Analytica Chimica Acta 755 (2012) 17–27

Contents lists available at SciVerse ScienceDirect

Analytica Chimica Acta

journal homepage: www.elsevier.com/locate/aca

Photon level chemical classification using digital compressive detection

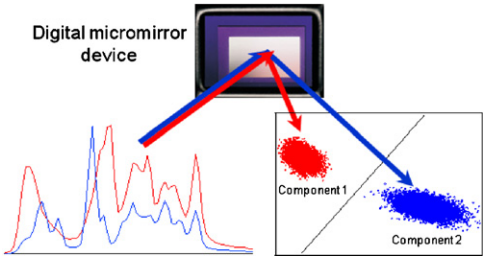
David S. Wilcox[a,*], Gregery T. Buzzard[b], Bradley J. Lucier[b], Ping Wang[a], Dor Ben-Amotz[a]

[a] Purdue University, Department of Chemistry, West Lafayette, IN, USA
[b] Purdue University, Department of Mathematics, West Lafayette, IN, USA

HIGHLIGHTS

- A new digital compressive detection strategy is developed.
- Chemical classification demonstrated using as few as ~10 photons.
- Binary filters are optimal when taking few measurements.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 27 July 2012
Received in revised form
21 September 2012
Accepted 3 October 2012
Available online 12 October 2012

Keywords:
Optimal binary filters
Compressive detection
Digital micromirror device
Classification
Total least squares

ABSTRACT

A key bottleneck to high-speed chemical analysis, including hyperspectral imaging and monitoring of dynamic chemical processes, is the time required to collect and analyze hyperspectral data. Here we describe, both theoretically and experimentally, a means of greatly speeding up the collection of such data using a new digital compressive detection strategy. Our results demonstrate that detecting as few as ~10 Raman scattered photons (in as little time as ~30 μs) can be sufficient to positively distinguish chemical species. This is achieved by measuring the Raman scattered light intensity transmitted through programmable binary optical filters designed to minimize the error in the chemical classification (or concentration) variables of interest. The theoretical results are implemented and validated using a digital compressive detection instrument that incorporates a 785 nm diode excitation laser, digital micromirror spatial light modulator, and photon counting photodiode detector. Samples consisting of pairs of liquids with different degrees of spectral overlap (including benzene/acetone and n-heptane/n-octane) are used to illustrate how the accuracy of the present digital compressive detection method depends on the correlation coefficients of the corresponding spectra. Comparisons of measured and predicted chemical classification score plots, as well as linear and non-linear discriminant analyses, demonstrate that this digital compressive detection strategy is Poisson photon noise limited and outperforms total least squares-based compressive detection with analog filters.

1. Introduction

Rapid identification and quantification of chemical species in complex mixtures is of importance to a wide range of applications in biology, medicine, manufacturing, and security. Multivariate statistical techniques combined with optical spectroscopies are increasingly employed in such applications for chemical component classification, calibration, and hyperspectral imaging. Here we show that a new *digital compressive detection* strategy can be used to facilitate rapid and accurate chemical classification based on the detection of as few as ~10 Raman scattered photons. Unlike previous full spectral or compressive detection/sensing methods, digital compressive detection utilizes binary optical filters that are optimized to minimize the resulting chemical classification uncertainty

* Corresponding author. Tel.: +1 765 494 5591.
E-mail address: wilcoxds@purdue.edu (D.S. Wilcox).

0003-2670/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.aca.2012.10.005

(as further explained below). Our strategy is shown to outperform previous full spectral and compressive chemical classification methods, including those based on analog rather than binary optical filters.

The present results build upon previous studies that have shown that chemometric techniques could be incorporated directly into the measurement process by using either static optical interference filters [1], or tunable liquid crystal [2–4] or micromirror [5–9] based multivariate optical elements built into the spectrometer hardware. The versatility of the tunable multivariate optical element permits these instruments to function as generalized spectrometers, capable of either full spectral acquisition or compressive detection using programmable optical filter functions. Here we demonstrate that our new digital compressive detection strategy facilitates the chemical classification of liquid samples with different degrees of spectral overlap with data collection times ranging from ~30 μs to ~5 ms. We further demonstrate how the resulting classification error varies both with the degree of spectral overlap and detection time. The variance and shape of the associated score plots are shown to entirely depend on photon counting statistics, as demonstrated using comparisons with detailed theoretical predictions as well as linear and non-linear discriminant analyses.

Conventional optical array (e.g., CCD) based spectrometers disperse light of different wavelengths onto N separate detectors in order to measure a spectrum. While this approach is advantageous for many applications, it has important drawbacks in the low-signal regime. For example, if we assume that a given chemical species emits over a fixed time period ~100 photons distributed over ~100 CCD pixels, then the resulting signal at each pixel would be well below the typical CCD read noise of a few counts per pixel. We show that digital compressive detection can be used to accurately classify chemical species under such conditions by using a single-channel photon counting detector to determine the total number of photons transmitted by binary optical filters optimized to distinguish the compounds of interest. On the other hand, it is also important to stress that the advantages of our compressive detection strategy are primarily restricted to this low signal regime, as a conventional CCD detection scheme may be preferential under high signal conditions (under which read and dark noise do not significantly impact the spectral signal-to-noise ratio).

Compressive detection is related in an interesting way to compressive data storage and multivariate chemometric methods. Chemometric techniques such as total least squares (TLS) [10], partial least squares (PLS) [10], principle component analysis (PCA) [10], and feature selection [11] may all be used to project N-channel spectral information onto a lower dimensional space defined by the projections (dot-product scores) of a measured spectrum onto a smaller number of M-axes. The latter M-scores may be used to classify, quantify, and compressively store the chemical information of interest. Compressive detection differs from the above procedures in that the M-scores are directly detected using the instrument hardware, rather than by post-processing full spectra. In other words, the intensity transmitted through each of the M filters represents a direct measure of the associated score, as it is equivalent to a dot-product of the collected light and the filter function. This mode of measurement benefits from Felgett's (or multiplex) signal-to-noise advantage, since the photons transmitted by each filter are detected on one channel, rather than being separately detected using N independent channels. The choice of filters to determine the M-axes may be considered as one type of feature selection. For a given measured spectrum, the set of all features could be defined to be the dot product of the spectrum with a vector having all entries between 0 and 1. We select M of these vectors as filters to minimize a particular measure of uncertainty in the quantities of interest. In a broad sense, the practice of integrating mathematics and hardware is precisely the goal of Integrated Sensing and Processing (ISP) [12], and so compressive detection (or the related multivariate optical computing concept [1]) may be considered a subset of ISP.

In this work we focus on compressive detection of emission-type spectra with binary optical filters and photon counting detection statistics. A key advantage of our digital compressive detection strategy derives from the fact that our binary filters are optimized to minimize the error in the classification score or component concentration, while previous compressive detection strategies focused on minimizing spectral differences. Our results demonstrate the surprising finding that results obtained using binary filters, optimized as described in this work, are capable of outperforming classification results obtained using analog filters and previous chemometric strategies.

2. Theory

In this section we construct a mathematical model that describes the measured response of a given filter in terms of assumed rates of photon emission for specified spectra. We then reformulate the problem of choosing filters to minimize the expected squared error in estimating these rates as a constrained optimization problem. We end this section by discussing some theoretical and practical observations about the solution of this problem.

2.1. Background

The problem of finding optimal settings for measurements taken with either the digital micromirror device (DMD) or an analog-based spatial light modulator (SLM) falls squarely into the subfield of statistics known as *Design of Experiments*, specifically with so-called *linear models*. A mathematical description of the results in the field as of 1993 is given by Pukelsheim [13]. Because our data are fundamentally *photon counts*, modeled by Poisson random variables whose variances equal their means, our problem does not fit precisely into the framework he develops, but we adopt his point of view in what follows.

