A Computationally Efficient Discrete Backus–Gilbert Footprint-Matching Algorithm

Philip J. Stephens and Andrew S. Jones

***Abstract*—A computationally efficient discrete Backus–Gilbert (BG) method is derived that is appropriate for resolution-matching applications using oversampled data. The method builds upon existing BG methods and approximation techniques to create a modified set of BG coefficients. The method in its current form is restricted to a resolution-only minimization constraint, but in the future could be extended to use a simultaneous noise minimization constraint using a generalized singular value decomposition (GSVD) approach. A theoretical one-dimensional intercomparison is performed using a hypothetical sensor configuration. A comparison of the discrete BG (DBG) method with a nondiscrete BG method shows that the new approach can be 250% more efficient while maintaining similar accuracies. In addition, an SVD approximation increases the computational efficiencies an additional 43%–106%, depending upon the scene. Several quadrature methods were also tested. The results suggest that accuracy improvements are possible using customized quadrature in regions containing known physical data discontinuities (such as along coastlines in microwave imagery data). The ability to recompute the modified BG coefficients dynamically at lower computational cost makes this work applicable toward applications in which noise may vary, or where data observations are not available consistently (e.g., in radio frequency interference (RFI) contaminated environments).**

***Index Terms*—Geophysical measurements, imaging, radiometry, remote sensing.**

I. Introduction

THE BACKUS–GILBERT (BG) method [1]–[3] has been employed by various authors to spatially coregister and invert various data sets while accounting for different spatial and error propagation behaviors [4, pp. 160–169]. In particular, the earth science remote sensing community has employed BG for footprint-matching between various satellite data channels (usually in the microwave spectrum) [5]–[9]. Fundamentally, this is because microwave sensors typically use a single large antenna across a relatively wide spectrum. Depending on the particular sensor configuration (e.g., sampling rates and illumination characteristics), this results in footprints that may overlap considerably and have spatial half-power beamwidths that are substantially different from one another. The reader is referred to [6] for a more detailed technical explanation of the instrument design constraints that lead to this common microwave sensor condition.

The Defense Meteorological Satellite Program (DMSP) Special Sensor Microwave/Imager (SSM/I) instrument is an example of a sensor that exhibits this type of behavior [10]. For this sensor, the ratio of the effective field of views (EFOVs) at 85-H GHz and 19-H GHz is 4.6×3.3 (along-track and cross-track, respectively) between colocated sensor footprints. This can create a substantial problem with regards to the correct interpretation of the brightness temperatures, since nonuniform fields can exist within a single footprint. The challenge of matching various data footprints will tend to increase as the number of coincident channels is increased. Stogryn [5] and Poe [6] provide additional discussion and graphical depictions that might be helpful to readers unfamiliar with the application of BG theory to the footprint-matching problem. In the interest of brevity, this paper does not attempt to replicate that discussion.

The original BG work [3] provides a rigorous mathematical basis for the inversion of inaccurate data. Later, Stogryn [5] applied it to the specific problem of microwave footprint-matching and further developed concepts from BG that are the basis of most BG footprint-matching applications today. A key feature of the BG method is that it can be used effectively to trade instrument noise for spatial resolution and vice versa. This flexibility is a fundamental strength of the BG approach. The earth science remote sensing community has expanded upon the original BG theory through various specific applications. However, other research communities have implemented various versions and modifications of the original BG method for spatial data analysis work. In particular, while initially not very well known among mathematicians [11], the mathematics community has subsequently expanded the body of knowledge in this area through various techniques (e.g., [12] and [13]), some of which will be discussed and applied later in this paper. Specifically, this paper combines the work in [5] and [6] with the work in [13].

It should be noted that some features of the original Stogryn method have been simplified in this work to allow the Hansen methodology [13] to be applied more fully, in particular the singular value decomposition (SVD) optimizations of Section IV. Stogryn [5] introduces an optional additional minimization constraint on the noise amplification so that resolution and noise are both minimized simultaneously to various degrees by varying a parameter γ. For this work, the case where $\gamma = 1$ is used to simplify the intercomparison analysis. Physically, this corresponds to a pure resolution minimization constraint. The inclusion of $\gamma \neq 1$ would require the use of a generalized SVD (GSVD) optimization approach [13], since several key matrices are no

Manuscript received November 28, 2001; revised April 4, 2002. This work was supported by the Department of Defense Center for Geosciences/Atmospheric Research at Colorado State University under Cooperative Agreement DAAL01-98-2-0078 with the Army Research Laboratory.

P. J. Stephens is with the Department of Physics, Cambridge University, Cambridge CB2 1TN, U.K. (e-mail: ps327@cam.ac.uk).

A. S. Jones is with the Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins, CO 80523 USA (e-mail: jones@cira.colostate.edu).

Publisher Item Identifier 10.1109/TGRS.2002.802520.

longer purely diagonal and become interdependent on the gain function. Utilization of a GSVD approach is beyond the current scope of this work. Thus, the practical applications of the discrete BG (DBG) methodology as described should only be used where sufficient data overlap occurs (i.e., where noise amplification is not a serious concern [9]).

This work has practical implications for the utilization of BG methods within the earth sciences community. For example, a long-standing problem with the application of BG to earth science remote sensing has been the computational expense of calculating the coefficients necessary for the method [14]. Current applications typically assume that the sensor and noise contributions are stable and that the coefficients can be assumed static. However, in an era of increasing radio frequency interference (RFI), the relatively benign radiometric operating conditions that the remote sensing community has enjoyed may be part of a passing era. Thus, methods that are more dynamic are needed to cope with such possible changes that threaten the performance of more traditional BG implementations. Also, certain computationally intensive applications involving remote sensing may impose demanding computational restrictions that traditional BG methods are not able to accommodate. In this vein, a new discrete DBG method is created that is computationally more efficient and operationally flexible in its configuration. In the new method, it will be shown that computational performance can be dynamically traded for method accuracy. This allows the method to expend CPU cycles where the spatial data analysis is most critical and vice versa.

In particular, the DBG method explicitly specifies the integration approach, which facilitates several optimization techniques. This paper specifically addresses two optimizations: 1) the diagonalization of several matrices following the Hansen integration approach and 2) the SVD approximation technique. It should be noted that the optimization improvements are compounding effects, in that the matrix diagonalization increases the computational performance by more than 250% for some scenes (see Section III), while the SVD performance gains are in addition to that increase in performance (see Section V). Several other optimizations remain unexplored that could exploit the flexibility of the DBG integration form; these include

1) adaptive grid methods;
2) customized or dynamic quadrature rules;
3) other possibilities.

Our purpose in this paper is to introduce the mathematical theory and optimization techniques that allow for this type of flexibility and efficiency, and to apply it to several simulated data set series for a theoretical one-dimensional sensor configuration. Several interesting intercomparisons are possible with the theoretical sensor configuration studies, but ideally, observational studies for specific sensors would also be performed in the future. This work is thus a foundational component for future work. Using theoretical simulations, the current work explores the impact of several integration quadrature rules. Aspects of the regularization behaviors are also discussed.

II. Theoretical Background and Derivation

As mentioned previously, the new BG method is based on [5] and [6] using the approach of [13]. Before deriving the new method, a common mathematical framework is introduced to reconcile the notational differences employed by the former works. All of the methods find a set of coefficients, $\mathbf{a}$, that provide a linear combination of the measurements to estimate the value at a point x_0. These coefficients are used to define an averaging kernel or interpolated gain function

$$A(x, x_0) = \sum_{i=1}^{M} a_i(x_0) G_i(x) \tag{1}$$

where $G_i(x)$ is the gain function centered at x_i, evaluated at position x, and i is the ith measurement out of M total measurements. If it were possible, the ideal interpolated gain function would be a delta function. By convolving the spatially integrated measurement $T_{\text{meas}}(x)$, with the interpolated gain function, an estimate of the measurement at a point x_0, $T_{\text{est}}(x_0)$, can be made

$$T_{\text{est}}(x_0) = \int A(x, x_0) T_{\text{meas}}(x) dx. \tag{2}$$

As an approximation to a typical microwave antenna beam pattern, this paper uses a truncated cosine function as the gain function $G_i(x) = \cos(x)$, for $0 \leq x \leq \pi/2$. In practice, the prelaunch antenna gain patterns or empirically estimated antenna gain patterns would be employed. The gain function represents oversampled data if i is simply a discretized version of the continuous spatial variable x. The spatial data overlap can be modified by adjusting the gain function width.

A. BG Method

The coefficients we are seeking (a_i) are given by [3, eq. (4.20)]

$$\mathbf{a} = \frac{\mathbf{S^{-1}u}}{\mathbf{u^T S^{-1} u}} \tag{3}$$

where $\mathbf{S}$ is an $M \times M$ matrix, and $\mathbf{u}$ is a vector of length M, with the condition $\mathbf{a} \cdot \mathbf{u} = 1$ imposed to keep the averaging kernel A unimodular (integration area equal to one), i.e.,

$$\int A dx = 1. \tag{4}$$

For (1) to be true, [i.e., its linear independence maintained with regards to the gain function $G_i(x)$], $\mathbf{S}$ must be symmetric and positive definite. It should be pointed out that if G_i were a delta function, then at point x_i, $a_j = 0$ for all $j \neq i$, $\mathbf{S}$ and $\mathbf{u}$ are defined as [3]

$$S_{ij} = \int G_i(x) G_j(x) J(x, x_0) dx \tag{5}$$

and

$$u_i = \int G_i(x) dx \tag{6}$$

where

S_{ij} ith and jth element of an $M \times M$ matrix;
M total number of measurements in both spatial directions;
u_i ith element of a vector of length M;
J criteria function or penalty function.

