

TOMOGRAPHIC RECONSTRUCTION OF ATMOSPHERIC DENSITY WITH MUMFORD-SHAH FUNCTIONALS

David Ren Lara Waldrop Farzad Kamalabadi

Department of Electrical and Computer Engineering and Coordinated Science Laboratory
University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

ABSTRACT

Knowledge of the three-dimensional spatial structure of Earth's uppermost atmosphere is necessary both to understand its role as a dynamic buffer against the solar-driven environment of interplanetary space as well as to assess the rate of its permanent escape from Earth's gravity through evaporation. The only available means of inferring atmospheric structure at these altitudes is through space-based remote sensing of solar radiation that is resonantly scattered or fluoresced by the ambient atoms. In this paper, the resultant tomographic image formation problem is formulated as an edge-preserving reconstruction algorithm based on the framework originally proposed by Mumford & Shah. Statistical interpretation of this reconstruction solution is formulated in the context of MAP estimation. The numerical results illustrate that the proposed reconstruction algorithm is capable of obtaining physically meaningful solutions that are superior to previous results formulated based on parametric assumptions on the unknown density.

1. INTRODUCTION

Knowledge of the three-dimensional (3-D) spatial structure of Earth's uppermost atmosphere, which extends from several hundred to several tens of thousand kilometers, is necessary both to understand its role as a dynamic buffer against the solar-driven environment of interplanetary space as well as to assess the rate of its permanent escape from Earth's gravity through evaporation. Elemental composition in this boundary region is dominated by atomic hydrogen, H, which forms a gravitationally-bound cloud known as the geocorona. The 3-D structure of the H geocorona exhibits large-scale spatial asymmetries and regional depletions as well as temporal variability on timescales ranging from days to years.

The only available means of inferring atmospheric structure at these altitudes is through space-based remote sensing of solar radiation that is resonantly scattered or fluoresced by the geocoronal H atoms. Beyond $\sim 3 R_E$ ($1 R_E = 1$ Earth radius = 6371 km), where the number density of H atoms is relatively low, scattering events are infrequent enough that every photon detected can be assumed to have scattered exactly once. In this "optically thin" approximation, geocoronal

emission intensity at 121.6 nm (known as Ly_α), denoted here by y , measured from a planetocentric vantage \mathbf{r} along a look direction $\hat{\mathbf{n}}$, is related to the number density of the emitters, N , as follows:

$$y_i(\mathbf{r}_i, \hat{\mathbf{n}}_i) = k(\hat{\mathbf{n}}_i) \int_0^{l_\infty[\hat{\mathbf{n}}_i]} N(\mathbf{r}') dl \quad (1)$$

where the integration is over distance $l \equiv |\mathbf{r}' - \mathbf{r}_i|$ along the line of sight, k is a look-angle dependent coefficient that is a function of solar source strength and can be assumed time-independent, and $i \in \{1, 2, \dots, I\}$ is the observation index [1]. The common observation geometry is 3-D, so $\mathbf{r}, \hat{\mathbf{n}} \in \mathbb{R}^3$. Although " ∞ " is used in the integration limit, the H density field can be assumed to have finite support of $\sim 30 R_E$ [2].

Several satellite missions have been deployed over the past several decades in order to estimate the three-dimensional H density field N from an ensemble of photometric observations of Ly_α . While the measurement position \mathbf{r} varies slowly as the satellite transits the region, the photon detector typically is mounted on a rotating or otherwise spatially-scanning platform. As a result, the look direction $\hat{\mathbf{n}}$ of the column emission measurements varies significantly from sample to sample and provides common-volume sensing over some finite duration of data acquisition.

Historical attempts to infer the three-dimensional structure of the geocorona from such data have been parametric, based on an assumed functional form (e.g., [3, 4]). This model-dependent approach precludes the detection of small scale spatial structuring, such as polar depletions or localized enhancements in the H source population, and thus is insufficient to fulfill the needs of the scientific community.

In this work, we present a new means of reconstructing the geocoronal H density field through a tomographic inversion of optically thin emission intensity data. This approach avoids the dependence on parametric formulations of the unknown H density distribution. Because the tomography concept is not dependent on the dimension of the observation geometry, we illustrate the technique here in a 2-D framework without loss of generality.