The problem we consider is based on a hard model of linear additive spectra. Essentially we want to determine the concentrations of various chemical species from combined known spectra; this is sometimes known as the supervised spectral unmixing problem. See [14,15] for very general introductions to this problem. The work most closely related to ours is [16,17]; while they use a Poisson model similar to ours (with error terms), they assume full spectrum measurements and hence do not consider the choice of optimal filters for multiplex measurements. Another related paper is [18], which considers the problem of parameter estimation from Poisson observations in the case that there are more parameters than measurements. This latter paper also includes an informative discussion on other approaches to Poisson estimation.

2.2. Formulating the estimation problem

Assume that our sample consists of various amounts of chemical species from a known list, called $S_1, S_2, \ldots, S_n$. (These S_j's are not the spectra, just labels for the different compounds.) In a given experiment, the stream of photons counted by a detector from each S_j can be modeled as a Poisson process with rate parameter $\overline{\Lambda}_j$, where $\overline{\Lambda}_j$ has units photon s^{-1}. All the photons from the various S_js are mixed together in the stream of photons coming from the sample, and we are going to use the pattern of energies of the photons to "unmix" the photons and so estimate all the rate parameters $\overline{\Lambda}_j$.

Each $\overline{\Lambda}_j$ will depend on, among other things, the amount of S_j in the sample. We assume that we know the rate at which a unit amount of S_j gives off photons, so if we can estimate $\overline{\Lambda}_j$, we can

estimate the amount of compound S_j in the sample. We emphasize that the measurements in our experiment are determined by the rate parameters $\overline{\Lambda}_j$, and that is what we are going to estimate; the amounts of the various S_js may be calculated from the $\overline{\Lambda}_j$s in a post-processing step.

So, the number of photons emitted in an interval of time of length t of species S_j is a Poisson random variable with mean $t\overline{\Lambda}_j$.

The wavelength, or energy, of each photon observed in the experiments can be labeled with an integer $i \in \{1, \ldots, N\}$; N is the total number of energy bins, or wavelength channels, in the detection system. Assume that we know that the probability that a photon from species S_j has label i is given by P_{ij}, so $\sum_{i=1}^{N} P_{ij} = 1$. In other words, the P_{ij}, $i = 1, \ldots, N$ form the spectrum of the jth compound, normalized so that the sum (or integrated area of the spectrum) is 1. Then the stream of labeled photons emanating from a sample are modeled by a vector Poisson process with rates

$$P\overline{\Lambda},$$

where $\overline{\Lambda} = (\overline{\Lambda}_1, \ldots, \overline{\Lambda}_n)^T$, and $P = (P_{ij})$. (Here and later, superscript T denotes "transpose.") If we run the experiment for time t then the number of photons with label i entering our instrument from all chemical species has a Poisson distribution with mean

$$t(P\overline{\Lambda})_i = t\sum_{\ell=1}^{n} P_{i\ell}\Lambda_\ell.$$

We assume that the number of wavelength channels, N, is greater than the number of chemical species n, and that the columns of P are linearly independent. This is a real limitation when there are many possible chemical species to detect.

Now that we have a model for the rate of photon emission with label i that we might measure, we shall describe the mathematical model for the measuring device. We may consider taking n independent measurements, one for each possible chemical species. (This will be generalized later.) In the jth measurement, we can set in our optical filter the transmittance of all photons with energy level i to be a number F_{ij} with $0 \le F_{ij} \le 1$; i.e., the probability that in the jth measurement a photon with energy label i is counted is F_{ij}. If $F_{ij} = 1$, then in the jth measurement all photons with energy level i are passed through to the detector and counted; if $F_{ij} = 0$, then in the jth measurement all photons with energy level i are blocked from the detector – none of them are counted. Our observation in the jth measurement is the total photon count, summed over all energy levels i, from observing the photon stream for T_{jj} seconds, which will be a Poisson random variable with mean

$$T_{jj}\sum_{i=1}^{N} F_{ij}\left(\sum_{\ell=1}^{n} P_{i\ell}\Lambda_\ell\right) = T_{jj}\sum_{i=1}^{N}\sum_{\ell=1}^{n} F_{ij}P_{i\ell}\Lambda_\ell.$$

(We use a double subscript on T_{jj} because we shall make these numbers the diagonal of a matrix T.) The columns of the matrix $F = (F_{ij})$ are *filters*, and the entries of F can be chosen as we wish, since they are parameters of our measuring device.

For a DMD, we can choose only $F_{ij} = 0$ or $F_{ij} = 1$, while for an analog SLM we can in principle choose any $0 \le F_{ij} \le 1$. While it seems that analog SLMs offer more flexibility, we shall see later that optimal filters are *nearly digital*, in that nearly all the entries F_{ij}, $i = 1, \ldots, N$, are either 0 or 1.

We denote by $\mathbf{x}$ our complete observation, a vector of n independent Poisson random variables with means and variances given by the vector

$$TF^TP\overline{\Lambda}, \tag{1}$$

where T is the $n \times n$ diagonal matrix with diagonal entries $T_{11}, \ldots, T_{nn}$. We assume that F is chosen so that F^TP is invertible (which is possible since P has full rank).

If we denote by $\hat{\mathbf{x}}$ a sample from this random variable, then our estimate $\hat{\Lambda}$ of the true rates $\overline{\Lambda}$ is given by

$$\hat{\Lambda} = (TF^TP)^{-1}\hat{\mathbf{x}} = (F^TP)^{-1}T^{-1}\hat{\mathbf{x}}. \tag{2}$$

We note that the expected value of $\hat{\Lambda}$ satisfies

$$E(\hat{\Lambda}) = (F^TP)^{-1}T^{-1}E(\hat{\mathbf{x}}) = (F^TP)^{-1}T^{-1}(TF^TP\overline{\Lambda}) = \overline{\Lambda}.$$

Our first goal is to derive an effective expression for the expected squared error of $\hat{\Lambda}$ as an estimator of $\overline{\Lambda}$,

$$E\left(\|\hat{\Lambda} - \overline{\Lambda}\|^2\right) = \sum_j E((\hat{\Lambda}_j - \overline{\Lambda}_j)^2) = \sum_j \mathrm{Var}(\hat{\Lambda}_j);$$

this is the sum of the expected squared errors of all $\hat{\Lambda}_j$s as estimators of the photon count rates $\overline{\Lambda}_j$ of the j chemical species S_j. We shall then look at the problem of *minimizing* this sum.

We remark that our analysis allows models for which there are more measurements than chemical compounds, in which case our estimate will be

$$\hat{\Lambda} = BT^{-1}\hat{\mathbf{x}},$$

where B is a left inverse of F^TP, i.e., $B(F^TP) = I$. If we denote by M the total number of measurements ($M \ge n$), then F will be an $N \times M$ matrix, and T will be an $M \times M$ matrix.

In a later paper, we give an example of two synthetic spectra for which three filters perform better (in the sense of having lower expected square errors) than the optimal choice of a pair of filters. We do not know how often such examples would come up with real chemical spectra.

2.3. Optimizing the estimation problems

To design our filters we can assume, without loss of generality, that the total measurement time, $\sum_j T_{jj}$, is 1, so that T_{jj} can be interpreted as the fraction of the measurement time that the jth filter is applied. When the total measurement time in an experiment is τ seconds, then one takes a measurement with the jth filter for a period of τT_{jj} seconds. With this normalization, our estimator in an experiment of total measurement time τ is

$$\hat{\Lambda} = B(\tau T)^{-1}\hat{\mathbf{x}} = \tau^{-1}BT^{-1}\hat{\mathbf{x}}$$

and the variance of our estimator is inversely proportional to τ (because the variance of $\hat{\mathbf{x}}$ is proportional to τ).

We write $A = F^TP$ and let B be a left inverse of A, i.e., $BA = I$ (so $B = A^{-1}$ if A is a square matrix). Then the previous section leads to the following problem:

Problem 2.1. For a given $\overline{\Lambda}$ and $N \times n$ matrix P of normalized spectra, find a number M (the number of measurements), an $M \times M$ diagonal matrix, T (the times for all measurements), an $N \times M$ filter matrix, F, and an $n \times M$ estimator matrix B to minimize

$$E(\|BT^{-1}\hat{\mathbf{x}} - \overline{\Lambda}\|^2) \tag{3}$$

subject to

$$BA = I, \quad A = F^TP, \quad \sum_i T_{ii} = 1,$$

where $0 \le F_{ij} \le 1$ and $T_{ii} > 0$.

