# Detection of unknown gas-phase chemical plumes in hyperspectral imagery

James Theiler and Brendt Wohlberg

Los Alamos National Laboratory, Los Alamos, NM 87545

## ABSTRACT

Gas-phase chemical plumes exhibit, particularly in the infrared, distinctive emission signatures as a function of wavelength. Hyperspectral imagery can exploit this distinctiveness to detect specific chemicals, even at low concentrations, using matched filters that are tailored both the the specific structure of the chemical signature and to the statistics of the background clutter. But what if the chemical species is unknown? One can apply matched filters to a long list of candidate chemicals (or chemical mixtures), or one can treat the problem as one of anomaly detection. In this case, however, the anomalous signals of interest are not completely unknown. Gas spectra are generically sparse (absorbing or emitting at only a few wavelengths), and this property can be exploited to enhance the sensitivity of anomaly detection algorithms. This paper investigates the utility of sparse signal anomaly detection for the problem of finding plumes of gas with unknown chemistry in hyperspectral imagery.

**Keywords:** hyperspectral, target detection, anomaly detection, sparse modeling

## 1. INTRODUCTION

What hyperspectral imagery promises is a *blessing* of dimensionality. With so many distinct spectral channels (typically over a hundred, and sometimes several hundreds), exquisite spectral discrimination is possible. Among the more dramatic success stories is the hyperspectral detection of gas-phase chemical plumes at astonishingly low concentrations.[1–12]

The difference between an on-plume and an off-plume spectrum can be very small, but when the chemistry of the plume is known, then that small difference has a fairly well-defined "direction" – and in a high-dimensional space, this direction can be very distinctive, and a matched filter can be very effective at detecting the presence of a given chemical plume. But to do this, one needs to know the gas spectrum: a detector that is designed for for carbon monoxide will not be appropriate for hydrogen sulfide. In many operational scenarios, one does not know *a priori* what chemicals (and/or mixtures of chemicals) are present, and the search is in some ways more like anomaly detection than target detection. This invites the question: to what extent can the tools of anomaly detection be brought to bear on the problem of gas detection, when the chemistry of the gas is not known.

In the Appendix, a selection of two dozen gas spectra are shown; it is evident that these are quite distinct from each other, but they also exhibit a common property that can potentially be leveraged to enhance the performance of our anomaly detectors. These spectra are *sparse*: the absorption coefficient is zero or near-zero for the vast majority of the wavelengths in the spectrum. This invites the question that led to this study: can this spectral sparsity be exploited to improve anomaly-detection approaches to the blind gas detection problem? Following up on preliminary work reported previously,[13] we extend and further develop the spectrally sparse anomaly detector, and suggest that it may be useful for detecting chemical plumes even when we do not know the chemical species in the plume.

In Section 2, we formulate the problem of spectrally sparse anomaly detection, starting with a "derivation" of the RX-style Mahalanobis distance detector, and then extending it both by considering restrictions on the target signature (mainly, that it be sparse) and by using more flexible models of background clutter. Section 3 describes a series of numerical experiments using artificial plumes on real hyperspectral backgrounds, and comparing variants of spectrally sparse detectors. We conclude in Section 4, but follow that up with an Appendix that includes many examples of chemical absorption spectra.

Emails: jt@lanl.gov, brendt@lanl.gov

## 2. SPECTRALLY SPARSE ANOMALY DETECTION: FORMULATION OF THE PROBLEM

Consider the hypothesis testing framework for detecting gas-phase chemical plumes in hyperspectral imagery. Let $\mathbf{x}$ correspond to the vector-valued measurement (*e.g.,* radiance, or – after some processing – reflectance) at a pixel. Each component of $\mathbf{x}$ is the measurement at a given spectral channel. If $\mathbf{s}$ is the spectral signature of a specific gas, then for the case of an absorptive plume,* we can write

$$\mathbf{x} = \mathbf{z} \cdot \exp(-\theta \mathbf{s}) \tag{1}$$

where $\mathbf{z}$ is the background (what the measurement would be in the absence of plume), $\theta$ is a measure of gas concentration, and '$\cdot$' indicates elementwise multiplication of vectors. For low concentration gases, there is a linear approximation

$$\mathbf{x} = \mathbf{z} - \theta \mathbf{z} \cdot \mathbf{s} \tag{2}$$

which is often further approximated by

$$\mathbf{x} = \mathbf{z} - \theta \mathbf{z}_o \cdot \mathbf{s} = \mathbf{z} + \mathbf{t} \tag{3}$$

where $\mathbf{z}_o$ is the mean background spectrum over the image, and $\mathbf{t} = -\theta \mathbf{z}_o \cdot \mathbf{s}$ is the effectively additive target vector. Note that $\mathbf{x}$ and $\mathbf{z}$ vary from pixel to pixel in the image, but $\mathbf{s}$ and $\mathbf{z}_o$ are constant over the image.

Let $p(\mathbf{z})$ correspond to a likelihood model for the gas-free background. To optimally test for the existence of a target signal $\mathbf{t}$, we consider the binary hypothesis test

$$\mathcal{H}_o : \ \mathbf{x} = \mathbf{z} \qquad \sim p(\mathbf{x}) \tag{4}$$

$$\mathcal{H}_1 : \ \mathbf{x} = \mathbf{z} + \mathbf{t} \qquad \sim p(\mathbf{x} - \mathbf{t}) \tag{5}$$

which leads to a likelihood ratio detector:

$$\mathcal{L}(\mathbf{x}) = \frac{\ell(\mathbf{x} \mid \mathcal{H}_1)}{\ell(\mathbf{x} \mid \mathcal{H}_o)} = \frac{p(\mathbf{x} - \mathbf{t})}{p(\mathbf{x})}. \tag{6}$$

The simplest likelihood model is a Gaussian with mean $\boldsymbol{\mu}$ and covariance $R$, usually taken to be the sample mean and sample covariance over the image; this leads to:

$$\mathcal{L}(\mathbf{x}) = \frac{\exp\left[-(\mathbf{x} - \mathbf{t} - \boldsymbol{\mu})^T R^{-1} (\mathbf{x} - \mathbf{t} - \boldsymbol{\mu})/2\right]}{\exp\left[-(\mathbf{x} - \boldsymbol{\mu})^T R^{-1} (\mathbf{x} - \boldsymbol{\mu})/2\right]}, \tag{7}$$

or

$$2 \log \mathcal{L}(\mathbf{x}) = \left[(\mathbf{x} - \boldsymbol{\mu})^T R^{-1} (\mathbf{x} - \boldsymbol{\mu})\right] - \left[(\mathbf{x} - \boldsymbol{\mu} - \mathbf{t})^T R^{-1} (\mathbf{x} - \boldsymbol{\mu} - \mathbf{t})\right] \tag{8}$$

$$= 2\mathbf{t}^T R^{-1} (\mathbf{x} - \boldsymbol{\mu}) - \mathbf{t}^T R^{-1} \mathbf{t}. \tag{9}$$

which is equivalent to the adaptive matched filter:[14] $\mathcal{M}(\mathbf{x}) = \mathbf{t}^T R^{-1} \mathbf{x}$. Although it is not highlighted in this derivation, the adaptive matched filter is the same for any $\theta > 0$, which makes it a good detector for any strength plume. It is by using the adaptive matched filter (and in some cases going beyond the matched filter[15,16]), that gas detection has been so successful. However, it is still required that the chemistry (*i.e.,* the plume signature $\mathbf{s}$) be known.