This function adjusts how much influence or weight nearby points have on the solution. In the original BG work, J was set

to be $(x_0 - x)^2$. It was later changed to an arbitrary function in both Hansen's and Stogryn's work. The larger J is at a point, the less that value is weighted in the solution (due to the inverse of S). Poe [6] set J to unity in order to put all points on an equal footing (i.e., no sidelobe distortions or corrections are used).

It is important to understand that the integration techniques within the BG method are not predetermined. In some application-specific instances, functional forms of the antenna gain patterns are used, and in others, discretized gain tables are used. The integration technique is left to the BG method implementer. The DBG method changes this, in that the integration for several terms is explicit, and then subsequent optimization is performed. It is important to note that the number of discrete integration intervals N is only referring to the DBG approach. In this work, a functional form for the gain was used for the intercomparisons. This corresponds to perfect knowledge of the antenna gain function; thus, differences between methods are due primarily to either explicit approximations made in the method (e.g., SVD approximation) or integration techniques (e.g., the DBG method explicitly specifies the integration technique but not the quadrature weights). In order to systematically analyze the behaviors of the DBG method, the Stogryn method is used as a "control" in the analysis that follows. Thus, the use of perfect knowledge favors the Stogryn method in regards to accuracy due to the relatively crude state of the DBG integration. This makes the accuracy comparisons conservative.

B. Discrete BG Method

Hansen discretizes the BG method using the simple rule [13]

$$\int G_i(x)dx \approx \sum_{k=1}^{N} w_k G_i(x_k) \tag{7}$$

where N is the number of discrete integration intervals, and w_k are the integration weights. This allows the BG method to be expressed as a product of vectors and matrices (see Appendix A for a description of the Hansen method). By combining the Hansen discretization with the Stogryn minimization constraints and conditions (see Appendix B) a modified set of coefficients a_i can be defined. The following derivation proceeds down that path and produces the DBG method.

Applying the Hansen discretization of the integration space to the Stogryn minimization constraint with $\gamma = 1$ yields

$$Q_0 = \sum_{k=1}^{N} \left[\sum_{i=1}^{M} a_i G_i(x_k) - F(x_k, x_0) \right]^2 \times J(x_k, x_0) w_k \tag{8}$$

where F is a feature extraction function chosen to produce the desired analysis (the reader is referred to [5] for discussion related to the feature extraction function's use and design). Unless otherwise specified, F was set equal to the gain function centered at x_0. The Stogryn constraint (8) is minimized with respect to $\mathbf{a}$, which results in the following relationship:

$$\sum_{i=1}^{M} a_i \sum_{k=1}^{N} G_{kk} w_k (J(x_0))_{kk} G_{ik} = \sum_{k=1}^{N} G_{kk} (J(x_0))_{kk} (F(x_0))_{kk} w_k \tag{9}$$

where the Hansen definitions are used for $(G)_{ij}$, $(J(x_0))_{ij}$, and $(W)_{ij}$ (see Appendix A), with the addition of the following feature extraction term definition:

$$(F(x_0))_{ij} = F(x_j, x_0)\delta_{ij}. \tag{10}$$

The Hansen matrices for $\mathbf{G}$, $\mathbf{J}$, and $\mathbf{W}$ are defined (see Appendix A), with the addition of a new matrix $\mathbf{F}$ which is defined as an $N \times N$ diagonal matrix (10). Thus, (9) becomes

$$\mathbf{Sa} = \mathbf{v} \tag{11}$$

where $\mathbf{S}$ and $\mathbf{v}$ are defined as

$$\mathbf{S} = \mathbf{G}(\mathbf{J}(x_0))\mathbf{W}\mathbf{G}^{\mathbf{T}} \tag{12}$$

and

$$\mathbf{v} = \mathbf{GJFWe}. \tag{13}$$

The extremum condition for $\mathbf{a}$ is applied; the Hansen "$\mathbf{u}$" is replaced by $\mathbf{v}$, and β is set to one to match the Stogryn form. This yields the following solution for $\mathbf{a}$ with the extremum condition

$$\mathbf{a_{St}} = (\mathbf{S} + \lambda^2 \mathbf{I})^{-\mathbf{1}} \mathbf{v} \tag{14}$$

where λ is a chosen regularization parameter [15].

A combination of the BG coefficients ($\mathbf{a_{BG}}$) and the Stogryn coefficients ($\mathbf{a_{St}}$) are used so that $\mathbf{a} \cdot \mathbf{u} = 1$ is true (see Appendix B). Therefore, the modified Hansen BG coefficients with the Stogryn minimization constraints can be written in matrix form as

$$\mathbf{a}(x_0) = \mathbf{M}\left[\mathbf{v} + \frac{1 - \mathbf{u^T M v}}{\mathbf{u^T M u}}\mathbf{u}\right] \tag{15}$$

where

$$\mathbf{M} = (\mathbf{S} + \lambda^2 \mathbf{I})^{-\mathbf{1}}. \tag{16}$$

The regularization parameter acts to unbias the estimates, which can occur if the gain function is unnormalized (not the averaging gain function which must be normalized). While most antenna gain patterns are typically presented in a normalized form, an unnormalized gain function can occur for some types of objective analysis functions, especially for those that are created on an ad hoc basis. Thus, the approach taken was to make the system as robust as possible and to account for the possibility of unnormalized gain functions to enhance the generality of the algorithm. For some gain function/scene configuration combinations, local biases can be intrinsic to the system. In our work, it was found to be most apparent for certain repetitive spatial structures or patterns. Since BG methods are part of a larger class of regularization methods that exhibit this behavior [13], this effect is common to all BG methods. To mitigate concerns about possible biases, it is recommended that simulated BG analyses be performed using typical scenes and patterns to determine the root-mean-square (rms) impact of such periodic data.

It should be noted that while there is an analytical form of the error propagation via the BG coefficients, given some assumptions about data independence, this approach is deficient in respect to scene-dependent analysis, since it assumes that the BG

theory is a perfect solution, which it is not. The various minimization constraints imposed on the system only create "optimal" BG coefficients. Numerical noise in the various inversions and integration method errors would not be accounted for by performing such a simplified analysis. In light of that limitation, it was determined that an rms scene-dependent analysis would provide the most insight into the spatial filter behaviors that include the scene-dependent behaviors.

III. Method Intercomparison Results

The Stogryn BG method [5] and the DBG method are applied to three simulated truth scenes to explore the various behaviors of the methods. The truth scenes are

A) constant temperature, $T = 200$ K;
B) same as A) but with $T = 300$ K for all $x > 0$;
C) same as B) but with $T = (200 + 50 \sin x)$ K, for all $x > 0$.

The truth scenes were chosen to represent various extremes. Scene A represents the simplest case and tests the methods for correct implementation details that manifest themselves as biases and random noise behaviors. Scene B is the most harsh example, in that a large discontinuity is present in the scene. The range of 100 K is representative of the magnitude of the brightness temperature discontinuity observed when transitioning from ocean to land surfaces in the littoral regions. As will be discussed later, this case exhibits some of the more significant differences in behavior. The last idealized scene, Scene C, represents a contrast between a homogeneous region and a region with significant spatial variations. This scene, while still significantly idealized, presents a stronger challenge to the BG methods than Scene A, and yields insights into the abilities of the methods to perform in regions where periodicity of the scene may be important.

In the following simulations, a hypothetical microwave sensor is assumed to have a known antenna gain function that is represented using a truncated cosine function. Each measurement is "observed" at each whole interval position (e.g., $k = 1, 2, \ldots, M$). It should be noted that the design and configuration of the hypothetical sensor used in the following simulations is not the focus of this work. The hypothetical sensor is merely meant to exercise the various BG methods in a manner in which the methods themselves can be objectively evaluated.

A. Stogryn Method Results

The Stogryn method is applied to three simulated truth scenes (A–C) where 5 K of random noise has been added, thus simulating a relatively noisy hypothetical microwave sensor. The estimated temperatures from the Stogryn method are shown in Fig. 1(a)–(c), using a penalty function of $J(x, x_0) = 1$, which is a reasonable choice for satellite footprint-matching applications [5], [6]. In all of the following simulations, the feature extraction function F is set to be the gain function centered at x_0. In Fig. 1(d)–(f), the results are shown for the case when the penalty function is $J(x, x_0) = x(x - x_0)^2$. Simulation results are presented in Table I for Scenes (A–C) as rms differences from the "truth" scene. Aside from the conclusion that the method is performing correctly for all simulated scenes, there is an obvious scene distinction in regards to the rms performances (see Table I). For example, the Scene A Stogryn method estimate (with $J(x, x_0) = 1$) exhibits an rms of 0.71 K, while the rms results from Scenes B and C are 12.2 K and 4.4 K, respectively. Thus, results from Scene A are considerably better in terms of rms performance. The reason for this behavior is that all points in Scene A share a common mean. The theoretical justification for this statement will be discussed in more detail in the lateral boundary condition discussions that follow. The rms performance for Scene A is substantially lower than the initial 5 K random noise component that was added; hence, the spatial overlap of the gain function is working to the method's advantage. In comparison to the other scenes, Scene A is the easiest, since it has no spatial structures; Scene B is the most difficult to estimate correctly, since it has a single large discontinuity at $x = 0$; and Scene C is the next most difficult case of the three. Scene C is not as difficult as Scene B, since the discontinuity at $x = 0$ within Scene C has a finite slope. The periodicity of Scene C for $x > 0$ is well represented by the Stogryn method.

