Nonlinear signal contamination effects for gaseous plume detection in hyperspectral imagery

James Theiler, Bernard R. Foy, and Andrew M. Fraser

Los Alamos National Laboratory, Los Alamos, NM 87544

ABSTRACT

When a matched filter is used for detecting a weak target in a cluttered background (such as a gaseous plume in a hyperspectral image), it is important that the background clutter be well-characterized. A statistical characterization can be obtained from the off-plume pixels of a hyperspectral image, but if on-plume pixels are inadvertently included, then that background characterization will be contaminated. In broad area search scenarios, where detection is the central aim, it is by definition unknown which pixels in the scene are off-plume, so some contamination is inevitable. In general, the contaminated background degrades the ability of the matched-filter to detect that signal. This could be a practical problem in plume detection. A linear analysis suggests that the effect is limited, and actually vanishes in some cases. In this study, we take into account the Beer's Law nonlinearity of plume absorption, and we investigate the effect of that nonlinearity on the signal contamination.

Keywords: matched filter, hyperspectral imagery, adaptive signal detection, clutter, signal contamination

1. INTRODUCTION

Hyperspectral images, by providing a high-resolution spectrum at each pixel, enable the detection of weak signals[1] – whether they are thin gaseous plumes or small subpixel targets – even in cluttered scenes. This sensitivity arises from the large number of spectral channels, and the potential for appropriate combinations of these spectral channels to simultaneously enhance the signal and suppress the background. Suppressing the background is key, but that suppression requires that the background be sufficiently well characterized in the first place. For the adaptive matched filter,[2] this characterization is in terms of the covariance of the background clutter, which is generally estimated from background data (*i.e.,* off-plume pixels, or "secondary vectors"). But it is not always known *a priori* whether a given pixel or vector really is in the background, or whether it includes the desired signal. Since it may not be possible to avoid this contamination of the background, it is important to understand its effect, particularly on the performance of algorithms for detecting the signal of interest. This problem has been investigated by a number of authors,[3–5] but usually in the domain where the signal is linearly superimposed on the background. In this paper, we will investigate the effects of nonlinearity on the covariance contamination problem.

1.1. Derivation of main plume formula

A downward-looking sensor operating in the long-wavelength infrared (LWIR) can detect a plume by analyzing how the radiance emitted and absorbed by the plume contrasts with the radiation emitted from the ground. For a sensor that operates at shorter wavelengths, the radiation is predominantly reflected solar, and the effect of the plume is predominantly absorptive. In this section, we will derive an expression that covers both cases, but in our examples we will concentrate on the purely absorptive case.

Following the derivations in Young[6] and in Foy *et al.*,[7] we note that the off-plume contributions to the observed radiance include thermal emission from atmosphere and instrument* ($\mathbf{L}^{\text{path}}$) and radiance from the ground ($\mathbf{L}^{\text{gnd}}$) attenuated by atmospheric absorption ($\boldsymbol{\tau}^{\text{atm}}$). That is,

$$\mathbf{L}^{\text{off-plume}} = \mathbf{L}^{\text{path}} + \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}}, \tag{1}$$

Email: {jt,bfoy,afraser}@lanl.gov

*Instrumental and other sources of noise are included in this term as well.

LA-UR-06-1996 Submitted to: *Proc. SPIE* **6233** (2006)

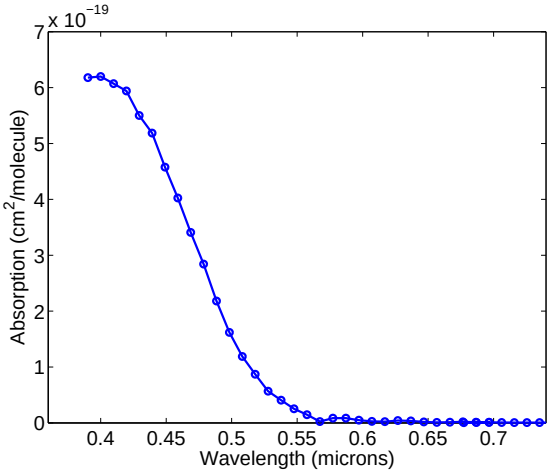

Figure 1. Absorption spectrum $\boldsymbol{\alpha}$ for NO_2 as a function of wavelength, λ, over a range that includes the first forty AVIRIS channels; for the remaining channels, NO_2 does not absorb ($\alpha = 0$). The full range of the AVIRIS spectral channels is from 0.4μ to 2.5μ.

where each variable nominally depends on the wavelength λ of the radiation, and varies from pixel to pixel in the scene. (In practice, however, $\mathbf{L}^{\text{path}}$ and $\boldsymbol{\tau}^{\text{atm}}$ do not generally vary with position over a single image.) The on-plume radiance is augmented by direct thermal emission from the plume ($\mathbf{L}^{\text{plume}}$), and reduced by absorption of the ground radiance due to the plume (given by transmissivity $\boldsymbol{\tau}^{\text{plume}}$). So the total radiance is given by

$$\mathbf{L} = \mathbf{L}^{\text{path}} + \mathbf{L}^{\text{plume}}\boldsymbol{\tau}^{\text{atm}} + \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}}\boldsymbol{\tau}^{\text{plume}}. \tag{2}$$

The thermal emission from the plume is a product of its emissivity $\boldsymbol{\epsilon}^{\text{plume}} = 1 - \boldsymbol{\tau}^{\text{plume}}$ and its blackbody radiation $\mathbf{B}(T^{\text{plume}})$. That is, $\mathbf{L}^{\text{plume}} = \mathbf{B}(T^{\text{plume}})(1 - \boldsymbol{\tau}^{\text{plume}})$, and it leads to

$$\mathbf{L} = \left[\mathbf{L}^{\text{path}} + \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}}\right] - \left[\mathbf{L}^{\text{gnd}} - \mathbf{B}(T^{\text{plume}})\right]\boldsymbol{\tau}^{\text{atm}}(1 - \boldsymbol{\tau}^{\text{plume}}). \tag{3}$$

Beer's Law[8] says that the transmission of light through a diffuse material decreases exponentially with the amount of material. That is,

$$\boldsymbol{\tau}^{\text{plume}} = \exp(-n_c\boldsymbol{\alpha}) \tag{4}$$

where $\boldsymbol{\alpha}$ is the absorption spectrum[†] for the material (*e.g.,* see Fig. 1), and n_c is the column density; for plumes, column density is generally expressed in units of parts-per-million-meters (ppm-m). This leads to

$$\mathbf{L} = \mathbf{L}^{\text{path}} + \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}} - \left[\mathbf{L}^{\text{gnd}} - \mathbf{B}(T^{\text{plume}})\right]\boldsymbol{\tau}^{\text{atm}}(1 - \exp(-n_c\boldsymbol{\alpha})), \tag{5}$$

which expresses the observed radiance $\mathbf{L}$, in terms of the ground clutter, atmosphere, and plume properties.