But we are going to use this derivation as a starting point, and introduce some variants that are appropriate for the unknown gas problem.

The simplest (and least effective) treatment makes no assumptions at all about the target signal $\mathbf{t}$. Going back to the likelihood ratio in Eq. (6), we will treat $\mathbf{t}$ as a nuisance parameter and employ the *generalized* likelihood ratio test†

$$\mathcal{L}(\mathbf{x}) = \frac{\max_{\mathbf{t}} p(\mathbf{x} - \mathbf{t})}{p(\mathbf{x})}. \tag{10}$$

*In the thermal infrared, plumes can be either absorptive or emissive, and more sophisticated radiative transfer models are appropriate.[2,4,5,8]

†Although the generalized likelihood ratio test (GLRT) is widely used,[17,18] there is no guarantee that it is optimal, and it in some cases produces detectors that are inadmissible.[19] This has led to interest in other composite hypothesis testing approaches,[20–24] some of which also fail the test of admissibility.[25]

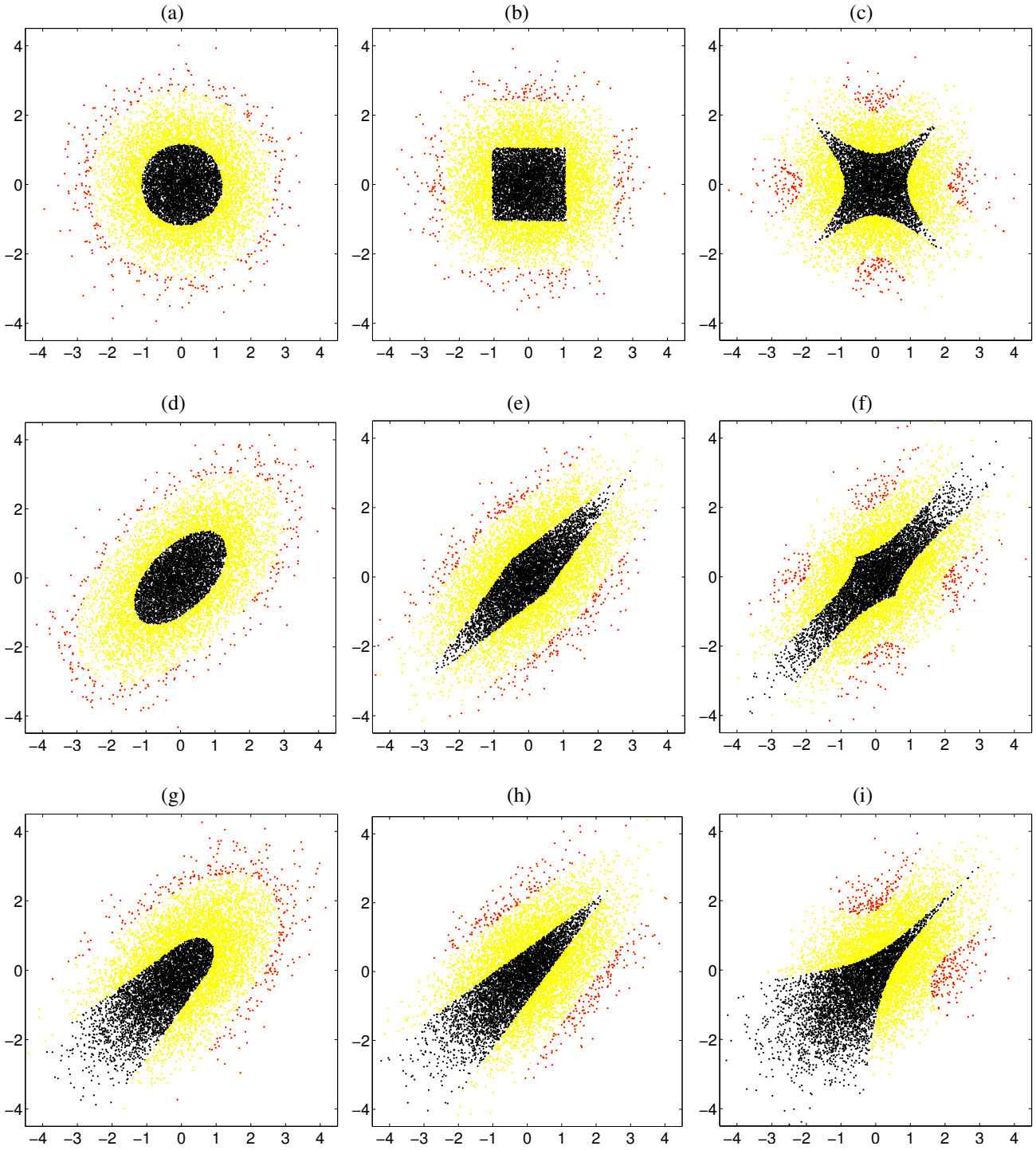



Figure 1. In this simple $d = 2$ dimensional model, we plot the most anomalous points (top three percent) in red, the least anomalous points (below median) in black, and the rest in yellow. A total of $10^4$ points are drawn at random from a Gaussian distribution. In (a-c) the distribution is Gaussian with covariance given by the identity; in (d-i) the Gaussian has a correlation of 0.3 in the off-diagonal component of the covariance matrix: the bottom row (g-i) restricts the anomalous target $\mathbf{t}$ to positive components. We take $k = d = 2$ in the leftmost panels (a,d,g); and for this non-sparse case, we observe the standard RX algorithm and in (a,d), with elliptical contours of anomalousness. For the sparse case ($k = 1$), we see diamond-shaped contours when the background is assumed to be Gaussian (b,e,h), but those diamonds become nonconvex as the background model becomes heavy-tailed (c,f,i). The heavy-tailed model is a multivariate-$t$ with $\nu = 3$.

(a) Indian Pines

(b) Cooke City

(c) Florida coast

Figure 2. **(a)** The 200-channel Indian Pines dataset[27] was cropped to $111 \times 111$ avoid some of the strong single-pixel anomalies that overpowered the rest of the image. **(b)** The 126 channel HyMap image of Cooke City, Montana, is the full $280 \times 800$ pixel blind reflectance dataset from the RIT blind test experiment.[28] **(c)** This AVIRIS dataset[29,30] was taken of the Florida coastline, near Cape Canaveral, flightline `f960323t01p02_r04_sc01`. This clip is $150 \times 500$ pixels, and all 224 channels were retained.

