An alternative concept for SMOS-HR: unfolding the brightness temperature map by along-the-track inversion of the Van Cittert-Zernike equation

Max Dunitz, Hugo Marsan, Clement Monnier, Eric Anterrieu, Francois Cabot, Ali Khazaal, Nemesio Rodriguez-Fernandez, Bernard Rouge, Yann Kerr, Jean-Michel Morel, et al.

An alternative concept for SMOS-HR: unfolding the brightness temperature map by along-the-track inversion of the Van Cittert-Zernike equation

Max Dunitz* 0000-0002-3244-1244
Hugo Marsan* 0000-0003-1179-3432
Clément Monnier* 0000-0002-1700-8709
Eric Anterrieu† 0000-0002-8825-9652
François Cabot† francois.cabot@cesbio.cnes.fr
Ali Khazaal† khazaala@cesbio.cnes.fr
Nemesio Rodriguez-Fernandez† 0000-0003-3796-149X
Bernard Rougé† rougebe@gmail.com
Yann Kerr† 0000-0001-6352-1717
Jean-Michel Morel* 0000-0002-6108-897X
Miguel Colom* 0000-0003-2636-0656

***Abstract*—Launched in 2009, the SMOS satellite [1] produces observations of brightness temperature through passive microwave radiometry in the protected radio astronomy portion of the L-band. Although a given point on the Earth's surface may be visible for 100 consecutive correlator integration times, brightness temperature measurements are produced one correlation period at a time, producing temperature maps that are both noisy and folded due to the antenna array's undersampling of the u-v frequency plane.**

For a potential successor of SMOS, SMOS-HR [2], we show that a global inversion of the observation model (based on the Van Cittert-Zernike theorem) across the orbital trace could simultaneously unfold and denoise the observations. To this end, we take advantage of the shift-invariance of the inversion problem in geodesic coordinates.

I. The global inversion in geodesic coordinates associated with the satellite's trajectory

To introduce the mathematical method in a slightly simplified geometric setup, we model the Earth's surface as a sphere of radius R and suppose the satellite orbits a great circle[1] at altitude h. We call the satellite's trace the curve drawn by the subsatellite point, that is, the point of the Earth at nadir. With these assumptions, the trace is a closed geodesic. We arbitrarily fix a point of reference on the trace; for any point along the trace, we denote by y its geodetic distance along the trace to the reference point, that is, the arc length of this geodesic divided by R. Since the satellite moves at constant speed, y is proportional to the time elapsed since the satellite last passed over the reference point. Define the x-axis as a great circle passing through the trace orthogonally at the reference point at a right angle and x as the angular distance to the trace along that great circle. We may think of y as the ordinate and x as the abscissa of such a geographical coordinate system. This coordinate system is thereby determined by the satellite's trace and a reference point on that trace: the abscissa's magnitude equals the curvilinear distance to the trace as one moves across the Earth's surface along a line perpendicular to the trace, and the ordinate equals the curvilinear distance to the line orthogonal to the trace. We say that the satellite is "in position $(0, y)$" when its subsatellite point is $(0, y)$ with respect to these coordinates, and we call (x, y) the "geodesic coordinates associated with the trajectory."

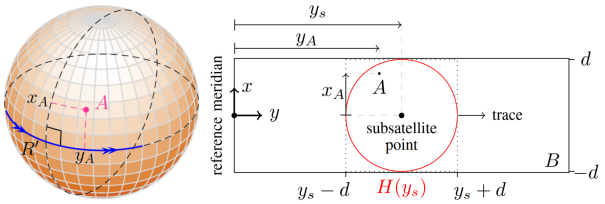

Fig. 1. Left: The spherical model S of the Earth. The satellite's trace, given in blue, locally approximates a great circle. R' is a reference point along the trace, with respect to which geodesic coordinates associated with the trajectory may be defined. For any point $A \in B$, the visible band of S, its geodesic angular coordinates are the angular distances across-track (x_A) and along-track (y_A) to R'. Right: The geodesic coordinates in $\mathbb{R}^2$. The set of points visible when the satellite is in position $(0, y_s)$ is given by $H(y_s) = \{(x, y) \in [-d, d] \times [y_s - d, y_s + d] \mid \cos(x)\cos(y - y_s) \geq \frac{R}{R+h}\}$. Points near the center of the band $B = \cup_{y \in \mathbb{R}} H(y)$ (those with $|x| \approx 0$) are visible in many snaphots from a variety of incidence angles $\theta(x, y; y_s) = \sin^{-1} \frac{R}{\rho}\sqrt{1 - \cos^2 x \cos^2(y - y_s)}$ (where the distance from $p(x, y)$ to the satellite in position $(0, y_s)$ is $\rho(x, y; y_s) = \sqrt{R^2 + (R+h)^2 - 2R(R+h)\cos(x)\cos(y - y_s)}$), whereas those with $|x|$ near d are visible in few snapshots, at exclusively high incidence angles, near $\sin^{-1}\left(\frac{R}{R+h}\right)$.

As the subsatellite point traces a geodesic on the spherical surface S, in the geodesic coordinate plane $\mathbb{R}^2$, the set of points visible to the satellite forms the band $B = [-d, d] \times \mathbb{R}$, where $d = R\cos^{-1}\left(\frac{R}{R+h}\right)$ is the horizon in geodesic coordinates. The application $p : B \to S$ locates the point with those coordinates on Earth's surface. The point $p(x, y) \in S$ has a signed geodesic distance to the trace equal to x. The point on the trace closest to $p(x, y)$ is $p(0, y)$. From Napier's laws of spherical trigonometry, elements of the set $H(y') \subsetneq [-d, d] \times [y' - d, y' + d] \subsetneq B$ of points

*Centre Borelli, ENS Paris-Saclay, Gif-sur-Yvette, France.
†Centre d'Etudes Spatiales de la Biosphère (CESBIO), Toulouse, France.

