Absorptive Weak Plume Detection on Gaussian and Non-Gaussian Background Clutter

Theiler, James Patrick

To be published in: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

DOI to publisher's version: 10.1109/JSTARS.2021.3093820

Permalink to record: http://permalink.lanl.gov/object/view?what=info:lanl-repo/lareport/LA-UR-21-24152

Absorptive Weak Plume Detection on Gaussian and Non-Gaussian Background Clutter

James Theiler

***Abstract*—For additive signals on Gaussian clutter, the optimal detector is a linear matched filter that is adapted to the known signal and the covariance of the background. This adaptive matched filter is widely used for gas-phase plume detection, even though the effect of the plume on the background is not strictly additive. Here, a derivation of the matched filter for a strictly absorptive plume produces, in the weak plume limit, a quadratic filter. This quadratic matched filter is extended in two ways: an elliptically-contoured multivariate t distribution is used to generalize the Gaussian background clutter, and a generalized likelihood ratio test detector is derived to extend applicability to stronger plumes. In addition to detectors whose purpose is to identify presence versus absence of a plume, expressions are also derived for estimating plume strength. The performance of these various detectors is evaluated by implanting simulated plume into background images that are either real hyperspectral images or simulated images based on different (Gaussian, multivariate t, and lognormal) clutter distributions.**

***Index Terms*—Adaptive signal detection, clutter, composite hypothesis testing, elliptically-contoured distribution, gas-phase plume detection, generalized likelihood ratio test, hyperspectral imagery, lognormal distribution, matched filter.**

I. Introduction

ONE of the triumphs of modern hyperspectral imaging is the remote detection of low-concentration gas-phase plumes on cluttered backgrounds [1]–[10]. Such plumes are often very nearly invisible, and the difference between an ON-plume pixel spectrum and an OFF-plume spectrum may be very small. But the gas absorption spectrum is known to high precision, and because the observations are made at multiple (often hundreds of) wavelengths, it is possible to detect the presence of these weak plumes even in cluttered backgrounds.

Plumes can be both absorptive and emissive, and although emissive spectra in the thermal infrared can be extremely discriminating, detection at visible wavelengths is much cheaper, and requires less specialized hardware. The focus here is on detection of plumes that are purely absorptive.

A modification of the additive target model is used to describe the physics of absorptive plume detection in the visible (and near-ultraviolet) wavelengths of hyperspectral imagery. The physics in this regime is simple—exponential Beer's Law absorption [11]—but it is not linear. Conventional lore holds that for a sufficiently weak plume, the exponential is approximately linear, and so a linear matched filter is appropriate. But even for this simple model, the target is not truly additive even in the weak plume limit.

This problem also provides an exercise in composite hypothesis testing. For such problems, a direct likelihood ratio test cannot be employed, because there is a nuisance parameter (in this case, the plume concentration) whose value is not *a priori* known. The usual approach in this situation is to employ the generalized likelihood ratio test (GLRT), but in the weak plume limit the locally most powerful (LMP) test seems to make sense. It bears remarking that more general Bayesian methods [12] (of which LMP is a special case) or Clairvoyant Fusion [13]–[15] (for which GLRT is a special case) can also be considered.

In the IGARSS conference paper [16] that this article follows, the background was assumed to be Gaussian, and closed-form solutions were derived for both the LMP and GLRT detectors. Here, the background is extended to the multivariate t-distributed background (of which the Gaussian is a special case), and new closed-form expressions are derived. A lognormal background model is also considered, based on the observation that treating the data in log space reduces the exponential attenuation to an additive target model. Finally, numerical experiments are carried out in order to evaluate the relative performance of these various detectors in different background scenarios.

A. Organization and Notation

Section II derives the basic equations for the absorptive plume, including the "classic" matched filter and a clairvoyant detector that is theoretically optimal when the background distribution is specified. Section III begins with the assumption that the background is Gaussian, and derives both a quadratic matched filter based on the LMP and a GLRT detector, the latter of which requires a maximum likelihood estimate of the plume strength, which is of interest in its own right (though it is not the focus of this article). In Section IV, the Gaussian results are generalized to the case of a multivariate t-distributed background, and in Section V a lognormal background is considered (and this is generalized to a log-multivariate t-distributed background). Implanted plumes on real and simulated hyperspectral backgrounds are employed in Section VI to evaluate the various detectors that have been derived in this article. Section VII concludes with a summary of the main points of this article. Some further remarks

Manuscript received April 28, 2021; revised June 7, 2021; accepted June 27, 2021. Date of publication June 30, 2021; date of current version July 20, 2021. This work was supported by a NASA-sponsored project on "Compact high-resolution trace-gas hyperspectral imagers, with agile on-board processing, for trace gas monitoring" under Grant N7-INVEST17-0010, led by S. P. Love at Los Alamos National Laboratory.

The author is with the Space Data Science, and Systems Group, Los Alamos National Laboratory, Los Alamos, NM 87545 USA (e-mail: jt@lanl.gov).

Digital Object Identifier 10.1109/JSTARS.2021.3093820

TABLE I
BASIC NOTATION USED IN THIS EXPOSITION

Name	Meaning
d	Number of spectral channels
λ	wavelength index
$\mathbf{z}$	Pixel spectrum in the absence of plume
$\mathbf{x}$	Measured pixel spectrum
$\mathbf{t}$	Gas absorption spectrum
T	Matrix form of gas absorption: $\mathrm{diag}(\mathbf{t})$
τ	Scalar measure of absorption strength: $\mathrm{Trace}(T)$
a	Plume strength
$\boldsymbol{\mu}$	Mean of pixel spectra
R	Covariance matrix of pixel spectra
ν	Parameter used in multivariate t-distribution

TABLE II
DEFINITIONS OF SOME USEFUL INTERMEDIATE QUANTITIES

Expression	Definition	Equation in text
$\mathcal{Q}(\mathbf{x})$	$-(T\mathbf{x})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})$	Eq. (8)
$\mathcal{A}(\mathbf{x})$	$(\mathbf{x}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})$	Eq. (14)
$\mathcal{E}(\mathbf{x})$	$(T\mathbf{x})'R^{-1}(T\mathbf{x})+(T^2\mathbf{x})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})$	Eq. (16)
$\mathcal{F}_\nu(\mathbf{x})$	$\sqrt{(\nu-1)/[\nu-2)+\mathcal{A}(\mathbf{x})]}$	Eq. (24)
Θ_ν	$(\nu-1)\tau/(\nu+d)$	Eq. (32)
a_o	$1/\|R^{-1/2}T\boldsymbol{\mu}\|$	Eq. (46)

on "characteristic" plume strength, and on why the LMP-based detectors fare so poorly, are included in the Appendix.

Table I provides a brief summary of the main symbols used in the text. Because the search for closed-form solutions often requires unwieldy algebra, various intermediate quantities were introduced; they are defined, as they arise, in the text, and are summarized in Table II.

II. ABSORPTIVE PLUME

For an absorptive plume, we have from Beer's Law that the radiance observed at some wavelength λ is given by $x_\lambda = z_\lambda \exp(-at_\lambda)$, where z_λ is the radiance that would observed in the absence of plume, t_λ is the absorption coefficient of the plume gas, and a is the plume strength. For a sensor with d wavelengths, we can express this in vector form, with d-dimensional vectors $\mathbf{x}$, $\mathbf{z}$, and $\mathbf{t}$, whose components correspond, respectively, to x_λ, z_λ, and t_λ:

$$\mathbf{x} = \exp(-aT)\mathbf{z} \tag{1}$$

where $T = \mathrm{diag}(\mathbf{t})$ is a diagonal matrix whose diagonal elements $T_{\lambda\lambda}$ are the absorption coefficients t_λ.

A. Linear Adaptive Matched Filter (AMF)

The classic AMF [17]–[19] was originally applied to radar signal detection, but is widely used for plume detection. Its derivation requires an additive model; i.e., a model of the form $\mathbf{x} = \mathbf{z} + a\mathbf{t}^*$ for some *constant* signature $\mathbf{t}^*$. For this purely additive model, assuming a Gaussian distribution for the OFF-plume clutter $\mathbf{z}$, the optimal (indeed, uniformly most powerful [12]) detector is given by $\mathcal{D}_{\mathrm{AMF}}(\mathbf{x}) = \mathbf{t}^{*\prime}R^{-1}(\mathbf{x}-\boldsymbol{\mu})$, where $\mathbf{x}$ is the measured spectrum at a pixel (which may or may not have plume present), and the prime symbol ($'$) indicates transpose.

To first order in a (i.e., assuming a weak plume), (1) looks like $\mathbf{x} \approx \mathbf{z} - aT\mathbf{z}$, which suggests $\mathbf{t}^* = -T\mathbf{z} \approx -T\mathbf{x}$, but this expression for t^* is not constant, so the additive matched filter is not applicable. A naive yet popular approximation (e.g., see [5]–[7]) takes $\mathbf{t}^* \propto \mathbf{t}$; this is based on the argument that the background varies much more slowly with wavelength than the gas spectrum does, and so the approximation is that the background is spectrally flat. A better approximation takes $\mathbf{t}^* = -T\boldsymbol{\mu}$, where $\boldsymbol{\mu} = \langle\, \mathbf{x} \,\rangle$ is the mean background over the image. In this absorptive plume context, then, the AMF becomes

$$\mathcal{D}_{T\mu\text{-AMF}}(\mathbf{x}) = -(T\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu}). \tag{2}$$

B. Likelihood Ratio and Clairvoyant Detector

Because (1) is *not* truly additive, the aim in this section is to derive a detector that accounts for the multiplicative nature of plume absorption. We begin by writing $P_{\mathrm{bkg}}(\mathbf{z})$ as the OFF-plume background distribution and $P_{\mathrm{plume}}(\mathbf{x})$ as the distribution of pixel values when plume is present. We can express the relationship between these two using the usual formula for change-of-variables in probability distributions

$$P_{\mathrm{plume}}(\mathbf{x}) = P_{\mathrm{bkg}}(\mathbf{z})\left|\frac{d\mathbf{z}}{d\mathbf{x}}\right| = P_{\mathrm{bkg}}(\exp(aT)\mathbf{x})\,|\exp(aT)| \tag{3}$$

where $|\cdot|$ indicates the determinant. Note that the determinant of the exponential can be expressed as the exponential of the trace (since T is diagonal, this is readily verified)

$$\begin{aligned}|\exp(aT)| &= \prod_\lambda \exp(at_\lambda) = \exp\left(a\sum_\lambda t_\lambda\right)\\ &= \exp(a\,\mathrm{Trace}(T)) = e^{a\tau}\end{aligned} \tag{4}$$

where $\tau = \mathrm{Trace}(T)$. Thus, the likelihood ratio for a pixel $\mathbf{x}$ is given by

$$\mathcal{L}(a,\mathbf{x}) = \frac{P_{\mathrm{plume}}(\mathbf{x})}{P_{\mathrm{bkg}}(\mathbf{x})} = \frac{P_{\mathrm{bkg}}(\exp(aT)\mathbf{x})\,\exp(a\tau)}{P_{\mathrm{bkg}}(\mathbf{x})}. \tag{5}$$

If we somehow *did* know the plume strength a (while at the same time *not* knowing whether or not there even was a plume), then this would be the optimal detector of that plume. It is known as the *clairvoyant* detector [20] and is useful both conceptually (as a starting point for detectors that do not require knowledge of plume strength) and experimentally (as a best-case baseline against which to compare other detectors).