The Stogryn method results for $J(x, x_0) = (x - x_0)^2$ [Fig. 1(d)–(f)] are similar to the $J(x, x_0) = 1$ simulations (see also Table I). A major distinction between the results is the lateral boundary condition (LBC) behaviors. In particular, Scene B shows much improved LBC behaviors for the $J(x, x_0) = (x - x_0)^2$ case, and Scene C exhibits a slight improvement for only the rightmost LBC estimate. This is exhibited in the rms values as well (see Table I), with a 4-K rms improvement for Scene B, which is the extreme example. However, the rms results for Scene C were slightly worse. This behavior suggests that a modified J functional form could be used as the data approach the edges of a scan. For example, the center of the scan could implement the $J(x, x_0) = 1$ form and transition to the $J(x, x_0) = (x - x_0)^2$ form near the edges. An alternate solution would be to discard the data near the LBC if $J(x, x_0) = 1$.

B. DBG Method Results

Similarly, the DBG method results using penalty functions of $J(x, x_0) = 1$ and $J(x, x_0) = (x - x_0)^2$ are presented in Fig. 2 (corresponding to Fig. 1). Again, a simple comparison demonstrates that the DBG method is also able to estimate the truth scenes with some fidelity. The LBC behaviors in Fig. 2 are a significant issue for the $J(x, x_0) = 1$ case, as was shown for the Stogryn results in Fig. 1. Like the Stogryn behaviors, the DBG LBC effect is minimized for Scenes A and B with the choice of $J(x, x_0) = (x - x_0)^2$. This suggests that this is a fundamental BG behavior with the Stogryn minimization constraints. An rms comparison between the DBG method and the Stogryn method (see Table I) shows that the DBG method is outperforming the Stogryn method for the simplest scenes (A–C) and is underperforming it for the most challenging scene (B). However, it should be noted that the DBG method simulations are for an integration configuration where $N = M = 50$, i.e., the number of measurements equal the number of discrete integration intervals. Under

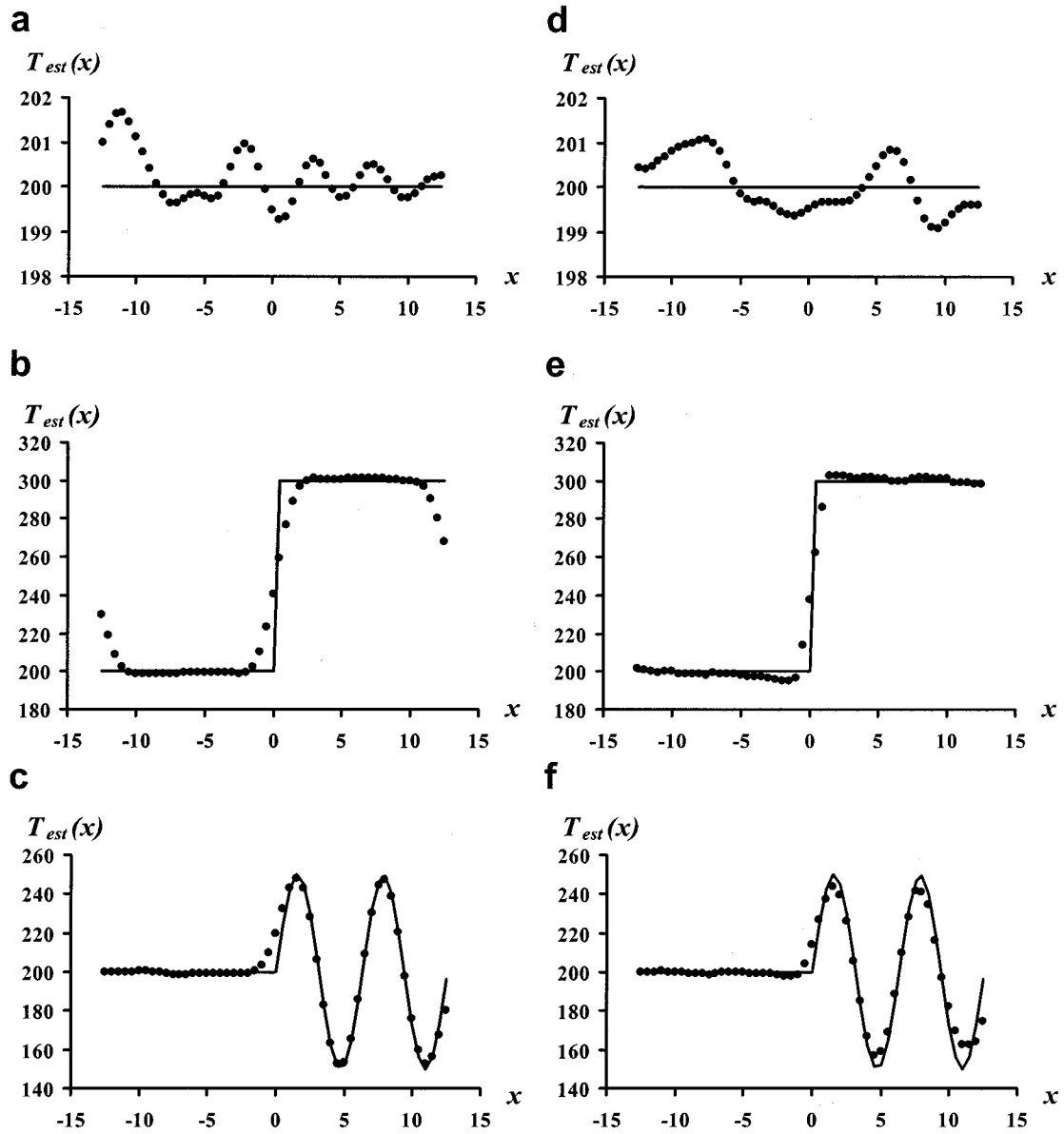

Fig. 1. Three scenes as estimated by the Stogryn method (dots) with $J(x, x_0) = 1$, with a random measurement error of 5 K imposed on the "true" scene functions (solid lines) which are (a) constant temperature $T = 200$ K, (b) the same as (a) but with $T = 300$ K, for all $x > 0$, and (c) the same as (a) but with $T = (200 + 50 \sin x)$ K, for all $x > 0$. (d)–(f) are the results (dots) that correspond to (a)–(c), respectively, except that $J = (x - x_0)^2$.

TABLE I
SIMULATION RESULTS FOR SCENES A–C WITH $M = 50$ AND $N = 50$

Method	Penalty Function $J(x_0)$	Scene A Homogenous Case		Scene B Step Function Case		Scene C Sine Function Case	
		RMS (K)	Std. Dev. (K)	RMS (K)	Std. Dev. (K)	RMS (K)	Std. Dev. (K)
Stogryn	1	0.71	0.14	12.2	0.08	4.4	0.08
DBG		0.61	0.14	12.3	0.07	4.5	0.15
Stogryn	$(x - x_0)^2$	0.51	0.14	8.2	0.06	5.8	0.13
DBG		0.33	0.06	10.4	0.08	5.3	0.13

these conditions, the discontinuity at $x = 0$ of Scene B is approximated by a single value within the discretized integration. Therefore, it is expected that Scene B should be the most challenging scene for the DBG method.

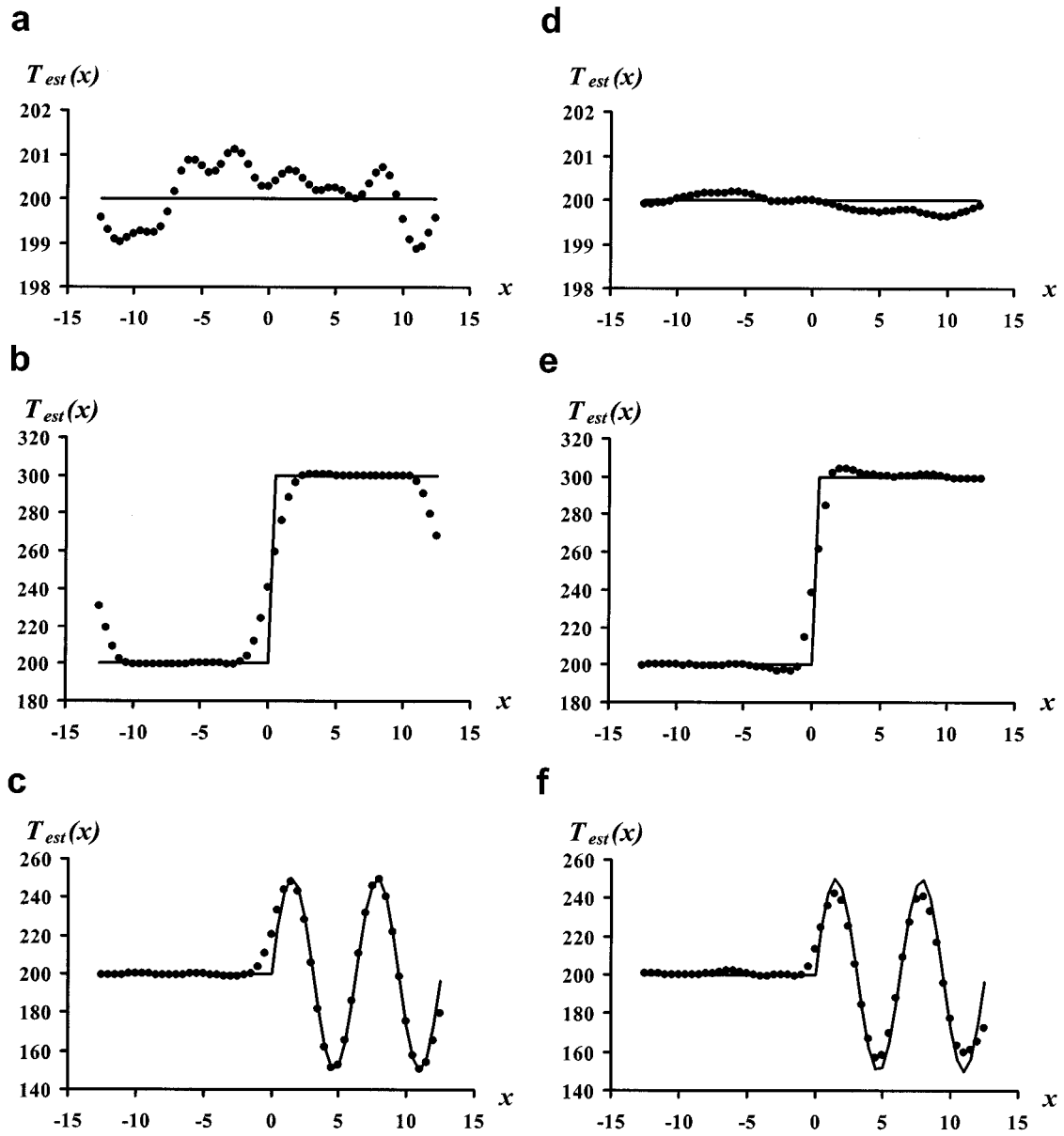

Fig. 2. Same as Fig. 1, except as estimated by the DBG method, where $N = M = 50$.