Our first goal is to find an explicit formula for (3).

We denote by $\mathbf{e}_j$ a column vector whose jth component is 1 and whose other components are 0. The jth component of the random variable $\mathbf{x}$, $x_j = \mathbf{e}_j^T\mathbf{x}$, is Poisson with mean and variance equal to

$$T_{jj}\mathbf{e}_j^T F^T P\overline{\Lambda} = T_{jj}(A\overline{\Lambda})_j,$$

since $F^TP = A$ by definition. So the random variable ϵ_j defined by

$$\epsilon_j = \frac{\mathbf{x}_j}{T_{jj}} - (A\overline{\Lambda})_j$$

has mean 0 and variance $T_{jj}^{-1}(A\overline{\Lambda})_j$. In other words,

$$\epsilon = T^{-1}\mathbf{x} - A\overline{\Lambda}$$

is a vector of independent random variables, the jth component of which has mean 0 and variance $T_{jj}^{-1}(A\overline{\Lambda})_j$. Thus, since $BA = I$,

$$B\epsilon = BT^{-1}\mathbf{x} - \overline{\Lambda}$$

is a vector random variable; the ith component of $B\epsilon$, $\sum_j b_{ij}\epsilon_j$, has mean 0 and variance

$$\sum_j b_{ij}^2 \mathrm{Var}(\epsilon_j) = \sum_j b_{ij}^2 T_{jj}^{-1}(A\overline{\Lambda})_j$$

Thus we have the useful formula

$$E(|(BT^{-1}\mathbf{x})_i - \overline{\Lambda}_i|^2) = \sum_j b_{ij}^2 T_{jj}^{-1}(A\overline{\Lambda})_j.$$

This expression is of interest in its own right, because one might want to ignore the error in estimating some components $\overline{\Lambda}_i$ of $\overline{\Lambda}$, so-called *nuisance parameters* [16]. A good example of this is the intensity of a fixed background spectrum, showing up in the device even when there is no sample; we can leave out of the following sum the estimate of the error of the background (see Appendix A for further discussion).

If we now sum over all i, we get

$$\begin{aligned} E\left(\|BT^{-1}\mathbf{x} - \overline{\Lambda}\|^2\right) &= \sum_i\sum_j b_{ij}^2 T_{jj}^{-1}(A\overline{\Lambda})_j \\ &= \sum_j T_{jj}^{-1}(A\overline{\Lambda})_j \sum_i b_{ij}^2 \\ &= \sum_j \frac{1}{T_{jj}}(A\overline{\Lambda})_j \|B\mathbf{e}_j\|^2. \end{aligned} \tag{4}$$

For fixed A, B, and $\overline{\Lambda}$, one can find the optimal measurement times T_{ii}; the following result follows from the Cauchy–Schwarz inequality.

Theorem 2.1. *For fixed A and B, the optimal values of T_{ii} in Problem* 2.1 *are given by*

$$T_{ii} = \frac{\|B\mathbf{e}_i\|\sqrt{(A\overline{\Lambda})_i}}{\sum_j \|B\mathbf{e}_j\|\sqrt{(A\overline{\Lambda})_j}}. \tag{5}$$

Thus, we can reformulate Problem 2.1 by replacing (3) with the right hand side of (4) and replacing T_{ii} by the right hand side in (5). This gives

$$E(\|BT^{-1}\hat{\mathbf{x}} - \overline{\Lambda}\|^2) = \left(\sum_i \|B\mathbf{e}_i\|\sqrt{(A\overline{\Lambda})_i}\right)^2. \tag{6}$$

Instead of minimizing the square, we can equivalently minimize the value itself. This gives the following.

Problem 2.2. For a given $\overline{\Lambda}$ and $N \times n$ matrix P of normalized spectra, determine the number of measurements, M, an $N \times M$ filter matrix, F, and an $n \times M$ matrix, B, to minimize

$$\sum_{i=1}^{M} \|B\mathbf{e}_i\|\sqrt{(A\overline{\Lambda})_i} \tag{7}$$

subject to

$$BA = I, \qquad A = F^TP, \quad 0 \le F_{ij} \le 1 \text{ for all } i, j.$$

We note the following property of optimal filters F, proved in [19].

Theorem 2.2. *If a (possibly nonoptimal) value of M, the number of measurements, is fixed, then the optimal filters obtained as the solution to Problem* 2.2 *can be chosen to consist of only* 0*s and* 1*s, except for at most $n-1$ entries strictly between* 0 *and* 1 *in each filter.*

Thus, one could say that the optimal filters are *almost binary*, or *almost digital*, in that one can take nearly all the channels of each filter to be either completely open (with $F_{ij} = 1$, i.e., full transmittance) or closed (with $F_{ij} = 0$, i.e., no transmittance). For example, to distinguish two spectra, we can choose two optimal filters such that only one channel in each filter is not digital, and presumably making that one channel in each filter 0 or 1 would leave one with filters that are nearly optimal. In our experiments we rounded each non-binary computed filter component F_{ij} to the nearer of 0 or 1.

We conjecture that for the optimal number of measurements, M, the optimal filters can be chosen to be completely digital.

We note that for *fixed* $\overline{\Lambda}$, A, and T, the minimizing B is the standard generalized least-squares minimizer as given in the following theorem:

Theorem 2.3 (generalized least squares). *Fix $\overline{\Lambda}$, A, and T as in Problem* 2.1 *and assume that $D = \mathrm{diag}(A\overline{\Lambda})$, the diagonal matrix whose diagonal entries are the components of the vector $A\overline{\Lambda}$, is invertible. Then the optimal B in Problem* 2.1, *which minimizes* (4), *is given by*

$$B = (A^TD^{-1}TA)^{-1}A^TD^{-1}.$$

Note that this theorem does not provide the optimal B for Problem 2.2 even when A is known (except when A is invertible, in which case $B = A^{-1}$), since T depends on B.

2.4. Computations

We address here some practical questions about computations based on the previous formulas.

Given $A = F^TP$ and $\overline{\Lambda}$, finding the B that minimizes (7) is a standard problem in convex analysis [19]. Similarly, given A, B, and $\overline{\Lambda}$, formula (5) gives explicit values of T_{ii}. So both matrices B and T are functions of A and $\overline{\Lambda}$.

If we ignore the dependence of (7) on $\overline{\Lambda}$, the main problem is determining the matrix $A = F^TP$ that minimizes (7); this consists of determining both the optimal number of measurement filters and the entries of each filter. We suggest, as a practical matter, to use the minimal number of filters, which is the number of possible chemical species in the sample, $M = n$. Determining A then reduces to minimizing (7) over all "feasible" matrices A; a matrix A is feasible if it can be written as $A = F^TP$ with $0 \le F_{ij} \le 1$ for all i and j.

The problem of finding the optimal A is a nonlinear, nonsmooth, nonconvex problem. Except in the case of two chemical species and two filters, we know of no algorithm that is guaranteed to find such a minimum. On the other hand, when the dimension of A is relatively small (we have tested up to ten chemical species and ten filters), the general routines from Matlab™, FMINCON for example, seem to perform adequately.