[1]At the resolution scale of SMOS, conversion between spherical and oblate elliptical models of the Earth calls for only a subpixel correction, and the extent to which the orbital spiral deviates from a great circle across a segment of the trace corresponding to a single global inversion (less than the orbital period) is minor and can profitably be ignored for the proof of concept that constitutes our present work.

$p(x, y)$ visible to a satellite in position $(0, y')$ must satisfy $\cos(x)\cos(y - y') \geq \frac{R}{R+h}$.

The geodesic coordinate system is depicted in Figure 1.

A. The observation model in geodesic coordinates

Radio interferometric measurements, such as the SMOS Level 1A data, are called *visibilities* and consist of the complex correlations between the signals received on each pair of antennas. The Van Cittert-Zernike theorem, which states that there is a Fourier transform relationship between visibilities and the intensity image of radio sources, permits us to produce these images from the visibilities, either by taking an inverse transform and removing instrument artifacts from the "dirty" intensity image, or by inverting the system of equations associated with the observation. The latter approach, though more computationally expensive, is better suited to our application.

We use a simplified version[2] of the observation model from Corbella et al. [5], which makes use of the Rayleigh-Jeans approximation of the Planck equation to give the visibilities in terms of the brightness temperature. For an instrument with M antennas, the visibility V_{kl} between antennas k and l may be written

$$V_{kl} = \int\limits_{||\boldsymbol{\xi}||\leq 1} \mathbf{T}'(\boldsymbol{\xi}) h_{kl}(\boldsymbol{\xi}) e^{-2\pi i \langle \mathbf{u}_{kl}, \boldsymbol{\xi} \rangle} \, \mathrm{d}S(\boldsymbol{\xi}) \,, \qquad (1)$$

for $k, l \in \mathbb{Z}/M\mathbb{Z}$, where $\boldsymbol{\xi}$ is the vector of cosine directions (ξ, η)—sometimes given (l, m)—along two orthogonal directions of the planar antenna array; $h_{kl}(\boldsymbol{\xi}) = A_k(\boldsymbol{\xi})\overline{A_l}(\boldsymbol{\xi}) e^{-2\pi i \langle \mathbf{u}_{kl}, \boldsymbol{\xi} \rangle}$ is the product of the kth antenna's radiation pattern,[3] the conjugate of the lth antenna's radiation pattern, and the baseline-dependent complex exponential term; $\mathrm{d}S(\boldsymbol{\xi}) = \frac{\mathrm{d}\xi \mathrm{d}\eta}{\sqrt{1 - ||\boldsymbol{\xi}||^2}}$ is the solid angle differential; $\boldsymbol{u}_{kl}$ is the 2D vector in wavelengths from antenna k to l in the u-v plane; $\langle \cdot \rangle$ and $||\cdot||$ are the Euclidean inner product and norm over $\mathbb{R}^n$; and $\mathbf{T}'$ maps each direction (ξ, η) to its brightness temperature. Equation (1) can be understood as a Fourier transform if we extend the product $\mathbf{T}'(\boldsymbol{\xi}) h_{kl}(\boldsymbol{\xi})$ to all of $\mathbb{R}^2$ and take its value to be zero at undefined directions (those for which $||\boldsymbol{\xi}|| > 1$).

When the image of sky brightness temperatures is known, we can work with the corrected visibilities obtained by subtracting out the effect of radio sources in the sky and space, thereby shrinking the domain of integration to $||\boldsymbol{\xi}|| \leq c$ where $c < 1$ is the cosine-direction radius associated with the observation horizon. With our spherical model of Earth, $c = \frac{R}{R+h}$ and the visibilities observed by our satellite in position $(0, y_s)$ is given by integrating over $H(y_s)$:

$$V_{kl}(y_s) = \int\limits_{||\boldsymbol{\xi}||\leq c} \mathbf{T}(x(\boldsymbol{\xi}), y_s + y(\boldsymbol{\xi})) h_{kl}(\boldsymbol{\xi}) \, \mathrm{d}S(\boldsymbol{\xi}) \,, \qquad (2)$$

where $\mathbf{T}(x, y)$ gives the brightness temperature at $p(x, y)$ (the second argument understood to be relative to the reference meridian) and each direction $\boldsymbol{\xi}$ that intersects with Earth's surface does so at geodesic coordinates *relative to the subsatellite point* of $(x(\boldsymbol{\xi}), y(\boldsymbol{\xi}))$.

B. Modeling brightness temperature

In the L-band, observed land and sea surface brightness temperature is responsive to soil moisture and ocean salinity. Using the brightness temperature, surface temperature, incidence angle, and the type of surface under surveillance, these ECVs can be recovered using theoretical models. The use of such models in the context of satellite monitoring poses overfitting issues due to the low spatial resolution, leading to high heterogeneity of environmental conditions across each pixel.

