Robust Raman Spectral Decomposition with Wavenumber Shifts Parametric Modelling

Mehrdad Yaghoobi,
School of Engineering, Institute for Digital Communications,
The University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh EH9 3JL, UK
Email: m.yaghoobi-vaighan@ed.ac.uk

***Abstract*—Nonlinearity in Raman spectral mixtures caused by wavenumber shifts, has been investigated in this paper. The spectral shifts are mainly caused by the existence of multiple chemicals in the mixtures, with complex molecular interactions, which can change the spectral features of each constituent. While such non-linear behaviour may be negligible in some mixtures, it may lead to incorrect identification of chemicals in some instances. We investigate some real spectra and demonstrate the nature of such nonlinearity in Raman spectra. We then mathematically formulate such spectral behaviour and present an approach to compensate the nonlinearity artifacts. The nonlinearity has been modelled as a smooth transition in a parametric space, which can be locally modelled using first order approximation. Such a first order approximation can be translated to some augmented spectral libraries to be used with a linear generative model. A convex sparse approximation program, with nonlinearity considerations, has finally been introduced to decompose the spectral mixtures. Such decomposition has been used for chemical fingerprinting and quantification. The effect of new approach has been demonstrated with some real and synthetic spectra.**

***Index Terms*—Raman Spectral Decomposition, Spectral Wavenumber Shifts, Nonlinear Sparse Decomposition, First Order Approximation and parametric modelling of functional spaces.**

I. INTRODUCTION

Spectral processing for chemical identification is a popular technique with numerous applications in Pharmaceutical, Defence, Oil and Gas industries. The identification of an unknown chemical in a mixture would be more challenging, when its concentration is low and it has small signature compared to other components of the mixture. Some decomposition techniques have been proposed to separate the spectra to the elementary components to be able to characterise the mixture [1], [2]. The main idea is to use a given set of reference spectra, as the building blocks. Such a collection of reference spectra is called a library and it is usually digitised. The library can be saved as the columns of a matrix $\mathbf{M} \in \mathbb{R}^{d \times N}$, where d is the number of different wavenumbers and N is the number of spectra in the library. Let the measured spectral mixture be $\mathbf{y} \in \mathbb{R}^d$. An approach to describe $\mathbf{y}$ with the library elements is to use a generative model as follows[1],

$$\mathbf{y} = \phi(\mathbf{M}, \boldsymbol{\alpha}) + \mathbf{b} + \boldsymbol{\omega} \tag{1}$$

where function ϕ is generally a non-linear mixing model, $\boldsymbol{\alpha} = [\alpha_i]_{i=1:N} \in \mathbb{R}_+^N$ is a vector with nonnegative values, presenting the contributions of the spectra, $\boldsymbol{b} \in \mathbb{R}_+^N$ is the deterministic artifact and $\boldsymbol{\omega} \in \mathbb{R}^N$ is the measurement noise. The non-linearity of ϕ may be reflected in deviating the library elements from the original location in $\mathbb{R}^d$ to some neighbourhood, which causes the coefficients $\boldsymbol{\alpha}$ changing in the actual representation. In this work, we only consider the effect of nonlinearity on the library elements, which is due to spectral shifts of Raman peaks. Such spectral wavenumber shifts may cause incorrect selection of more library spectra to compensate the shift, or quantification.

If the deterministic noise is morphologically different to the spectra, it can be separated/subtracted using some correction techniques [4]–[6]. In the most popular case, the generative model is approximately linear in the form of [7]:

$$\begin{aligned} \mathbf{y} &= \mathbf{M}\boldsymbol{\alpha} + \boldsymbol{\omega} \\ &= \sum_{i=1}^{N} \alpha_i \mathbf{m}_i + \boldsymbol{\omega}. \end{aligned} \tag{2}$$

The task is now how robustly decompose $\mathbf{y}$, using elements of $\mathbf{M}$, and find an $\boldsymbol{\alpha}$ with only few elements. The sparsity of the decomposition is related to the fact that most mixtures of interests only have a few chemicals. A fast greedy sparse approximation based algorithm has been proposed in [7] to decompose the spectra with non-negative weighted-sum of pure spectra in $\mathbf{M}$.