1) LBCs: As seen in Figs. 1 and 2, the LBC results for the step function scene [Scene B] are not well behaved when $J(x, x_0) = 1$. The LBC artifacts can be removed by increasing the input range, but not the output range. This has the same effect as discarding the estimates near the edges of the scan. For example, if the boundaries are moved further out (e.g., to $x = \pm 15$), the estimates are computed for the range $x = \pm 12.5$. This procedure effectively removes the LBC issues and, for the case when $J(x, x_0) = 1$, reduces the rms values when applied to the previous simulations (see Table II). The results for Scene B are presented in Fig. 3. Clearly, the LBC effects can be a significant problem if not handled correctly. As mentioned previously, an ideal solution would be to transition the functional form of J to a nonunity-based form near the edges of the scan, thus retaining the desirable $J(x, x_0) = 1$ analysis properties for footprint-matching applications.

2) Special LBC Considerations for $J(x, x_0) = 1$: It is important to see that setting $J(x, x_0) = 1$ in the original BG derivation does not give a spatially dependent estimate of the brightness temperatures. It is seen from (5) that the $\mathbf{S}$ matrix is independent of x_0 when J is constant. This makes the coefficients $\mathbf{a}$ (3), also independent of x_0. If the coefficients are independent of x_0, then the estimated temperature (2) will also be independent of x_0, and the BG derivation becomes nothing more then an averaging technique. The LBC artifacts arise because the $\mathbf{a_{BG}}$ term (see Appendix B) gives the mean of the input, while the $\mathbf{a_{St}}$ term produces the deviations from the mean, providing the correct estimations for the Stogryn minimization constraints. Again, if $\mathbf{S}$ is independent of x_0, then the only dependence for $\mathbf{a_{St}}$ comes from $\mathbf{v}$. When F is defined to be the gain function, this term decreases rapidly near the boundaries. This rapid decrease results in a decrease in the $\mathbf{a_{St}}$ term. That is why the LBC estimates approach the mean for the $J(x, x_0) = 1$ cases shown in Figs. 1 and 2.

3) Computational Performance versus Accuracy Considerations: Having shown that the LBC effects are a product of

TABLE II
SIMULATION RESULTS FOR SCENES A–C WITH $M = 100$ AND $N = 100$, AND WITH AN LBC SPATIAL FILTER APPLIED (CORRESPONDS TO FIGS. 4–6)

Method	Penalty Function $J(x_0)$	Scene A Homogenous Case RMS (K)	Scene A Homogenous Case Std. Dev. (K)	Scene B Step Function Case RMS (K)	Scene B Step Function Case Std. Dev. (K)	Scene C Sine Function Case RMS (K)	Scene C Sine Function Case Std. Dev. (K)
Stogryn	1	0.64	0.07	9.9	0.03	3.6	0.06
DBG		0.65	0.07	9.9	0.06	3.6	0.12

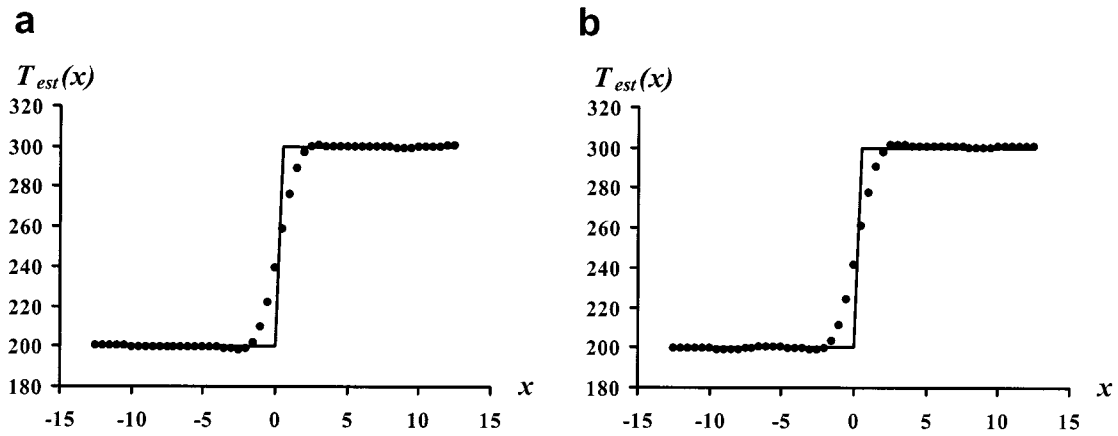

Fig. 3. Same as panel (b) of Figs. 1 and 2, except (a) is the DBG method and (b) is the Stogryn method. These results have the LBCs moved further away from the region of analysis by the filtering procedure discussed in the text.

the Stogryn minimization constraints applied with the condition that $J(x, x_0) = 1$, the biggest practical benefit of setting J to unity, other than its physical intuitiveness for sensor applications, is that $\mathbf{S}$ only needs to be calculated once, and can be reused for each estimate, which decreases the computational time. This optimization was not implemented for the results presented in this paper. Thus, all time intercomparisons within this paper are also generally applicable for other J definitions.

4) Configuration Behaviors: It should be noted that the results presented are for an oversampled sensor configuration; thus, the intersection of the neighboring gain functions at the interval midpoint is $G(x, 0.5)\mid_{x=0} = \cos(0.5) \cong 0.88$, and thus the strength of the overlap is not small. An alternative configuration was also used to simulate the placement of measurements at every other position, with $G(x/2, 0.5)\mid_{x=0} = \cos(1.0) \cong 0.54$. In this configuration, the DBG rms errors of the step function study increased by 43% for the DBG method, while the Stogryn rms errors increased by only 13%. The improved performance of the Stogryn method is expected, due to that method's increased integration sophistication. The DBG method's discrete integration approach was effectively reduced in resolving ability by a factor of two (since N was held constant and the gain function was twice as narrow), thus the more obvious degradation of performance. Since the resolution minimization constraint is imposed without accompanying noise minimization, the results for Scene B were most affected. It is expected that the results would continue to degrade as the overlap strength is decreased. For a particular observational sensor configuration and targeted applications, a more in-depth analysis would be appropriate.

The DBG method is quite flexible. Since the number of integration points is adjustable, this allows for faster, but less accurate, implementations of the technique. This would be particularly useful in the presence of large signal-to-noise ratios or when computational requirements are for a particularly fast implementation. The Stogryn implementation is optimized using library routines for numerical integration and matrix inversion. To limit the integration computational cost, the integration was optimized, and appropriate error thresholds were used. The DBG method also uses the same matrix inversion routine. However, due to its diagonal matrix form, matrix multiplications in the DBG method can be further optimized. This is one of the primary computational advantages of the DBG method.

A series of simulations were performed in which the number of integration intervals used within the DBG method N was modified to determine the computational cost of the method versus the method's rms accuracy. All results shown in Figs. 4–6 have the LBC windowing filter applied to remove the LBC artifacts for the $J(x, x_0) = 1$ cases. In each of the figures, the number of measures M is set to 100. The number of integration points N is adjustable. The Stogryn method was used as a control and was not modified from its base configuration.

Fig. 4 presents results for the uniform scene (Scene A). The DBG method rms values are generally lower than the Stogryn rms values. This is likely due to a more optimal integration pattern for the random distribution, since the DBG integration is evenly spaced. The numerically optimized integration weights used in the Stogryn method may also be amplifying the effect of

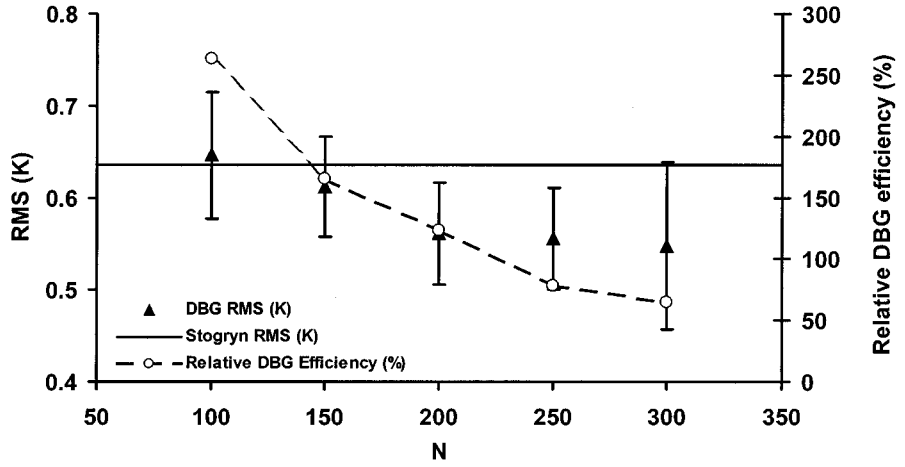

Fig. 4. Root-mean-square performance (K) for the uniform scene simulations (Scene A) for the DBG method (DBG) and the Stogryn method versus the number of integration points (N) are presented for the case where $J = 1$. The lateral boundary condition filter described in the text has been applied to the results shown. Also shown is the relative DBG computational efficiency (%) as compared to the Stogyrn computational costs.