The best way to model this brightness temperature, we argue, when observing radiation emitted from a mixture of many environments, each with its own highly parameterized physical model, is to jettison the local physics in favor of a much lower-variance model with few parameters. Thus, we propose a model where each point $p(x, y)$ has at most m parameters associated with it, $\alpha_1(x, y), \ldots, \alpha_m(x, y)$. Since points $p(x, y)$ with low x-coordinate values are observed at a greater diversity of incidence angles than those with high x-coordinate values, the number of parameters—and basis functions—may vary with the value of the across-track coordinate, x. These parameters are inputs to a model that governs the relationship between the brightness temperature and the incidence angle[4] θ from $p(x, y)$ to the satellite. The incidence angle $\theta(\boldsymbol{\xi})$ is related to the received brightness temperature for a variety of reasons, most notably by the amount of sky through which radiation must pass en route to the satellite and through the larger pixel sizes, in surface area—and thus potentially more heterogeneous observations—near the observation horizon.

This generic brightness temperature model reads

$$\begin{aligned} \mathbf{T}(x, y, \theta) = f_0(\theta) + \alpha_1(x, y) f_1(\theta) + \ldots \\ + \alpha_m(x, y) f_m(\theta) = \langle \alpha(x, y), f(\theta) \rangle, \end{aligned} \qquad (3)$$

where $f = (f_i)_{i=0 \cdots m}$ are the basis functions of the model.

When the satellite is in position $(0, y)$, at each cosine direction $\boldsymbol{\xi}$ below horizon ($||\boldsymbol{\xi}|| < c$), the satellite observes the point $p(x(\boldsymbol{\xi}), y + y(\boldsymbol{\xi}))$. This means, physically, that the surface of Earth at $p(x, y)$ emits its brightness temperature toward the satellite in the angle $\theta(\boldsymbol{\xi})$.

C. Decomposing a global inversion of the visibilities into inversions of each along-the-track frequency

We substitute our temperature model from Equation (3) into Equation (2) and tweak the limits of integration to

[2] We use Equation 7 in the reference. Here we ignore the polarization properties of the electric field. We do not incorporate coupling from the incident wave (**a** in [5]) into the observation model; to do so, we could simply subtract out from the brightness temperatures $\mathbf{T}$ in Equation (1) the physical temperature of the instrument, which for isolated antennas is the noise temperature, as is done by Corbella et al. in their Equation 18. For simplicity, we also ignore the fringe washing function, which arises from the bandpass filter before the correlator, and which can be folded into h_{kl}. The complete observation model, directing attention to the coupling and filtering, is presented in Equations 29-30 of [5]. The method presented in the present article can assimilate these details without trouble.

[3] Or rather, the radiation pattern of the entire instrument experienced at the kth antenna's port.

[4] In situations where radial symmetry assumption introduces too much bias, we might instead use basis functions $f(\boldsymbol{\xi})$ that depend on both cosine directions in an arbitrary fashion, rather than as mediated through the incidence angle, that is, rather than $f(\theta(\boldsymbol{\xi}))$.

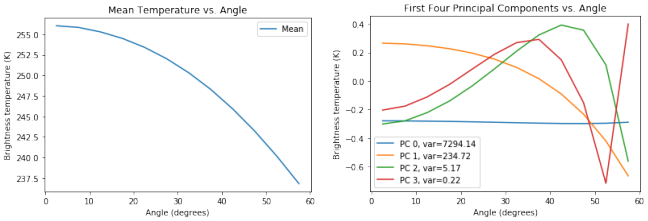

Fig. 2. Parametric representation of brightness temperatures trained using level 3 SMOS data taken from several ascending and descending orbits over several days, over 12 incidence angle bins (each bin in [6] besides those centered at 40° and 62.5°). These data give, for each bin, vertical and horizontal polarization brightness temperature, as well as the RMSE between the modeled and observed brightness temperatures. Using principal component analysis (PCA), we verified that binned angular brightness temperature data are highly redundant. Just four principal components explain almost all the variance. The PCA model given in the two images above was used in this paper to express the brightness temperature of a point $p(x, y)$ as a function of the incidence angle θ at which it is observed and its brightness temperature parameters $\alpha(x, y)$: $\mathbf{T}(x, y, \theta) = T_{mean}(\theta) + \sum_{i=0}^{3} \sqrt{\text{var}_{PCi}} T_{PCi}(\theta)\alpha_i(x, y)$, where $T_{mean}(\theta)$ (left image) is the mean temperature of pixels observed at incidence angle θ and the $T_{PCi}(\theta)$ (right image) are the normalized (orthonormal on $\mathbb{R}^{12}$, not over the set of points H') basis functions (covariance matrix eigenvectors visualized as temperature curves rather than just vectors in $\mathbb{R}^{12}$). The first basis function is relatively constant (thus, α_1 adjusts pixel brightness across all angles); the second tracks the mean temperature curve; and the remaining allow adjustments to temperature in particular incidence angle bands.

indicate these visibilities have corrected by subtracting out the radiation in $c < ||\boldsymbol{\xi}|| \leq 1$.

The contributions to the visibilities from the visible surface of Earth, therefore, can be expressed as

$$V_{kl}(y_s) = \int\limits_{||\boldsymbol{\xi}||\leq c} \langle \alpha(x(\boldsymbol{\xi}), y_s + y(\boldsymbol{\xi})), f(\theta(\boldsymbol{\xi}))\rangle h_{kl}(\boldsymbol{\xi})\, \mathrm{d}S(\boldsymbol{\xi})\,. \tag{4}$$

Thanks to the satellite's yaw corrections, the acquisition geometry does not vary across the orbit. In other words, relative to the subsatellite point, $H(y_s)$ describes the same set of points as the subsatellite point moves along the trace. The set $H' = \{(x, y - y_s) \mid (x, y) \in H(y_s)\}$ is stationary. This implies that the model of Equation (4) is stationary along the trace, and this manifests itself formally in the fact that it is a convolution with respect to the variable y_s.