Raman spectral mixtures are not always an exact weighted sum of library spectra. With an example, it has been demonstrated here that the spectra may have some wavenumber shifts in the mixtures. We chose a mixture of simple chemicals where one chemical has sharp Raman features and the other has broad and weak Raman features, which would be cancelled using baseline correction methods. The baseline corrected spectra of the mixture has been shown in figure 1 with the blue line. The spectra can be approximately represented using the constituent spectra shown in figure 1 with the red line. The original peak has been shifted for 12 cm^{-1}, which is better demonstrated in the zoomed area. This behaviour is a well-known behaviour and the amount of shift is related to the concentration of each

---

[1] Here, we assume that the library of spectra is complete and has all the elementary spectra. If there is an unknown element, we need to consider separating the non-representable residual [3], which is out of scope of this paper.

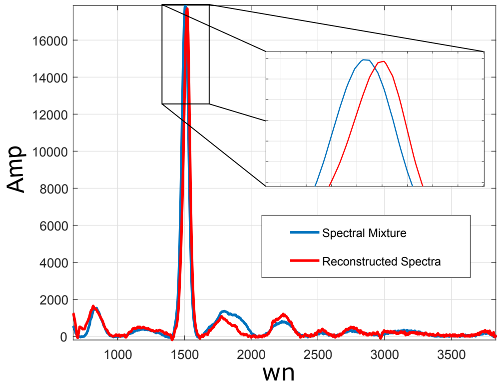

Fig. 1. The wavenumber shift artifact, a) spectral mixture and b) approximated spectra with reference spectra.

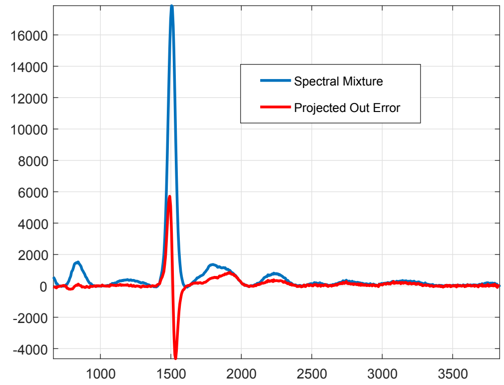

Fig. 2. The spectrum with wavenumber shift (blue) and the error after projecting out the corresponding reference library element (red)

chemical in the mixture [8]. To understand the error due to linear representations, such a representation has been plotted in figure 2 in red. The oscillating behaviour of the error around 1500 cm^{-1} is directly related to the spectral shift.

The reduction of such an error can naturally helps us in the complex mixtures deconvolution and provides better quantification. [9] treats the shift alignment as a complex discrete optimisation problem and proposes to solve the problem with the genetic algorithm. While this may be useful in an offline spectra processing, for large size problem and computationally restricted spectroscopy, this method is not suitable. [10] has a different approach and uses a local variational approximation for the shifts and considers within the cell resolution shifts, *i.e.* very small shifts in modern Raman spectroscopy. As this approach does not incorporate a prior knowledge about the spectra, *e.g.* sparsity, and only compensates small local shifts, it has a limited application in Raman spectroscopy.

In this paper we introduce a parametric model for the spectral peaks, which enables us to mathematically describe the behaviour of local peak shifts using a modified first order approximation. A robust sparse spectral decomposition is also introduced to incorporate such parametric model, which will be then solved using a simple convex program.

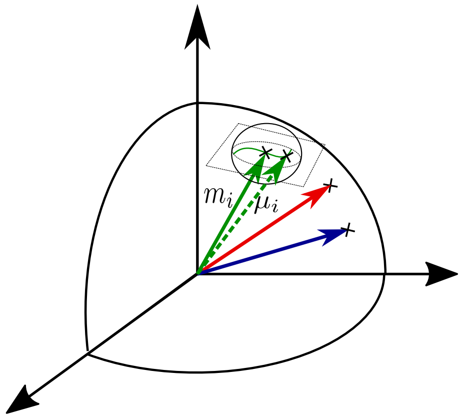

Fig. 3. Three dimensional representation of the shift in the tangent space of a unit sphere at a reference spectral vector.