For the Gaussian model, we can write this (cf. Eq. (8))

$$2\log\mathcal{L}(\mathbf{x}) = \left[(\mathbf{x}-\boldsymbol{\mu})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu})\right] - \min_{\mathbf{t}}\left[(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})\right] \tag{11}$$

And with no restrictions at all on $\mathbf{t}$, we can minimize the expression on the right with $\mathbf{t} = \mathbf{x}-\boldsymbol{\mu}$, which yields the classical Mahalanobis distance anomaly detector (also known as global RX[26]):

$$\mathcal{A}(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu}). \tag{12}$$

## 2.1 Restrictions on the target signal

The most casual inspection of the gaseous absorption signatures in the Appendix indicates two properties that are common to all of them. The most striking property is that these signatures are spectrally sparse. For most of the spectral channels, the absorption coefficient is zero or very nearly zero. There is a second property which is so obvious that it seems entirely unremarkable; and that is that the signatures are non-negative at every wavelength. This second property leads to a considerable restriction on the target signal $\mathbf{t}$; every component must have the same sign (positive if the plume is in emission; negative if it is in absorption).

In general, we can impose restrictions on the target signal $\mathbf{t}$ by introducing the set $\mathcal{T}$ of permissible targets. Then, we can re-write Eq. (10) in terms of a constrained optimization:

$$\mathcal{L}(\mathbf{x}) = \frac{\max_{\mathbf{t}\in\mathcal{T}} p(\mathbf{x}-\mathbf{t})}{p(\mathbf{x})}. \tag{13}$$

which for Gaussian models becomes

$$2\log\mathcal{L}(\mathbf{x}) = \left[(\mathbf{x}-\boldsymbol{\mu})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu})\right] - \min_{\mathbf{t}\in\mathcal{T}}\left[(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})\right] \tag{14}$$

The positivity (or negativity) restriction on $\mathbf{t}$ is straightforward to exploit. Let $\mathcal{T}$ be the set of vectors with all-positive (or all-negative) components, and then Eq. (14) becomes a quadratic programming problem (which, however, must be solved at every pixel).

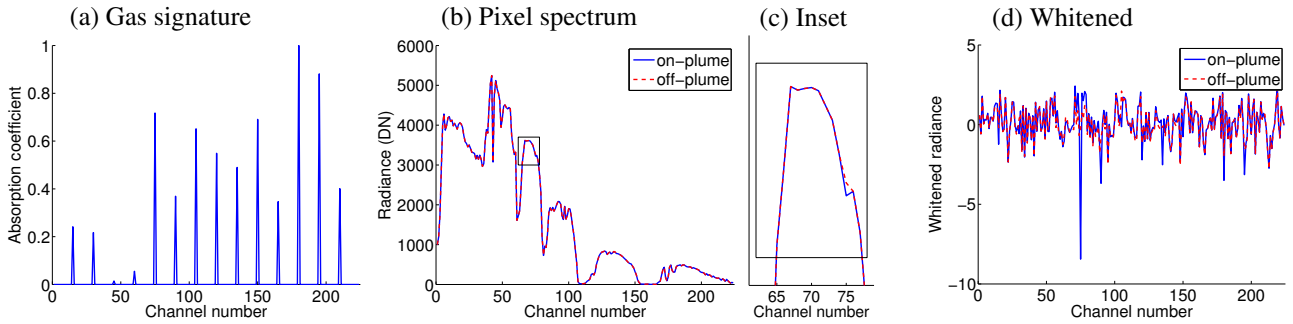



Figure 3. **(a)** Artificial chemical signature used to test spaRX algorithms; **(b)** A typical pixel spectrum in the Florida coast image; shown are the spectra with and without the effect of the plume. The difference is smaller than the thickness of the line used to make the plot. To see the size of the difference, the small box at about channel 70 is replotted in **(c)**; the actual difference is seen in channel 75. **(d)** After whitening, the difference between on-plume and off-plume is much more evident.

One way to impose a sparsity constraint is to only permit $\mathbf{t} \in \mathcal{T}_k$ where $\mathcal{T}_k$ is the set of all vectors with at most $k$ nonzero components. This leads to a sparse RX (or "spaRX") detector. If we write $W = R^{-1/2}$ as the whitening transform, and $\widetilde{\mathbf{x}} = W(\mathbf{x} - \boldsymbol{\mu})$ as the whitened pixel, then Eq. (14) can be written

$$2 \log \mathcal{L}(\mathbf{x}) = ||\widetilde{\mathbf{x}}||^2 - \min_{\mathbf{t} \in \mathcal{T}_k} ||\widetilde{\mathbf{x}} - W\mathbf{t}||^2 \tag{15}$$

The minimization in this case has the flavor of sparse coding in which one seeks to approximate the vector $\widetilde{\mathbf{x}}$ with a linear combination of $k$ columns of the dictionary $W$. Since there are so many combinations of $k$ components, particularly when the number of spectral channels is large, full optimization is impractical. For most of the experiments reported here, we employed an orthogonal matching pursuit algorithm[31] which greedily incorporated columns of $W$, one item at a time, each time providing the best match to the residual, and followed by re-fitting of coefficients.

However, we also investigated more principled convex optimization approximation approaches. Here, we replace what amounts to an $\ell_0$ constraint on $\mathbf{t}$ (namely, that $||\mathbf{t}||_0 = \sum_i |t_i|^0 \leq k$, where $t_i$ is the $i$th component of $\mathbf{t}$) with an $\ell_1$ constraint (namely, that $||\mathbf{t}||_1 = \sum_i |t_i| \leq \tau$, for some $\tau$). This is a variant of the lasso[32] with a weighted $\ell_2$ objective function, and can either be solved using the SPGL1 code,[33] or via an efficient ADMM algorithm[34] by expressing the constraint as a penalty using the indicator function of the constraint set.

We remark that the sparse anomaly detection problem that we are formulating here differs from some others that have appeared in the literature. What we are seeking is an additive anomaly with sparse components, not a sparse matched filter,[35] nor a sparse set of features features to describe $\mathbf{x}$ followed by a kernel-based anomaly detector to in that low-dimensional space.[36] This approach is closer to that of a low-rank plus sparse decomposition,[37,38] but in place of a low-rank subspace, we have a full-rank distribution $p(\mathbf{x})$.

## 2.2 Heavy-tailed elliptically-contoured distributions

The Gaussian distribution is a convenient model, but it is not an accurate one. In particular, the tails of the Gaussian distribution do not match the distributions observed in real data. This especially matters for problems in anomaly detection, where false alarm rates are required to be low, even at the expense of lower detection rates.