1.2. Variants of the plume formula

It will be convenient to encode the effect of the Beer's Law nonlinearity in the following definition:

$$\boldsymbol{\gamma} \equiv \frac{1 - \exp(-n_c\boldsymbol{\alpha})}{n_c} \tag{6}$$

Here, $\boldsymbol{\gamma}$ is a kind of distorted absorption curve; in the weak (small n_c) limit, we have $\boldsymbol{\gamma} = \boldsymbol{\alpha}$. This enables us to rewrite the total radiance in Eq. (5) as

$$\mathbf{L} = \mathbf{L}^{\text{path}} + \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}} - n_c\left[\mathbf{L}^{\text{gnd}} - \mathbf{B}(T^{\text{plume}})\right]\boldsymbol{\tau}^{\text{atm}}\boldsymbol{\gamma}. \tag{7}$$

This expression appears to be linear in n_c, but that's because the nonlinearity is hidden in Eq. (6). Still, if one is seeking a plume of a strength that is known *a priori*, let's say n_o, then one can define a fixed signature $\boldsymbol{\gamma} = [1 - \exp(-n_o\boldsymbol{\alpha})]/n_o$ and use a linear detector against Eq. (7).

[†]Unfortunately, the prevailing convention is to express $\boldsymbol{\alpha}$ in units for which $\log_{10}\boldsymbol{\tau}^{\text{plume}} = -n_c\boldsymbol{\alpha}$. This introduces an extra factor of $\log_e 10 \approx 2.3026$ that is built into the units of the absorption $\boldsymbol{\alpha}$, but otherwise does not affect our discussion.

At visible wavelengths, $\mathbf{L}^{\text{path}}$ and $\mathbf{B}(T^{\text{plume}})$ effectively vanish. There is no emission from the plume, but it is still absorptive, and in this absorptive-only situation, the expression in Eq. (5) reduces to

$$\begin{aligned} \mathbf{L} &= \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}}\exp(-n_c\boldsymbol{\alpha}) && (8)\\ &= \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}}(1-n_c\boldsymbol{\gamma})\,. && (9)\end{aligned}$$

This simplified expression also neglects atmospheric scattering and sensor noise, which are present at shorter wavelengths, but these are usually dominated by ground clutter variability.

2. LINEAR ALGEBRA

For the purposes of mathematical abstraction, we can write the linearized plume equation in terms of an additive signal on a cluttered background. In particular, we can write

$$\mathbf{r} = \mathbf{z} - \epsilon\mathbf{b} \tag{10}$$

where the vector $\mathbf{r} \in \mathbb{R}^d$ is the observed radiance[‡]; each component of the vector corresponds to a measurement at a given wavelength. Similarly, $\mathbf{z} \in \mathbb{R}^d$ is the vector-valued contribution to the radiance that is due to background clutter. The scalar ϵ corresponds to the strength of the signal of interest (*i.e.,* the plume), and the vector $\mathbf{b} \in \mathbb{R}^d$ is the spectral signature of the signal.

We can associate the terms in Eq. (10) with the linearized physical plume equation given in Eq. (7):

$$\begin{aligned} \mathbf{r} &\leftrightarrow \mathbf{L} && (11)\\ \mathbf{z} &\leftrightarrow \mathbf{L}^{\text{path}} + \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}} && (12)\\ \epsilon &\leftrightarrow n_c && (13)\\ \mathbf{b} &\leftrightarrow \left[\mathbf{L}^{\text{gnd}} - \mathbf{B}(T^{\text{plume}})\right]\boldsymbol{\tau}^{\text{atm}}\boldsymbol{\gamma} && (14)\end{aligned}$$

The first three associations are straightforward, but the signature $\mathbf{b}$ in Eq. (14) is more complicated. We will consider two simplifications: a linear signal model in which $\mathbf{b}$ is assumed to be a known, constant vector; and an absorptive plume model, in which $\mathbf{b}$ depends on the background $\mathbf{z}$.

For the linear signal model, we choose a constant $\mathbf{b}$ which approximates $\left[\mathbf{L}^{\text{gnd}} - \mathbf{B}(T^{\text{plume}})\right]\boldsymbol{\tau}^{\text{atm}}\boldsymbol{\gamma}$ over the scene. A natural choice is the mean: $\mathbf{b} = \left\langle\left[\mathbf{L}^{\text{gnd}} - \mathbf{B}(T^{\text{plume}})\right]\boldsymbol{\tau}^{\text{atm}}\boldsymbol{\gamma}\right\rangle$, although in practice simpler approximations are often used; one common choice is $\mathbf{b} = \boldsymbol{\tau}^{\text{atm}}\boldsymbol{\alpha}$. In the absorptive-only situation, we relate Eq. (10) and Eq. (9), by modifying two associations:

$$\begin{aligned} \mathbf{z} &\leftrightarrow \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}} && (15)\\ \mathbf{b} &\leftrightarrow \mathbf{L}^{\text{gnd}}\boldsymbol{\tau}^{\text{atm}}\boldsymbol{\gamma} && (16)\end{aligned}$$

In the interests of maintaining a formal linear algebraic formulation, we define a "signature matrix" Γ, derived from the chemical absorption α_λ, which is a diagonal matrix whose diagonal elements are given by

$$\Gamma_{\lambda\lambda} = \gamma_\lambda = \frac{1-\exp(-\epsilon\alpha_\lambda)}{\epsilon}. \tag{17}$$

where γ_λ are the components of the vector $\boldsymbol{\gamma}$ defined in Eq. (6). We will also write

$$\Gamma_o = \lim_{\epsilon\to 0}\Gamma; \tag{18}$$

in particular, $\Gamma_{o,\lambda\lambda} = \alpha_\lambda$. Now, we can write $\mathbf{r}$ in matrix-vector notation:

$$\mathbf{r} = \mathbf{z} - \epsilon\Gamma\mathbf{z} \tag{19}$$

Comparison with Eq. (10) identifies $\Gamma\mathbf{z}$ as the plume signature; unlike the linear signal model, however, this is a variable signature that depends on the (unknown) background $\mathbf{z}$.

One naive way to address the variability of the $\Gamma\mathbf{z}$ term is to treat it like the linear signal model, and use $\mathbf{b} = \langle\Gamma\mathbf{z}\rangle = \Gamma\boldsymbol{\mu}$. That leads to the matched filter $\mathbf{q}$ given by $\mathbf{q} = K^{-1}\Gamma\boldsymbol{\mu}$. This for example was done in Ref. 5, though without any formal justification.

[‡]We treat $\mathbf{r}$ as radiance, but at visible wavelengths, much of the formalism could be applied to reflectance as well.

2.1. Linear filter

Having translated the plume detection problem to a linear algebra problem enables us to employ the formalism of matched filters. A linear filter is a vector $\mathbf{q} \in \mathbb{R}^d$ that is applied to a pixel vector $\mathbf{r}$ to produce a scalar $\mathbf{q}^T\mathbf{r}$ whose value indicates the presence (and/or strength) of the plume at that pixel. A good filter will emphasize the signal while suppressing the background; that is,

$$\mathbf{q}^T\mathbf{r} = \underbrace{\mathbf{q}^T\mathbf{z}}_{\text{clutter}} - \underbrace{\epsilon\mathbf{q}^T\mathbf{b}}_{\text{signal}} \tag{20}$$

and the aim is to choose $\mathbf{q}$ so that $\mathbf{q}^T\mathbf{z}$ is small but $\mathbf{q}^T\mathbf{b}$ is large.