II. SPECTRAL SHIFTS AND NONLINEAR SIGNAL MODELLING

Let the mixture spectrum $\mathbf{y} \in \mathbb{R}_+^d$ and the library spectra $\mathbf{m}_i \in \mathbb{R}_+^d, \forall i, 1 \leq i \leq N$ be baseline corrected and no coefficient non-linearity be in the generative model (1). The vectors $\mathbf{y}$ and $\mathbf{m}_i$'s live in a single orthant out of 2^d orthants. $\mathbf{m}_i$'s also locate on the surface of unit sphere, see figure 3. As a result of spectral shift, each m_i can rotate within a small neighbourhood. Let the local admissible set around $\mathbf{m}_i$ be called $\mathcal{M}_i$ and each two $\mathcal{M}_i$ and $\mathcal{M}_j$, for $i \neq j$, be disjoint, *i.e.* $\mathcal{M}_i \cap \mathcal{M}_j = \varnothing$. We can then rewrite the baseline corrected formulation of (1) as follows,

$$\begin{aligned}\mathbf{y} &= \phi(\mathbf{M}, \boldsymbol{\alpha}) + \boldsymbol{\omega} \\ &= \sum_{\substack{i \\ \boldsymbol{\mu}_i \in \mathcal{M}}} \alpha_i \boldsymbol{\mu}_i + \boldsymbol{\omega},\end{aligned} \tag{3}$$

where $\boldsymbol{\mu}_i$ is a signal in $\mathcal{M}_i$ and $\mathcal{M} = \cup_{i=1}^N \mathcal{M}_i$. Let $\boldsymbol{f}_i(\Delta) : \mathbb{R}^r \to \mathcal{M}_i$ be the function mapping a Euclidean space dimension r to $\mathcal{M}_i$, with the following properties for each $i, 1 \leq i \leq N$,

$$\begin{aligned}\boldsymbol{\mu}_i &= \boldsymbol{f}_i(\Delta), \\ \mathbf{m}_i &= \boldsymbol{f}_i(\mathbf{0}).\end{aligned} \tag{4}$$

$\boldsymbol{\mu}_i$ can then be locally approximated using first order Taylor's expansion of continuous $f_i(\Delta)$ around a point Δ_0 as follows,

$$\boldsymbol{\mu}_i \approx \boldsymbol{f}_i(\Delta_0) + (\Delta - \Delta_0) \frac{\partial}{\partial \Delta} \boldsymbol{f}_i(\Delta_0). \tag{5}$$

As we know $\boldsymbol{f}_i(\Delta_0) = \mathbf{m}_i$ at $\Delta_0 = \mathbf{0}$, it is usually easier to approximate $\boldsymbol{\mu}_i$'s around zero as follows,

$$\boldsymbol{\mu}_i \approx \mathbf{m}_i + \Delta \frac{\partial}{\partial \Delta} \boldsymbol{f}_i(\mathbf{0}). \tag{6}$$

This approximation can be loose for large Δ. We later show how we can have a better approximation of $\boldsymbol{\mu}_i$ using the expansion around $\Delta/2$.

A. Spectral Mixtures with Wavenumber Shifts

Raman spectra generally consists of some large peaks and noise type spectra, see for example the spectra in figure 1. If we simplify the structure of peaks and represent each peak, or the peak which mainly have spectral shifts, with some (differentiable) kernel functions $\boldsymbol{g}_\tau(\nu)$, where ν is the wavenumber and τ is the location of the peak. In this setting, we can write a spectrum as follows,

$$\boldsymbol{f}_i(\Delta) = \overline{\mathbf{m}}_i + \sum_{\tau \in T_i} \boldsymbol{g}_\tau(\nu - \Delta_\tau), \tag{7}$$

where $\overline{m}_i$, T_i and $\Delta = \{\Delta_\tau\}_{\tau \in T_i}$ are respectively the residual spectrum, *i.e.* spectrum without peaks, the locations of peaks and the set of all shifts. While T_i is normally a non-empty set for Raman spectra, Δ is set to zero for the reference spectra in the library. Δ is non-zero in the spectral mixtures with wavenumber shifts.