III. GAUSSIAN BACKGROUND

For a Gaussian background, we have $P_{\mathrm{bkg}}(\mathbf{x}) \propto \exp[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})]$; incorporating this expression in (5), and taking logarithms, we have

$$\begin{aligned}\log\mathcal{L}(a,\mathbf{x}) &= \log P_{\mathrm{bkg}}(\exp(aT)\mathbf{x}) + a\tau - \log P_{\mathrm{bkg}}(\mathbf{x})\\ &= -\tfrac{1}{2}(\exp(aT)\mathbf{x}-\boldsymbol{\mu})'R^{-1}(\exp(aT)\mathbf{x}-\boldsymbol{\mu})\\ &\quad + a\tau\\ &\quad + \tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})\end{aligned} \tag{6}$$

which provides a clairvoyant detector of absorptive plume on a Gaussian background. And although our interest in this exposition is with weak plumes, note that (6) is optimal for all positive values of a.

A. Quadratic Matched Filter (QMF)

Because we care about *weak* plumes, we will derive an LMP detector [20] that is optimal for $a \to 0$. Note that a direct substitution of $a = 0$ into (6) leads to $\log \mathcal{L}(0, \mathbf{x}) = 0$, which cannot be used as a detector. But we can always take a monotonic transform of any detector without affecting its performance, and as long as $a > 0$ (which is the case for absorptive plumes), the transform $\log \mathcal{L}(a, \mathbf{x})/a$ is monotonic. We, therefore, define

$$\mathcal{D}_{\text{QAMF}}(\mathbf{x}) = \lim_{a \to 0} \frac{\log \mathcal{L}(a, \mathbf{x})}{a} = \frac{\partial}{\partial a} \log \mathcal{L}(a, \mathbf{x}) \Big|_{a=0} \tag{7}$$

with the expression on the right evaluated at $a = 0$. Define

$$\begin{aligned} \mathcal{Q}(\mathbf{x}) &= \lim_{a \to 0} \frac{\partial}{\partial a} \left[-\tfrac{1}{2} (\exp(aT)\mathbf{x} - \boldsymbol{\mu})' R^{-1} (\exp(aT)\mathbf{x} - \boldsymbol{\mu}) \right] \\ &= \lim_{a \to 0} \left[-\frac{\partial}{\partial a} (\exp(aT)\mathbf{x} - \boldsymbol{\mu})' \right] R^{-1} (\exp(aT)\mathbf{x} - \boldsymbol{\mu}) \\ &= \lim_{a \to 0} \left[-\exp(aT) T\mathbf{x} \right]' R^{-1} (\exp(aT)\mathbf{x} - \boldsymbol{\mu}) \\ &= -(T\mathbf{x})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) \end{aligned} \tag{8}$$

as a quadratic expression of interest, and observe that

$$\mathcal{D}_{\text{QAMF}}(\mathbf{x}) = \lim_{a \to 0} \frac{\log \mathcal{L}(a, \mathbf{x})}{a} = \mathcal{Q}(\mathbf{x}) + \tau \tag{9}$$

is the "quadratic matched filter." Note that the additive term τ has no influence on the performance of the QMF as a detector, and could safely be discarded. One advantage of keeping it in the definition, however, is that the expected value of $\mathcal{D}_{\text{QAMF}}(\mathbf{z})$ over non-plume pixels is zero. As a further aside, comparison with (2) shows that the quadratic $\mathcal{Q}(\mathbf{x})$ looks like the linear matched filter, but the match is to $T\mathbf{x}$ instead of $T\boldsymbol{\mu}$.

B. Generalized Likelihood Ratio Test

The GLRT formulation recognizes the dependence of the detector on plume strength a. The abovementioned LMP formulation considered the small a limit; by contrast, the GLRT formulation takes two steps: first an estimated plume strength $\widehat{a}$ is computed and then the likelihood is evaluated at that estimate. These steps can be combined into a single expression that defines the GLRT for any background distribution P_{bkg}

$$\mathcal{D}^*_{\text{GLRT}}(\mathbf{x}) = \frac{\max_a P_{\text{bkg}} \left(\exp(aT)\mathbf{x} \right) \exp(a\tau)}{P_{\text{bkg}}(\mathbf{x})}. \tag{10}$$

The maximum likelihood estimator for the plume strength $\widehat{a}$ is the value of a that achieves the maximum in (10)

$$\widehat{a} = \text{argmax}_a P_{\text{bkg}} \left(\exp(aT)\mathbf{x} \right) \exp(a\tau) / P_{\text{bkg}}(\mathbf{x}). \tag{11}$$

For the Gaussian distribution, the logarithm provides a convenient monotonic transform

$$\mathcal{D}_{\text{GLRT}}(\mathbf{x}) = 2 \log \mathcal{D}^*_{\text{GLRT}}(\mathbf{x})$$

$$= \max_a \left[2 \log P_{\text{bkg}} \left(\exp(aT)\mathbf{x} \right) + 2a\tau - 2 \log P_{\text{bkg}}(\mathbf{x}) \right] \tag{12}$$

$$\begin{aligned} &= \max_a \left[- \left(\exp(aT)\mathbf{x} - \boldsymbol{\mu} \right)' R^{-1} \left(\exp(aT)\mathbf{x} - \boldsymbol{\mu} \right) \right. \\ &\qquad \left. + 2a\tau + (\mathbf{x} - \boldsymbol{\mu})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right]. \end{aligned} \tag{13}$$

Note that (13) is a transcendental equation, and an exact closed-form solution is beyond the algebraic skills of this author. However, since we are interested in small a, we can approximate the solution using a Taylor series expansion up to quadratic terms in a. First, let us define

$$\mathcal{A}(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) \tag{14}$$

which is the Mahalanobis distance from $\mathbf{x}$ to the centroid of the distribution, and the basis of the RX anomaly detector [21]. Note that

$$\begin{aligned} \mathcal{A}\left(\exp(a\,T)\mathbf{x} \right) &= \mathcal{A}\left(\mathbf{x} + a\,T\mathbf{x} + \tfrac{1}{2} a^2 T^2 \mathbf{x} + O(a^3) \right) \\ &= (\mathbf{x} - \boldsymbol{\mu})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) + 2a (T\mathbf{x})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) \\ &\quad + a^2 (T\mathbf{x})' R^{-1} (T\mathbf{x}) + a^2 (T^2\mathbf{x})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) + O(a^3) \\ &= \mathcal{A}(\mathbf{x}) - 2a\mathcal{Q}(\mathbf{x}) + a^2 \mathcal{E}(\mathbf{x}) + O(a^3) \end{aligned} \tag{15}$$

where $\mathcal{Q}(\mathbf{x})$ is defined in (8) and

$$\mathcal{E}(\mathbf{x}) = (T\mathbf{x})' R^{-1} T\mathbf{x} + (T^2\mathbf{x})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) \tag{16}$$

is defined here for convenience. By combining (13)–(15), we have

$$\begin{aligned} \mathcal{D}_{\text{GLRT}}(\mathbf{x}) &= \max_a \left[-\mathcal{A}(\exp(aT)\mathbf{x}) + 2a\tau + \mathcal{A}(\mathbf{x}) \right] \\ &= \max_a \left[2a \left(\mathcal{Q}(\mathbf{x}) + \tau \right) - a^2 \mathcal{E}(\mathbf{x}) + O(a^3) \right]. \end{aligned} \tag{17}$$

In this formulation, (17) is maximized at $a = \widehat{a}$, where

$$\widehat{a} = \frac{\mathcal{Q}(\mathbf{x}) + \tau}{\mathcal{E}(\mathbf{x})} = \frac{-(T\mathbf{x})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) + \tau}{(T\mathbf{x})' R^{-1} T\mathbf{x} + (T^2\mathbf{x})' R^{-1} (\mathbf{x} - \boldsymbol{\mu})} \tag{18}$$

though we should properly restrict $\widehat{a} \geq 0$; i.e., reset negative quantities to zero: $\widehat{a} \leftarrow \max(0, \widehat{a})$.

Substituting this $\widehat{a}$ into (17) leads to the GLRT detector

$$\begin{aligned} \mathcal{D}_{\text{GLRT}}(\mathbf{x}) &= \frac{\mathcal{Q}(\mathbf{x}) + \tau}{\sqrt{\mathcal{E}(\mathbf{x})}} \\ &= \frac{-(T\mathbf{x})' R^{-1} (\mathbf{x} - \boldsymbol{\mu}) + \tau}{\sqrt{(T\mathbf{x})' R^{-1} T\mathbf{x} + (T\mathbf{x})' T R^{-1} (\mathbf{x} - \boldsymbol{\mu})}}. \end{aligned} \tag{19}$$

1) Brief Remark on Quadratic Detector Functions: In (17), we see a monotonic rescaling of the likelihood ratio expressed in terms of a Taylor series of the form $[2aA - a^2 B + O(a^3)]$. This is a pattern will arise again in Section IV-D, and is fairly common in general. Taking the derivative and setting to zero produces $\widehat{a} = A/B$, but restricting to nonnegative a makes that $\widehat{a} = \max(0, A)/B$. Inserting this expression back into the quadratic $[2aA - a^2 B]$ gives a GLRT that looks like $[\max(0, A)^2 / B]$. We can further rescale with the monotonic $f(x) = \text{sign}(x)\sqrt{|x|}$, and that leads to $\max(0, A)/\sqrt{B}$. As a detector, though, we can cheat and simply write $A/\sqrt{B}$ since negative values of the detector will not count as detections,

anyway. And weirdly, the cheat is often better in the region of the ROC curve we do not usually care about, getting better results in the high false alarm rate regime. (One author has interpreted this weirdness as a flaw in the GLRT itself [13].)

2) Some Remarks on Albedo-Corrected (AC) Estimates of Plume Strength: In Foote *et al.* [22], an AC estimate of $\widehat{a}$ is suggested. To explain it, we begin with a non-AC (NAC) estimate, based on the matched filter in (2). From the expression $\mathbf{x} = \mathbf{z} - aT\boldsymbol{\mu}$, where $\mathbf{z}$ is the plume-free pixel spectrum, assumed to be distributed as a Gaussian with mean $\boldsymbol{\mu}$ and covariance R, we can obtain a best estimate of a at a pixel $\mathbf{x}$

$$\widehat{a}_{\text{NAC}} = \frac{-(T\boldsymbol{\mu})' R^{-1}(\mathbf{x} - \boldsymbol{\mu})}{(T\boldsymbol{\mu})' R^{-1}(T\boldsymbol{\mu})}. \tag{20}$$

Based on the observation that this estimate is biased high for high-albedo ground pixels, Foote *et al.* [22] proposed an AC estimate

$$\widehat{a}_{\text{AC}} = \frac{1}{r} \times \widehat{a}_{\text{NAC}} = \frac{1}{r} \times \frac{-(T\boldsymbol{\mu})' R^{-1}(\mathbf{x} - \boldsymbol{\mu})}{(T\boldsymbol{\mu})' R^{-1}(T\boldsymbol{\mu})} \tag{21}$$

where the scalar factor r is given by $r = \mathbf{x}'\boldsymbol{\mu}/\boldsymbol{\mu}'\boldsymbol{\mu}$.