2.2. Adaptive matched filter

The adaptive matched filter can be derived by optimizing the signal-to-clutter ratio (SCR). The magnitude of the signal is $|\epsilon\mathbf{q}^T\mathbf{b}|^2$; and the magnitude of the clutter is the variance of $\mathbf{q}^T\mathbf{z}$, which can be expressed in terms of the mean and covariance of the background clutter $\mathbf{z}$. First, compute the average $\boldsymbol{\mu}_o = \langle\mathbf{z}\rangle$, and covariance matrix

$$K_o = \left\langle (\mathbf{z}-\boldsymbol{\mu}_o)(\mathbf{z}-\boldsymbol{\mu}_o)^T \right\rangle. \tag{21}$$

The variance of the matched filter image over the plume-free scene is

$$\begin{aligned} \mathrm{Var}(\mathbf{q}^T\mathbf{z}) &= \left\langle |\mathbf{q}^T(\mathbf{z}-\boldsymbol{\mu}_o)|^2 \right\rangle = \left\langle \mathbf{q}^T(\mathbf{z}-\boldsymbol{\mu}_o)(\mathbf{z}-\boldsymbol{\mu}_o)^T\mathbf{q} \right\rangle = \mathbf{q}^T \left\langle (\mathbf{z}-\boldsymbol{\mu}_o)(\mathbf{z}-\boldsymbol{\mu}_o)^T \right\rangle \mathbf{q} \\ &= \mathbf{q}^T K_o \mathbf{q}, \end{aligned} \tag{22}$$

and the signal-to-clutter is given by

$$\mathrm{SCR} = \frac{|\epsilon\mathbf{q}^T\mathbf{b}|^2}{\mathrm{Var}(\mathbf{q}^T\mathbf{z})} = \frac{\epsilon^2\mathbf{q}^T\mathbf{b}\mathbf{b}^T\mathbf{q}}{\mathbf{q}^T K_o \mathbf{q}}. \tag{23}$$

To optimize the SCR, we will introduce a change of variable:

$$\mathbf{w} = K_o^{1/2}\mathbf{q}. \tag{24}$$

Then $\mathbf{q} = K_o^{-1/2}\mathbf{w}$, and

$$\mathrm{SCR} = \frac{\epsilon^2|\mathbf{w}^T K_o^{-1/2}\mathbf{b}|^2}{\mathbf{w}^T\mathbf{w}}. \tag{25}$$

Since the denominator depends only on the magnitude of $\mathbf{w}$, optimizing the numerator is achieved when $\mathbf{w} \propto K_o^{-1/2}\mathbf{b}$. It follows that

$$\mathbf{q} = K_o^{-1/2}\mathbf{w} \propto K_o^{-1/2}(K_o^{-1/2}\mathbf{b}) = K_o^{-1}\mathbf{b} \tag{26}$$

is the matched filter that optimizes the SCR. In fact, since SCR is invariant to the magnitude of $\mathbf{q}$, any scalar multiple of $K_o^{-1}\mathbf{b}$ is also optimal. Since it is "adaptive" to the covariance K_o of the background, this is called the adaptive matched filter (AMF).[2]

If the background distribution is Gaussian, then the AMF can also be derived from a direct likelihood ratio test; when the plume strength ϵ is not known *a priori*, a generalized likelihood ratio test gives essentially[§] the same filter.[9,10]

Scharf[11] further shows that this linear matched filter is the optimal ("uniformly most powerful") detector when the background is Gaussian. For non-Gaussian distributions (such as elliptically-contoured distributions,[12] also known as spherically-invariant random processes[13–15]), the optimal detector is typically nonlinear[16,17]; however, if restricted to linear detectors, even for non-Gaussian distributions, the best linear detector is often very nearly equal to the adaptive matched filter.[18]

[§]In the original derivation by Kelly,[9] there is a small $O(1/N)$ correction; for hyperspectral data, it has little influence.

3. EFFECT OF PLUME CONTAMINATION: THE LINEAR CASE

If the plume is present in the scene from which the background statistics (mean $\boldsymbol{\mu}$ and covariance K) are computed, then the estimates of those statistics will be biased, and this will affect the performance of the matched filter. In particular, we will have

$$\boldsymbol{\mu} \equiv \langle \mathbf{r} \rangle = \langle \mathbf{z} \rangle - \langle \epsilon \mathbf{b} \rangle = \boldsymbol{\mu}_o - \langle \epsilon \rangle \mathbf{b}, \tag{27}$$

and

$$K \equiv \left\langle (\mathbf{r} - \boldsymbol{\mu})(\mathbf{r} - \boldsymbol{\mu})^T \right\rangle. \tag{28}$$

But from Eq. (10) and Eq. (27),

$$\mathbf{r} - \boldsymbol{\mu} = \mathbf{z} - \epsilon \mathbf{b} - (\boldsymbol{\mu}_o - \langle \epsilon \rangle \mathbf{b}) = (\mathbf{z} - \boldsymbol{\mu}_o) - (\epsilon - \langle \epsilon \rangle)\mathbf{b}, \tag{29}$$

so

$$\begin{aligned} K &= \left\langle (\mathbf{z} - \boldsymbol{\mu}_o - (\epsilon - \langle \epsilon \rangle)\mathbf{b}) \, (\mathbf{z} - \boldsymbol{\mu}_o - (\epsilon - \langle \epsilon \rangle)\mathbf{b})^T \right\rangle & (30) \\ &= \left\langle (\mathbf{z} - \boldsymbol{\mu}_o)(\mathbf{z} - \boldsymbol{\mu}_o)^T \right\rangle \\ &\quad - \left\langle (\epsilon - \langle \epsilon \rangle) \left[\mathbf{b}(\mathbf{z} - \boldsymbol{\mu}_o)^T + (\mathbf{z} - \boldsymbol{\mu}_o)\mathbf{b}^T \right] \right\rangle \\ &\quad + \left\langle (\epsilon - \langle \epsilon \rangle)^2 \right\rangle \mathbf{b}\mathbf{b}^T. & (31) \end{aligned}$$

We define a measure of how plume strength varies over the entire image:

$$\varepsilon^2 \equiv \left\langle (\epsilon - \langle \epsilon \rangle)^2 \right\rangle. \tag{32}$$

Note, for instance, that a plume with uniform strength over the entire scene has $\varepsilon^2 = 0$; even if it is a strong plume, it is virtually invisible. Following Ref. 5, we will also define an expression for the correlation, over the scene, of plume strength and background spectrum.

$$\boldsymbol{\zeta} = \left\langle \left(\frac{\epsilon - \langle \epsilon \rangle}{\varepsilon} \right) (\mathbf{z} - \boldsymbol{\mu}_o) \right\rangle. \tag{33}$$

From the definitions in Eqs. (21,32,33), we can write

$$K = K_o - \varepsilon \mathbf{b}\boldsymbol{\zeta}^T - \varepsilon \boldsymbol{\zeta}\mathbf{b}^T + \varepsilon^2 \mathbf{b}\mathbf{b}^T. \tag{34}$$

In general we expect $\boldsymbol{\zeta}$ to be small since there is no *a priori* reason to expect the plume strength to be correlated with the background spectrum. On the other hand, for typical plumes in typical scenes, we expect some "residual" nonzero correlation. But one obtains qualitatively different behavior for the contaminated covariance, depending on whether or not $\boldsymbol{\zeta}$ is strictly zero.