Model (7) can be jointly used with the generative signal model (3), subject to differentiability of the functions $\boldsymbol{g}_\tau$ w.r.t. Δ_τ. When the Raman spectral peaks are approximately symmetric, a Gaussian kernel can be a reasonable function to model such peaks. Let a Gaussian kernel $\boldsymbol{g}_\tau(\nu) := \beta_\sigma \exp(-(\nu - \tau)^2/2\sigma^2)$ be fitted to the spectral peaks, where $\beta_\sigma = 1/\sigma\sqrt{2\pi}$ is the normalisation factor and σ is the standard deviation.

The traditional first order approximation (5) is accurate, when the deviation Δ_0 is small. We here introduce a new approximation which is practically more accurate for linear representation of $\mathbf{y}$. The idea is to approximate $\boldsymbol{f}_i(\boldsymbol{\Delta}_0)$, where $\boldsymbol{\Delta}_0$ is the desired shift, around $\boldsymbol{\Delta}_0/2$, to reduce the approximation error. Using the Taylor's series (5) with $\boldsymbol{\Delta}_0 \leftarrow \boldsymbol{\Delta}_0/2$, we find the following approximation for $\boldsymbol{\mu}_i = \boldsymbol{f}_i(\boldsymbol{\Delta})$,

$$\boldsymbol{f}_i(\Delta) \approx \boldsymbol{f}_i(\Delta_0/2) + (\Delta - \Delta_0/2)\frac{\partial}{\partial \Delta}\boldsymbol{f}_i(\Delta_0/2). \tag{8}$$

For the spectral shift model, we can drive the following expansion,

$$\boldsymbol{f}_i(\boldsymbol{\Delta}) \approx \boldsymbol{f}_i(\frac{\boldsymbol{\Delta_0}}{2}) + \sum_{\tau \in T_i} (\Delta_\tau - \frac{\Delta_{0\tau}}{2})\frac{\partial}{\partial \Delta}\boldsymbol{g}_\tau(\nu - \Delta)|_{\Delta = \Delta_{0\tau}/2} \tag{9}$$

where $\Delta_{0\tau}$ is the element of Δ_0 relating to centre of $\boldsymbol{g}_\tau$, τ. We now evaluate $\boldsymbol{f}_i(\boldsymbol{\Delta})$ on $\mathbf{0}$ and $\boldsymbol{\Delta}_0$,

$$\begin{aligned}
\boldsymbol{f}_i(\mathbf{0}) &\approx \boldsymbol{f}_i(\frac{\boldsymbol{\Delta_0}}{2}) - \sum_{\tau \in T_i} \frac{\Delta_{0\tau}}{2}\frac{\partial}{\partial \Delta}\boldsymbol{g}_\tau(\nu - \Delta)|_{\Delta = \Delta_{0\tau}/2} \\
\boldsymbol{f}_i(\boldsymbol{\Delta}_0) &\approx \boldsymbol{f}_i(\frac{\boldsymbol{\Delta_0}}{2}) + \sum_{\tau \in T_i} \frac{\Delta_{0\tau}}{2}\frac{\partial}{\partial \Delta}\boldsymbol{g}_\tau(\nu - \Delta)|_{\Delta = \Delta_{0\tau}/2}.
\end{aligned} \tag{10}$$

Replacing $\boldsymbol{f}_i(\frac{\boldsymbol{\Delta}_0}{2})$ in the second approximation with its value found from the first and considering the Gaussian kernel, we derive an explicit expression for $\boldsymbol{\mu}_i$ as follows,

$$\begin{aligned}
\boldsymbol{\mu}_i &= \boldsymbol{f}_i(\boldsymbol{\Delta}_0) \\
&\approx \boldsymbol{f}_i(\mathbf{0}) + \sum_{\tau \in T_i} \Delta_{0\tau}\frac{\partial}{\partial \Delta}\boldsymbol{g}_\tau(\nu - \Delta)|_{\Delta = \Delta_{0\tau}/2} \\
&= \mathbf{m}_i - \sum_{\tau \in T_i} \frac{\beta_{\sigma_\tau}}{\sigma_\tau^2}\Delta_{0\tau}\frac{\partial}{\partial \Delta}\boldsymbol{g}_\tau(\nu - \Delta)|_{\Delta = \Delta_{0\tau}/2}.
\end{aligned} \tag{11}$$

