Practical Methods for Rapid and Accurate Computation of Interferometric Spectra for Remote Sensing Applications

Christopher D. Barnet, John M. Blaisdell, and Joel Susskind

***Abstract*—The apodization of an interferogram corresponds to a linear transformation in spectral space between unapodized and apodized radiances. Many apodization functions have well-behaved numerical inverse transformations, and we show an analytic inverse for the Hamming apodization function. The inverse transformation has many practical uses for remote sensing applications and can also be used theoretically to show the equivalence between unapodized spectra and properly apodized spectra. The inverse transformation, which is a representation of the discrete convolution theorem, can be used to readily convert computed apodized spectra to spectra computed for other symmetric apodization functions (including unapodized), which may have poorer characteristics with regard to calculating channel-transmittance parameters or radiances. We also show a quantitative method for comparing apodization functions of different mathematical forms.**

***Index Terms*—Convolution, cosine transforms, digital filters, discrete cosine transforms, optical interferometry, optical transfer functions, remote sensing, satellite applications.**

I. Introduction

THE NEXT generation of infrared (IR) atmospheric sounders is being designed for high-spectral resolving power ($\nu/\Delta\nu \simeq 1200$) and high SNR. NASA's Earth Observing System (EOS) will use a grating IR instrument (AIRS) with ≈2400 spectral channels, together with two microwave instruments (AMSU and HSB), to retrieve the geophysical state of the atmosphere [1]. Interferometers have also been proposed for these applications. For example, the European meteorological satellite program (EUMETSAT) plans to use IASI [2], with ≈8000 spectral channels in the IR and a maximum optical-path difference of $L = 2$ cm. These advanced IR sounders have a channel SNR on the order of 500. Both spectral characteristics and SNR are critical to retrieval accuracy.

For an unapodized (boxcar apodized) interferometer, the channel spectral response function (CSRF), given by the cosine transform of the apodization function, contains a $\sin(y)/y$ function, where y is proportional to the frequency separation from the channel center. This function has large side-lobes that alternate in sign and fall off slowly with increasing frequency separation. Typically, interferograms are apodized (the interferogram is multiplied by a function which, in effect, smooths the spectrum) to produce a CSRF that is localized and has small side-lobes. The apodization process modifies the channel signal and noise characteristics in both beneficial and detrimental ways. We begin in Section II with a brief review of the spectral characteristics of an interferometer without apodization.

A number of publications [3], [4] have dealt with optimal apodization functions for use in various applications. In this paper, we will investigate the effects of the apodization process in the context of atmospheric sounding. The motivation for this paper is the proposed use of unapodized measurements for atmospheric sounding purposes [5]–[7]. Throughout this paper, methodology for the computation of unapodized radiances from geophysical parameters is discussed in the sounding context.

Interferometers have previously been used for remote sensing of the atmospheres of Earth (e.g., NIMBUS-7) and other planets (e.g., IRIS on Voyager), but the fast integration time required for EOS and EUMETSAT places new demands on data processing. Both EOS and EUMETSAT will have to produce individual soundings in approximately 0.1 s. All sounding algorithms implicitly or explicitly require computation of instrumental radiances from the atmospheric state. The traditional method of doing fast atmospheric soundings is to perform radiative transfer calculations using a precalculated transmittance model for each spectral channel of the instrument [8]–[11], which parameterizes the channel-averaged transmittance as a function of atmospheric conditions. This procedure, called the forward computation, is described in Section III, in which we argue that a localized CSRF is a requirement for existing rapid and accurate channel-transmittance models and radiance calculations.

In Section IV, we show that a linear relationship exists between complete sets of channel radiances computed using any of a family of cosine apodization functions, and in Section V, we show that an inverse transformation exists for these apodization functions. The inverse transformation can be used to compute unapodized radiances quickly and accurately from radiances computed with an apodized forward model. The existence of a reversible linear operator proves that there is equivalent information content between cosine apodized and unapodized radiances. In Section VI, we demonstrate that the same equivalence holds between unapodized radiances and any of a family of more general well-behaved apodization function. In Section VII, we

Manuscript received September 28, 1998; revised January 11, 1999. This work was supported by the Integrated Program Office (IPO), Silver Spring, MD, as part of an internal concept study of using an interferometer for the IR sounder in the National Polar-Orbiting Operational Environmental Satellite System (NPOESS).

C. D. Barnet and J. M. Blaisdell are with the General Sciences Corporation, a subsidiary of Science Applications International Corporation, Beltsville, MD 20705-2675 USA (e-mail: cbarnet@spectra.gsfc.nasa.gov).

J. Susskind is with NASA Goddard Space Flight Center, Code 910, Greenbelt, MD 20771 USA.

Publisher Item Identifier S 0196-2892(00)00013-9.

discuss the effects of apodization as they relate to sounding applications and show that apodization should have no effect on retrieval results. The only differences that might arise in retrieval results are due to possible advantages of one form or the other with regard to implementation of certain retrieval algorithms. In Section VIII, we analyze many apodization functions found in the literature and discuss which are reasonable functions to use, both as a means to compute radiances for any apodization function, including unapodized radiances, and in the retrieval process itself.

II. Unapodized Channel Spectral-Response Function

In this paper, we assume that the interferograms are perfectly calibrated and all other instrument effects (e.g., phase errors, self-apodization, etc.) have been removed. The raw data product from an interferometer, the interferogram, is a cosine transform of the incoming radiance. One obtains the channel i radiance $R(i)$ by taking the cosine transform of the product of the interferogram $I(\delta)$ with an apodization function $A(\delta)$. The cosine transform has finite limits, since the interferometer has a finite maximum optical-path difference L:

$$R(i) = \frac{\displaystyle\int_{\delta=-\infty}^{\infty} A(\delta)\cdot I(\delta)\cdot \cos(2\pi(\nu-\nu_i)\cdot\delta)\cdot d\delta}{\displaystyle\int_{\delta=-\infty}^{\infty} A(\delta)\cdot d\delta}. \quad (1)$$

In the interferogram domain, the unapodized (or boxcar) apodization function is defined as

$$\begin{aligned} A_U(\delta) &= 1 \quad \text{for } |\delta| \le L \\ &= 0 \quad \text{for } |\delta| > L \end{aligned} \quad (2)$$

where δ is the optical-path difference. The CSRF is the cosine transform of the apodization function. The CSRF for an unapodized interferometer ϕ_U is equal to

$$\begin{aligned} \phi_U(\nu,\nu_i) &= \frac{\sin(y)}{y} + \frac{\sin(z)}{z} \\ &\equiv \operatorname{sinc}(y) + \operatorname{sinc}(z) \end{aligned} \quad (3)$$

where

y $= 2\pi L\cdot(\nu-\nu_i)$
z $= 2\pi L\cdot(\nu+\nu_i)$
ν_i channel i center frequency
ν frequency.

The unapodized channel i radiance $R_U(i)$ also can be given by a convolution of the CSRF with the monochromatic radiance, $R(\nu)$ at the entrance to the interferometer

$$R_U(i) = \frac{\displaystyle\int \phi_U(\nu,\nu_i)R(\nu)\,d\nu}{\displaystyle\int \phi_U(\nu,\nu_i)\,d\nu}. \quad (4)$$

Typically, the $\operatorname{sinc}(z)$ term in (3) is ignored. For a smooth spectrum, the term is insignificant. However, for a spectrum that has structure on the order of the sampling, the term can become significant due to the integration across the entire band. For typical terrestrial spectra, numerical simulations of a high signal-to-noise remote sounding instrument (e.g., IASI) indicate this term is $\approx$0.05% near 650 cm^2 or $\approx$1/4 of the instrumental noise. Resonant spectral structure (trace gases, etc., with a line spacing of $\approx 1/L$) would have larger significance. Therefore, we will retain this term for completeness in these discussions.

The Nyquist sampling theorem states that the optimal sampling of channels is such that the channel spacing is $\delta\nu = 1/(2L)$ in the frequency domain [12]. No additional information is gained by sampling the interferogram at a higher rate (although oversampling is used to reduce out-of-band aliasing). However, information is lost if the interferogram is sampled at a lower rate. The resulting unapodized spectrum is given as an array of radiance values $R_U(i)$ where i is the channel-index number.

The $\operatorname{sinc}(y)$ function has large side-lobes that alternate in sign between the zeros of the function spaced at $y = \pm n\pi$. The first four side-lobes have heights of -21.7%, $+12.8\%$, -9.1%, and $+7.1\%$ with respect to the central lobe. The full-width-half-maximum (FWHM) of the $\operatorname{sinc}(y)$ function is equal to

$$\text{FWHM}_U = \frac{0.603355}{L} \simeq \frac{1.2}{2L}. \quad (5)$$

Only 45% of the area of the unapodized CSRF comes from the central lobe. The remainder of the contribution to the radiance comes primarily from the first few side-lobes, but nonnegligible contributions arise from very distant frequencies within the bandpass, as the heights of the unapodized CSRF side-lobes are still above 1% at 30 zeros ($\nu = \nu_i \pm 30/L$) from the central lobe.

The use of nonlocalized, unapodized radiances can produce complications in the retrieval of geophysical parameters. For multispectral retrievals (e.g., combining microwave and IR radiances) it is convenient, but not necessary, to represent radiances as brightness temperatures (i.e., the temperature of a blackbody with the same radiance.) For unapodized spectra, brightness temperature is a meaningless concept due to the distortion caused by negative side-lobes that can produce negative-channel radiances. The unapodized CSRF also produces severe complications in the development of efficient and accurate methods to compute channel radiances, such as the use of channel-averaged transmittance functions. This will be discussed in the next section.

III. Rapid-Transmittance Algoithms

A physically-based retrieval algorithm is a procedure that minimizes the difference between the measured-channel radiances and radiances calculated from an estimate of the geophysical state (temperature, moisture, etc.). Susskind *et al.*, [13] demonstrate that physically-based retrievals, using a carefully selected subset of the channels, can significantly improve upon a regression-based first guess for the AIRS/AMSU system. In general, the physically-based retrieval process is highly nonlinear, and convergence may require many iterations.