3.1. Remark on the $\boldsymbol{\zeta} = 0$ case

When $\boldsymbol{\zeta} = \mathbf{0}$, we have $K = K_o + \varepsilon^2 \mathbf{b}\mathbf{b}^T$, so the covariance *is* contaminated. But this contamination has absolutely no effect on the matched filter, even for large ε. That is because the matched filter vectors $\mathbf{q}_o = K_o^{-1}\mathbf{b}$ and $\mathbf{q} = K^{-1}\mathbf{b} = (K_o + \varepsilon^2 \mathbf{b}\mathbf{b}^T)^{-1}\mathbf{b}$ are parallel. This can be seen by noting that $K_o \mathbf{q}_o = K\mathbf{q} = \mathbf{b}$ and that

$$K_o \mathbf{q} = (K - \varepsilon^2 \mathbf{b}\mathbf{b}^T)\mathbf{q} = K\mathbf{q} - \varepsilon^2 \mathbf{b}(\mathbf{b}^T\mathbf{q}) = \left[1 - \varepsilon^2 (\mathbf{b}^T\mathbf{q})\right] \mathbf{b} = \theta \mathbf{b}. \tag{35}$$

where $\theta = 1 - \varepsilon^2(\mathbf{b}^T\mathbf{q})$ is a scalar. Multiply both sides by K_o^{-1}, and we get $\mathbf{q} = K_o^{-1}(\theta \mathbf{b}) = \theta(K_o^{-1}\mathbf{b}) = \theta \mathbf{q}_o$, and so $\mathbf{q}$ and $\mathbf{q}_o$ are parallel. It follows that the signal-to-clutter for the contaminated and uncontaminated matched filters will be the same.¶

¶There is occasional confusion over the fact that the scalar θ in the above equations can alter the signal level by a considerable amount, even when ε is small; but that scalar alters the clutter by the same amount, so there is no effect on SCR or on ROC-curve performance.

3.2. Nonzero $\boldsymbol{\zeta}$ in the linear regime

For nonzero $\boldsymbol{\zeta}$, the effect of covariance contamination is to degrade the performance of the matched filter. Since $\boldsymbol{\zeta}$ is generally small, the degradation is fairly minor. In particular, we showed in Ref. 5 that

$$\mathrm{SCR} = \frac{\epsilon^2 \mathbf{b} K_o^{-1} \mathbf{b}}{1 + \varepsilon^2 \left[(\mathbf{b} K_o^{-1} \mathbf{b})(\boldsymbol{\zeta}^T K_o^{-1} \boldsymbol{\zeta}) - (\mathbf{b} K_o^{-1} \boldsymbol{\zeta})^2 \right]}. \tag{36}$$

That is, SCR increases monotonically with plume strength, eventually saturating at a value that is inversely proportional to the magnitude of $\boldsymbol{\zeta}$. Note also, that the manifestation of nonzero $\boldsymbol{\zeta}$ is in terms of the dimensionless quantities $\boldsymbol{\zeta}^T K_o^{-1} \boldsymbol{\zeta}$ and $(\mathbf{b} K_o^{-1} \boldsymbol{\zeta})^2 / \left[(\mathbf{b} K_o^{-1} \mathbf{b})(\boldsymbol{\zeta}^T K_o^{-1} \boldsymbol{\zeta}) \right]$.

4. SIMULATED PLUME

The experiments that follow are based on simulated plumes imposed on real data. We start with 224-channel hyperspectral data from the AVIRIS sensor[19,20] (see Fig. 2(a)). This is what we use for our background $\mathbf{z}$ (equivalently: $\mathbf{L}^{\mathrm{gnd}} \boldsymbol{\tau}^{\mathrm{atm}}$). To generate a hysperspectral dataset with a plume, we first generate a plume profile, shown in in Fig. 2(b), which is maximally concentrated at a point toward the upper right of the image, and then diffuses away from that concentration while at the same time being blown toward the bottom of the image.[‖] Then, using the gas absorption spectrum[21] for NO_2 (shown in Fig. 1) for $\boldsymbol{\alpha}$, we make two different plume datasets:

$$\text{Beer's Law:} \quad r_\lambda = z_\lambda \exp(-\epsilon \alpha_\lambda) \qquad \text{Beer's Law vector:} \quad \mathbf{r} = \mathbf{z} - \epsilon \Gamma \mathbf{z} \tag{37}$$

$$\text{Linearized:} \quad r_\lambda = z_\lambda - \epsilon \alpha_\lambda \langle z_\lambda \rangle \qquad \text{Linearized vector:} \quad \mathbf{r} = \mathbf{z} - \epsilon \mathbf{b} \tag{38}$$

with $\mathbf{b} = \Gamma_o \boldsymbol{\mu}_o$. We do both nonlinear and linear plume simulations because we want to distinguish which aspects of the matched filter performance are due to nonlinearity, and which (by contrast) are also observed in the linear regime.

We also consider a matched-pair plume simulation (see Fig. 3(b)); here the scene is duplicated, and one duplicate has the original data and the other has plume imposed at a uniform plume concentration. Although this is less realistic looking than the plume shown in Fig. 2(b), it provides for a more controlled experiment. An important feature of this construction is that the correlation defined in Eq. (33) is strictly zero: $\boldsymbol{\zeta} = \mathbf{0}$. Another important (and arguably unrealistic) feature is that the area of the plume in this simulation is half the area of the whole scene.

Table 1 shows various parameters that characterize the two plume simulations.

5. EXPERIMENTS

The first and most straightforward experiment is to observe and measure the signal-to-clutter ratio (SCR) for various matched filters as a function of plume concentration. In particular, we want to compare the matched filter $\mathbf{q}_o = K_o^{-1} \mathbf{b}$ obtained from the uncontaminated covariance matrix K_o, against the matched filter $\mathbf{q} = K^{-1} \mathbf{b}$ obtained from the contaminated K.

The way we estimate signal-to-clutter in these experiments is with the mean value of the matched-filter signal over the on-plume pixel, that value squared, divided by the variance of the off-plume pixels.

$$\mathrm{SCR} = \frac{\left\langle \mathbf{q}^T (\mathbf{r} - \boldsymbol{\mu}_o) \right\rangle^2_{\text{on-plume}}}{\mathrm{Var}(\mathbf{q}^T (\mathbf{r} - \boldsymbol{\mu}_o))_{\text{off-plume}}} \tag{39}$$

For most of these experiments, we plot this estimate of SCR on the y-axis. On the x-axis, we plot the plume strength ϵ associated with the strongest plume pixel in the plume. Each point on the curve corresponds to a

[‖] When we refer to the overall plume strength as a single value, ϵ, we are referring to the strongest part of the plume in the image.