We observe that the $\widehat{a}$ defined in (18) has a similar AC character, although albedo correction was never explicitly incorporated into it; rather, the form of the expression in (18) arose naturally from the maximum likelihood estimate, based on exponential absorption.

IV. Elliptically Contoured (EC) Multivariate t-Distributed Background

The detectors above for GLRT in (19) and LMP in (9), as well as the linear matched filter in (2), were all derived under the assumption that the background distribution is Gaussian. In this section, these expressions are generalized for elliptically-contoured fat-tailed backgrounds given by the multivariate t distribution, which is often more descriptive of actual hyperspectral clutter [23], [24]. The multivariate t is a fatter-tailed distribution than the Gaussian, but like the Gaussian, depends on $\mathbf{x}$ though the elliptical term $\mathcal{A}(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\mu})' R^{-1}(\mathbf{x} - \boldsymbol{\mu})$. It is given by

$$P_{\text{bkg}}(\mathbf{x}) = c\left[(\nu - 2) + \mathcal{A}(\mathbf{x})\right]^{-\frac{d+\nu}{2}} \tag{22}$$

where d is the dimension of $\mathbf{x}$ and c is a constant prefactor. The scalar parameter ν characterizes how fat the tails are. The limit $\nu \to \infty$ corresponds to the Gaussian, and smaller values of ν describe distributions with fatter tails. As long as $\nu > 2$, the second moment exists (and so the covariance is well defined). We can estimate ν from the data using either an exceedance plot [24] or a moment method [25].

The extension of AMF to elliptically-contoured multivariate t-distributed backgrounds was developed in [26]

$$\mathcal{D}_{\text{EC}}(\nu, \mathbf{x}) = \mathcal{F}_\nu(\mathbf{x})\, \mathcal{D}_{\text{AMF}}(\mathbf{x}) \tag{23}$$

where

$$\mathcal{F}_\nu(\mathbf{x}) = \sqrt{\frac{(\nu - 1)}{(\nu - 2) + \mathcal{A}(\mathbf{x})}} \tag{24}$$

can be thought of as a kind of "fattening" factor. Observe that in the $\nu \to \infty$ limit, we have that $\mathcal{F}_\nu = 1$, and (23) reverts to the standard AMF in (2). At the other extreme, as $\nu \to 2$, we obtain the ACE detector (which is usually derived under other assumptions [27]–[29])

$$\mathcal{D}_{\text{ACE}}(\mathbf{x}) = \mathcal{F}_2(\mathbf{x})\, \mathcal{D}_{\text{AMF}}(\mathbf{x}) = \mathcal{D}_{\text{AMF}}(\mathbf{x}) / \sqrt{\mathcal{A}(\mathbf{x})}. \tag{25}$$

The extension of the quadratic matched filter to this fatter-tailed background begins with the likelihood ratio

$$\begin{aligned}\mathcal{L}(a, \mathbf{x}) = \frac{P_{\text{plume}}(\mathbf{x})}{P_{\text{bkg}}(\mathbf{x})} &= \frac{P_{\text{bkg}}(\exp(a\,T)\mathbf{x})\exp(a\tau)}{P_{\text{bkg}}(\mathbf{x})} \\ &= \frac{\left[\nu - 2 + \mathcal{A}(\exp(a\,T)\mathbf{x})\right]^{-\frac{d+\nu}{2}}}{\left[\nu - 2 + \mathcal{A}(\mathbf{x})\right]^{-\frac{d+\nu}{2}}} \times \exp(a\tau).\end{aligned} \tag{26}$$

Thus,

$$\begin{aligned}\mathcal{L}(a, \mathbf{x})^{-2/(\nu+d)} &= \frac{(\nu - 2) + \mathcal{A}(\exp(a\,T)\mathbf{x})}{(\nu - 2) + \mathcal{A}(\mathbf{x})} \times \exp\left(\frac{-2a\tau}{\nu + d}\right) \\ &= \left(1 + \frac{\mathcal{A}(\exp(a\,T)\mathbf{x}) - \mathcal{A}(\mathbf{x})}{(\nu - 2) + \mathcal{A}(\mathbf{x})}\right) \times \exp\left(\frac{-2a\tau}{\nu + d}\right) \\ &= \left(1 + \frac{\mathcal{F}_\nu(\mathbf{x})^2\left[\mathcal{A}(\exp(a\,T)\mathbf{x}) - \mathcal{A}(\mathbf{x})\right]}{\nu - 1}\right) \\ &\quad \times \exp\left(\frac{-2a\tau}{\nu + d}\right).\end{aligned} \tag{27}$$

A. Clairvoyant Detector for EC Distribution

If a is known, then the optimal detector is given by a monotonic rescaling of the likelihood ratio. Here, we can write

$$\mathcal{D}(a, \mathbf{x}) = (\nu - 1)\left(1 - \exp\left(\frac{2a\tau}{\nu + d}\right)\mathcal{L}(a, \mathbf{x})^{-2/(\nu+d)}\right) \tag{28}$$

as the monotonic transform, and then use (27) for $\mathcal{L}(a, \mathbf{x})^{-2/(\nu+d)}$ to obtain the clairvoyant detector

$$\mathcal{D}(a, \mathbf{x}) = \mathcal{F}_\nu^2(\mathbf{x})\left[\mathcal{A}(\mathbf{x}) - \mathcal{A}(\exp(a\,T)\mathbf{x})\right]. \tag{29}$$

B. Weak Plume Regime (Small a)

To derive the LMP and GLRT solutions for weak plumes, we need to express the likelihood ratio in terms of small a, and this will be done with Taylor series formula that expresses these quantities as low-order polynomials in a. Recall from (15) that we can write $\mathcal{A}(\exp(a\,T)\mathbf{x}) - \mathcal{A}(\mathbf{x})$ as a polynomial in a with coefficients that depend on $\mathcal{Q}(\mathbf{x})$ and $\mathcal{E}(\mathbf{x})$. We can also expand the exponential

$$\exp\left(\frac{-2a\tau}{\nu + d}\right) = 1 - \frac{2a\tau}{\nu + d} + \frac{2a^2\tau^2}{(\nu + d)^2} + O(a^3) \tag{30}$$

so the product in (27) becomes

$$\mathcal{L}(a, \mathbf{x})^{-2/(\nu+d)} = 1 - \frac{2a}{\nu - 1}\left[\mathcal{F}_\nu^2(\mathbf{x})\mathcal{Q}(\mathbf{x}) + \Theta_\nu\right]$$

TABLE III
TAXONOMY OF EXPRESSIONS FOR ABSORPTIVE PLUME DETECTORS

	Gaussian ($\nu \to \infty$)	EC	Fat ($\nu \to 2$)
Additive	$\mathcal{D}_{\text{AMF}}(\mathbf{x}) = -(T\boldsymbol{\mu})' R^{-1}(\mathbf{x} - \boldsymbol{\mu})$	$\mathcal{F}_\nu(\mathbf{x})\mathcal{D}_{\text{AMF}}(\mathbf{x})$	$\mathcal{D}_{\text{AMF}}(\mathbf{x})/\sqrt{\mathcal{A}(\mathbf{x})}$
LMP	$\mathcal{D}_{\text{QAMF}}(\mathbf{x}) = \mathcal{Q}(\mathbf{x}) + \tau$	$\mathcal{F}_\nu^2(\mathbf{x})\mathcal{Q}(\mathbf{x})$	$\mathcal{Q}(\mathbf{x})/\mathcal{A}(\mathbf{x})$
$\widehat{a}$	$\widehat{a} = (\mathcal{Q}(\mathbf{x}) + \tau)/\mathcal{E}(\mathbf{x})$	See Eq. (36)	Eq. (36) with $\nu = 2$
GLRT	$\mathcal{D}_{\text{GLRT}}(\mathbf{x}) = (\mathcal{Q}(\mathbf{x}) + \tau)/\sqrt{\mathcal{E}(\mathbf{x})}$	See Eq. (37)	See Eq. (38)
log	$\mathcal{D}_{\text{log-AMF}}(\mathbf{x}) = \mathbf{t}' \widetilde{R}^{-1}(\log \mathbf{x} - \widetilde{\boldsymbol{\mu}})$	$\widetilde{\mathcal{F}_\nu}(\log \mathbf{x})\, \mathcal{D}_{\text{log-AMF}}(\mathbf{x})$	$\mathcal{D}_{\text{log-AMF}}(\mathbf{x})/\sqrt{\widetilde{\mathcal{A}}(\log \mathbf{x})}$
Clairvoyant	$\mathcal{D}_{\text{C}}(\mathbf{x}) = \mathcal{A}(\mathbf{x}) - \mathcal{A}(\exp(a\,T)\mathbf{x})$	$\mathcal{F}_\nu^2(\mathbf{x})\mathcal{D}_{\text{C}}(\mathbf{x})$	$\mathcal{D}_{\text{C}}(\mathbf{x})/\mathcal{A}(\mathbf{x})$

TABLE IV
HYPERSPECTRAL DATASETS USED FOR EVALUATION OF ALGORITHMS

Dataset	Bands d	Parameter $\widehat{\nu}$	Plume Strength a/a_o	Plume Strength $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$
(a) OMI SO_2	320	5.75	5.21	0.0033
(b) OMI NO_2	320	3.76	6.02	0.0052
(c) AVIRIS 2016 CH_4	212	5.81	5.56	0.0026
(d) AVIRIS 2019 CH_4	213	6.44	4.57	0.0025

$$+ \frac{a^2}{\nu - 1}\left[\mathcal{F}_\nu^2(\mathbf{x})\mathcal{E}(\mathbf{x}) + \frac{4\mathcal{Q}(\mathbf{x})\Theta_\nu}{\nu - 1} + \frac{2\Theta_\nu^2}{\nu - 1}\right] + O(a^3) \tag{31}$$

with

$$\Theta_\nu = \frac{(\nu - 1)\tau}{\nu + d}. \tag{32}$$

So now a monotonic rescaling of the likelihood can be expressed

$$\begin{aligned} D(a, \mathbf{x}) &= (\nu - 1)(1 - \mathcal{L}(a, \mathbf{x})^{-2/(\nu + d)}) \\ &= 2a\left[\mathcal{F}_\nu^2(\mathbf{x})\mathcal{Q}(\mathbf{x}) + \Theta_\nu\right] \\ &\quad - a^2\left[\mathcal{F}_\nu^2(\mathbf{x})\mathcal{E}(\mathbf{x}) + \frac{4\mathcal{Q}(\mathbf{x})\Theta_\nu + 2\Theta_\nu^2}{\nu - 1}\right] + O(a^3). \end{aligned} \tag{33}$$

C. LMP Detector for EC Distribution

At this point, we can define the LMP detector in terms of the limit

$$\mathcal{D}_{\text{QEC}}(\mathbf{x}) = \lim_{a \to 0} \frac{\mathcal{D}(a, \mathbf{x})}{2a} - \Theta_\nu = \mathcal{F}_\nu^2(\mathbf{x})\, \mathcal{Q}(\mathbf{x}). \tag{34}$$