For the Gaussian kernels, this approximation can be read as the following,

$$\boldsymbol{\mu}_i \approx \mathbf{m}_i - \sum_{\tau \in T_i} \Delta_{0\tau}\boldsymbol{g}'(\nu - \tau - \frac{\Delta_{0\tau}}{2}, \sigma_\tau). \tag{12}$$

By comparing (12) and traditional Taylor's approximation around zero, we realise the difference which is on the extra shift of $\frac{\Delta_{0\tau}}{2}$ in the new approximation. This term makes the approximation significantly more accurate for large spectral shifts $\boldsymbol{\Delta}$. This fact has been demonstrated in figure 4, where the approximation based on (12) has been plotted with dash-dotted green curves.

The spectral mixture $\mathbf{y}$ can now be represented as follows,

$$\mathbf{y} \approx \sum_i \alpha_i \mathbf{m}_i - \sum_i \sum_{\tau \in T_i} \alpha_i \Delta_{0\tau}\boldsymbol{g}'(\nu - \tau - \frac{\Delta_{0\tau}}{2}, \sigma_\tau) + \boldsymbol{\omega}. \tag{13}$$

This approximation is ”not” a linear function of shift values $\boldsymbol{\Delta}_0$, which can be potentially a problem in linear sparse approximations. The approximation can be discretised w.r.t. $\boldsymbol{\Delta_0}$ and the model becomes linear, considering possible shifts. Let the possible shifts $\Delta_{0\tau}, \forall\tau$ be noted by $\mathcal{D} = \{j\delta\}_{j \in \mathcal{J}}$, where $\mathcal{J}$ is a signed or unsigned set of integers including 0, *e.g.* $\mathcal{J} = \{-2, -1, 0, 1, 2\}$. We can then rewrite (13) as follows,

$$\mathbf{y} \approx \sum_i \alpha_i \mathbf{m}_i - \sum_i \sum_{\tau \in T_i} \alpha_i \sum_{j \in \mathcal{J}} \theta_{i,j} j\delta \boldsymbol{g}'(\nu - \tau - \frac{j\delta}{2}, \sigma_\tau) + \boldsymbol{\omega}, \tag{14}$$

where $\theta_{i,j} = \{0, 1\}$ and $\forall i, \sum_{j \in \mathcal{J}} \theta_{i,j} = 1$. If the constraint on $\theta_{i,j}$ is ignored, we can write $\mathbf{y}$ as follows,

$$\mathbf{y} \approx \sum_i \alpha_i \mathbf{m}_i - \sum_i \sum_{\tau \in T_i} \sum_{j \in \mathcal{J}} \gamma_{i,j,\tau}\boldsymbol{g}'(\nu - \tau - \frac{j\delta}{2}, \sigma_\tau) + \boldsymbol{\omega}, \tag{15}$$

where $\gamma_{i,j,\tau} = j\alpha_i\theta_{i,j}\delta$. The generative model for $\mathbf{y}$ is now a linear model based on α_i and $\gamma_{i,j,\tau}$. By relaxation of the constraint on $\theta_{i,j}$, we may find some representations which have non-vanishing $\gamma_{i,j,\tau}$'s. To indirectly encourage $\gamma_{i,j,\tau}$ become zero, we need some sparsity enforcing mechanism, which is the subject of next section.

III. ROBUST SPARSE DECOMPOSITION WITH AUGMENTED LIBRARY

Most mixture of interests are composed of few chemicals, which justifies sparsity of the coefficient vector [3]. The modified models for $\mathbf{y}$ in (15) have some extra terms to