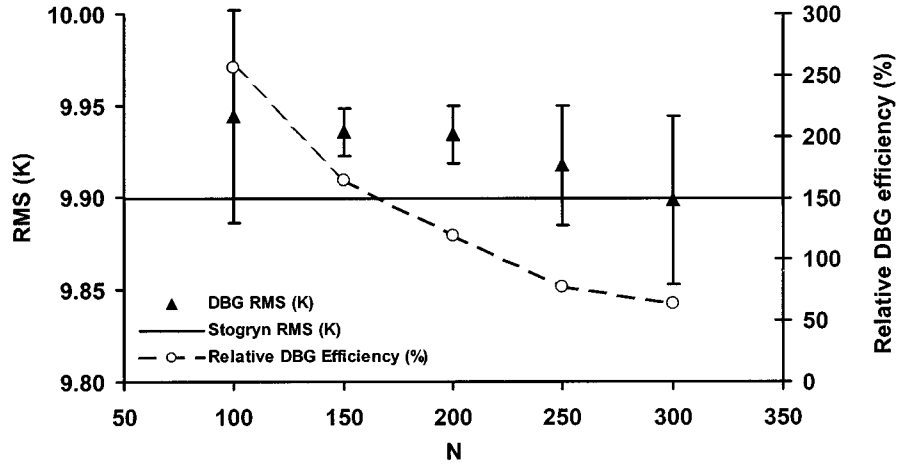

Fig. 5. Same as Fig. 4, except for the step function scene (Scene B).

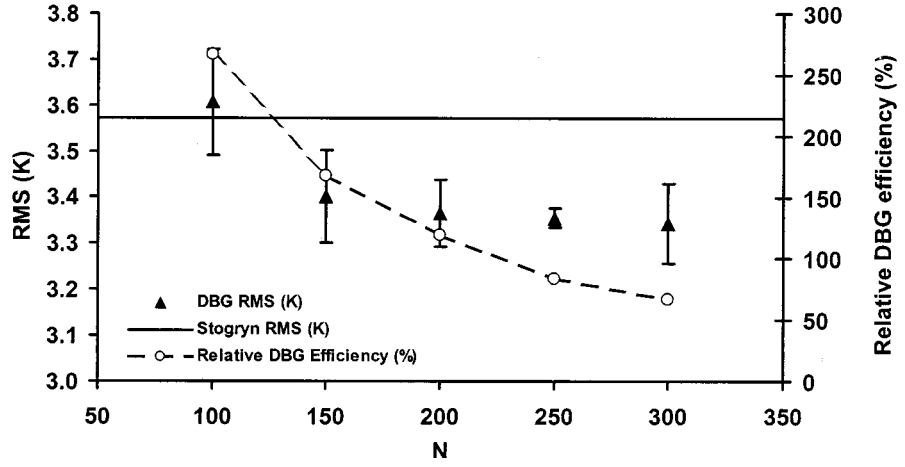

Fig. 6. Same as Fig. 4, except for the sine function scene (Scene C).

the various random patterns within the data set series. The effect of the uniform placement of integration intervals is even more pronounced for the case when $J(x, x_0) = (x - x_0)^2$, where the DBG method is approximately twice as accurate as the Stogryn method (see Table I). This suggests that the DBG method's integration pattern flexibility could be used to exploit known structures or features within data sets. The trend in the DBG results (see Fig. 4) is toward lower rms values with larger N. This corresponds to improved integration accuracy with larger N. The relative CPU consumption for the DBG method increases with

N as $\sim N^2$. The computational costs between the methods are approximately equivalent at $N \cong 220$.

The step function of Scene B is more of a challenge for both the DBG method and the Stogryn method (see Fig. 5). Both the DBG rms values and the Stogryn rms results are greater than the original noise. The DBG rms results are not significantly improved by going to larger values of N. As can be seen in Fig. 2, most errors are concentrated at the discontinuity. It is rather obvious that the step function is the harshest test of the DBG method. The computational cost for the DBG method is approximately the same as for the simple scene (Scene A1) results presented in Fig. 4.

The results for Scene C (see Fig. 6) are similar to that of Scene B (see Fig. 5), but with reduced severity. The rms ranges are below the 5 K random noise level, but Scene C rms values are still larger than the rms values for Scene A. Most error is again concentrated at the discontinuity at $x = 0$. The DBG method rms tendency is to decrease with greater values of N; hence, additional integration intervals are proving to be somewhat helpful with this particular scene. Again, the computational cost behaviors to process Scene C are similar to that of Scenes A and B.

IV. SVD Analysis

The SVD of a general real $M \times N$ matrix allows the target matrix to be separated into left and right singular vectors with the definition of appropriate singular values [16]. The practical consequence of this is that it allows the matrix to be handled via summations, and additional approximations become available to increase the overall performance of the method. The following SVD derivation is a modification of [13] to use the new DBG method with the Stogryn minimization constraints.

A new matrix and vector $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{e}}$ are defined as

$$\tilde{\mathbf{G}} = \mathbf{G}\mathbf{J}^{1/2}\mathbf{W}^{1/2} \tag{17}$$

$$\tilde{\mathbf{e}} = \mathbf{J}^{-1/2}\mathbf{W}^{1/2}\mathbf{e} \tag{18}$$

where $\tilde{\mathbf{G}}$ is an $M \times N$ matrix, and $\tilde{\mathbf{e}}$ is a vector of length N. This allows the following form for $\mathbf{S}$, which subsequently allows $(\mathbf{S} + \lambda^2\mathbf{I})^{-1}$ to be written in a notation more suitable for the SVD technique

$$\mathbf{S} = \tilde{\mathbf{G}}\tilde{\mathbf{G}}^{\mathbf{T}} \tag{19}$$

$$\mathbf{u} = \tilde{\mathbf{G}}\check{\mathbf{e}} \tag{20}$$

and

$$\mathbf{v} = \tilde{\mathbf{G}}\mathbf{J}\mathbf{F}\check{\mathbf{e}}. \tag{21}$$

SVD is now used to decompose $\tilde{\mathbf{G}}$ into

$$\tilde{\mathbf{G}} = \mathbf{B}\mathbf{\Sigma}\mathbf{C}^{\mathbf{T}} \tag{22}$$

with $\mathbf{B}^{\mathbf{T}}\mathbf{B} = \mathbf{I}$, $\mathbf{C}^{\mathbf{T}}\mathbf{C} = \mathbf{I}$, and $\Sigma = \mathrm{diag}(\sigma_i)$, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)} \geq 0$ are the singular values of $\mathbf{\Sigma}$ [17]. If the noise covariance matrix differs from $\mathbf{I}$, then a GSVD analysis could be performed [18]–[21]. A portion of the numerator term of (15) is defined such that

$$\mathbf{p}(x_0) = (\mathbf{S} + \lambda^2\mathbf{I})^{-1}\mathbf{u} = \mathbf{B}(\mathbf{\Sigma}\mathbf{\Sigma}^{\mathbf{T}} + \lambda^2\mathbf{I})^{-1}\mathbf{B}^{\mathbf{T}}\mathbf{B}\mathbf{\Sigma}\mathbf{C}^{\mathbf{T}}\tilde{\mathbf{e}} \tag{23}$$

where (20) and (22) have been substituted into the expression. Through SVD, (23) can be expressed as

$$\mathbf{p}(x_0) = \sum_i \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}\frac{\mathbf{c}_i^{\mathbf{T}}\tilde{\mathbf{e}}}{\sigma_i}\mathbf{b}_i \tag{24}$$

where $\mathbf{p}$ is a vector of length M, and $\mathbf{c}_i$ and $\mathbf{b}_i$ are the right and left singular vectors of $\tilde{\mathbf{G}}$ with length N and M, respectively, where $\mathbf{b}_i$ are the eigenvectors of $(\tilde{\mathbf{G}}\tilde{\mathbf{G}}^{\mathbf{T}} + \lambda^2\mathbf{I})$. This leads to the following solution for $\mathbf{a}$:

$$\mathbf{a} = \mathbf{r}(x_0) + \frac{1 - \mathbf{p}(x_0)^{\mathbf{T}}\mathbf{r}(x_0)}{\mathbf{p}(x_0)^{\mathbf{T}}\tilde{\mathbf{G}}\check{\mathbf{e}}}\mathbf{p}(x_0) \tag{25}$$

where

$$\mathbf{p}(x_0)^{\mathbf{T}}\tilde{\mathbf{G}}\check{\mathbf{e}} = \sum_i \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}\left(\mathbf{c}_i^{\mathbf{T}}\check{\mathbf{e}}\right)^2 \tag{26}$$

$$\mathbf{r}(x_0) = \sum_i \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}\frac{\mathbf{c}_i^{\mathbf{T}}\dot{\mathbf{e}}}{\sigma_i}\mathbf{b}_i \tag{27}$$

and

$$\dot{\mathbf{e}} = \mathbf{F}\mathbf{J}\check{\mathbf{e}}. \tag{28}$$

Equations (24)–(28) represent the SVD form of the DBG method.