The ACE-ification of QMF is obtained by taking the $\nu \to 2$ limit of QEC

$$\mathcal{D}_{\text{QACE}}(\mathbf{x}) = \mathcal{Q}(\mathbf{x})/\mathcal{A}(\mathbf{x}) \tag{35}$$

D. GLRT Detector for EC Distribution

While the LMP detector required only terms linear in a, the GLRT detector requires the quadratic terms as well. From the formula for $\mathcal{D}(a, \mathbf{x})$ in (33), which can be expressed as $2aA - a^2B + O(a^3)$, we can see that this expression is maximized at

$$\widehat{a} = \frac{A}{B} = \frac{\mathcal{F}_\nu^2(\mathbf{x})\mathcal{Q}(\mathbf{x}) + \Theta_\nu}{\mathcal{F}_\nu^2(\mathbf{x})\mathcal{E}(\mathbf{x}) + \frac{4\mathcal{Q}(\mathbf{x})\Theta_\nu + 2\Theta_\nu^2}{\nu - 1}}. \tag{36}$$

Substituting $\widehat{a}$ into $\mathcal{D}(a, \mathbf{x})$ gives an expression of the form A^2/B. Following the same argument we used for the Gaussian case, namely that the combination of a square root and the imposition of $\widehat{a} \geq 0$ yields the form $A/\sqrt{B}$, we have

$$\mathcal{D}_{\text{GLRT-EC}}(\mathbf{x}) = \frac{\mathcal{F}_\nu^2(\mathbf{x})\mathcal{Q}(\mathbf{x}) + \Theta_\nu}{\sqrt{\mathcal{F}_\nu^2(\mathbf{x})\mathcal{E}(\mathbf{x}) + \frac{4\mathcal{Q}(\mathbf{x})\Theta_\nu + 2\Theta_\nu^2}{\nu - 1}}}. \tag{37}$$

Observe that in the $\nu \to \infty$ limit, $\mathcal{F}_\nu \to 1$ and $\Theta_\nu \to \tau$, which reduces to $\widehat{a} = (\mathcal{Q}(\mathbf{x}) + \tau)/\mathcal{E}(\mathbf{x})$, in agreement with (18) for a Gaussian background. In this limit, (37) similarly reduces to $(\mathcal{Q}(\mathbf{x}) + \tau)/\sqrt{\mathcal{E}(\mathbf{x})}$, which is the form in (19).

The $\nu \to 2$ limit is also of interest; this is the ACE-ification limit, and here it leads to

$$\begin{aligned} \mathcal{D}_{\text{GLRT-ACE}}(\mathbf{x}) &= \lim_{\nu \to 2} \mathcal{D}_{\text{GLRT-EC}}(\mathbf{x}) \\ &= \frac{\mathcal{Q}(\mathbf{x})/\mathcal{A}(\mathbf{x}) + \Theta_2}{\sqrt{\mathcal{E}(\mathbf{x})/\mathcal{A}(\mathbf{x}) + 4\mathcal{Q}(\mathbf{x})\Theta_2 + 2\Theta_2^2}} \end{aligned} \tag{38}$$

with $\Theta_2 = \tau/(d + 2)$.

V. LOGNORMAL BACKGROUND

A number of authors have recognized that exponential attenuation becomes linear when you are working in log space [30]–[34]. That is to say, (1) becomes

$$\log \mathbf{x} = \log \mathbf{z} - a\mathbf{t}. \tag{39}$$

In particular, Schaum [34] suggested the detector

$$\mathcal{D}_{\text{log-AMF}}(\mathbf{x}) = \mathbf{t}' \widetilde{R}^{-1}(\log \mathbf{x} - \widetilde{\boldsymbol{\mu}}) \tag{40}$$

where $\widetilde{\boldsymbol{\mu}}$ and $\widetilde{R}$ are the mean and covariance, respectively, of the background in log space. If the background is lognormal (that is: if the background is Gaussian in log space), then this log matched filter is the uniformly most powerful (UMP) detector [34].

Two natural extensions to this expression consider the case that the log space is not Gaussian but multivariate t-distributed. These expressions do not share the UMP property of the log-AMF detector, even in when the background *is* accurately modeled by a multivariate t, but they are better fits to these more general background distributions. Here, following (23), we can write

$$\mathcal{D}_{\text{log-EC}}(\mathbf{x}) = \widetilde{\mathcal{F}_\nu}(\log \mathbf{x})\, \mathcal{D}_{\text{log-AMF}}(\mathbf{x}) \tag{41}$$

and the $\nu \to 2$ limit provides

$$\mathcal{D}_{\text{log-ACE}}(\mathbf{x}) = \mathcal{D}_{\text{log-AMF}}(\mathbf{x})/\widetilde{\mathcal{A}}(\log \mathbf{x}). \tag{42}$$

TABLE V
PERFORMANCE OF VARIOUS DETECTORS OF PLUMES THAT ARE IMPLANTED ON REAL HYPERSPECTRAL DATA. BOLDFACE VALUES INDICATE THE BEST (OR NEARLY BEST) NONCLAIRVOYANT PERFORMANCE FOR EACH STATISTIC

(a) OMI SO_2

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.11893	0.09018	0.26088
$\mathbf{t}$-ACE	0.08527	0.08511	0.23474
$\mathbf{t}$-EC	0.08564	0.08513	0.23508
$T\mu$-AMF	0.01707	0.04856	0.16003
$T\mu$-ACE	0.00561	0.04152	0.13030
$T\mu$-EC	0.00564	0.04161	0.13064
QAMF	0.14044	0.07835	0.27786
QACE	0.07594	0.06334	0.22074
QEC	0.07674	0.06355	0.22157
GLRT-AMF	0.04280	0.03254	0.16442
GLRT-ACE	0.01618	0.02333	0.10177
GLRT-EC	0.01636	0.02333	0.10239
log-AMF	0.01604	0.02122	**0.01941**
log-ACE	**0.00525**	**0.01600**	0.04044
log-EC	**0.00524**	**0.01597**	0.04022
Clairvoyant-AMF	0.00512	0.01363	0.08572
Clairvoyant-ACE	0.00370	0.01187	0.08411
Clairvoyant-EC	0.00374	0.01189	0.08415

(b) OMI NO_2

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.03983	0.04045	0.17865
$\mathbf{t}$-ACE	0.01389	0.02709	0.11207
$\mathbf{t}$-EC	0.01400	0.02715	0.11234
$T\mu$-AMF	0.00282	0.01879	0.07615
$T\mu$-ACE	0.00070	0.01222	0.04527
$T\mu$-EC	0.00069	0.01226	0.04549
QAMF	0.05871	0.04501	0.22295
QACE	0.00807	0.02181	0.09330
QEC	0.00820	0.02193	0.09399
GLRT-AMF	0.00552	0.00916	0.05368
GLRT-ACE	0.00071	0.01102	0.04359
GLRT-EC	0.00070	0.01095	0.04367
log-AMF	0.00497	0.00588	**0.00779**
log-ACE	**0.00012**	**0.00483**	0.02074
log-EC	**0.00012**	**0.00480**	0.02059
Clairvoyant-AMF	0.00031	0.00319	0.01052
Clairvoyant-ACE	0.00059	0.00347	0.01804
Clairvoyant-EC	0.00058	0.00345	0.01791

(c) AVIRIS-NG 2016 CH_4

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.00755	0.01539	0.07983
$\mathbf{t}$-ACE	0.00151	0.01019	0.04858
$\mathbf{t}$-EC	0.00154	0.01025	0.04900
$T\mu$-AMF	0.00450	0.00942	0.04910
$T\mu$-ACE	0.00045	0.00521	0.02535
$T\mu$-EC	0.00046	0.00526	0.02563
QAMF	0.07289	0.04048	0.23644
QACE	0.02901	0.02409	0.14727
QEC	0.02963	0.02436	0.14929
GLRT-AMF	0.00301	0.00688	0.03439
GLRT-ACE	0.00040	0.00474	0.02299
GLRT-EC	0.00042	0.00476	0.02313
log-AMF	0.00244	**0.00333**	**0.00576**
log-ACE	**0.00029**	0.00390	0.01698
log-EC	**0.00029**	0.00387	0.01681
Clairvoyant-AMF	0.00067	0.00179	0.00547
Clairvoyant-ACE	0.00020	0.00122	0.00387
Clairvoyant-EC	0.00021	0.00122	0.00388

(d) AVIRIS-NG 2019 CH_4

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.00455	0.00772	0.02923
$\mathbf{t}$-ACE	0.00163	0.00614	0.02621
$\mathbf{t}$-EC	0.00162	0.00613	0.02617
$T\mu$-AMF	0.00219	0.00367	0.01156
$T\mu$-ACE	**0.00027**	0.00245	0.01030
$T\mu$-EC	**0.00028**	0.00245	0.01028
QAMF	0.01481	0.01366	0.04569
QACE	0.00283	0.00718	0.03298
QEC	0.00289	0.00719	0.03293
GLRT-AMF	0.00180	0.00327	0.01056
GLRT-ACE	**0.00028**	**0.00231**	0.00992
GLRT-EC	**0.00028**	**0.00231**	0.00991
log-AMF	0.00147	0.00246	**0.00464**
log-ACE	**0.00033**	0.00282	0.01195
log-EC	**0.00033**	0.00280	0.01189
Clairvoyant-AMF	0.00066	0.00158	0.00412
Clairvoyant-ACE	0.00018	0.00095	0.00287
Clairvoyant-EC	0.00019	0.00096	0.00287

Here, $\widetilde{\mathcal{F}_\nu}$ and $\widetilde{\mathcal{A}}$ are essentially the same as $\mathcal{F}_\nu$ and $\mathcal{A}$ [defined, respectively, in (24) and (14)], but using $\widetilde{\boldsymbol{\mu}}$ and $\widetilde{R}$ in place of $\boldsymbol{\mu}$ and R.

A taxonomy of all of these detectors is shown in Table III. (Note that many of these detectors are expressed in terms of intermediate quantities, which are summarized in Table II.)

VI. NUMERICAL EVALUATION

One way to measure the quality of a detection algorithm is to implant a plume into a hyperspectral image, and see if the algorithm can find it. We will do that here, but in a formalized way that makes two copies of the hyperspectral data. The first copy is untouched, while the second copy has plume implanted in every pixel. Thus, if $\mathbf{z}$ is the pixel in the first (presumably plume-free) copy, then the corresponding pixel in the ON-plume copy is $\mathbf{x} = \exp(-aT)\mathbf{z}$. The mean $\boldsymbol{\mu}$ and covariance R are estimated from the OFF-plume data (in other words, we are neglecting contamination effects, arguing that in an operational scenario, the ON-plume pixels will be rare). This is a kind of matched-pair approach [35], [36] that Rotman calls the two histogram method [37]. The obtained detector is applied to both ON-plume and OFF-plume pixels and from these a ROC curve is computed. Three statistics of interest to us are as follows:

1) false alarm rate at threshold with detection rate of 80% (FAR@DR80);
2) detection rate at threshold with false alarm rate of 0.05 (DR@FAR05);
3) area under the ROC curve (AUC).