To take advantage of this fact, we apply to this equation the partial Fourier transform with respect to y_s. For a function of two variables $u(x, y)$, this transform is defined by

$$\widetilde{u}(x, \omega) := \int_{-\infty}^{\infty} u(x, y)e^{-i\omega y}\, \mathrm{d}y\,.$$

Now the partial Fourier transform of the corrected visibilities, $\widetilde{V}_{kl}(\omega) = \int_{-\infty}^{\infty} V_{kl}(y)e^{-i\omega y}\, \mathrm{d}y$, can be written in terms of the partial Fourier transform of the parameters, $\widetilde{\alpha}(x, \omega) = \int_{-\infty}^{\infty} \alpha(x, y)e^{-i\omega y}\, \mathrm{d}y$. Letting

$$q(x, y) = h_{kl}(\boldsymbol{\xi}(x, y))(1, f_1(\theta(x, y)), \ldots, f_m(\theta(x, y))),$$

we see, through a simple change of variable $y' = y_s + y(\boldsymbol{\xi})$ and the stationarity of the geometry, that

$$\begin{aligned}\widetilde{V}_{kl}(\omega) &= \int_{-\infty}^{\infty} \left(\iint\limits_{H(y_s)} \langle \alpha(x, y_s + y), q(x, y)\rangle\, \mathrm{d}S \right) e^{-i\omega y_s}\, \mathrm{d}y_s \\ &= \iint\limits_{||\boldsymbol{\xi}||<c} \langle \widetilde{\alpha}(x(\boldsymbol{\xi}), \omega)e^{-i\omega y(\boldsymbol{\xi})}, q(x(\boldsymbol{\xi}), y(\boldsymbol{\xi}))\rangle\, \mathrm{d}S\,(\boldsymbol{\xi}).\end{aligned} \tag{5}$$

A global inversion of Equation (5) to recover the brightness temperature parameters from all visibilities measured along an orbit can therefore be decomposed into smaller inversions, recovering independent subsets of the parameters. For each orbital frequency $\omega \in \mathbb{R}$, we can invert Equation (5) to recover the set of transformed parameters $\widetilde{\alpha}(x(\boldsymbol{\xi}), \omega)$. More precisely, for each satellite position y, we are given M^2 visibilities, derived from the M^2 equations (given in Equation (4)), which we seek to invert to estimate the brightness temperature parameters $\alpha(x, y)$. After transforming Equation (4) into Equation (5), we still have M^2 equations, with the M^2 transformed visibilities known. However, these equations now have, as unknowns, the functions $x \to \widetilde{\alpha}(x)$. They can now be resolved independently for each ω! Having done this, we can apply the inverse partial Fourier transform to $\widetilde{\alpha}(x, \omega)$ to arrive at $\alpha(x, y)$.

D. Discretization

We pass to the discrete setting because the visibilities are computed not in a continuous fashion but rather once per correlator integration period and because we wish to invert Equation (5) over a finite segment of the orbital trace (small enough to resemble a great circle). The above analysis, based on the invariance of the acquisition geometry across the orbital segment, readily informs the discrete setting. The partial Fourier transform trick becomes the partial discrete Fourier transform trick.

Inverting Equation (5) requires discretizing the set of directions that intersect Earth to form a Riemann sum over the set of $\boldsymbol{\xi}$ for which $||\boldsymbol{\xi}|| \leq c$. Equivalently, this can be seen as discretizing $H' = \{(x(\boldsymbol{\xi}), y(\boldsymbol{\xi})) \mid ||\boldsymbol{\xi}|| \leq c\}$ and ultimately the band B. After subtracting the contribution of the mean temperature function over the scene $f_0(\theta)$ from the visibilities, discretizing the integral in Equation (5) relates the partial Fourier transform $\widetilde{V}_{kl}$ of the visibilities, evaluated at each orbital frequency ω, to the partial Fourier transform $\widetilde{\alpha}(x, \omega)$ of the maps of brightness temperature parameters via a matrix multiplication.

Let $N_{H'}$ be the number of distinct points in our sampling of H' and N_{lat} be the number of distinct values of $x(\boldsymbol{\xi})$. Inverting Equation (5) involves computing the pseudo-inverse of an $M^2 \times (m \cdot N_{H'})$ matrix, referred as the G matrix.[5]

E. Comparison with similar approaches in radio astronomy

In 1979, Ekers and Rots [3] introduced into the radio interferometry literature the idea of taking the Fourier

[5]Since ξ and η each depend on both geodesic coordinates x and y (via $\rho(x, y)$), to integrate out the y and render the equation purely in terms of the unknown $x \to \widetilde{\alpha}(x)$, we must rewrite $\mathrm{d}S(\boldsymbol{\xi})$ as $\mathrm{d}S(x, y) = \frac{R^2 \cos(x)(R-(R+h)\cos(x)\cos(y))}{\rho(x,y)^3}\, \mathrm{d}y\, \mathrm{d}x$. When we do this, the G matrix shrinks to $M^2 \times (m \cdot N_{lat})$, where $N_{lat} = O(\sqrt{N_{H'}})$.