Because the detectors we have described so far have all been derived from a background model that is an arbitrary distribution $p(\mathbf{x})$, we are free to employ non-Gaussian distributions in our sparse anomaly detectors.

Elliptically contoured (EC) distributions provide an attractive alternative to the Gaussian, and their utility for hyperspectral data in particular has long been argued by Manolakis *et al.*[39,40] A number of classic algorithms in hyperspectral data analysis have been derived using a Gaussian distribution as a starting point; replacing the Gaussian with an EC distribution can produce more effective variants.[41–43] And that is just what we propose to do in this subsection.

For an EC distribution, we can write

$$p(\mathbf{x}) = c|R|^{-1/2}\, h\left((\mathbf{x} - \boldsymbol{\mu})^T R^{-1} (\mathbf{x} - \boldsymbol{\mu})\right) \tag{16}$$

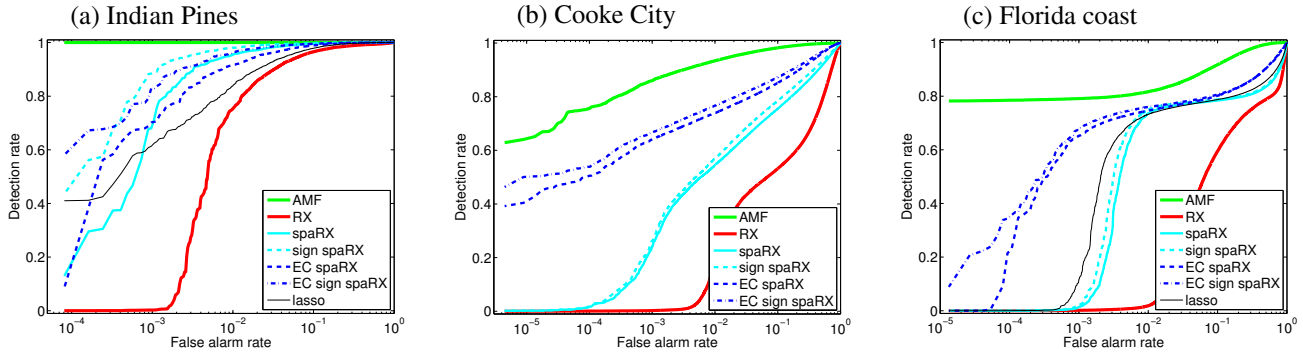



Figure 4. ROC curves compare the performance of different variants of spaRX to AMF (which requires that the gas signature be known) and to global RX (which makes no assumptions at all about the gas signature). For all the spaRX variants in these plots, we used $k = 2$. The plain spaRX assumes that the target signal $\mathbf{t}$ is sparse and that the background is Gaussian; the "sign spaRX" permits only negative values for $\mathbf{t}$; the "EC spaRX" uses an elliptically contoured background model, and "EC sign spaRX" uses the EC background and the signed target signal. The "lasso" curves (when available), were obtained by optimizing Eq. (14) subject to a constraint on the $\ell_1$ metric of $\mathbf{t}$.

where $c$ is a normalizing constant, $|R|$ is the determinant of the covariance matrix, and $h$ is a scalar-valued function of a non-negative scalar argument; we will further assume that $h$ decreases monotonically with increasing argument. For instance, $h(\xi) = \exp(-\xi/2)$ is the Gaussian distribution, and

$$h(\nu;\xi) = \left(1 + \frac{\xi}{\nu - 2}\right)^{-(d+\nu)/2} \tag{17}$$

is the multivariate-$t$ distribution with $\nu$ degrees of freedom, where $d$ is the dimension (or number of spectral channels).[40] The $\nu \to \infty$ limit leads back to the Gaussian, but for finite $\nu$, this distribution has fatter tails than the Gaussian. It is furthermore fairly straightforward to estimate $\nu$ from the data.[40,43]

Our constrained optimization in Eq. (14) becomes, for these more general distributions

$$\mathcal{A}(\mathbf{x}) = H\big(\,(\mathbf{x}-\boldsymbol{\mu})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu})\,\big) - \min_{\mathbf{t}\in\mathcal{T}} H\big(\,(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})\,\big) \tag{18}$$

where

$$H(\xi) = \frac{2}{d+\nu}\log h(\xi) = -\log\left(1 + \frac{\xi}{\nu-2}\right). \tag{19}$$

One place where EC distributions do *not* help is the straightforward anomaly detection problem. Here, $\mathcal{T}$ is the set of *all* possible vectors $\mathbf{t}$, and the minimum is achieved by $\mathbf{t} = \mathbf{x} - \boldsymbol{\mu}$, which leads to

$$\mathcal{A}(\mathbf{x}) = H\big(\,(\mathbf{x}-\boldsymbol{\mu})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu})\,\big). \tag{20}$$

but since $H$ is a monotonic function, this is equivalent to the Mahalanobis distance.

As soon as we begin placing restrictions on $\mathbf{t}$, however, the choice of $H$ does affect the anomaly detector. The detector in Eq. (18) is qualitatively different from the detector in Eq. (14), and has different contours of anomalousness. Compare Fig. 1(b,e,h) to Fig. 1(c,f,i) for an example in the simple case of two spectral channels. The algorithm complexity, however, is not substantially affected because the minimization problem (which is the most significant contributor to algorithm complexity) is the same for Gaussian as for general EC distributions. That's because $H$ is monotonic, and therefore

$$\min_{\mathbf{t}\in\mathcal{T}} H\big(\,(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})\,\big) = H\left(\min_{\mathbf{t}\in\mathcal{T}}\left[(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})^T R^{-1}(\mathbf{x}-\boldsymbol{\mu}-\mathbf{t})\right]\right). \tag{21}$$

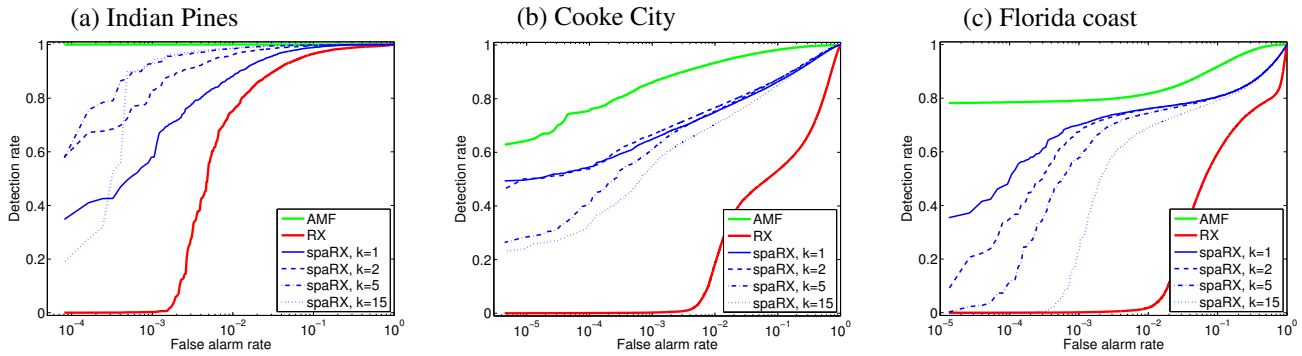



Figure 5. ROC curves show the performance of sparse RX ("spaRX") compared to RX and to AMF, which depends on knowing the gas signature. The spaRX curves assume negatively signed anomalies, and an EC background distribution (multivariate-$t$, with with $\nu$ estimated from the data). Different values of $k$ are used in the matching pursuit estimation process. The best choice of $k$ is generally seen to be smaller than the number of peaks in the chemical signatures (which are shown in Fig. 3).