The first statistic is more appropriate for most detection scenarios (where low false alarm rates are crucial), but the second statistic provides an alternative that might be relevant in scenarios for which detections are crucial and some false alarms can be tolerated. The AUC is widely employed, and arguably

TABLE VI
IMPLANTED PLUMES IN GAUSSIAN BACKGROUNDS. BOLDFACE VALUES INDICATE THE BEST (OR NEARLY BEST) NONCLAIRVOYANT PERFORMANCE FOR EACH STATISTIC. HERE, WE EXPECT (AND OBSERVE) THAT THE AMF AND EC VARIANTS OF THE VARIOUS DETECTORS PERFORM THE BEST, THOUGH OFTEN ONLY MARGINALLY SO. FOR THESE DATASETS, THE GLRT-BASED DETECTORS ARE SUBSTANTIALLY BETTER THAN THE $T\boldsymbol{\mu}$-BASED MATCHED FILTERS, WHILE THE LOG-BASED DETECTORS ARE MUCH WORSE

(a) OMI SO_2

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.11318	0.10770	0.25312
$\mathbf{t}$-ACE	0.11728	0.10843	0.25776
$\mathbf{t}$-EC	0.11318	0.10770	0.25312
$T\mu$-AMF	0.01004	0.06988	0.14770
$T\mu$-ACE	0.01081	0.07004	0.14884
$T\mu$-EC	0.01004	0.06988	0.14770
QAMF	0.03115	0.02897	0.16010
QACE	0.03414	0.02979	0.16558
QEC	0.03115	0.02897	0.16010
GLRT-AMF	**0.00016**	**0.00606**	**0.02849**
GLRT-ACE	**0.00016**	**0.00614**	0.02867
GLRT-EC	**0.00016**	**0.00606**	**0.02849**
log-AMF	0.71157	0.46984	0.94950
log-ACE	0.72343	0.47328	0.94603
log-EC	0.72492	0.47318	0.94638
Clairvoyant-AMF	0.00006	0.00139	0.00540
Clairvoyant-ACE	0.00008	0.00140	0.00554
Clairvoyant-EC	0.00006	0.00139	0.00540

(b) OMI NO_2

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.04882	0.08216	0.19871
$\mathbf{t}$-ACE	0.05046	0.08243	0.20050
$\mathbf{t}$-EC	0.04882	0.08216	0.19871
$T\mu$-AMF	0.00330	0.07040	0.13352
$T\mu$-ACE	0.00330	0.07036	0.13348
$T\mu$-EC	0.00330	0.07040	0.13352
QAMF	0.08729	0.06047	0.23913
QACE	0.08965	0.06114	0.24263
QEC	0.08729	0.06047	0.23913
GLRT-AMF	**0.00110**	**0.02237**	**0.08441**
GLRT-ACE	**0.00111**	0.02380	0.08711
GLRT-EC	**0.00110**	**0.02237**	**0.08441**
log-AMF	0.24037	0.20314	0.94614
log-ACE	0.24141	0.16836	0.76399
log-EC	0.23987	0.18953	0.90676
Clairvoyant-AMF	0.00040	0.00499	0.03457
Clairvoyant-ACE	0.00039	0.00504	0.03480
Clairvoyant-EC	0.00040	0.00499	0.03457

(c) AVIRIS-NG 2016 CH_4

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.00194	0.03043	0.08826
$\mathbf{t}$-ACE	0.00225	0.03057	0.08910
$\mathbf{t}$-EC	0.00194	0.03043	0.08826
$T\mu$-AMF	**0.00044**	0.02388	0.06850
$T\mu$-ACE	0.00051	0.02393	0.06885
$T\mu$-EC	**0.00044**	0.02388	0.06850
QAMF	0.03311	0.04237	0.17340
QACE	0.03779	0.04362	0.18023
QEC	0.03311	0.04237	0.17340
GLRT-AMF	**0.00040**	**0.01671**	**0.06113**
GLRT-ACE	**0.00044**	**0.01675**	0.06143
GLRT-EC	**0.00040**	**0.01671**	**0.06113**
log-AMF	0.29354	0.18789	0.76637
log-ACE	0.29698	0.17510	0.67367
log-EC	0.29658	0.17532	0.67593
Clairvoyant-AMF	0.00017	0.00321	0.01940
Clairvoyant-ACE	0.00018	0.00324	0.01946
Clairvoyant-EC	0.00017	0.00321	0.01940

(d) AVIRIS-NG 2019 CH_4

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.00431	0.01655	0.07043
$\mathbf{t}$-ACE	0.00476	0.01683	0.07182
$\mathbf{t}$-EC	0.00431	0.01655	0.07043
$T\mu$-AMF	**0.00092**	0.00933	0.03973
$T\mu$-ACE	0.00103	0.00942	0.04014
$T\mu$-EC	**0.00092**	0.00933	0.03973
QAMF	0.01522	0.02327	0.11504
QACE	0.01808	0.02440	0.12115
QEC	0.01522	0.02327	0.11504
GLRT-AMF	**0.00087**	**0.00866**	**0.03792**
GLRT-ACE	**0.00095**	**0.00871**	0.03813
GLRT-EC	**0.00087**	**0.00866**	**0.03792**
log-AMF	0.33403	0.19260	0.68247
log-ACE	0.34014	0.18823	0.64933
log-EC	0.34000	0.18822	0.64952
Clairvoyant-AMF	0.00063	0.00283	0.01361
Clairvoyant-ACE	0.00071	0.00287	0.01377
Clairvoyant-EC	0.00063	0.00283	0.01361

provides a kind of compromise between the first two, except that it places most of its weight on performance at high false alarm rates, so it is not usually a good choice in target detection applications. All of these statistics are scalar values between 0 and 1; in the Tables, we have listed FAR, 1-AUC, and 1-DR so that smaller values are always better.

As described in Table IV, four hyperspectral datasets, and three different gas species (NO_2, SO_2, and CH_4), were used to evaluate the algorithms. These datasets are as follows.

a) An SO_2 plume is implanted into an image obtained from the Ozone Monitoring Instrument (OMI) [32], and then resampled to match the wavelengths available on the Nanosat Atmospheric Chemistry Hyperspectral Observation System [38], [39] to produce a 320 band image with 320×1444 pixels.
b) An NO_2 plume implanted into a different OMI scene [40], and again resampled to produce a $320 \times 320 \times 1444$ pixel image.
c) A plume of CH_4 is implanted into a a 598×1200 chip of an AVIRIS-NG image of a landfill in India (scene `ang20160211t075004`) [41]; although there are 425 total channels in the scene, we only use the 212 longer-wavelength channels, that match the methane absorption spectrum.
d) A CH_4 plume is implanted into a $213 \times 598 \times 1500$ chip of an AVIRIS-NG image acquired over an oil and natural gas field near Carlsbad, New Mexico, USA (scene `ang20191023t151141`) [41].

In all four cases, the plume strength was chosen to be several times the characteristic strength a_o, defined in (46). The direct

TABLE VII
IMPLANTED PLUMES IN MULTIVARIATE t-DISTRIBUTED BACKGROUNDS, WITH $\nu = 3.0$. BOLDFACE VALUES INDICATE THE BEST (OR NEARLY BEST) NONCLAIRVOYANT PERFORMANCE FOR EACH STATISTIC. HERE, WE EXPECT (AND OBSERVE) THAT THE EC AND ACE VARIANTS OF THE DETECTORS TO BE BETTER, AND THEY ARE OFTEN SUBSTANTIALLY BETTER THAN THEIR AMF COUNTERPARTS. AS WITH THE GAUSSIAN BACKGROUND, THE LOG-BASED DETECTORS PERFORM POORLY

(a) OMI SO_2

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.02547	0.06596	0.13971
$\mathbf{t}$-ACE	0.01137	0.06199	0.14508
$\mathbf{t}$-EC	0.01116	0.06193	0.14464
$T\mu$-AMF	0.00561	0.04294	0.07309
$T\mu$-ACE	**0.00003**	0.04168	0.08283
$T\mu$-EC	**0.00003**	0.04165	0.08261
QAMF	0.01575	0.01791	0.07259
QACE	0.02066	0.02644	0.14307
QEC	0.01979	0.02611	0.14122
GLRT-AMF	0.00163	0.00391	**0.00917**
GLRT-ACE	**0.00000**	**0.00333**	0.01553
GLRT-EC	**0.00000**	**0.00331**	0.01539
log-AMF	0.62577	0.42527	0.94436
log-ACE	0.61906	0.42711	0.93704
log-EC	0.62118	0.42888	0.93694
Clairvoyant-AMF	0.00035	0.00191	0.00544
Clairvoyant-ACE	0.00000	0.00041	0.00068
Clairvoyant-EC	0.00000	0.00041	0.00068

(b) OMI NO_2

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.01305	0.05050	0.09814
$\mathbf{t}$-ACE	0.00183	0.04813	0.10869
$\mathbf{t}$-EC	0.00177	0.04808	0.10834
$T\mu$-AMF	0.00342	0.04324	0.06724
$T\mu$-ACE	**0.00000**	0.04226	0.07504
$T\mu$-EC	**0.00000**	0.04224	0.07488
QAMF	0.01462	0.03857	0.11397
QACE	0.02885	0.04396	0.17166
QEC	0.02788	0.04373	0.17043
GLRT-AMF	0.00103	0.01911	**0.03655**
GLRT-ACE	**0.00000**	**0.01318**	0.04676
GLRT-EC	**0.00000**	**0.01317**	0.04660
log-AMF	0.09564	0.11096	0.88316
log-ACE	0.08109	0.07092	0.31299
log-EC	0.08958	0.09338	0.62267
Clairvoyant-AMF	0.00019	0.00492	0.02121
Clairvoyant-ACE	0.00000	0.00143	0.00632
Clairvoyant-EC	0.00000	0.00143	0.00631

(c) AVIRIS-NG 2016 CH_4

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.00428	0.02424	0.04222
$\mathbf{t}$-ACE	**0.00002**	0.02395	0.05713
$\mathbf{t}$-EC	**0.00002**	0.02390	0.05675
$T\mu$-AMF	0.00307	0.02100	0.03489
$T\mu$-ACE	**0.00000**	0.02090	0.04812
$T\mu$-EC	**0.00000**	0.02085	0.04783
QAMF	0.00916	0.02648	0.07234
QACE	0.02209	0.03858	0.15239
QEC	0.02046	0.03807	0.14988
GLRT-AMF	0.00155	0.01619	**0.02996**
GLRT-ACE	**0.00000**	**0.01213**	0.04078
GLRT-EC	**0.00000**	**0.01210**	0.04048
log-AMF	0.02579	0.04005	0.03309
log-ACE	0.00614	0.03412	0.08959
log-EC	0.00620	0.02969	0.06215
Clairvoyant-AMF	0.00062	0.00354	0.01186
Clairvoyant-ACE	0.00000	0.00117	0.00389
Clairvoyant-EC	0.00000	0.00118	0.00389

(d) AVIRIS-NG 2019 CH_4

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.00679	0.01728	0.03205
$\mathbf{t}$-ACE	0.00023	0.01723	0.05544
$\mathbf{t}$-EC	0.00021	0.01714	0.05486
$T\mu$-AMF	0.00423	0.01296	**0.02207**
$T\mu$-ACE	**0.00002**	0.01321	0.04123
$T\mu$-EC	**0.00001**	0.01314	0.04082
QAMF	0.00652	0.02071	0.04731
QACE	0.02416	0.03534	0.15178
QEC	0.02231	0.03472	0.14856
GLRT-AMF	0.00186	0.01542	0.02562
GLRT-ACE	**0.00001**	**0.01099**	0.03897
GLRT-EC	**0.00001**	**0.01096**	0.03862
log-AMF	0.02173	0.03153	0.03552
log-ACE	0.00516	0.03141	0.09613
log-EC	0.00506	0.02628	0.07257
Clairvoyant-AMF	0.00113	0.00599	0.01476
Clairvoyant-ACE	0.00000	0.00135	0.00519
Clairvoyant-EC	0.00000	0.00135	0.00520

effect of the plumes on the spectra $\mathbf{x}$, however, was quite small, typically less than about half a percent.