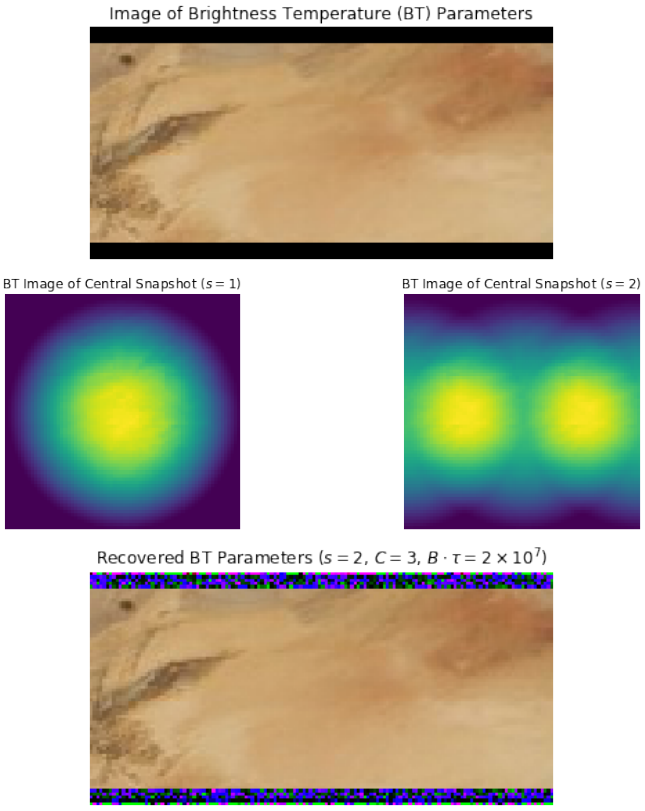

Fig. 3. At the top, we have a three-channel image of brightness temperature parameters. When the subsatellite point lies in the center of this image, the map of brightness temperature, uniformly sampled and occupying the square $[-1, 1] \times [-1, 1]$ in the ξ-η plane, is as given in the center row, left column. However, due to undersampling of the visibilities in the u-v plane, a direct inversion of the visibilities to recover the brightness temperature map associated with the snapshot yields a folded image, as in the image in the center row, right column. This folding invades the entire field of view; the brightness temperature parameters are everywhere indeterminate from a single snapshot alone. Only by inverting a system of snapshots—either by a direct inversion of the convolution mechanism (one large, invertible matrix), or by using the "partial Fourier transform" trick to decompose this computation orbital frequency-by-orbital frequency—can we recover an image of the brightness temperature parameters, as in the image in the bottom row. Note that the top and bottom five rows of the parameter image, which correspond to the sky and space (modeled here as a constant 3 K from any angle), are poorly reconstructed because in most snapshots they are not visible ($||\boldsymbol{\xi}|| > 1$) and even when they are visible the modified brightness temperature (that is, the brightness temperature scaled by the antenna product and obliquity factor) is extremely low.

transform along the direction cosines corresponding to the pointing directions of an interferometer moving between phase centers. Pety and Rodriguez-Fernandez [4] used this technique in wide field imaging, fusing images from multiple pointings via a convolution in the u-v plane. Owing to the different geometry of our application (with all but the shortest orbital segments), such an approach to image fusing is not possible here. Our approach differs not just in the geometry but in the way images are fused, via the transform itself and not via a convolution in the transform domain.

II. Numerical validation of unfolding using the "partial Fourier transform" trick

In this section, we demonstrate numerically that the "partial Fourier transform trick" along with the global inversion strategy can simultaneously denoise a map of brightness temperature parameters and expand the alias-free field of view.

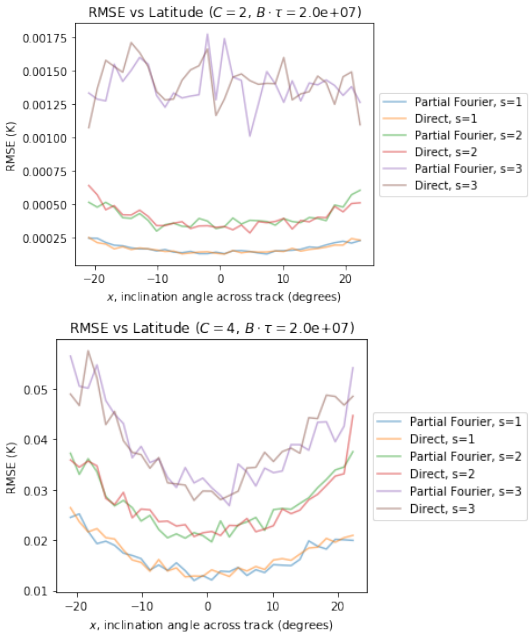

Fig. 4. A comparison of the direct and partial Fourier inversion methods, with the SMOS noise model and the swath from Figure 5. Across all geodesic coordinates x, reconstruction error is higher for a brightness temperature model with four principal components (bottom) than it is for one with two (top). Error also increases with the factor s by which the visibilities were decimated. Reassuringly, reconstruction error did not depend on the choice of inversion technique. Note that, in these figures, the set of visibilities each method used to invert had different noise samples though of the same variance. When given the same set of visibilities, with identical noise samples, images of parameters recovered via the two methods are identical, up to a precision of between 10^{-13} and 10^{-10} (see the second image in Figure 5), depending on the probed values of brightness temperature parameters C as well as the decimation factor s and correlation bandwidth-time product $B \cdot \tau$—and thus the noise level.

For scenarios where the modified brightness temperatures and visibilities form a DFT pair, we can decimate the visibilities along v-axis to introduce a highly structured pattern of aliasing that can be directly modeled and unfolded.[6] Though there is no alias-free portion of each brightness temperature snapshot, we have enough snapshots in our simulation to recover the brightness temperature parameters via two global inversion techniques: either by directly unfolding the linear system that acts on the brightness temperature parameters to generate the folded brightness temperature snapshots, or by using the partial Fourier trick introduced in the previous section.