2. FORWARD MODEL

In 2-D, we define the sampling variables $\mathbf{r}, \hat{\mathbf{n}} \in \mathbb{R}^2$. To solve for N from observations $y_i(\mathbf{r}_i, \hat{\mathbf{n}}_i)$ in (1), we use a polar grid to discretize the region of interest into J polar rectangles which are separated by constant radius and angle. For sufficiently large J , variation in the H-density within a given polar rectangle is negligible such that the density can be assumed to be constant. In that case, we project the function $N(\mathbf{r})$ onto J discrete orthonormal basis functions $\{N_j(\mathbf{r})\}_{j=1}^J$ such that:

$$N(\mathbf{r}) = \sum_{j=1}^J x_j N_j(\mathbf{r}) \quad (2)$$

and (1) becomes:

$$y_i(\mathbf{r}_i, \hat{\mathbf{n}}_i) = \sum_{j=1}^J \left[k(\hat{\mathbf{n}}_i) \int_0^{l_\infty[\hat{\mathbf{n}}_i]} N_j(\mathbf{r}') dl \right] x_j \quad (3)$$

Basis functions can be constructed in an arbitrary way, and we use standard basis functions in polar coordinates:

$$N_j(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} \in P_j \\ 0 & \text{else} \end{cases} \quad (4)$$

where $P_j = \{\mathbf{r} \in \mathbb{R}^2 : r_{start}^j \leq \|\mathbf{r}\| \leq r_{end}^j \ \& \ \theta_{start}^j \leq \arg(\mathbf{r}) \leq \theta_{end}^j\}$. Since the term in the square brackets in (3) is known *a-priori*, we can express (1) as:

$$y_i(\mathbf{r}_i, \hat{\mathbf{n}}_i) = \sum_{j=1}^J L_{ij} x_j \quad (5)$$

which is equivalent to the algebraic form: $\mathbf{y} = \mathbf{L}\mathbf{x}$ where \mathbf{y} is a $[I \times 1]$ vector of measurements y_i , \mathbf{x} is a $[J \times 1]$ vector of polar rectangles x_j , and \mathbf{L} is a $[I \times J]$ observation matrix. Denoting \mathbf{w} as a $[I \times 1]$ noise vector associated with each measurement, the forward model becomes:

$$\mathbf{y} = \mathbf{L}\mathbf{x} + \mathbf{w} \quad (6)$$

The inherent constraints on the observing geometry dictate a non-uniform limited-angle measurement acquisition. Furthermore, depending on the observation frequency, typically $I \neq J$, such that the observation matrix \mathbf{L} does not have full rank. Thus, existence, uniqueness, and stability of the solutions are not guaranteed, causing the inverse problem to be ill-posed. An example of such a limited-angle observation scenario is depicted in Figure 1.

3. INVERSE MODEL

3.1. Variational Method

As discussed above, rank-deficient property of the observation matrix \mathbf{L} results in an ill-posed tomography problem.

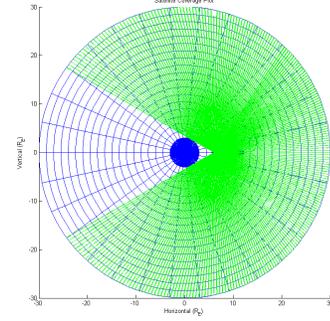


Fig. 1: Observation geometry depicting individual lines-of-sight for individual measurements from a rotating, transiting satellite.

Therefore, direct inversion to reconstruct the field is not applicable. Since noise is inherent in the measurements and the problem is ill-conditioned, least squares solution is inadequate and regularization is needed to ensure solution stability. The most commonly used technique to improve conditioning of the problem, Tikhonov regularization [5], has the quadratic objective function: $E(\mathbf{x}) = \|\mathbf{y} - \mathbf{L}\mathbf{x}\|_2^2 + \alpha^2 \|\mathbf{D}\mathbf{x}\|_2^2$ where $\|\cdot\|_p$ is the l_p norm operator, and \mathbf{D} is a first order discrete gradient operator. This variational form is equivalent to a Maximum-*a-posteriori* estimate $\hat{\mathbf{x}}_{MAP}$, assuming $\mathbf{y}|\mathbf{x} \sim \mathcal{N}(\mathbf{L}\mathbf{x}, \mathbf{I})$, $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, (\mathbf{D}^T \mathbf{D})^{-1})$. The solution $\hat{\mathbf{x}}$ is obtained under the implicit assumption that the underlying field varies smoothly and thus produces globally smooth undesirable results when the underlying field contains spatial gradients as is the case in the atmospheric H density field [6, 7, 8].