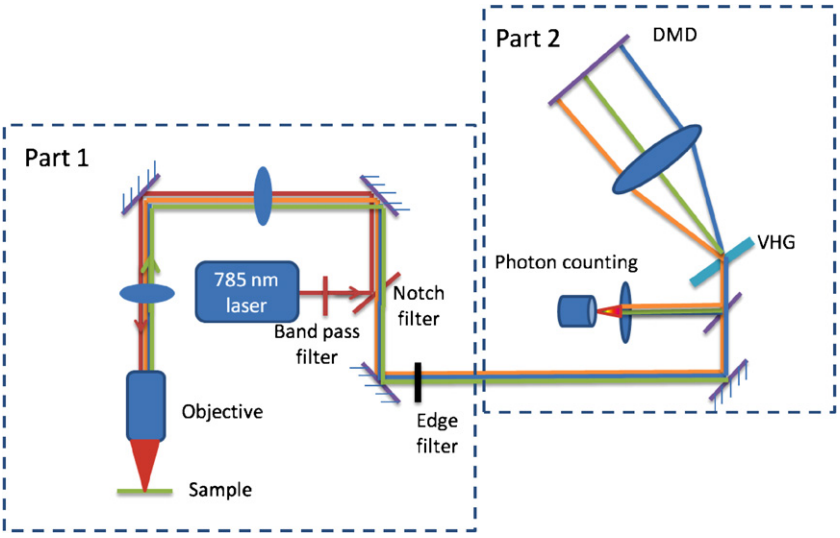


Fig. 1. Schematic of the DMD-based near infrared digital compressive detection instrument.

As for which vector $\overline{\Lambda}$ we should use in (7), we believe that a practical set of filters F can be designed assuming that the pure component emission rates are normalized to the same value,

$$\overline{\Lambda}_i = \overline{\Lambda}_j \tag{8}$$

for all i and j, i.e., we design measurement filters F to minimize the error in estimating a mixture where the rate of photons emitted by all chemical species are the same. Setting $\overline{\Lambda} = (1, 1, \ldots, 1)^T$ suffices. This determines $A = F^T P$, B, and T. Matlab software to determine OB filters is available on request. See www.math.purdue.edu/~buzzard/software/ for more details.

3. Experimental

3.1. Experimental apparatus

The compressive detection spectrometer, shown in Fig. 1, employs a Raman backscattering collection geometry. Part 1 is similar to that described in [2]. The excitation source is a 785 nm single mode laser (Innovative Photonic Solutions). After passing through a laser-line bandpass filter (Semrock, LL01-785-12.5), the laser is focused onto the sample with a NIR lens (Olympus, LMPlan IR, 20×). The Raman scattering is collected and separated from the laser Rayleigh scattering with a dichroic mirror (Semrock, LPD01-785RS-25) and a 785 nm notch filter (Semrock, NF03-785E-25).

The Raman scattered light is then sent to Part 2, where it is first filtered with a 900 nm shortpass filter (Thorlabs, FES0900) and subsequently directed to a volume holographic grating (1200 L mm^{-1}, center wavelength 830 nm, Edmund Optics, 48–590). The window of the dispersed light is ~200–1700 cm^{-1} with a spectral resolution of 30 cm^{-1} (this resolution is limited by the beam quality and hence the image of the diode laser focal spot size, which spans approximately 15 mirrors on the surface of the DMD). The light is collimated with an achromatic lens with a focal length of f = 50 mm (Thorlabs, AC254-050-B) and focused onto the DMD (Texas Instruments, DLP Discovery 4000). The DMD consists of 1920 × 1080 aluminum mirrors (10.8 μm pitch) that can tilt ±12° relative to the flat state of the array, controlled by an interface card (DLP D4000, Texas Instruments). All 1080 mirrors in each rows of the array are set to the same angle, and the 1920 columns are divided into adjacent groupings – e.g., if we want to divide the energy of the photons into 128 "bins", then groups of 15 adjacent columns are set in unison. The DMD is mounted at an angle such that the −12° mirror position directs photons back with a vertical offset of ~1° below the incident light in order to spatially separate the incident and reflected photons. The latter photons are recombined in a second pass through the holographic grating, and focused onto a fiber optic cable that is connected to a photodiode photon counting module (PerkinElmer, SPCMCD2969PE). The photon counting module has a dark count rate of ~200 photons s^{-1} and no read noise. A TTL pulse is output by the photon counter as each photon is detected, and the pulses are counted in a USB data acquisition (DAQ) card (National Instruments, USB-6212BNC). Integration timing is controlled by setting the sampling rate and number of samples to acquire with the DAQ card in Labview 2009.

Binary filter functions (F), optimal times (T), and the estimator (B) were generated from the spectra of all pure components (see Section 3.2 for more information) using functions from Matlab 7.13 R2011b. The input binary optical filter function determined which mirrors will point toward the detector (assigned a value of 1) or point away (assigned a value of 0). The binary (0–1) mathematical filters are configured to the DMD through Labview software (Texas Instruments, DDC4100, Load Blocks.vi) that sets blocks of mirrors on the DMD array corresponding to different wavelengths to the appropriate ±12° position. Labview scripts were used to sequentially apply the filters and integrate for the corresponding times, to store the raw photon counts, and to calculate the photon rates. Linear and quadratic discriminant analyses were performed in Matlab 7.13 R2011b. Data was further processed and plotted in Igor Pro 6.04.

3.2. Constructing filters

Generating accurate filters for a given application requires high signal-to-noise training spectra of each of the components of interest. Measuring full spectra with the DMD is achieved by notch scanning. This is done by sequentially directing one mirror (or a small set of mirrors) toward the detector (with all other mirrors directed away) and counting the number of photons detected at each notch position. Notch scanning measurements were performed using 1 s per notch to obtain spectra with a signal-to-noise ratio of ~500:1. A background spectrum is present in all of our training spectra, arising from the interaction of the excitation laser and the intervening optical elements. We have implemented two compressive detection strategies for removing this background. The first method involves measuring the background (with no sample)

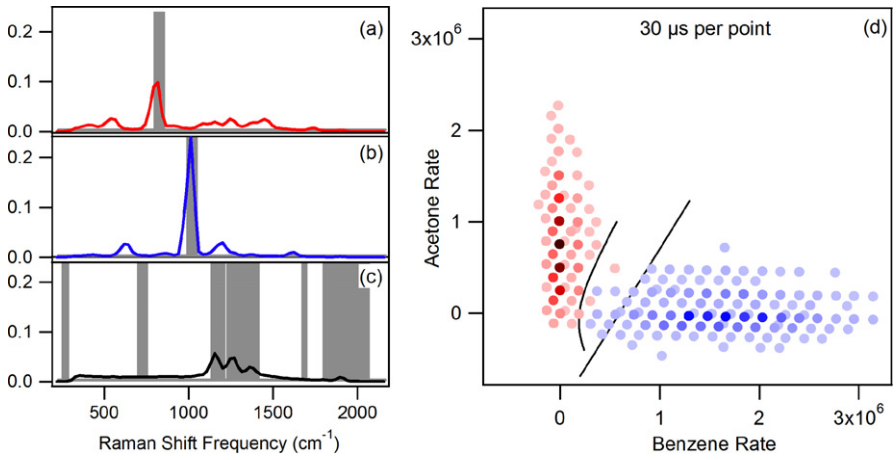


Fig. 2. Classification of minimally overlapping spectra. The Raman spectra of acetone, benzene, and the background are shown in panels (a)–(c), respectively (note that panels (a) and (b) include the background spectrum of panel (c)). The shaded regions denote the OB wavelengths that are directed toward the detector. The score plot in panel (d) shows the classification of acetone (red) and benzene (blue) using both linear and quadratic discriminants. Each point was obtained using three OB filters applied for a total integration time of 30 µs. The darkness of each colored disk represents the number of times each pair of photon rates was obtained in 1000 independent measurements (per cloud) – the lightest disks indicate one result each, the darkest acetone disk indicates 143 results, the darkest benzene disk indicates 74 results. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

and subtracting it from the spectrum of each sample. The background spectrum and background-free sample spectra are then each treated as independent components. Alternatively, since the background is a fixed, unwanted component, we have modified the optimization problem to minimize the variance of only the chemical components of interest (see Appendix A). In this method, the training spectra are used as measured (i.e., they include the background), and the background is treated as an additional independent component. Linear combinations of the resulting optimal digital filters are used to differentiate the photon counts originating from the variable components and the fixed background (see Appendix A). All of the results presented in this work have been obtained using the latter approach.