If the Planck function $B_\nu(T)$ does not change appreciably over the width of the CSRF, the monochromatic radiative-transfer equation can be represented by a channel-averaged radiative-transfer equation (e.g., see McMillin *et al.* [6]), in

which a channel-averaged transmittance to space $\tau_i(z)$ for channel i is computed by

$$\tau_i(z) \equiv \frac{\int F(\nu)\phi(\nu, \nu_i)\tau_\nu(z)\,d\nu}{\int F(\nu)\phi(\nu, \nu_i)\,d\nu} \tag{6}$$

where $\tau_\nu(z)$ is the monochromatic transmittance from a level at altitude z. An interferometer has optical filters, electronic band-pass filters, and a finite-detector response, which all limit the extent over which the integral must be computed. $F(\nu)$ is the effective filter response of the interferometer, filters, and detectors.

The integral over frequency is time consuming. For this reason, it is customary to construct a model (for each channel of the spectrometer) of the logarithm of atmospheric transmittance $\log(\tau_i(z))$ as a function of slant-path angle and the vertical distribution of temperature, moisture, ozone, and other gases, with precalculated coefficients based on monochromatic line-by-line calculations for an ensemble of profiles. The model, called a rapid-transmittance algorithm (RTA), can enhance the execution time of a retrieval algorithm by many orders of magnitude compared to integrating over monochromatic calculations for each profile.

Monochromatic atmospheric transmittance is a physical quantity that decreases monotonically with height and must have values between zero and one [i.e., $0 \leq -\log(\tau_\nu(z)) \leq \infty$]. State-of-the-art channel-averaged transmittance models (e.g., [8]–[11]) implicitly assume $\tau_i(z)$ is also a physical quantity with the same properties as the monochromatic transmittances. Therefore, current channel-transmittance models can only be used if the following is true.

1) The CSRF is sufficiently localized.
2) The CSRF's negative side-lobes are sufficiently small to ensure that the computed layer channel transmittance is always physical.

The unapodized CSRF does not satisfy either criterion for using channel transmittance. In fact, channel transmittance is not a meaningful concept for unapodized spectra both because the Planck function varies significantly over a CSRF and because channel transmittances can be nonphysical. McMillin *et al.* [6] develop a piecewise approach to model unapodized radiances as a weighted sum of radiances computed for a number of localized "channels." We will describe a different approach in terms of computed-Hamming apodized radiances, in Section V. In the next section, we review the Hamming function in this context.

IV. Cosine-Apodization Functions

Many apodization functions can be applied to an interferogram that will localize the CSRF and satisfy the RTA requirements in Section III. It is desirable to be able to compute radiances of any arbitrary symmetric apodization, including unapodized and self-apodized (if the self-apodization function is symmetric). We will show that a cosine-apodization function can be used for this purpose. The cosine-apodization function is given by

$$A_C(\delta) = \left\{ \begin{array}{ll} (1-2a) + 2a \cdot \cos\left(\pi \dfrac{\delta}{L}\right), & \text{for } |\delta| \leq L \\ 0, & \text{for } |\delta| > L \end{array} \right\} \tag{7}$$

which has values of $A_C(\delta = 0) = 1$ and $A_C(\delta = L) = 1 - 4a$.

The CSRF is given by the cosine transform of (7) and is equal to

$$\phi_C(y, z) = \frac{\sin(y)}{y} \cdot \left[(1-2a) + \frac{2a \cdot y^2}{\pi^2 - y^2}\right] + \frac{\sin(z)}{z} \cdot \left[(1-2a) + \frac{2a \cdot z^2}{\pi^2 - z^2}\right]. \tag{8}$$

Apodized radiances can be shown to be equivalent to a three-point running mean of unapodized radiances if the Nyquist channel spacing of $\delta\nu = 1/(2L)$ is used

$$R_C(i) = (1-2a) \cdot R_U(i) + a\left[R_U(i-1) + R_U(i+1)\right]. \tag{9}$$

Hamming [14] found an optimum value of a, which minimized the first side-lobe of the CSRF. He also showed that the optimum value of a was a function of the number of points in the spectrum. However, the optimum value of a converged to 0.23 for more than 100 points. In Fig. 1, the height of the maximum side-lobe relative to the height of the central lobe is shown as a function of a for a large number of points. The number of the side-lobe with the maximum height is shown along the top of the figure, and at $a = 0.23$ it can be seen that all side-lobes are suppressed to less than 1%, with the fourth side-lobe being the largest.

There are indeterminate points in the CSRF at $y = \pm\pi$, for which $\phi_C(y \to \pm\pi, z) \to 2a/(1-2a)$. Zeros of the CSRF occur at $y = \pm n\pi$. However, n starts with the second zero ($n = 1$ is the indeterminate point). An additional set of zeroes exists at $y = \pm\pi\sqrt{(1-2a)/(1-4a)}$. The ratio of the FWHM of the CSRF with respect to FWHM_U increases with increasing values of the parameter a, as shown in Fig. 2. When $a = 0.23$, corresponding to the Hamming function, the FWHM_H is 50.4% larger than FWHM_U.

The Hamming function is a reasonable function to use for a fast transmittance model because 1) the side-lobes are suppressed to less than 1% and hence, channel-averaged transmittances should behave in a physical way, and 2) the FWHM is relatively small. If one instead chose $a = 0.20$, the FWHM would be only 40% larger than the unapodized value, but the largest side-lobe would be larger by about a factor of three compared to the Hamming CSRF.

One consequence of apodization is that the noise becomes correlated among neighboring channels. This is discussed in more detail in Section VI, (31) and (32). For $a = 0.23$, the three-point running filter (0.23, 0.54, 0.23) reduces the noise in a given channel by a factor of 1.6. However, noise in adjacent channels is correlated by 62.5%, and noise in channels separated by $\Delta\nu = 1/L$ is correlated by 13.3%. This must be accounted for if the data are to be used optimally in the retrieval process. This complication is one factor that has led to the consideration of use of unapodized spectra in the retrieval process [5].

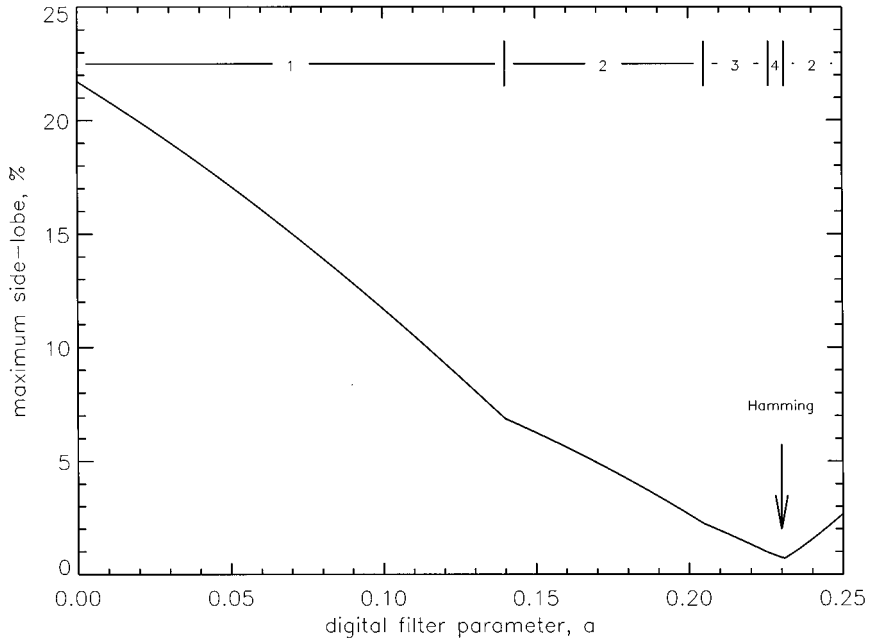

Fig. 1. Ratio of largest side-lobe's height to the height of the central lobe for cosine apodization CSRF's versus the digital-filter parameter a. The Hamming CSRF ($a = 0.23$) has side-lobes that are suppressed to below 1%. Along the top of the figure is the number of the largest side-lobe. For Hamming, the fourth side-lobe is the largest.

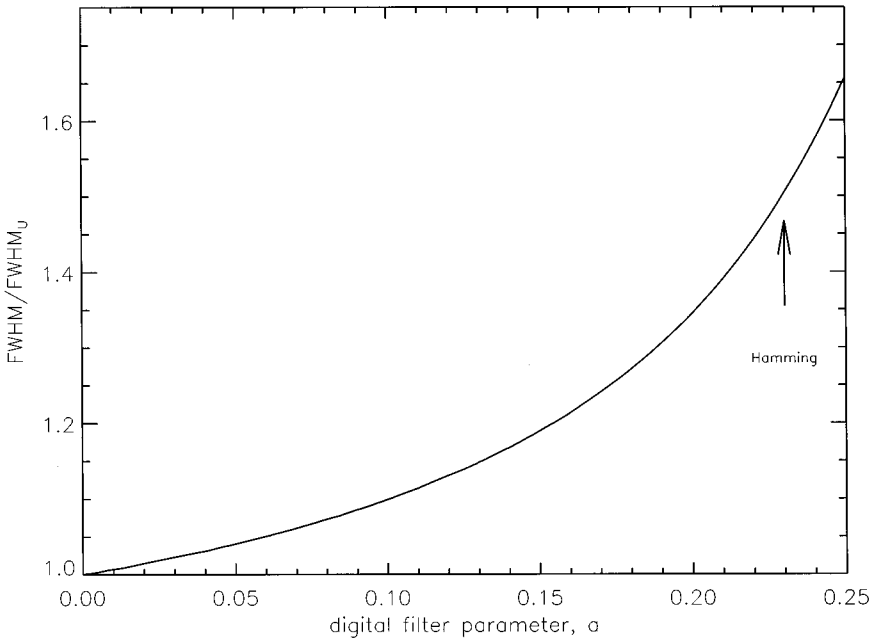

Fig. 2. Ratio of the FWHM of the cosine-apodized CSRF to the FWHM$_U$ versus the digital-filter parameter a. As a increases, the FWHM increases. The Hamming function's FWHM is 50% larger than the FWHM$_U$ of the $\mathrm{sinc}(y)$ function.