To overcome this limitation, an edge-preserving reconstruction algorithm is needed. The physical knowledge of our problem motivates Mumford-Shah (MS) functional [9] which has the form:

$$\hat{\mathbf{u}} = \arg \min_{\mathbf{u}} \int_{\Omega} (\mathbf{u} - \mathbf{f})^2 dx + \alpha_1^2 \int_{\Omega \setminus \Gamma} |\nabla \mathbf{u}|^2 dx + \alpha_2^2 |\Gamma| \quad (7)$$

where $\Omega \subset \mathbb{R}^2$ is the image domain, \mathbf{f} is an image field, and $\Gamma \subset \Omega$ is the segmentation of Ω . In this functional, the minimizer $\hat{\mathbf{u}}$ represents the piecewise smooth approximation of the original image \mathbf{f} . An edge set is used to prevent the algorithm from smoothing the edges in the image. Similarly, the first term in equation (7) is the data fidelity term, which requires image \mathbf{u} to be close to image field \mathbf{f} . The second term is the regularization term, which requires smoothness in the non-boundary part. In other words, the gradient of \mathbf{u} is added up only in the region of $\Omega \setminus \Gamma$. Therefore, changes of value in the region $\Omega \setminus \Gamma$ will be penalized. Along the edge/boundary, any arbitrary change is allowed. The third term is the second regularization term, preventing excessive edges in the reconstruction.

Due to the non-differentiability of (7), the discrete Ambrosio & Tortorelli approximation [10] of the MS functional

is used:

$$(\hat{\mathbf{x}}, \hat{\mathbf{s}}) = \arg \min_{\mathbf{x}, \mathbf{s} \in \mathbb{R}^J} \|\mathbf{y} - \mathbf{L}\mathbf{x}\|_2^2 + \alpha_1^2 \|\mathbf{D}\mathbf{x}\|_{\mathbf{W}_s}^2 + \frac{1}{2} \left(\rho \|\mathbf{D}\mathbf{s}\|_2^2 + \frac{\|\mathbf{s}\|_2^2}{\rho} \right) \quad (8)$$

where

$$\mathbf{W}_s = \text{diag}((1 - s)^2) \quad (9)$$

where $\mathbf{s} \in \mathbb{R}^J$ is an edge field corresponding to \mathbf{x} , and $\rho > 0$ is a parameter such that as $\rho \rightarrow 0$, functional (8) approaches functional (7).

Instead of having an edge set Γ , a smooth varying ‘‘edginess’’ (i.e. $s_i \in [0, 1]$) is used for the corresponding $x_i, \forall i \in \{1, \dots, n\}$. In the first regularization term, $\|\cdot\|_{\mathbf{W}_s}$ is a weighted sum, such that if an element x_i is more likely to be the edge, it will be penalized less than other elements, allowing the algorithm to preserve edges. The second regularization term demands the smoothness of the edge field s itself. Having introduced \mathbf{W}_s , which is dependent on \mathbf{s} , the objective function (8) becomes non-linear. There are many ways to minimize a non-linear functional. For instance, algorithms such as Levenberg-Marquardt Algorithm can be applied. However, convergence and computation time is not guaranteed depending on the choice of initial point. Another method is the fixed point iteration (coordinate descent method) [11], in which we fix \mathbf{x} and minimize $E(\mathbf{x}, \mathbf{s})$ with respect to \mathbf{s} , and then, we fix \mathbf{s} and construct \mathbf{W}_s to minimize $E(\mathbf{x}, \mathbf{s})$ with respect to \mathbf{x} . Namely, we set $\frac{\partial}{\partial \mathbf{x}} E(\mathbf{x}, \mathbf{s}) = \mathbf{0}$ and $\frac{\partial}{\partial \mathbf{s}} E(\mathbf{x}, \mathbf{s}) = \mathbf{0}$; the minimum has to satisfy:

$$(\mathbf{L}^T \mathbf{L} + \alpha_1^2 \mathbf{D}^T \mathbf{W}_s \mathbf{D}) \hat{\mathbf{x}} = \mathbf{L}^T \mathbf{y} \quad (10)$$