Table V shows the performance of the various algorithms in detecting plume that has been implanted on the hyperspectral data. In all four cases, we observe that $T\boldsymbol{\mu}$ is always better than $\mathbf{t}$ as a target vector. The QMF family of detectors, based on the LMP hypothesis, is in all cases worse than the matched filters based on the $T\boldsymbol{\mu}$ targets. Although LMP is optimal for weak plumes, that optimality does not hold up when the plumes are strong enough to be detectable at reasonable false alarm rates (see Appendix A-B). The GLRT detectors are almost always better than the matched filter detectors (and the Clairvoyant detectors are usually better still, but that detector requires *a priori* knowledge of plume strength, and in practice that is not available). For these real datasets, the backgrounds were fatter tailed than Gaussian, and the multivariate t distribution provided a better model, as is evidenced by the (often considerably) better performance obtained by the ACE and EC variants of these detectors. That the ACE and EC detectors were nearly identical is a consequence of the $\widehat{\nu} \ll d$ that was observed for these datasets. The log-based detectors, especially log-ACE and log-EC, perform fairly well on this data, even outperforming (in the OMI NO_2 case) the clairvoyant detectors; that can happen because the clairvoyant detectors are optimized for specific background distributions (Gaussian for AMF, multivariate t for EC) that may not correspond to the actual distribution of the real hyperspectral data.

The log-transformed data provided interesting results. The log-AMF detector, which would be UMP if the background were lognormal [34], was never as good as the GLRT-ACE or GLRT-EC detectors. But the log-ACE and log-EC detectors provided excellent performance, outperforming even the GLRT-ACE and

TABLE VIII
IMPLANTED PLUMES IN LOGNORMAL BACKGROUND. BOLDFACE VALUES INDICATE THE BEST (OR NEARLY BEST) NONCLAIRVOYANT PERFORMANCE FOR EACH STATISTIC. AS PREDICTED, THE LOG-AMF IS OPTIMAL AGAINST THIS BACKGROUND. THE *OTHER* AMF-BASED DETECTORS, HOWEVER, ARE NOT AS GOOD AS THE EC AND ACE VARIANTS

(a) OMI SO_2

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.36872	0.21368	0.75528
$\mathbf{t}$-ACE	0.34334	0.18783	0.64374
$\mathbf{t}$-EC	0.34380	0.18806	0.64492
$T\mu$-AMF	0.29800	0.17470	0.69792
$T\mu$-ACE	0.25362	0.14296	0.54207
$T\mu$-EC	0.25418	0.14318	0.54316
QAMF	0.36392	0.21461	0.77243
QACE	0.30146	0.16721	0.61189
QEC	0.30262	0.16779	0.61451
GLRT-AMF	0.44301	0.29732	0.93017
GLRT-ACE	0.28480	0.16518	0.63493
GLRT-EC	0.28609	0.16608	0.63854
log-AMF	**0.03981**	**0.03311**	**0.17067**
log-ACE	0.04024	0.03335	0.17224
log-EC	**0.03981**	**0.03311**	**0.17067**
Clairvoyant-AMF	0.28590	0.18374	0.66012
Clairvoyant-ACE	0.28076	0.15920	0.57173
Clairvoyant-EC	0.28093	0.15937	0.57196

(b) OMI NO_2

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.13485	0.08147	0.47370
$\mathbf{t}$-ACE	0.06022	0.04587	0.22710
$\mathbf{t}$-EC	0.05987	0.04571	0.22607
$T\mu$-AMF	0.05695	0.03547	0.22608
$T\mu$-ACE	0.00113	0.00315	0.01373
$T\mu$-EC	0.00120	0.00332	0.01459
QAMF	0.10658	0.05992	0.41387
QACE	0.00250	0.00468	0.02118
QEC	0.00276	0.00505	0.02346
GLRT-AMF	0.02602	0.01744	0.06235
GLRT-ACE	0.00141	0.00330	0.01414
GLRT-EC	0.00143	0.00335	0.01431
log-AMF	**0.00048**	**0.00166**	**0.00621**
log-ACE	**0.00053**	**0.00171**	**0.00630**
log-EC	**0.00048**	**0.00166**	**0.00621**
Clairvoyant-AMF	0.00411	0.01179	0.03380
Clairvoyant-ACE	0.00127	0.00296	0.01198
Clairvoyant-EC	0.00129	0.00297	0.01194

(c) AVIRIS-NG 2016 CH_4

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.03906	0.02398	0.14933
$\mathbf{t}$-ACE	0.00374	0.00520	0.02278
$\mathbf{t}$-EC	0.00395	0.00533	0.02332
$T\mu$-AMF	0.02424	0.01576	0.08321
$T\mu$-ACE	0.00075	0.00205	0.00754
$T\mu$-EC	0.00082	0.00211	0.00774
QAMF	0.08747	0.05097	0.35074
QACE	0.00450	0.00729	0.03727
QEC	0.00497	0.00769	0.03971
GLRT-AMF	0.01531	0.00983	0.02669
GLRT-ACE	0.00066	0.00204	0.00770
GLRT-EC	0.00070	0.00208	0.00782
log-AMF	**0.00013**	**0.00075**	**0.00218**
log-ACE	**0.00017**	**0.00081**	0.00229
log-EC	**0.00013**	**0.00075**	**0.00218**
Clairvoyant-AMF	0.00199	0.00383	0.01108
Clairvoyant-ACE	0.00045	0.00153	0.00574
Clairvoyant-EC	0.00047	0.00155	0.00585

(d) AVIRIS-NG 2019 CH_4

Detector	FAR@DR80	1-AUC	1-DR@FAR05
$\mathbf{t}$-AMF	0.01457	0.01143	0.05130
$\mathbf{t}$-ACE	0.00323	0.00538	0.02386
$\mathbf{t}$-EC	0.00324	0.00538	0.02382
$T\mu$-AMF	0.00609	0.00564	0.01978
$T\mu$-ACE	0.00065	0.00183	0.00651
$T\mu$-EC	0.00066	0.00183	0.00654
QAMF	0.02981	0.01947	0.11121
QACE	0.00154	0.00290	0.01065
QEC	0.00155	0.00291	0.01071
GLRT-AMF	0.00443	0.00432	0.01268
GLRT-ACE	0.00068	0.00183	0.00644
GLRT-EC	0.00067	0.00183	0.00645
log-AMF	**0.00039**	**0.00145**	**0.00503**
log-ACE	**0.00047**	**0.00153**	0.00523
log-EC	**0.00039**	**0.00145**	**0.00503**
Clairvoyant-AMF	0.00151	0.00286	0.00877
Clairvoyant-ACE	0.00058	0.00167	0.00588
Clairvoyant-EC	0.00058	0.00167	0.00588

GLRT-EC detectors in three of the four cases (and coming very close in the fourth case).

In Tables VI–VIII, the background data are simulated from known distributions, with mean and covariance adjusted to match the mean and covariance of the original datasets. At a minimum, this provides a sanity check that the algorithms that are optimized for those distributions indeed perform best in those cases. Beyond that, however, these experiments also provide a measure of stability of the various algorithms to different background distributions.

Table VI uses a Gaussian background distribution, and in this case we see—as expected—that the AMF variants of the different detectors achieve the best statistics. But we also see that the ACE variants are nearly as good. For these datasets, the GLRT-based detectors are substantially better than the linear $T\boldsymbol{\mu}$-based matched filters. On the other hand, the log-based detectors are much worse (the FAR@DR80 statistic is hundreds of times larger for the log-based detectors, compared to the GLRT detectors).

Table VII uses a multivariate t-distributed background with very small parameter ($\nu = 3.0$) corresponding to a very fat-tailed elliptically-contoured distribution. Here, the ACE and EC detectors are best, and the AMF-based detectors are often much worse. The log-based detectors generally fare poorly here, as well, though they are not as bad as they were against Gaussian backgrounds (except for the OMI background with SO_2 plume). Interestingly, and somewhat counter-intuitively, the lowest false alarm rates for these fat-tailed backgrounds are lower than the

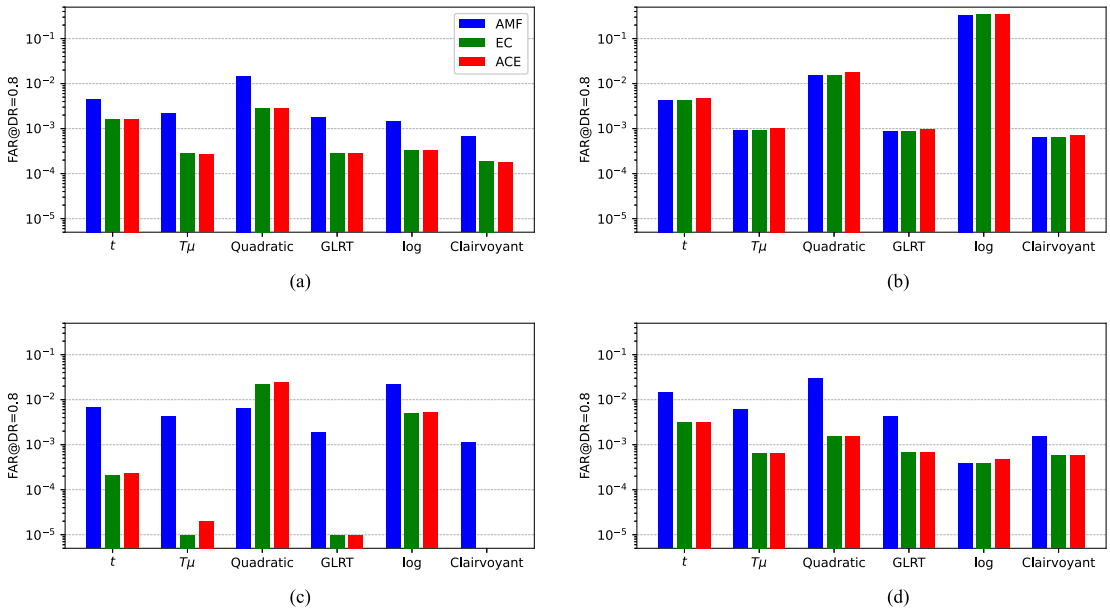

Fig. 1. Comparison of performance (based on the false alarm rate when detection rate is 80%) for different detectors against implanted CH_4 plumes on different backgrounds. (a) Hyperspectral background is based on the 2019 AVIRIS-NG image over Carlsbad, New Mexico [data set (d) in the text, and in Tables V–VIII]. (b) Simulated Gaussian background with same mean and covariance as the AVIRIS-NG data. (c) Simulated multivariate t-distributed background; note, the missing bars for Clairvoyant-EC and Clairvoyant-ACE correspond to zero false alarms. (d) Simulated lognormal background. For a Gaussian background, we expect the AMF to perform best, and although that is what we observe, we also observe that the EC and ACE variants exhibit nearly identical performance. Against the multivariate t background, the EC variants are (as expected) best, and against the lognormal background, the log-AMF detector is (again, as predicted by theory) best. Against the real hyperspectral background, we see the EC and ACE variants consistently and substantially outperforming their AMF counterparts. We also see the simple $T\boldsymbol{\mu}$-based matched filters nearly identical to the GLRT, which in turn is nearly as good as the clairvoyant detectors. And the log-EC and log-ACE detectors are also seen to do fairly well. Those log-based detectors, however, perform very poorly against the simulated Gaussian and multivariate t backgrounds. In every case, we see that the $\mathbf{t}$-based matched filters are outperformed by the $T\boldsymbol{\mu}$-based detectors, and the LMP-based Quadratic detectors do not fare well at all.

lowest false alarm rates on Gaussian backgrounds (as reported in Table VI), even though the plume strength is the same in both cases.