In the upper panel of Fig. 3, we show the CSRF's for $\mathrm{sinc}(y)$, Hamming, and a Gaussian with FWHM = FWHM$_H$. In the lower panel of Fig. 3, we show the apodization functions corresponding to the CSRF's in the upper panel. The cosine transform of a Gaussian is a Gaussian with a FWHM comparable to that of the Hamming apodization function. In Fig. 3, we also show the cosine transform of a typical terrestrial spectrum for a bandpass of 600 to 1200 wavenumbers. In comparing apodization functions, the unapodized and apodized have the same coverage in optical delay, but the apodization function weights the large optical delay values less. This might appear to be a loss in information content. However, the next section will show that information content only depends on the value of L, not on the apodization function. The Gaussian apodization function has similar qualities to the Hamming function but has significant values extending to larger L. This implies that more information content exists in the Gaussian CSRF, and that the small side-lobes of the Hamming CSRF actually degrade the resolution compared to the Gaussian function, so that the effective resolution of the interferometer is actually slightly poorer than FWHM$_H$.

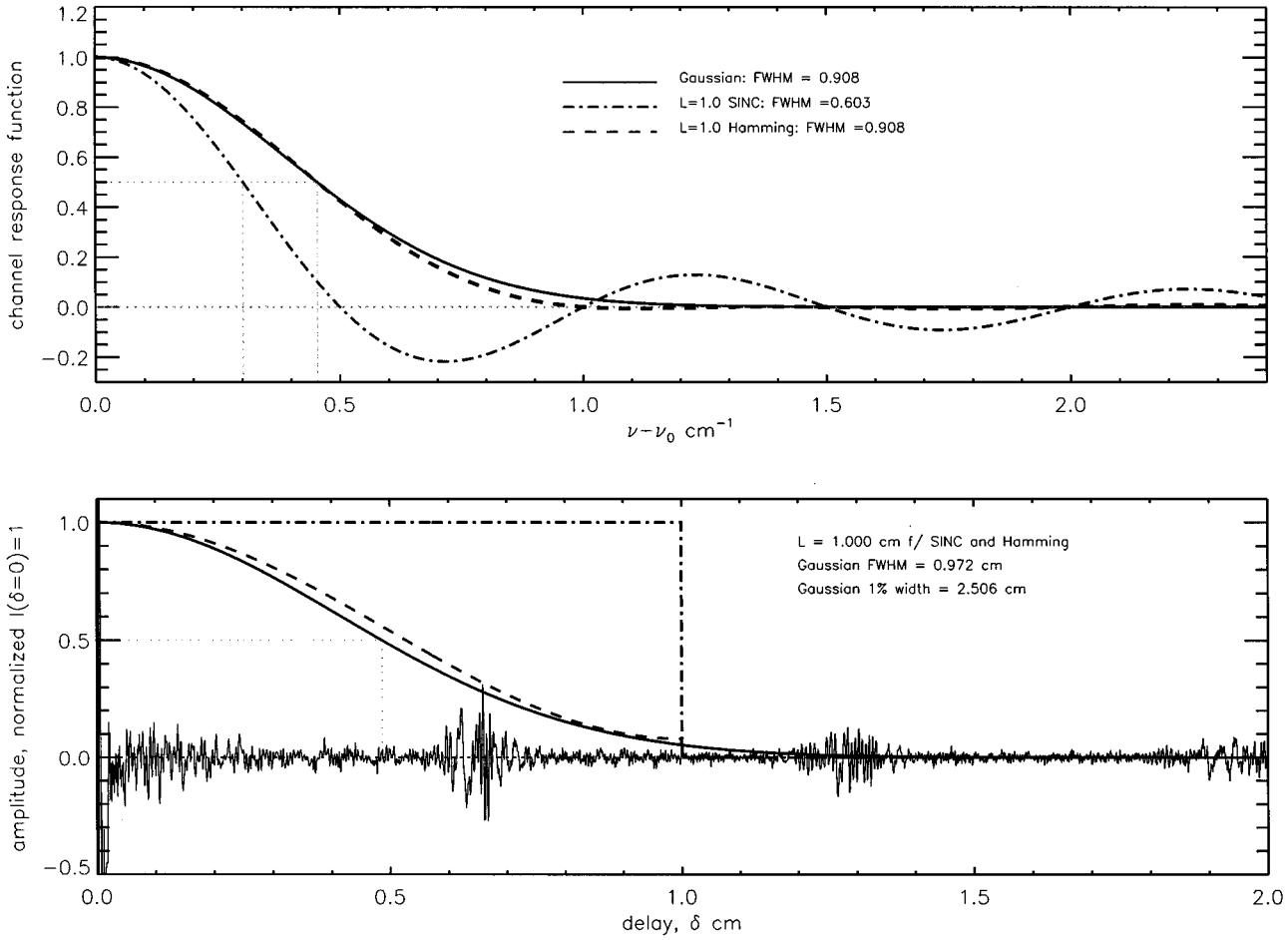

Fig. 3. The upper panel shows three CSRF's: the Gaussian function with the same FWHM ($0.908/L$) as the Hamming CSRF, the Hamming function, and the $\mathrm{sinc}(y)$ function for $L = 1$ cm. Since these functions are symmetric, only the positive side is shown. The cosine transforms of these CSRF's are shown in the lower panel. The $\mathrm{sinc}(y)$ function transforms to a boxcar function of width $\pm L$. The Hamming CSRF transforms to an apodization function equal to $0.54 + 0.46 \cdot \cos(\pi\delta/L)$. The Gaussian CSRF transforms to a Gaussian-apodization function. Also shown is the cosine transform of a typical midlatitude terrestrial spectrum for a bandpass of 600 to 1200 wavenumbers. The structures at $\delta = 0.65$ and $\delta = 1.3$ are due to the strong CO_2 absorption in this bandpass. Water vapor and ozone lines also contribute to the information content within this interferogram.

V. Equivalence Between Cosine-Apodized and Unapodized Radiances

The cosine-apodized spectrum $\vec{R}_C = R_C(i)$ is related to the unapodized spectrum according to (9) and can be expressed in matrix form

$$\vec{R}_C = \mathbf{M} \cdot \vec{R}_U. \tag{10}$$

The cosine filter given by (9) and (10) can be represented as a tridiagonal matrix

$$\mathbf{M} = (1 - 2a) \cdot \mathbf{J} \tag{11}$$

where $\mathbf{J}$ has the form

$$\mathbf{J} = \begin{pmatrix} 1 & b & 0 & 0 & 0 & \cdots \\ b & 1 & b & 0 & 0 & \cdots \\ 0 & b & 1 & b & 0 & \cdots \\ 0 & 0 & b & 1 & b & \cdots \\ 0 & 0 & 0 & b & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \tag{12}$$

and where b has the value $a/(1 - 2a)$. We show that this matrix can be inverted to provide a unique and well-behaved transformation from apodized radiances to unapodized radiances

$$\vec{R}_U = \mathbf{M^{-1}} \cdot \vec{R}_C \tag{13}$$

where $\mathbf{M^{-1}}$ is determined strictly from b, provided that $b < 0.5$ (i.e., $a < 0.25$).

For a small number of channels (fewer than about 50), the matrix $\mathbf{J}$ is readily invertible by numerical means. However, for a large number of channels, the inverse, calculated by standard numerical techniques, may be unreliable due to accumulated computational roundoff and a determinant approaching machine zero. Hence, an analytic inverse of the matrix is desirable.

The inverse of a matrix (if it exists) can be written as the classical adjoint of the matrix divided by the determinant of the matrix. The adjoint is the transpose of the matrix of cofactors, where the cofactor of a particular element is the determinant of the matrix composed of the rows and columns not containing that particular element. For this class of matrices, as the dimension increases, we rapidly approach values that are computationally 0/0. However, the limits exist and can be found.

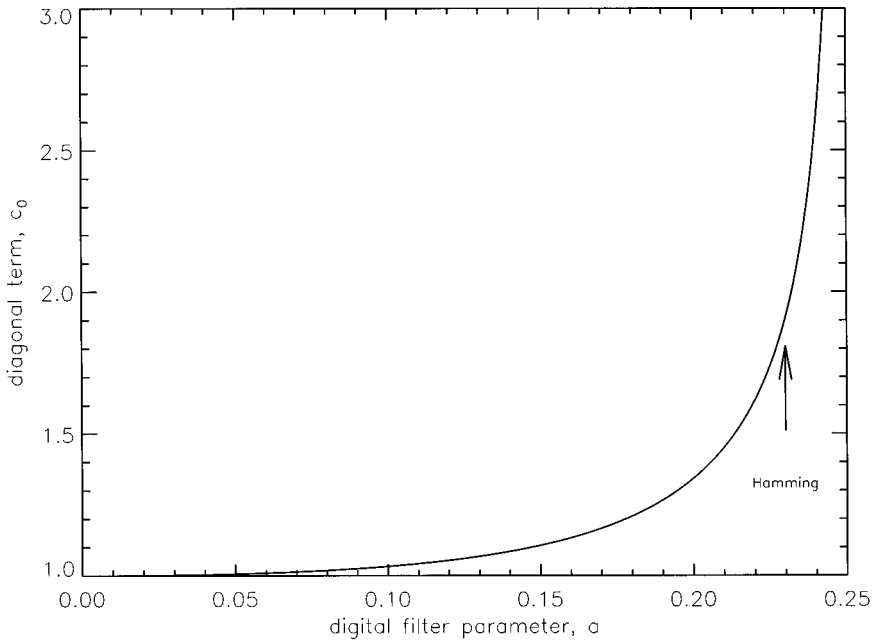

Fig. 4. Value of the diagonal matrix element c_0 of the inverse-transformation matrix $\mathbf{M}^{-1}$ versus the digital-filter parameter a. This curve was generated using (19).