$$(\mathbf{W}_x + \alpha_2^2 \mathbf{D}^T \mathbf{D}) \hat{\mathbf{s}} = \mathbf{W}_x \mathbf{z} \quad (11)$$

where

$$\mathbf{W}_x = \text{diag}(\alpha_1^2 (\mathbf{D}\mathbf{x})^2 + \alpha_3^2) \quad (12)$$

$$\mathbf{z} = \mathbf{W}_x^{-1} (\alpha_1^2 (\mathbf{D}\mathbf{x})^2) \quad (13)$$

For $\mathbf{v} \in \mathbb{R}^{n \times 1}$, \mathbf{v}^2 is defined to be the element wise square. The fixed point iteration converges to a local minimum in a linear convergence rate [12]. Then, to solve for (10) and (11), we can use pre-conditioned conjugate gradient iterative method, thus solving the overall equation (8).

3.2. Maximum-a-posteriori (MAP) Estimation

In this subsection we formulate the solution of MS functional as a solution to a Maximum-a-posteriori problem. The solution is:

$$(\hat{\mathbf{x}}, \hat{\mathbf{s}}) = \arg \max_{\mathbf{x}, \mathbf{s} \in \mathbb{R}^J} (p(\mathbf{y}|\mathbf{x}, \mathbf{s}) p(\mathbf{x}, \mathbf{s})) \quad (14)$$

Assuming Bayesian statistics for all image, edge field, and noise, we have:

$$\begin{aligned} (\hat{\mathbf{x}}, \hat{\mathbf{s}}) &= \arg \min_{\mathbf{x}, \mathbf{s} \in \mathbb{R}^J} (-\log p(\mathbf{y}|\mathbf{x}, \mathbf{s}) - \log p(\mathbf{x}, \mathbf{s})) \\ &= \arg \min_{\mathbf{x}, \mathbf{s} \in \mathbb{R}^J} \|\mathbf{y} - \mathbf{L}\mathbf{x}\|_2^2 + \underbrace{\|\mathbf{D}\mathbf{x}\|_{\mathbf{W}_s}^2 + \|\mathbf{D}\mathbf{s}\|_2^2 + \|\mathbf{s}\|_2^2}_{\propto -\log p(\mathbf{x}, \mathbf{s})} \end{aligned} \quad (15)$$

For simplicity, and without loss of generality, we drop the coefficients in (8). When expanding the regularization terms in (15), denoted by R , they become:

$$\begin{aligned} R &= \|\mathbf{D}\mathbf{x}\|_{\mathbf{W}_s}^2 + \|\mathbf{D}\mathbf{s}\|_2^2 + \|\mathbf{s}\|_2^2 \\ &= \mathbf{x}^T \mathbf{D}^T \mathbf{W}_s \mathbf{D} \mathbf{x} + \mathbf{s}^T \mathbf{D}^T \mathbf{D} \mathbf{s} + \mathbf{s}^T \mathbf{s} \end{aligned} \quad (16)$$

Now, we define:

$$\begin{aligned} \Sigma_1^{-1} &= \mathbf{D}^T \mathbf{W}_s \mathbf{D} \\ \Sigma_2^{-1} &= \mathbf{D}^T \mathbf{D} + \mathbf{I}_J \\ \mathbf{v} &= \begin{bmatrix} \mathbf{x} \\ \mathbf{s} \end{bmatrix} \end{aligned} \quad (17)$$

By combining all terms in (16),

$$R = \mathbf{v}^T \Sigma^{-1} \mathbf{v} \quad (18)$$

where we define $\Sigma^{-1} = \begin{bmatrix} \Sigma_1^{-1} & \mathbf{0} \\ \mathbf{0} & \Sigma_2^{-1} \end{bmatrix}$. Therefore,

$$p(\mathbf{v}) = p(\mathbf{x}, \mathbf{s}) = k^{-1} \exp\left(-\frac{1}{2} \|\mathbf{v}\|_{\Sigma^{-1}}^2\right) \quad (19)$$

where k is a normalization constant.

If \mathbf{v} is an image with an edge field, the prior favors those that are piece-wise smooth and possess limited amount of edges. A future direction is to incorporate not only the general ‘‘characteristics’’ of the underlying image, but also to learn the statistical distribution or moments of the priors based on available training data.