Finally, Table VIII uses a lognormal background. The log of this data will have the same mean and covariance as the original data in log space. As predicted by theory, the log-based detectors are best, and in particular the log-AMF lives up to its theoretical optimality in this situation. In three of the four cases (b)–(d), we see that the GLRT-EC detector, while not as good as the log-AMF, was still competitive, with only a few times as many false alarms at a fixed detection rate.

Fig. 1 illustrates the main results in these four tables, restricting consideration to a single statistic (FAR@DR80) and to case (d) abovementioned: a CH_4 plume implanted in four different backgrounds with the same mean and covariance as the 2019 AVIRIS-NG dataset.

Numerical results are necessarily anecdotal and, therefore, always somewhat tentative, but here are some general patterns that can be discerned from these simulations:

1) There can be orders-of-magnitude differences in the performance of different plume detectors.
2) One of the greatest sources of this variability in performance is differences in the background model, and how well the model matches the actual background.
3) When including the clairvoyant detectors in these comparisons, it is important to remember that what they are clairvoyant about is the plume strength, not the background distribution. This explains, for instance, how the clairvoyant-AMF (which assumes a Gaussian background) can fare so poorly in a lognormal background.
4) Although the AMF detectors are provably optimal when the background is Gaussian, they are only slightly better than the EC and ACE detectors in that scenario. And for non-Gaussian backgrounds, the AMF detectors are often substantially worse.
5) Although EC detectors are theoretically better than both AMF and ACE detectors, adapting themselves to the degree of fat-tailed-ness in the background distribution, they were found in these experiments to be only marginally better than the ACE detectors.
6) The log-AMF detector, although it has nice theoretical properties when the background distribution is lognormal [34], does not perform well under other background distributions. On the other hand, the log-EC and log-ACE detectors, although they also fare poorly with Gaussian and multivariate t-distributed backgrounds, did surprisingly well in the four real datasets.
7) Although the GLRT detectors are designed to be better than the simple matched-filters, we find that the $T\boldsymbol{\mu}$-ACE detector is very often competitive with the best GLRT detector for a given scenario. This bears some emphasis because that detector is algebraically very simple, but still different from the more widely employed $T\boldsymbol{\mu}$-AMF detector.

8) The t-based matched filters were in virtually every case outperformed by the Tmu-based equivalents.
9) The LMP-based "quadratic matched filters" did not perform well at all, and are not recommended for practical use. Appendix A-B discusses this from a theoretical perspective.

VII. Conclusion

It is often, if informally, asserted that a linear filter is appropriate for gas-phase plume detection because the exponential in Beer's law becomes linear in the weak plume limit. But a more careful derivation shows that the linear matched filter is not strictly appropriate even in the limit as plume strength goes to zero.

Furthermore, this weak plume limit is itself not quite appropriate if it implies that the plume is so weak that it is undetectable. The LMP solution that corresponds to the weak plume limit ($a \rightarrow 0$) and that leads to the quadratic matched filter is empirically found to be less effective than the GLRT solution (indeed, it is often less effective than the linear matched filter) when the plume is strong enough to be detected with reasonably small false alarm rate.

In deriving, the GLRT solutions, estimates of plume strength $\widehat{a}$ are made along the way, and it is observed in Section III-B2 that these estimates provide a built-in albedo correction, similar in flavor to the albedo correction in [22]. A useful future study would be to evaluate the relative performance of these different estimators.

As computers become faster and cheaper, the need for closed-form solutions has arguably diminished. Particularly with modern machine learning tools, it is possible to produce detectors that are adapted for virtually any physical situation that can be modeled [35], [36], [42]; there is, therefore, less need for linear approximations or simplifying assumptions. But even in this environment, simple expressions do still have advantages. They are useful when computation *is* limited (e.g., on small space-borne platforms [38], [39]), or when intuitive interpretation is desired.

Appendix A
Characteristic Plume Strength a_o

The detectability of a plume certainly depends on its strength a, but it also depends on the absorption spectrum $\mathbf{t}$, and the mean μ and covariance R of the clutter. For a given $\mathbf{t}$, μ, and R, we seek a characteristic plume strength a_o that corresponds roughly to a minimum detectable quantity. The actual MDQ will depend on $\mathbf{x}$ as well (or, more properly, on $\mathbf{z}$, the pixel spectrum in the absence of plume), and of course on the level of false alarm rate that you can tolerate with your detection.

We will go back to the approximation $\mathbf{x} = \mathbf{z} - aT\boldsymbol{\mu}$. The detector $(T\boldsymbol{\mu})'R^{-1}(\mathbf{x} - \boldsymbol{\mu})$ will, in the absence of plume ($\mathbf{x} = \mathbf{z}$), be a scalar random variable with mean zero and variance given by

$$\begin{aligned}\text{Variance}\left((T\boldsymbol{\mu})'R^{-1}(\mathbf{z}-\boldsymbol{\mu})\right) &= \left\langle (T\boldsymbol{\mu})'R^{-1}(\mathbf{z}-\boldsymbol{\mu})(\mathbf{z}-\boldsymbol{\mu})'R^{-1}(T\boldsymbol{\mu})\right\rangle \\ &= (T\boldsymbol{\mu})'R^{-1}(T\boldsymbol{\mu}). \end{aligned} \tag{43}$$

In the presence of plume of magnitude a, we have a mean value for the detector of

$$\begin{aligned}\text{Mean}\left((T\boldsymbol{\mu})'R^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) &= \left\langle (T\boldsymbol{\mu})'R^{-1}(\mathbf{z}-aT\boldsymbol{\mu}-\boldsymbol{\mu})\right\rangle \\ &= -a(T\boldsymbol{\mu})'R^{-1}(T\boldsymbol{\mu}). \end{aligned} \tag{44}$$

To obtain an "n sigma" detection requires

$$|\text{Mean}| = n\sqrt{\text{Variance}} \tag{45}$$

which is to say that $a(T\boldsymbol{\mu})'R^{-1}(T\boldsymbol{\mu}) = n\sqrt{(T\boldsymbol{\mu})'R^{-1}(T\boldsymbol{\mu})}$. That is: $a = n/\sqrt{(T\boldsymbol{\mu})'R^{-1}(T\boldsymbol{\mu})}$. We will, therefore, take our characteristic plume strength to be

$$a_o = 1/\sqrt{(T\boldsymbol{\mu})'R^{-1}(T\boldsymbol{\mu})} = 1/\|R^{-1/2}T\boldsymbol{\mu}\| \tag{46}$$

and recognize that we will need plumes that are several times stronger than this in order to get detections at low false alarm rate.

Perhaps the main use of a_o is to inform our choice of a in simulations assessing algorithms for detecting weak plumes. We may simulate multiple scenarios (e.g., different gases on different backgrounds), but our baseline comparison is in terms of the characteristic a_o for the different scenarios.

A. Remark on the Strength of Weak Plumes

It bears remarking that although a plume whose strength is on the order of the characteristic strength a_o is just detectable, it is still a weak plume. Consider a typical pixel, for which $\mathbf{x} = \boldsymbol{\mu}$, and notice that the magnitude of the difference between OFF-plume and ON-plume is given by $\|\exp(-a_oT)\boldsymbol{\mu} - \boldsymbol{\mu}\|$. A natural comparison is the magnitude of the pixel itself, which is $\|\boldsymbol{\mu}\|$, and the ratio is given by

$$\begin{aligned}\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} &\sim \frac{\|\exp(-a_oT)\boldsymbol{\mu}-\boldsymbol{\mu}\|}{\|\boldsymbol{\mu}\|} \approx \frac{\|a_oT\boldsymbol{\mu}\|}{\|\boldsymbol{\mu}\|} \\ &\sim \frac{\|T\boldsymbol{\mu}\|}{\|\boldsymbol{\mu}\|} \times \frac{1}{\|R^{-1/2}T\boldsymbol{\mu}\|}. \end{aligned} \tag{47}$$

A slightly (but only slightly) more accurate derivation yields

$$\begin{aligned}\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} &\sim \frac{\|\exp(-a_oT)\mathbf{x}-\mathbf{x}\|}{\|\mathbf{x}\|} \approx \frac{\|a_oT\mathbf{x}\|}{\|\mathbf{x}\|} \\ &\sim \sqrt{\frac{\|T\boldsymbol{\mu}\|^2 + \mathbf{t}'R\mathbf{t}}{\|\boldsymbol{\mu}\|^2 + \text{Trace}(R)}} \times \frac{1}{\|R^{-1/2}T\boldsymbol{\mu}\|}. \end{aligned} \tag{48}$$

We find in practice this ratio is often much–much less than one, even when the plume to clutter ratio in the matched filter space is of order one.