Let us denote the determinant of a tridiagonal matrix of the form $\mathbf{J}$ and of dimension n as D_n. It can be shown that D_n for any n can be expressed by

$$D_n = \sum_{k=0}^{[n/2]} (-1)^k \frac{(n-k)!}{k!(n-2k)!}\, b^{2k} \tag{14}$$

where the brackets indicate the largest integer less than or equal to $n/2$. The first few terms of this series are

$$D_n = 1 - (n-1)b^2 + \frac{n^2-5n+6}{2}\, b^4 - \frac{n^3-12n^2+47n-60}{6}\, b^6 + \cdots. \tag{15}$$

The pth diagonal matrix element of $\mathbf{J}^{-1}$ is given by the cofactor of the pth diagonal element of a matrix of dimension n divided by the determinant of the matrix

$$\mathbf{J}^{-1}_{p,p} = \frac{D_{p-1}D_{n-p}}{D_n}. \tag{16}$$

The first off-diagonal matrix element of $\mathbf{J}^{-1}$ is given by the cofactor of that element divided by the determinant

$$\mathbf{J}^{-1}_{p,p+1} = -b\,\frac{D_{p-1}D_{n-p-1}}{D_n} \tag{17}$$

and, in general, any element of the matrix can be computed by

$$\mathbf{J}^{-1}_{p,p+q} = (-b)^q\,\frac{D_{p-1}D_{n-p-q}}{D_n} \tag{18}$$

where q is the distance from the diagonal. The matrix elements given by (18) depend on the dimension of the matrix, given by the number of channels. Equation (18) is exact but cumbersome to evaluate exactly for a large number of channels.

An excellent approximation can be made by taking the limit for a large number of channels. For the limit of the central portion of a matrix with a large dimension $2m+1$ corresponding to $2m+1$ channels, (16)–(18) converge to expressions that are only functions of b. The central diagonal element, defined here as c_0, can be expressed as

$$c_0 = \lim_{m\to\infty} \frac{D_m D_m}{D_{2m+1}} \tag{19}$$

which is equal to

$$c_0 = \sum_{k=0}^{\infty} \frac{(2k)!}{(k!)^2}\, b^{2k}. \tag{20}$$

The dependence of c_0 on a is shown in Fig. 4. For a well-behaved inverse, the matrix elements should not be too large. The inverse becomes ill-conditioned as a approaches 0.25, in which case $A_C(L) = 0$. In the Hamming case ($a = 0.23$), c_0 converges (with more than 45 channels) to within 10^{-12} to a value

$$c_0 = 1.909188309204\cdots. \tag{21}$$

The ratio of the first off-diagonal matrix element of $\mathbf{J}^{-1}$ to the diagonal matrix element of $\mathbf{J}^{-1}$, as well as the ratio of the $(n+1)$th off-diagonal term to the nth off-diagonal term, is given by $r = -bD_{m-1}/D_m$. In the large number of channels limit $m \to \infty$, this ratio converges to a constant, and (15) can be used in conjunction with direct polynomial division to get

$$r = 1 - b^2 - b^4 - 2b^6 - 5b^8 - 14b^{10} - 42b^{12} - 132b^{14} - \cdots. \tag{22}$$

The dependence of r on a is shown in Fig. 5. For $a = 0.23$, (22) converges to within 10^{-12} for $m > 45$ to a value of

$$r = -0.5590375815769\cdots. \tag{23}$$

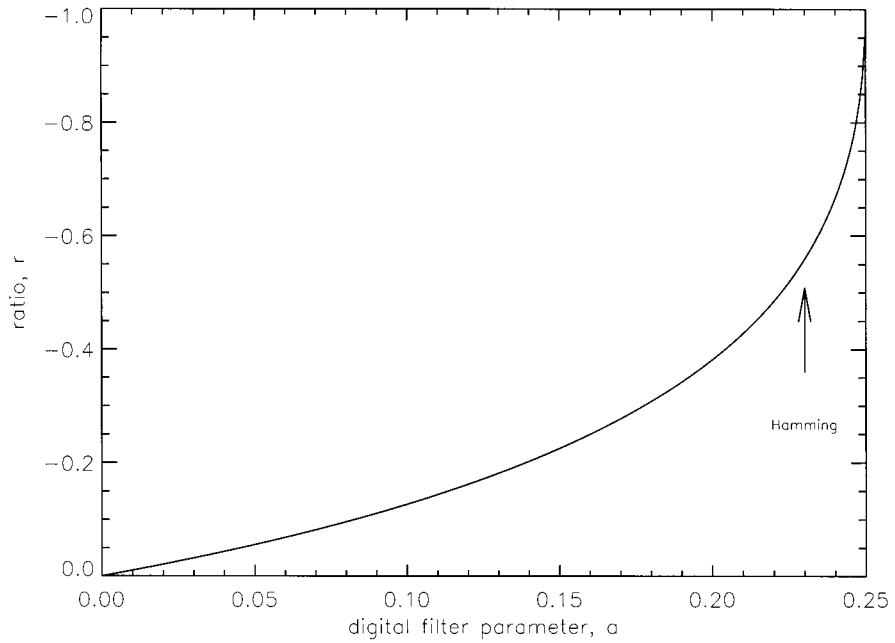

Fig. 5. Value of the matrix-element ratio r of the inverse-transformation matrix $\mathbf{M}^{-1}$ versus the digital-filter parameter a. This curve was generated using (22).

This value of r makes the matrix localized to the diagonal, falling off by roughly a factor of two per channel and alternating in sign. For high-precision calculations, all terms of the matrix are important.

Using (12)–(22), one can now express unapodized channel radiances in terms of a linear combination of cosine-apodized radiances if all the channels within a band are used. Each line of the inverse transformation can be expressed as

$$R_U(i) = \frac{c_0}{(1-2a)} \cdot \left[R_C(i) + \sum_{j=1}^{N} r^j \cdot (R_C(i+j) + R_C(i-j)) \right] \quad (24)$$

where all the parameters can be expressed in terms of a. In (24), radiances in out of band channels are defined as zero, $R_C(i \leq 0) \equiv 0$, and $R_C(i > N) \equiv 0$. As a numerical test of (24), we have computed both Hamming-apodized radiances and unapodized radiances using appropriate convolutions of monochromatic radiances. We also computed unapodized radiances from Hamming-apodized radiances using (24) with appropriate values of a, c_0, and r, and the results (including noise) agreed to the machine accuracy of one part in 10^7.

Equations (9) and (24) express the set of unapodized radiances and cosine-apodized radiances as linear combinations of each other. Hence, the information content of the complete set of radiances, including noise effects, is identical. This implies that retrieval results using unapodized radiances, or cosine apodized radiances with $0 < a < 0.25$ should produce identical results, provided all the channels are used. Furthermore, it implies that while the width of the central lobe of the CSRF associated with an apodized interferogram is larger than the CSRF associated with an unapodized interferogram, there is no loss in "resolution" as both sets of radiances contain equivalent information. Equation (5) has been used to describe the effective spectral resolution of an interferometer (e.g., see [15]). Such a definition can be used to compare the resolution of one interferometer to another with a different value of L. This definition is misleading, however, if one attempts to compare the spectral resolution of an interferometer to that of an instrument for which the CSRF does not have side-lobes.

Equation (24) allows one to compute unapodized radiances for a given channel in terms of apodized radiances, which are easier to model in terms of an RTA. For greater than 50 channels, the approximation will be highly accurate and much better than measurement uncertainties, except for a few channels near the edge of the band. The error at the edge of the matrix can be minimized by calculation of about 21 channel radiances beyond the instrument's band. This requires calculating apodized radiances for channels that will not be used. However, this is still much simpler than any other highly accurate method for computation of unapodized radiances of which we are aware.

For an operational instrument (18) divided by $(1-2a)$ can be used to precompute $\mathbf{M}^{-1}$ exactly for the specific number of channels. The conversion from apodized radiances to unapodized radiances would involve a single matrix multiply. This operator is identical to taking the inverse-cosine transform of the computed apodized radiances, dividing the result by the apodization function, and taking the cosine transform of that result to produce unapodized radiances. In some applications, the matrix method would be simpler and faster.

An alternative approach to McMillin *et al.* [6] requires five radiance calculations for each channel in the band to generate accurate unapodized radiances. This approach could be complex and time consuming operationally, especially if only a subset of the unapodized radiances will be used.

The one caveat is that the Hamming channel radiances have to be calculated with sufficient accuracy. McMillin *et al.* [6] have shown that Hamming radiances, calculated using the approximation in (6), agree with Hamming radiances calculated with

channel-averaged monochromatic transmittances to within one part in 10^5. Applying our analytical-matrix method to compute unapodized radiances from the computed Hamming radiances magnifies this error by a factor of two. This would introduce an insignificant error into computed instrumental radiances for an instrument with an SNR on the order of 500.

VI. Higher-Order Cosine Apodization Functions

Use of apodization functions other than simple cosine apodization functions may be optimal for different applications. All continuous symmetric apodization functions can be expanded into a cosine series, which can be used to express radiances of these apodization functions as linear combinations of unapodized radiances. This linear operator is a powerful tool to compare the properties of apodization functions with different mathematical forms. In this section, we will develop the cosine expansion and show its utility. In addition, we will show that the cosine expansion can be used to compute apodized radiances directly from Hamming-apodized radiances, since the Hamming-apodization function has a well-behaved inverse.

The general form of the cosine expansion with J terms can be written as an apodization function of the form

$$A_G(\delta) = a_0 + 2 \cdot \sum_{j=1}^{J-1} a_j \cdot \cos\left(j \cdot \pi \frac{\delta}{L}\right), \quad \text{for } |\delta| \le L \quad (25)$$

where the first coefficient is constrained so that $A(\delta = 0) = 1$ and

$$a_0 \equiv \left(1 - 2 \sum_{j=1}^{J-1} a_j\right). \quad (26)$$

The CSRF is equal to the cosine transform of (25) and can be written generally as

$$\phi_G(y, z) = \frac{\sin(y)}{y} \cdot \left[a_0 + \sum_{j=1}^{J-1} 2a_j \cdot \frac{(-1)^j \cdot y^2}{y^2 - (j\pi)^2}\right] + \frac{\sin(z)}{z} \cdot \left[a_0 + \sum_{j=1}^{J-1} 2a_j \cdot \frac{(-1)^j \cdot z^2}{z^2 - (j\pi)^2}\right]. \quad (27)$$

These functions are normalized such that $\phi_G(y = 0) = 1$. It can be easily shown that for apodized spectra sampled at the Nyquist spacing of $\delta y = \pi$, the radiances are a linear combination of the unapodized radiances

$$R_G(i) = a_0 \cdot R_U(i) + \sum_{j=1}^{J-1} a_j \cdot (R_U(i-j) + R_U(i+j)) \quad (28)$$

which can be expressed in matrix form as

$$R_G = \mathbf{M_G} \cdot R_U \quad (29)$$

where $\mathbf{M_G}_{i,\, i \pm j} = a_j$.