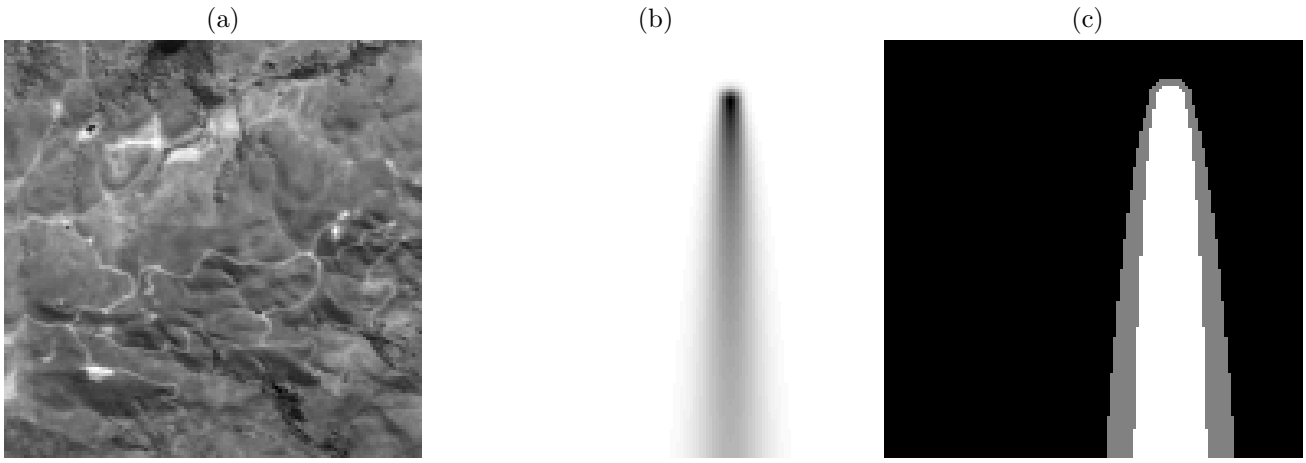

Figure 2. (a) 128×128 chip of an AVIRIS scene – this is a broadband image obtained by summing all 224 of the spectral channels. **(b)** Profile of the plume that is artificially imposed on the hyperspectral data in panel (a). **(c)** The plume mask distinguishes on-plume (white) from off-plume (black) pixels. The gray area specifies pixels which are not categorized as either on-plume or off-plume; these are pixels on the outer fringe of the plume, and corresponds to pixels where the plume column density is less then ten percent of the maximum column density in the plume.

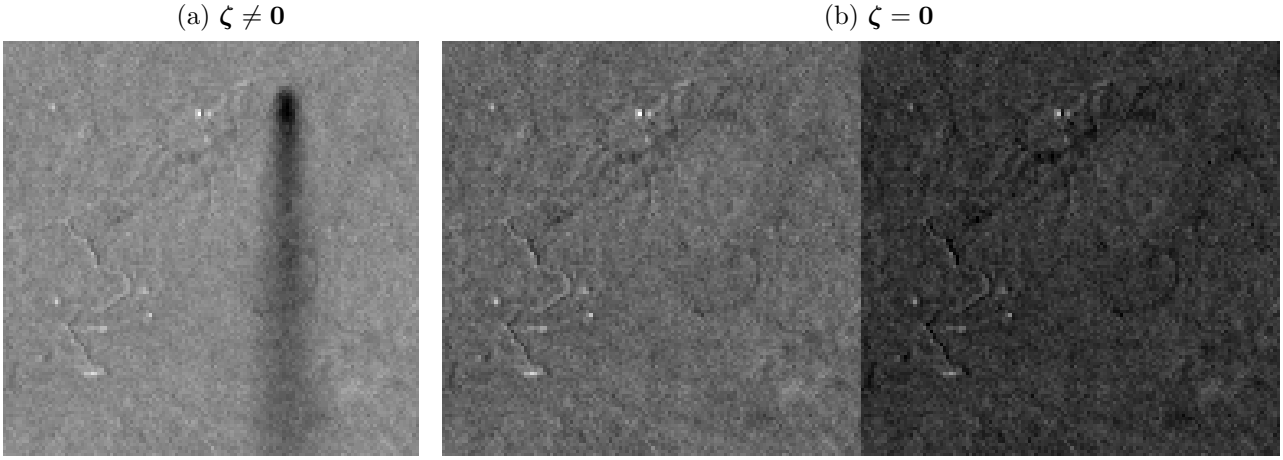

Figure 3. (a) Adaptive matched filter image of the scene shows the plume as darker pixels. The plume strength here is 35 ppm-m at maximum; its average over the on-plume pixels is approximately 10 ppm-m. **(b)** Adaptive matched filter image of a matched-pair plume; the artificial plume has a uniform concentration 10 ppm-m, and extends over the right half of the image. This right half is a pixel-by-pixel copy of the left half of the image, but with the effect of the plume imposed.

Table 1. Properties of the two simulated plumes. Measures of plume strength are relative to the strength ϵ of the plume at its strongest point. Here, $\langle\epsilon\rangle_{\text{plume}}$ is the average strength of the plume over the on-plume pixels, while $\langle\epsilon\rangle_{\text{image}}$ is the average strength over the whole image. The plume variance ε is defined in Eq. (32). The plume correlation $\boldsymbol{\zeta}$ is defined in Eq. (33), but the values shown in the table are expressions that include scaling by the covariance matrix to produce dimensionless values.

	Measures of plume strengh				Measures of plume correlation	
	ϵ	$\langle\epsilon\rangle_{\text{plume}}$	$\langle\epsilon\rangle_{\text{image}}$	ε	$\boldsymbol{\zeta}^T K^{-1}\boldsymbol{\zeta}$	$\dfrac{\boldsymbol{\zeta}^T K^{-1}\mathbf{b}}{\sqrt{(\boldsymbol{\zeta}^T K^{-1}\boldsymbol{\zeta})(\mathbf{b}^T K^{-1}\mathbf{b})}}$
Shaped Plume, Fig. 3(a)	1	0.2857	0.0384	0.1052	0.1198	0.0488
Matched-pair, Fig. 3(b)	1	1.0000	0.5000	0.5000	0.0000	——

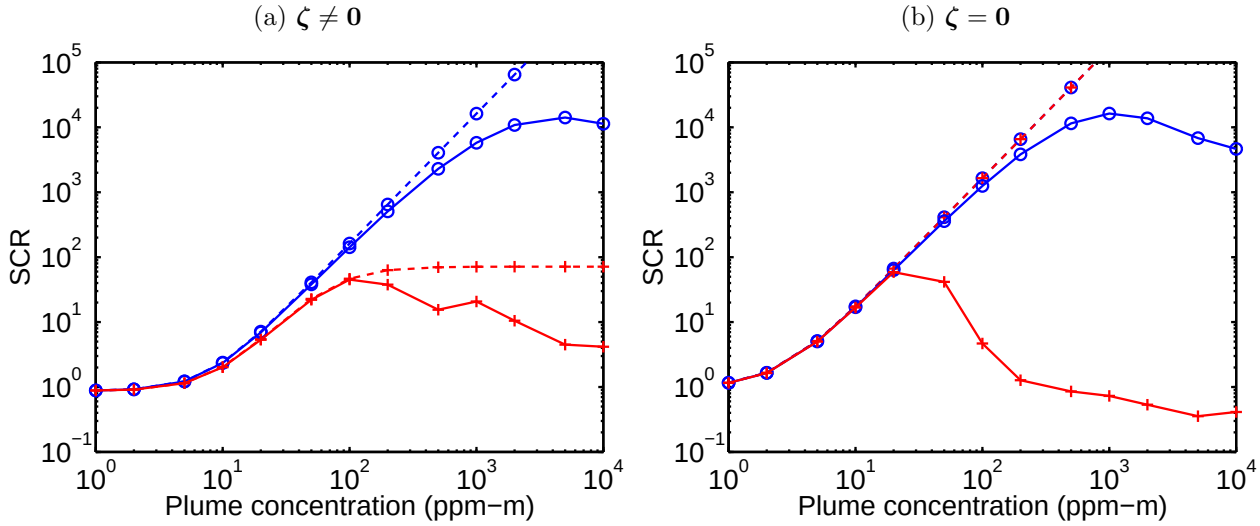

Figure 4. SCR as a function of plume concentration, for the plumes shown in Fig. 3. Dashed line is for the linear plume simulation; solid line is for Beer's Law nonlinearity. Circles (o) indicate uncontaminated covariance matrix; pluses (+) are for contaminated covariance matrix. In both cases the same signature $\mathbf{b} = \langle\mathbf{z}\rangle\,\boldsymbol{\alpha}$ was used, with $\langle\mathbf{z}\rangle$ in both cases uncontaminated. Note that even for the uncontaminated nonlinear plume (solid with circles), the SCR is nonmonotonic – a stronger plume is not necessarily more visible. **(b)** In the matched-pair case, with $\boldsymbol{\zeta} = \mathbf{0}$, we find that the linear plume is utterly unaffected by contamination.