In summary, the SVD form of the DBG method converts the matrix form of the DBG method into a specialized group of summations. As will be shown later, this results in significant computational savings. It should be noted that the SVD computational savings are in addition to the computational savings that result from the inherent form of the DBG method over traditional methods. The SVD form of the DBG is an explicit optimization technique that specifically exploits the DBG diagonal matrix form.

V. SVD Results

The summations in the SVD of the DBG method can be truncated at any desired point. Fig. 7 shows a sequence of results in which additional terms are progressively added to the SVD summation, for the case where $M = 100$. It can be seen that the effect is similar to that of a Fourier series. The first few terms give the general average and scaling of the spatial features, and additional terms refine the spatial structure of the results.

The rms performances of the SVD results are calculated for three simulated truth scenes as before. The results (see Table III) are defined as ratios in which the SVD DBG rms results are normalized relative to the non-SVD DBG rms results. Corresponding standard deviations are also calculated. Root-mean-square ratio results less than one indicate an improvement using the SVD approach. The computational cost ratio is defined similarly, with the cost ratio being defined as the SVD cost relative to the non-SVD cost. The computational costs are independent of the specific scene. These results show that the rms errors decrease as the number of summation terms is increased. Table III also shows a linear relationship between computational costs and the number of summation terms, which is expected. In practice, 20% of the terms generally gave sufficient structure, low rms values, as well as minimizing computational costs for the

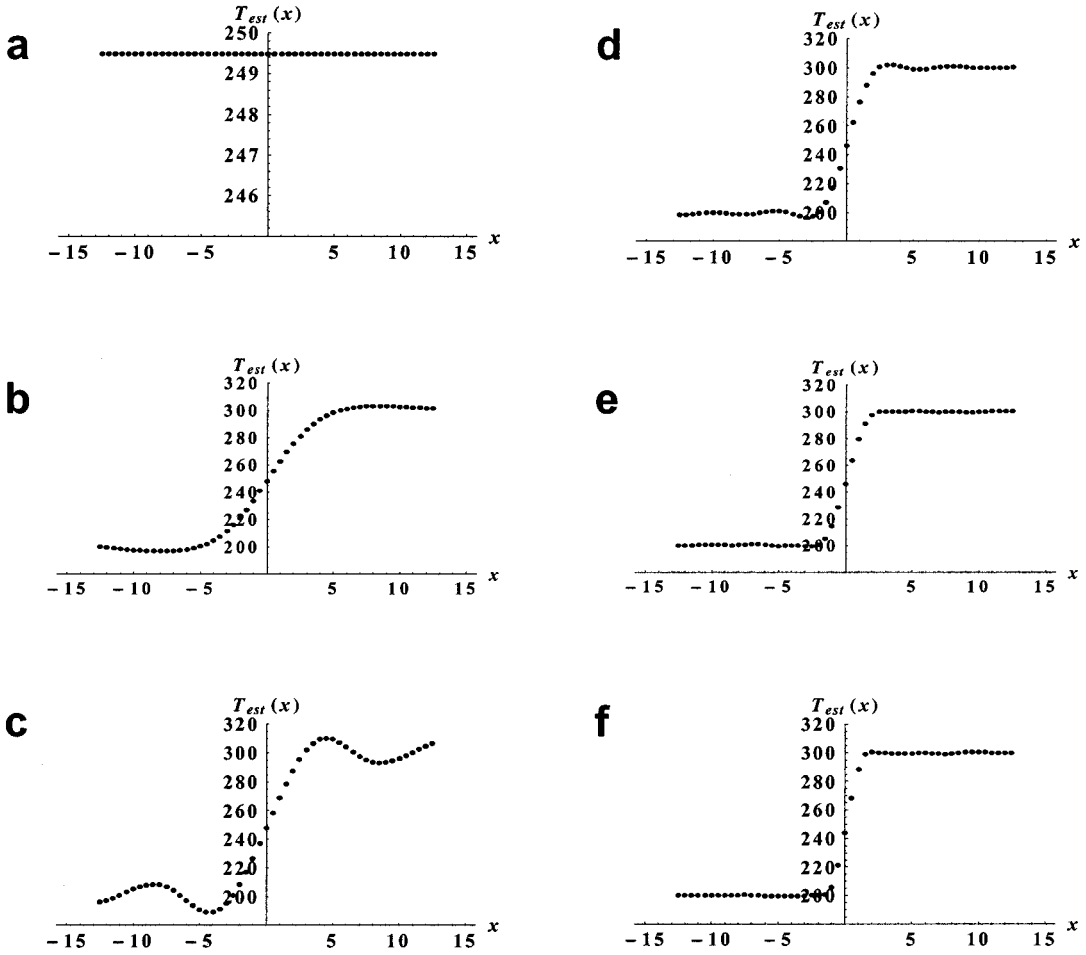

Fig. 7. Summations in the SVD of the DBG method can be truncated at any desired point with reduced fidelity. A sequence of results is shown in which additional terms are progressively added to the SVD summation, where $M = 100$, and (a) has 1% of M terms, (b) 3%, (c) 5%, (d) 10%, (e) 15%, and (f) 100%.

TABLE III
SVD DBG RMS PERFORMANCE RELATIVE TO THE NON-SVD DBG RMS PERFORMANCE

	Uniform		Step		Sine		Cost	
%M	RMS	σ	RMS	σ	RMS	σ	AVERAGE	σ
100	0.50	0.09	0.991	0.003	1.30	0.03	0.699	0.009
50	0.47	0.05	0.993	0.004	1.31	0.03	0.580	0.006
20	0.43	0.07	0.992	0.004	1.31	0.03	0.519	0.010
10	0.43	0.07	1.062	0.004	1.46	0.03	0.486	0.002

particular simulated scenes explored here. However, it should be pointed out that performance was dependent on the particular scene. For the uniform scene, all SVD rms results were significantly improved, with rms values less than 50% of the non-SVD rms results, while the step function scene results were only marginally improved. The sine wave scene was the most difficult scene for the SVD approach and resulted in errors that were 30% greater than the non-SVD results. The choice of where to truncate the SVD series will in general be dependent on the specific application and its associated error tolerance requirements. It is also worth noting that if $J(x, x_0)$ is constant, then the SVD of $\tilde{\mathbf{G}}$ only needs to be performed once. This improves the run times by a factor of three or more and does not change the output.

VI. QUADRATURE SENSITIVITIES

In this work, the integrations in the BG method were discretized using (7). Since (7) uses the same set of points x_k for the integral approximation, and is independent of the choice of the discrete gain positions x_i, the total summation for each gain function may actually differ slightly, though the nondiscretized integral would not. This can lead to certain areas being emphasized more or less than other areas. This emphasis can be especially useful to improve the estimates at certain locations. The quadrature points x_k and weights w_k can be selected using simple integration estimates, such as the trapezoidal rule [22], or by more advanced techniques, like Gaussian quadrature, or

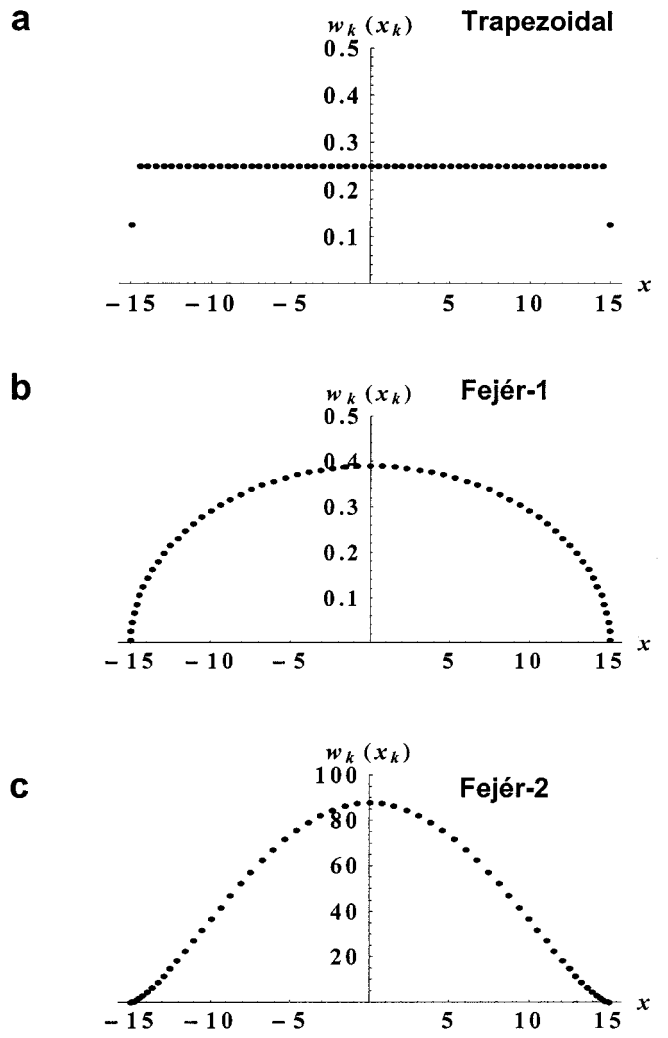

Fig. 8. Quadrature weights w_k and positions x_k for (a) trapezoidal, (b) Fejér-1, and (c) Fejér-2 quadrature.