The fact that the apodized radiances are a linear combination of the unapodized radiances allows calculation of the noise correlation in adjacent apodized channels. Equation (29) represents a $2J-1$ point running mean with weights, w_k, given by

$$w_k = a_{|k|}, \qquad \text{for} \quad k = -(J-1), J-1. \quad (30)$$

For the Hamming-apodization function ($J = 2, a_0 = 0.54, a_1 = 0.23$), the three running-mean coefficients are $w_{-1} = 0.23, w_0 = 0.54$, and $w_1 = 0.23$. In general, the random channel noise is reduced by a factor of

$$f = \left[\sum_{k=1-J}^{J-1} w_k^2\right]^{-(1/2)} \quad (31)$$

and it can be shown that the channel noise correlation for channels separated by $\pm n$ Nyquist spacings is given by

$$C_n = f^2 \sum_{k=1-J}^{J-1-n} w_k \cdot w_{k+n}, \qquad \text{for} \quad n = 1, 2J-2. \quad (32)$$

The noise-reduction factor and the noise-correlation matrix for apodized radiances are important properties for remote-sensing applications. The cosine transform of apodization functions is useful to compute the noise-correlation matrix for apodization functions of different mathematical forms. The equations for noise correlation and noise reduction have been verified numerically on ensembles of random-noise spectra for a number of commonly used apodization functions to be discussed in Section VIII.

The Hamming-filter function has the property that it is a good apodization function for the purpose of a fast transmittance model, and it also has a simple inverse with well-behaved matrix elements. The fact that it is a three-point running mean minimizes the effect of channel-correlated noise on the retrieval process, since only noise in channels ±2 Nyquist spacings is correlated. Equations (13) and (29) can be used to transform Hamming radiances R_H to radiances of any apodization function

$$R_G = \mathbf{M_G} \cdot \mathbf{M_H^{-1}} \cdot R_H \quad (33)$$

where $\mathbf{M_H^{-1}}$ is the cosine inverse matrix discussed in Section V with $a = 0.23$. Therefore, an RTA for Hamming-apodized channel transmittances is useful for calculation of radiances for any apodization function, including instrumental self-apodization.

If $\mathbf{M_G^{-1}}$ exists, we can also write

$$R_U = \mathbf{M_G^{-1}} \cdot R_G. \quad (34)$$

This implies that for any well-behaved apodization function such that $\mathbf{M_G^{-1}}$ exists, the set of apodized radiances R_G can be used to compute radiances of any apodization function. A necessary condition for $\mathbf{M_G^{-1}}$ to exist is that $A_G(\delta)$ be positive for all δ. We have demonstrated numerically that $\mathbf{M_G^{-1}}$ exists for a number of apodization functions. We have not been able to construct an analytic expression for $\mathbf{M_G^{-1}}$ when $J > 2$, and in general, the elements of $\mathbf{M_G^{-1}}$ are much larger than those of $\mathbf{M_H^{-1}}$. Large values of $\mathbf{M_G^{-1}}$ amplify the effects of errors in

R_G and are therefore undesirable. In general, this makes A_G less practical than A_H for use in computing radiances for other apodization functions.

VII. Use of the Matrix Operator in Remote Sensing

In this section, we discuss a number of issues relevant to remote-sensing applications in which the transformation matrices $\mathbf{M}$ and $\mathbf{M}^{-1}$ play a significant role.

A. Transformation of Instrument Noise

An important realization for the transformation of radiances with noise is that the noise-free radiance can be transformed separately from the noise spectrum. We can consider that an unapodized measurement consists of a noise-free component $R_U(i)$ with a specific noise $\delta R_U(i)$ added to the noise-free component. Since the apodized radiance R_G is derived from the unapodized radiance

$$R_G(i) = \mathbf{M_G} \cdot R_U(i) \tag{35}$$

then it can be shown that the apodized noise is derived with the identical transform

$$\delta R_G(i) = \mathbf{M_G} \cdot \delta R_U(i). \tag{36}$$

In this way, it is obvious that the effect of apodization is identical for the radiance and the noise, and that the information content within a complete set of channels is unaffected by the apodization process. Choice of apodization may make a difference in results if only a subset of channels are used to analyze the data, or if a nonlinear combination of channels is used. In computing radiances to test the effect of apodization on retrieval accuracy, we assume that the unapodized channel noise is random and uncorrelated and the signal-to-noise values are appropriate for the instrumental characteristics.

For a given unapodized noise spectrum $\delta R_U(i)$ the apodized noise spectrum is computed exactly using (36). We use this linear-reversible transform in Section VII-B to show that regression retrievals are insensitive to the apodization used.

In the physical retrieval process, we only have knowledge of the statistical noise estimate [usually given as the noise-equivalent difference radiance $NE\Delta N(i)$, and we account for the noise reduction and correlation of apodized channel noise using (31) and (32). This will be illustrated in Section VII-C.

B. Equivalence of Linear-Regression Retrievals

The equivalence of apodized and unapodized retrieval results has been discussed previously [16], [17] in a noise-free context. The current formulation also addresses the role of noise on the equivalence. The linear operators developed above can be used to understand the effects of apodization on atmospheric retrieval algorithms with noisy radiances. We define the apodized noisy radiance $\tilde{R}_G(i, n)$ to be composed of the noise-free apodized radiance $R_G(i, n)$ plus an apodized noise spectrum $\delta R_G(i, n)$ that is both a function of channel i and case n

$$\tilde{R}_G(i, n) \equiv R_G(i, n) + \delta R_G(i, n). \tag{37}$$

A regression-retrieval algorithm determines the statistical relationship between radiances and geophysical parameters (e.g., vertical temperature profile, vertical moisture profile, vertical ozone profile, surface parameters, etc.). In the regression-retrieval methodology, a training ensemble is used to determine a regression matrix. The radiances $\tilde{R}_G(i, n)$ for channel i and case n are usually computed using the complete nonlinear radiative transfer using an apodized RTA from the geophysical state $X(k, n)$ where k is an index for a specific geophysical component of X and then a case-specific noise $\delta R_G(i, n)$ is added .

In practice, the training ensemble is a set of radiance measurements $\hat{R}(i, n)$ and colocated measurements of the geophysical state $X(k, n)$ (e.g., radiosondes, lidar, etc.). In this case, $\tilde{R}_G(i, n)$ represents the measured apodized radiances, and $\delta R_G(i, n)$ is the case-specific noise.

It is common to define an average geophysical condition $\overline{X}(k) \equiv \sum_{n=1}^{N} X(k, n)/N$ and compute the mean radiance of an ensemble of N profiles $\overline{R}(i) = \sum_{n=1}^{N} \tilde{R}(i, n)/N$ and write the regression relationship for a set of K geophysical parameters and apodized radiances as

$$\begin{aligned} X(k, n) - \overline{X}(k) &= W_G(k, i) \cdot \left(\tilde{R}_G(i, n) - \overline{R_G}(i)\right) \\ &= W_G(k, i) \cdot \Delta R_G(i, n). \end{aligned} \tag{38}$$

The regression matrix is found by solving (38) for W_G using the training ensemble of cases

$$\begin{aligned} W_G(k, i) = &\left(X(k, n) - \overline{X}(k)\right) \cdot \sum_j \Delta R'_G(j, n) \\ &\cdot \left(\sum_m \Delta R_G(j, m) \cdot \Delta R'_G(m, i)\right)^{-1}. \end{aligned} \tag{39}$$

We have used j and m to represent indices in summation computations for channel and case, respectively. The regression matrix is used to solve for the geophysical state that best matches a set of independently measured radiances $\hat{R}(i)$. For an apodized-regression matrix, the regression-retrieval equation is given by

$$\hat{X}_G(k) = \overline{X}(k) + \sum_i W_G(k, i) \cdot \left(\hat{R}_G(i) - \overline{R_G}(i)\right). \tag{40}$$

A similar development could have been made for unapodized radiances. It is easy to show that the regression matrix for unapodized, noise-free radiances is related to the regression matrix for apodized, noise-free radiances according to

$$W_U(k, i) = \sum_j W_G(k, j) \cdot \mathbf{M_G}(j, i) \tag{41}$$

and the resulting unapodized-regression retrieval for noisy radiances is given by

$$\hat{X}_U(k) = \overline{X}(k) + \sum_i W_U(k, i) \cdot \left(\hat{R}_U(i) - \overline{R_U}(i)\right). \tag{42}$$

The measured-unapodized radiances with noise are related to the apodized radiances with noise via $\mathbf{M_G^{-1}}$. Therefore, the results of a regression retrieval are insensitive to apodization

$$\hat{X}_G(k) = \hat{X}_U(k) \tag{43}$$

provided all the channels are used in the retrieval process.

C. Transformation of Derivatives and Noise Covariance Matrices

In physical-retrieval algorithms, the matrix of partial derivatives of the radiance with respect to the lth geophysical parameter being solved for $X(l)$, called the Jacobian

$$S(i, l) = \frac{\partial R(i)}{\partial X(l)} \tag{44}$$

must be calculated and iterated [13]. In addition, the optimal channel-noise covariance matrix of a physical retrieval should include uncertainties in the calculated radiances $E(i, k)$ due to uncertainties in the kth parameter $\delta X(k)$ which is held constant (i.e., not solved for) in a given step of the process. These are proportional to the product of the sensitivity of the channel radiance to changes in a given parameter not solved for and the uncertainty in that parameter

$$E(i, k) = \frac{\partial R(i)}{\partial X(k)} \delta X(k). \tag{45}$$

The noise covariance $N_{i,j}$ is then constructed from the instrumental noise-covariance function $\tilde{N}_{i,j}$ and the computational noise estimates as follows

$$N_{i,j} = \tilde{N}_{i,j} + \sum_{k=1}^{K} E(i, k) \cdot E(j, k). \tag{46}$$

For unapodized radiances, the instrumental noise-covariance matrix is a diagonal matrix (assuming there are no other correlated noise sources within the instrument), while for apodized radiance, the matrix can be constructed from the instrumental noise equivalent delta radiance $NE\Delta N(i)$ and the noise-correlation terms given in (31) and (32)

$$\tilde{N}_{i,j} = (f \cdot NE\Delta N(i)) \cdot C_{i-j} \cdot (f \cdot NE\Delta N(j)). \tag{47}$$

Susskind *et al.* [13] take advantage of the narrow highly localized channels of AIRS to select a subset of channels that simultaneously have large sensitivity to parameters being solved for, $S(i, l)$, and low sensitivity to other parameters, $E(i, k)$.