different experiment, in which a plume of a given strength is simulated over the scene, the matched-filter is applied, and the SCR is estimated.

In Fig. 4 we compare the performance of these two matched filters on both linear and nonlinear Beer's Law plumes. Not surprisingly, the linear and nonlinear behavior is nearly identical for low plume strengths, but as the plume gets stronger, we find that the linear plumes are more readily detected than the more realistic nonlinear plumes. Also (and again, not surprisingly), the uncontaminated matched filter generally outperforms the contaminated filter. The one exception is the linear matched-pair plume; here, we have $\boldsymbol{\zeta} = \mathbf{0}$ and, as noted in Section 3.1, the contaminated and uncontaminated matched filters are parallel and have identical performance. For the more realistically shaped ($\boldsymbol{\zeta} \neq \mathbf{0}$) but still linearized plume, we see that the SCR does saturate with increasing plume strength, but (unlike the nonlinear plumes) it never decreases with increasing plume strength. A notable feature of the nonlinear plumes is that as the plume strength goes beyond a certain level, the SCR actually starts decreasing with increasing plume strength. The plume is so strong that it starts to become invisible again!

Fig. 5 illustrates the effect of a slightly more realistic contamination model. Here the contamination of the

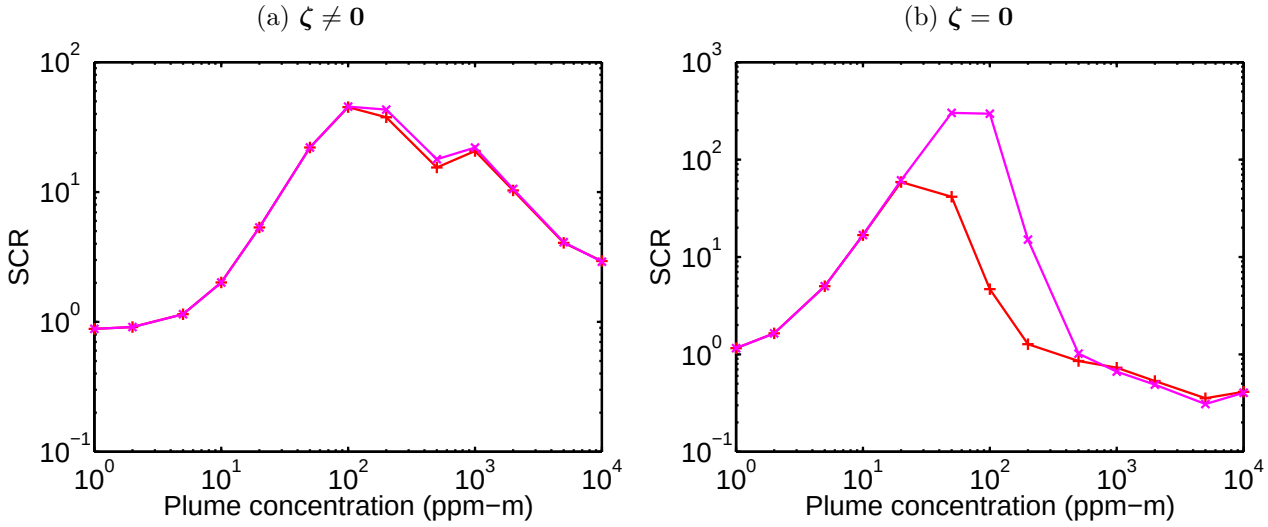

Figure 5. SCR as a function of plume concentration, for the plume shown in Fig. 2. Solid plus (+) corresponds to the curves in Fig. 4, in which contamination is applied only the the covariance K. The solid cross (×) corresponds to contamination on both the covariance K and the signature $\mathbf{b} = \langle \mathbf{z} \rangle \, \boldsymbol{\alpha}$.

plume affects not only the estimate of covariance, but it affects the estimate of the plume signature as well. These plots show that the effect of the two kinds of contamination is generally quite similar. The effect is much larger in the matched-pair plume, but in that case, the plume is such a large part of the scene, that the contamination really is substantial.

In Fig. 6, we show what the matched-filter images look like for both the uncontaminated and the contaminated matched filters. The uncontaminated matched filter has high SCR and produces an image in which the plume is strongly detected. However, even for the uncontaminated filter, the effect of nonlinearity is evident in that the central (and strongest part) of the plume is actually less dark than the plume toward the outer edges. This subtle effect is hugely magnified for the contaminated filter image in Fig. 6(b); here the outer wispy edges of the plume are well-detected, but the central part is not.

Using $\boldsymbol{\gamma} = [1 - \exp(-\epsilon \boldsymbol{\alpha})]/\epsilon$ in place of $\boldsymbol{\alpha}$, the effect of nonlinearity can be ameliorated. The effect of this nonlinearity, for the completely uncontaminated case, is seen in Fig. 7. Here, SCR as a function of plume strength is plotted for the standard matched filter $\mathbf{q}_o = K_o^{-1} \mathbf{b}_o$, where $\mathbf{b}_o = \langle \mathbf{z} \rangle \, \boldsymbol{\alpha}$, and for a modified matched filter, which still uses the uncontaminated covariance, but for which the signature is adapted to the nonlinearity: $\mathbf{q} = K_o^{-1} \mathbf{b}$, where $\mathbf{b} = \langle \mathbf{z} \rangle \, \boldsymbol{\gamma}$. The modified plume signature gives better performance, but the effect is small. However, when using the contaminated covariance, also seen in Fig. 7, the benefits of using the modified signature are far more substantial, and they are evident at much weaker plume strengths.

Fig. 7 illustrates that better detection can be achieved if the signature is adapted to the expected strength of the plume. Here, using $\boldsymbol{\gamma}$, as defined in Eq. (6) in place of $\boldsymbol{\alpha}$, one can alleviate some of the effects of Beer's Law nonlinearity. Making this change requires knowing (or estimating) the plume strength beforehand.