Fejér quadrature [23]. Here we compare, both graphically and statistically, the results of using three different quadrature rules on a step function (200 K $x < 0$, 300 K $x > 0$). The gain function used is a truncated cosine function with width of 2π. The width is wider than the prior analysis configuration to enhance the region of transition from 200 K to 300 K. This allows the intercomparison to be more pronounced.

The three different quadrature weights are as follows.

1) *Trapezoidal:* d for all points except the end points, which are $d/2$, where d is the separation between neighboring points [22] (all points are equally separated).
2) *Fejér-1:* Consists of a Fejér weight function of 1, using the Gauss points of the Chebyschev weight function, $(1-(x/b)^2)^{1/2}$, where b is the maximum value of x [23].
3) *Fejér-2:* Consists of a Fejér weight function of $(x-a)(b-x)$, again using the Gauss points of the Chebyschev weight function $(1-(x/b)^2)^{1/2}$, where a is the minimum value of x and b is the maximum value of x [23].

Graphical depictions of the three weighting functions are shown in Fig. 8. It can be seen from these weighting functions that the Fejér-2 quadrature has a much sharper peak at $x = 0$ than either of the other quadrature methods.

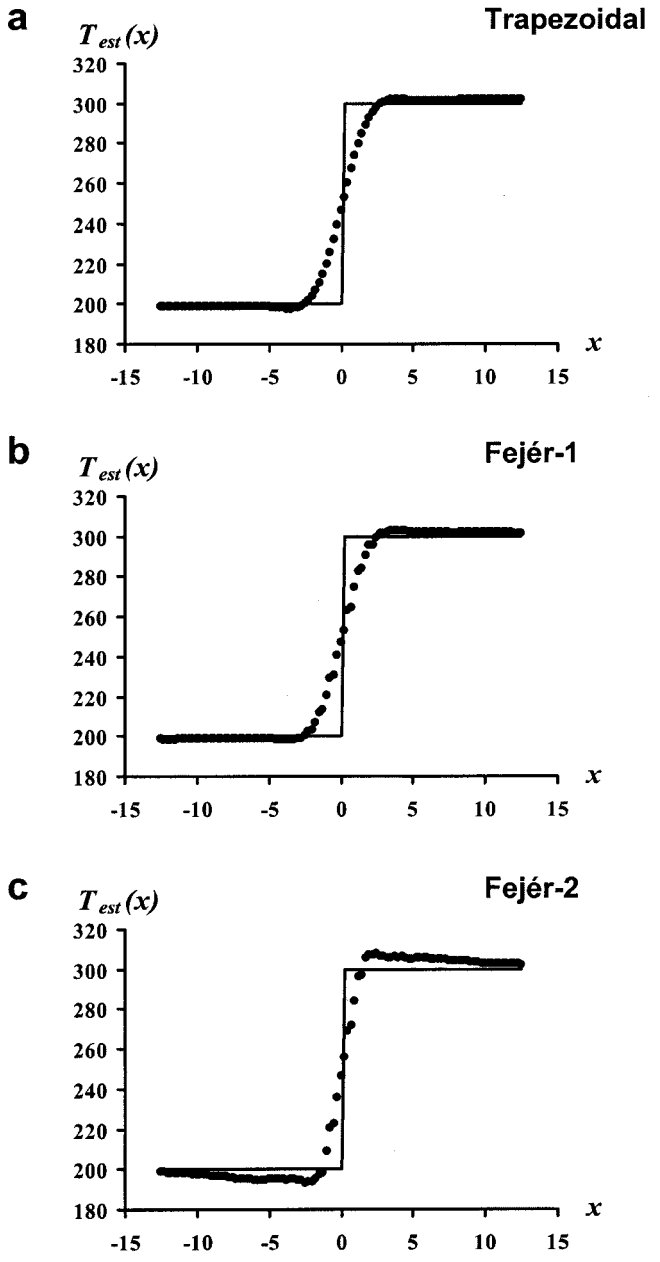

Fig. 9. DBG temperature estimates (dots) of the step function scene (solid lines) using (a) trapezoidal, (b) Fejér-1, and (c) Fejér-2 quadrature.

DBG-estimated temperature results using each quadrature rule are presented in Fig. 9. The Fejér-2 quadrature results have a much sharper slope at the discontinuity than the other quadrature methods. This increase in slope definition is exchanged for a decrease in the accuracy of the uniform parts of the scene. The rms results for each quadrature method are presented in Table IV for the three idealized scenes. It is important to note that the choice of the quadrature technique can improve the rms performance. This is significant and suggests that customized quadrature methods could be employed along known physical data discontinuities (such as along coastlines in microwave imagery data).

TABLE IV
DBG RMS PERFORMANCE USING THREE DIFFERENT QUADRATURE RULES FOR THREE SIMULATED SCENES

Scene	Trapezoidal		Fejér-1		Fejér–2	
	RMS	σ	RMS	σ	RMS	σ
Uniform	0.40	0.09	0.43	0.11	0.50	0.09
Step	11.24	0.04	11.37	0.04	10.26	0.09
Sine	10.00	0.08	10.00	0.07	6.11	0.10

VII. CONCLUSIONS

A DBG method has been created that is a modification of the Hansen BG method [13] to include the Stogryn minimization constraints. Several behaviors were observed. In particular, the LBC rms behaviors, for the $J(x, x_0) = 1$ case for both the Stogryn and the DBG methods, were explained as a consequence of the Stogryn minimization. It is suggested that alternative forms for J should be explored to minimize this effect, so that data near the scan edge are not simply discarded. A composite functional form for J in which the center of the scan implements a $J(x, x_0) = 1$ functional, and toward the edges of the domain the $J(x, x_0) = (x - x_0)^2$ form becomes dominant, was suggested as a possible candidate. Additionally, the rms performance of each method was shown to be dependent on the individual scene structure. The DBG method with regularly spaced integration weights was able to outperform the traditional Stogryn method for a uniform scene with random noise. For the more challenging scenes, the Stogryn method performed slightly better due to the application of a more sophisticated spatial integration method. As a practical consequence, the DBG method can be employed as a less costly footprint-matching algorithm. If necessary to increase the overall method's performance, modifications to the distribution and number of integration weights could be made. In a similar vein, if conditions warrant, the Stogryn method could be applied for structures that are particularly difficult to spatially resolve. Additional work is needed to examine the criteria and conditions for when such a switch between methods should be made.

The DBG method was also optimized using an SVD method. The result is increased computational efficiencies between 43%–106% while maintaining similar accuracies. By incorporating the Stogryn γ parameter [5], the optimization approach is also able to incorporate the simultaneous minimization of resolution and noise; however, doing so would require using a GSVD approach. In this paper, the additional complexity of a GSVD method was avoided, but it deserves future exploration.

The examination of several alternative quadrature methods demonstrates the sensitivity of the DBG method to the choice of the quadrature rules employed, and suggests that further optimization methods may yield additional increased performance, if some *a priori* information exists regarding the overall scene structure. In addition, the regularization sensitivity studies indicate the potential for improper regularization when used with unnormalized gain functions. This is a caution worth noting, in that future analysis functions may diverge from traditional normalized behaviors. This work also demonstrates a formalism to identify and constrain such possible biases. In practice, when using unnormalized gain functions, it is suggested that simulations be used to verify the accuracy of the method with the particular gain function.

The flexibility of the DBG method allows it to trade computational cost for accuracy in most conditions, thus lending it to several challenging research application areas. In particular, the use of a more flexible method would serve well in applications where the BG coefficients need to be routinely recalculated, depending on conditions, such as in an RFI-contaminated environment. The flexibility of the DBG method also allowed for several of the optimizations to be performed in a rather straightforward manner. Many additional optimizations are likely possible.

Future work will include investigating additional computational enhancements and testing the scope and validity of those assumptions. Additional research will also be conducted to exploit the flexibility of the DBG method for dynamic resolution and noise adjustments based on the signal-to-noise content. The new work will be applicable toward dynamic RFI minimization techniques, as well as dynamic enhancement of littoral features (based on content). It is our hope that the flexible framework of the DBG method will foster additional BG implementations that resolve some of these remaining challenges.

APPENDIX A
HANSEN METHODOLOGY

As mentioned in the main text, the BG method is discretized using the simple rule [13]

$$\int G_i(x)dx \approx \sum_{k=1}^{N} w_k G_i(x_k) \tag{A1}$$

where N is the number of discrete integration intervals, and w_k are the integration weights. Hansen does not necessarily define x_k to be the same set of points as x_i used in the kernel or gain function G_i. The set of points x_k can be adjusted to increase the accuracy or speed. Discretization of the BG integration space is accomplished by substituting (7) into the $\mathbf{S}$ matrix (5) and $\mathbf{u}$ vector (6) BG definitions. The DBG definitions of $\mathbf{S}$ and $\mathbf{u}$ then take the following form:

$$\begin{aligned} S_{ij} &= \int G_i(x)G_j(x)J(x, x_0)dx \\ &\approx \sum_{k=1}^{N} G_i(x_k)G_j(x_k)J(x_k, x_0)w_k \end{aligned} \tag{A2}$$

and

$$u_i = \int G_i(x)dx \approx \sum_{k=1}^{N} G_i(x_k)w_k. \tag{A3}$$

If an identity vector $\mathbf{e}$ of length N is defined, then the summations over k can be forced so that a matrix element form is obtained, so that

$$S_{ij} = G_{ik}G_{jk}(J(x_0))_{kk}W_{kk}e_k \tag{A4}$$

and

$$u_i = G_{ik}W_{kk}e_k \tag{A5}$$

where the following matrix element definitions are used [13]:

$$(G)_{ij} = G_i(x_j), \tag{A6}$$

$$(J(x_0))_{ij} = J(x_j, x_0)\delta_{ij} \tag{A7}$$

and

$$(W)_{ij} = w_j\delta_{ij} \tag{A8}$$

where δ_{ij} is the Kronecker delta. Since $\mathbf{J}$ and $\mathbf{W}$ are diagonal matrices by definition [(A7) and (A8)], (A4) and (A5) can be expressed in matrix notation as

$$\mathbf{S} = \mathbf{G}\mathbf{J}(x_0)\mathbf{W}\mathbf{G}^{\mathbf{T}} \tag{A9}$$

and

$$\mathbf{u} = \mathbf{G}\mathbf{W}\mathbf{e} \tag{A10}$$

where

$\mathbf{S}$	$M \times M$ matrix;
$\mathbf{G}$	$M \times N$ matrix;
$\mathbf{J}$ and $\mathbf{W}$	$N \times N$ diagonal matrices;
$\mathbf{e}$ and $\mathbf{u}$	vectors of length N and M, respectively.