S_H, E_H, and R_H for Hamming apodized spectra can all be calculated in a straightforward manner using the channel-averaged transmittance approach [13]. These can then be transformed to appropriate values for spectra using any apodization (R_G) or unapodized spectra (R_U), according to

$$S_G(i, l) = \sum_j \left[\mathbf{M_G}\mathbf{M_H^{-1}}\right](i, j) \cdot S_H(j, l) \tag{48}$$

and

$$E_G(i, k) = \sum_j \left[\mathbf{M_G}\mathbf{M_H^{-1}}\right](i, j) \cdot E_H(j, k) \tag{49}$$

where $\mathbf{M_G}$ would be a unit matrix for unapodized derivatives. Therefore, even if one wants to use a different apodization, the highly localized Hamming apodization is an important intermediate tool for use in computing the necessary terms for use in a physical-retrieval algorithm. If an apodization function is both a good function to use for the RTA and also has a well-behaved numerical inverse matrix, then it can be used instead of the Hamming for the computation of the R_G, S_G, and C_G in (48) and (49).

There may be a particular advantage to an approach using either apodized or unapodized radiances in performing physical retrievals if either 1) results improve when a subset of channels is used in the retrieval, or 2) results do not change appreciably if only a subset of channels is used. In case 1), use of apodized or unapodized radiances may prove beneficial if better results can be obtained with a subset of channels using one or the other approach. In case 2), there will be no change in accuracy, but an improvement in computational efficiency if one approach produces an accurate solution using fewer channels than the other. Equations (48) and (49) give an easy way to compute the necessary terms for a spectrum of any apodization function, including unapodized. However, in this regard, it should be realized that calculations using apodized radiances need only be done for the subset of apodized radiances used in the retrieval process, if the same apodization is used for the RTA. On the other hand, for retrievals using differently apodized (or unapodized) radiances (44)–(49) must be computed using all the channel radiances.

VIII. Comparison of Characteristics of Different Apodization Functions

We will discuss the characteristics of apodization functions in the literature and compare their attributes to those of the Hamming-apodization function. We will utilize the cosine transformation of the apodization function to provide a quantitative method to assess the utility of an apodization function of any mathematical form. In Section VIII-A, we will consider these apodization functions for use as a basis for constructing apodized channel radiances, which can then be transformed to radiances using different apodization functions for their use in the retrieval process, analogous to (33). In Section VIII-B, we will discuss the functions for direct use in a retrieval.

A. Assessment of Apodization Functions for Use in Computing Other Apodization Functions

In this section, we will define the functions and discuss them in terms of transforming to other apodization functions. As discussed previously, for the purpose of modeling channel radiances, one wants a highly localized function with small sidelobes. To convert to other apodization functions, we require a well-behaved inverse relating the apodized radiances computed with the RTA to unapodized radiances.

In Fig. 6, the ratio of the area of the central lobe of the CSRF to the total area within the CSRF is shown as a function of the ratio of the FWHM to FWHM_U [defined by (5)] for a number of common apodization functions. The symbols and notations used in this figure will be discussed later. For RTA purposes, it is desirable to optimize the tradeoff between FWHM and spectral purity (ideally, one would like values toward the upper left corner of Fig. 6), as well as to keep the apodization function from becoming small at $A(L)$ (thus stabilizing the inverse matrix).

Any symmetric apodization function can be decomposed into a cosine expansion. In Table I, we summarize the first four cosine expansion terms for a number of apodization functions that

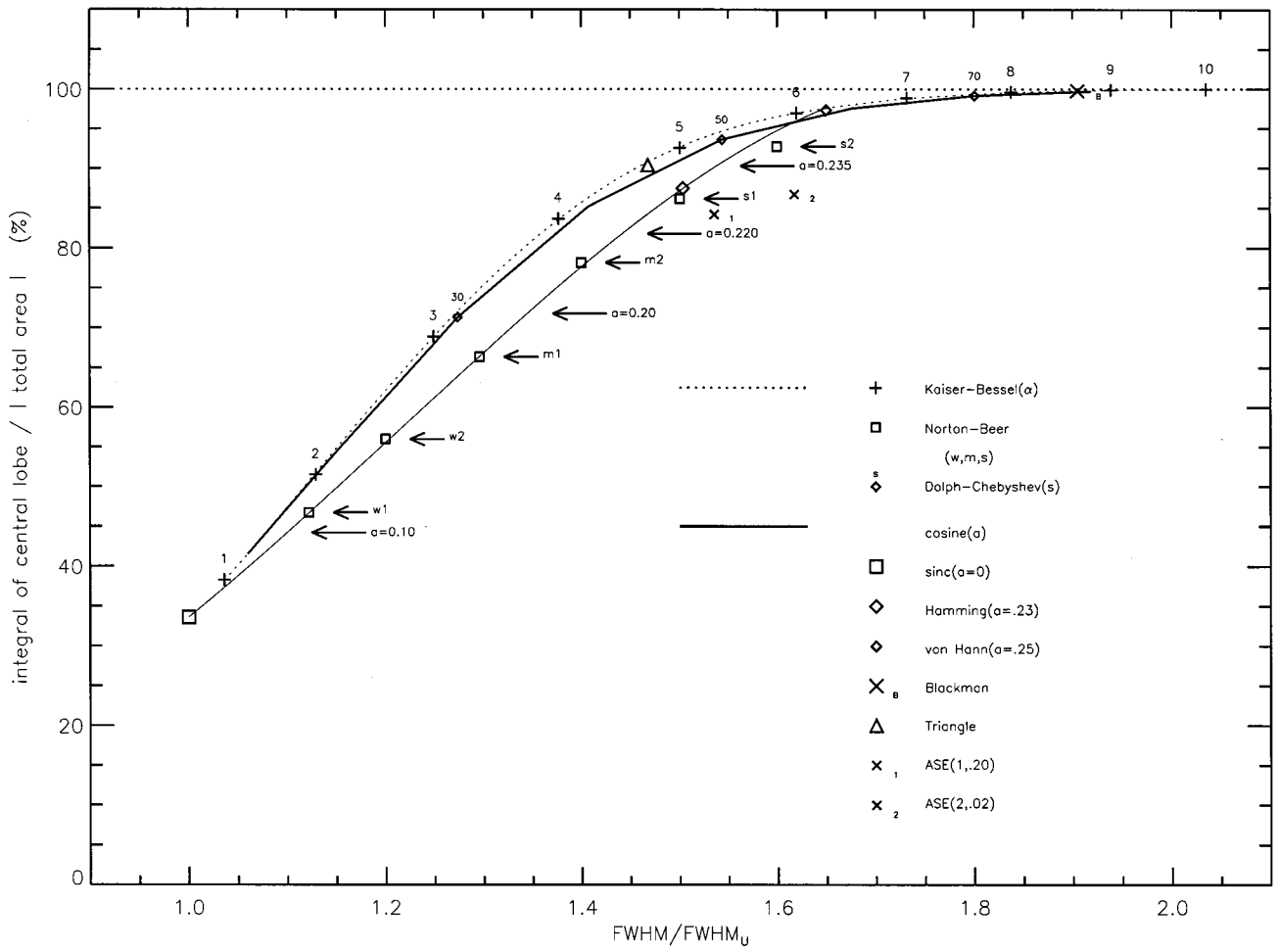

Fig. 6. Ratio of the area of the central lobe to the total area of the CSRF is shown in percent as a function of the ratio of the FWHM to FWHM$_U$ for various apodization functions used in remote-sensing applications. This is a measure of the effect of the side-lobes. See Section VIII for explanation of the symbols.

have been used in remote-sensing applications. These terms define the off-diagonal nature of the $\mathbf{M}$ matrix and help to compare apodization functions of various mathematical forms. We will discuss the functions in the order they are given in Table I.

The $\mathrm{sinc}(y)$ function, Hamming function ($a = 0.23$), and von Hann function ($a = 0.25$) are denoted with a box and large and small diamond symbols, respectively. These are all part of the family of $J = 2$ cosine-apodization functions, which are shown as functions of a by the solid line in Fig. 6.

The Blackman [18] function is a $J = 3$ cosine-apodization function with the coefficients shown in Table I. The Blackman function's side-lobes (shown as an X in Fig. 6) are significantly smaller than those of the Hamming function (shown as a diamond in Fig. 6), which results in a more localized response function. This function has a zero at $\delta = L$, which makes the inverse transformation ill-conditioned, and it cannot be used as a basis to generate radiances for other apodization functions.

A triangle-apodization function (sometimes referred to as a Bartlett-apodization function [14], [19] and shown as a triangle in Fig. 6) has a CSRF equal to $\mathrm{sinc}^2(y) + \mathrm{sinc}^2(z)$. The side-lobes are all positive, and they fall off much faster than for the Hamming apodization. These characteristics make this function a good function to use for the RTA. However, the fact that all the side-lobes are positive makes the far side-lobes more significant to the total radiance, since there is no cancellation. This function is nonlinear and has a zero at $\delta = L$. Therefore, there is no linear transformation between unapodized and apodized radiances, and it cannot be used to simulate radiances of other apodization functions.