## 3. RESULTS

We performed experiments using simulated chemical plumes on three different background hyperspectral images, shown in Fig. 2. The plume signatures were generated with nonzero values at every fifteen spectral channels, with coefficients generated randomly, and scaled by the largest coefficient. This process led to the spectra shown in Fig. 3.

We used a matched-pair approach to evaluate the performance of the detection algorithms. For each image, a duplicate image was made, and an absorptive chemical plume was simulated on every pixel in the duplicate image, using Eq. (1) with $\theta = 0.02$ and $\mathbf{s}$ shown in Fig. 3. The non-plume image was used for "training" (that is, for computing the mean and covariance), and the detector was applied to both the non-plume and plume images. For a given threshold, the number of detections in the non-plume image provides a false alarm rate, and the number of detection in the plume image gives a detection rate. A receiver-operating characteristic (ROC) curve plots detection rate as a function of false alarm rate.

As Fig. 4 shows, all of the spaRX algorithms outperformed the plain RX algorithm. Both the signed-spaRX and EC-spaRX algorithms generally outperformed the original spaRX, and EC-signed-spaRX was generally the best of them all. The lasso performance was mixed. It consistently outperformed RX, and was generally competitive with the other spaRX algorithms, at least in the low false-alarm rate regime. But it was not the best detector, and in our current implementation, it is by far the slowest detector, and choosing the parameter $\tau$ is much less intuitive than choosing the parameter $k$ that was used by the matching pursuit algorithms.

In Fig. 5 we investigated how performance varied with choice of parameter $k$. This is a free parameter in our algorithm, and we have no *a priori* guidance for its choice, since we do not know which gas might be present in the scene. It nominally corresponds to the number of nonzero channels in our absorption spectrum, but what this experiment shows is that RX is outperformed over a wide range of $k$ values, and that the best values tend to be smaller than the number of nonzero channels. Since the matching pursuit algorithms are faster for small $k$ anyway, we chose $k = 2$ as our default parameter. Fig. 6 confirms that the enhanced performance of the EC variants of spaRX are due to the heavy-tailed background.

## 4. CONCLUSION

We investigated the utility of spectrally sparse anomaly detection[13] to the problem of blind gas detection in hyperspectral imagery. Since we do not know the chemical species of the gas in the scene, this amounts to a problem in anomaly detection. But while anomalies are traditionally treated as completely unknown, we do know something about gas absorption spectra. Our aim in this study has been to leverage the observation that these spectra tend to be sparse, with most spectral channels exhibiting zero or near-zero absorption. Exploiting this sparsity leads to the sparse RX (or spaRX) algorithm.

In addition to the basic spaRX algorithm, we considered two potential improvements. One addresses a property of the target signal, and one addresses a property of the background clutter. The "sign spaRX" algorithm exploits the knowledge that the sign of the components of $\mathbf{t}$ are all either positive (if plume is emissive) or negative (if plume is absorptive). The "EC spaRX" algorithm derives the expression for anomalousness from the likelihood ratio using an elliptically-contoured distribution in place of a Gaussian; specifically, we employed the multivariate-$t$ distribution with the parameter $\nu$ estimated

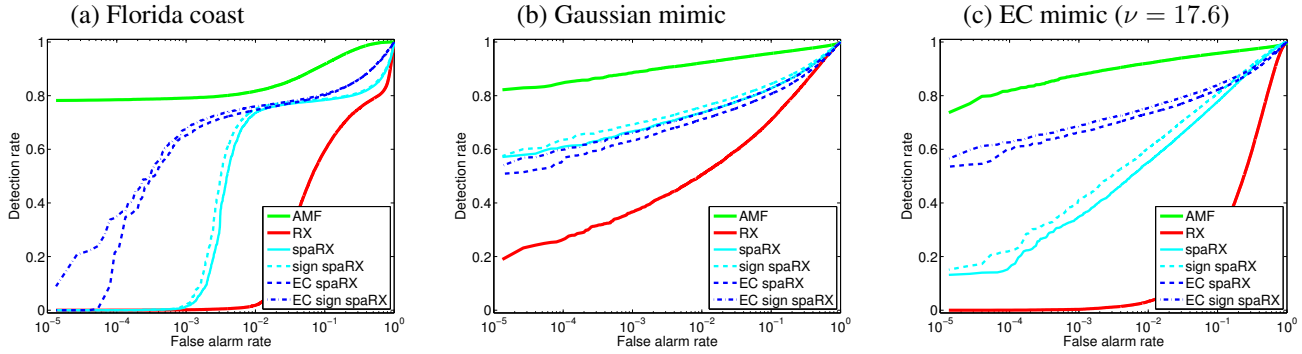



Figure 6. To assess to role of the background distribution in the relative performance of the spaRX and EC-spaRX algorithms, we considered three datasets. **(a)** The first is the orignal Florida coast dataset, and this plot is the same as Fig. 4(c). **(b)** This dataset is multivariate Gaussian with the same covariance as the Florida coast data. We see here that the EC enhancement to the spaRX algorithm does not improve its performance. (Overall, we see that performance is better on Gaussian data than on the original data; this further illustrates that the Gaussian approximation is not a close approximation.) **(c)** This dataset is multivariate-$t$ with the same covariance matrix and the same $\nu$ as the orignal Florida coast data. In contrast to what was observed in (b), the EC-spaRX algorithms substantially outperform the non-EC spaRX.

from the data. One can furthermore employ both of these improvements, leading to the "EC sign spaRX" algorithm, which generally was best among the spaRX variants. But for the Cooke City and Florida coast images, in particular, the EC gain was far larger than the signed signal gain.