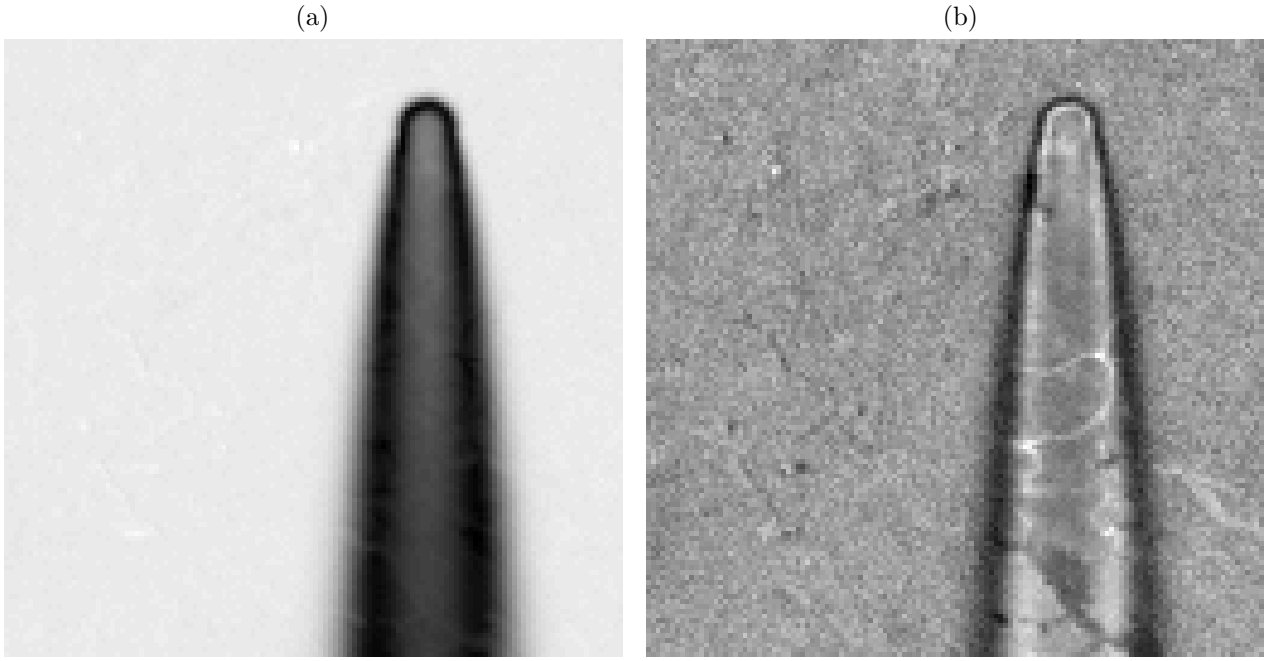

Figure 6. Effect of covariance contamination in the case of extreme plume concentration (10^4 ppm-m). Shown are matched-filter images, for **(a)** uncontaminated covariance matrix and signature, and **(b)** contaminated covariance matrix and signature. In both plumes, the strongest parts of the plume show supressed detection, compared to weaker parts of the plume – this is a subtle effect for the uncontaminated matched filter in (a), and a dramatic effect for the contaminated matched filter in (b); in the latter case, it is only the wispy envelope of the plume that is strongly detected.

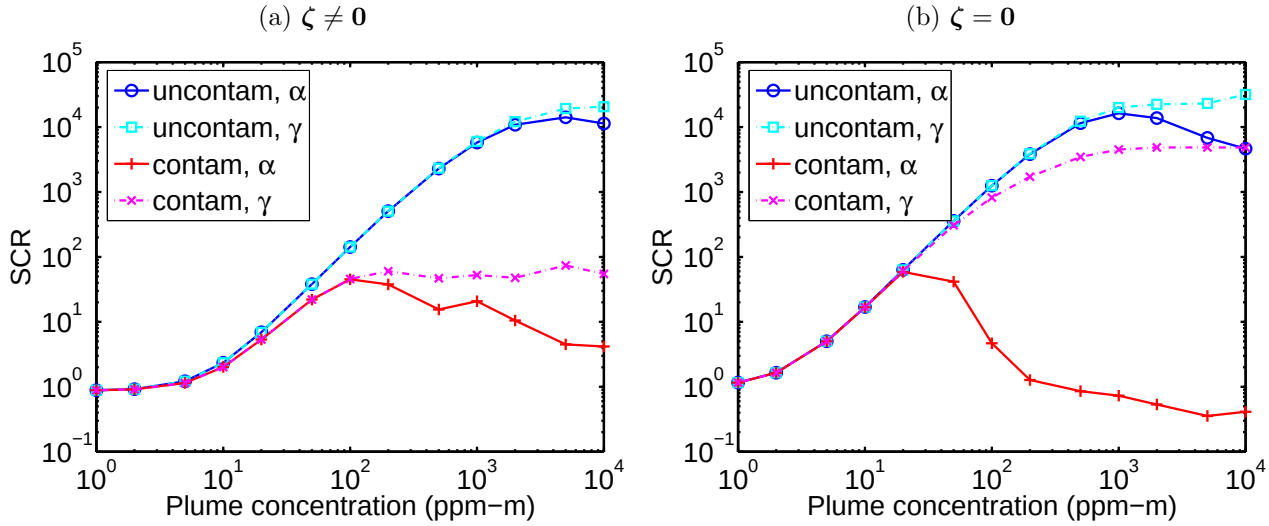

Figure 7. SCR as a function of plume concentration, comparing the use of $\boldsymbol{\alpha}$ (solid lines) with $\boldsymbol{\gamma} = [1 - \exp(-\epsilon\boldsymbol{\alpha})]/\epsilon$ (dash-dotted lines), using the average plume strength for ϵ. The top curves are based on the matched filter from the uncontaminated covariance; for these curves, we see that $\boldsymbol{\gamma}$ (squares (□)) gives better performance than $\boldsymbol{\alpha}$ (circles (o)), but the difference is small, even for very strong plumes. The bottom curves are based on the contaminated covariance; here, again, $\boldsymbol{\gamma}$ (crosses (×)) outperforms $\boldsymbol{\alpha}$ (pluses (+)); but in this case, the difference is far more substantial, and is evident at much lower plume strengths.