The Norton–Beer [3], [20] apodization functions (denoted as "NB" in Table I) are given by power series

$$A_{\mathrm{NB}}(\delta) = \sum_{j=0}^{4} C(j) \cdot \left(1 - (\delta/L)^2\right)^j \tag{50}$$

with different sets of optimized coefficients $C(j)$, denoted as w, m, and s for their weak, medium, and strong apodization character. Different sets of coefficients were published by Norton and Beer in their first paper ("1") [3] and their second paper ("2") [20]. These functions, shown as small boxes in Fig. 6, lie very close to the solid line. The Norton–Beer "weak" and "medium" apodization functions are less desirable for the channel-transmittance calculations for high signal-to-noise instruments due to their larger side-lobes. The "strong" apodization function in the original paper [3] ("s1" in Fig. 6) has similar characteristics to the Hamming function. The higher-order terms in the cosine expansion of the Norton–Beer function are significant and this complicates the analytic form

of the matrix inversion. Even if an analytic inverse matrix could be found, there is no advantage to using these functions for computing radiances for other apodization functions, compared to the simple cosine apodization.

The Kaiser–Bessel apodization function [14], [21] (denoted as "KB" in Table I) is a modified Bessel function I_0 with an adjustable parameter α and is given by

$$A_{\mathrm{KB}}(\delta) = \frac{I_o\left(\alpha\sqrt{1-(\delta/L)^2}\right)}{I_0(\alpha)}. \tag{51}$$

The Kaiser–Bessel function is shown as a function of α by the dotted line in Fig. 6, with plus symbols showing the specific values corresponding to the values in Table I. In most cases, these functions are more localized than other apodization functions with the same FWHM (i.e., the side-lobes decrease more rapidly away from the central lobe). For $\alpha = 5$, the Kaiser–Bessel FWHM is approximately equal to FWHM_H and only the first few side-lobes are larger. For smaller α, the side-lobes are too large for use in a RTA, and for larger α values the inverse matrix tends to become unstable due to the small values at $A(L)$. No analytic form of the inverse matrix is known for these functions; however, a numerical inverse matrix can be found for a given number of channels. Therefore, the $\alpha \simeq 5$ Kaiser–Bessel function would be a very good function to use for conversion to other apodized radiances because the apodized function is more localized than the Hamming function and a numerical inverse can be found.

The Dolph–Chebyshev apodization function [12, p. 183], denoted as "DC" in Table I and shown as small diamond symbols in Fig. 6, is a function of the parameter s, which is the side-lobe suppression in decibels. The form of the DC apodization function is given in Ward [22]. This function is similar to the Kaiser–Bessel functions, in that the apodization is continuously tunable with the parameter s. The Dolph–Chebyshev apodization function has characteristics which are similar to the Kaiser–Bessel; the $s = 50$ dB function is similar to the Kaiser–Bessel $\alpha \simeq 5.2$ and the same conclusions hold for this function as well.

Finally, the autoregressive spectral estimator (ASE) apodization function of Amato *et al.* [4] is written in terms of two parameters, p and λ, which are found by minimizing radiance residuals with respect to a rectangular CSRF for an ensemble of spectra with a given signal-to-noise. The form of their equation is

$$A_{\mathrm{ASE}}(\delta) = \frac{1}{1+\lambda\cdot\left(\dfrac{2\pi\delta}{L}\right)^{2\cdot p}}. \tag{52}$$

Results are shown for two sets of (p, λ). The first function (1, 0.25) was deduced from the figures in [4] and the second function (2, 0.02) was deduced from the figures in [7]. The ASE apodization functions (denoted as x_1 and x_2 in Fig. 6) fall to the right of the cosine-apodization functions. This implies that they have no relative advantage in computation of the RTA. These functions also require many terms of the cosine expansion and are difficult matrices to invert. However, the inverse matrix can be found numerically. Amato *et al.* [7] used their ASE function to demonstrate the equivalence between unapodized and apodized retrievals, which can be understood in the context of Section VII-B since the numerical-inverse matrix exists.

In general, if the apodization function is weaker than Hamming, it is a poorer function to use for an RTA, due to the large side-lobes. The strong apodization functions are generally poorer candidates for the inverse transformation, because these functions all have a smaller value at maximum OPD ($A(L) \rightarrow 0$). Thus, the inverse becomes large or ill-conditioned. For many operational applications, the functions that are more localized, that have a FWHM similar to Hamming (i.e., the functions to the upper right in Fig. 6), and that also have well-behaved numerical-inverse matrices such as Kaiser–Bessel $\alpha \simeq 5$ and Dolph–Chebyshev $s \simeq 50$ dB should be considered. The Hamming function is the only function for which we have found an analytic inverse transformation, and it is a reasonable function to use for RTA's.

B. Assessment of Apodization Functions for Direct Use in Remote-Sensing Applications

In this section, we address the selection of an apodization function for the specific use of solving the inverse problem to retrieve geophysical parameters from the observed apodized radiances. In direct-retrieval applications, the accuracy of the RTA is still of utmost concern. However, the ability to convert the apodized radiances to another form of apodized radiances is unnecessary. The selection of an apodization function depends on two primary attributes of the retrieval process.

1) If the retrieval algorithm utilizes a subset of channels, then the spectral purity must be reasonable.
2) If the retrieval algorithm is sensitive to the channel-noise correlation, then the correlation should be localized and small.

In Table II, we summarize the noise-correlation attributes of the apodization functions shown in Table I. The fact that the instrument noise is correlated should not be of great concern, because many aspects of the retrieval noise covariance are already highly correlated (e.g., see Susskind *et al.* [13]).

The $J = 3$ cosine-apodization functions have more extensive channel-noise correlation than comparable $J = 2$ cosine-apodization functions. For example, the Blackman function's side-lobes are significantly smaller than those of the Hamming function, which results in a more localized response function. However, it has a significantly increased FWHM. When a subset of channels is used, broader channels lessen the ability to find desirable channels. Therefore, the Blackman function could substantially degrade the retrieval performance. In addition, the increased noise correlation could complicate the retrieval process and increase retrieval errors if not properly accounted for. Other $J = 3$ functions can be found that approximate the functions described below and therefore would have similar characteristics (e.g., using the first three coefficients listed in Table I).

As discussed earlier, the triangle-apodization function is a reasonable function to use for the RTA. However, the channel noise is highly correlated across many neighboring channels and must be accounted for properly. We do not recommend this function for retrieval applications.

TABLE I
COSINE-EXPANSION COEFFICIENTS FOR APODIZATION FUNCTIONS USED FOR REMOTE SENSING

function	a_0	a_1	a_2	a_3
Hamming	.540	.230	0	0
von Hann	.500	.250	0	0
Triangle	.508	.203	0	.023
Blackman	.420	.250	.04	0
NB w_1	.778	.115	-.004	-.0002
NB w_2	.701	.156	-.006	.0002
NB m_1	.634	.189	-.006	-.0008
NB m_2	.586	.214	-.008	-.00007
NB s_1	.534	.227	.006	.0002
NB s_2	.503	.239	.010	-.0002
KB α =1	.928	.043	-.010	.0045
KB α =2	.795	.119	-.024	.0101
KB α =3	.684	.176	-.026	.0108
KB α =4	.604	.211	-.018	.0081
KB α =5	.545	.231	-.006	.0046
KB α =6	.500	.241	.008	.0017
KB α =7	.465	.247	.021	$3{\cdot}10^{-5}$
KB α =8	.435	.249	.033	-.0005
KB α =9	.412	.249	.045	.0001
KB α =10	.391	.248	.055	.0015
DC s=30 dB	.663	.184	-.022	.0092
DC s=50 dB	.525	.234	.002	.0029
DC s=70 dB	.444	.248	.031	-.0002
ASE p=1,λ=.20	.437	.196	.055	.0216
ASE p=2,λ=.02	.460	.261	.044	-.0167

TABLE II
APODIZATION-NOISE FACTORS: f IS THE NOISE-REDUCTION FACTOR COMPUTED USING [], AND THE NOISE-CORRELATION FACTORS GIVEN IN PERCENT AND COMPUTED USING (32). 24 COSINE-EXPANSION TERMS WERE USED IN THE CALCULATION

function	f	$C_1(\%)$	$C_2(\%)$	$C_3(\%)$
Hamming	1.5863	62.51	13.31	
von Hann	1.6330	66.67	16.67	
Triangle	1.7147	60.38	14.90	6.71
Blackman	1.8119	75.51	31.55	6.57
NB w_1	1.2581	28.18	1.09	-.19
NB w_2	1.3611	40.04	2.83	-.30
NB m_1	1.4531	50.16	5.90	-.59
NB m_2	1.5141	56.93	8.35	-.75
NB s_1	1.6039	63.09	14.86	.70
NB s_2	1.6487	66.64	18.13	1.18
KB α =1	1.0749	9.14	-1.92	0.83
KB α =2	1.2285	27.68	-3.11	1.34
KB α =3	1.3712	43.54	0.06	.62
KB α =4	1.4838	54.30	5.80	.01
KB α =5	1.5746	61.61	12.12	.23
KB α =6	1.6513	66.85	18.20	1.30
KB α =7	1.7183	70.81	23.76	3.00
KB α =8	1.7782	73.91	28.76	5.13
KB α =9	1.8324	76.42	33.24	7.51
KB α =10	1.8822	78.48	37.24	10.0
DC s=30 dB	1.4025	46.39	1.77	.406
DC s=50 dB	1.6108	64.03	15.08	.758
DC s=70 dB	1.7616	73.09	27.48	4.60
ASE p=1,λ=.20	1.9074	71.07	34.89	15.62
ASE p=2,λ=.02	1.6845	74.41	28.09	-.22

The Norton–Beer apodization functions have characteristics similar to the $J = 2$ cosine functions, and the s_1 and s_2 functions are both reasonable functions to use for an RTA for high signal-to-noise instruments. The channel correlation of noise is similar to Hamming but extends farther due to the higher-order cosine expansion.

In most cases, the Kaiser–Bessel functions are more localized than other apodization functions with the same FWHM, as seen by the larger contribution of the central lobe in Fig. 6. The quickly vanishing side-lobes are an inherent advantage to the RTA, and this makes the Kaiser–Bessel $\alpha \approx 5$ function a better choice than the Hamming function for direct use in retrievals. For larger α values, the FWHM and channel correlation become larger, and this function is less suitable for algorithms that use a subset of channels.