4. NUMERICAL EXPERIMENTS

4.1. Synthetic data

In principle, it is possible to design an observation (acquisition) geometry and sampling scheme that would yield optimal reconstruction results with respect to some optimality criterion for a given algorithm. In practice, however, due to the prohibitive cost of a dedicated space mission, it is more realistic to take advantage of a mission of opportunity with a predetermined orbital configuration. Such an observation geometry, which is based on the trans-lunar trajectory of the upcoming Explorer Mission 1 (EM-1) spacecraft, is used in the numerical experiments presented in this section.

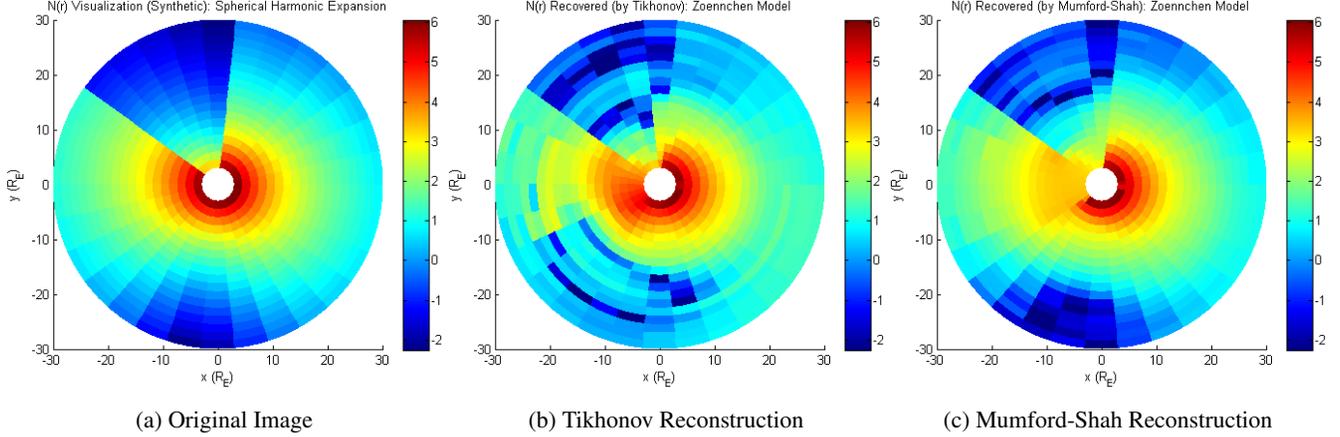


Fig. 2: These are the original image (left), image reconstructed using Tikhonov regularization (middle), and image reconstructed using Mumford-Shah regularization (right).

The ensemble of simulated Ly_{α} emission intensity data is acquired as the satellite transits the 2-D density field from $6-12 R_E$, along a straight line oriented radially away from Earth. The observing positions, denoted \mathbf{r}_i in (1), are spaced along this trajectory every 630 km. At a given position, 300 measurements are acquired along lines-of-sight evenly spaced in the 2-D plane every 1.2° , corresponding to a 1 second integration by a detector that is rotating with a 5 minute period. These sampling parameters are consistent with typical space-based Ly_{α} detector capabilities and satisfy threshold requirements for sufficient photon counting rates. Figure 1 depicts the sampling geometry as green lines-of-sight emanating from the viewing positions. Because the formulation of (1) assumes optically thin conditions which only exist beyond $3 R_E$, lines-of-sight passing within this radius are not included in the sample set.

The synthetic H density data is based on the empirical model reported by Zoenchen et al. [4], which was derived as a parametric fit of optically thin Ly_{α} emission intensity data acquired from the orbiting NASA TWINS satellites. The [4] model assumes that the functional form of the 3-D distribution at a given distance from Earth is a spherical harmonic expansion up to the second order whose coefficients decay exponentially with increasing distance from Earth. Our synthetic model corresponds to the reported parametric fit evaluated in the 2-D meridional plane at the center points of the polar rectangles.