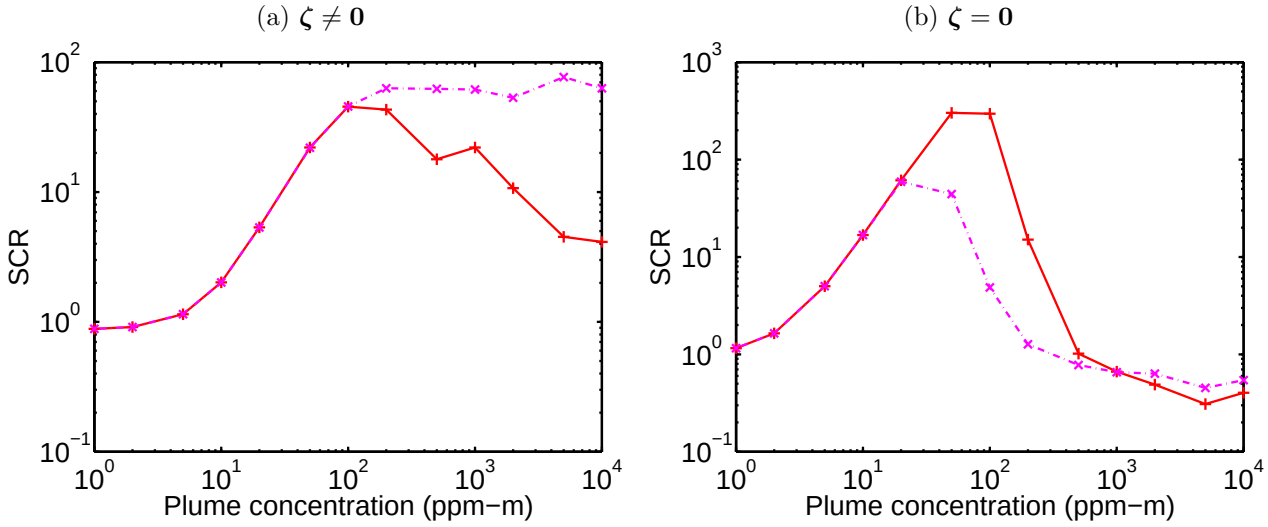

Figure 8. Same as Fig. 7(c,d), but with contamination in both the covariance K and in the estimated mean spectrum $\langle \mathbf{z} \rangle$. SCR is plotted as a function of plume concentration, comparing the use of $\boldsymbol{\alpha}$ (solid pluses (+)) with $\boldsymbol{\gamma} = [1 - \exp(-\epsilon\boldsymbol{\alpha})]/\epsilon$ (dash-dotted crosses (×)). **(a)** For the smaller plume, the nonlinear correction $\boldsymbol{\gamma}$ again provided a substantial advantage. **(b)** For the matched-pair simulation, where the contamination is much greater, we see an anomalous result; the nonlinear correction $\boldsymbol{\gamma}$ does not improve the detection.

6. CONCLUSIONS

Although the magnitude of the matched filter can exhibit extreme sensitivity due to even weak contamination of the covariance matrix with the signal of interest, the actual effect on detectability, at least for weak plumes, is usually small. In particular, the linear analysis[5] indicates that the effect depends only on the residual correlation $\boldsymbol{\zeta}$ between the plume strength and the background spectrum. The effect of a nonzero $\boldsymbol{\zeta}$ is to supress the signal-to-clutter ratio; however, as the signal gets stronger (and as a consequence, the contamination also gets stronger), the resultant SCR still increases. If $\boldsymbol{\zeta}$ strictly vanishes, then the contamination has no effect whatever.

Here, we have explored the case of strong plumes, where the effect of the plume is not a simple linear addition but, because of Beer's Law, has a nonlinear effect on the background. In this regime, the effect of the contamination is compounded with an effect due to signal mismatch, with the result that nonmonotonic variation of SCR with signal strength can be observed. That is, as the signal becomes stronger, it becomes less visible.

This reversal in visibility as a function of plume strength can be understood in terms of deviations of $\boldsymbol{\gamma}$ from $\boldsymbol{\alpha}$ as the plume strength increases. The effect is fairly subtle for the uncontaminated matched filter (shown qualitatively in Fig. 6(a), and quantitatively in the upper curves of Fig. 7). But the effect appears greatly magnified for the contaminated matched filter. Here, the nonmonotonicity of SCR with plume strength is both a larger effect, and one that is evident at weaker plume strengths, than is seen for the uncontaminated matched filter. Fig. 6(b) shows an extemely concentrated plume (10^4 ppm-m) that is nearly invisible; and the lower curves of Fig. 7 show a marked nonmonotonicity.

In high-dimensional spaces, the adaptive matched filter is a very sensitive instrument, capable of detecting trace quantities of gaseous chemicals in large scenes. That sensitivity, however, can be a weakness, when it is applied to very strong plumes. In one case from the literature,[22] the matched filter failed to detect a plume that was so strong it dominated the third principal component. We speculate that this was cased by the contamination of the covariance by the existence of the plume in the scene, combined with the nonlinearity of Beer's Law absorption.

Acknowledgements

We thank Brad Henderson and Herb Fry for useful conversations on this topic. This project was funded by the Laboratory Directed Research and Development (LDRD) program at Los Alamos.

REFERENCES

1. D. Manolakis and G.Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Magazine* **19**(1), pp. 29–43, 2002.
2. I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerospace and Electronic Systems* **10**, pp. 853–863, 1974.
3. C. G. Khatri and C. R. Rao, "Effects of estimated noise covariance matrix in optimal signal detection," *IEEE Trans. Acoustics, Speech, and Signal Processing* **35**, pp. 671–679, 1987.
4. K. Gerlach, "The effects of signal contamination on two adaptive detectors," *IEEE Trans. Aerospace and Electronic Systems* **31**, pp. 297–309, 1995.
5. J. Theiler and B. R. Foy, "Effect of signal contamination in matched-filter detection of the signal on a cluttered background," *IEEE Geoscience and Remote Sensing Letters* **3**, pp. 98–102, 2006.
6. 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.
7. 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.
8. A. Beer, "Bestimmung der absorption des rothen lichts in farbigen flussigketiten," *Ann. Physik* **86**, pp. 78–88, 1852.
9. E. J. Kelly, "An adaptive detection algorithm," *IEEE Trans. Aerospace and Electronic Systems* **22**, pp. 115–127, 1986.
10. 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.
11. L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Processing* **42**, pp. 2146–2156, 1994.
12. D. Manolakis, D. Marden, J. Kerekes, and G. Shaw, "On the statistics of hyperspectral imaging data," *Proc. SPIE* **4381**, pp. 308–316, 2001.
13. M. Rangaswamy, P. Chakravarthi, D. Weiner, L. Cai, H. Wang, and A. Ozturk, "Signal detection in correlated and non-Gaussian radar clutter," Tech. Rep. RL-TR-93-79, Rome Laboratory, 1993.
14. M. Rangaswamy, D. Weiner, and A. Ozturk, "Non-Gaussian random vector identification using spherically invariant random processes," *IEEE Trans. Aerospance and Electronic Systems* **29**, pp. 111–124, 1993.
15. M. Rangaswamy, "Spherically invariant random processes for modeling non-Gaussian radar clutter," in *Proc. 27th Asilomar Conference on Signals, Systems, and Computers, vol 2*, pp. 1106–1110, 1993.
16. 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.
17. A. Schaum and E. Allman, "Detection in hyperspectral imagery based on non-Gaussian elliptically contoured distributions," *Proc. MSS (Military Sensing Symposium) on Camouflage, Concealment, and Deception* , 2006. To appear.
18. B. R. Foy, J. Theiler, and A. M. Fraser, "Unreasonable effectiveness of the adaptive matched filter," *Proc. MSS (Military Sensing Symposia) Passive Sensors Conference* , 2006. To appear.
19. 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.
20. AVIRIS Free Standard Data Products, Jet Propulsion Laboratory (JPL), National Aeronautics and Space Administration (NASA). `http://aviris.jpl.nasa.gov/html/aviris.freedata.html`.
21. A. C. Vandaele, C. Hermans, P. C. Simon, M. Carleer, R. Colin, S. Fally, M. F. Merienne, A. Jenouvrier, and B. Coquart, "Measurements of the NO_2 absorption cross-section from 42000 cm^{-1} to 10000 cm^{-1} (238-1000 nm) at 220 K and 294 K," *Journal of Quantitative Spectroscopy and Radiative Transfer* **59**, pp. 171–184, 1998.
22. D. W. Messinger, "Gaseous plume detection in hyperspectral images: a comparison of methods," *Proc. SPIE* **5425**, pp. 592–603, 2004.