The Dolph–Chebyshev apodization function with $s \approx 50$ dB is similar to the Kaiser–Bessel function with $\alpha \approx 5$. These functions also appear to be a good choice for use in retrievals.

The ASE apodization functions suffer from large-channel noise correlation relative to the Hamming function and do not appear to have significant advantages over other tunable functions, including the simple cosine functions.

In general, weak apodization functions may have smaller interchannel noise correlations, but they are poorer functions to use for the RTA due to the side-lobes. An apodization function that is stronger than Hamming apodization tends to be a good function for the RTA, but has a broader central lobe, which could be detrimental when a subset of channels is used. Also, the noise is more strongly correlated with the neighboring channels for the stronger apodization functions, and proper handling of the noise-covariance matrix is more critical within the retrieval.

IX. CONCLUSIONS

We show that apodization functions can be quantitatively compared using cosine expansions. In remote sensing applications requiring the use of RTA's, it is advantageous for some algorithms that the CSRF should have a combination of highest spectral purity and narrowest central lobe. In this context, we

determined that the Kaiser–Bessel ($\alpha \simeq 5$) and Dolph–Chebyshev ($s \simeq 50$ dB) functions were the best apodization functions for use with RTA's, although simple functions such as Hamming were also suitable choices. Other algorithms may prefer to use highly nonlocalized CSRF's, such as those arising from unapodized spectra. They pose practical complications for the rapid and accurate computation of channel radiances, however. We developed a simple methodology for accurate and efficient computation of radiances and other operators for any symmetric apodization functions, including unapodized spectra.

Unapodized radiances measured at the Nyquist sampling can be transformed to apodized radiances with a matrix multiplication. For the Hamming apodization function, an analytic inverse of this matrix $\mathbf{M}_{\mathbf{H}}^{-1}$ exists and is well-behaved such that retrieval algorithms can calculate Hamming radiances, Jacobians, and noise-covariance matrices, and convert them to analogous terms for any apodization, including unapodized radiances. For some other functions, the inverse matrix can be determined numerically and used for these purposes. In general, the functions must be highly localized so that an accurate RTA can be used to compute the apodized radiances.

For functions in which the inverse transformation exists, the sets of apodized and unapodized radiances, including noise effects, contain equivalent information. We also show that apodization does not affect the accuracy of regression retrievals for any apodization in which $\mathbf{M}_{\mathbf{G}}^{-1}$ exists, if all the channels are used within the retrieval. The same result should hold true for physically-based retrievals if all the channels are used. This demonstrates that the larger FWHM resulting from apodization relative to unapodized FWHM is not an indication that apodization has degraded the spectral resolution of the instrument.

Retrievals using a subset of channels may produce equivalent, or possibly improved, results compared to those using all channels. For retrieval algorithms in which a subset of channels is used, there may be an advantage of one function over another in the sense that the efficiency of the retrieval algorithm would depend on the efficiency of the computation of radiances, possibly using different numbers of channels.

REFERENCES

[1] H. H. Aumann and R. J. Pagano, "Atmospheric infrared sounder on Earth observing system," *Opt. Eng.*, vol. 33, no. 3, pp. 776–784, Mar. 1994.

[2] F. Cayla, "IASI infrared interferometer for operations and research," in *High Spectral Resolution Infrared Remote Sensing for Earth's Weather and Climate Studies*, A. Chedin, M. Chahine, and N. Scott, Eds. Berlin: Springer Verlag, 1993, vol. 9, pp. 9–19.

[3] R. H. Norton and R. Beer , "New apodizing functions for Fourier spectroscopy," *J. Opt. Soc. Amer.*, vol. 66, no. 3, pp. 259–264, Mar. 1976.

[4] U. Amato, V. Cuomo, and C. Serio, "An advanced optimal spectral estimation algorithm in Fourier spectroscopy with application to remote sensing of the atmosphere," *J. Appl. Meteorol.*, vol. 32, pp. 1508–1520, Sept. 1993.

[5] IASI Science Team and NPOESS Sounder Community, private communication, 1999.

[6] L. M. McMillin, M. D. Goldberg, H. Ding, J. Susskind, and C. D. Barnet, "A forward calculation for interferometers: Method and validation," *Appl. Opt.*, vol. 37, pp. 3059–3068, May 1997.

[7] U. Amato, V. Cuomo, and C. Serio, "Inverting interferogram signals to retrieve atmospheric temperature profiles," in *Tech. Proc. 7th Int. TOVS Study Conf.*, J. Eyre, Ed.. Igls, Austria, Feb. 10–16, 1993, pp. 1–10.

[8] S. Hannon, L. L. Strow, and W. W. McMillan, "Atmospheric infrared fast transmittance models: A comparison of two approaches," *SPIE, Ver. 2830*, vol. 94, p. 105, 1996.

[9] L. M. McMillin, L. J. Crone, and T. J. Kleespies, "Atmospheric transmittance of an absorbing gas. 5. Improvements to the OPTRAN approach," *Appl. Opt.*, vol. 34, pp. 8396–8399, Dec. 1995.

[10] J. R. Eyre and H. M. Woolf, "Transmittance of atmospheric gases in the microwave region: A fast model," *Appl. Opt.*, vol. 27, pp. 3244–3249, Aug. 1988.

[11] J. Susskind, J. Rosenfield, and D. Reuter, "An accurate radiative transfer model for use in the direct physical inversion of HIRS2 and MSU temperature sounding data," *J. Geophys. Res.*, vol. 88, pp. 8550–8568, Oct. 1983.

[12] E. O. Brigham, *The Fast Fourier Transform and its Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[13] J. Susskind, C. Barnet, and J. Blaisdell, "Determination of atmospheric and surface parameters from simulated AIRS/AMSU sounding data: Retrieval and cloud clearing methodology," *Adv. Space Res.*, vol. 21, no. 3, pp. 369–384, 1998.

[14] R. W. Hamming, *Digital Filters*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1977.

[15] R. J. Bell, *Introductory Fourier Transform Spectroscopy*. New York: Academic, 1972.

[16] U. Amato, I. De Feis, V. Cuomo, and C. Serio, "Regularization methods to solve inverse problems: An investigation in the context of Fourier spectroscopy from satellite," in *Proc. 5th Workshop Atmospheric Science from Space using Fourier Transform Spectroscopy*, Tokyo, Japan, Nov. 30–Dec. 2, 1994.

[17] P. Van Delst and H. Revercomb, private communication, Madison, WI, 1999.

[18] R. B. Blackman and J. W. Tukey, *The Measurement of Power Spectra*. New York: Dover, 1959.

[19] M. S. Bartlett, "Periodogram analysis and continuous spectra," *Biometrika*, vol. 37, pp. 1–16, 1950.

[20] R. H. Norton and R. Beer, "Errata," *J. Opt. Soc. Amer.*, vol. 67, p. 419, Mar., 1977.

[21] F. F. Kuo and J. F. Kaiser, *System Analysis by Digital Computer*. New York: Wiley, 1966.

[22] H. R. Ward, "Properties of Dolph–Chebyshev weighting functions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-9, p. 785, Sept. 1973.

Christopher D. Barnet was born in Staten Island, NY, on October 30, 1952, and received the B.S. degree in electronics technology in 1976 and the M.S. degree in solid state physics in 1978 from Northern Illinois University (NIU), DeKalb. He received the M.S. degree and the Ph.D. degree in astronomy in 1988 and 1990, respectively, from New Mexico State University, Las Cruces.

In 1979, he held a research position at NIU studying the superconducting properties of tertiary compounds. From 1981 to 1986, he was with Lockheed Engineering and Management Company, Las Cruces, NM, and Nicolet's FTIR Instrument Division, Madison, WI. His NRC postdoctoral associateship with NASA Goddard Space Flight Center (GSFC) from 1990 to 1992 utilized coupled radiative-dynamic models to understand IR and visible observations of the outer planets' stratospheres and tropospheres made by the Voyager Spacecraft's IRIS interferometer and Hubble Space Telescope (HST). From 1992 to 1995, he was with the Institute for Space and Terrestrial Science, Toronto, Ont., Canada, using the HST to study the outer planets' atmospheres using visible and ultraviolet observations and multiscattering radiative-transfer models. Since 1995, he has been with Space Applications International Corporation/General Sciences Corporation (SAIC/GSC), Beltsville, MD, on advanced algorithms for terrestrial infrared and microwave remote sensing.

John M. Blaisdell was born in Buffalo, NY, on July 12, 1956. He received the B.A. degree in mathematics and the B.S. degree in physics from the University of Rochester, Rochester, NY, in 1978. He received the M.S. and Ph.D. degrees in physics in 1978 and 1980, respectively, from the University of Illinois, Urbana. His dissertation research covered electronic structure of metal oxide semiconductors in the vicinity of surface defects.

From 1982 to 1983, he worked in the Chemistry Department, Johns Hopkins University, Baltimore, MD, where he studied the electronic structure of molecular crystals. He has been with Space Applications International Corporation/General Sciences Corporation (SAIC/GSC), Beltsville, MD, since 1983, working on space and atmospheric applications. Since 1993, he has been implementing scientific algorithms for atmospheric sounding applications.

Joel Susskind was born in Brooklyn, NY, on April 9, 1944. He received the A.B. degree from Columbia Univ., New York, in 1964 and the Ph.D. degree in physical chemistry from the University of California, Berkeley, in 1968, where he did research in experimental and theoretical microwave gas phase spectroscopy.

He did postdoctoral research in high resolution IR gas phase spectroscopy in the Chemistry Department, University of Minnesota, Minneapolis, from 1968 to 1970 and in the Physics Department, Florida State University, Tallahassee, from 1970 to 1972. He did research at the NASA Goddard Institute for Space Studies, Greenbelt, MD, from 1973 to 1977, involving atmospheric spectroscopy and radiative transfer and development of methodology to analyze atmospheric sounding data from a suite of IR and microwave sounding instruments. He is currently a Senior Scientist in the Laboratory for Atmospheres, NASA Goddard Space Flight Center, where he has worked since 1977, doing research in IR and microwave remote sensing. He also is currently the NASA Polar Orbiting Environmental Satellite Project Scientist and a member of the AIRS Science Team.