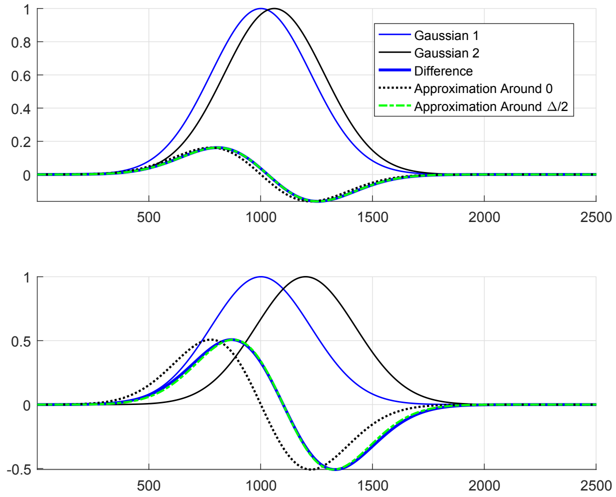

Fig. 4. Gaussian difference models: small shift (top panel) and medium shift (bottom panel)

compensate the wavenumber shifts. As we briefly explained, we also need to have sparse vectors $\gamma_{i,j,\tau}$ at most with the number of shiftable peaks for non-vanishing element α_i, *i.e.* $\sum_{\forall i, \alpha_i \neq 0} |T_i|$, where $|T_i|$ is the cardinality of set T_i. It would be easier to relax this strict sparsity condition on $\gamma_{i,\tau}$ and $\gamma_{i,j,\tau}$ and apply a general sparsity penalty on the linear representations. In this setting, the aim is to sparsely represent $\mathbf{y}$, with the following model,

$$\begin{aligned}\mathbf{y} &= \mathbf{M}\boldsymbol{\alpha} + \mathbf{A}\boldsymbol{\gamma} + \boldsymbol{\omega} \\ &= \underbrace{\begin{bmatrix}\mathbf{M} & \mathbf{A}\end{bmatrix}}_{\mathbf{M}^+}\underbrace{\begin{bmatrix}\boldsymbol{\alpha} \\ \boldsymbol{\gamma}\end{bmatrix}}_{\boldsymbol{\alpha}^+} + \boldsymbol{\omega}\end{aligned} \qquad (16)$$

where $\mathbf{A} \in \mathbb{R}^{d \times M}$ is the matrix related to the discretised linear functional$\boldsymbol{g}'(\nu - \tau - \frac{j\delta}{2}, \sigma_\delta)$, and $\boldsymbol{\alpha}^+ \in \mathbb{R}_+^{N+M}$. $\mathbf{M}^+$ is the augmented library with some extra functions modelling the spectral shifts. Such a linear generative model can be used by various sparse decomposition methods, including iterative optimisation, greedy and iterative reweighting methods. The convex formulation for finding sparse vectors is very popular as we practically find the global solution in a polynomial time. This formulation is based on using a sparsity promoting penalty like ℓ_1 and solve a program which assures the fidelity of the representation, as follows:

$$\boldsymbol{\alpha}^{+*} = \mathrm{argmin}_{\boldsymbol{\alpha}^+ \in \mathbb{R}_{\geq 0}} \|\mathbf{y} - \mathbf{M}^+\boldsymbol{\alpha}^+\|_2^2 + \lambda \|\boldsymbol{\alpha}^+\|_1, \qquad (17)$$

where λ is the Lagrange parameter which controls the noise level of the representation. Various algorithms have been presented to efficiently solve (17), including optimal first order methods, also called Nesterov's methods [11], and alternative direction of multipliers methods (ADMM), see [12] for a review.

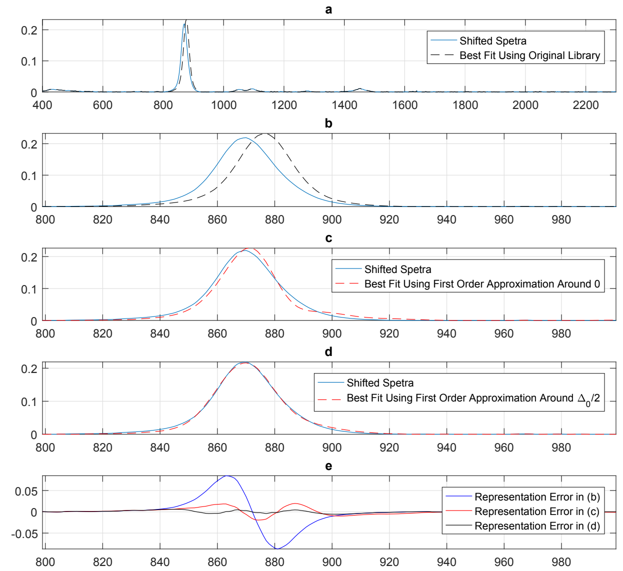

Fig. 5. Spectral mixture with 7 wavenumber shifts and estimated spectrum using original library (a), zoomed on the shifted peak (b), corrected reconstructed spectrum using first order approximation around zero (c), corrected reconstructed spectrum using first order approximation around $\Delta_0/2$, with maximum shift $\Delta_0 = 30$(d) and the errors (e).