Once the training spectra are measured, they are normalized to unit area and are used as input to generate filters satisfying the constraints of Problem 2.2; we refer to filters generated in this way as optimal binary (OB) filters. Alternatively, the instrument shown in Fig. 1 may utilize various other types of filters. For example, analog filters of any spectral shape may be produced by performing measurements using 7 binary filters in order to obtain a grayscale with 128 transmittance levels at each wavelength channel (see Appendix B for more details). Such filters may be used to implement TLS multivariate spectral analysis, which we have done in the following two ways. The first method, referred to as TLS1, utilizes analog filters whose shapes are the same as the Raman spectra of each of the chemical components of interest (as measured using notch scanning). The second method, referred to as TLS2, utilizes linear combinations of the above filter functions to produce filters that directly measure the classification score for each component. Since TLS2 filters include both positive and negative features, two filters are used to independently measure the positive and negative features, which are subsequently combined to obtain the TLS2 filter signal. Further details regarding the construction of OB filters and analog TLS1/TLS2 filters are provided in Appendix C, and examples of the three types of filters are presented in Section 4.

4. Results and discussion

In order to critically test and quantitatively compare the digital OB and analog TLS1/TLS2 compressive detection strategies we have performed measurements using three pairs of organic liquids with varying degrees of spectral overlap. The first pair of liquids, acetone and benzene, have the least similar spectra (with a correlation coefficient of 0.12). The second pair of liquids, n-hexane and methylcyclohexane, have more significantly overlapped spectra (with a correlation coefficient of 0.71). The third pair of liquids, n-heptane and n-octane, have very highly overlapped spectra (with a correlation coefficient of 0.99). In addition to testing the influence of spectral overlap, we compare the resulting score distributions with theoretical predictions, including both linear and non-linear discriminant analyses.

4.1. Classification of minimally overlapping spectra

The spectra of acetone and benzene, shown in Fig. 2, are well separated, and in fact nearly orthogonal to each other as the angle between the corresponding normalized spectral vectors is ~84° (and thus their dot-product is 0.12). Note that all the spectra include the background spectrum shown in Fig. 2(c). The corresponding OB filter functions (obtained as described in Appendix A) are indicated by the gray regions in Fig. 2(a)–(c); these are the regions in which DMD mirrors are directed toward the detector to distinguish acetone and benzene from each other as well as from the background signal.

In this minimally overlapping case, each of the OB filters reflects photons toward the detector that originate primarily from the corresponding component. We say that each filter is "associated with" one spectra. The filters associated with acetone and benzene appear reasonable in that each wavelength is (approximately) assigned to a filter where the probability of that wavelength appearing in a spectrum is highest (see Fig. 2). However, there is no simple heuristic for determining the mirror positions of *optimal* binary filters. We start the optimization using a heuristic that assigns to each component spectrum a corresponding filter that has 1s in positions of relatively high photon count for that spectrum. However, the error from these filters is significantly higher than that found by the optimization routine. Moreover, having more filters than component spectra can decrease the recovery error, but there is no evident heuristic for finding such filters. We note the background is treated as a nuisance parameter (Appendix A), and the resulting filters are somewhat unintuitive.

The data points in the score plot in Fig. 2(d) represent the resulting component photon count rates ($\hat{\Lambda} = (F^T P)^{-1} T^{-1} \hat{\mathbf{x}}$) computed directly from the measured number of photons emerging from each

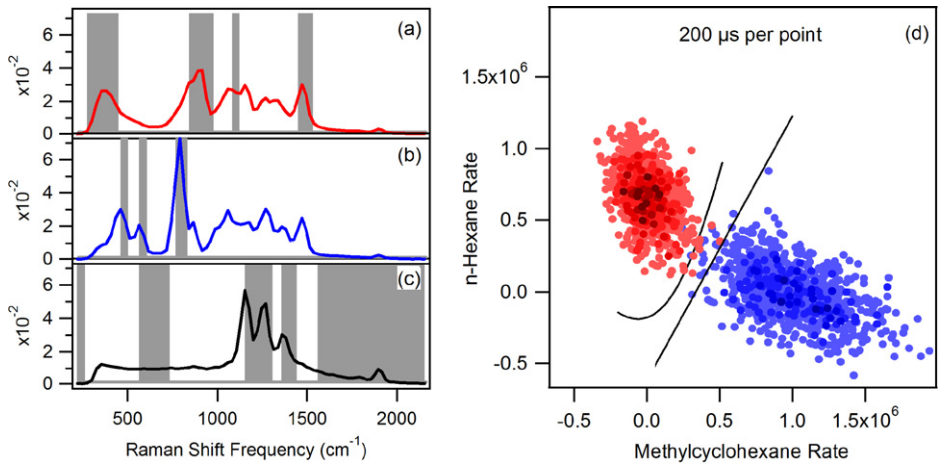


Fig. 3. Classification of moderately overlapping spectra. The Raman spectra of n-hexane (plus background), methylcyclohexane (plus background), and the background are shown in panels (a)–(c), respectively. The shaded regions denote the wavelengths directed toward the detector, as determined by the OB filters. Note that the filter associated with the background differs in this experiment from the filter in the previous experiment, even though the backgrounds are the same. This illustrates that filters depend on the interaction of all spectra in an experiment. The score plot in panel (d) shows the classification of n-hexane (red) and methylcyclohexane (blue) using both linear and quadratic discriminants. Each point was obtained using three OB filters applied for a total integration time of 200 μs. The darkness of each colored disk represents the number of times each pair of photon rates was obtained in 1000 independent measurements (per cloud) – the lightest disks indicate one result each, the darkest n-hexane disk indicates 5 results, the darkest methylcyclohexane disk indicates 4 results. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

filter ($\hat{\mathbf{x}}$), as specified by Eq. (2) (in this data set, the values of T were $T_{11} = 0.5289$, $T_{22} = 0.4195$, and $T_{33} = 0.0516$ for acetone, benzene, and the background, respectively). The score plots shown in Fig. 2(d) demonstrate that, when using OB filters, the photon rates can classify acetone and benzene in a total integration time of only 30 μs (for all three filters). This figure illustrates clearly that the points in a data cloud lie on a geometrical lattice arising from the fact that only whole numbers of photons are detected. The lattice points are well separated in this example because the number of detected photons, $\hat{x}_j$, was generally less than 10 photons per filter. Consequently, many of the 1000 independent measurements (per cloud) resulted in coincident triples of $\hat{x}_j$, and thus coincident pairs of $\hat{\Lambda}_j$ in Fig. 2(d). This figure also illustrates that we allow components of $\hat{\Lambda}$ to be negative, which is necessary as we require $E(\hat{\Lambda})$, the expected value of $\hat{\Lambda}$, to equal $\overline{\Lambda}$, i.e., we require that $\hat{\Lambda}$ be an *unbiased* estimate of $\overline{\Lambda}$. If we had chosen to set all negative components of $\hat{\Lambda}$ summarily to zero, then $\hat{\Lambda}$ would be a *biased* estimator of $\overline{\Lambda}$.

Linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) were employed to find a classification boundary and crossing rate (using results obtained with a total integration time of 100 μs). Though neither of LDA's assumptions (equal group covariances and normally distributed probability densities) were met with this data set, a reasonable classification boundary is nevertheless found, with 0% error (boundary crossing) for acetone and 2.2% for benzene. Relaxing the equal group covariance assumption with QDA results in a better non-linear classification boundary. The error rate in this case becomes 0.3% for acetone and 0% for benzene. These results clearly demonstrate that digital compressive detection can accurately classify these two compounds using only a handful of detected photons.