Observe that the covariance R typically has its maximum eigen-direction along a vector that is roughly parallel to $\boldsymbol{\mu}$; but since the data values are positive, we can say that the largest eigenvalue is roughly bounded by $\|\boldsymbol{\mu}\|^2$, and that most eigenvalues tend to be quite a bit smaller than that. Thus, the smallest eigenvalue of R^{-1} is of order $1/\|\boldsymbol{\mu}\|^2$ with most eigenvalues quite a bit larger than that. Thus, for most vectors $\mathbf{u}$, we expect $\mathbf{u}'R^{-1}\mathbf{u}/\mathbf{u}'\mathbf{u} \gg 1/\|\boldsymbol{\mu}\|^2$. Taking $\mathbf{u} = T\boldsymbol{\mu}$, we obtain:

$(T\boldsymbol{\mu})'R^{-1}T\boldsymbol{\mu}/\|T\boldsymbol{\mu}\|^2 \gg 1/\|\boldsymbol{\mu}\|^2$ from which it follows:

$$\frac{\|T\boldsymbol{\mu}\|}{\|\boldsymbol{\mu}\|} \times \frac{1}{\|R^{-1/2}T\boldsymbol{\mu}\|} \ll 1. \tag{49}$$

That is: $\|\Delta\mathbf{x}\| \ll \|\mathbf{x}\|$, which means that the effect of the $a = a_o$ plume is small. Numerically, we find with the datasets used here, that this ratio is less than one in a thousand. (Specifically, $\|\Delta\mathbf{x}\|/\|\mathbf{x}\| = 0.00087$ for the for OMI data with the NO_2 plume, 0.00063 for OMI data with the SO_2 plume, 0.00047 for the 2016 AVIRIS-NG data with CH_4 plume, and 0.00054 for the 2019 CH_4 data. The values in Table IV are somewhat larger than this because $a > a_o$ for those plumes.) Thus, the multiplying effect of matched filters is considerable here. Plumes that are virtually invisible in the raw data can be readily evident in matched filter projections.

B. Remark on the Inadequacy of the QMF (And of the LMP Generally)

This notion of a characteristic plume strength also allows us to explain why the QMF, which is "optimized" for weak plumes, does such a poor job of detecting weak plumes. In the derivation of QMF, we neglect the $O(a^2)$ terms on grounds that a is small for weak plumes. But small compared to what? In particular, one of the quadratic terms that we neglected, as seen in (17), was $a^2(T\mathbf{x})'R^{-1}(T\mathbf{x})$. But in terms of the characteristic plume strength a_o, we see that

$$\left\langle\, a^2(T\mathbf{x})'R^{-1}(T\mathbf{x})\,\right\rangle > a^2(T\boldsymbol{\mu})'R^{-1}(T\boldsymbol{\mu}) = a^2/a_o^2. \tag{50}$$

If our "weak" plume is at least strong enough to be detected, then this term is greater than 1, and can hardly be assumed to approach zero. Put another way, the QMF detector is optimized for plumes that are too weak (i.e., for $a \ll a_o$) to be effectively detected.

ACKNOWLEDGMENT

The author is grateful to Alan Schaum for valuable insights and useful discussions and, in particular, for emphasizing the importance of the $|d\mathbf{z}/d\mathbf{x}|$ factor in (3). He also thanks Bernard Foy, with whom he had many insightful discussions over the years, for help acquiring and interpreting the OMI data and Philip Dennison for help understanding the AVIRIS-NG data at the NASA-JPL website [41]. The author also thanks both these organizations for making their data freely available.

REFERENCES

[1] 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," *Appl. Opt.*, vol. 35, pp. 2802–2809, 1996.

[2] S. P. Love, F. Goff, D. Counce, C. Siebe, and H. Delgado, "Passive infrared spectroscopy of the eruption plume at popocatépetl volcano, Mexico," *Nature*, vol. 396, pp. 563–567, 1998.

[3] S. P. Love *et al.*, "Passive infrared spectroscopic remote sensing of volcanic gases: Ground-based studies at White Island and Ruapehu, New Zealand, and Popocatépetl, Mexico," in *Remote Sensing of Active Volcanism (ser. Geophysical Monograph)*, P. Mouginis-Mark, J. Crisp, and J. Fink, Eds., vol. 116. Washington, DC, USA: Amer. Geophysical Union, 2000, pp. 117–138.

[4] S. J. Young, "Detection and quantification of gases in industrial-stack plumes using thermal-infrared hyperspectral imaging," the aerospace corporation, Tech. Rep. ATR-2002(8407)-1, 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." in *Proc. SPIE*, vol. 4722, pp. 98–109, 2002.

[6] D. G. Manolakis and F. M. D'Amico, "A taxonomy of algorithms for chemical vapor detection with hyperspectral imaging spectroscopy," *Proc. SPIE*, vol. 5795, pp. 125–133, 2005.

[7] J. Theiler, B. R. Foy, and A. M. Fraser, "Characterizing non-gaussian clutter and detecting weak gaseous plumes in hyperspectral imagery," *Proc. SPIE*, vol. 5806, pp. 182–193, 2005.

[8] D. R. Thompson *et al.*, "Real-time remote detection and measurement for airborne imaging spectroscopy: A case study with methane," *Atmospheric Meas. Techn.*, vol. 8, pp. 4383–4397, 2015.

[9] C. Frankenberg *et al.*, "Airborne methane remote measurements reveal heavy-tail flux distribution in four corners region," in *Proc. Nat. Acad. Sci.*, vol. 113, pp. 9734–9739, 2016.

[10] K. N. Buckland, S. J. Young, E. R. Keim, B. R. Johnson, P. D. Johnson, and D. M. Tratt, "Tracking and quantification of gaseous chemical plumes from anthropogenic emission sources within the los angeles basin," *Remote Sens. Environ.*, vol. 201, pp. 275–296, 2017.

[11] A. Beer, "Bestimmung der absorption des rothen lichts in farbigen flussigketiten," *Ann. Physik*, vol. 86, pp. 78–88, 1852.

[12] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*. New York, NY, USA: Springer, 2005.

[13] A. Schaum, "Continuum fusion: A theory of inference, with applications to hyperspectral detection," *Opt. Exp.*, vol. 18, pp. 8171–8181, 2010.

[14] J. Theiler, "Confusion and clairvoyance: Some remarks on the composite hypothesis testing problem," in *Proc. SPIE*, vol. 8390, 2012, Art. no. 839003.

[15] A. Schaum, "Clairvoyant fusion: A new methodology for designing robust detection algorithms," in *Proc. SPIE*, vol. 10004, 2016, Art. no. 100040C.

[16] J. Theiler and A. Schaum, "Some closed-form expressions for absorptive plume detection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp., 2020*, pp. 1786–1789.

[17] E. J. Kelly, "Performance of an adaptive detection algorithm: Rejection of unwanted signals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, no. 2, pp. 122–133, Mar. 1989.

[18] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-10, no. 6, pp. 853–863, Nov. 1974.

[19] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, vol. II. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.

[21] I. S. Reed and X. Yu, "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1760–1770, Oct. 1990.

[22] M. D. Foote *et al.*, "Fast and accurate retrieval of methane concentration from imaging spectrometer data using sparsity prior," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6480–6492, Sep. 2020.

[23] D. Manolakis, D. Marden, J. Kerekes, and G. Shaw, "On the statistics of hyperspectral imaging data," *Proc. SPIE*, vol. 4381, pp. 308–316, 2001.

[24] D. B. Marden and D. Manolakis, "Using elliptically contoured distributions to model hyperspectral imaging data and generate statistically similar synthetic data," in *Proc. SPIE*, vol. 5425, pp. 558–572, 2004.

[25] J. Theiler, C. Scovel, B. Wohlberg, and B. R. Foy, "Elliptically-contoured distributions for anomalous change detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 271–275, Apr. 2010.

[26] J. Theiler and B. R. Foy, "EC-GLRT: Detecting weak plumes in non-gaussian hyperspectral clutter using an elliptically-contoured generalized likelihood ratio test," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2008, pp. I-221–I-224.

[27] L. L. Scharf and L. T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, vol. 2, 1996, pp. 1114–1117.

[28] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 427–438, Feb. 2005.

[29] D. Manolakis, M. Pieper, E. Truslow, T. Cooley, M. Brueggeman, and S. Lipson, "The remarkable success of adaptive cosine estimator in hyperspectral target detection," *Proc. SPIE*, vol. 8743, 2013, Art. no. 874302.

[30] J. Li, C. C. Schmidt, J. P. Nelson III, T. J. Schmit, and W. Menzel, "Estimation of total atmospheric ozone from GOES sounder radiances with high temporal resolution," *J. Atmospheric Ocean. Technol.*, vol. 18, pp. 157–168, 2001.
[31] N. A. Krotkov, S. A. Carn, A. J. Krueger, P. K. Chartia, and K. Yang, "Band residual difference algorithm for retrieval of SO_2 from the aura ozone monitoring instrument (OMI)," *Trans. Geosci. Remote Sens.*, vol. 44, pp. 1259–1266, 2006.
[32] K. Yang, N. A. Krotkov, A. J. Krueger, S. A. Carn, P. K. Bhartia, and P. F. Levelt, "Retrieval of large volcanic SO_2 columns from the aura ozone monitoring instrument: Comparison and limitations," *J. Geophysical Res.*, vol. 112, 2007, Art. no. D 24S43.
[33] C. Li, J. Joiner, N. A. Krotkov, and P. K. Bhartia, "A fast and sensitive new satellite SO_2 retrieval algorithm based on principal component analysis: Application to the ozone monitoring instrument," *Geophysical Res. Lett.*, vol. 40, pp. 6314–6318, 2013.
[34] A. Schaum, "A uniformly most powerful detector of gas plumes against a cluttered background," *Remote Sens. Environ.*, vol. 260, 2021, Art. no. 112443.
[35] J. Theiler, "Matched-pair machine learning," *Technometrics*, vol. 55, pp. 536–547, 2013.
[36] J. Theiler, "Transductive and matched-pair machine learning for difficult target detection problems," *Proc. SPIE*, vol. 9088, 2014, Art. no. 90880E.
[37] M. Bar-Tal and S. R. Rotman, "Performance measurement in point target detection," *Infrared Phys. Technol.*, vol. 37, pp. 231–238, 1996.
[38] S. P. Love *et al.*, "High-resolution hyperspectral imaging of dilute gases from cubesat platforms," in *Proc. Amer. Geophysical Union, Fall Meeting*, 2018, pp. A41K–3107.
[39] J. Theiler, B. R. Foy, C. Safi, and S. P. Love, "Onboard CubeSat data processing for hyperspectral detection of chemical plumes," in *Proc. SPIE*, vol. 10644, 2018, Art no. 1064405.
[40] A. K. Mebust, A. R. Russell, R. C. Hudman, L. C. Valin, and R. C. Cohen, "Characterization of wildfire NO_x emissions using MODIS fire radiative power and OMI tropospheric NO_2 columns," *Atmospheric Chem. Phys.*, vol. 11, pp. 5839–5851, 2011.
[41] NASA Jet Propulsion Laboratory, "Benchmark dataset for methane and carbon dioxide plumes," Accessed: Mar. 2021. [Online]. Available: https://avirisng.jpl.nasa.gov/benchmark_methane_carbon_dioxide.html
[42] A. Ziemann, M. Kucer, and J. Theiler, "A machine learning approach to hyperspectral detection of solid targets," in *Proc. SPIE*, vol. 10644, 2018, Art. no. 1064404.

James Theiler received the Ph.D. degree in physics from the California Institute of Technology, Pasadena, CA, USA, in 1987.

He held appointments at UCSD, MIT Lincoln Laboratory, Los Alamos National Laboratory, and the Santa Fe Institute. He joined technical staff at Los Alamos in 1994, and was named a Laboratory Fellow in 2005. His professional interests include statistical modeling, machine learning, image processing, and remote sensing.

Dr. Theiler is currently a Senior Area Editor for IEEE TRANSACTIONS ON COMPUTATIONAL IMAGING.