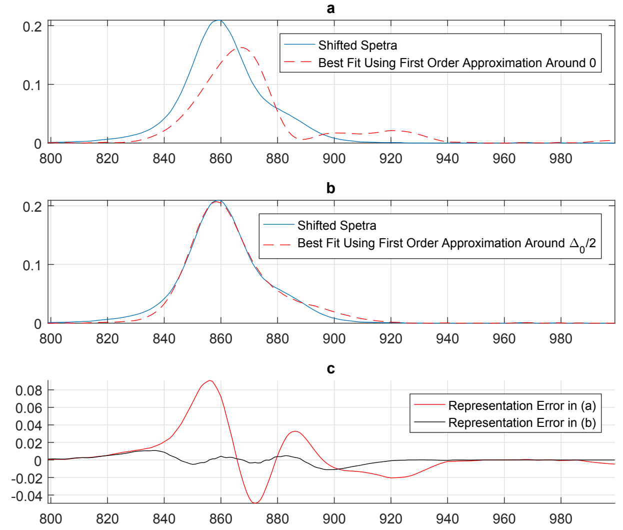

Fig. 6. Corrected reconstructed spectrum after 17 wavenumber shifts using first order approximation, around zero (c), around $\Delta_0/2$, with maximum shift $\Delta_0 = 30$ (b) and the errors (c).

IV. SIMULATIONS

The sparse decomposition of Raman spectra, with a library of 29 spectra, has been used to demonstrate the capabilities of the proposed robust method in improving the spectral fingerprinting performances. The maximum shift for formulation (15) has set to be 30 wavenumbers (wn) and $\delta = 5$. For a comprehensive comparison between the robust and standard sparse spectral decomposition methods, we synthetically generated spectra with local spectral shifts, using real spectral measurements. A real spectral pair was selected in our experiments, which are prone to spectral shifts. The peak of one spectrum was separated from the smooth part, shifted and added in a new wavenumber location. We can then synthetically generate arbitrary shifts in mixtures with

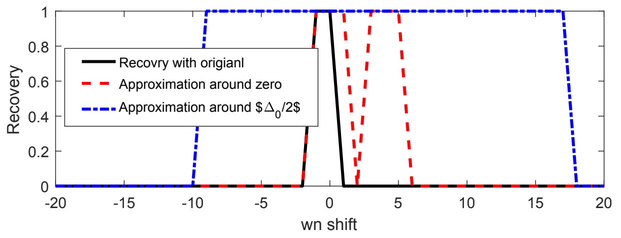

Fig. 7. Recovery with origianl (solid black), first order approximation around zero (dashed red) and first order approximation around $\Delta_0/2$ (blue dash-dotted).

desired contributions of constituents. We generated a shifted spectra with 7 and 17 wavenumber spectral shifts, and a mixture by combining %85 / %15 ratio of the spectra. We have demonstrated the effect of shift compensation methods by showing the spectra and residual errors in figures 5 and 6. The mixture and best fitted spectra, using original pure spectral library, have been plotted in 5.a and, the zoomed version in, 5.b. The spectral shifts are clear and the fitted spectra using augmented libraries are plotted in 5.c and 5.d for small shift 7, and 6.a and 6.b for large shift 17. While the advantage of using latter technique, which is more reliable for large shifts, is clear in these plots, we have shown the representation errors in 5.e and 6.c.

The robust sparse spectral decomposition for spectral fingerprinting has been challenged in the next experiment. We altered spectral shifts and applied sparse decomposition with original and two augmented libraries. The correct fingerprinting has been indicated with value 1 in figure 7. It is clearly demonstrated in this experiment that the approximation around $\Delta_0/2$ gives a more robust fingerprinting, with the expense of using a larger augmented library, which is caused by the discretisation of the shift parameter.