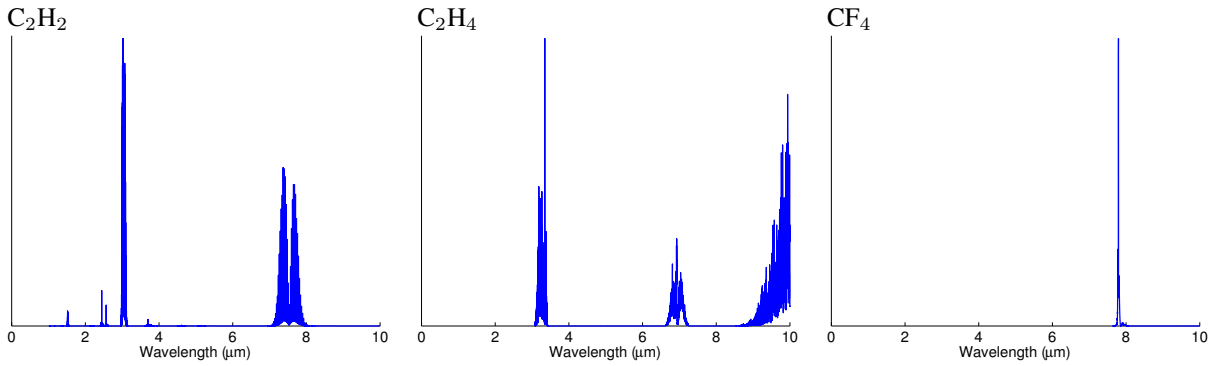
Our spaRX algorithm, and its variants, all begin with the global RX algorithm as the underlying anomaly detector. In global RX, the target-free pixel is estimated by mean over the whole image, and its natural variability is characterized by the global covariance matrix.

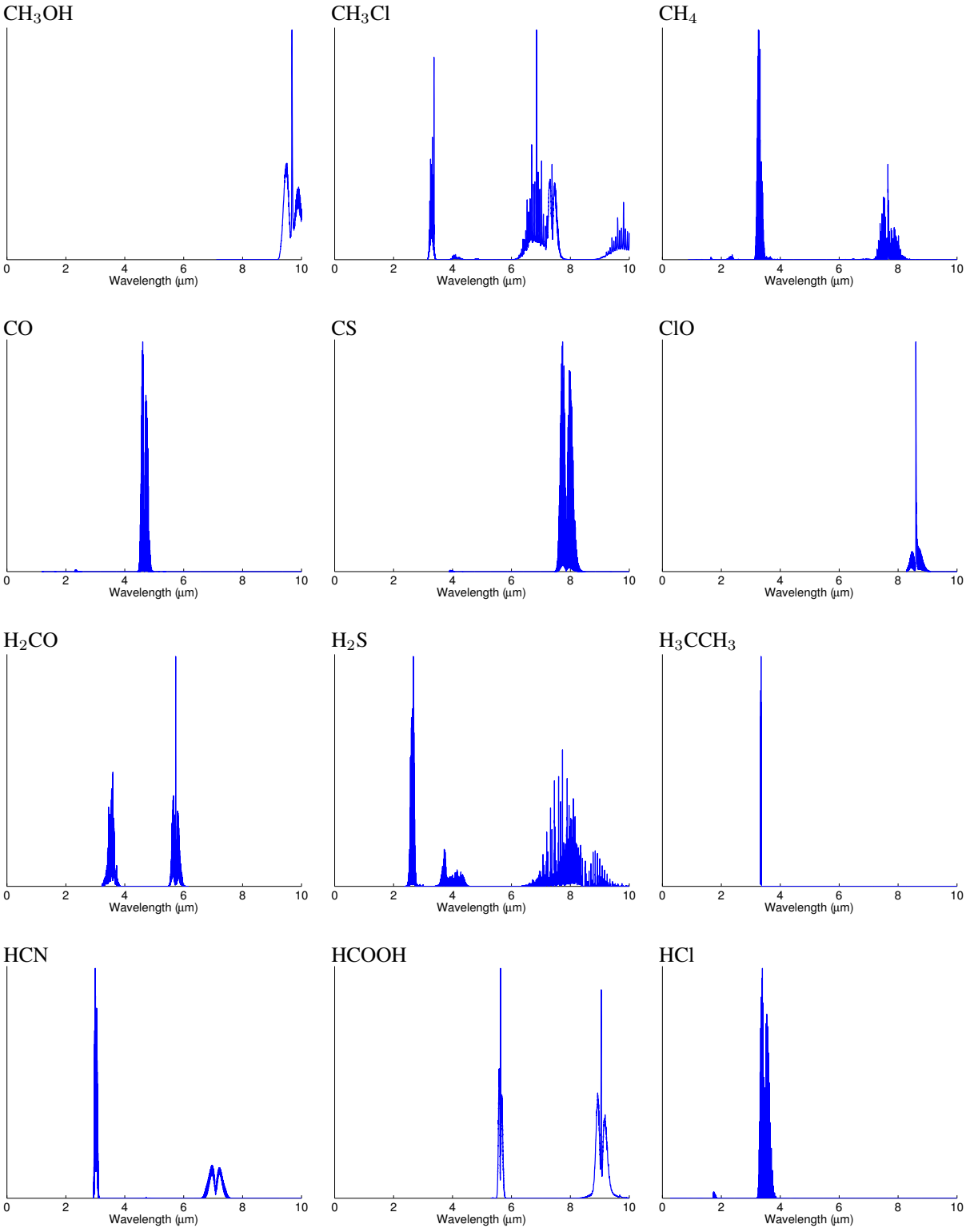
Most modern anomaly detection algorithms[44–46] (including the original RX[26]) use the mean over a local annulus to estimate the target-free pixel, and this can provide a substantial improvement. Locally adaptive covariance estimators have also been suggested.[26,44,45,47,48] But these are most appropriate for spatially compact targets, and that may or may not be appropriate in the gas detection problem. Also in our comparisons of relative performance of spaRX and AMF, these local methods would also improve the AMF results.