4.2. Classification of moderately overlapping spectra

The more significantly overlapped spectra of n-hexane and methylcyclohexane (and the background) are shown in Fig. 3. Although the spectral vectors of these two compounds are separated by an angle of only ∼45° (and thus have a dot product of 0.71), the spectra clearly have different shapes. The gray regions in Fig. 3 again show the corresponding OB filters. Note that unlike the spectra in Section 4.1, when using OB filters many of the detected photons come from multiple components. As a consequence of the increased degree of overlap, a longer total integration time of 200 μs is required to accurately distinguish the chemical components. The fractional times per filter were 0.3746, 0.5085, and 0.1169 for the three filters associated with n-hexane, methylcyclohexane, and the background, respectively. The score plots in Fig. 3(d) again contain 1000 points per cloud. In this case the average observed photon counts $\hat{\mathbf{x}}$ were of the order of 25 photons per filter measurement (see Eq. (2)), resulting in many fewer coincident pairs of $\hat{\Lambda}$.

The n-hexane/methylcyclohexane score plots were again analyzed using both LDA and QDA (with a training set collected using a total integration time of 100 ms). The group covariances of the n-hexane and methylcyclohexane rates are more similar than those of benzene and acetone, and consequently LDA and QDA provided similar classification results. QDA was slightly better at classifying the 200 μs total integration data correctly, with an error rate of 0.9% for n-hexane and 0.2% for methylcyclohexane vs LDA's error rate of 0% for n-hexane and 1.5% for methylcyclohexane.

4.3. Classification of highly overlapping spectra

In this section we apply the OB theorem to the more challenging classification of two linear alkanes, n-heptane and n-octane. From the spectra of n-heptane and n-octane, shown in Fig. 4, it is clear that there are very few regions where one component dominates (the dot product of the two normalized spectral vectors is 0.99, which corresponds to an angle of ∼8° between the two vectors). At many wavelengths, there is a near 50:50 probability of detecting a photon from n-heptane or n-octane. The wavelengths of greatest variance in the alkane spectra yield at most a ∼40:60 probability of detecting a photon from one component or the other. Nevertheless, by solving Problem 2.2 we obtain the resulting OB filters shown in Fig. 4. The score plots in Fig. 4(d) were obtained using a total integration time of 5 ms and fractional times of 0.4378, 0.5176, 0.0446 for the filters associated with n-heptane, n-octane, and the background, respectively. Due to the high degree of spectral overlap, considerably more photons (an average of 200 per filter measurement) were needed for minimally overlapping score distributions.

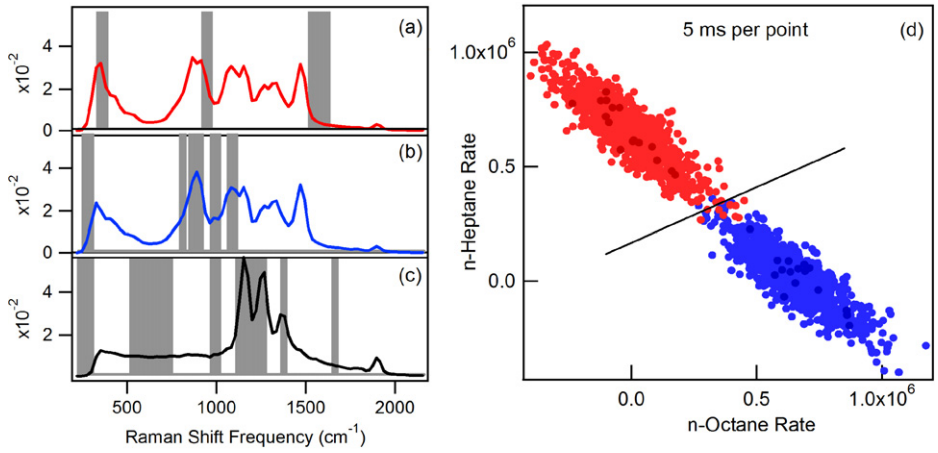


Fig. 4. Classification of highly overlapping spectra. The Raman spectra of n-heptane (plus background), n-octane (plus background), and the background are shown in panels (a)–(c), respectively. The shaded regions denote the wavelengths directed toward the detector, as determined by the OB filters. The score plot in panel (d) shows the classification of n-heptane (red) and n-octane (blue) using a linear discriminant. Each point was obtained using three OB filters applied for a total integration time of 5 ms. The darkness of each colored disk represents the number of times each pair of photon rates was obtained in 1000 independent measurements (per cloud) – the lightest disks indicate one result each, the darkest disk for both n-heptane and n-octane indicates 2 results. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

The LDA classification boundary was obtained using a training set obtained in a total integration time of 100 ms, and gives an error rate of 0.9% for n-heptane and 0.3% for n-octane; QDA gives the same results because the group covariances are equal.

4.4. Classification as a function of time

Sections 4.1–4.3 have shown the minimum amount of time required to classify spectra with varying degrees of overlap to within ~1% error. Fig. 5 shows how the classification score plots for n-hexane and methylcyclohexane vary as a function of integration time, ranging from (a) 0.3 ms to (d) 100 ms. When longer integration times are employed, the variance of each distribution decreases as the inverse of the integration time and the associated misclassification rate essentially vanishes. The shapes and sizes of the score plots are entirely determined by photon counting (Poisson) statistics (as demonstrated in the following sub-section).

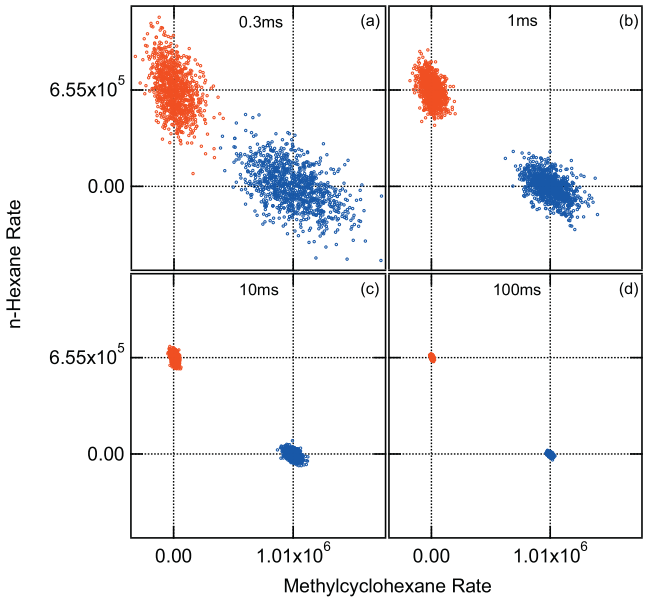


Fig. 5. The variance of the experimentally measured photon rates for pure n-hexane (red) and pure methylcyclohexane (blue) decreases as the total measurement time increases from (a) 0.3 ms to (d) 100 ms. The misclassification rate vanishes at longer times, as the 1000 points in each cloud cluster tightly together. The horizontal and vertical lines in each panel indicate the corresponding mean photon detection rates. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

4.5. Comparison with theory

Fig. 6 shows both the experimentally measured and theoretically simulated score plot for the n-hexane/methylcyclohexane system at 1ms total integration time. The mean rates, $\overline{\Lambda}$, used in the theoretical predictions were set to the experimental average $\hat{\Lambda}$ values obtained from 10,000 independent measurements of n-hexane and methylcyclohexane. Thus, the OB filters, optimal times, normalized training spectra (including the background), and approximated $\overline{\Lambda}$ were the only input parameters. It was assumed that the filters were perfectly square (binary) functions, and that the photons emerging through the filters had ideal Poisson distributions. The results shown in Fig. 6 clearly show the excellent agreement between the theoretical (black) and experimental (colored) measurement.

We applied the Kolmogorov–Smirnov statistical test to determine whether the experimental outputs of the filters are indeed distributed as Poisson random variables. We took 10,000 measurement triples (with each triple totaling 1 ms measurement time) of both n-hexane and methylcyclohexane in the experimental apparatus. Using this data, we tested whether the six sets of random samples are distributed as Poisson random variables with the sample means. The tests show that one cannot reject the hypotheses that the experimental output of the filters is Poisson with the sample means, even at 20% level of significance. It is possible, however, that the sample means may differ from what the theory predicts; such a comparison is left to a later study. Preliminary results suggest this discrepancy is due to photon leakage from DMD mirrors corresponding to longer Raman scattered wavelengths.