A. Obtaining an explicit model of along-track folding

We systematically decimate the visibilities to induce a regular pattern of aliasing, which we call "folding." We can obtain an explicit model of the aliasing because decimation and folding are transform pairs. In particular, decimating the DFT of an M-point signal by a factor of s, where s divides M, induces s-fold folding in the time-domain signal.

[6]It is not essential for the partial Fourier method that the undersampling of baselines be limited to the along-track direction. Undersampling of visibilities in a regular manner only along the v-axis ensures that folding in the brightness temperature snapshots occurs only along latitude lines. This simplification enables the decomposition of the global direct inversion into a series of latitude-by-latitude inversions, making the direct inversion technique easier to specify and more computationally tractable.

Lemma II.1. *Let x be an M-point discrete signal and X its M-point DFT. Suppose s divides M. Consider the $\frac{M}{s}$-point discrete signal x_f whose DFT is X, decimated by the factor s: for $\omega \in \mathbb{Z}/\frac{M}{s}\mathbb{Z}$, $X_f[\omega] = X[s\omega]$. We may write x_f as follows:*

$$x_f[m] = x[m] + x[m+s] + \ldots + x\big[m + \frac{s-1}{s}M\big]$$

for $m \in \mathbb{Z}/\frac{M}{s}\mathbb{Z}$.

Let I be a C-channel, $M \times N$ image of the C brightness temperature parameters α of each point p of a visible band of the Earth's surface. We suppose the subsatellite point begins at the middle of the leftmost edge of the image[7] and that the grid of sampled directions moves one pixel to the right each snapshot, as in Figure 1.[8] The kth image of brightness temperatures observed, then, is $T_k = \sum_{c=0}^{C-1} I_k^c \odot Q_c$, where I_k^c is the image of the cth parameter of the kth snapshot[9] and Q_c is the Hadamard product ($\odot$) of the antenna radiation patterns,[10] the cth principal component of the brightness temperature model scaled by its root variance, and the modified obliquity factor.[11] The $M \times M$ image of visibilities of the kth snapshot are produced by taking the two-dimensional, M-point DFT of the temperature image T_k, then setting each column to zero except for columns $0, s, \ldots, \frac{(s-1)M}{s}$. To this $M \times M$ image of visibilities, we add complex additive white Gaussian noise with standard deviation $\sigma = V_{DC}/\sqrt{2B\tau}$, where V_{DC} is the visibility corresponding to the $(0,0)$ baseline, B is the bandwidth and τ the correlation time. (We typically use SMOS values $B = 20$ MHz and $\tau = 1$ second, though in Figure 5 we reduced the time-bandwidth averaging by a factor of 100 to produce distinguishable brightness temperature curves.)

From the set of noisy, decimated simulated visibilities associated with any snapshot number k, we can recover the $M \times M$ image of observed brightness temperatures $\widetilde{T}_k$ by taking the two-dimensional, M-point inverse DFT. As the rightmost image in Figure 3 indicates, a highly structured—and, when adequately observed, invertible—form of aliasing invades the full $M \times M$ image of brightness temperatures. The brightness temperature parameters of the pixels visible in this snapshot can be recovered only through a global inversion that uses multiple observations. In effect, we have sampled the visibilities at $\frac{M^2}{s}$ distinct baselines, across N snapshots, generated from brightness temperatures governed by CMN unknown parameters. With $N = 2M$, we have $\frac{2}{s}M^3$ equations and $2CM^2$ unknowns, which can be inverted latitude by latitude (the direct global inversion) or orbital frequency by orbital frequency (the partial Fourier inversion), but not snapshot by snapshot (as in SMOS) due to the lack of alias-free field of view.

B. The direct global inversion

Fix row number r and consider the one-dimensional, M-point, $\frac{M}{s}$-periodic signal $S_k^r = \widetilde{T}_k[r,:]$ of observed brightness temperatures[12] observed in row r in snapshot k. By Lemma II.1 and linearity of expectation, the expected value of this is a convolution operating on T_k—and, indeed, the rows of I. For $m \in \mathbb{Z}/\frac{M}{s}\mathbb{Z}$,

$$\begin{aligned} S_k^r[m] &= T_k[r,m] + T_k[r,m+s] + \ldots + T_k[r,m+\tfrac{s-1}{s}M] \\ &= \textstyle\sum_{c=0}^{C-1} \big(I[r,k+m,c]Q_c[r,m] + \ldots + \\ &\qquad I[r,k+m+\tfrac{(s-1)M}{s} \pmod N, c]Q_c[r,m+\tfrac{(s-1)M}{s}] \big). \end{aligned}$$

Since the subsatellite point and grid of sampled directions (ξ, η) move one pixel to the right each correlation time, by making a periodic extension of the image Earth beyond its boundary in each row, the visibilities are produced via a cyclic convolution of the values of the brightness temperature along the row.

The inversion of a row can be accomplished by taking the pseudoinverse of an $\frac{NM}{s} \times CN$ matrix $\boldsymbol{G} = \begin{pmatrix} G_{0,0} & \ldots & G_{0,C-1} \\ \vdots & \ddots & \vdots \\ G_{\frac{M}{s}-1,0} & \ldots & G_{\frac{M}{s}-1,C-1} \end{pmatrix}$, an $\frac{M}{s} \times C$ block matrix composed of $N \times N$ circulant matrix blocks $G_{m,c}$ that capture the contribution of the cth brightness temperature parameter to mth pixel in row r of the kth folded snapshot, $S_k^r[m]$, for snapshot $k \in \mathbb{Z}/N\mathbb{Z}$.