The last change that is made by [13] is to apply the extremum condition. This leads to the following solution for $\mathbf{a}$ at position x_0 [combining (A9) and (A10) with (3)]:

$$\mathbf{a}(x_0) = \frac{(\mathbf{S} + \lambda^2\mathbf{I})^{-1}\mathbf{G}\mathbf{W}\mathbf{e}}{(\mathbf{G}\mathbf{W}\mathbf{e})^{\mathbf{T}}(\mathbf{S} + \lambda^2\mathbf{I})^{-1}\mathbf{G}\mathbf{W}\mathbf{e}} \tag{A11}$$

where λ is a chosen regularization parameter [15], and $\mathbf{I}$ is an identity matrix. In practice, numerical techniques are employed to find an optimal λ. Since $\mathbf{S}$ may be singular, depending on the choice of $\mathbf{J}(x_0)$, selection of an optimal λ is a matter of iteratively converging to a value that provides error minimization.

Appendix B
Stogryn Methodology

Stogryn [5] adds another criterion to the BG method. Stogryn presents the idea of minimizing, with respect to $\mathbf{a}$, another function of the form

$$Q_0 = \int \left[\sum_{i=1}^{M} a_iG_i(x) - F(x, x_0)\right]^2 J(x, x_0)dx \tag{B1}$$

with a normalization contraint

$$\int \sum_{i=1}^{M} a_iG_i(x)dx = 1 \tag{B2}$$

where F is a feature extraction function. This leads to the condition

$$\int \sum_{i=1}^{M} a_iG_i(x)G_j(x)J(x, x_0)dx = \int G_j(x)F(x, x_0)J(x, x_0)dx. \tag{B3}$$

Equation (B3) can be expressed more compactly by defining a vector $\mathbf{v}$ with the following elements:

$$v_j = \int G_j(x)F(x, x_0)J(x, x_0)dx; \tag{B4}$$

substituting (5) into (B3), and rearranging the summation and integration order, and converting to matrix form produces

$$\mathbf{a_{St}} = \mathbf{S^{-1}}\mathbf{v} \tag{B5}$$

which is the new condition of the Stogryn minimization. The final solution is a combination of BG coefficients (3), $\mathbf{a_{BG}}$, and Stogryn coefficients (B5), $\mathbf{a_{St}}$, so that $\mathbf{a} \cdot \mathbf{u} = 1$ is valid. The result is of the form $\mathbf{a} = \mathbf{a_{BG}} + \mathbf{a_{St}} + f(\mathbf{u}, \mathbf{v}, \mathbf{a_{St}}, \mathbf{a_{BG}})$. Taking $\mathbf{a} \cdot \mathbf{u}$ and realizing that the first term is already equal to 1, then $\mathbf{a_{St}} \cdot \mathbf{u} = -f(\mathbf{u}, \mathbf{v}, \mathbf{a_{St}}, \mathbf{a_{BG}}) \cdot \mathbf{u}$ must also be true, which results in the following expression for $\mathbf{a}$:

$$\mathbf{a} = \mathbf{a_{BG}} + \mathbf{a_{St}} - (\mathbf{u} \cdot \mathbf{a_{St}})\mathbf{a_{BG}}. \tag{B6}$$

Substituting (3) and (B5) into (B6) produces the BG coefficients with the Stogryn minimization constraints imposed [i.e., (B1) and (B2)]

$$\mathbf{a}(x_0) = \mathbf{S^{-1}}\left[\mathbf{v} + \frac{1 - \mathbf{u^T}\mathbf{S^{-1}}\mathbf{v}}{\mathbf{u^T}\mathbf{S^{-1}}\mathbf{u}}\mathbf{u}\right]. \tag{B7}$$

Acknowledgment

The authors would like to thank T. Greenwald for a helpful review.

References

[1] G. Backus and F. Gilbert, "Numerical applications of a formalism for geophysical inverse problems," *Geophys. J. R. Astron. Soc.*, vol. 13, pp. 247–276, 1967.

[2] ——, "The resolving power of gross earth data," *Geophys. J. R. Astron. Soc.*, vol. 16, pp. 169–205, 1968.

[3] ——, "Uniqueness in the inversion of inaccurate gross earth data," *Philos. Trans. R. Soc. London A, Math. Phys. Sci.*, vol. 266, no. 1173, pp. 123–192, 1970.

[4] S. Twomey, *Introduction to the Mathematics of Inverse Remote Sensing and Indirect Measurements*. New York: Dover, 1996.

[5] A. Stogryn, "Estimates of brightness temperatures from scanning radiometer data," *IEEE Trans. Antennas Propagat.*, vol. AP-26, pp. 720–726, May 1976.

[6] G. Poe, "Optimum interpolation of imaging microwave radiometer data," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 800–810, Sept. 1990.

[7] M. R. Farrar and E. A. Smith, "Spatial resolution enhancement of terrestrial features using deconvolved SSM/I microwave brightness temperatures," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 349–355, Mar. 1992.

[8] W. D. Robinson, C. Kummerow, and W. S. Olson, "A technique for enhancing and matching the resolution of microwave measurements from the SSM/I," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 419–428, May 1992.

[9] R. Bennartz, "Optimal convolution of AMSU-B to AMSU-A," *J. Atmos. Ocean. Technol.*, vol. 17, pp. 1215–1225, 2000.

[10] J. P. Hollinger, J. L. Peirce, and G. A. Poe, "SSM/I instrument evaluation," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 781–790, Sept. 1990.
[11] A. Kirsch, B. Schomburg, and G. Berendt, "The Backus–Gilbert method," *Inverse Problems*, vol. 4, pp. 771–783, 1988.
[12] R. Snieder, "An extension of Backus–Gilbert theory to nonlinear inverse problems," *Inverse Problems*, vol. 7, pp. 409–433, 1991.
[13] P. Hansen, "The Backus–Gilbert method: SVD analysis and fast implementation," *Inverse Problems*, vol. 10, pp. 895–904, 1994.
[14] J. F. Galantowicz and A. W. England, "The michigan earth grid: Description, registration method for SSM/I data and derivative map projections," Radiation Lab., Univ. of Michigan, Ann Arbor, MI, 027 396-2-T, 1991.
[15] R. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*. Norwell, MA: Kluwer, 1991.
[16] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK User's Guide*. Philidelphia, PA: SIAM, 1999.
[17] J. W. Demmel and W. Kahan, "Accurate singular values of bidiagonal matricies," *SIAM J. Sci. Stat. Comput.*, vol. 11, pp. 873–912, 1990.
[18] C. Paige, "Computing the generalized singular value decomposition," *SIAM J. Sci. Stat.*, vol. 7, pp. 1126–1146, 1986.
[19] Z. Bai and J. W. Demmel, "Computing the generalized singular value decomposition," *SIAM, J. Sci. Comput.*, vol. 14, pp. 1464–1486, 1993.
[20] Z. Bai and H. Zha, "A new preprocessing algorithm for the computation of the generalized singular value decomposition," *SIAM J. Sci. Comput.*, vol. 14, pp. 1007–1012, 1993.
[21] J. Christensen-Dalsgaard, P. C. Hansen, and M. Thompson, "GSVD analysis of heiloseismic inversions," *Mon. Not. R. Astron. Soc.*, vol. 264, pp. 541–564, 1993.
[22] S. D. Conte and C. de Boor, *Elementary Numerical Analysis: An Algorithmic Approach*. New York: McGraw-Hill, 1980.
[23] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*. New York: Academic, 1975.

Philip J. Stephens was born in Fort Collins, CO, in 1979. He received the B.S. degree in physics and the B.S. degree in computer science from Colorado State University, Fort Collins, in 2001. He is currently pursuing the Ph.D. degree in physics at Cambridge University, Cambridge, U.K.

Currently, he is working on Monte Carlo collider simulations for upcoming Large Hadron Collider experiments at the European Organization for Nuclear Research (CERN) facilities. In 2001, he was with the Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, at which time he worked on distributed systems and footprint-matching algorithms.

Mr. Stephens is the recipient of a Cambridge Overseas Trust Scholarship.

Andrew S. Jones was born in Mattoon, IL, in 1964. He received the B.S. degree in physics from Eastern Illinois University, Charleston, in 1986, and the M.S. and Ph.D. degrees in atmospheric science from Colorado State University, Fort Collins, in 1988 and 1996, respectively.

He is currently a Research Scientist with the Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, where he is working on satellite data assimilation techniques to retrieve soil moisture properties, the development of cross-sensor satellite data fusion techniques, and satellite measurements of microwave surface emissivity and other surface properties.

Dr. Jones is a Member of the American Geophysical Union, the American Institute of Physics, and the American Meteorological Society.