4.6. Comparison with TLS

The classification performance obtained using OB filters is found to be better than that obtained using TLS, as illustrated using the n-hexane/methylcyclohexane results shown in Fig. 7. As discussed in Section 3.2 and Appendix C, two methods have been used to implement TLS filters in our compressive detection instrument: TLS1

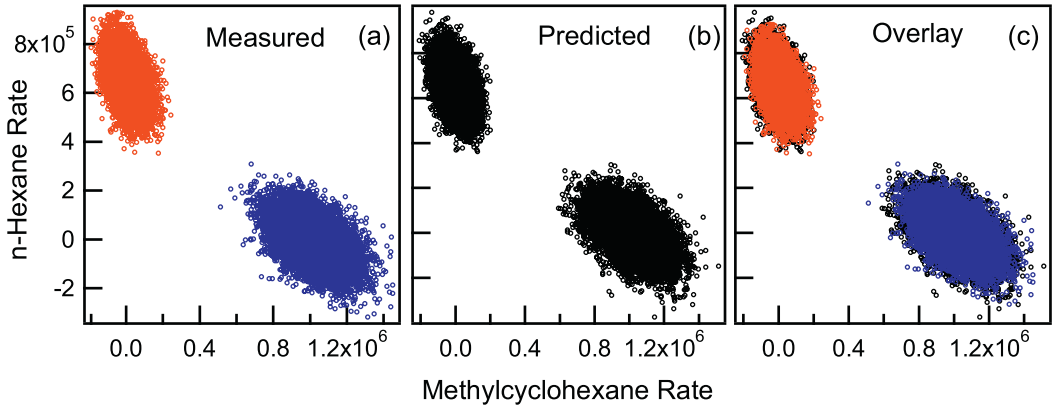


Fig. 6. Comparison between the experimental and theoretical variances and shapes of the photon count distributions for n-hexane and methylcyclohexane. Each cloud contains 10,000 points. The panels show (a) the experimental distributions, (b) the theoretical distributions, and (c) the overlay of experimental and theoretical distributions, which reveals that they are in excellent agreement.

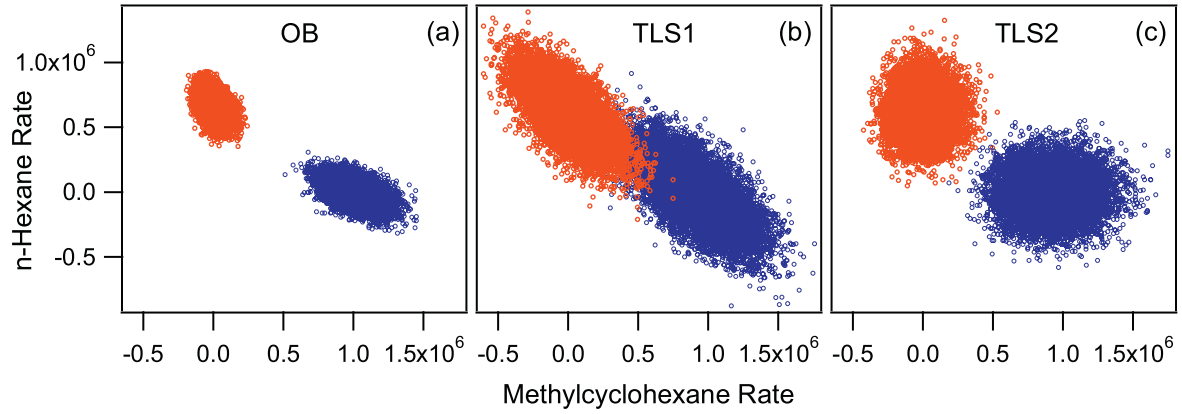


Fig. 7. Compressive detection results obtained using (a) OB filters are clearly better than results obtained using either implementation of total least squares, (b) TLS1, or (c) TLS2. Each cloud contains 10,000 data points.

(which uses spectral equivalent transmission functions) and TLS2 (which uses linear combinations of TLS1 filters to directly measure individual component photon rates). Fig. 7 compares the measured distributions of n-hexane and methylcyclohexane integrated for a total time of 1 ms using filters obtained from the OB theorem, TLS1, and TLS2. The OB filters clearly outperformed either implementation of TLS, while TLS2 outperformed TLS1 (as indicated by the larger variance of the latter score distribution). Recall that TLS2 requires twice as many DMD measurements as TLS1 in order to independently measure the positive and negative features of each TLS2 filter (as further explained in Appendix C). The fractional integration times per TLS1 filter were ~333 μs, and the corresponding TLS2 filters were applied for ~0.167 μs. We have found that TLS2 is a better method than TLS1 for classifying any moderate to highly overlapping spectra. The improved performance obtained using TLS2 is linked to the fact that the TLS2 filters are better focused on the more highly overlapping regions of the component spectra. However, results obtained using the OB filters invariably outperform both TLS1 and TLS2.

5. Conclusion and future work

A new digital compressive detection strategy has been implemented in this work. We have developed a method for finding optimal binary filters that can be realized using a digital micromirror device based compressive detection spectrometer. The binary filters are optimal in the sense that the sum of the variance of the estimated component photon count rates is minimized. The ability to minimize the latter objective arises from the freedom to choose which binary mirrors are open or closed, as well as optimal integration times per filter.

We have tested and validated this new digital compressive detection strategy by classifying liquids with various degrees of spectral overlap. We found that as few as 10–25 photons per measurement were required to accurately classify low to moderately overlapping spectra with an error of less than 1% with total measurement times ranging from tens to hundreds of microseconds. For the highly overlapping case of two linear alkanes, accurate classification was obtained by detecting ~200 photons, collected in a few milliseconds. Data acquisition times approaching these timescales are not accessible using conventional, CCD-based Raman spectroscopy, thus highlighting the power of compressive detection. OB filters were also shown to require fewer photons for accurate classification than TLS filters.

In future publications we intend to explore several extensions of the present paper. For example, the classification results we present here can be extended to *quantification* results with slightly longer measurement times (cf. Fig. 5). Moreover, we plan to extend the "nuisance variable" approach, used here to remove the background signal, to remove fluorescence that may vary from sample to sample. In addition, the speed of our results facilitates real time *chemical imaging* measurements, by using compressive detection to classify pure compounds (or mixtures) present at each spatial point of a sample. Finally, because Partial Least Squares (PLS) or Principal Component Analysis (PCA) scores are simply the inner products of the sample spectrum with various filters, we can

implement such calculations in the hardware directly (as we did for TLS filters).

Acknowledgments

This work was supported in part by grants from the National Science Foundation (Award # IDBR 0754740 to DSW, PW, and DBA) and the Simons Foundation (Award #209418 and #229816 to BJL).

Appendix A. Dealing with a fixed background

We say a signal is a "fixed background" if it is present in every measured spectrum, does not significantly vary in spectral shape from one measurement to another, and scales linearly with exposure time. In this appendix we describe how the theory in Section 2 can be extended to remove fixed backgrounds.

To illustrate our method, we assume that there are two possible chemicals S_1 and S_2 in our sample, and that a background signal is present in all measurements. With no chemical sample in the equipment, we measure a normalized spectrum P_b (b for "background"); P_b is a column vector whose entries sum to 1. We repeat this process with pure samples of S_1 and S_2 in the equipment to obtain normalized spectra P_{1+b} and P_{2+b}. We also measure the total photons counted over one second with no sample, S_1 alone, and S_2 alone – we denote these quantities by Σ_b, Σ_{1+b}, and Σ_{2+b}, respectively.

From these initial measurements, we solve Problem 2.2 with the matrix $P = [P_1\ P_{1+b}\ P_{2+b}]$, $M = 3$, and $\overline{\Lambda} = (1, 1, 1)^T$ to obtain OB filters F, the matrix T whose diagonal gives relative measurement times for each filter, and resulting matrices $A = F^T P$ and $B = A^{-1}$.