$G_{m,c}$ can be specified by giving its first row, which has s nonzero terms at positions $m, m + \frac{M}{s}, \ldots, m + \frac{s-1}{s}M$:

$$(\mathbf{0}_m Q_c[m] \mathbf{0}_{\frac{M}{s}-1} Q_c[m+\tfrac{M}{s}] \mathbf{0}_{\frac{M}{s}-1} \ldots Q_c[m+\tfrac{(s-1)M}{s}] \mathbf{0}_{\frac{M}{s}-1-m} \mathbf{0}_{N-M}),$$

where $\mathbf{0}_k = \underbrace{(0, \ldots, 0)}_{k \text{ times}}$ is the vector of k 0s.

C. The "partial Fourier" global inversion

Taking the partial Fourier transform of the visibilities,[13] we get

$$\begin{aligned} \widehat{V}_k[m,n,\omega] &= \sum_{n=0}^{N-1} V_k[m,n] e^{\frac{-2\pi i n\omega}{N}} \\ &= \sum_{c=0}^{C-1} \sum_{q,r=0}^{M-1} Q_c[m,l] e^{\frac{-2\pi i(qm+rns)}{M}} e^{\frac{2\pi i r\omega}{N}} \Big(\sum_{t=r}^{r+N-1} I[q,t,c] e^{\frac{-2\pi i t\omega}{N}} \Big) \\ &= \sum_{c=0}^{C-1} \sum_{q=0}^{M-1} \widehat{I}_{partial}[q,\omega,c] e^{\frac{-2\pi i q}{M}} \Big(\sum_{r=0}^{M-1} Q_c[q,r] e^{-2\pi i r(ns/M - \omega/N)} \Big). \end{aligned}$$

Now expressing the inversion of the transformed visibilities associated with a single orbital frequency ω, the $\frac{M^2}{s} \times CM$ $\boldsymbol{G}$ matrix is made up of $\frac{M}{s} \times C$ circulant

[7] More precisely, the trace of the subsatellite point passes through the latitude parallel $x = 0$, which is taken to move across row $\frac{M-1}{2}$ of the image (between rows if M is even). The subsatellite point's initial position, assumed to coincide with the reference longitude $y = 0$, overlies the zeroth column of the image if M is odd and lies half a pixel to the left of that column otherwise.

[8] This assumption is not consistent with DFT sampling. Recall that the DFT sampling was chosen chiefly so as to induce a highly regular pattern of folding in snapshots of brightness temperatures recovered from a single-snapshot inversion of an undersampled set of visibilities.

[9] With DFT sampling, this is the kth $M \times M$ image of parameters that lie within the unit square in the ξ-η plane:

$$I_k^c = \begin{cases} I\big[:, k - \frac{M}{2} \pmod N : k + \frac{M}{2} \pmod N, c\big] & \text{if } M \text{ is even;} \\ I\big[:, k - \frac{M-1}{2} \pmod N : k + \frac{M+1}{2} \pmod N, c\big] & \text{if not.} \end{cases}$$

(Index slices are left-inclusive and right-exclusive.)

[10] For both antennas we use simulated patterns based on the SMOS X polarization.

[11] That is, $\frac{1}{\sqrt{1-\xi^2-\eta^2}}$ where $||(\xi,\eta)|| < 1$ and 1 otherwise (where $||(\xi,\eta)||_\infty \leq 1$).

[12] Remember, we have subtracted out the mean contribution to the visibilities before decimating and inverting to form the brightness temperature snapshots S_k^r.

[13] With the contribution of $f_0(\theta)$ already subtracted out of each visibility, or equivalently from the partial Fourier transformed visibilities at the orbital frequency $\omega = 0$.

matrix blocks $G_{r,c}$ whose transpose's (also circulant) first row contains the M elements of $\widehat{Q}_c^{ZP}[:, 2ns - \omega \pmod N]$ under the permutation on $\{0, \ldots, M-1\}$ given by $\pi : n \mapsto M - n \pmod M$. Here $\widehat{Q}_c^{ZP}$ is the two-dimensional, M-point by N-point DFT of the zero-padded Q_c^{ZP}, defined by

$$Q_c^{ZP}[i,j] = \begin{cases} Q_c[i,j] \text{ if } j \leq M-1, \\ 0 \text{ otherwise.} \end{cases}$$

D. Comparison of methods

As Figure 4 shows, both methods successfully invert simulated images of brightness temperature parameters using a realistic noise model. In fact, Figure 5 shows that they produce almost identical results, in high-noise scenarios. SVD truncation error explains the differences between the two methods, and it was insignificant for the DFT sampling scenario and antenna patterns tested.

III. Conclusion

We proposed a global inversion technique for the equation that expresses the Van Cittert-Zernike theorem. Using the invariance of the acquisition geometry, we can express the partial discrete Fourier transform of the visibilities, taken along the satellite's trace, in terms of the corresponding partial discrete Fourier transform of the image of brightness temperature parameters, sampled uniformly in latitude and longitude relative to the satellite's trace. The image of brightness temperature parameters can therefore be inverted one frequency at a time. In this way, a global inversion, with its ability to expand the field of view by overcoming folding and to reduce noise, is broken down into a set of smaller inversions that can be performed independently.