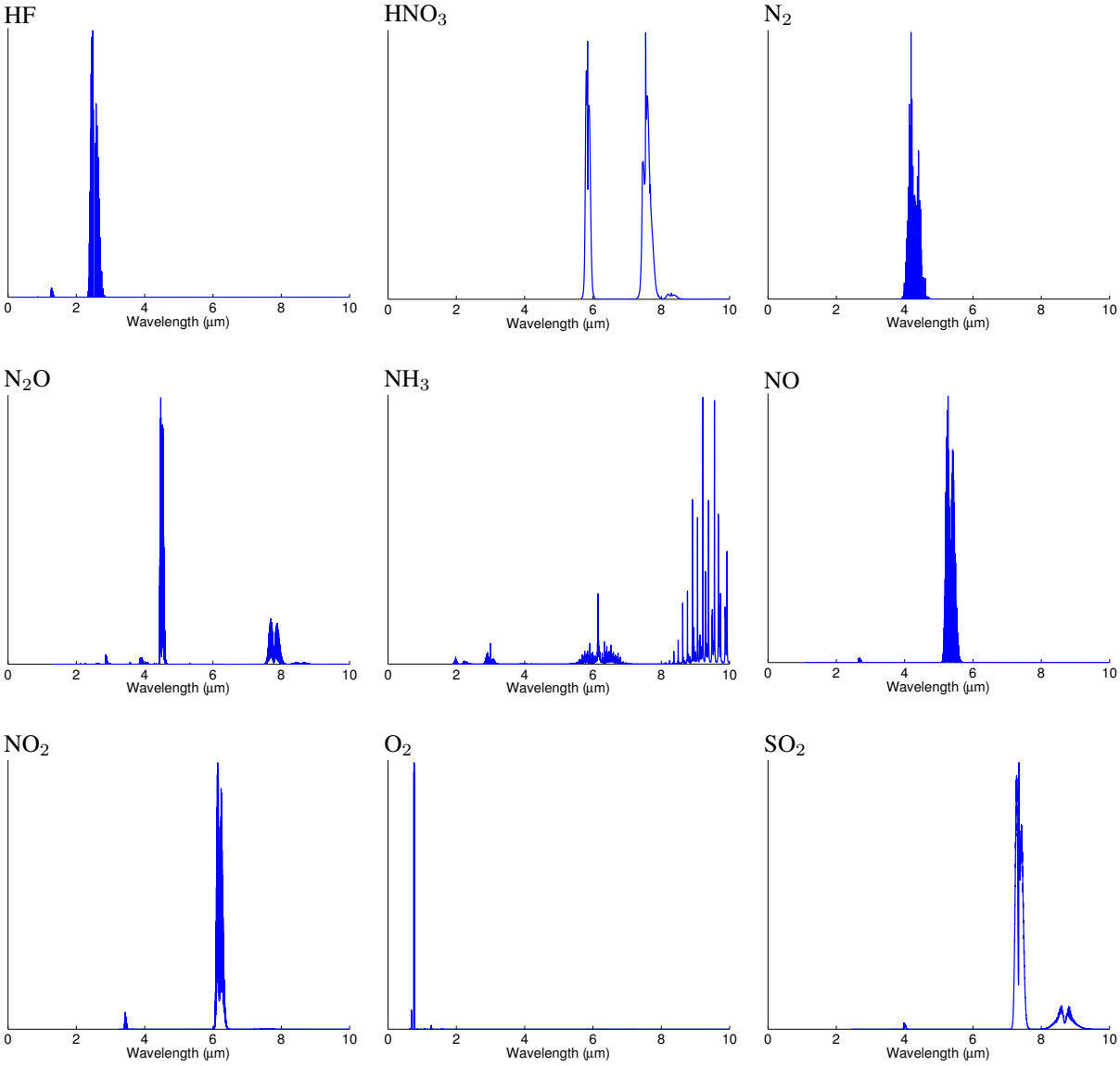
The substantial performance enhancement provided by the EC variant of spaRX suggests that a closer look at different ways of characterizing the heavy tails[49,50] might prove valuable. Other potential extensions are to kernelized anomaly detectors[51] or to anomalous change detection.[52–54]

## APPENDIX: GAS SPECTRA

In this appendix, we plot a number of examples of gas absorption spectra. These spectra were drawn from the HITRAN database.[55] Relative absorption coefficients are shown as a function of wavelength. These spectra are quite distinct from each other, which is why hyperspectral imagery is such an effective tool for chemical identification of plumes. Furthermore, for virtually all of the gases, the absorption coefficient is nonzero for only a sparse subset of the wavelengths.




$CH_3OH$
$CH_3Cl$
$CH_4$
CO
CS
ClO
$H_2CO$
$H_2S$
$H_3CCH_3$
HCN
HCOOH
HCl
0 2 4 6 8 10
Wavelength (μm)




## ACKNOWLEDGMENTS

This work was supported by the Laboratory Directed Research and Development (LDRD) program at Los Alamos National Laboratory.

## REFERENCES

1. C. L. Bennett, M. R. Carter, D. J. Fields, and F. D. Lee, "Infrared hyperspectral imaging results from vapor plume experiments," *Proc. SPIE* **2480**, pp. 435–444, 1995.
2. A. Hayden, E. Niple, and B. Boyce, "Determination of trace-gas amounts in plumes by the use of orthogonal digital filtering of thermal-emission spectra," *Applied Optics* **35**, pp. 2802–2809, 1996.
3. P. V. Villeneuve, H. A. Fry, J. Theiler, B. W. Smith, and A. D. Stocker, "Improved matched-filter detection techniques," *Proc. SPIE* **3753**, pp. 278–285, 1999.
4. S. J. Young, "Detection and quantification of gases in industrial-stack plumes using thermal-infrared hyperspectral imaging," Tech. Rep. ATR-2002(8407)-1, The Aerospace Corporation, 2002.