V. CONCLUSION

We presented a new method for robust sparse spectral decompositions, and the application to chemical fingerprinting. The new approach is based on allowing the pure chemical spectra in the reference library to slightly move on the surface of some spheres and model the rotation with first order Taylor's series around a middle point. This technique has proposed to increase the robustness of the sparse chemical fingerprinting [7] to the local spectral shifts. Such a robustness to the local spectral shifts can be crucial for fingerprinting of chemicals with small trace, which is of interest in defence, security and quality control. The simulation results demonstrate that the proposed technique can be used for Raman spectroscopy with a wide range of local spectral shifts nonlinearities.

We had access to a few spectral mixtures with nonlinear artefacts. However, we synthetically generated spectral mixtures to evaluate the algorithms. Our experiments with real mixtures confirm the results presented here. A more comprehensive study using a richer set of real mixtures has been left for the future work. We also incorporated a small library of spectra and allowed shifts in a single spectral pair. Having a larger augmented library will affect fingerprinting results, due to higher library coherence. Such a study was also left for the future work.

VI. ACKNOWLEDGEMENT

This work was supported in part by the University Defence Research Collaboration (UDRC) for signal processing in the networked battlespace, number P/K014277/1. It was also supported by the UK Defence Science and Technology Laboratory (Dstl) through enabling contracts ED-TIN2-4 and ED-TIN2-6 and through the provision of Raman spectra. The author also wants to thank Rhea Clewes, Dstl, for her constructive comments and suggestions.

REFERENCES

[1] M. Pelletier, "Quantitative analysis using Raman spectrometry," *Applied spectroscopy*, vol. 57, no. 1, pp. 20A–42A, 2003.

[2] D. M. Haaland and E. V. Thomas, "Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information," *Anal. chem*, vol. 60, no. 11, pp. 1193–1202, 1988.

[3] D. Wu, M. Yaghoobi, S. Kelly, M. Davies, and R. Clewes, "A sparse regularized model for raman spectral analysis," in *Sensor Signal Processing for Defence (SSPD), 2014*. IEEE, 2014, pp. 1–5.

[4] V. Mazet, C. Carteret, D. Brie, J. Idier, and B. Humbert, "Background removal from spectra by designing and minimising a non-quadratic cost function," *Chemometrics and Intelligent Laboratory Systems*, vol. 76, no. 2, pp. 121 – 133, 2005.

[5] Z.-M. Zhang, S. Chen, and Y.-Z. Liang, "Baseline correction using adaptive iteratively reweighted penalized least squares," *Analyst*, vol. 135, no. 5, pp. 1138–1146, 2010.

[6] X. Ning, I. W. Selesnick, and L. Duval, "Chromatogram baseline estimation and denoising using sparsity (BEADS)," *Chemometrics and Intelligent Laboratory Systems*, vol. 139, pp. 156–167, 2014.

[7] M. Yaghoobi, D. Wu, R. Clewes, and M. E. Davies, "Fast Sparse Raman Spectral Unmixing for Chemical Fingerprinting and Quantification," Edinburgh, Sep. 2016.

[8] H. Witjes, M. Van den Brink, W. Melssen, and L. Buydens, "Automatic correction of peak shifts in Raman spectra before PLS regression," *Chemometrics and Intelligent Laboratory Systems*, vol. 52, no. 1, pp. 105–116, 2000.

[9] J. Forshed, I. Schuppe-Koistinen, and S. P. Jacobsson, "Peak alignment of NMR signals by means of a genetic algorithm," *Analytica Chimica Acta*, vol. 487, no. 2, pp. 189–199, 2003.

[10] K. A. Veselkov, J. C. Lindon, T. M. Ebbels, D. Crockford, V. V. Volynkin, E. Holmes, D. B. Davies, and J. K. Nicholson, "Recursive segment-wise peak alignment of biological 1h NMR spectra for improved metabolic biomarker recovery," *Analytical chemistry*, vol. 81, no. 1, pp. 56–66, 2008.

[11] Y. Nesterov, "Gradient Methods for Minimizing Composite Objective Function," CORE Discussion Paper, Tech. Rep. 2007/76, 2007.

[12] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.