Each photon from the background is associated either to the background alone or to one of the samples S_i plus the background. So for *any* concentration of chemical S_i, the rate of photons coming from S_i alone is given by

$$\overline{\Lambda}_i = \left(\frac{\Sigma_{i+b} - \Sigma_b}{\Sigma_{i+b}}\right)\overline{\Lambda}_{i+b} = \left(1 - \frac{\Sigma_b}{\Sigma_{i+b}}\right)\overline{\Lambda}_{i+b}.$$

Thus, if we have computed estimates $\hat{\Lambda}_{1+b}$ and $\hat{\Lambda}_{2+b}$ of the rates of photons counted from S_1 plus the background and S_2 plus the background, then we estimate $\overline{\Lambda}_i$ by

$$\hat{\Lambda}_i = \left(1 - \frac{\Sigma_b}{\Sigma_{i+b}}\right)\hat{\Lambda}_{i+b}.$$

So $E(\hat{\Lambda}_i) = \overline{\Lambda}_i$ and

$$\mathrm{Var}(\hat{\Lambda}_i) = \left(1 - \frac{\Sigma_b}{\Sigma_{i+b}}\right)^2 \mathrm{Var}(\hat{\Lambda}_{i+b}) < \mathrm{Var}(\hat{\Lambda}_{i+b}).$$

Appendix B. Generating analog patterns on the digital micromirror device

The DMD mirrors may either be set on (toward the detector) or off (away from the detector). In order to produce analog patterns, some mirrors must be turned on longer than others over the period of photon integration. Since the mirror switching time of our DMD is ~30 ms, the dead time for 128 grayscales (or 128 separate binary filters) is ~4 s. Thus, the most efficient way to produce grayscale patterns is to switch the mirror positions as few times as possible. Here we show that n binary filters can be combined to reproduce one filter with 2^n grayscales by varying the integration time of each of the n filters as follows. For simplicity, consider the case where one wants $128 = 2^7$ grayscales in 128 ms. Given an arbitrary analog transmission function, the shape is first scaled to have a maximum intensity of 127 so that each mirror is assigned a different transmission intensity between 0 and 127. The grayscale transmission intensities at each filter entry are then expressed as a binary number. The first of 7 binary filters is obtained by turning on all mirrors for which the corresponding grayscale entry has the 64 (2^6) bit set; all other mirrors are turned off. This filter is measured for 64 ms. The second binary filter is obtained by turning on all mirrors for which the corresponding grayscale entry has the 32 (2^5) bit set; all others are turned off. This filter is measured for 32 ms. The third binary filter is obtained similarly using all entries that have the 16 (2^4) bit set; this filter is measured for 16 ms, etc. In other words, all entries with grayscales from 32 to 63 and from 96 to 127 have the mirror turned on in the second step; in the third step, all entries with grayscales from 16 to 31, 48 to 63, 80 to 95, and 112 to 127 have the mirror turned on; etc. The n measurements are then summed in a post-processing step to reproduce the effect of applying the original grayscale filter for 128 ms. Since the mirrors are switched 7 times, there is a delay of 210 ms associated with such measurements. However, a faster DMD interface could reduce the delay to less than 1 ms per filter.

Appendix C. Implementing TLS

TLS filters can be implemented in the compressive detection spectrometer in two different ways. In Section 2, we described the application of filters F_{ij} with all non-negative entries, corresponding to a transmission of 0–100% at each wavelength (or in the case of binary filters, 0 or 100%). In the first version of TLS, which we refer to as TLS1, the jth filter F_{ij} is the same shape as the jth spectra (scaled to have a maximum intensity of 1 or transmission of 100%). TLS1 filters are analog, but all non-negative (ranging from 0% to 100% transmission) at each wavelength. Therefore, the rates from TLS may be calculated from Eq. (2), where $B = (F^T P)^{-1}$ and F are TLS1 filters. Computing the rates from $\hat{\mathbf{x}}$ using this method requires that all n filters must be applied to estimate the photon rates of n components.

The second version of TLS, denoted TLS2, uses linear combinations of TLS1 filters. It follows from the solution of Eq. (2) that the transmission function defined by the product $(F^T P)^{-1} F^T$ may be used as an alternative way to obtain estimates of $\overline{\Lambda}$. Note that the filters F are the same spectral equivalent filters as in TLS1, but the product yields TLS2 filters. TLS2 filters can measure the photon rates of fewer than n components, since each filter is orthogonal to all others (i.e., each filter only "sees" one component and is blind to all others). However, the filter entries have intensities that can be positive or negative. Filters with negative transmission intensity are not physical, but this can be handled by writing a general filter as the difference between two filters, each with only positive entries. Each of these filters is then scaled to have a maximum transmission of 100%. Measuring with each of these filters, then rescaling and taking the difference in post-processing mimics the measurement with the corresponding TLS2 filter. While this method requires $2n$ filters to measure information from all n components, it can also be used to measure photon rates from fewer than n components. In other words, if we are only interested in determining the intensity of a single component in a mixture then we only need to obtain the signal from the one TLS2 filter function pertaining to that component (which required only measuring the number of photons detected using the corresponding positive and negative filters). See Section 4.6 for a comparison of TLS1, TLS2, and OB filters.

References

[1] M.P. Nelson, J.F. Aust, J.A. Dobrowolski, P.G. Verly, M.L. Myrick, Anal. Chem. 70 (1998) 73–82.
[2] B.M. Davis, A.J. Hemphill, D.C. Maltas, M.A. Zipper, P. Wang, D. Ben-Amotz, Anal. Chem. 83 (2011) 5086–5092.

[3] W.C. Sweatt, C.A. Boye, S.M. Gentry, M.R. Descour, B.R. Stallard, C.L. Grotbeck, Imaging Spectrometry IV Volume 3438 of Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE), 1998, pp. 98–106.
[4] N. Uzunbajakava, P. de Peinder, G.W. t Hooft, A.T.M. van Gogh, Anal. Chem. 78 (2006) 7302–7308.
[5] Z.J. Smith, S. Strombom, S. Wachsmann-Hogiu, Opt. Express 19 (2011) 16950–16962.
[6] Q.S. Hanley, P.J. Verveer, T.M. Jovin, Appl. Spectrosc. 52 (1998) 783–789.
[7] R.A. DeVerse, R.M. Hammaker, W.G. Fateley, J.A. Graham, J.D. Tate, Am. Lab. 30 (1998) 112S.
[8] N.T. Quyen, E. Da Silva, N.Q. Dao, M.D. Jouan, Appl. Spectrosc. 62 (2008) 273–278.
[9] E.P. Wagner, B.W. Smith, S. Madden, J.D. Winefordner, M. Mignardi, Appl. Spectrosc. 49 (1995) 1715–1719.
[10] K. Varmuza, P. Filzmoser, Introduction to Multivariate Statistical Analysis in Chemometrics, CRC Press/Taylor and Francis, Boca Raton, FL, 2009.
[11] M. Dash, H. Liu, Intell. Data Anal. 1 (1997) 131–156.
[12] C.E. Priebe, D.J. Marchette, D.M. Healyl, Modern Signal Processing, Volume 46, MSRI Publications, 2003.
[13] F. Pukelsheim, Optimal design of experiments, Volume 50 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006, Reprint of the 1993 original.
[14] J.M. Bioucas-Dias, A. Plaza, in: Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, volume 7830 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2010.
[15] N. Keshave, Lincoln Lab. J. 14 (2003) 55–78.
[16] R.D. Palkki, A.D. Lanterman, Opt. Eng. 49 (2010) 113601.
[17] B. Drake, J. Kim, M. Mallick, H. Park, 13th Conference on Information Fusion (FUSION), 2010, pp. 1–8.
[18] Z. Harmany, R. Marcia, R. Willett, IEEE Trans. Image Process. 21 (2012) 1084–1096.
[19] G.T. Buzzard, B.J. Lucier (2012), in preparation.