5. B. R. Foy, R. R. Petrin, C. R. Quick, T. Shimada, and J. J. Tiee, "Comparisons between hyperspectral passive and multispectral active sensor measurements.," *Proc. SPIE* **4722**, pp. 98–109, 2002.
6. B. R. Foy, H. A. Fry, and B. D. McVey, "Approaches to chemical plume detection in hyperspectral infrared imaging," Tech. Rep. LA-CP-03-0103, Los Alamos National Laboratory, 2003.
7. B. R. Foy and J. Theiler, "Scene analysis and detection in thermal infrared remote sensing using Independent Component Analysis," *Proc. SPIE* **5439**, pp. 131–139, 2004.
8. J. Theiler, B. R. Foy, and A. M. Fraser, "Characterizing non-Gaussian clutter and detecting weak gaseous plumes in hyperspectral imagery," *Proc. SPIE* **5806**, pp. 182–193, 2005.
9. D. G. Manolakis and F. M. D'Amico, "A taxonomy of algorithms for chemical vapor detection with hyperspectral imaging spectroscopy," *Proc. SPIE* **5795**, pp. 125–133, 2005.
10. D. Manolakis, "Signal processing algorithms for hyperspectral remote sensing of chemical plumes," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* , pp. 1857–1860, 2008.
11. C. M. Gittins, "Detection and characterization of chemical vapor fugitive emissions by nonlinear optimal estimation: theory and simulation," *Applied Optics* **48**, pp. 4545–4561, 2009.
12. S. Niu, S. E. Golowich, V. K. Ingle, and D. G. Manolakis, "Design and evaluation of robust matched filters for chemical agent detection," *Proc. SPIE* **8186**, p. 81860R, 2011.
13. J. Theiler and B. Wohlberg, "Detection of spectrally sparse anomalies in hyperspectral imagery," *Proc. IEEE Southwest Symposium on Image Analysis and Interpretation* , pp. 117–120, 2012.
14. F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerospace and Electronic Systems* **28**, pp. 208–216, 1992.
15. A. Schaum and E. Allman, "Nonlinear signature-based hyperspectral target detection: Beyond the matched filter," *Proc. MSS (Military Sensing Symposium) on Camouflage, Concealment, and Deception* , 2004.
16. J. Theiler, B. R. Foy, and A. M. Fraser, "Beyond the adaptive matched filter: nonlinear detectors for weak signals in high-dimensional clutter," *Proc. SPIE* **6565**, p. 656503, 2007.
17. E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerospace and Electronic Systems* **22**, pp. 115–127, 1986.
18. S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. II, Prentice Hall, New Jersey, 1998.
19. E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, Springer, New York, 2005.
20. A. Schaum, "Hyperspectral target detection using a Bayesian likelihood ratio test," *Proc. IEEE Aerospace Conference* **3**, pp. 1537–1540, 2002.
21. A. Schaum, "Continuum fusion: a theory of inference, with applications to hyperspectral detection," *Optics Express* **18**, pp. 8171–8181, 2010.
22. A. P. Schaum and B. J. Daniel, "Continuum fusion methods of spectral detection," *Optical Engineering* **51**, p. 111718, 2012.
23. P. Bajorski, "Min-max detection fusion for hyperspectral images," *Proc. 3rd IEEE Worskhop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)* , 2011.
24. J. Theiler, "Formulation for min-max clairvoyant fusion based on monotonic recalibration of statistics," *Optical Engineering* **51**, p. 111714, 2012.
25. J. Theiler, "Confusion and clairvoyance: some remarks on the composite hypothesis testing problem," *Proc. SPIE* **8390**, p. 839003, 2012.
26. I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoustics, Speech, and Signal Processing* **38**, pp. 1760–1770, 1990.
27. D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*, John Wiley & Sons, 2003.
28. D. Snyder, J. Kerekes, I. Fairweather, R. Crabtree, J. Shive, and S. Hager, "Development of a web-based application to evaluate target finding algorithms," *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* **2**, pp. 915–918, 2008.
29. G. Vane, R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen, and W. M. Porter, "The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)," *Remote Sensing of the Environment* **44**, pp. 127–143, 1993.
30. AVIRIS Free Standard Data Products, Jet Propulsion Laboratory (JPL), National Aeronautics and Space Administration (NASA). `http://aviris.jpl.nasa.gov/html/aviris.freedata.html`.
31. S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing* **41**, pp. 3397–3415, 1993.

32. R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)* **58**(1), pp. 267–288, 1996.
33. E. van den Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing* **31**(2), pp. 890–912, 2008.
34. 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* **3**(1), pp. 1–122, 2010.
35. J. Theiler and K. Glocer, "Sparse linear filters for detection and classification in hyperspectral imagery," *Proc. SPIE* **6233**, p. 62330H, 2006.
36. A. Banerjee, R. Juang, J. Broadwater, and P. Burlina, "Sparse feature extraction for support vector data description applications," *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* , pp. 4236–4239, 2010.
37. E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *J. ACM* **58**, pp. 11:1–11:37, 2011.
38. B. Wohlberg, R. Chartrand, and J. Theiler, "Local principal component pursuit for nonlinear datasets," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* , pp. 3925–3928, 2012.
39. D. Manolakis, D. Marden, J. Kerekes, and G. Shaw, "On the statistics of hyperspectral imaging data," *Proc. SPIE* **4381**, pp. 308–316, 2001.
40. D. B. Marden and D. Manolakis, "Modeling hyperspectral imaging data," *Proc. SPIE* **5093**, pp. 253–262, 2003.
41. A. Schaum, E. Allman, J. Kershenstein, and D. Alexa, "Hyperspectral change detection in high clutter using elliptically contoured distributions," *Proc. SPIE* **6565**, p. 656515, 2007.
42. J. Theiler and B. R. Foy, "EC-GLRT: Detecting weak plumes in non-Gaussian hyperspectral clutter using an elliptically-contoured generalized likelihood ratio test," *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* , p. I:221, 2008.
43. J. Theiler, C. Scovel, B. Wohlberg, and B. R. Foy, "Elliptically-contoured distributions for anomalous change detection in hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters* **7**, pp. 271–275, 2010.
44. D. W. J. Stein, S. G. Beaven, L. E. Hoff, E. M. Winter, A. P. Schaum, and A. D. Stocker, "Anomaly detection from hyperspectral imagery," *IEEE Signal Processing Magazine* **19**, pp. 58–69, Jan 2002.
45. S. Matteoli, M. Diani, and G. Corsini, "A tutorial overview of anomaly detection in hyperspectral images," *IEEE A&E Systems Magazine* **25**, pp. 5–27, 2010.
46. N. Gorelnik, H. Yehudai, and S. R. Rotman, "Anomaly detection in non-stationary backgrounds," *Proc. 2nd IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)* , 2010.
47. C. E. Caefer, J. Silverman, O. Orthal, D. Antonelli, Y. Sharoni, and S. R. Rotman, "Improved covariance matrices for point target detection in hyperspectral data," *Optical Engineering* **47**, p. 076402, 2008.
48. L. Bachega, J. Theiler, and C. A. Bouman, "Evaluating and improving local hyperspectral anomaly detectors," *Proc. 40th IEEE Applied Imagery Pattern Recognition (AIPR) Workshop* , 2011.
49. S. M. Adler-Golden, "Improved hyperspectral anomaly detection in heavy-tailed backgrounds," *Proc. 1st IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)* , 2009.
50. J. Theiler and D. Hush, "Statistics for characterizing data on the periphery," *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* , pp. 4764–4767, 2010.
51. H. Kwon and N. M. Nasrabadi, "Kernel RX: a new nonlinear anomaly detector," *Proc. SPIE* **5806**, pp. 35–46, 2005.
52. A. Schaum and E. Allman, "Advanced algorithms for autonomous hyperspectral change detection," *IEEE Applied Imagery Pattern Recognition (AIPR) Workshop: Emerging technologies and applications for imagery pattern recognition* **33**, pp. 33–38, 2005.
53. M. T. Eismann, J. Meola, A. D. Stocker, S. G. Beaven, and A. P. Schaum, "Airborne hyperspectral detection of small changes," *Applied Optics* **47**, pp. F27–F45, 2008.
54. J. Theiler, "Quantitative comparison of quadratic covariance-based anomalous change detectors," *Applied Optics* **47**, pp. F12–F26, 2008.
55. L. Rothman and S. Tashkun, "HITRAN on the web." Harvard-Smithsonian Center for Astrophysics (CFA), Cambridge, MA, USA; and V.E. Zuev Insitute of Atmosperic Optics (IAO), Tomsk, Russia. `http://hitran.iao.ru/`.