In addition to the smoothly varying synthetic model, we also consider the presence of a regional density depletion in order to compare the edge preserving properties of Tikhonov and Mumford-Shah regularization. Densities in the depletion region are reduced by 90% relative to their un-modified values. The annulus of $3R_E \leq r \leq 30R_E$ is divided into $18 \times 30 = 540$ polar rectangles, as shown in Figure 2a. The resolution is chosen in such a way so that it is not too high to cause overfitting, and it is not too low to be unable to repre-

sent the underlying image.

4.2. Reconstruction Results

Figure 2 shows the reconstruction results using the Mumford-Shah method in comparison with Tikhonov regularization and illustrates its superior performance when significant structure is present in the underlying density field. Figure 3 shows the residual errors of the two reconstruction techniques. In general, the error of estimation is smaller in the region where the satellite’s trajectory allows dense sampling.

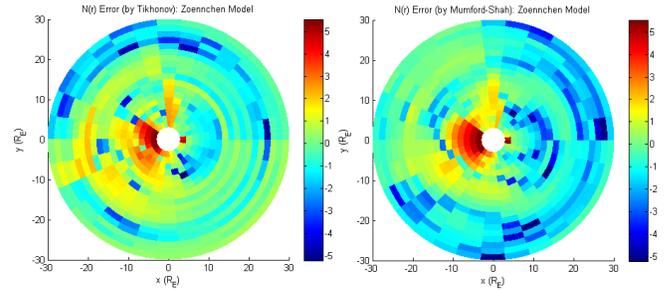


Fig. 3: Plots of log of Residuals when reconstructing using Tikhonov and Mumford-Shah.

5. CONCLUSION

The reconstruction results using MS regularization in (7) shows strong ability to reveal edges in the underlying unknown field, which is not practical by using parametric models. Therefore, this algorithm allows model-independent tomographic inversion to reconstruct geocoronal H density field. Future work will include learning prior distribution of the field and incorporating this information in the reconstruction process.

6. REFERENCES

- [1] R.R. Meier, "Ultraviolet spectroscopy and remote sensing of the upper atmosphere," *Space Sci. Rev.*, vol. 58, pp. 1–185, 1991.
- [2] R. R. Hodges, Jr., "Monte Carlo simulation of the terrestrial hydrogen exosphere," *Journal of Geophysical Research*, vol. 99, pp. 23229, Dec. 1994.
- [3] J Bailey and M Gruntman, "Experimental study of exospheric hydrogen atom distributions by lyman-alpha detectors on the twins mission," *Journal of Geophysical Research: Space Physics (1978–2012)*, vol. 116, no. A9, 2011.
- [4] JH Zoennchen, U Nass, and HJ Fahr, "Exospheric hydrogen density distributions for equinox and summer solstice observed with twins1/2 during solar minimum," *Ann. Geophys*, vol. 31, pp. 513, 2013.
- [5] Tikhonov A., "Solution of incorrectly formulated problems and the regularization method," *Soviet Math. Dokl.*, vol. 5, pp. 1035/1038, 1963.
- [6] G. E. Thomas and A. Vidal-Madjar, "Latitude variations of exospheric hydrogen and the polar wind," *Planetary and Space Science*, vol. 26, pp. 873–882, 1978.
- [7] T. E. Moore, A. P. Biddle, J. H. Waite, Jr., and T. L. Killeen, "Auroral zone effects on hydrogen geocorona structure and variability," *Planetary and Space Science*, vol. 33, pp. 499.
- [8] J Bailey and M Gruntman, "Observations of exosphere variations during geomagnetic storms," *JGeophysical Research Letters*, vol. 40, pp. 1907–1911, 2013.
- [9] David Mumford and Jayant Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on pure and applied mathematics*, vol. 42, no. 5, pp. 577–685, 1989.
- [10] Luigi Ambrosio and Vincenzo Maria Tortorelli, "Approximation of functional depending on jumps by elliptic functional via t-convergence," *Communications on Pure and Applied Mathematics*, vol. 43, no. 8, pp. 999–1036, 1990.
- [11] Stephen J. Wright, "Coordinate descent algorithms," *Math. Program.*, vol. 151, no. 1, pp. 3–34, June 2015.
- [12] J. C. Bezdek, R. J. Hathaway, R. E. Howard, C. A. Wilson, and M. P. Windham, "Local convergence analysis of a grouped variable version of coordinate descent," *J. Optim. Theory Appl.*, vol. 54, no. 3, pp. 471–477, Sept. 1987.