m+1)\delta_t) \Big] \tag{B28}$$

after using equations (1) and (2), setting $(R, \delta) = (R_t, \delta_t)$, and factorizing. Now, by expanding the right-hand side into two summations, we obtain $\partial \tilde{y}(x; R, \delta)/\partial R = S_1 + S_2$, where

$$S_1 = R_t^{-1} \sum_{m=0}^{\infty} (-1)^m m R_t^m f_1(x + m\delta_t) \tag{B29}$$

and

$$S_2 = \sum_{m=0}^{\infty} (-1)^m m R_t^m f_1(x + (m+1)\delta_t). \tag{B30}$$

S_2 can be expressed as

$$S_2 = R_t^{-1} \sum_{m=1}^{\infty} (-1)^m (m-1) R_t^{m-1} f_1(x + m\delta_t). \tag{B31}$$
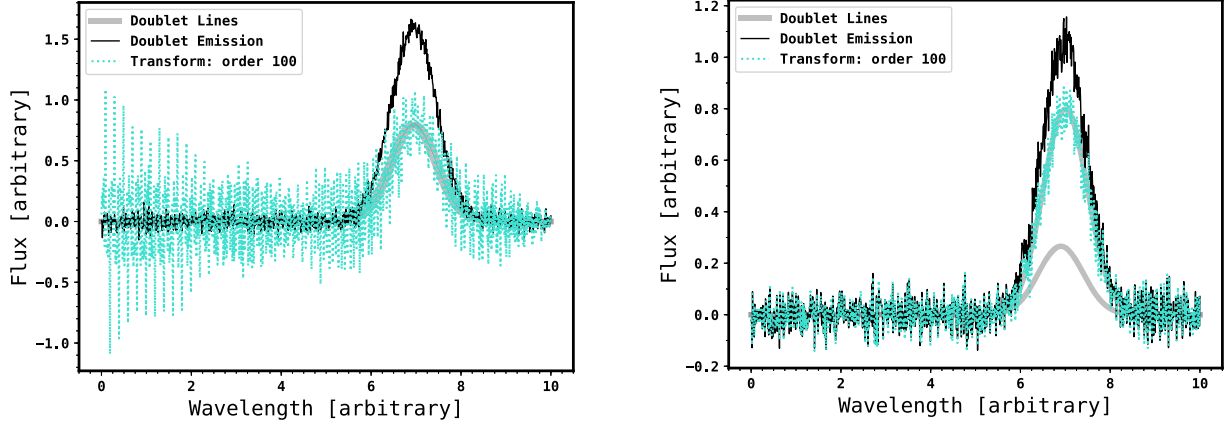
Figure B1. Simulations showing the effect of noise on a transform with a line ratio of $R = 1$ for order $= 100$. Left-hand panel: A doublet comprised of two Gaussians with $R = 1$, $\sigma = 0.5$, and $\delta = 0.1$. Divergence of the noise is noticeable over a large number of iterations in this case. The right-hand panel shows the same noise level and line separation, but with a line ratio of $R = 1/3$. The noise does not diverge in this case.

Hence we obtain

$$\frac{\partial \tilde{y}(x; R, \delta)}{\partial R} = S_1 + S_2 = R_t^{-1} \sum_{m=1}^{\infty} (-1)^m R_t^m f_1(x + m\delta_t) \quad \text{(B32)}$$

after combining the two sums and exploiting the fact that the $m = 0$ term in S_1 is 0. Moving the $m = 0$ term in equation (3) to the left-hand side and applying the resulting equation to the above equation, we obtain

$$\frac{\partial \tilde{y}(x; R, \delta)}{\partial R} = \frac{1}{R_t} \left[\tilde{f}_1(x; R_t, \delta_t) - f_1(x) \right]. \quad \text{(B33)}$$

We now turn to $\partial \tilde{y}(x; R, \delta)/\partial \delta$. Taking the derivative of equation (3) with respect to δ gives

$$\frac{\partial \tilde{y}(x; R, \delta)}{\partial \delta} = \sum_{m=0}^{\infty} (-1)^m m R^m y'(x + m\delta), \quad \text{(B34)}$$

where recall that $y'(u)$ denotes the value of $\partial y(x)/\partial x$ evaluated at $x = u$. This becomes

$$\frac{\partial \tilde{y}(x; R, \delta)}{\partial \delta} = \sum_{m=0}^{\infty} (-1)^m m R_t^m \left[f_1'(x + m\delta_t) \right.$$
$$\left. + R_t f_1'(x + (m + 1)\delta_t) \right] \quad \text{(B35)}$$

after using equations (1) and (2), and setting $(R, \delta) = (R_t, \delta_t)$. Note that the right-hand side of the above equation resembles equation (B29). The difference is that the above has $f_1'(x)$ instead of $f_1(x)$, and is missing the factor R_t^{-1} present in equation (B29). With this in mind, following the same procedure we used above to derive equation (B34) from equation (B29), we obtain

$$\frac{\partial \tilde{y}(x; R, \delta)}{\partial R} = \tilde{f}_1'(x; R_t, \delta_t) - f_1'(x). \quad \text{(B36)}$$

Finally, equation (B26) results from substituting equations (B34) and (B37) into equation (B27).

This paper has been typeset from a TeX/LaTeX file prepared by the author.