This concept was validated numerically by introducing a regular, invertible pattern of folding in an orbital scenario whose sampling permits the brightness temperatures sampled in the ξ-η plane and the visibilities sampled in u-v plane to form a DFT pair. Brightness temperature parameters recovered from the visibilities via the partial Fourier inversion agreed with those obtained through a direct inversion of the folding mechanism. Future work will evaluate this technique with arbitrary, irregular baseline sampling in the u-v plane to aid the design of future satellite instruments.

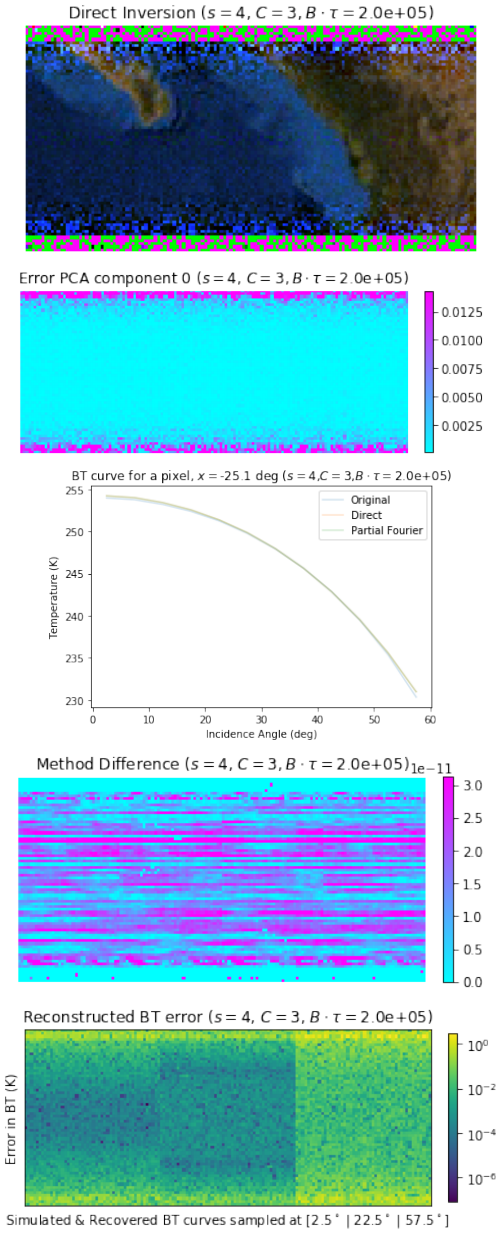

Fig. 5. The top image displays three brightness temperature (BT) parameters ($C = 3$), recovered via the direct inversion from a series of observations of highly undersampled visibilities ($s = 4$), in an unrealistically high-noise scenario (the $B\tau$ product is 100 times that of SMOS). The recovered parameters are written into each channel of this RGB image, overstating the reconstruction error: the third principal component has 10^{-3} the variance of the first, but the two contribute equally to this RGB image! (Hence the noisy blue channel.) Below that, we see the reconstruction error on the first principal component parameters for the portion of the image on the Earth's surface (that is, without the top and bottom five rows). Even at the horizon, where the signal-to-noise ratio is at its lowest, the recovered parameters give reasonable temperature curves, as the third image illustrates. The direct inversion and the partial Fourier inversion give, for each pixel, nearly identical BT curves, and these two curves agree with the curve resulting from the pixel's actual BT parameters. The fourth image displays the pixelwise difference in the second principal component between the two global inversion methods, which is less than 10^{-10}. The final image shows the error (in Kelvin) in the reconstructed BT curve for each pixel on Earth's surface, sampled at 2.5° (left third of the image), 22.5° (middle third), and 57.5° (right third). The error, at a given incidence angle, is lowest at latitudes that are observed at that incidence angle the most. At low incidence angles, pixels near the trace show the lowest error. The reconstruction error at intermediate incidence angles is lowest in intermediate latitude bands, whose pixels were observed at intermediate incidence angles in the most snapshots. All pixels were observed at 57.5° roughly as often (and with similar antenna pattern values), except those at the most extreme latitudes.

References

[1] É. Anterrieu et al., "Preliminary system studies on a high-resolution SMOS follow-on: SMOS-HR," in *IEEE Intl. Geosci. Remote Sens. Symp.*, 2019, pp. 8451-8454.

[2] Y. Kerr et al. "The SMOS mission: New tool for monitoring key elements of the global water cycle." *Proc. IEEE*, vol. 98, no. 5, pp. 666-687, 2010.

[3] R. D. Ekers and A. H. Rots. "Image Formation from Coherence Functions in Astronomy." *Proc. IAU Coll*, vol. 49, pp. 61-64, 1979.

[4] J. Pety and N. Rodriguez-Fernandez. "Revisiting the theory of interferometric wide-field synthesis." *Astron. & Astrophys.*, vol. 517, pp. A12, 2010.

[5] I. Corbella et al. "The visibility function in interferometric aperture synthesis radiometry." *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1677-1682, 2004.

[6] A. Al Bitar et al. "The global SMOS Level 3 daily soil moisture and brightness temperature maps." *Earth Sys. Sci. Data*, vol. 9, no. 1, pp. 293-315, 2017.

[7] P.H. van Cittert. "Die Wahrscheinliche Schwingungsverteilung in Einer von Einer Lichtquelle Direkt Oder Mittels Einer Linse Beleuchteten Ebene." *Physica*, vol. 1, no. 1-6, pp. 201-210, 1934.

[8] F. Zernike. "The concept of degree of coherence and its application to optical problems." *Physica*, vol. 5, no. 8, pp. 785-795, 1938.